

Lecture Notes in Networks and Systems 1317

Akshi Kumar
Abhishek Swaroop
Pancham Shukla *Editors*

Proceedings of Fourth International Conference on Computing and Communication Networks

ICCCN 2024, Volume 1

 Springer

Lecture Notes in Networks and Systems

Volume 1317

Series Editor

Janusz Kacprzyk , Systems Research Institute, Polish Academy of Sciences,
Warsaw, Poland

Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA,
School of Electrical and Computer Engineering—FEEC, University of
Campinas—UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering,
Bogazici University, Istanbul, Türkiye

Derong Liu, Department of Electrical and Computer Engineering, University of
Illinois at Chicago, Chicago, USA

Institute of Automation, Chinese Academy of Sciences, Beijing, China

Witold Pedrycz, Department of Electrical and Computer Engineering, University of
Alberta, Alberta, Canada

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Marios M. Polycarpou, Department of Electrical and Computer Engineering,
KIOS Research Center for Intelligent Systems and Networks, University of Cyprus,
Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary

Jun Wang, Department of Computer Science, City University of Hong Kong,
Kowloon, Hong Kong

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Akshi Kumar · Abhishek Swaroop ·
Pancham Shukla
Editors

Proceedings of Fourth International Conference on Computing and Communication Networks

ICCCN 2024, Volume 1

Editors

Akshi Kumar 
Department of Computing
Goldsmiths University of London
London, UK

Pancham Shukla
Department of Computing
Faculty of Engineering
Imperial College London
London, UK

Abhishek Swaroop
Department of Information Technology
Bhagwan Parshuram Institute
of Technology
New Delhi, Delhi, India

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-96-3941-0

ISBN 978-981-96-3942-7 (eBook)

<https://doi.org/10.1007/978-981-96-3942-7>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

ICCCNet-2024 Committees

ICCCNet-2024 Steering Committee

General Chairs

Prof. Dr. Omer Rana, Cardiff University, UK

Prof. Dr. Yang Xiao, The University of Alabama, USA

Honorary Chairs

Prof. Dr. Vincenzo Piuri, University of Milan, Italy

Prof. Dr. Janusz Kacprzyk, FIEEE, Polish Academy of Sciences, Poland

Prof. Dr. Manu Malek, EiC Computer and Electrical Engineering, Stevens Institute of Technology, USA

Dr. Jon G. Hall, EiC Expert Systems (WILEY), The Open University, UK

Conference Chairs

Dr. Ali Kashif Bashir, Manchester Metropolitan University, UK

Prof. Dr. Darren Dancey, Manchester Metropolitan University, UK

Prof. Abhishek Swaroop, Bhagwan Parshuram Institute of Technology, India

Technical Program Chairs

Dr. Rajkumar Rathore, Cardiff Metropolitan University, UK
Fatma Bassyouni, National Research Centre, Cairo, Egypt
Aditya Khamparia, Babasaheb Bhimrao Ambedkar University, India

Conveners

Dr. Akshi Kumar, Godsmith University of London, UK
Dr. Pancham Shukla, London Metropolitan University, London, UK
Dr. Utku Kose, Suleyman Demirel University, Isparta, Turkey
Dr. Harpreet Singh Arora, Tel-Aviv University, Israel

Publicity Chairs

Prof. Dr. Valentina Emilia Balas, Aurel Vlaicu University of Arad, Romania
Prof. Dr. Zdzislaw Polokowski, Jan Wyzykowski University, Polkowice, Poland
Prof. Dr. Vijay Singh Rathore, IIS (Deemed to be University), Jaipur
Dr. Hamid Reza Boveiri, Sama College, Islamic Azad University, Shoushtar Branch, Iran
Dr. Syed Bilal Shah, Darul Hekma University, Saudi Arabia
Dr. Umesh Gupta, Bennett University, India
Dr. Bilal Ozturk, Software Engineering Department, Istanbul Aydin University, Istanbul, Turkey

ICCCNet-2024 Advisory Committee and Technical Program Committee

Prof. Dr. Vincenzo Piuri, University of Milan, Italy
Prof. Dr. Valentina Emilia Balas, Aurel Vlaicu University of Arad, Romania
Prof. Dr. Aboul Ella Hassanien, Cairo University, Egypt
Prof. Dr. Vineet Kansal, ProVC, AKTU; Director, IET, Lucknow, India
Prof. Dr. João Manuel R. S. Tavares, Universidade do Porto (FEUP), Portugal
Prof. Dr. Neeraj Kumar, Thapar Institute of Engineering and Technology, Patiala, Punjab, India
Prof. Dr. Zdzislaw Polkowski, Jan Wyzykowski University, Poland
Prof. Dr. George A. Tsihrintzis, University of Piraeus, Greece
Prof. Dr. Arun Sharma, Indira Gandhi Delhi Technical University for Women, India

Prof. Dr. Abhishek Swaroop, Bhagwan Parshuram Institute of Technology, India
Prof. Dr. Giorgos Karagiannidis, Aristotle University of Thessaloniki, Greece
Prof. Dr. Sheng-Lung Peng, National Dong Swa University, Taiwan
Dr. Dijana Oreski, University of Zagreb, Varazdin, Croatia
Dr. Ahmed Zobaa, Brunel University London, UK
Prof. Dr. A. K. Singh, National Institute of Technology Kurukshetra, India
Prof. Dr. Anil Kumar Ahlawat, KIET Group of Institutes, India
Dr. Jafar A. Alzubi, Al-Balqa Applied University, Salt, Jordan
Prof. Dr. Alex Norta, Tallinn University of Technology, Estonia
Dr. Utku Kose, Suleyman Demirel University, Isparta, Turkey
Prof. Dr. Isabel De la Torre Díez, University of Valladolid, Spain
Dr. Oana Geman, Universitatea Stefan cel Mare din Suceava, Romania
Dr. Varun G. Menon, SCMS School of Engineering and Technology, Kochi, India
Dr. Mohammad Shojafar, University of Surrey, UK
Dr. Muhammad Habib ur Rehman, King's College London, UK
Dr. Irfan Mehmood, University of Bradford, UK
Dr. Muddesar Iqbal, London South Bank University, UK
Dr. Shahid Mumtaz, Institute de Telecommunication, Portugal
Dr. Waleed Ejaz, Lakehead University, Canada
Dr. Anish Jindal, University of Essex, UK
Dr. Suresh Chavhan, Vellore Institute of Technology, Vellore, India
Dr. Gagangeet Singh Aujla, Durham University, UK
Dr. Sachin Kumar, South Ural State University, Chelyabinsk, Russian Federation
Dr. Prayag Tiwari, Aalto University, Finland
Dr. Pradeep Malik, KIIT University, India
Dr. Akash Kumar Bhoi, Sikkim Manipal Institute of Technology, India
Dr. Hamid Reza Boveiri, Sama College, IAU, Shoushtar Branch, Shoushtar, Iran
Dr. Sahil Garg, École de technologie supérieure, Université du Québec, Montreal, Canada
Dr. Gulshan Shrivastava, Sharda University, Greater Noida, India
Dr. Gabriella Casalino, University of Bari, Italy
Dr. Aditya Khamparia, Babasaheb Bhimrao Ambedkar University, India
Dr. Amit Kumar Jaiswal, University of Leeds, UK
Dr. Qianqian Xie, University of Manchester, Manchester, UK
Dr. Yousaf Bin Zikria, Yeungnam University, South Korea
Dr. Francesco Piccialli, University of Naples Federico II, Italy
Dr. Ashiq Anjum, University of Leicester, UK
Prof. Dr. Nuno M. Garcia, University of Beira Interior, Covilhã, Portugal
Prof. Dr. Kashif Saleem, Universiti Teknologi Malaysia, Riyadh, Saudi Arabia
Dr. Le Hoang Son, University of Dannang, Vietnam
Dr. Jaafar Alghazo, Virginia Military Institute, Lexington, VA, USA
Dr. Jalil Piran, Sejong University, South Korea
Dr. Hari Mohan Pandey, Edge Hill University, UK
Prof. Dr. P. Sanjeevikumar, Aarhus University, Herning, Denmark
Dr. Kemal Polat, Abant Izzet Baysal University, Turkey

Dr. Juhriyansyah Dalle, Universitas Lambung Mangkurat, Indonesia
Dr. Ahmed Elngar, Beni-Suef University, Egypt
Dr. Prajoy Podder, Institute of Information and Communication Technology, BUET, Dhaka, Bangladesh
Dr. M. Rubaiyat Hossain Mondal, Institute of Information and Communication Technology, BUET, Dhaka, Bangladesh
Dr. Sarada Prasad Gochhayat, Old Dominion University, USA
Dr. Daniel Nogueira, University of Minho, Brazil
Dr. Khan Muhammad, Sejong University, South Korea
Dr. Yenumula B. Reddy, Grambling State University, Louisiana, USA
Dr. Chandran Venkatesan, KPR Institute of Engineering and Technology, India
Dr. Alireza Jolfaei, Macquarie University, Australia
Dr. Souvik Ganguli, Thapar Institute of Engineering and Technology, India
Dr. Flah Aymen, National School of Engineering of Gabes, Tunisia
Prof. Dr. Placido Rogerio Pinheiro, University of Fortaleza, Brazil
Dr. Daniela Clara Moraru, University of Luxembourg, Luxembourg
Dr. Gautam Srivastava, Brandon University, Canada
Dr. Vassilis C. Gerogiannis, University of Thessaly, Greece
Prof. Dr. B. S. Manoj, Indian Institute of Space Science and Technology, India
Dr. M^a Luz Castro Pena, Universidade da Coruña, Spain
Dr. Ilya Levin, Tel Aviv University, Israel
Dr. Muhibul Haque Bhuyan, Southeast University, Bangladesh
Prof. Dr. Med Salim Bouhlef, Lab SETIT; Sfax University, Tunisia
Dr. Mamoun Alazab, Charles Darwin University, Australia
Dr. Lalit Garg, University of Malta, Msida, Malta
Dr. Arij Naser Abougrien, University of Tripoli, Libya
Dr. Sherif Mohamed Ismail, Egyptian German Academy
Prof. Dr. Vijay Singh Rathore, IIS Deemed to be University, India
Dr. Aslanbek Naziev, Ryazan State University named after S. A. Esenin, Russia
Dr. Mwaffaq Ootom, Yarmouk University, Jordan
Dr. Ahmed A. Ewees, Damietta University, Egypt
Dr. Iwan Adhicandra, University of Sydney, Australia
Prof. Dr. Meng Li, Hefei University of Technology, China
Dr. Korhan Cengiz, Trakya University, Turkey
Dr. Muhammad Bilal, Hankuk University of Foreign Studies, South Korea
Dr. RR Venkatesha Prasad, TU Delft, The Netherlands
Dr. Özge Korkmaz, Malatya Turgut Özal University, Turkey
Dr. Alexander Fedorov, Rostov State University of Economics, Russia
Prof. Dr. Alfredo Grieco, Politecnico di Bari, Italy
Prof. Dr. Quoc-Viet Pham, Pusan National University, South Korea
Dr. Enkeleda Lula, University Haxhi Zeka, Peja, Kosovo
Dr. Fides del Castillo, De La Salle University, Philippines
Dr. Dr. Houda Chihi, Innov'COM Lab of Sup'COM, Tunisia
Prof. Dr. Tu Nguyen, Kennesaw State University, Kennesaw, USA
Prof. Dr. Christos Douligeris, University of Piraeus, Greece

Dr. Dr. Surbhi Bhatia, King Faisal University, Saudi Arabia

Dr. Feras M. Awaysheh, Tartu University, Delta Research Center, Estonia

Dr. Assunta Di Vaio, University of Naples “Parthenope” (Italy)

Dr. Mehdi Gheisari, Southern University of Science and Technology, Shenzhen, Guangdong Province, P. R. China

Dr. Suleyman Eken, Kocaeli University, Turkey

Preface

We hereby are delighted to announce that Manchester Metropolitan University, Manchester, UK, has hosted the eagerly awaited and much coveted International Conference on Computing and Communication Networks (ICCCNet-2024) in hybrid mode during October 17–18, 2024. The conference has attracted many high-quality submissions and stimulates the cutting-edge research discussions among many academic pioneering researchers, scientists, industrial engineers, students with the reception of papers including more than 4800 authors from different parts of the world. The committee of professionals dedicated toward the conference is striving to achieve a high-quality technical program with tracks on Networks and Computing Technologies, Advances in Artificial Intelligence and Machine Learning, Security and Privacy, Emerging Topics in 5G/6G Communication Systems, Cyber physical Systems, Emerging Trends in Data Analytics, Cyber Security for Industry 4.0 and Smart and Sustainable Environmental Systems. All the tracks chosen in the conference are interrelated and are very famous among present-day research community. Therefore, a lot of research is happening in the above-mentioned tracks and their related sub-areas. More than 1150 full-length papers have been received, among which the contributions are focused on theoretical, computer simulation-based research and laboratory-scale experiments. Among these manuscripts, 304 papers have been included in the Springer proceedings after a thorough two-stage review and editing process. All the manuscripts submitted to the ICCCN-2024 were peer-reviewed by at least two independent reviewers, who were provided with a detailed review pro forma. The comments from the reviewers were communicated to the authors, who incorporated the suggestions in their revised manuscripts. The recommendations of two reviewers were taken into consideration while selecting a manuscript for inclusion in the proceedings. The exhaustiveness of the review process is evident, given the large number of articles received addressing a wide range of research areas. The stringent review process ensured that each published manuscript met the rigorous academic and scientific standards. It is an exalting experience to finally see that these elite contributions materialize into six volumes as ICCCN-2024 proceedings by Springer entitled “International Conference on Computing and Communication Networks”.

ICCCNet-2024 invited two keynote speakers, who are eminent researchers in the field of computer science and engineering, from different parts of the world. In addition to the plenary sessions on each day of the conference, twelve (four in physical mode and twenty six in online mode) concurrent technical sessions are held every day to assure the oral presentation of around 304 accepted papers. Keynote speakers and session chair(s) for each of the concurrent sessions have been leading researchers from the thematic area of the session. A technical exhibition is held during all the 2 days of the conference, which has put on display the latest technologies, expositions, ideas and presentations.

An international conference of such magnitude and release of the ICCCNNet-2024 proceedings by Springer has been the remarkable outcome of the untiring efforts of the entire organizing team. The success of an event undoubtedly involves the painstaking efforts of several contributors at different stages, dictated by their devotion and sincerity. Fortunately, since the beginning of its journey, ICCCNNet-2024 has received support and contributions from every corner. We thank them all who have wished the best for ICCCNNet-2024 and contributed by any means toward its success. The edited proceedings volume by Springer would not have been possible without the perseverance of all the steering, advisory and technical program committee members.

All the contributing authors owe thanks from the organizers of ICCCNNet-2024 for their interest and exceptional articles. We would also like to thank the authors of the papers for adhering to the time schedule and for incorporating the review comments. We wish to extend my heartfelt acknowledgment to the authors, peer-reviewers, committee members and production staff whose diligent work put shape to the ICCCNNet-2024 proceedings. We especially want to thank our dedicated team of peer-reviewers who volunteered for the arduous and tedious step of quality checking and critique on the submitted manuscripts. The management, faculties, administrative and support staff of the university have always been extending their services whenever needed, for which we remain thankful to them.

Lastly, we would like to thank Springer for accepting our proposal for publishing the ICCCNNet-2024 conference proceedings. Help received from Mr. Aninda Bose, the acquisition senior editor, in the process has been very useful.

Organizers and Conveners

ICCCNet-2024

Akshi Kumar

Abhishek Swaroop

Pancham Shukla

London, UK

Delhi, India

London, UK

About This Book

International Conference on Computing and Communication Networks (ICCCNet-2024) was held on 17–18th October 2024 at Manchester Metropolitan University, Manchester, UK. The conference has attracted many high-quality submissions and stimulate the cutting-edge research discussions among many academic pioneering researchers, scientists, industrial engineers, students with the reception of papers including more than 4800 authors from different parts of the world. Only 304 papers have been accepted and registered with an acceptance ratio of 24% to be published in six volumes of prestigious springer *Lecture Notes in Networks and Systems* (LNNS) series.

Contents

Voice and Speech in Parkinson's Disease: Advancements in Early Diagnosis and Treatment	1
K. M. Varsha and Akhilesh Verma	
An Efficient Approach for Decentralized Web Messaging Using WebRTC	17
Manisha Kasar, Trupti Suryawanshi, Pranoti Kavimandan, Snehal Chaudhari, Shreekala Deshmukh, and Pramod Jadhav	
Face Identification and Categorization with and Without Mask Utilizing MTCNN and OpenCV for Accessing Bank Locker Facility	33
R. Monica Lakshmi, K. Ramar, Nikita Devendran, L. P. Deepika, and S. Mounika	
Software Engineering Approach for Automated Vehicle Detection and Classification on Remote Sensing Images	51
Mohammed Assiri	
A Portable and Robust Smart Mobile Prototype for Millet Crop Productivity Assessment	69
Akansha Tiwari, Vaibhav Kumar, Soumya Sahoo, Sushruta Mishra, and Mohammed Al-Farouni	
Prognostic Analysis of Meteorological Data for a Small Solar Farm Using Machine Learning	81
Phuoc Quy Phong Nguyen, Prabhu Paramasivam, Huu Cuong Le, Duy Tan Nguyen, Van Phuc Nguyen, and Thanh Nam Dang	
Feature Selection Using Hybrid Opposition-Based Whale Optimization Algorithm	99
Parul, Charu Gupta, and Devendra Kumar Tayal	

ECG Classification System with Novel Spike-Driven Simulated Neural Network	117
Deepa Jose, Fahima Niaz, K. Kanaga Durga, and Suchismita Singh	
Performance Optimization and Application Exploration Based on Intelligent Algorithm and IT Technology in Cloud Computing Task Scheduling System	129
Ying Ding	
Diabetic Retinopathy Diagnosis Using Convolutional Neural Networks	141
Vahini Siruvoru, Nampally Vijay Kumar, Venkata Satya Sai Sagiraju, Jayanth Venavenka, and Jarupula Hrithika	
Wearable Technology, A Smart Step Towards Healthcare—A Review	153
Niharika Nagpal, Anshika Sharma, and Divya Agarwal	
Revolutionizing Food Supply Chains: An AI and ML-Driven Model for Enhanced Quality Control and Traceability	163
Arjun Kumar Chandrasekar, G. Anitha, and Vigneswaran Narayanamurthy	
Intelligent Irrigation Mechanism that Uses Machine Learning and Artificial Intelligence to Improve the Sensor Network Topology	177
K. P. Sriram, P. Kola Sujatha, M. Robinson Joel, and M. Gokuldhev	
Algorithm Bias Detection System Based on Artificial Intelligence	187
Jiayi Zhang, Sha Yang, and Haoping Guo	
An Intelligent Probability-Based Supervised Localized Framework for Cardiologist Recommendation Using Predictive Analytics Approach	197
Suvra Mukherjee, Shruti Dutta, Kunal Anand, Lu Wang, and Bharati Rathore	
Optimizing the Performance of Diabetes Risk Prediction Using Ensemble Learning Techniques	209
Sukurat Salam, Glen Hopkinson, Christopher G. Udomboso, and Serifat Folorunso	
A Secure Data Dissemination in VANET Using Trust and Random Forest Algorithm	225
C. Balakumar and B. Akashpriyan	
Cognitive Disorder Detection Used for Early Autism Screening Using Deep Learning	239
R. Gunasundari and G. Jeyabharathi	

A Precise Modelling Endeavour for the Emission of Biodiesel Powered Diesel Engine Using Gene Expression Programming	253
Tien Han Nguyen, Anh Quan Nguyen, Duy Tan Nguyen, Minh Thai Duong, and Van Vang Le	
Analysis and Prediction of Water Access and Hygiene on Resource Data Using Machine Learning	267
R. Gunasundari and D. Sakthivel	
SDN-Enabled Secure Data Communication in 5G VANETS Using ECC Algorithm	283
C. Balakumar and U. Priyatharshini	
Trust and Reinforcement Learning-Based Security Model (TRLS) for Mitigating Black Hole Attacks in Vehicular Ad Hoc Networks	297
C. Balakumar and G. Surya	
X-ray Classification for Content-Based Medical Image Retrieval	311
S. Subashree, E. Saranya, M. Chinnadurai, and M. Rajakumaran	
A Node to Node Sensory Cloud-Driven Smart Drip Irrigation Model with Relayed Connectivity Structure	327
Sakshi Priya, Pranav Prakash, Priyansh Kumar Singh, Sushruta Mishra, and Muntather Almusawi	
Enhanced Cyberbullying Detection Using Named Entity Recognition in Online Social Networks	339
Suresh Sankaranarayanan, K. Meenakshi, and K. Sandhiya	
Hybrid Deep Learning Mechanism for Brain Tumor Detection	351
Vahini Siruvoru, Nampally Vijay Kumar, Dasari Rini Ruha, Boorla Asritha, and Sahithi Polunati	
Spot the Fake Legs: Synthetic Generating Medical Images Using Latent Diffusion Model (SGMs)	361
Sura Ahmed Abd Alasidi, Khawla Hussein Ali, and Zaid Ameen Abduljabbar	
Research on the Design of Computer Security Protection System Based on Machine Learning	373
Haiming Zhang, Yuqi Wang, and Junpu Cai	
LLMs in Healthcare: Exploring Language Models for Dental, Medical, and Mental Health with LangChain	383
S. Krishnaveni, B. Jothi, S. Sivamohan, Gautham Brijesh, S. Aiswarya, Louie Allen, and Syed Faizan Fiaz	

Harvesting Insights: A YOLOv5 Template Design for Real-Time Pesticide Recommendation with Reduced Computational Complexity	399
Swati Kapoor, Ipsita Majhi, Gyan Prakash Dash, Sushruta Mishra, and Haider Alabdeli	
Segmentation of Tapioca Plants Point Cloud Data from Irregular Ground Surfaces	411
S. Prince Sahaya Brighty, R. Anuradha, and M. Sandeep	
Denoising Diffusion Probabilistic Model with Fashion MNIST	427
Palak Singh, Pratyush Priyam, Ayush Yadav, Aaron Issac Abraham, Sudeshna Datta, Nimish Singh, and Pranjal Kumar	
The World State of Artificial Intelligence Based on the Analysis of Research Papers and Patents	443
Adriana Baravalle, Ignacio del Carril, and Gabriela Robiolo	
A Novel Factorized with Label Smoothing Cognitive Architectural Framework for Apple Leaf Disease Prognosis	455
Arindam Kanrar, Swayantan Maji, V. Sanjay, Sushruta Mishra, and Hassan M. Al-Jawahry	
Blood Pressure Estimation from Photoplethysmogram Signals: An EdgeML Approach	465
Nithin Aditya Pradeepkumar and Krishna Karthik Bhaskar	
Secure and Verify Exchange Data in VANET Using Blockchain	475
Zahraa Sh. Alzaidi, Ali A. Yassin, Zaid Ameen Abduljabbar, and Vincent Omollo Nyangaresi	
Hyperparameter Tuning for Machine Learning Algorithms Used for Crops Recommendation System	489
Humam R. Aljburi and Jamal Mustafa Al-Tuwaijari	
Intelligent Digital Industrial Internet of Things Network Security Monitoring Based on Edge Cloud Computing	501
Yijian Qiu and Chuang Wang	
Visualizing the Alarming Surge: Mapping and Analyzing Suicide Patterns Across India	511
P. S. Metkewar, Syed Irfan Yaqoob, Shubhalaxmi Joshi, Anshul Mishra, and Ishita Kalbandhe	
Astroturfing: Understanding and Addressing the Influence of Fake Grassroots Movements Across Diverse Domains on Social Media Platforms	527
Akshata Balasaheb Badade and Rajesh Kumar Dhanaraj	

An Improved DDOS Attack Preventing Model for Securing Cloud Computing Utilizing Hybrid Deep Learning Algorithms	539
Abdulbasit Alazzawi and Omer Salam	
PlantPath: Deep Learning-Based Plant Pathogen Detection	553
Kooturu Kanishk Reddy, Sujal Mangesh Limje, T. Y. J. Naga Malleswari, and S. Ushasukhanya	
Sentiment Extraction and Mining of Textual Content in Social Networking Space Using Context-Based Intelligent Retrieval System	567
Aastha Sahani, Oindrella Chatterjee, Tiansheng Yang, Bharati Rathore, V. Sanjay, Urvash Shende, and Sushruta Mishra	
Artificial Intelligence and Tissue Biomarkers: Benefits, Risks, and Prospects for Pathology	579
Mohamed Bobor Jalloh, Manbeer Singh Sidhu, and Vivek Kumar Garg	
Audio Sentiment Analyzer for Mental Health Support Hotline	591
Sneha Maurya, Shivani Saxena, Shreya Shukla, Sukriti Maurya, and Akhilesh Verma	
The Application of Artificial Intelligence Algorithms in Computer Network Information Security Protection and Evaluation	605
Qianqian Yue	
An Optimal EEG Features Extraction Methods for Autism Detection	619
Hajir Ammar Hatim and Zaid Abdi Alkareem Alyasseri	
Web Scraping of Scientific Literature From SCOPUS Repositories of Iraqi Universities	633
Zahraa Taufeeq Abdul-sahib Al-Madhhachi and Salma A. Mahmood	
A Space-Saving Technique in SDN for Identifying Top-k Super-Spreaders Using P4-Enabled Switches	659
Ali Nadim Alhaj, Wilson Naik Bhukya, and Rajendra Prasad Lal	
Assessment of Advancements in Online Examination Proctoring Systems	675
Sangeeta Lamba and Neelam Sharma	
Drug Discovery Using Variational Quantum EigenSolver	691
S. Surya Prakash, P. Banu Priya, Sudarshan Someneni, Udendhran, and Arnab Banerjee	
Biomimetic Swarm-Based UAV Collision Avoidance Control	705
Wenfei Dai, Zhiyi Wang, and Chaijinyang Ding	

5G-SHE: NIST-ASCON-Based Security Framework for 5G-Enabled Smart Healthcare Ecosystem	719
Vaishali Meena and Gaurav Indra	
Design and Application of Wearable Motion Training and Monitoring System Based on Convolutional Neural Network	731
Mei Hu and Chang Chen	

About the Editors

Dr. Akshi Kumar is currently a senior lecturer (associate professor) and Director-Post graduate Research (PGR) in the Department of Computing at Goldsmiths, University of London, UK. My academic path spans from a Doctorate in the domain of Web Mining at the University of Delhi, India to a post-doctoral on Explainable AI for NLP from Brazil. I've held pivotal positions at Manchester Metropolitan University, UK as well as at NSUT and DTU in New Delhi, India. Recognized as an "Exceptional Promise in AI and Data Science" by the Royal Academy of Engineering, UK in 2022, I've received international accolades, including being named in Stanford's "Top 2% Scientist" list in 2021 and 2022. I've successfully guided 6 doctorates and 33 Master theses. Serving as an Associate and Guest Editor for impactful journals, my research encompasses Affective Computing, Social media and Network analytics, Natural Language Processing (NLP), and AI-driven healthcare.

Prof. Abhishek Swaroop completed his B.Tech. (CSE) from GBP University of Agriculture and Technology, M.Tech. from Punjabi University Patiala, and Ph.D. from NIT Kurukshetra. He has Industrial experience of 8 years in organizations like Usha Rectifier Corporations and Envirotech Instruments Pvt. Limited. He has 22 years of teaching experience. He has served in reputed educational institutions such as Jaypee Institute of Information Technology, Noida, Sharda University Greater Noida and Galgotias University Greater Noida. He has served at various administrative positions such as Head of the Department, Division Chair, NBA coordinator for the university, Head Training and Placements. Currently, he is serving as Professor and HoD department of Information Technology in Bhagwan Parshuram Institute of Technology, Rohini, and Delhi. He is actively engaged in research. He has more than 60 quality publications out of which 8 are SCI and 16 scopus. Two of his Ph.D. Scholars have completed his Ph.D. from NIT Kurukshetra and IIT Dhanbad. He is currently supervising one Ph.D. student from AKTU Lucknow. He has guided 10 M.Tech., Dissertations, and numerous other Projects. He has authored 3 books and 5 book chapters. He had been part of the organizing committee of three IEEE conferences, three Springer conferences (ICICC-2018) as technical program chair. He was the convener of recently organized national conference NCCIN-2K18. He is

member of various professional societies like CSI and ACM and editorial board of various reputed journals.

Dr. Pancham Shukla holds Ph.D. in Electrical and Electronic Engineering from Imperial College London, M.Phil. in Electronic and Electrical Engineering from University of Strathclyde (Glasgow, UK), ME in Electronic and Computer Engineering and BE in Electronic Engineering from Sardar Patel University (India).

Prior to joining London Met in 2007, Dr. Shukla worked as a Research Fellow in The Applied Digital Signal and Image Processing Research Centre (ADSIP) at the University of Central Lancashire and as a Research Associate in Communications and Signal Processing (CSP) group at Imperial College London. While pursuing his M.Phil. in Centre for Signal and Image Processing (CeSIP) at the University of Strathclyde, Dr. Shukla was involved in tutoring postgraduate courses in mathematical signal processing. Dr. Shukla has been involved in teaching and research in India and UK since 1996.

Voice and Speech in Parkinson's Disease: Advancements in Early Diagnosis and Treatment



K. M. Varsha and Akhilesh Verma

Abstract In this review, we are looking at Parkinson's Disease, specifically how it impacts on speech and voice functions that include hypokinetic dysarthria which affects both communication and emotional expression. This paper highlights the importance of voice analysis during early stages of PD through a focus on Artificial Intelligence (AI) and Machine Learning (ML) for PD diagnosis. It also examines speech changes in PD, such as diminished pitch variability and lower articulatory precision, and underlies gender specific contrasts. With increased global incidence rates of PD presentation with both motor and non-motor symptoms like sleep issues among others, this study argues for voice changes to be signs of early stage PD. In general, its purpose is to improve diagnostic techniques as well as treatments holistically. One of the main issues addressed in this article is the role played by speech analysis and voices in identifying people who have been affected with Parkinson disease at an early stage; especially when it comes to constant monitoring of symptoms or clinical decision-making based on information provided by smartphones. The recommendations also highlight special therapies for enhancing sociability and psychological well-being among those suffering from PD. Finally, the study suggests further researches and technological improvements concerning symptomatic speeches caused by Parkinson's disease with an emphasis on making affordable home-based monitoring devices that would be integrated into healthcare facilities.

Keywords Parkinson's disease (PD) · Voice disorder detection · Speech signal analysis · Pattern recognition

K. M. Varsha · A. Verma (✉)
Ajay Kumar Garg Engineering College, Ghaziabad, India
e-mail: vermaakhilesh@akgec.ac.in

K. M. Varsha
e-mail: varsha2210005m@akgec.ac.in

1 Introduction

Parkinson's disease (PD) emerges as a prominent thread in the intricate tapestry of neurological conditions, weaving a complex pattern of symptoms and challenges. Characterized by the progressive degeneration of nerve cells in the brain's substantia nigra, PD stands as a prevalent disorder, following Alzheimer's in its incidence [1, 2]. It predominantly affects the elderly, marking its presence in 0.5–1% of those aged 65–69 and escalating to 1–3% in individuals over 80 years [3, 4], intertwining a deficiency in dopamine with a spectrum of motor and non-motor impairments [5].

1.1 *Understanding and Diagnosing Voice and Speech Disorders in Parkinson's Disease and Other Neurological Conditions*

In the realm of neurological disorders, conditions such as Parkinson's Disease (PD) and Adductor Spasmodic Dysphonia (ASD) severely impact voice quality and control. Through advancements in artificial intelligence (AI), this study differentiates between healthy voices and those affected by PD and ASD [6]. This analysis, which links historical data with advanced modeling, highlights the potential for remote health care and supports early disease detection [7]. Hypokinetic dysarthria, which primarily affects speech, is frequently caused by PD. This study focuses on the application of machine learning (ML) in Parkinson's Disease, addressing challenges such as the need for comprehensive multimodal approaches and the pursuit of model interpretability [8–11]. Assessing speech alterations in PD patients, including reduced pitch and articulation, is essential for the early detection and understanding of voice patterns associated with the disorder.

In addition, Bradykinesia and resting tremors are defining features of Parkinson's disease, a condition caused by the gradual depletion of dopaminergic neurons [12, 13]. This progressive disorder affects motor skills and everyday activities globally, with roughly one million people diagnosed in the United States alone [14, 15]. Researchers are increasingly examining voice alterations as early indicators of Parkinson's disease, emphasizing the necessity of a thorough grasp of the disease's causes to advance diagnostic methods and therapeutic strategies [16–19]. This focus also underscores the significance of AI and machine learning in deepening our comprehension of these intricate neurological conditions.

In an effort to provide a comprehensive understanding of Parkinson's disease and its multifaceted impact, it is crucial to integrate recent research findings with a structured overview of the disease's symptoms. The exploration by [54] features key side effects of Parkinson's sickness, including engine side effects like quakes, unbending nature, and bradykinesia, as well as non-engine side effects, e.g., mental changes, mind-set problems, and rest aggravations. These side effects, sorted into engine and non-engine types, add to the intricacy and testing nature of the condition.

Table 1 Classification of Parkinson’s disease symptoms: motor and non-motor manifestations

Category	Symptom
Motor symptoms	Tremors (typically at rest), bradykinesia (slowness of movement), rigidity (stiffness of the limbs), postural instability
Non-motor symptoms	Cognitive impairment, depression, anxiety, sleep disturbances (insomnia, REM sleep behavior disorder), fatigue, loss of sense of smell (anosmia), autonomic dysfunction (orthostatic hypotension, constipation), speech, and swallowing difficulties

This synthesis of research and symptomatic categorization is instrumental in illustrating the broad spectrum of challenges faced by individuals with Parkinson’s disease. It emphasizes the importance of adopting a holistic approach in both diagnosis and treatment. By understanding the wide range of symptoms—both motor and non-motor—it becomes clear that Parkinson’s disease is a condition with diverse manifestations, necessitating tailored and comprehensive management strategies (Table 1).

1.2 Exploring Speech Patterns, Diagnostic Techniques, and Life Quality Enhancement in Parkinson’s Disease

Parkinson’s Sickness: The deficiency of dopamine-delivering neurons in Parkinson’s Illness (PD) appears as unmistakable engine side effects and hypokinetic dysarthria, which adjusts discourse designs with critical orientation explicit contrasts [20]. Neuroimaging studies have started to examine the one of a kind discourse qualities of ladies with PD, including more limited stops and expanded subharmonics [21]. Besides, the effect of PD on discourse designs presents different difficulties [22]. Disrupted muscle coordination is the root cause of these speech impairments, which fall into the phonatory, articulatory, and prosodic categories [23, 24]. AI-based diagnostic methods can improve the accuracy of PD diagnoses by identifying these altered speech characteristics [25, 26].

Emphasizing the utility of smartphone-based assessments, this review underscores the importance of using voice and speech as biomarkers in PD [27–29]. Analyzing speech abnormalities aids not only in early detection but also in the continuous monitoring of PD, contributing to the development of diagnostic tools and interventions [30–33]. Moreover, PD’s speech disorders significantly affect quality of life, reflecting the disease’s severity. Addressing these disorders through targeted therapy can maintain social connections and improve adherence to medical advice, thus playing a crucial role in enhancing the emotional well-being and overall quality of life for individuals with PD. This comprehensive approach highlights the significance of speech and voice analysis in understanding and managing Parkinson’s Disease, aligning medical and therapeutic strategies with patients’ daily experiences.

To further elucidate the array of interventions available for speech disorders in PD, the following table provides a detailed overview. Entitled “Overview of Therapeutic Interventions for Speech Disorders in Parkinson’s Disease: Descriptions, Effectiveness, and Limitations,” it presents a spectrum of treatments, from speech therapy and medication to innovative technological solutions, and evaluates their effectiveness and limitations, along with corresponding references. This table serves as a valuable resource for understanding the multifaceted approaches employed in managing speech-related symptoms in PD (Table 2).

Table 2 Overview of therapeutic interventions for speech disorders in Parkinson’s disease: descriptions, effectiveness, and limitations

Intervention	Description	Effectiveness	Limitations	References
Speech therapy	Targeted programs addressing speech challenges in Parkinson’s	Improves speech quality	Variable outcomes, ongoing therapy may be needed	[8, 23, 42, 16, 61, 62]
Medication	Pharmacological interventions for Parkinson’s symptoms	Alleviates motor symptoms impacting speech	Variable response, side effects, may not target speech specifically	[8, 39, 61]
Surgical interventions	Deep brain stimulation (DBS) for motor symptoms	Shows promise in managing motor symptoms influencing speech	Variable outcomes, invasive with potential risks	[23, 16, 61]
Multidisciplinary approaches	Collaborative care from various specialists	Holistic care tailored to individual needs	Coordination challenges limited access	[8]
Alternative therapies	Creative interventions like singing therapy	Positively impacts communication	Limited evidence, variable responses, accessibility issues	[23]
Adaptive technologies	Technological aids like voice amplifiers, communication apps	Enhances communication	Accessibility and affordability issues, may not address underlying issues	[17, 23, 52]

2 Innovative Approaches in Diagnosing Parkinson's Disease

This section reviews the literature on Parkinson's disease (PD) and its impact on speech and voice, highlighting machine learning algorithms and their diagnostic accuracy for PD-related communication effects. Research from 2020 to 2024, sourced from databases such as IEEE, Springer, Google Scholar, and Science Direct, underpins this analysis.

2.1 *Advancements in Diagnosing Parkinson's Disease: Machine Learning, Speech Assessment, and Acoustic Biomarkers*

Diagnosis improvements for Parkinson's Disease (PD) are focusing more on two systems. This means they combine feature extraction and processes in neat ways. There are popular tools currently in use, such as the Voice Assessment Instrument and the Unified Parkinson's Disease Rating Scale (UPDRS) [6, 17, 18, 34]. These tools integrate Mel Frequency Cepstral Coefficients (MFCCs) with machine learning models like Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs). This work highlights early hypokinetic dysarthria, which is a significant indicator of the disease [15, 35].

Telemonitoring is becoming really useful too! It helps assess PD severity, especially using sustained vowel sounds [36]. A study by [37] lays out a thorough framework called Feature Alignment Methodology (FAM). It looks at speech signals using Mel Frequency Cepstral Coefficients (MFCC), and then it trains a Gaussian Mixture Model (GMM)-Hidden Markov Model (HMM). This is super important for better vocal analysis accuracy for people with PD.

Did you know? Over 90% of PD patients deal with voice issues [38]. Using chi-square ranking alongside optimal subset selection can really boost detection accuracy [39]. Acoustic measures help to tell apart different neurological diseases. Some techniques can reach up to 90% accuracy when looking at vocal disorders related to PD [40, 41]. Early detection gets a boost from various speaking tasks and Voice Onset Time (VOT) analysis [42].

Wavelet Neural Networks (WNNs) and Evolutionary Wavelet Neural Networks (EWNNs) are leading the way. They make it easier to optimize machine learning parameters [39, 43]. New articulatory biomarkers tied to Parkinson's dysarthria show around 85% accuracy [45]. Plus, there is an interesting framework that uses Artificial Neural Networks and SVMs for spotting early signs of PD [2]. These innovations highlight how tech is coming together with voice and speech analysis. More accurate early detection and better disease management are on the horizon.

2.2 Utilizing Acoustic Analysis and Machine Learning for Speech Characterization and Parkinson's Disease Diagnosis

Research into how PD affects speech shows major changes in phonation, articulation, and prosody. Focusing on machine learning along with statistical analysis helps get a clear picture of within-subject variability. These studies really matter for making accurate diagnoses [44, 46–49]. Computer-based methods using different machine learning strategies assist in medical decision-making. They shine especially bright when it comes to diagnosing PD through speech data [50].

Acoustic analysis looks carefully at hypokinetic dysarthria in PD. It explores neural links and how treatments can change speech patterns too [51]. Early stage effects on speech production get studied using short-time power spectrum estimation [52]. Long-term evaluations give us important insights into how the disease progresses [53].

There is also a lot of vocal feature extraction happening! This includes looking at indicators of dysphonia and vocal formant parameters. These studies help us understand how PD affects speech impairments better [7]. Recognizing feature learning techniques plays a big role in early diagnosis for PD [43, 55–57]. When classifying data, SVMs and Random Forests (RF) are used together with hybrid intelligent systems achieving high accuracy in detecting PD [3, 58–60]. This whole approach really shows how technology can help improve diagnosis for those with Parkinson's.

3 Integrating Technological Advancements and Psychosocial Care in Managing Parkinson's Disease Speech Disorders

Managing speech issues related to Parkinson's Disease (PD) is changing a lot, thanks to new and better psychosocial support. A introduces cool things like BioDigit Home and Digit Speech, which use Automatic Speech (ASR) and machine. These help assess speech in PD patients accurately. These tools help doctors diagnose understand the severity of speech disorders. That means better plans for treatment! Technology is growing too! We now have wearable devices, speech therapy apps, and voice recognition tools made just for PD patients. The goal is to facilitate their communication [63, 64]. At the same time, a lot of attention is paid to people with PD's mental and emotional health.

This features how people could confront different discourse difficulties because of PD. Also, Telehealth administrations, Augmentative and Elective Correspondence (AAC) applications, and computer generated reality treatment are truly supporting relational abilities and generally speaking personal satisfaction for people with PD [27, 61]. Changes in how they speak can really affect how they feel inside. It can lead to things like feeling left out or having low self-esteem [19, 21, 26]. But with

the right therapy and technology aimed at these challenges, PD patients can enjoy better social interactions and improved quality of life [4, 16, 20, 33]. Recent studies have shown that voice analysis can effectively tell apart PD patients from healthy individuals. This opens doors to useful screening tools [22, 33].

There is also talk about fresh ways to research these issues using different methods that look closely at changes in the same people over time [65, 66]. Studies exploring how jaw and tongue movement work together in people with PD show high accuracy while spotting speech problems. They show how good simple tech can be here [67]. Efforts focused on boosting the mental well-being of those facing speech challenges—like counseling and community involvement—greatly help improve their self-esteem and communication abilities. This really enhances the quality of life for many [68, 69]. This whole approach shows how important teamwork is between tech and social care when managing speech issues due to PD. Both areas are key in making life better for everyone dealing with this condition.

4 Speech Analysis in Parkinson's Disease: Articulation, Features, and Classification Approaches

The analyzing speech is super important for catching PD early and tracking its progress too! Research has looked at data from new cases as well as early-onset and late-onset PD patients alongside control groups. This helps reveal the tricky speech problems that come with PD. For example, Rusz et al. [70] studied some features like harmonic-to-noise ratio (HNR) and vocal onset time (VOT). They managed to differentiate between 100 newly diagnosed PD patients and healthy controls with an impressive accuracy score of 0.86 for men and 0.93 for women using logistic regression (LR).

The assortment of discourse highlights examined—from essential ones like power and pitch (F0) to more complicated time-recurrence highlights—shows exactly how mind boggling PD-related discourse issues are! Enlists et al. [72] arrived at an incredible close to 100% exactness utilizing convolutional brain organizations (CNN) joined with brief time frame Fourier change (STFT) highlights. In the interim, Mama et al. [76] investigated profound learning strategies on two datasets accomplishing noteworthy exactnesses of 98.49% and 99.67% Clearly, computer based intelligence has exceptional likely in focusing on talk regarding PD (Table 3).

Various assessments by Rusz et al. [72, 73, and 74] as well as Karan et al. [76, 77, and 78] meticulously described specific subjects like hypokinetic dysarthria and the impacts of drug on discourse in PD patients. These assessments shed light on the talk issues related with PD and stress how state-of-the-art assessment can uphold its underlying area and overhaul treatment decisions. Overall, this study gives us a better understanding of how Parkinson's disease affects communication and how effective speech analysis can be in diagnosing and monitoring the disease. It clears the way for more effective patient consideration by explicitly focusing on them.

Table 3 Comprehensive overview of speech and phonation analysis for PD diagnosis

Study focus	Accuracy	Pros	Cons	References
Articulation and features	Up to 99.67%	Offers detailed insights into speech impair ments. Can analyze effects of gender and medication	Requires comprehensive datasets and may involve complex analysis	[70–78]
Classification approaches	Up to 99%	Utilizes advanced machine learning models effectively. High diagnostic accuracy	May incur high computational costs and requires extensive training data	[70, 71 and 75]
Phonation analysis	Up to 95%	Non-invasive with potential for early diagnosis and continuous monitoring	Limited in capturing the full spectrum of speech impairments in PD	[79–84]
Monitoring and severity assessment	Up to 99% for females, 98% for males	Facilitates the development of tools for assessing disease severity over time. Gender-specific analysis enhances accuracy	Dependent on longitudinal data collection and consistency	[83, 84]

5 Advances in Phonation Analysis for Parkinson’s Disease Diagnosis and Monitoring

Phonation analysis has emerged as a crucial, non-invasive technique for assessing Parkinson’s Disease (PD), offering deep insights into its neurological impact on vocal functions. Leveraging datasets like TIMIT, Sakar, and DNSH, and employing features from the Multi-Dimensional Voice Program, baseline characteristics, and advanced dysphonia metrics, researchers have charted the phonatory alterations linked to PD. Quite, Zhang et al. [79] used an inventive AI strategy to recognize PD patients from solid controls with up to 95% exactness on the Sakar dataset and 86.7% on the DNSH dataset, flagging the methodology’s diagnostic potential. Moreover, Vandana et al. [80] found that drug states were connected with head repeat (F0) varieties in male PD patients, featuring the impact that treatment has on vocal attributes. Using a comprehensive set of features and Support Vector Machines (SVM), Tai and others [81] achieved an accuracy of 88 percent on the mPower dataset, demonstrating the value of conventional machine learning. Suphinnapong et al. [82] further shown critical phonatory boundary contrasts between PD patients and controls, giving objective separation markers (Table 4).

Past demonstrative accuracy, phonation investigation has made ready for instruments pointed toward continuous PD checking and seriousness appraisal. Naranjo et al. e.g. [83] used information expansion to adjust sickness movement following,

Table 4 Summary of recent advances in speech and phonation analysis for Parkinson’s disease

Study focus	Key findings	References
Articulation and features	Studies across various datasets show significant potential of machine learning in speech analysis for PD, using features like HNR, VOT, and advanced time-frequency features. Gender-specific differences and the impact of dopaminergic medication were noted	[70–78]
Classification approaches	The application of LR, CNN, and novel deep learning models demonstrated high diagnostic accuracies, highlighting the role of sophisticated classifiers in identifying PD-affected speech	[70, 71 and 75]
Phonation analysis	Employing novel machine learning methods and comprehensive feature sets, studies achieved high accuracies in diagnosing PD from vocal features, indicating the importance of phonation analysis in early diagnosis and monitoring	[79–84]
Monitoring and severity assessment	Phonation features were effectively used for continuous monitoring and assessing disease severity, with gender specific analysis enhancing diagnostic potential	[83, 84]

while Meghraoui et al. [84] demonstrated the value of orientation explicit examination in improving symptomatic exactness by announcing close to amazing acknowledgment rates in identifying PD dysphonia. These developments, which highlight the role of phonation analysis in the development of sophisticated, non-invasive Parkinson’s disease diagnostic and monitoring tools, reinforce the synergy between technology and clinical insights in disease management.

6 Challenges, Solutions, and Future Directions in PD

Diagnosis and Treatment via Speech and Voice Analysis

Early diagnosis and treatment of Parkinson’s Disease (PD) using speech and voice analysis encounter challenges like the nuanced nature of speech impairments and variability in patient responses. Overcoming these requires new diagnostic methods, patient-centric tools, and leveraging emerging technologies.

6.1 Core Challenges

Accurately detecting subtle speech changes—tone, pitch, cadence—as early PD indicators is difficult due to patient variability. Additionally, it is still challenging to locate demonstrative devices, particularly in low-asset regions.

6.2 Strategic Solutions

Improvements include:

- High-level Diagnostics Instruments: -Creating devices that focus too much on resources and are irrelevant, making use of cells for greater accessibility, and paying careful attention.

6.3 Future Directions

Research should:

- Merge Technologies: Analyze the genuine capacities of computerized reasoning (simulated intelligence) in talk examination for novel suggestive and really looking at methodologies. Interdisciplinary Research: Combine experiences in computational semantics, discourse pathology, and nervous system science for a comprehensive comprehension of PD's effects on discourse. Prosperity for the Near and Dear: Undeniable level inclination affirmation developments are urgent for the far reaching organization of PD patients' personal well-being.

6.4 Conclusion

A diverse approach that incorporates technological advancement and cross-disciplinary research is required for using speech analysis to navigate the complexities of PD diagnosis and treatment. These systems vow to further develop PD care, featuring the requirement for continuous innovative work.

References

1. Cai Z et al (2018) An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach. Comput Math Methods Med 2018. <https://doi.org/10.1155/2018/2396952>

2. Parisi L, RaviChandran N, Manaog ML (2018) Feature-driven machine learning to improve early diagnosis of Parkinson's disease. *Expert Syst Appl* 110:182–190. <https://doi.org/10.1016/J.ESWA.2018.06.003>
3. Kodali M, Kadiri SR, Alku P (2024) Automatic classification of the severity level of Parkinson's disease: a comparison of speaking tasks, features, and classifiers. *Comput Speech Lang* 83:101548. <https://doi.org/10.1016/J.CSL.2023.101548>
4. Tunc HC et al (2020) Estimation of Parkinson's disease severity using speech features and extreme gradient boosting. *Med Biol Eng Comput* 58(11):2757–2773. <https://doi.org/10.1007/S11517-020-02250-5>
5. Rusz J, Cmejla R, Ruzickova H, Ruzicka E (2011) Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am* 129(1):350–367. <https://doi.org/10.1121/1.3514381>
6. Cesarini V et al (2023) Voice disorder multi-class classification for the distinction of Parkinson's disease and adductor spasmodic dysphonia. *Appl Sci* 13(15). <https://doi.org/10.3390/APP13158562>
7. Costantini G et al (2023) Artificial intelligence-based voice assessment of patients with Parkinson's disease off and on treatment: machine vs. deep-learning comparison. *Sensors* 23(4). <https://doi.org/10.3390/S23042293>
8. Guatelli R, Aubin V, Mora M, Naranjo-Torres J, Mora-Olivari A (2023) Detection of Parkinson's disease based on spectrograms of voice recordings and extreme learning machine random weight neural networks. *Eng Appl Artif Intell* 125:106700. <https://doi.org/10.1016/J.ENGAPPAI.2023.106700>
9. Amato F, Saggio G, Cesarini V, Olmo G, Costantini G (2023) Machine learning-and statistical-based voice analysis of Parkinson's disease patients: a survey. *Expert Syst Appl* 219:119651. <https://doi.org/10.1016/J.ESWA.2023.119651>
10. Junaid M, Ali S, Eid F, El-Sappagh S, Abuhmed T (2023) Explainable machine learning models based on multimodal time-series data for the early detection of Parkinson's disease. *Comput Methods Programs Biomed* 234:107495. <https://doi.org/10.1016/J.CMPB.2023.107495>
11. Olivares A et al (2023) Perceptual and qualitative voice alterations detected by GIRBAS in patients with Parkinson's disease: is there a relation with lung function and oxygenation? *Aging Clin Exp Res* 35(3):633–638. <https://doi.org/10.1007/S40520-022-02324-4/METRICS>
12. Galaz Z et al (2016) Degree of Parkinson's disease severity estimation based on speech signal processing. In: 2016 39th International Conference on Telecommunications and Signal Processing (TSP) Nov 2016, pp 503–506. <https://doi.org/10.1109/TSP.2016.7760930>
13. Das R (2010) A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Syst Appl* 37(2):1568–1572. <https://doi.org/10.1016/J.ESWA.2009.06.040>
14. Kang K et al (2023) Utilizing speech analysis to differentiate progressive supranuclear palsy from Parkinson's disease. *Parkinsonism Relat Disord* 115:105835. <https://doi.org/10.1016/J.PARKRELDIS.2023.105835>
15. Abedinzadeh Torghabeh F, Hosseini SA, Ahmadi Moghadam E (2023) Enhancing Parkinson's disease severity assessment through voice-based wavelet scattering, optimized model selection, and weighted majority voting. *Med Nov Technol Devices* 20:100266. <https://doi.org/10.1016/J.MEDNTD.2023.100266>
16. Ma A, Lau KK, Thyagarajan D (2020) Voice changes in Parkinson's disease: what are they telling us? *J Clin Neurosci* 72:1–7. <https://doi.org/10.1016/J.JOCN.2019.12.029>
17. Ngo QC, Motin MA, Pah ND, Drotár P, Kempster P, Kumar D (2022) Computerized analysis of speech and voice for Parkinson's disease: a systematic review. *Comput Methods Programs Biomed* 226:107133. <https://doi.org/10.1016/J.CMPB.2022.107133>
18. Dash S (2022) A systematic review of adaptive machine learning techniques for early detection of Parkinson's disease. *Artif Intell Neurol Disord* 361–385. <https://doi.org/10.1016/B978-0-323-90277-9.00018-3>
19. Kumari R, Ramachandran P (2023) A review on early diagnosis of parkinson's disease using speech signal parameters based on machine learning technique. *Lect Notes Electr Eng* 966:217–233. <https://doi.org/10.1007/978-981-19-8338-2-18/COVER>

20. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K (2020) Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech. *Biomed Signal Process Control* 61. <https://doi.org/10.1016/J.BSPC.2020.102050>
21. Rusz J, Tykalova T, Novotny M, Zogala D, Ružićka E, Dušek P (2022) Automated speech analysis in early untreated Parkinson's disease: relation to gender and dopaminergic transporter imaging. *Eur J Neurol* 29(1):81–90. <https://doi.org/10.1111/ENE.15099>
22. Arora S, Baghai-Ravary L, Tsanas A (2019) Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. *J Acoust Soc Am* 145(5):2871–2884. <https://doi.org/10.1121/1.5100272>
23. Di Pietro DA et al (2022) Voice alterations, dysarthria and respiratory derangements in patients with Parkinson's disease. *J Speech Lang Hear Res* 65(10):3749–3757. <https://doi.org/10.1044/2022-jslhr-21-00539>
24. Moro-Velazquez L, Gomez-Garcia JA, Arias-Londono JD, Dehak N, Godino-Llorente JI (2021) Advances in Parkinson's disease detection and assessment using voice and speech: a review of the articulatory and phonatory aspects. *Biomed Signal Process Control* 66. <https://doi.org/10.1016/J.BSPC.2021.102418>
25. Fayad R et al (2021) Vocal test analysis for assessing Parkinson's disease at early stage. *Int Conf Adv Biomed Eng ICABME 2021-October*:171–174. <https://doi.org/10.1109/ICABME53305.2021.9604891>
26. Moro-Velazquez L, Gomez-Garcia J, Dehak N, Godino-Llorente JI (2021) New tools for the differential evaluation of Parkinson's disease using voice and speech processing. <https://doi.org/10.21437/IBERSPEECH.2021-36>
27. Zhang YN (2017) Can a smartphone diagnose Parkinson disease? A deep neural network method and Telediagnosis system implementation. *Parkinsons Dis* 2017. <https://doi.org/10.1155/2017/6209703>
28. Arora S, Lo C, Hu M, Tsanas A (2021) Smartphone speech testing for symptom assessment in rapid eye movement sleep behavior disorder and Parkinson's disease. *IEEE Access* 9:44813–44824. <https://doi.org/10.1109/ACCESS.2021.3057715>
29. Arora S, Tsanas A (2021) Assessing Parkinson's disease at scale using telephone-recorded speech: insights from the Parkinson's voice initiative. *Diagnostics* 11(10). <https://doi.org/10.3390/DIAGNOSTICS11101892>
30. Vitacca M et al (2021) Exercise intolerance and oxygen desaturation in patients with Parkinson's disease: triggers for respiratory rehabilitation? *Int J Env. Res Public Heal* 18(23):12298. <https://doi.org/10.3390/ijerph182312298>
31. Amato F, Borzi L, Olmo G, Artusi CA, Imbalzano G, Lopiano L (2021) Speech impairment in Parkinson's disease: acoustic analysis of unvoiced consonants in Italian native speakers. *IEEE Access* 9:166370–166381. <https://doi.org/10.1109/AC-CESS.2021.3135626>
32. Amato F, Borzi L, Olmo G, Orozco-Arroyave JR (2021) An algorithm for Parkinson's disease speech classification based on isolated words analysis. *Heal Inf Sci Syst* 9. <https://doi.org/10.1007/s13755-021-00162-8>
33. Kadiri SR, Kethireddy R, Alku P (2020) Parkinson's disease detection from speech using single frequency filtering cepstral coefficients. *Proc Annu Conf Int Speech Commun Assoc INTERSPEECH 2020-October*:4971–4975. <https://doi.org/10.21437/INTERSPEECH.2020-3197>
34. Dastgheib ZA, Lithgow B, Moussavi Z (2012) Diagnosis of Parkinson's disease using electro-vestibulography. *Med Biol Eng Comput* 50(5):483–491. <https://doi.org/10.1007/S11517-012-0890-Z>
35. Harel BT, Cannizzaro MS, Cohen H, Reilly N, Snyder PJ (2004) Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment. *J Neurolinguistics* 17(6):439–453. <https://doi.org/10.1016/J.JNEUROLING.2004.06.001>
36. Ali L, Zhu C, Zhang Z, Liu Y (2019) Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network. *IEEE J Transl Eng Heal Med* 7. <https://doi.org/10.1109/JTEHM.2019.2940900>

37. Moro-Velazquez L et al (2019) A forced gaussians based methodology for the differential evaluation of Parkinson's disease by means of speech processing. *Biomed Signal Process Control* 48:205–220. <https://doi.org/10.1016/J.BSPC.2018.10.020>
38. Bru'ckl M, Ghio A, Viallet F (2018) Measurement of tremor in the voices of speakers with Parkinson's disease. *Procedia Comput Sci* 128:47–54. <https://doi.org/10.1016/J.PROCS.2018.03.007>
39. Ali L, Zhu C, Zhou M, Liu Y (2019) Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection. *Expert Syst Appl* 137:22–28. <https://doi.org/10.1016/j.eswa.2019.06.052>
40. Gillivan-Murphy P, Miller N, Carding P (2019) Voice tremor in Parkinson's disease: an acoustic study. *J Voice* 33(4):526–535. <https://doi.org/10.1016/J.JVOICE.2017.12.010>
41. Benba A, Jilbab A, Sandabad S, Hammouch A (2019) Voice signal processing for detecting possible early signs of Parkinson's disease in patients with rapid eye movement sleep behavior disorder. *Int J Speech Technol* 22(1):121–129. <https://doi.org/10.1007/S10772-018-09588-0>
42. Montan'a D, Campos-Roca Y, Pe'rez CJ (2018) A diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease. *Comput Methods Programs Biomed* 154:89–97. <https://doi.org/10.1016/J.CMPB.2017.11.010>
43. Wu K, Zhang D, Lu G, Guo Z (2018) Learning acoustic features to detect Parkinson's disease. *Neurocomputing* 318:102–108. <https://doi.org/10.1016/J.NEUCOM.2018.08.036>
44. Naranjo L, Pe'rez CJ, Mart'ın J, Campos-Roca Y (2017) A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. *Comput Methods Programs Biomed* 142:147–156. <https://doi.org/10.1016/J.CMPB.2017.02.019>
45. Godino-Llorente JI, Shattuck-Hufnagel S, Choi JY, Moro-Vela'zquez L, Go'mez-Garc'ia JA (2017) Towards the identification of Idiopathic Parkinson's disease from the speech. New articulatory kinetic biomarkers. *PLoS One* 12(12). <https://doi.org/10.1371/JOURNAL.PONE.0189583>
46. Bocklet T, Steidl S, No'th E, Skodda S (2013) Automatic evaluation of Parkinson's speech—acoustic, prosodic and voice related cues. *Proc Annu Conf Int Speech Commun Assoc INTERSPEECH* 1149–1153. <https://doi.org/10.21437/INTERSPEECH.2013-313>
47. Postuma RB, Lang AE, Gagnon JF, Pelletier A, Montplaisir JY (2012) How does Parkinsonism start? Prodromal Parkinsonism motor changes in idiopathic REM sleep behaviour disorder. *Brain* 135(6):1860–1870. <https://doi.org/10.1093/BRAIN/AWS093>
48. Orozco-Arroyave JR et al (2016) Automatic detection of Parkinson's disease in running speech spoken in three different languages. *J Acoust Soc Am* 139(1):481–500. <https://doi.org/10.1121/1.4939739>
49. Benba A, Jilbab A, Hammouch A (2016) Analysis of multiple types of voice recordings in cepstral domain using MFCC for discriminating between patients with Parkinson's disease and healthy people. *Int J Speech Technol* 19(3):449–456. <https://doi.org/10.1007/S10772-016-9338-4>
50. Gu'ru'ler H (2017) A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method. *Neural Comput Appl* 28(7):1657–1666. <https://doi.org/10.1007/S00521-015-2142-2>
51. Brabenec L, Mekyska J, Galaz Z, Rektorova I (2017) Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *J Neural Transm* 124(3):303–334. <https://doi.org/10.1007/S00702-017-1676-0>
52. Upadhyia SS, Cheeran AN, Nirmal JH (2018) Thomson multitaper MFCC and PLP voice features for early detection of Parkinson disease. *Biomed Signal Process Control* 46:293–301. <https://doi.org/10.1016/J.BSPC.2018.07.019>
53. Arias-Vergara T, Va'squez-Correa JC, Orozco-Arroyave JR, No'th E (2018) Speaker models for monitoring Parkinson's disease progression considering different communication channels and acoustic conditions. *Speech Commun* 101:11–25. <https://doi.org/10.1016/j.specom.2018.05.007>

54. Jankovic J (2008) Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry* 79(4):368–376. <https://doi.org/10.1136/JNNP.2007.131045>
55. Verde L, De Pietro G, Sannino G (2018) Voice disorder identification by using machine learning techniques. *IEEE Access* 6:16246–16255. <https://doi.org/10.1109/AC-CESS.2018.2816338>
56. Arias-Vergara T, Vázquez-Correa JC, Orozco-Arroyave JR (2017) Parkinson's disease and aging: analysis of their effect in phonation and articulation of speech. *Cognit Comput* 9(6):731–748. <https://doi.org/10.1007/S12559-017-9497-X>
57. Kim Y (2017) Acoustic characteristics of fricatives /s/ and /ʃ/ produced by speakers with parkinson's disease. *Clin Arch Commun Disord* 2(1):7–14. <https://doi.org/10.21849/CACD.2016.00080>
58. Benba A, Jilbab A, Hammouch A (2016) Voice assessments for detecting patients with Parkinson's diseases using PCA and NPCA. *Int J Speech Technol* 19(4):743–754. <https://doi.org/10.1007/S10772-016-9367-Z>
59. Hariharan M, Polat K, Sindhu R (2014) A new hybrid intelligent system for accurate detection of Parkinson's disease. *Comput Methods Programs Biomed* 113(3):904–913. <https://doi.org/10.1016/J.CMPB.2014.01.004>
60. Arora S et al (2015) Detecting and monitoring the symptoms of Parkinson's disease using smart-phones: a pilot study. *Park Relat Disord* 21(6):650–653. <https://doi.org/10.1016/J.PAR.KRELDIS.2015.02.026>
61. Chenausky K, MacAuslan J, Goldhor R (2011) Acoustic analysis of PD speech. *Parkinsons Dis*. <https://doi.org/10.4061/2011/435232>
62. Hartelins L, Svensson P (1994) Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: a survey. *Folia Phoniatr Logop* 46(1):9–17. <https://doi.org/10.1159/000266286>
63. Borz'i L et al (2019) Home monitoring of motor fluctuations in Parkinson's disease patients. *J Reliab Intell Environ* 5(3): 145–162. <https://doi.org/10.1007/S40860-019-00086-X>
64. Cantu'rk I, Karabiber F (2016) A machine learning system for the diagnosis of Parkinson's disease from speech signals and its application to multiple speech signal types. *Arab J Sci Eng* 41(12):5049–5059. <https://doi.org/10.1007/S13369-016-2206-3>
65. Naranjo L, Pe'rez CJ, Mart'ın J (2017) Addressing voice recording replications for tracking Parkinson's disease progression. *Med Biol Eng Comput* 55(3):365–373. <https://doi.org/10.1007/S11517-016-1512-Y>
66. Naranjo L, Pe'rez CJ, Campos-Roca Y, Mart'ın J (2016) Addressing voice recording replications for Parkinson's disease detection. *Expert Syst Appl* 46:286–292. <https://doi.org/10.1016/J.ESWA.2015.10.034>
67. Zhang HH et al (2016) Classification of Parkinson's disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples. *Biomed Eng Online* 15(1). <https://doi.org/10.1186/S12938-016-0242-6>
68. Go'mez-Vilda P et al (2017) Parkinson disease detection from speech articulation neuromechanics. *Front Neuroinform* 11. <https://doi.org/10.3389/FNINF.2017.00056>
69. Van Den Eeden SK et al (2003) Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. *Am J Epidemiol* 157(11):1015–1022. <https://doi.org/10.1093/AJE/KWG068>
70. Rusz J, Tykalova' T, Novotny' M, Ru'z'ic'ka E, Dus'ek P (2022) Automated speech analysis in early untreated Parkinson's disease: relation to gender and dopaminergic transporter imaging. *Eur J Neurol* 29(1):81–90
71. Hires' M, Gazda M, Drota'r P, Pah ND, Motin MA, Kumar DK (2022) Convolutional neural network ensemble for Parkinson's disease detection from voice recordings. *Comput Biol Med* 141:105021
72. Senthinathan A, Adams S, Page AD, Jog M (2021) Speech intensity response to altered intensity feedback in individuals with Parkinson's disease. *J Speech Lang Hear Res* 64(6s):2261–2275
73. Rusz J, Tykalova' T, Novotny' M, Zogala D, Sonka K, Ruzicka E, Dusek P (2021) Defining speech subtypes in de novo Parkinson disease: response to long-term levodopa therapy. *Neurology* 97(21):e2124–e2135

74. Rusz J, Hlavnic̣ka J, Novotný M, Tykalova T, Pelletier A, Montplaisir J, Gagnon JF, Duṣek P, Galbiati A, Marelli S, Timm PC, Teigen LN, Janzen A, Habibi M, Stefani A, Holzknecht E, Seppi K, Evangelista E, Rassu AL, Dauvilliers Y, Họgl B, Oertel W, St Louis EK, Ferini-Strambi L, Rụz'ic̣ka E, Postuma RB, Ṣonka K (2021) Speech biomarkers in rapid eye movement sleep behavior disorder and Parkinson disease. *Ann Neurol* 90(1):62–75
75. Ma J, Zhang Y, Li Y, Zhou L, Qin L, Zeng Y, Wang P, Lei Y (2021) Deep dual-side learning ensemble model for Parkinson speech recognition. *Biomed Signal Process Control* 69 Article 102849
76. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K (2021) Non-negative matrix factorization-based time-frequency feature extraction of voice signal for Parkinson's disease prediction. *Comput Speech Lang* 69:101216
77. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K (2020) Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech. *Biomed Signal Process Control* 61:102050
78. Karan B, Sahu SS, Mahto K (2020) Parkinson disease prediction using intrinsic mode function based features from speech signal. *Biocybern Biomed Eng* 40(1):249–264
79. Zhang X, Ma J, Li Y, Wang P, Liu Y (2021) Few-shot learning of Parkinson's disease speech data with optimal convolution sparse kernel transfer learning. *Biomed Signal Process Control* 69, Article 102850
80. Vandana VP, Darshini JK, Vikram VH, Nitish K, Kumar PP, Ravi Y (2021) Speech characteristics of patients with Parkinson's disease-does dopaminergic medications have a role? *J Neurosci Rural Pract* 12(4):673–679
81. Tai YC, Bryan PG, Loayza F, Pela'ez E (2021) A voice analysis approach for recognizing Parkinson's disease patterns. *IFAC-PapersOnLine* 54(15):382–387
82. Azadi H, Akbarzadeh TM, Shoeibi A, Kobravi HR (2021) Evaluating the effect of Parkinson's disease on jitter and shimmer speech features. *Adv Biomed Res* 10 Article 54
83. Naranjo L, Pe'rez CJ, Campos-Roca Y (2021) Monitoring Parkinson's disease progression based on recorded speech with missing ordinal responses and replicated covariates. *Comput Biol Med* 134 Article 104503
84. Meghraoui D, Boudraa B, Merazi-Meksen T, Go'mez Vilda P (2021) A novel pre-processing technique in pathologic voice detection: application to Parkinson's disease phonation. *Biomed Signal Process Control* 68 Article 102604

An Efficient Approach for Decentralized Web Messaging Using WebRTC



Manisha Kasar, Trupti Suryawanshi, Pranoti Kavimandan,
Snehal Chaudhari, Shreekala Deshmukh, and Pramod Jadhav

Abstract Decentralized messaging is a system of communication where the messages are sent and received without any base of operations that dominates the flow of the messages. However, the network is established on a distributed system, where nodes are used to help store as well as transmit data. In comparison with the centralized systems, the decentralized messaging has stronger security and privacy features. The decentralized messaging systems cannot be easily censored since there is no single point of control, which can simply block or limit the access to the system. Before that, blockchain was utilized and it was found to have disadvantages as will be discussed herewith; for example, blockchains may suffer from high latency which implies that it takes a greater number of moments before one gets to confirm the transaction and get it incorporated into the chain. It can be costly, time consuming, may require much computation and power to sustain the network. Because of the issues present in centralized messaging, we are using decentralized messaging through Web Real-Time Communication WebRTC a system that allows real-time direct communication between users over the Internet with no central server. As in the case of WebRTC for decentralizing the messaging system, it allows the users to just directly

M. Kasar (✉) · T. Suryawanshi

Department of Computer Science and Engineering, Bharati Vidyapeeth (Deemed to Be University) College of Engineering, Pune, India
e-mail: mmkasar@bvucoep.edu.in

T. Suryawanshi

e-mail: tv Suryawanshi@bvucoep.edu.in

P. Kavimandan · P. Jadhav

Department of Computer Science and Business Systems, Bharati Vidyapeeth (Deemed to Be University) College of Engineering, Pune, India
e-mail: pskavimandan@bvucoep.edu.in

S. Chaudhari · S. Deshmukh

Department of Information Technology, Bharati Vidyapeeth (Deemed to Be University) College of Engineering, Pune, India
e-mail: sdchaudhary@bvucoep.edu.in

S. Deshmukh

e-mail: smdeshmukh@bvucoep.edu.in

converse with each other using their browser or mobile application. WebRTC has extremely low latency which is a much better and secure method of how user can communicate in real time.

Keywords Real-time communication · Networking · Socket.io · Node.js

1 Introduction

Everyday, people communicate globally with the use of the Internet, as communication is an integral part of their lifestyle. Today, different chat systems work on centralized systems, which means all the data is centralized. Therefore, causing a major problem that is if the central server goes down then the entire network fails, and due to this major drawback is that there can be a loss of user's information and resources which is stored on the centralized server or even there can be a leak of user's chat information that is stored on the server [1]. The enormous benefits such as improved privacy, anonymity, security, transparency, integrity, lack of third parties influenced us to research this and come up with a better and efficient solution. The integration of WebSocket, WebRTC, and Socket.IO can enhance the web chat experience by providing real-time and interactive communication between customers and chat agents. WebSocket is a protocol that helps in bi-directional communication between the client and the server [2] It thus does away with the requirement of the continuous exchange of HTTP request and response that interrupts constant communication between the client and server. WebRTC is a technology seen as a platform that enables peer-to-peer real time and high-quality audio, and video communications within a browser [3]. It offers an operator-client real-time multimedia communication, where customers and chat agents can share voice and video calls inside the web chat [4, 5]. Socket. IO library provides synchronized client communication with a server in real time. It is used to perform the exchange of information between the customer and the chat agent and creates a communication medium that allows the transfer of text messages in real time. The idea will be to integrate WebSocket, WebRTC, and Socket into a single application. With IO, we can provide our customers with real time and interactive web chat. It can also support voice and video chat, history of the chat session and customer details that help provide efficient and effective customer service [6, 7]. The suggested work has the following contributions:

1. The suggested approach creates a direct end-to-end encrypted, peer-to-peer connection between the web browsers of users created using RTC peer connection API.
2. The proposed approach is tested on a TURN server provided by Open Relay Project as a fallback mechanism when users cannot connect directly, here REST API is used.
3. The proposed approach makes use of websockets to connect to the ICE server hosted on render.com to get active users list.

4. Reliable session establishment capability works on multiple platforms and devices,
5. It works reliably on real-time data channels.

The following sub-sections of the paper is organized as follows. Section 2 discusses about the state-of-the-art study done by the researchers, while Sect. 3 is focused on the proposed methodology with different frameworks and technologies used to create the decentralized system architecture. Comparative analysis of earlier methods and the proposed model is discussed in Sect. 4. Section 5 concludes the work and is followed by the future scope in the last Sect. 6.

2 Literature Review

Decentralization would involve exploring the problems with centralized communication systems, such as the centralization of power and the risk of data breaches. In addition, it would involve how decentralized messaging aims to address these issues. A literature review of decentralized web messaging would likely include a discussion of its potential benefits and applications, such as enhancing privacy and security for users, as well as its potential to support censorship-resistant communication and freedom of expression. The review would likely cover the challenges facing decentralized messaging, such as scalability, usability, and adoption, and discuss ongoing study and implementation efforts aimed at decreasing these hurdles. Scholars have tried to develop Decentralized Web Messaging. For instance, Abhishek and Chaitanya proposed a decentralized chat application using blockchain technology. In their framework, overall user information is collected on a block interconnected to other blocks creating a chain. Additionally, if a hacker tries to modify the information in the block, he or she will have to modify all copies of that block on the whole blockchain network, which can be quite challenging [8]. This application uses the INFURA API for the deployment of smart contracts on the Ethereum blockchain network [9].

Mohamad Abdullah proposed a fully secure individual user database method which was the objective of the work. A security element was built into user registration, user login, message exchange, and database storage [10]. Uma et al. have worked on a decentralized chat application by using Gun.js, a decentralized graph database that is real-time has low latency, and also has security, encryption, and authorization for the data [11]. GABRIEL ELLEBRINK's proposed Open and Non-proprietary Decentralized Messaging Protocol states that it presents a decentralized messaging protocol that can operate without a centralized entity in charge. He stated that the study has shown that the presented messaging protocol is relatively inexpensive to operate, particularly compared to using the Ethereum blockchain [12]. Khalkar and Dhake suggested a topic titled 'About the Investigation on the Decentralized Chat Applications using Blockchain Technology'. A decentralized application approach (dApp) is used by child companies, which stores user data in blocks that are linked

together. Furthermore, it incorporates highly secure encryption and hash functions, which make data stored in the blocks virtually invisible [13].

As given in Table 1, the authors have tried to compare earlier proposed methods that are related to centralized and decentralized approaches. In addition to secure messaging, blockchains for decentralized messaging, blockchain messaging applications in education, and WebRTC/Web Sockets for decentralized messaging, this book discusses issues like decentralized messaging using WebRTC and Web Sockets. The following table contains an index of the main points about the development of a decentralized chat application presented in the provided references. It brings out the main issues of concern discussed in each of the research papers. However, the authors have not used some issues that are associated with the transfer of real-time data, efficiency, data integrity, and the amount of computational resource that are required in the blockchain technologies. Thus, to remove all the restrictions pointed out, it is necessary to propose a computationally less expensive method. The suggested approach is defined after the section which is presented following.

3 Proposed Methodology

The suggested work erases the main issue of central servers that are inherent to messengers such as Instagram and WhatsApp. We make use of peer-to-peer (P2P) connection with web real-time communication (WebRTC) for direct and secure data sharing between users. A central signaling server in the network has a relatively small responsibility at the start of the connection and contains the agreement about how traffic is to be routed. When connected the users are able to send data directly to another user directly through web real-time communication webrtc through the browser. For the purpose of identifying the public IP addresses of users, who are behind a Network Address Translator (NAT), a STUN server is used. That is why it makes much sense to decentralize the mechanism and develops the functions for users and privacy which is maximally deprived when the central server is the only key point of interest.

Figure 1 describes the working process of the suggested approach. The secure web chat application starts with user login/signup and configuration. Upon initiating a chat session, a signaling server facilitates peer discovery. If users are behind a NAT, a STUN server helps to find their public IP addresses. When possible RTConnect is made to the other peer in case information will have to be relayed by a TURN server. On a successful establishment of the connections, several data of RTC are used in the establishment of a real-time communication. The first part deals with the connection and the second part are the disconnection that gives message to user regarding connection complications and the last part is the logout of the user.

The web sockets are used for client-sever data transfer since it provides a full-duplex communication channel and unique identifiers for peers are implemented using a Js library called uuidv4. The centralized server does not store any data permanently but only relies on the new nodes (users). Data between clients is transferred

Table 1 Comparative analysis of earlier proposed methods

Ref.	Method	Advantages	Disadvantages
[14]	Comparison of centralized and decentralized systems	Explains centralized and decentralized architectures	Not specific to chat applications and efficiency is poor
[15, 16]	WebRTC based video conferencing system	Supports virtual learning, peer-to-peer conference sessions	Not specifically for decentralized chat and not applicable for real-time data transfer
[17]	WebRTC design patterns	Improves scalability, performance, and efficiency of WebRTC applications	Not specific to chat applications and accuracy is poor
[18]	WebRTC signaling via web sockets	Provides foundation for WebRTC signaling	Not a complete chat application and integrity is poor
[19]	Framework for a decentralized web	Solves data integrity, centralization, and control issues	Not specific to chat applications
[20, 21]	Real-time communication over blockchain	Enables secure news and media sharing	Not a complete chat application and efficiency is also poor
[22]	Kademlia network into WebRTC	Reduces need for centralized signaling service	Not a complete chat application and unable to access real-time data
[23]	WebRTC chat and video communication	Peer-to-peer technology for real-world application	Not specific to decentralized chat
[24]	Anonymous decentralized ephemeral chat application	Enables anonymous communication using IPFS	Not able to access real-time data also efficiency is poor
[25]	Secure messaging	Evaluates framework security, ease of use, and adoption	Not a complete chat application
[26]	Secure instant messaging design	Secure instant messaging application with privacy	Not specific to decentralized chat
[27]	Security agreement for decentralized secure group messaging	Implements secure group messaging in decentralized networks	Not a complete chat application also accuracy is poor
[28]	Blockchain for decentralization of internet	Decentralized methods with different internet features	Not specific to chat applications
[29]	Exploration of blockchain technology in education	Potential application of blockchain in education	Not specific to chat applications also unable to access real-time data

(continued)

Table 1 (continued)

Ref.	Method	Advantages	Disadvantages
[30]	Decentralized approach to messaging using blockchain technology	• Uses Ethereum smart contracts for secure messaging	• Large amount of computational time require
[31–33]	Decentralized web messaging system using WebRTC	• Peer-to-peer network for establishing connections	• Efficiency is poor
[34]	fybrChat: a distributed chat application for secure P2P messaging	• Secure peer-to-peer messaging	• Large amount of computational time require and accuracy is poor
[35]	Web sockets	• Real-time communication between web browsers and servers	• Not specific to chat applications
[36]	Multi-user web chat application using Node.js and Socket.io	• Create rooms and chat with anyone	• Not specific to decentralized chat
[37]	WebRTC APIs	• Provides multimedia transmission services	• Not a complete chat application also large amount of computational time require

through JSON objects. For implementing video calls we use socket.io providing a real-time peer-to-peer connection. Moreover, it also provides security against any malicious attacks as well as data leaks which have become common in recent years.

3.1 Architecture

The architecture is given in Fig. 2. This figure depicts how clients discover and connect with the help of a signaling server.

3.2 Algorithm of Proposed Work

1. **Login/Configure:** User logs in, sets preferences (messaging/video call).
2. **Peer Discovery:** Signaling server finds active users and facilitates connection. STUN helps discover public IP if behind NAT.
3. **Connection & Data:**
 - Direct P2P connection (WebRTC) for data transfer (if possible).
 - TURN server as backup relay (if direct connection fails).
 - WebSockets for client–server communication (full-duplex channel).
 - Data format: JSON objects.

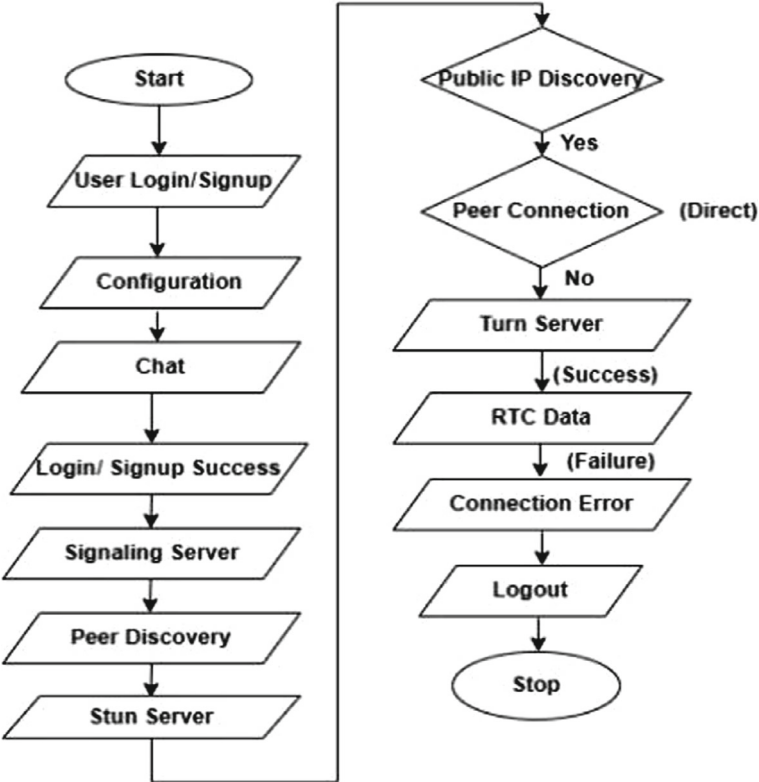


Fig. 1 Working flow of proposed system

Fig. 2 System architecture

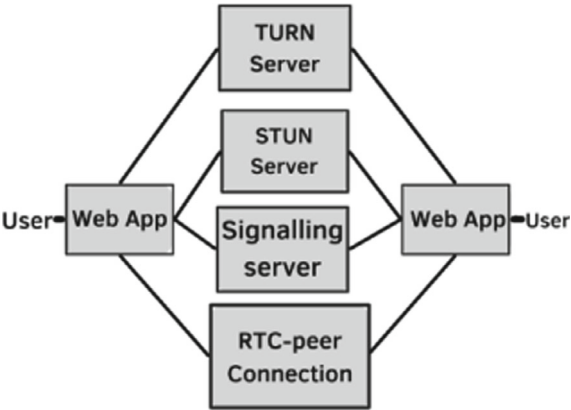


Fig. 3 CORS

```
cors: {
  origin: "https://dmessagevc.onrender.com",
  methods: ["GET", "POST"],
  allowedHeaders: "*"
},
```

4. **Communication:** Real-time communication with error handling.
5. **Logout:** User logs out after the session.

3.3 Signaling Server

The task of finding active users is done by a signaling server which acts as a middleman to establish a connection. Here we use a WebSocket to implement this server where the ICE candidates transmit signaling information as JSON objects to the server. The Cross-Origin Resource Sharing (CORS) is implemented to prevent malicious attackers from assessing the information shown in Fig. 3.

3.4 Session Traversal Utilities for Nat (STUN) Server

As most of the clients are behind NAT, their public IP address is unknown; this task of discovering the public facing IP address of these devices is done by STUN server that returns the IP address and port. The author implemented a public STUN server provided by Google as shown in Fig. 4.

```
const configuration = {
  iceServers: [
    { url: "stun:stun.1.google.com:19302" }, //stun server
    {
      url: "turn:turn.anyfirewall.com:443?transport=tcp",
      credential: " ",
      username: " ",
    }, //turn server
  ],
};
```

Fig. 4 STUN, TURN configuration

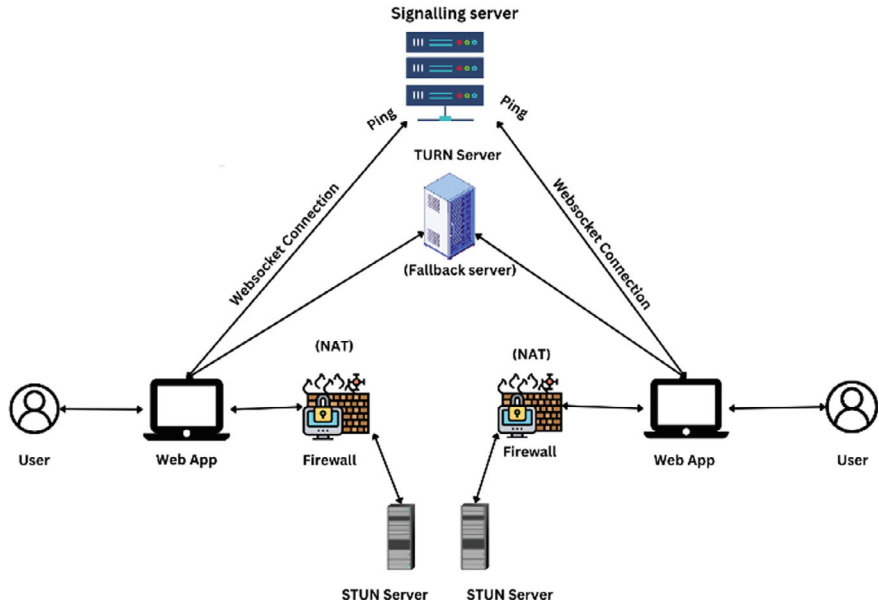


Fig. 5 Network architecture

3.5 Traversal Using Relays Around Nat (TURN) Server

This server is used as a safe fallback mechanism where the connection cannot be established. In such scenarios, turn server is used to relay the data between peers. A private turn server is implemented for this as shown by Fig. 3.

Figure 5 shows a secure web chat application’s architecture. It facilitates direct communication between users via peer-to-peer (P2P) connections. A signaling server helps establish these connections, while a STUN server aids in finding user IP addresses if behind a NAT. If a direct P2P connection isn’t possible, a TURN server acts as a backup relay. Websockets enable real-time data exchange throughout the process. This design prioritizes secure and efficient P2P communication.

3.6 Rtc Peer Connection

Each user logged in on the web app generates Interactive Connectivity Establishment (Ice) candidates and sends it. These Ice candidates have the configuration information like protocol and routing data necessary to transmit data to peers. It is compatible with all the widely used browsers making it reliable to use. These Ice candidates are then to create an RTC peer connection that acts as a peer-to-peer webRTC connection

interface between the local device and the remote peer. It provides various methods for connection establishment and connection status.

4 Results and Discussion

The experiments conducted on decentralized web messaging using WebRTC generally focus on evaluating the performance and feasibility of using WebRTC for such systems. Here are some of the key findings from these experiments:

Low latency and high reliability: WebRTC-based decentralized messaging systems offer low latency and high reliability due to their peer-to-peer communication architecture. This means that messages are delivered quickly and reliably without requiring a centralized server. Due to peer-to-peer communication, WebRTC-based messaging systems consume less network bandwidth than traditional client–server architectures. This can be a significant advantage for users who have limited bandwidth.

Scalability: WebRTC-based messaging systems can scale well since they do not require a centralized server to maintain communication. This means that as the number of users increases, the system can still maintain its performance. Since WebRTC can be used with most modern web browsers and mobile devices, it can be used to develop cross-platform decentralized messaging systems.

Privacy and security: WebRTC provides end-to-end encryption for communication, making it suitable for developing secure and private messaging systems.

Figure 6 shows snapshot of RTC peers unique identifier RTC peers unique identifier. This secure chat app uses P2P connections for direct communication. A signaling server helps initiate connections, while STUN and TURN servers provide backup mechanisms for complex network scenarios. Figure 7 depicts a secure web chat app's architecture. It prioritizes direct communication between users via P2P connections.

Figure 8 shows a secure web chat app's architecture. It uses a signaling server to initiate direct connections between users for real-time chat. If a direct connection can't be established, STUN and TURN servers help route data. Figure 9 shows video chat paper of proposed approach. It prioritizes direct communication between users via P2P connections. A signaling server helps establish these connections, while STUN and TURN servers act as backups for situations where a direct P2P connection isn't possible. Overall, the experiments conducted on decentralized web messaging using WebRTC demonstrate that WebRTC is a promising technology for developing decentralized messaging systems. The performance characteristics of WebRTC-based messaging systems make them suitable for use in a wide range of applications, including chat applications, file-sharing applications, and more. The security and privacy features of WebRTC also make it a suitable technology for developing secure and private messaging systems.

```
PROBLEMS  OUTPUT  TERMINAL  DEBUG CONSOLE

Terminate batch job (Y/N)? y
PS A:\DMessage_VC_CHAT\Server> npm start

> VideoChat@1.0.0 start
> node server.js

VC server is running on port 5000
Cw0sH19PjLP6BIT5AAAB
I2GXa8z4PkjLYWnrAAAD
kR9uBDzKzYYWqRnfAAAF
FlsBAKvEw_r9h3M6AAAH
OELb1GddYzPzSiLAAAJ
█
```

Fig. 6 RTC peers unique identifier

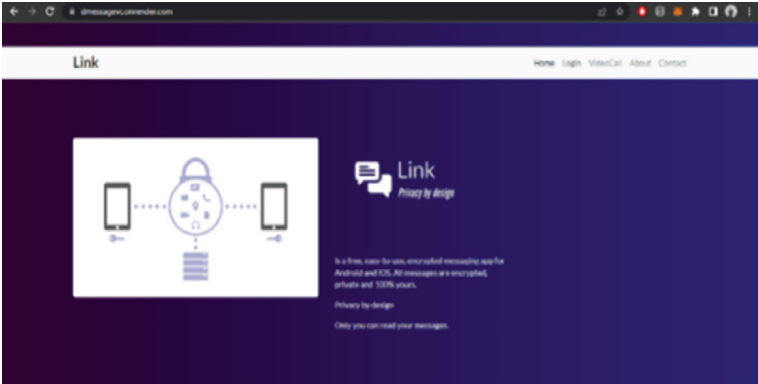


Fig. 7 Homepage

5 Comparative Analysis

Table 2 explains the comparative analysis of several methods and components of different research papers. The first column shows the paper number also called as the reference number of the paper. The second column lists the titles of research papers published by various authors. The third column highlights the benefits of the methodologies discussed in these papers, often focusing on unique features that set them apart from others. Finally, the fourth column outlines the limitations or drawbacks of each paper’s proposed approach, noting areas where improvements could be made with more advanced frameworks and technologies.

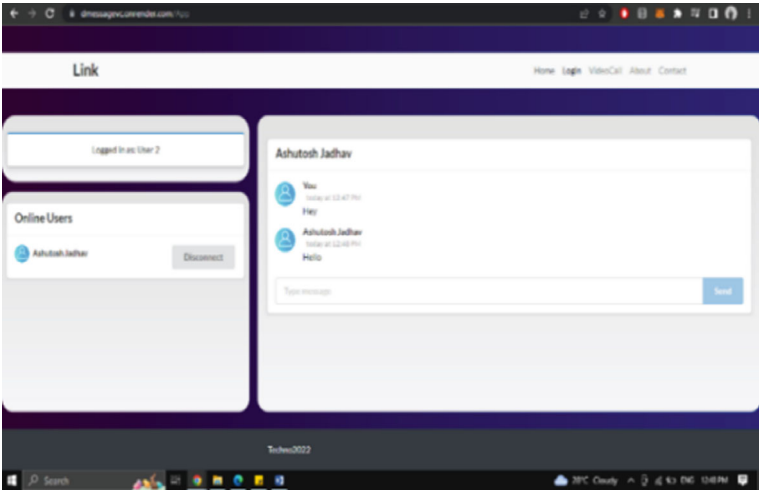


Fig. 8 Messaging page

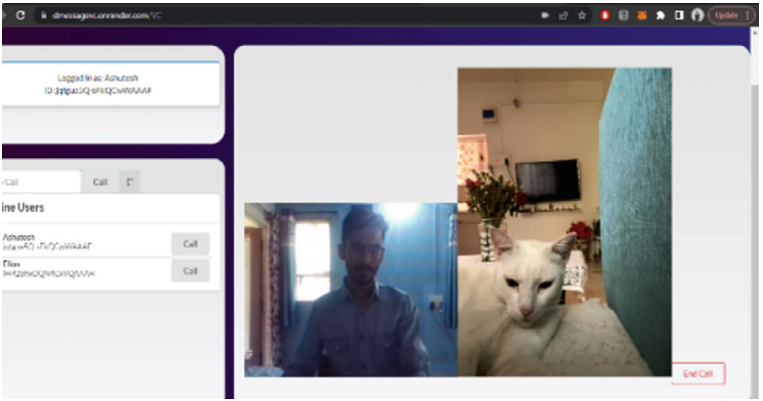


Fig. 9 Video call page

Figure 10 shows comparison on different papers which uses blockchain and Ethereum to implement messaging which takes up more transfer times as compared to WebRTC and Socket.io framework used in our methodology making it more efficient.

Table 3 explains the comparative analysis of several parameters that are based on centralized and decentralized server. The first column represents the parameters based on which differences are obtained. The second column illustrates the how using a decentralized network provides the mentioned advantages. The third column depicts the cons of using the centralized network. However, decentralized web messaging offers a secure, censorship-resistant alternative to traditional chat with user control at its core. Nevertheless, scalability, usability, and adoption remain an issue, its potential

Table 2 Comparative analysis

Ref	Name of the paper	Advantages	Limitation(s)
[8]	Decentralized chat application using blockchain technology	It uses secure encryption and hashing algorithms	It uses secure encryption and hashing algorithms
[9]	bChat: a decentralized chat application	The stored data can be encrypted making it impervious to theft	It is expensive to store data on the Ethereum blockchain
[11]	Decentralized chat application	Robust and secure peer-to-peer data storage	Limitation of svelte is lack of IDE support
[30]	A decentralized approach to messaging using blockchain technology	1. Unlimited storage 2. Server free zone	Robust and secure peer-to-peer data storage
[13]	Decentralized secure messaging application using blockchain technology	The list of active users is stored in the form of a local hash table containing the Ip addresses and port number of the connected users	The list of active users is stored in the form of a local hash table containing the Ip addresses and port number of the connected users
[10]	A secure chat application based on pure peer-to-peer architecture	User registrations are done among themselves. A two way authentication protocol is implemented	Files and folders cannot be centrally backed up
Proposed work	Proposed methodology decentralized web messaging using webRTC and socket.io	Reliable session establishment Works on multiple platforms and devices. Robust and reliable real-time data channels	When WebRTC is deployed without additional technologies, scale, and quality are limited

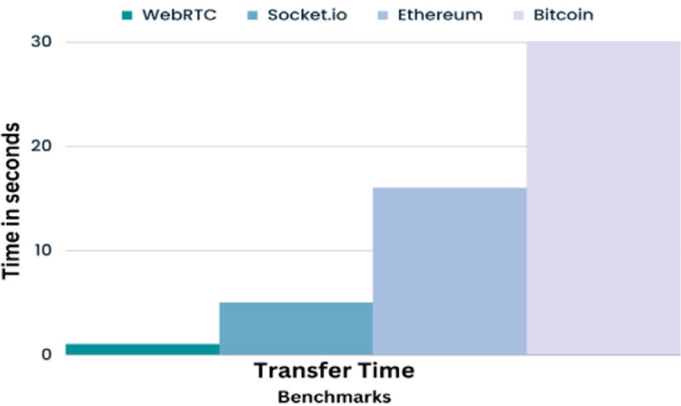


Fig. 10 Comparison between Socket.IO versus Blockchain versus Ethereum

Table 3 Comparison between centralized and decentralized networks

Parameters	Decentralized	Centralized
Security	Store data in a network of nodes with everyone having equal access	Centralized systems are easier to hack since they store data in a single central storage
Scalability	Easier to scale since they leverage scalability through network coordination	Centralized systems are harder to scale as or rely on single-point resource management
Control	Controlled by the users	Controlled by only one authority. This can result in biased decisions
Speed	Faster because network administrations create master nodes when user activity is high	Centralized systems are relatively slower as even during high user authority is responsible to maintain the system
Cost	Relatively cheaper to set up since they don't require a large team. Maintenance cost is high	Costlier to setup since they require a large team

for a more open communication future is undeniable. Advancements in scalability and persistent data storage can further empower this technology.

6 Conclusion and Future Scope

Decentralized web messaging is a burgeoning approach to online communication, empowering users by shifting control away from centralized servers. This technology prioritizes privacy, security, and resistance to censorship, offering a more open and democratic communication platform. However, challenges like scalability, usability, and user adoption need to be addressed. Despite these hurdles, the potential advantages are significant, suggesting a pivotal role for decentralized messaging in the future of online communication. Future advancements such as more scalable systems and persistent data storage via session instances or decentralized databases can further enhance its capabilities.

References

1. Ray PP (2023) Web3: a comprehensive review on background, technologies, applications, zero-trust architectures, challenges and future directions. *Internet Things Cyber-Phys Syst* 3:213–248. <https://doi.org/10.1016/j.iotcps.2023.05.003>
2. Berg DR et al (2022) WebRTC-based decentralized chat application with minimal latency. In: *International conference on intelligent innovations in engineering and technology (ICIET)*. Coimbatore, India, pp 210–215. <https://doi.org/10.1109/ICIET55458.2022.9967597>
3. Blum N, Lachapelle S, Alvestrand H (2021) WebRTC: real-time communication for the open web platform. *Commun ACM* 64(8):50–54. <https://doi.org/10.1145/3453182>

4. Shen H, Li J, Wang Q, Zhang W (2021) A survey on peer-to-peer communication technologies and applications. *IEEE Commun Surv Tutor* 20(3):2068–2101. <https://doi.org/10.1109/COMST.2018.2817050>
5. Nikose MS, Dosani S, Pardhi S, Nikode D, Jais A (2023) Real-time chat application. *Int J Sci Res Sci Eng Technol* 10(2):496–501(2023). <https://doi.org/10.32628/IJSRSET2310267>
6. Mardan A (2018) Real-time apps with WebSocket, Socket.IO, and DerbyJS: building real-world scalable web apps. *Pract Node js*. 307–330. https://doi.org/10.1007/978-1-4842-3039-8_9
7. Akyildiz IF, Wang X (2023) WebSockets: real-time communication protocol for web applications. *Comput Netw* 55(12):2702–2713. <https://doi.org/10.1016/j.comnet.2011.05.010>
8. Takale AP, Vaidya CV, Suresh S (2019) Decentralized chat application using blockchain technology. *Int J Res Eng Appl Manag Special issue*:2454–9150
9. Sourabh DR, Kapkoti K, Aggarwal S, Khanna A (2020) bChat: a decentralized chat application. *Int Res J Eng Technol* 07(05):2775–2780
10. Mohamad Afendee M et al (2015) A secure chat application based on pure peer-to-peer architecture. *J Comput Sci* 11(5):723–729. <https://doi.org/10.3844/jcssp.2015.723.729>
11. Thakur U, Chichmalkar A, Sambhare A, Chaturvedi A, Khuspare C (2022) Decentralized chat application. *Int Res J Eng Technol* 09(02):249–253
12. Ellebrink G (2022) An open and nonproprietary decentralized messaging protocol. *Int Res J Mod Eng Technol Sci*. Independent thesis Advanced level. 91. <https://doi.org/10.56726/IRJ METS30138>
13. Khalkar K, Dhake N, Tejas S (2023) Decentralized chat application using blockchain technology. *Int J Res Appl Sci Eng Technol* 11:813–816. <https://doi.org/10.22214/ijraset.2023.48650>
14. Sateesh N, Chandhana N, Sireesha P (2022) A trustworthy decentralized internet: an application of blockchain. *Turk J Comput Math* 13(03):419–429. <https://doi.org/10.17762/turcomat.v13i03.13009>
15. Kasetwar A, Balani N, Damwani D, Pandey A, Sheikh A, Khadse A (2022) A WebRTC based video conferencing system with screen sharing. *Int J Res Sci Eng Technol* 10(5):5061–5067. <https://doi.org/10.22214/ijraset.2022.43595>
16. Pasha M, Shahzad F, Ahmad A (2016) Analysis of challenges faced by WebRTC video conferencing and a remedial architecture. *Int J Comput Sci Inf Secur* 14(10):698–705. <https://doi.org/10.48550/arXiv.1701.09182>
17. Loof A, Holm S (2019) Peer to peer multimedia real-time communication system based on WebRTC technology. *Int J Hist Eng Technol* 2.9(7):125–130. <https://doi.org/10.22214/ijraset.2022.43595>
18. Cui J, Lin Z (2016) Research and implementation of webRTC signaling via web socket based for real-time multimedia communication. In: *International conference on computer science and automation engineering*. Sanya, Hainan, pp 374–380 <https://doi.org/10.2991/issue-15.2016.72>
19. Singh R, Donegan A, Tiwari H (2021) Framework for a decentralized web. In: *International telecommunication networks and application conference*, Melbourne, VIC, Australia, pp 1–7. <https://doi.org/10.1109/ITNAC50341.2020.9315032>
20. Dhyani H, Gupta T, Kumar P, Sharma A (2022) Real-time communication over blockchain. *Int J Res Sci Eng Technol* 10:1209–1215. <https://doi.org/10.22214/ijraset.2022.42479>
21. Mandal S, Kar AK, Gupta S (2023) Blockchain-based application security versus centralized and distributed data management systems—a comparative study. *Transf Diffus Adopt Next-Gener Digit Technol* 697:207–218. https://doi.org/10.1007/978-3-031-50188-3_18
22. Zou R (2022) Decentralized WebRTC P2P network using kademlia. 1–4. <https://doi.org/10.48550/arXiv.2206.07685>
23. Saeed AM, Edan MN (2022) Design and implementation of peer to peer video and chat communication. *Tech Soc Sci J* 37:142–152. <https://doi.org/10.47577/tssj.v37i1.7570>
24. Faraz K, Niraj M, Sagar R, Dhananjay D, Puja P (2020) Anonymous de-centralized ephemeral chat application using interplanetary file system. *ITM Web of Conferences* 32:02004. <https://doi.org/10.1051/itmconf/20203202004>

25. Unger N, Dechand S (2015) SOK: secure messaging. In: IEEE symposium on security and privacy. San Jose, CA, USA, pp 232–249. <https://doi.org/10.1109/SP.2015.22>
26. Ali RM, Alsaad SN (2020) Instant messaging security and privacy secure instant messenger design. IOP Conf Ser Mater Sci Eng 881:1–10. <https://doi.org/10.1088/1757-899X/881/1/012117>
27. Kleppmann M, Weidner M (2021) Key agreement for decentralized secure group messaging with strong security. University USA, UK. <https://doi.org/10.1145/3460120.3484542>
28. Zarrin J, Phang Wen H, Saheer Babu L, Zarrin B (2021) Blockchain for decentralized of internet prospects, trends and challenges. Clust Comput 24:2841–2866. <https://doi.org/10.1007/s10586-021-03301-8>
29. Chen G, Xu B, Lu M, Chen SN (2018) Exploring blockchain technology and its potential application for education. Smart Learn Environ 5(1):1–10. <https://doi.org/10.1186/s40561-017-0050-x>
30. Gautami T, Ahad MA, Casalino G (2023) A comprehensive review of blockchain technology: underlying principles and historical background with future challenges. Decis Anal J 9:100344. <https://doi.org/10.1016/j.dajour.2023.100344>
31. Dang NT, Tran HM, Nguyen SV, Maleszka M, Le H-D (2021) Sharing secured data on peer-to-peer applications using attribute-based encryption. J Inf Telecommun 5(4):440–459. <https://doi.org/10.1080/24751839.2021.1941574>
32. Khan AF, Akash A, Avinash R, Lokesh C (2018) WebRTC peer to peer learning. Int J Eng Res Technol 7(03):233–238. <https://doi.org/10.17577/IJERTV7IS030135>
33. Shreyas HN, Sharanabasappa, Sanjana JK (2021) Decentralized web messaging peer to peer multimedia real-time communication system based on WebRTC technology. Int J Hist Eng Technol 2.9(7):125–130
34. Halder D, Bhushan S, Shreya G, Kumar P (2022) fybrrChat: a distributed chat application for secure P2P messaging. <https://doi.org/10.48550/arXiv.2207.02487>
35. Gupta B, Vani MP (2018) An overview of web socket: the future of real-time communication. Int Res J Eng Technol 05(12):434–437
36. Kumar M, Thakur V, Gurjar D (2022) Multi-user web chat application using node.js and Socket.io. Int J Adv Res Science Commun Technol 2(4). <https://doi.org/10.48175/568>
37. Nayyef Z, Amaer S, Hussain Z (2019) Peer to peer multimedia real-time communication system based on WebRTC technology. Int J Hist Eng Technol 2.9(7):125–130

Face Identification and Categorization with and Without Mask Utilizing MTCNN and OpenCV for Accessing Bank Locker Facility



R. Monica Lakshmi, K. Ramar, Nikita Devendran, L. P. Deepika, and S. Mounika

Abstract Face recognition is an application-rich, widely-used technology. However, the rise of face masks during the COVID-19 epidemic poses a hurdle to precise facial recognition. Creating a facial recognition system that can recognize people even while they are wearing masks is the aim of this research. They will be able to open the locker once the face has been confirmed. If not, entry would be refused. First, faces in the stream are identified using face identification techniques such as MTCNN. Next, each recognized face's orientation and scale are normalized through the use of an alignment procedure. Feature extraction using a deep learning model, like Facenet, is the next stage. This model converts every aligned face into an embedding, which is a high-dimensional vector representation of the face's key traits and features. Finally, the embedded vectors are subjected to a softmax classifier to authenticate or categorize faces. Furthermore, activities such as face verification and grouping can be enabled by using similarity measures such as cosine similarity to assess the similarity of embedded vectors.

R. Monica Lakshmi (✉) · N. Devendran · L. P. Deepika · S. Mounika
Department of Computer Science and Business Systems, R.M.D. Engineering College,
Thiruvallur, Tamil Nadu, India
e-mail: monica.ramar@gmail.com

N. Devendran
e-mail: ucb20128@rmd.ac.in

L. P. Deepika
e-mail: ucb20111@rmd.ac.in

S. Mounika
e-mail: ucb20126@rmd.ac.in

K. Ramar
Department of Computer Science and Engineering, R.M.K. College of Engineering and
Technology, Pudukkottai, Chennai, India
e-mail: monica.ramar@gmail.com

Keywords MTCNN · Facenet · OpenCV · Classification · Face alignment · Face recognition

1 Introduction

Facial recognition and classification are critical components in modern security systems, especially in situations such as bank locker access. However, traditional facial recognition techniques are facing significant challenges as a result of the growing acceptance of mask wear for a variety of reasons, including health concerns. To address this issue, novel techniques are being investigated, including the use of OpenCV and the Multi-task Cascaded Convolutional Neural Network (MTCNN). These technologies enable the development of systems capable of correctly recognizing and classifying people, whether wearing masks or not. The project's goal is to develop a facial recognition system that can accurately identify people even when they are wearing masks. The system correctly links the recognized face to its matching identity in the database by using extensive matching algorithms based on deep learning. By providing dependable access control to bank locker facilities, this solution not only strengthens security measures but also highlights how flexible facial recognition systems are in constantly shifting contexts. Furthermore, these solutions take a forward-thinking approach to addressing modern difficulties by supporting both masked and unmasked scenarios, all the while upholding the integrity and effectiveness of the required security procedures.

2 Motivation

The widespread use of facial recognition systems in security infrastructure, particularly for bank locker access, emphasizes the importance of reliable identification methods. However, the rise of mask-wearing as a common practice, prompted by a variety of factors including health concerns, poses a significant challenge to traditional facial recognition techniques. The inability of these systems to correctly identify people wearing masks jeopardizes the overall security of such facilities. This challenge serves as the primary motivation for researching novel techniques that can overcome the challenges posed by mask-wearing scenarios, resulting in robust and dependable access control measures.

To meet this pressing need, there is a growing interest in using advanced technologies like OpenCV and the Multi-task Cascaded Convolutional Neural Network (MTCNN) to create novel solutions. These technologies present promising opportunities for improving facial recognition systems in order to effectively recognize and classify individuals, regardless of whether they are wearing masks. In order to address the issues raised by mask-wearing situations, it is possible to develop systems that

can accurately identify faces in a range of environmental conditions by utilizing the power of deep learning algorithms and sophisticated image processing techniques.

3 Main Contribution

A. Adapting Facial Recognition for Mask Usage

The project's adaptation of facial recognition technology to situations involving mask wear is one of its main contributions. Due to the fact that masks are worn for many different purposes, including health-related ones, traditional facial recognition algorithms have a very hard time correctly identifying people. Using OpenCV and the Multi-task Cascaded Convolutional Neural Network (MTCNN), our system takes on this challenge head-on. We developed a method that can reliably identify and classify people even when they are wearing masks, boosting the security systems' resilience in situations like bank locker access.

B. Creation of a Dependable Facial Recognition System

The development of an accurate and dependable facial recognition system is another noteworthy output of this project. To make sure that recognized faces are connected to their matching identities in the database, our system employs sophisticated matching algorithms built on deep learning. For security applications, accuracy must be prioritized. Our system ensures dependable access control by not only meeting but surpassing the strict requirements of contemporary security protocols.

C. Flexibility and Adaptability in Shifting Contexts

Additionally, our research highlights how important it is for facial recognition systems to be adaptive and flexible, particularly in situations where the surroundings are constantly changing. Our system is flexible enough to accommodate both masked and unmasked scenarios, allowing it to adjust to changing conditions without compromising security or efficacy. Our facial recognition technology keeps correctly identifying and authenticating them, guaranteeing the consistency of security protocols across a range of environments.

D. Strengthening Security Measures

Ultimately, the main objective of this project is to enhance security protocols through the implementation of a cutting-edge facial recognition technology. By offering trustworthy access management and lowering the risks of fraudulent or illegal entry, our solution strengthens the overall security posture of bank locker facilities. Additionally, we show how facial recognition systems may adjust to novel situations and still be a reliable means of safeguarding confidential information. This project serves as an example of how security technologies are constantly evolving and innovating.

4 Organization of the Paper

To make it easier to convey our research coherently, the paper is divided into multiple sections. We start with a review of the literature to look at what is known about face recognition systems, with an emphasis on issues with mask use. After that, we talk about the shortcomings of the systems as they stand, preparing the audience for the introduction of our suggested fix.

We discuss our approach's methodology, which combines the Multi-task Cascaded Convolutional Neural Network (MTCNN) with OpenCV. Here are the findings from in-depth testing and analysis, which shed light on how well our system functions. We also acknowledge the limitations of our current study and propose potential avenues for future research. We wrap off by summarizing our main conclusions and stressing the significance of modifying facial recognition software for situations in which masks are worn.

5 Literature Survey

[1] Rahul Baghel, Pallavi Pahadiya, Upendra Singh, "*Human Face Mask Identification using Deep Learning with OpenCV Techniques*", 7th International Conference on Communication and Electronics Systems (ICCES), 2022, pp. 1051–1057.

The WHO states that respiratory droplets and interpersonal contact are the two ways that the COVID-19 virus spreads. It is advised that people isolate themselves from others and use masks to stop the spread. To comply with regulatory requirements, a mask recognition system has been developed for video streams. Masks are detected using a deep learning method based on Python's TensorFlow, Keras, and PyTorch packages.

[2] Amjad Ali Khan, Shankar Yadav, Lalit Kumar, "Face Mask Detection Using OpenCV and Machine Learning", 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022, pp. 80–85.

Wearing protective face masks has become a new fashion trend that is widely accepted. As a result, we are developing an AI system capable of detecting whether or not a person is wearing a mask. Using this method, we can prevent the virus from spreading and protect the ecosystem. This project necessitates the installation of numpy, OpenCV, TensorFlow, and other learning tools, in addition to Jupyter Notebook.

[3] Tamilarasan Kumar, Rajendrane Rajmohan, Muthu Pavithra, "Automatic Face Mask Detection System in Public Transportation in Smart Cities Using IoT and Deep Learning", Expert Systems with Applications, Volume 36, Issue 2, Part 2.

To mitigate the effects of COVID-19, there is an increased demand for effective city management as the population grows. AI should be used to expand smart cities so that they can thrive. This article describes an IoT-based face mask detection system for

public transit. This system collects real-time data. The model detects faces and masks with minimal inference time and memory, meeting the trials' resource constraints. Additionally, the system's scalability allows for deployment in a variety of urban settings, ensuring widespread adherence to health protocols.

[4] Gopinath Pranav Bhargav, Kancharla Shridhar Reddy, Alekhya Viswanath, "Facemask Detection with Face Recognition and Alert System using MobileNetV2", 2nd International Conference on Intelligent and Cloud Computing (ICICC), 2022, vol. 286, pp. 77–87.

Donning a mask is a successful strategy for preventing COVID-19. We developed a web application that uses an integrated facemask detection and face recognition algorithm to continuously remind users to wear masks. The suggested method recognizes a person's face and determines if they are wearing a mask or not. Additionally, an automatically produced email is sent to that particular violator's personal email address reminding him to wear a mask.

[5] K. Yagna Sai Surya, T. Geetha Rani, B. K. Tripathy, "Social Distance Monitoring and Face Mask Detection Using Deep Learning", Computational Intelligence in Data Mining, 2022, pp.461–476.

Deep learning frameworks and Python can be used to create a reliable system for tracking social distancing and people capture. The technology can reliably determine whether people are wearing masks by utilizing deep learning models, which guarantees adherence to safety protocols. Regardless of the time of day, the system can continuously observe people in real time by using computer vision algorithms to monitor social distance measures. This technology is very useful in public spaces, shopping centers, and other places where maintaining safe distances is crucial to stopping the spread of illness.

[6] Ning Zhang, Junmin Luo, Wuqi Gao, "Research on Face Detection Technology Based on MTCNN", International Conference on Computer Networks, Electronics and Automation, 2021, pp. 154–158.

Multi-task Cascaded Convolutional Networks (MTCNN), a deep learning architecture, increase the precision and dependability of face identification systems. High performance is guaranteed even in challenging circumstances like shifting lighting and occlusions because of MTCNN's architecture, which integrates several functions like face detection and facial landmark localization. The technology is perfect for applications like social interaction monitoring, biometric verification, and security surveillance since it can process real-time video streams. This study demonstrates MTCNN's ability to improve face detection technology.

[7] Alyaa Abdelaziz, Hana Serag, Malak Ziad, Mariam Amr, "Face Recognition Using Python and OpenCV", International Journal of Scientific and Research Publications, 2024, Vol. 14, pp. 70–94.

Because human faces are dynamic, multidimensional systems, sophisticated recognition methods are required. The use of automated face recognition in identity verification, criminal identification, public safety, and intelligent self-driving cars has made it more and more popular in recent years. Research in this field began in the 1960s, resulting in the development of a variety of techniques, including local, holistic, and hybrid methods. Despite advances, researchers in computer vision

and pattern recognition continue to face challenges in developing robust and efficient face recognition algorithms. This paper gives an overview of face recognition, highlighting various techniques.

[8] Jayanth Vadlapati, Senthil Velan, Ewin Varghese, “Facial Recognition using the OpenCV Libraries of Python for the Pictures of Human Faces Wearing Face Masks during the COVID-19 Pandemic”, 2021, 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–5.

The technology can reliably determine whether people are wearing masks by utilizing deep learning difficulties in correctly detecting people who are donning masks. Conventional facial recognition algorithms rely largely on the capture of unique facial features such as the mouth, nose, and chin, which are frequently covered up, either completely or in part, by masks. As such, the performance of current facial recognition algorithms noticeably deteriorates in situations where mask use is common, due to the obscuring of critical facial features that these algorithms rely on for accurate identification.

This research delves into the application of computer technology in AI and Machine Learning, particularly in Image Processing and Pattern Recognition. These techniques are pivotal for recognizing real-world objects such as human faces. Utilizing Python’s diverse array of face recognition libraries, we’ve trained models capable of identifying individuals even when they are wearing masks, which obscure a significant portion of facial features. This technology is essential for applications in biometrics, video surveillance, and other fields where accurate and efficient face detection is crucial for enhancing security measures.

6 Existing System

In the state of face recognition systems today, facial feature analysis is the primary method of identification.

7 Proposed System

Developers can use face recognition algorithms to reliably identify human faces from a stream against a database of recognized faces by utilizing Python and its associated modules.

First, faces in the stream are identified and located using face detection techniques such as MTCNN. After the faces are identified, an alignment procedure is utilized to standardize the orientation and scale of every identified face, guaranteeing uniform input for the following phases.

After detecting and aligning faces, Facenet extracts features, converting each aligned face into an embedding—a high-dimensional vector capturing its unique traits. These embeddings facilitate effective comparison and recognition, serving as

compact and discriminative representations. Subsequently, softmax classifiers classify faces, while cosine similarity enables tasks such as face verification and grouping based on the similarity of embedded vectors.

8 Methodology of Approach

A. System Specifications

The software requirements are:

- Anaconda Distributor.
- Python - 3.7 or later versions.
- Text editor such as Notepad, or MS Word.
- A 64-bit OS with a $\times 64$ -based processor.
- Windows 7 or later versions.
- Visual Studio Code.

The hardware requirements are:

- Core i3 processor and above.
- 4 GB of RAM.
- 80 GB of hard disk capacity.
- 15 inch color monitor.
- Full HD webcam.
- 52 $\times$ CD-ROM drive.
- Architecture Diagram

Architecture diagram is a visual representation of software system components. The below diagram is the architecture of the system (Fig. 1).

The process of collecting data entails compiling an array of distinct facial picture datasets of each person, containing a range of positions, expressions, and lighting scenarios captured from the webcam's real-time feed.

Next, in order to increase model generalization, the data preprocessing module standardizes and improves the gathered images by resizing, normalizing, and enhancing them.

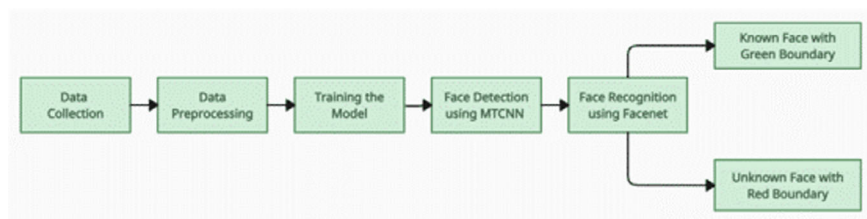


Fig. 1 Architecture diagram

During the training phase, the preprocessed data is fed into a recognition model (e.g., Facenet) so that it may learn to produce embeddings and extract discriminative features from faces.

MTCNN is used for face detection, precisely locating and extracting facial regions from pictures or video frames. After being identified, these faces are aligned and sent into Facenet so that embeddings may be extracted.

Lastly, the trained model is applied to real-time face recognition tasks, where people are identified or verified by comparing embedded vectors using methods such as cosine similarity.

C. Libraries and Frameworks

The libraries and frameworks used in this system are:

• Matplotlib	• Utils
• Keras	• Glob
• TQDM	• Argparse
• Scipy	• Numpy
• Pickle	• CV2

D. Data Collection

Data collection is a crucial step in developing a face recognition system. The process involves gathering relevant data to train the face detection, and face recognition algorithms. Collect a diverse dataset of face images representing different individuals. These images should cover variations in age, gender, ethnicity, and pipe for training. Rotation, translation, scaling, flipping, shearing, zooming, brightness and contrast alteration facial expressions. The webcam is used to take the pictures. To get realistic photos, place the subject squarely in front of the camera under well-lit conditions. To begin the collection process, a virtual environment needs to be built using the Anaconda prompt. Activate the environment once it has been built. As a consequence, the system gathers about 100 photographs. When a person is identified, their name will be presented in the directory under their respective name.

E. Data Preprocessing

Data preprocessing refers to the process of transforming raw data into a more suitable for analysis or machine learning tasks.

Generating fresh versions of the training images that represent the scenarios the model faces during inference is the aim of image data augmentation. The model gains the ability to handle various conditions by introducing these changes. It is frequently used in conjunction with Python libraries to add them all to the.

F. Training the Model

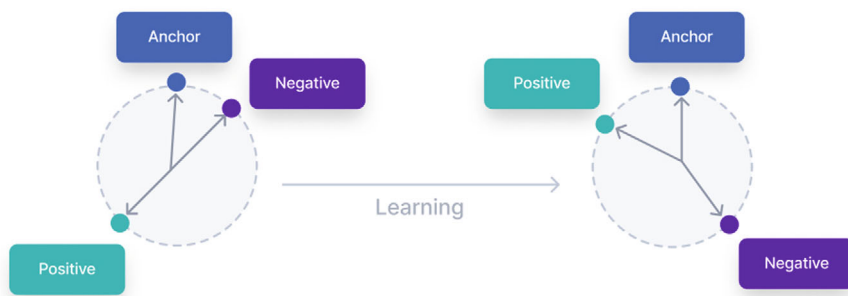


Fig. 2 Triplet loss in facenet

The process of using a machine learning algorithm to identify patterns and relationships within a dataset is referred to as “training the data.” A machine learning model is trained on a large collection of tagged face photos in the context of facial recognition. During training, the model learns to identify relevant features and assign them to appropriate labels.

“Training the data” is the process of using a machine learning algorithm to identify patterns and relationships in a dataset. A machine learning model is trained for facial recognition using a large set of annotated face images. During training, the model learns to recognize relevant features and associate them with the appropriate labels.

After alignment, the facial images are run through Facenet, an embeddings model. A face that has been taken from an image and represented numerically is referred to as embedding. A face recognition system converts an image into a compact numerical high-dimensional vector representation, or embedding, by using a deep learning model to extract features from the face (Fig. 2).

Depending on how this embedding is learned, similar faces are mapped close together in this space, while dissimilar faces are mapped farther apart. FaceNet’s use of a triplet loss function during training is one of its most significant advances. This loss function compares the similarity of three images: a negative picture (representing a different identification), a positive image (representing the same identity as the anchor), and an anchor image (representing the identity to be recognized).

The model’s training objective is to maximize the distance between the anchor’s embeddings and negative images while minimizing the distance between the anchor’s embeddings and positive images. This encourages the model to learn discriminative embeddings that are invariant to lighting, posture, and facial emotions.

G. Face Detection

Finding human faces in photos or videos is possible with computer vision techniques like face detection. It involves identifying and pointing out faces’ locations and presence. The coordinates of the smallest rectangles that can be drawn around faces are extracted to create a list of bounding boxes.

We use MTCNN for face detection. MTCNN uses a three-stage neural network detector. First, the image is resized several times to detect faces of various sizes.

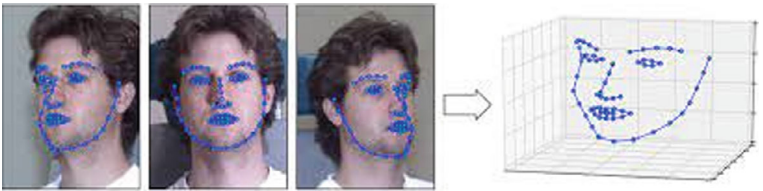


Fig. 3 Face alignment

The P-network (Proposal) scans the images and performs the first detection. The proposed regions are fed into the second network, the R-network (Refine), which filters detections to produce highly precise bounding boxes. MTCNN’s final stage is the O-network (Output), which refines all bounding boxes.

H. Face Alignment

The process of identifying and locating important facial landmarks on a face, followed by aligning or posing the image to a standard is known as face alignment. Typically, these landmarks consist of points like the corners of the mouth, nose tip, and eyes.

Face alignment is used to make sure that, despite changes in head position, facial expression, or illumination, facial features are positioned and aligned consistently in various photographs. Facial recognition algorithms can compare and match facial features more precisely for identification or verification when faces are aligned to a shared reference frame (Fig. 3).

I. Classification

A classifier is a machine that determines if a picture is positive or negative. To accurately identify a new image as a face or non-facial image, it is trained on hundreds of thousands of face and non-face photos. To find things, we employ the Haar Classifier, which instills a cascade operation from the images. Next, the traits are extracted from the image. Every characteristic is a single number that is obtained by deducting the total of the pixels in the model’s predicted output from the actual target values in the training dataset. It represents the discrepancy between the predicted and actual values and is used as a measure of how well the model is performing (Table 1).

Table 1 Observation of Haar

No. of faces	Execution time (s)	No. of faces detected	Accuracy (%)
5	0.141	5	100
10	0.055	9	90
15	0.11	12	80
20	0.369	19	95

The Haar value classification is as follows:

$$\begin{aligned} \text{Pixel value} &= (\text{Sum of the Dark Pixels} / \text{Number of Dark Pixels}) \\ &\quad - (\text{Sum of the Light Pixels} / \text{Number of Light Pixels}) \end{aligned}$$

The accuracy of the model's predictions on the training dataset is expressed as a percentage. Loss is the difference between the model's predicted output and the actual target values in the training dataset.

The output layer of a neural network uses the softmax function to convert raw scores or logits into probabilities. These probabilities represent the possibility of each class or identity appearing in the input data. It takes facial image features as input and outputs a vector of scores that represent various identities or classifications.

$$P(\text{class}_i) = \frac{e^{\text{score}_i}}{\sum_j e^{\text{score}_j}}$$

where $P(\text{class } i)$ is the probability of the i -th class, score i is the raw score or logit for the i -th class, e is the natural logarithm base, and j is the sum of the exponentials of all raw scores.

If two vectors are roughly pointing in the same direction, their cosine similarity can be calculated using the cosine of their angle. Cosine similarity measures the similarity between two faces by translating them into these numerical representations.

A value of 1 represents perfect similarity, -1 represents total dissimilarity, and values closer to 0 indicate a lack of meaningful resemblance. In face identification, choose the match that has the highest cosine similarity score.

The cosine similarity between the two feature vectors A and B , can be calculated by the following formula:

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

An epoch is one complete pass of the entire training dataset through the neural network during the training phase.

To reduce the error between the predicted outputs and the actual labels, the neural network uses optimization algorithms like gradient descent to modify its parameters (weights and biases) during training. Iteratively going through several epochs, this procedure is carried out until the model converges to a performance level that is suitable. The training dataset is split up into smaller batches throughout each epoch, which comprises several iterations. The batches are successively fed into the neural network, and the model's parameters are adjusted based on the accumulated gradients once all of the batches have been processed in an epoch.

The accuracy of a machine learning model during testing and training is displayed as the number of epochs increases on the graph. "Epochs" is the label on the x -axis of the graph, while "accuracy" is the label on the y -axis. The training accuracy in the

graph indicates that the model is perfectly fitting the training data because it starts at 1.0 and remains there throughout all epochs. When the number of epochs rises, the testing accuracy varies slightly from its initial value of 0.98 (Fig. 4).

J. Face Identification

A high-dimensional vector space is created by encoding faces into FaceNet, a facial recognition framework, so that similar faces are situated next to one another.

FaceNet recognizes faces in images and uses preprocessing to prepare the image before feeding it into the neural network to create an embedding, or numerical representation, of the face. The similarity between the faces is then ascertained by comparing these embeddings using cosine similarity or alternative distance metrics.

Finding and locating faces in each video stream frame is the initial step in the process. This is done using deep learning-based techniques like Haar cascades and convolutional neural networks. Faces may need to be aligned and normalized to a standard size and orientation after they are found in order to improve recognition accuracy. After normalizing, facial features are extracted from the aligned faces. Important facial features, like the locations of noteworthy landmarks or numerical

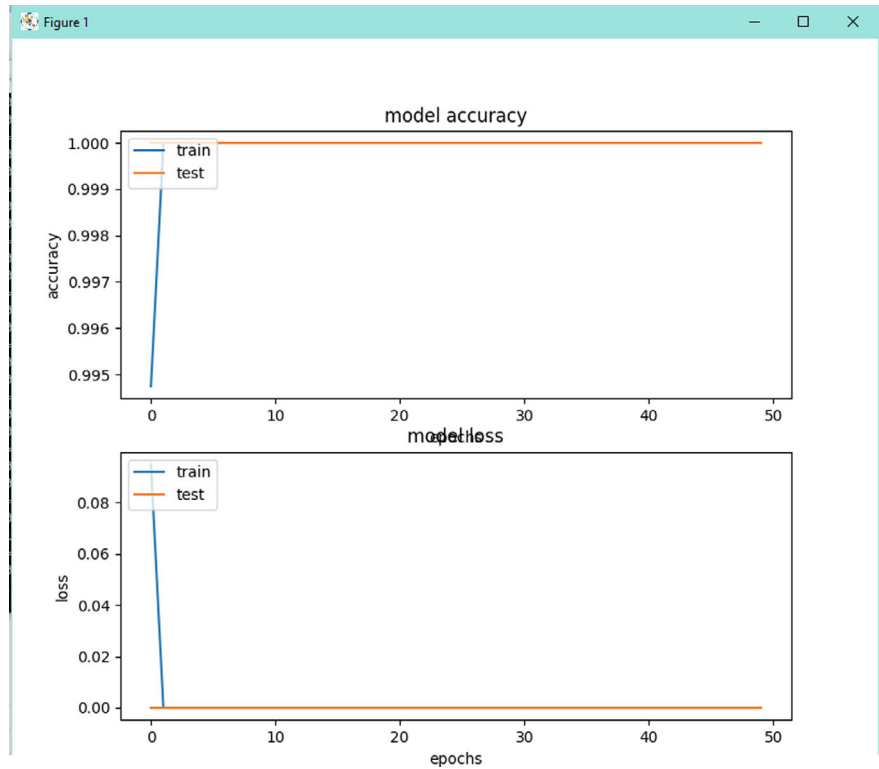


Fig. 4 Epoch calculation

representations of facial patterns, are recorded using this technique. Lastly, a database of recognized faces or templates is compared to the retrieved characteristics using similarity metrics like cosine similarity and Euclidean distance.

9 Result and Discussion

A. Elapsed Duration

It refers to the amount of time it takes for the system to process and analyze a given image or video frame to identify faces. When considering the presence of masks, traditional systems struggle to accurately identify individuals, requiring additional steps such as manual verification. In contrast, the new facial recognition systems have adapted to accommodate mask-wearing, effectively reducing the time (Fig. 5).

B. Accuracy

Accuracy refers to the ability of the system to correctly identify or verify individuals from their facial features. It is typically measured as the percentage of correct identifications out of the total number of attempts. Without masks, both systems exhibit comparable accuracy, but the new system demonstrates its effectiveness in the presence of facial coverings (Fig. 6).

C. F1 Score

It is a metric for assessing the performance of a facial recognition system. The $F1$ score is calculated by taking the harmonic mean of precision (the proportion of true positives among all positive instances) and recall. Individuals wearing masks have a significantly lower score due to the occlusion of facial features. With all-new facial recognition systems, the score remains strong even in masked scenarios (Fig. 7).

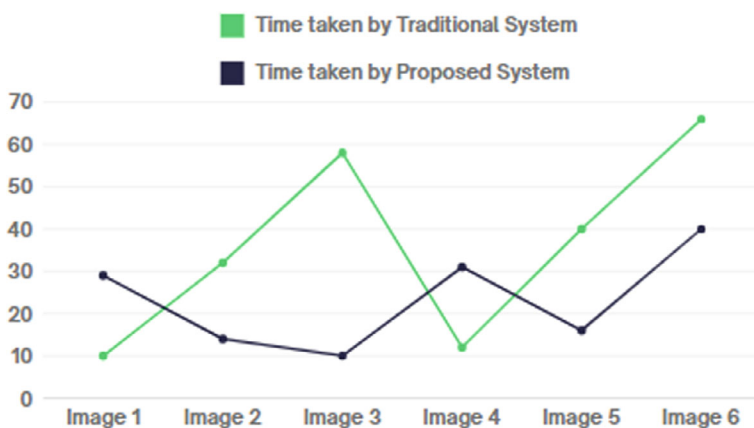


Fig. 5 Elapsed duration

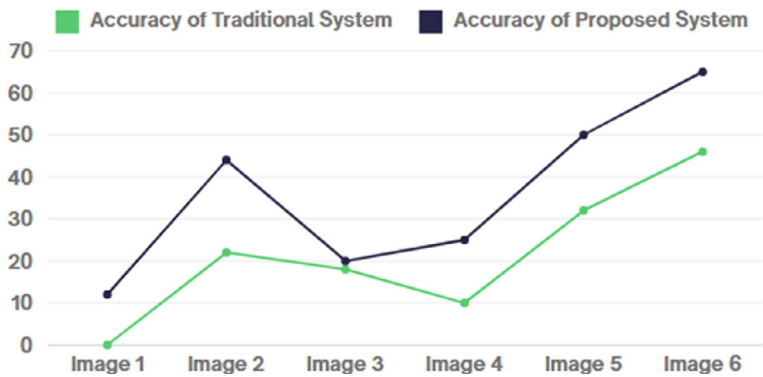


Fig. 6 Accuracy

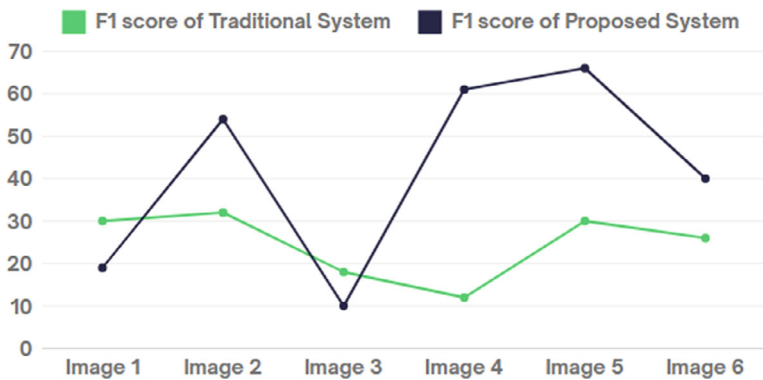


Fig. 7 F1 score

D. Resource Consumption

The quantity of computational resources needed for the system to function properly is referred to as resource consumption. Conventional facial recognition systems require a significant amount of computational power because they process the entire facial structure. By processing facial features even when partially obscured, the suggested system with mask detection algorithms can lower computational overhead (Fig. 8).

E. Environmental Resilience Capacity

It refers to the system’s ability to adapt and perform efficiently despite environmental factors. Traditional systems suffer in areas with variable illumination and partial obstacles, making them less resilient. The proposed system, which is specifically intended to respond to these problems, displays greater environmental resilience (Fig. 9).

F. Cross-Domain Performance

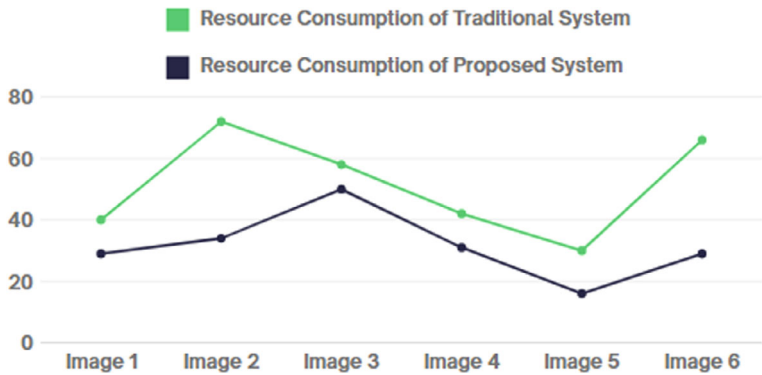


Fig. 8 Resource consumption

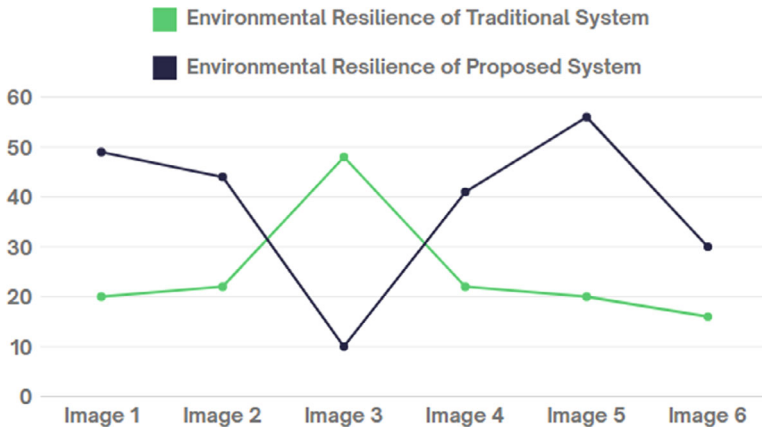


Fig. 9 Environmental resilience

Cross-domain performance is an assessment of a system’s effectiveness when deployed across multiple domains or applications. This evaluation is critical for determining how well the system functions in many situations and circumstances, such as surveillance, access control, and user authentication. This includes examining its capacity to produce consistent and trustworthy results across a wide range of contexts, datasets, and real-time situational use cases (Fig. 10).

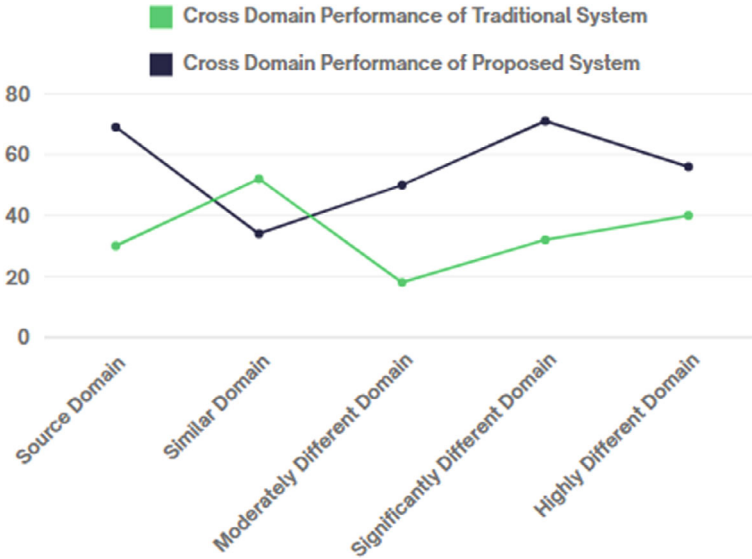


Fig. 10 Cross-domain performance

10 Limitation of Work

While our findings suggest promising advances in adapting facial recognition technology for mask-wearing scenarios, it is critical to recognize several limitations. To begin, our system’s performance may vary depending on factors such as lighting conditions, camera quality, and the fit of the mask worn by individuals. Our approach is effective in controlled environments, but its scalability and robustness in real-world settings with diverse populations and environmental conditions warrant further investigation. Furthermore, our system’s reliance on deep learning models presents computational challenges, especially in resource-constrained environments or applications that require real-time performance. Furthermore, the applicability of our findings to different demographic groups and cultural contexts requires careful consideration. Finally, ethical and privacy concerns about facial recognition technology require ongoing monitoring and scrutiny. Addressing these limitations will be critical to the continued development and deployment of facial recognition systems in changing scenarios.

11 Future Enhancements

- **Improved Mask Detection:** Improve the mask detection algorithm's accuracy and robustness by incorporating advanced machine learning techniques like deep learning models or ensemble methods.
- **Multi-Modal Fusion:** Investigate the use of multiple biometric modalities, such as facial features, voice recognition, gait analysis, and even thermal imaging, to improve identification accuracy.
- **Integration with Bank Locker Access System:** Create APIs or protocols for communication and integration between the face recognition module and the bank's locker access system, ensuring real-time authentication and authorization for individuals.
- **Real-time Monitoring and Logging:** Implement real-time monitoring of individuals entering the bank locker facility, including facial recognition and mask status. Log this information, along with timestamps, for auditing and ensuring security protocol compliance.
- **Scalability and Robustness:** Ensure that the system is scalable to handle a large number of users and reliable in real-world scenarios.

12 Conclusion

The face recognition system uses MTCNN and OpenCV to improve identification accuracy even when masks are used. It improves security and access control by combining mask detection and face recognition while maintaining reliability. Incorporating alternative biometric modalities improves identification in mask-wearing scenarios while preserving privacy by focusing on identification only when masks are not present. To summarize, this system enables accurate identification, increased security, and adaptability to mask-related challenges.

Software Engineering Approach for Automated Vehicle Detection and Classification on Remote Sensing Images



Mohammed Assiri

Abstract In recent times, the integration of the software engineering concepts with remote sensing technology has resulted in novel methods for vehicle classification, enabling different applications ranging from traffic management to environmental monitoring. Remote sensing images (RSI) become useful in several applications of intelligent transportation systems (ITS) like traffic prediction, vehicle classification, road classification, etc. Robust and accurate vehicle classification in ITS using RSI gains considerable attention using computer vision (CV) and deep learning (DL) approaches. However, the detection capability is restricted to the absence of well-annotated samples, particularly in dense crowded scenes. Vehicle detection and vehicle classification are the two problems commonly addressed by the CV and DL models. In this aspect, this study develops an optimal DL-based vehicle detection and classification model on ITS (ODLVDC-ITS) using RSI. The key objective of the ODLVDC-ITS method lies in the automated classification of vehicles present in the RSI. To achieve this, the ODLVDC-ITS method comprises three main processes namely vehicle detection, parameter tuning, and vehicle classification. Primarily, the ODLVDC-ITS technique employs YOLO-v5 object detector to identify the vehicles in the images and its hyperparameters can be tuned by coyote optimization algorithm. The ODLVDC-ITS technique exploits Dendritic Neural (DRN) technique for vehicle classification process. Lastly, the hyperparameter tuning of DRN model performed via Aquila optimization algorithm (AOA). The simulation outcomes of the ODLVDC-ITS method are tested on open access dataset and the outcomes are investigated under various measures. A widespread experimental analysis stated the promising performance of ODLVDC-ITS approach on vehicle classification process.

Keywords Remote sensing images · Intelligent transportation systems · Software engineering · Deep learning · Vehicle classification · Object detection

M. Assiri (✉)

Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

e-mail: m.assiri@psau.edu.sa

1 Introduction

In last decade, several research studies about the exploitation of remote sensing images (RSIs) in the design of intelligent transportation system (ITS) has been considerably raised [1]. ITS plays a major role in the design of smart cities, which needs to improve the quality of the people irrespective of the drastic increase in traffic density, delay, traffic jam, high fuel cost, toxic emissions, accidents, etc. [2]. Latest advances in information and communication technologies (ICT) in several domains like telecommunication, networking, hardware, etc., have posed new possibilities to develop a sustainable ITS [3]. The incorporation of ICT with the ITS enables to provide of effective, reliable, and safe travel. The accomplishment of effective ITS mainly relies on the environment applied for accessing, collecting, and processing information from the target. Recently, RSI becomes effective in the data collection process [4]. Various recent developments have allowed automatic modeling, and the understanding of the data is a fascinating research area in the RSI-based ITS [5].

With the accessibility of high-resolution RSI, the object detection process find useful. Although object detectors are available in the literature, object recognition in RSI remains a challenge due to high-resolution spatial images [6]. The vehicles can be viewed as tiny spots, which cannot be easily identified from the forefront regions. Classification of identified vehicles is important as it is not easier to differentiate large-sized and small-sized vehicles by human eye from high-resolution HIS [7]. The design of object detectors depending upon machine learning (ML) becomes familiar with respect to accuracy and speed over classification detectors with feature extraction process. It can be employed to classify and locate objects in the automated image recognition process using statistic and geometry features [8]. The benefit of openly expending images in the object detector applications enables to elimination of the manual deriving of the features. CNN remains an effective object detector because of the characteristics of straightway deriving the features and identifying the objects from the images [9]. Recently, excellent outcomes have been accomplished in the CNN to resolve the object detection problem via training the networks by the combination of local regional perception and feature extraction with classification task [10].

This study develops an optimal DL-based vehicle detection and classification model on ITS (ODLVDC-ITS) using RSI. Initially, the ODLVDC-ITS technique employs YOLO-v5 object detector to identify the vehicles in the images and its hyperparameters can be tuned by coyote optimization method. For vehicle classification process, the ODLVDC-ITS technique exploits Dendritic Neural (DRN) technique. Lastly, the hyperparameter tuning of DRN model performed via Aquila optimization algorithm (AOA). The experimental results of the ODLVDC-ITS approach are tested on open access dataset and the outcomes are inspected under various measures.

2 The Proposed Model

In this study, we have designed a new ODLVDC-ITS method for automated detection and classification of vehicles present in the RSI. The key objective of the ODLVDC-ITS technique lies in the automated classification of vehicles present in the RSI. The ODLVDC-ITS method comprises three most important processes such as YOLO-v5 based vehicle detection, COA based parameter tuning, DRN-based vehicle classification, and AOA based parameter optimization. Figure 1 represents the working process of ODLVDC-ITS algorithm.

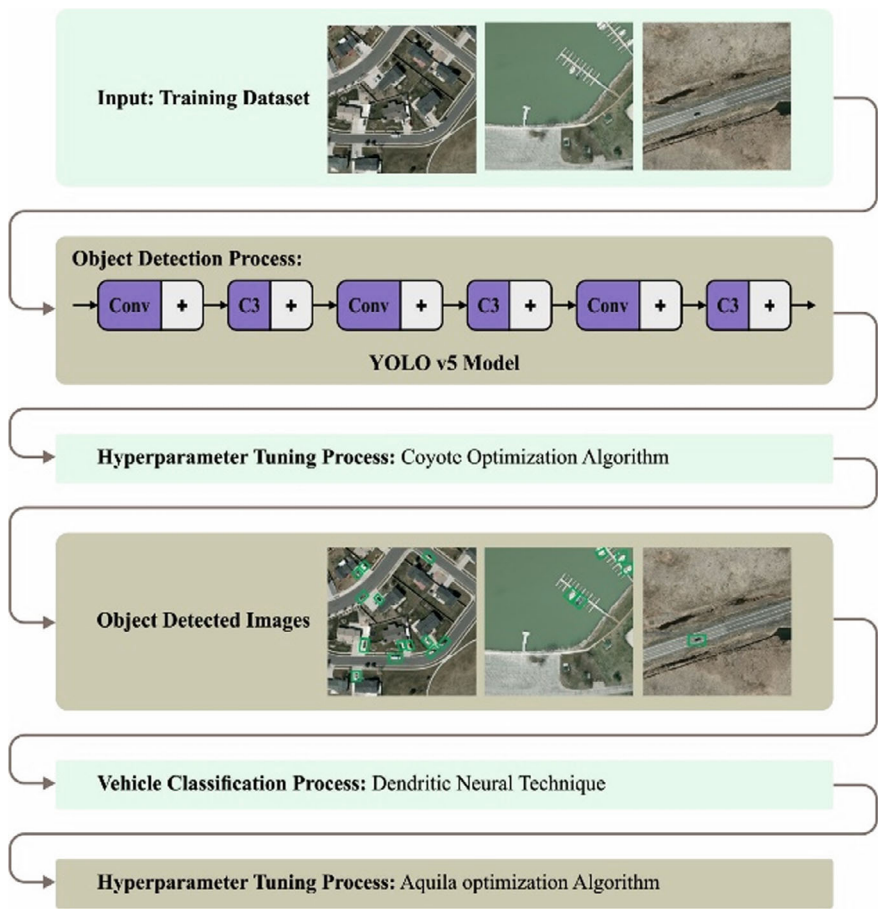


Fig. 1 Working flow of ODLVDC-ITS system

2.1 Module I: Vehicle Detection

For vehicle detection, the ODLVDC-ITS technique employs YOLO-v5 object detector. YOLO is a recently developed object detector, which mainly relies on exhausted sliding window and gives the outcome to the classification layer functioning at uniformly placed positions on the entire images or region proposal for generating bounding box which probably hold object [11]. The YOLO model considers object detection as a regression problem that offers the computation of coordinates of the bounding box and class probability of the box. The fundamental concept to design the YOLO model is the division of image into grid points that enhances the accuracy and decrease the running time for object detector. In short, the YOLO model comprises three major processes. At initial stage, the YOLO model receives an input image, resizes to predefined sizes, operates on convolutional network as a combined network, and lastly offers a threshold on the final detection by confidence score. In the presented ODLVDC-ITS technique, the YOLO-v5 architecture is exploited for object recognition purpose. The backbone module is used to eliminate features in different images at varying scales. Then, the FPN model is considered to combine the features in diverse scales and fed as input to the detection head, which recognizes the classes and generates bounding boxes. The loss function of the YOLO-v5 model can be presented using Eq. (1):

$$\text{loss} = l_{\text{box}} + l_{\text{cls}} + l_{\text{obj}}, \quad (1)$$

where bounding box regressor, classifier, and confidence loss function can be defined by l_{box} , l_{cls} , and l_{obj} correspondingly. The loss function of bounding box regressor loss function is equated as follows.

$$l_{\text{box}} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} b_j (2 - w_i \times h_i) \left[\left(x_i - x_i^{\wedge j} \right)^2 + \left(y_i - y_i^{\wedge j} \right)^2 + \left(w_i - w_i^{\wedge j} \right)^2 + \left(h_i - h_i^{\wedge j} \right)^2 \right]. \quad (2)$$

The classifier loss function is represented using Eq. (3):

$$l_{\text{cls}} = \lambda_{\text{class}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} \sum_{C \in \text{classes}} p_i(c) \log(\hat{p}_l(c)). \quad (3)$$

The confidence loss function is depicted by Eq. (4):

$$l_{\text{obj}} = \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{noobj}} (c_i - c_i^{\wedge l})^2 + \lambda_{\text{obj}} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} (c_i - c_i^{\wedge l})^2, \quad (4)$$

where λ_{coord} and λ_{class} are the position and the class loss coefficient, $\hat{w}$, $\hat{h}$ represent width and height of objects, and $\hat{x}$, $\hat{y}$ designates center coordinate points of the object. If the anchor box locates at (i, j) includes object, then $I_{i,j}^{\text{obj}}$ becomes 0. Here, $\hat{p}_l(c)$ denotes actual value of the class, and $p_i(c)$ implies the class possibility of the target. The length of two is equal to total number of class labels C .

To adjust the hyperparameters related to the YOLO-v5 method, the COA is used. The COA relies on adaptive behavior of the coyote via the exchange of coyote experiences along with the environment [12]. COY has an interesting method to balance exploitation and exploration. The process begins with population and coyote as nominee solution. The population size of coyote can be separated into a multiplication of n_{coy} coyotes per pack and n_{pak} number of packs. The social status of all the coyotes specifies the S candidate solution to optimization problems. The COY model begins by randomly allocating the coyotes to the pack at t time. The social condition per coyote characterizes the single solution referred to as:

$$S = (S_1.S_2.S_3 \dots S_D) \quad (5)$$

In Eq. (5), D indicates the problem dimension in the searching range. Initially, the coyote was randomly commenced within the prior searching space. Hence, the location of random solution is determined by the following equation:

$$S_{\text{coy},j}^{\text{pak},t} = (1 - r_j)_j + r_j + ub_j \quad (6)$$

- lb_j and ub_j denotes lower and upper limitations of searching space, $j \in 1.2 \dots D$, r_j specifies the random value in $[0,1]$.

The fitness value estimates adaptation of coyote to the social as follows:

$$\text{fit}_{\text{coy}}^{\text{pak},t} = f(S_{\text{coy}}^{\text{pak},t}) \quad (7)$$

Coyote has a tendency to abandon the present pack and join another pack. The alpha coyotes in another pack are ones who is the most popular with their surroundings.

The alpha might be evaluated for the reduction problem by the subsequent formula:

$$\text{alpha}^{\text{pak},t} = \left\{ S_{\text{coy}}^{\text{pak},t} \mid \arg_{\text{coy}} = \{1.2 \dots n_{\text{coy}}\} \min f(S_{\text{coy}}^{\text{pak},t}) \right\} \quad (8)$$

Once the coyotes share their social environment with other packs to improve the survivability of the pack. In this context, the new condition of the coyotes is upgraded as follows:

$$S_{\text{coy},j}^{\text{pak},t} = S_{\text{coy},j}^{\text{pak}} + r_1 * \delta_1 + r_2 * \delta_2 \quad (9)$$

where:

- δ_1 : the alpha coyote impacts the random coyote.

$$\delta_1 = \text{alpha}^{\text{pak},t} - S_{\text{coy},r_1}^{\text{pak},t}$$

- δ_2 : the space between average position of each coyote in the pack and other coyotes in pack.

$$\delta_2 = \text{avgd}^{\text{pak},t} - S_{\text{coy},r_2}^{\text{pak},t}$$

- The r_1 and r_2 are randomly generated numbers within [0, 1].

The fitness function of novel solution using following expression:

$$\text{newfit}_{\text{coy}}^{\text{pak},t} = f(\text{new}S_{\text{coy}}^{\text{pak},t}) \quad (10)$$

Once the new social condition is supposed to be superior to the prior, the coyote elects to maintain them. Or else, the latter continues to be used as follows:

$$S_{\text{coy}}^{\text{pak},t+1} = \begin{cases} \text{new}S_{\text{coy}}^{\text{pak},t} & \text{if } \text{newfit}_{\text{coy}}^{\text{pak},t} < \text{fit}_{\text{coy}}^{\text{pak},t} \\ S_{\text{coy}}^{\text{pak},t} & \text{otherwise} \end{cases} \quad (11)$$

Lastly, social state of coyotes itself adapted better and was selected and applied as a total solution to the problem.

2.2 Module II: Vehicle Classification

In this work, the detected vehicles can be categorized by using DRN model. Typically, the basic DRN model comprises four layers [13]. The initial layer can be represented as the synaptic layer and its aim is to accept the input dataset. Then, the activation function transfers the received input dataset to the following layer. The next layer is dendritic layer; the branch of this layer is used for transmitting data from input to membrane layers that is the following layer (third layer). The mathematical model for the abovementioned step is given below:

- (1) Synaptic layer: Eq. (12) is employed on the received input dataset:

$$D_{ij} = \frac{\omega_i}{1 + e^{-a\left(\frac{\omega_i \cdot x_i - \theta_i}{a_i}\right)}} \quad (12)$$

where χ_j denotes the i th input sample and D_{ij} is the value of i th synapse. Furthermore, α denotes a positive constant value, whereas w_{im} and θ_{im} denote the adjustable parameter.

- (2) Dendrite layer: This layer aggregates input dataset to initial layer. Then, input dataset is connected with each other nonlinearly. This might be tremendously significant while processing brain data. Equation (13) acts as a representation of these nonlinear relationships.

$$M_j = \prod_{i=1}^I y_{ij} \quad (13)$$

In Eq. (13), M_j denotes the m th dendritic branch output value.

- (3) Membrane layer: The input dataset from dendrite layer's branch is integrated into the membrane layer and it can be performed by using a summation:

$$S = \sum_{j=1}^J (u_j * M_j) \quad (14)$$

where u_j represents strength of dendritic branch and S indicates input of subsequent layer u_j indicates the parameter exploited in several methods to change value for addressing regression issues.

- (4) Soma layer: The sigmoid function can be applied as an activated function. Furthermore, if membrane surpasses threshold, cell body might be removed. The mathematical formula of these functions was shown below:

$$R = \frac{1}{1 + e^{-\alpha(S-v)}} \quad (15)$$

where R symbolizes the output layer, v and α are the positive constants.

Lastly, we used AOA for optimally selecting the hyperparameter of DRN algorithm. AOA is a new optimization technique based on Aquila's behavior in the environment [14]. The AOA procedure was classified into four classes: choosing the searching region by higher flying with the vertical stoops, which discovers within the separate searching space by contour flying with the shorter glide assaults, using within the converged search region by lower flying with the slower descent attacks, and pouncing by walking and catching the target. AO swaps shooting strategy for different prey and later strike the victim with its faster speed integrated with stronger feet and claws. The exploration and exploration stages are mathematically expressed as follows:

$$\begin{aligned} \text{ag}_1(t+1) &= \text{ag}_b(t) * \left[\left(\frac{1-t}{mn} \right) - r_1 \right] + \text{ag}_m(t) \\ \text{ag}_m(t+1) &= \frac{1}{N} \sum_N \text{ag}_i(t) \end{aligned} \quad (16)$$

In the exploration, the search can be directed by $\left(\frac{1-t}{mn}\right)$ Eq. (16). t and mn : the current iteration and the maximum iteration counter.

$$\begin{aligned} \text{ag}_2(t+1) &= \text{ag}_b(t) * (\text{LEVY}(D) + \text{ag}_R(t) - (y - r)) * r_2 \\ \text{LEVY}(D) &= 0.01 * u * \sigma * |V|^{\frac{-1}{\beta}} \end{aligned} \quad (17)$$

$$\sigma = 2^{\frac{(1-\beta)}{2}} \sin\left(\frac{\pi\beta}{2}\right) * \left(\frac{\Gamma(1+\beta)}{\beta * \Gamma\left(\frac{1+\beta}{2}\right)}\right) \quad (18)$$

$$x = r * \sin\left(\frac{300\pi - D_1}{200}\right), y = r * \cos\left(\frac{300\pi - D_1}{200}\right), r = r_3 + 0.0057D_1$$

The exploitation adjustment parameter is set to 0.1.

- ub and lb : upper and lower bounds of the search space.
- D : the dimensional size $\text{LEVY}(D)$ signifies the Levy flight function.
- $r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8$: random integer ranges from zero to one.
- D_1 : integer number within $[1, D]$.
- r : the number of search cycles from [1, 13].

Afterward defining approximate position of prey and vertically descending to launch the initial attacks, Aquila makes use of the selected region to approach the prey and attack it. Such behaviors are exemplified by the following:

$$\text{ag}_3(t+1) = \alpha(\text{ag}_b(t) - \text{ag}_m(t)) + \delta * (lb(1 - r_s) + r_{\text{sub}}) - r_4 \quad (19)$$

The Aquila hunt the victim and thereby attack the prey on the ground. This hunting behavior can be mathematically expressed as follows:

$$\text{ag}_4(t+1) = q(t) * \text{ag}_b(t) + (2r_8 - 1) \left(r_7 - r * x(t) - 2\text{LEVY}(D) \left(1 - \frac{t}{mn} \right) \right) \quad (20)$$

In Eq. (15), $q(t)$ is the quality function applied for equilibrating the search process.

Fitness selection is a vital issue in the AOA method. The solution encoder was utilized to assess the goodness of candidate results. Now, the accuracy values are the essential form used to design a FF.

$$\text{Fitness} = \max(P) \quad (21)$$

$$P = \frac{TP}{TP + FP}$$

(22)

From the expression, TP and FP are the true positive and the false positive values.

3 Experimental Validation

In this study, the vehicle classification outcomes of the ODLVDC-ITS method are tested under the benchmark VEDAI dataset [15, 16], comprising 3687 images with 9 classes as demonstrated in Table 1. Figure 2 shows the sample images of VEDAI dataset. Figure 3 represents the sample and detected images.

The confusion matrix of ODLVDC-ITS method under vehicle classification method is reported in Fig. 4. The results exhibit the enhanced outcome of the ODLVDC-ITS method with the classification of nine class labels.

Table 2 and Fig. 5 demonstrate an overall vehicle classification outcomes of the ODLVDC-ITS technique. The results are inspected under distinct number of epochs and class labels. On 500 epochs, the ODLVDC-ITS technique has obtained an average $accu_y$ of 99.37%, $prec_n$ of 96.17%, $reca_l$ of 92.57%, F_{score} of 94.16%, and MCC of 93.89%. Likewise, on 1000 epochs, the ODLVDC-ITS approach has attained an average $accu_y$ of 99.45%, $prec_n$ of 95.31%, $reca_l$ of 95.63%, F_{score} of 95.42%, and MCC of 95.12%. Likewise, on 1500 epochs, the ODLVDC-ITS algorithm has achieved an average $accu_y$ of 96.36%, $prec_n$ of 72.26%, $reca_l$ of 61.80%, F_{score} of 64.99%, and MCC of 63.88%. Finally, on 2000 epochs, the ODLVDC-ITS system has reached an average $accu_y$ of 99.37%, $prec_n$ of 96.98%, $reca_l$ of 92.74%, F_{score} of 94.58%, and MCC of 94.34%.

In Fig. 6, the TRACC and VLACC of ODLVDC-ITS technique are inspected on vehicle classifier outcome which is shown. The experimental value shows the

Table 1 Details of dataset

VEDAI dataset	
Class	No. of instances
Car	1340
Truck	300
Van	100
Pickup	950
Boat	170
Camping car	390
Other	200
Plane	47
Tractor	190
Total number of instances	3687



Fig. 2 Sample images VEDAI dataset

ODLVDC-ITS method has showed superior outcome with highest values of TRACC and VLACC. Notably, the ODLVDC-ITS method has reached maximal TRACC outcomes.

In Fig. 7, the TRLS and VLLS of ODLVDC-ITS technique are tested on vehicle classifier outcome is demonstrated. The experimental value exhibited the ODLVDC-ITS method has superior outcome with minimal values of TRLS and VLLS. Seemingly, the ODLVDC-ITS method has resulted in minimized VLLS outcomes.

A detailed PR examination of ODLVDC-ITS method under test database is demonstrated in Fig. 8. The figure exhibited the ODLVDC-ITS approach has resulted in improved values of PR values under all classes.

In Fig. 9, the brief ROC analysis of the ODLVDC-ITS technique under test data is shown. The outcomes specified the ODLVDC-ITS approach have showed its capability in classifying different classes.

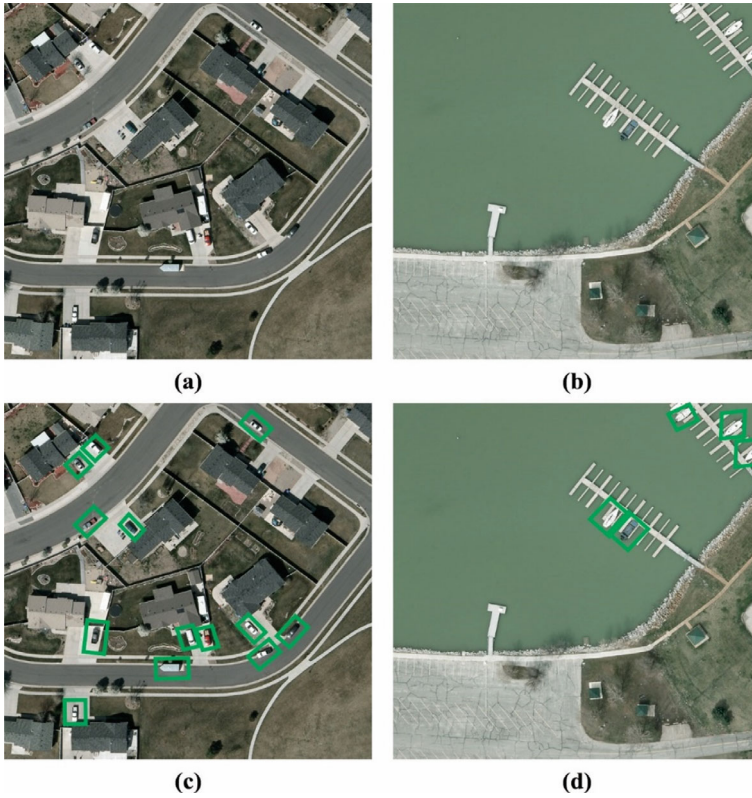


Fig. 3 a and b Sample images, c and d Detected images

In Table 3 and Fig. 10, a brief comparative analysis of the ODLVDC-ITS method with other models of vehicle classification process is demonstrated [17]. The outcomes stated that the ODLVDC-ITS method has superior outcomes over other approaches. For instance, based on $accu_y$, the ODLVDC-ITS technique obtains increasing $accu_y$ of 99.45% while the YOLO-V2, SSD, R-DFPN, RFCN, FADet, and ICN Models attain reducing $accu_y$ of 94.15%, 97.19%, 94.75%, 94.29%, 97.03%, and 94.23% correspondingly. In the meantime, based on $prec_n$, the ODLVDC-ITS technique obtains increasing $prec_n$ of 95.31%, whereas the YOLO-V2, SSD, R-DFPN, RFCN, FADet, and ICN approaches attain reducing $prec_n$ of 93.65%, 92.64%, 92.11%, 93.26%, 92.31% and 93.67% correspondingly. Eventually, based on $reca_l$, the ODLVDC-ITS method obtains increasing $reca_l$ of 95.63% while the YOLO-V2, SSD, R-DFPN, RFCN, FADet, and ICN methods attain reducing $reca_l$ of 93.57%, 92.88%, 92.72%, 92.82%, 92.06% and 93.65% correspondingly.

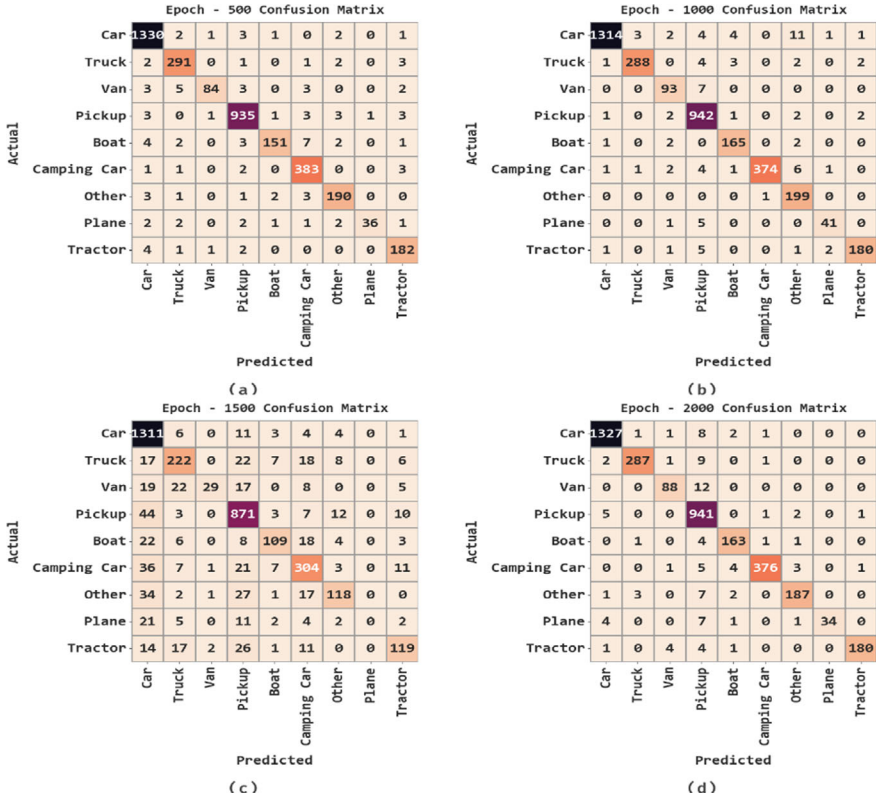


Fig. 4 Confusion matrices of ODLVDC-ITS algorithm, **a** Epoch 500, **b** Epoch 1000, **c** Epoch 1500, **d** Epoch 2000

These results confirmed the enhanced classification outcomes of the ODLVDC-ITS method on vehicle classification process using RSIs.

Table 2 Vehicle classifier outcome of ODLVDC-ITS method with different epochs

No. of epoch	Class	Accu _y	Prec _n	Rec _{al}	F _{score}	MCC
500	Car	99.13	98.37	99.25	98.81	98.13
	Truck	99.38	95.41	97.00	96.20	95.86
	Van	99.48	96.55	84.00	89.84	89.81
	Pickup	99.13	98.21	98.42	98.32	97.73
	Boat	99.35	96.79	88.82	92.64	92.39
	Camping car	99.32	95.51	98.21	96.84	96.47
	Other	99.43	94.53	95.00	94.76	94.46
	Plane	99.67	97.30	76.60	85.71	86.18
	Tractor	99.40	92.86	95.79	94.30	94.00
	Average	99.37	96.17	92.57	94.16	93.89
1000	Car	99.16	99.62	98.06	98.83	98.18
	Truck	99.57	98.63	96.00	97.30	97.07
	Van	99.54	90.29	93.00	91.63	91.40
	Pickup	99.00	97.01	99.16	98.07	97.41
	Boat	99.62	94.83	97.06	95.93	95.74
	Camping car	99.54	99.73	95.90	97.78	97.54
	Other	99.32	89.24	99.50	94.09	93.89
	Plane	99.73	91.11	87.23	89.13	89.01
	Tractor	99.59	97.30	94.74	96.00	95.80
	Average	99.45	95.31	95.63	95.42	95.12
1500	Car	93.60	86.36	97.84	91.74	87.00
	Truck	96.04	76.55	74.00	75.25	73.12
	Van	97.97	87.88	29.00	43.61	49.83
	Pickup	93.98	85.90	91.68	88.70	84.68
	Boat	97.69	81.95	64.12	71.95	71.34
	Camping car	95.31	77.75	77.95	77.85	75.22
	Other	96.88	78.15	59.00	67.24	66.35
	Plane	98.73	00.00	00.00	00.00	00.00
	Tractor	97.04	75.80	62.63	68.59	67.39
	Average	96.36	72.26	61.80	64.99	63.88
2000	Car	99.29	99.03	99.03	99.03	98.48
	Truck	99.51	98.29	95.67	96.96	96.70
	Van	99.48	92.63	88.00	90.26	90.02
	Pickup	98.24	94.38	99.05	96.66	95.52
	Boat	99.54	94.22	95.88	95.04	94.81
	Camping car	99.51	98.95	96.41	97.66	97.40
	Other	99.46	96.39	93.50	94.92	94.65

(continued)

Table 2 (continued)

No. of epoch	Class	Accu _y	Prec _n	Reca _l	F _{score}	MCC
	Plane	99.65	100.00	72.34	83.95	84.90
	Tractor	99.67	98.90	94.74	96.77	96.63
	Average	99.37	96.98	92.74	94.58	94.34

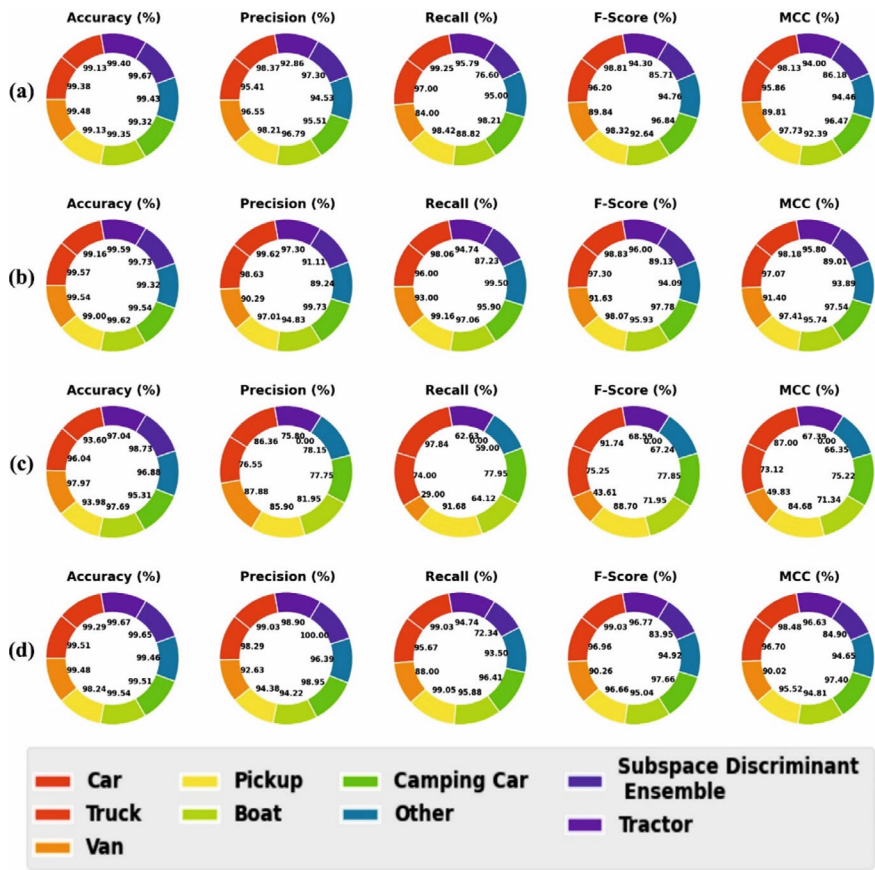


Fig. 5 Performance analysis of, a Epoch 500, b Epoch 1000, c Epoch 1500, d Epoch 2000

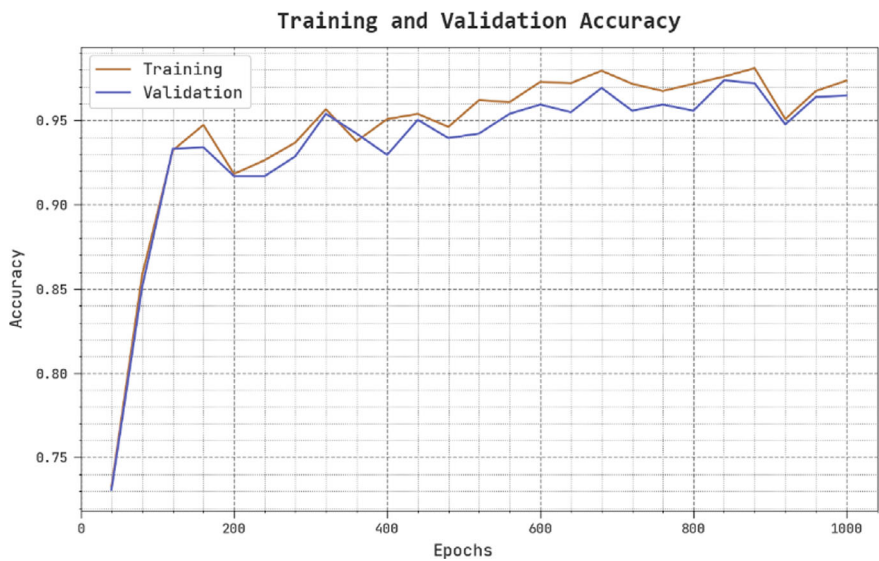


Fig. 6 TRACC and VLACC analysis of ODLVDC-ITS method

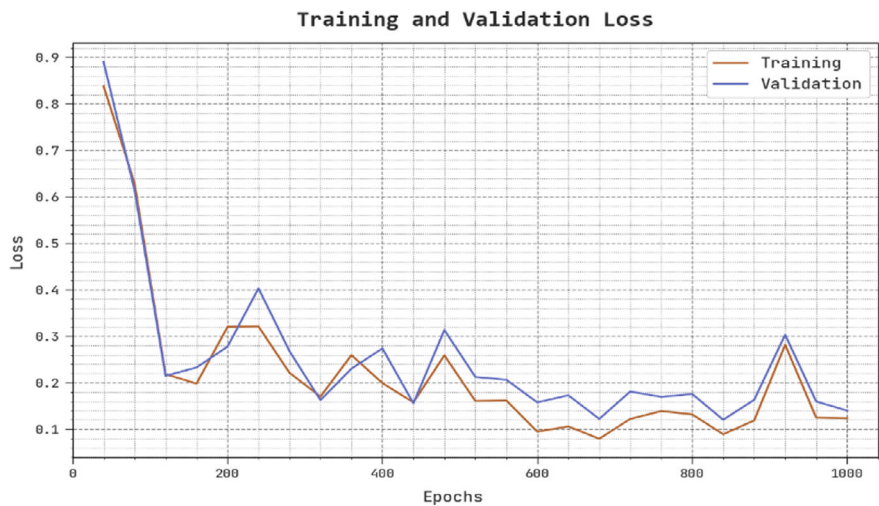


Fig. 7 TRLS and VLLS examination of ODLVDC-ITS method

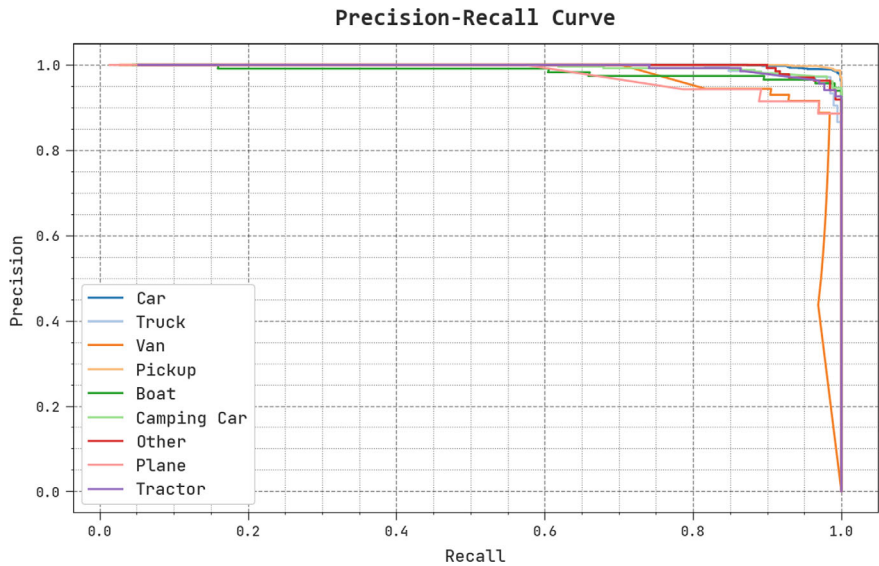


Fig. 8 PR curve of ODLVDC-ITS method

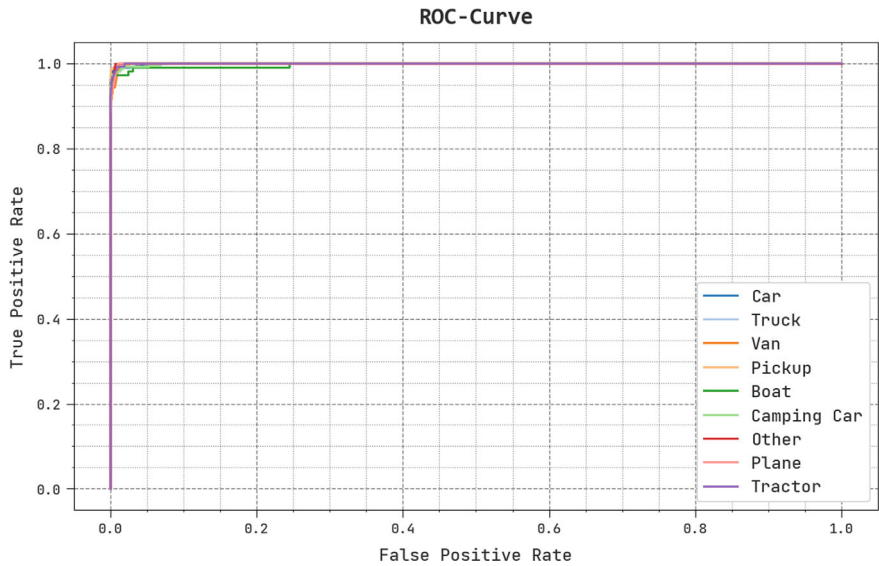


Fig. 9 ROC analysis of ODLVDC-ITS method

Table 3 Comparative analysis of ODLVDC-ITS technique with other approaches

Methods	Accu _y	Prec _n	Reca _l	F _{score}
ODLVDC-ITS	99.45	95.31	95.63	95.42
YOLO-V2	94.15	93.65	93.57	92.51
SSD Model	97.19	92.64	92.88	91.98
R-DFPN	94.75	92.11	92.72	90.41
RFCN	94.29	93.26	92.82	92.22
FADet	97.03	92.31	92.06	92.76
ICN model	94.23	93.67	93.65	93.31

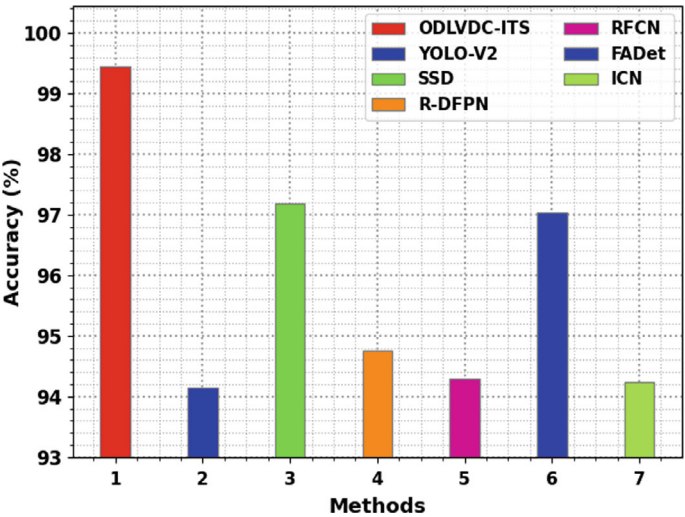


Fig. 10 Accu_y analysis of ODLVDC-ITS approach with other techniques

4 Conclusion

In this study, we have established a new ODLVDC-ITS method for automated recognition and classification of vehicles present in the RSI. The ODLVDC-ITS approach comprises three important processes such as vehicle detection, parameter tuning, and vehicle classification. Initially, the ODLVDC-ITS method employs YOLO-v5 object detector to identify the vehicles in the images and its hyperparameters can be tuned by coyote optimization algorithm. Next, the AOA with DRN technique is applied for vehicle classification process. The stimulation values of the ODLVDC-ITS model are tested on open access dataset and the outcomes are inspected under various measures. A widespread experimental analysis stated the promising performance of the ODLVDC-ITS method on vehicle classification algorithm. In the future,

the vehicle classification outcome of the ODLVDC-ITS method can be extended by advanced DL classifiers.

References

1. Chen Y, Qin R, Zhang G, Albanwan H (2021) Spatial temporal analysis of traffic patterns during the COVID-19 epidemic by vehicle detection using planet remote-sensing satellite images. *Remote Sens* 13(2):208
2. Valappil NK, Memon QA (2021) CNN-SVM based vehicle detection for UAV platform. *Int J Hybrid Intell Syst*, (Preprint), pp1–12
3. Chen H, Luo Y, Cao L, Zhang B, Guo G, Wang C, Li J, Ji R (2019) Generalized zero-shot vehicle detection in remote sensing imagery via coarse-to-fine framework. In: *IJCAI*, pp 687–693
4. Karnick S, Ghalib MR, Shankar A, Khapre S, Tayubi IA (2022) A novel method for vehicle detection in high-resolution aerial remote sensing images using YOLT approach. *Multimed Tools Appl* 1–16
5. Yan P, Liu X, Wang F, Yue C, Wang X (2021) LOVD: land vehicle detection in complex scenes of optical remote sensing image. *IEEE Trans Geosci Remote Sens* 60:1–13
6. Guo D, Zhu S (2019) Research on vehicle identification based on high resolution satellite remote sensing image. In: 2019 International conference on intelligent transportation, big data & smart city (ICITBS). IEEE, pp 62–65
7. Liu H, Ding Q, Hu Z, Chen X (2021) Remote sensing image vehicle detection based on pre-training and random-initialized fusion network. *IEEE Geosci Remote Sens Lett* 19:1–5
8. Lin HY, Tu KC, Li CY (2020) Vaid: an aerial image dataset for vehicle detection and classification. *IEEE Access* 8:212209–212219
9. Zhang J, Jia X, Hu J, Tan K (2021) Moving vehicle detection for remote sensing video surveillance with nonstationary satellite platform. *IEEE Trans Pattern Anal Mach Intell* 44(9):5185–5198
10. Koga Y, Miyazaki H, Shibasaki R (2018) A CNN-based method of vehicle detection from aerial images using hard example mining. *Remote Sens* 10(1):124
11. Wang J, Chen Y, Dong Z, Gao M (2022) Improved YOLOv5 network for real-time multi-scale traffic sign detection. *Neural Comput Appl* 1–13
12. Yuan Z, Wang W, Wang H, Yildizbasi A (2020) Developed coyote optimization algorithm and its application to optimal parameters estimation of PEMFC model. *Energy Rep* 6:1106–1117
13. Al-qaness MA, Ewees AA, Elaziz MA, Samak AH (2022) Wind power forecasting using optimized dendritic neural model based on seagull optimization algorithm and Aquila optimizer. *Energies* 15(24):9261
14. Taha AA, Abouroumia HO, Mohamed SA, Amar LA (2022) Enhancing the lifetime and energy efficiency of wireless sensor networks using Aquila optimizer algorithm. *Future Internet* 14(12):365
15. <https://downloads.greyc.fr/vedai/>
16. Razakarivony S, Jurie F (2016) Vehicle detection in aerial imagery: a small target detection benchmark. *J Vis Commun Image Represent* 34:187–203
17. Wang B, Gu Y (2020) An improved FBPNN-based detection network for vehicles in aerial images. *Sensors* 20(17):4709

A Portable and Robust Smart Mobile Prototype for Millet Crop Productivity Assessment



Akansha Tiwari, Vaibhav Kumar, Soumya Sahoo, Sushruta Mishra, and Mohammed Al-Farouni

Abstract The design aims to revise millet crop monitoring and productivity assessment through the integration of smartphone technology. By using the ubiquitous presence of smartphones, this innovative approach seeks to give growers and agrarian stakeholders with a stoner-friendly and accessible tool for directly assessing millet crop productivity. The model encompasses colorful factors, including data collection, analysis, and visualization, all seamlessly integrated into a smartphone operation. Through this operation, growers can capture pivotal data points related to millet civilization, similar as soil quality, rainfall conditions, pest infestations, and growth stages. These data are also reused using advanced algorithms and remote seeing ways to induce comprehensive assessments of crop productivity and health. The model's crucial features include real-time monitoring capabilities, prophetic analytics for early discovery of implicit issues, and customizable reporting functionalities. By empowering growers with practicable perceptivity deduced from smartphone-enabled technology, this design aims to enhance millet crop operation practices, optimize resource application, and eventually contribute to bettered yields and livelihoods in millet husbandry communities. By leveraging smartphone technology, remote sensing, and machine learning, our proposed model offers a cost-effective, scalable, and efficient solution for millet crop productivity assessment.

A. Tiwari · V. Kumar · S. Mishra (✉)

Kalinga Institute of Industrial Technology, Deemed to Be University, Bhubaneswar, India
e-mail: sushruta.mishrafcs@kiit.ac.in

S. Sahoo

Department of Computer Science and Engineering, C.V. Raman Global University, Bhubaneswar, India
e-mail: soumya.sahoo@cgu-odisha.ac.in

M. Al-Farouni

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University, Najaf, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Al Diwaniyah, Al Diwaniyah, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Babylon, Babylon, Iraq

This innovative approach has the potential to empower farmers, agricultural experts, and policymakers with actionable insights to enhance millet cultivation practices, optimize resource utilization, and improve food security in millet-growing regions.

Keywords Millets · Smart phone · Edge computing · Millets production · App development · Monitoring soil parameters

1 Introduction

Millet is a drought-tolerant crop that may thrive in dry, arid environments when other crops would fail. In addition, it is a nutrient-dense grain with a high fiber content and a high mineral content. It's a widely consumed grain variety across the globe. It is a staple cuisine throughout Asia and many other regions of the world. An estimated 1.2 billion people worldwide eat millet as part of their diet, according to the World Food Program. Millet crops have several important features that make them valuable in agriculture and nutrition. Millets can acclimatize to numerous ecological conditions and grow well in rain-fed conditions. In thirsty climates, they need little water, toxin and fungicides. Health-promoting grains: They've better micronutrients and bioactive flavonoids compared to other grains. Wheat has a low glycemic indicator (GI) and is also linked to precluding diabetes. They're a good source of minerals similar as iron, zinc and calcium. Millet is gluten-free and can be eaten by people with celiac complaint. Millet is effective in controlling and precluding hyperlipidemia and heart complaint. Studies have shown that millets help reduce weight, body mass indicator and blood pressure. In India, millets are frequently eaten with sap to add protein, increase amino acid content and ameliorate the overall insipidity of the protein. Millet-grounded value addition in the ready mess order is easy and accessible for the civic population. Millet has two uses as food and fodder, making it a better system of husbandry. Millet civilization helps reduce carbon footmark. Millet product has remained fairly stable over the once many times, with an estimated product of 28 million metric tons in 2020. Africa is one of the largest producers of Millet then comes Asia. India is the leading country in terms of consumption, while Niger and China are also important consumers. Millet production is also significant in countries like Burkina Faso, Mali, and Senegal. Although it is not a staple food crop in developed nations, it plays an essential role in the diets of many people living in developing countries. It can be grown in arid places with scanty rainfall where other crops cannot grow; thus millet is drought-tolerant crop unlike most cereals. Additionally, it is highly nutritive because it contains minerals and fiber.

Because of these causes, millet will remain a significant food crop that can be raised in the near future. In recent days, India has been increasing its production levels of millets. Millets have become economically unviable for farmers in India and thus they have reduced their planting areas for this crop which was touted as being drought-resistant. The Indian government has also included millet among crops to promote under its National Food Security Mission. These factors are responsible

for the expected increase in millet production in India in the near future. There are a number of distribution points for millets all over India and it happens to be one of the leading suppliers and exporters along with other major players such as Niger and Burkina Faso. Main millet producing states in India include Rajasthan, Maharashtra, Karnataka, Andhra Pradesh and Madhya Pradesh where many farmers grow it both for local consumption and international markets as well. Apart from main producers, there are also some regions where small quantity of this grain is grown but not enough to sustain themselves or any other market externally even at lower prices when compared to other grains like maize or beans. These regions include the countries of Uttar Pradesh, Bihar and Madhya Pradesh. Raising awareness about the nutritional benefits of millets is essential to promote their consumption and support millet farmers. Addressing these challenges requires concerted efforts from policy-makers, researchers, farmers, and other stakeholders to promote sustainable millet production practices, improve market access and value chains, enhance resilience to climate change, and raise awareness about the nutritional benefits of millets.

From Fig. 1, Area Rajasthan contributes the highest area under millet cultivation, accounting for 36% of the total area in India. Other major contributors include Karnataka (18.9%), Maharashtra (13.6%), Madhya Pradesh (8.9%), and Gujarat (4.6%). The remaining states collectively contribute less than 20% of the total area. Rajasthan also leads in millet production, contributing 26.3% of the total production in India.

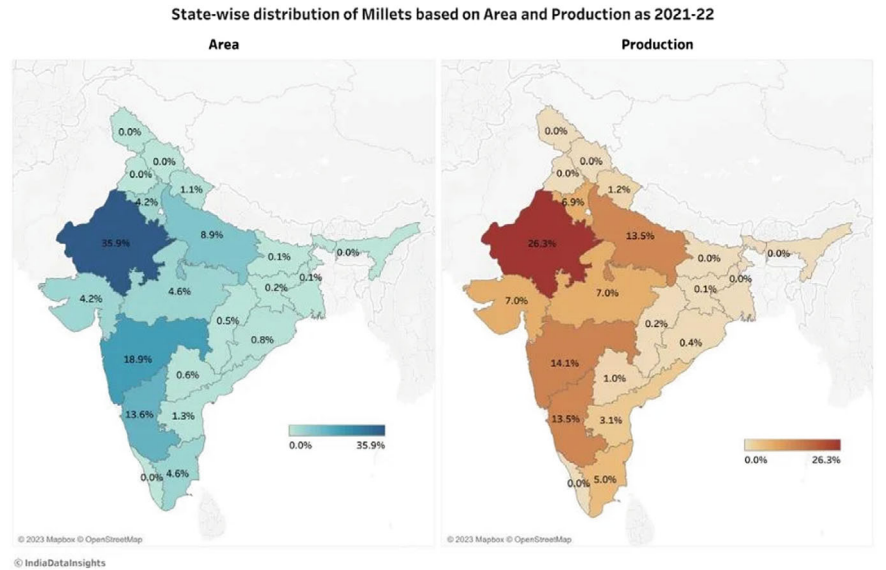


Fig. 1 State-wise production of millets

From Table 1, Karnataka follows closely with 14.1%, followed by Maharashtra (13.5%), Gujarat (7.0%), and Madhya Pradesh (4.2%). Similar to the area distribution, the remaining states contribute less than 20% of the total production.

Millets are sustainable options to achieve food and nutritive security. Millets cultivated and consumed traditionally in India displayed a decline in area and product except for bajra. These crops are getting wide attention because of their health benefits and terrain-friendly nature. This study examines if the force would be sufficient to meet the demands in the near future while also assaying the trends in the area, product, productivity, and consumption of millets in India at the public and state position. The study indicates that there will be a fat in the case of bajra and small millets, while in the case of jowar and ragi there will be a net deficiency in the coming times which is a matter of concern. There is a compass for accelerating the product in the short run

Table 1 State-wise production of millets

State/UT	2019–20			2020–21			2021–22*			Normal		
	A	P	Y	A	P	Y	A	P	Y	A	P	Y
Andhra Pradesh	0.22	0.22	1000	0.22	0.19	864	0.14	0.11	786	0.18	0.15	804
Assam	0.05	0.03	619	0.05	0.03	656	0.05	0.03	620	0.05	0.03	613
Bihar	0.02	0.02	755	0.02	0.02	753	0.02	0.01	754	0.04	0.03	752
Chhattisgarh	0.63	0.19	300	0.85	0.22	258	0.52	0.26	503	0.75	0.23	311
Gujarat	0.00	0.00	560	0.08	0.13	1541	0.12	0.23	1988	0.07	0.11	1535
Himachal Pradesh	0.05	0.05	890	0.02	0.02	972	0.02	0.02	972	0.04	0.03	872
Jammu and Kashmir	0.00	0.00	0	0.00	0.00	0	0.00	0.00	0.00	0.01	0.01	497
Karnataka	0.49	0.37	765	0.26	0.20	778	0.27	0.21	789	0.31	0.24	786
Kerala	0.00	0.00	754	0.00	0.00	745	0.00	0.00	1206	0.00	0.00	825
Madhya Pradesh	0.84	0.74	881	0.78	0.69	890	0.89	0.77	860	0.98	0.85	867
Maharashtra	0.41	0.22	536	0.37	0.17	452	0.44	0.17	396	0.42	0.20	466
Odisha	0.33	0.17	517	0.35	0.18	511	0.38	0.20	511	0.34	0.17	514
Rajasthan	0.10	0.05	518	0.07	0.04	660	0.07	0.04	633	0.09	0.06	717
Tamil Nadu	0.26	0.37	1444	0.24	0.31	1247	0.23	0.27	1179	0.24	0.32	1335
Telangana	0.01	0.02	1711	0.00	0.00	0	0.00	0.00	0.0	0.00	0.00	1711
Uttar Pradesh	0.08	0.06	728	0.12	0.09	765	0.07	0.05	714	0.08	0.06	723
Uttarakhand	0.53	0.71	1339	0.49	0.71	1449	0.47	0.73	1559	0.53	0.72	1368
West Bengal	0.06	0.06	907	0.01	0.00	502	0.02	0.02	1134	0.03	0.03	885
Other#	0.49	0.43	876	0.50	0.46	914	0.53	0.61	1162	0.48	0.48	993
All India	4.58	3.71	809	4.44	3.47	781	4.23	3.75	885	4.65	3.73	802

by perfecting the productivity without adding the area through depleting new kinds, expanding pukka seed distribution, and perfecting crop operation practices.

2 Related Works and Literature

As a result of this study, we have learned various techniques for collecting data, pre-processing, disease detection, classification and visualization methodologies used in it. In the conventional methodologies, the farmers do it themselves by distinguishing between plants that are sick and those that are healthy [1]. Devoid of the monitoring of elemental factors such as soil kind, dampness, temperature and macro and micro-nutrient levels in soils, plus the nutrient requirements of different growth stages in a crop plant. Even more so, these are time consuming measures that demand plenty of human input. Also, farmers require specialists to give them guidance on how to diagnose diseases in their plants. The adoption of edge computing has simplified the automation of plant disease detection and identification [2]. This author have suggested a model in this research for capturing of data. They employed the use of drone to capture these images over a wide area within a short time span. They utilized support vector machine (SVM) for classification of diseases in rice crops. However, their system did not provide for on-demand capture of real-time monitoring and predictions. In addition, IOT categorization of healthy and infected leaves was done by authors in [3]. Sensors were placed at different points to monitor soil quality, temperature and humidity. The camera was used to take photographs of the crop plants. In order to store and process captured data for real-time prediction, they put together the sensors interface with the Raspberry Pi camera. Images clustering was done using the K -means algorithm and pixel masking step was used to find out if the leaf is disease or healthy. Deep transfer learning also has a number of applications which were explored by another research group [4]. Detection of downy mildew disease in pearl millet was done by them using VGG-16 that is pre-trained [5, 6]. They said that from their experiments proved that deep transfer learning is able to extract the most critical information. It is possible to use these features of the pre-trained network for other purposes. These features are then reused by the pre-trained networks, which keep training on more data sets. The model works better this way. When it comes to fine-tuning a model based on its size and type of dataset, transfer learning plays a vital role as well. The authors also claimed that transfer learning is helpful to avoid overfitting and to improve the model's predictive capacity [7]. Our world has been revolutionized by the era of IoT. Precision agriculture is known as an eco-friendly, sustainable and profitable way to improve agricultural production and quality. For monitoring plants, infection detection approaches demand a huge amount of data and information. Analyze each cluster with alternative methods one by one to examine and indicate all possible results for dealing with such data [8]. For monitoring a grape disease detection within a closed circuit camera prototype, authors in [9] used Hidden Markov Model. In the agricultural field, sensors are deployed for temperature and moisture information gathering purposes. Then these data clusters

are transmitted to the server using Zig-Bee. HMC is often estimated and judged by farmers depending on their experience which is subjective due to psychological factors or the condition of neighboring farms. However, growth conditions (e.g., rice variety, time of planting and farming method) differ between paddies; therefore, it is hard to establish one right standard for assessment by experience alone. So also for managers of large-scale paddy fields such as agricultural contracts and farmer associations this means spending much time and human resources to go through each field's growth status surveys. This study used smartphones to gauge the GMC and anticipate the best HMC for scheduling a harvest in order to improve livelihoods by increasing income and reducing costs. Grain maturity starts with heading: gradually changing from green to golden yellow grain color. This is when carbohydrates are being filled into grains as a result of photosynthetic reactions; it involves four stages: milk, soft dough, hard dough and maturity [10]. The milk stage of growth is such that the grains develop and fill up with milky juice inside them. Grains in the soft dough stage are fuller, although they remain to be soft. Grains in the hard dough stage turn from green to yellow; they are fully filled and gradually dried, and the average GMC ranges between 25 and 30% as the OHD approaches. Most of them are filled and dried at maturity but late maturing ones are still in filling stage; over time, greenness on leaves and clumps slowly fades away while GMC in the main stem is about 15–18%, with an average GMC range lying between 18 and 21%. The operational implementation of the Pradhan Mantri Fasal Bima Yojana (PMFBY), a national crop insurance scheme in India, can be enhanced with advanced technology especially regarding accuracy and timeliness of crop yield estimates used by Gram Panchayat (GP) for determining yield losses. This study was carried out as a pilot test for PMFBY in the kharif season of 2018 where the Terrestrial Observation and Prediction System (TOPS)-based technologies were tested and deployed to estimate GP level crop yields of bajra (pearl millet, *Pennisetum glaucum*) in Firozabad District of Uttar Pradesh and rice (*Oryza sativa*) in Kendujhar District of Odisha. Crop extent was mapped using a combination of Synthetic Aperture Radar and optical data [11].

3 Proposed Model

The proposed model (Fig. 3) is a service model that uses a smartphone-based data collection system to provide farmers with information and resources on how to grow millet according to the soil type and climate of their area. The model consists of several components, which I am going to explain below. Firstly, the data gathered through a smartphone forms the basis of this model. Such data can include soil type, climate and other environmental factors that may affect growth of millet. Sensors and technologies integrated into smartphones are used in collecting such data hence enabling easy access to information for farmers about their crops as well as growing environment [12–14]. This information after gathering it is analyzed and then put into use by giving recommendations on how farmers should set up their crops for optimal growth. Here, the right millet variety will be selected based on the kind

4 Result Analysis

The line graph (Fig. 4) presents the soil fertility index of a millet crop as it changes within one week. The days of the week are shown on the x -axis, while the y -axis indicates fertility index. This index ranges from 0 to 100, where higher index values represent higher fertility. On Monday, the soil fertility index for millet crops was highest at 72 on the graph before falling to 62 by Sunday. The overall trend of this graph is downward but there are fluctuations in it is almost everyday. It is important to note that this graph only shows the soil fertility index which is just one measure of crop health and other factors such as water availability and temperature can also affect crop growth.

From the above graph (Fig. 5), throughout the week, the pH of the soil for a millet crop is shown on this graph. The x -axis for the graph represents each day of the week from Monday to Sunday while y -axis represents pH level as such. On a scale of 0–14, pH measures how acidic or basic a soil is; it is considered neutral at 7, anything below shows it is acidic and above 7 shows alkalinity.

The graph reveals that at Monday, the beginning time, millet crop's soil had a pH of 7.0 but by Sunday, it had come down to 6.5 showing that over this period the soil was slightly acidified.

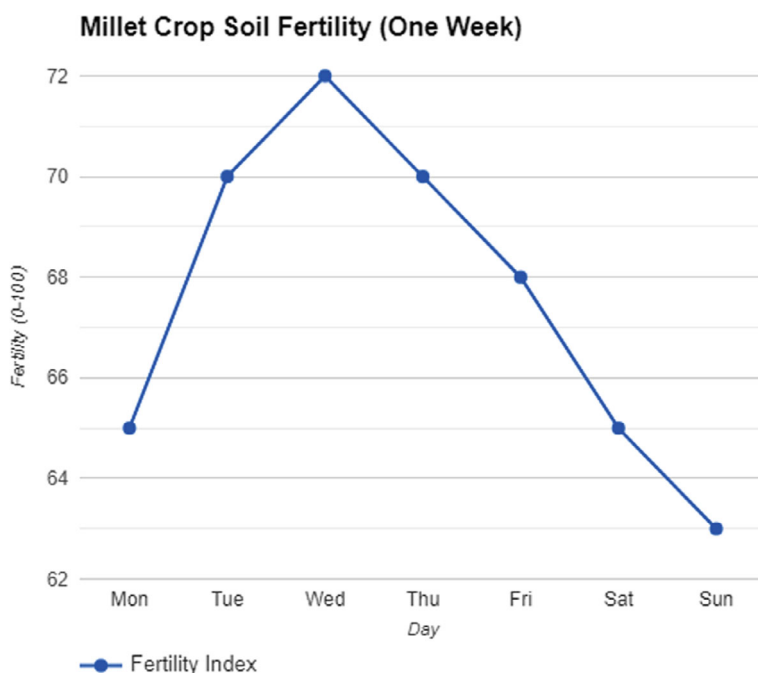


Fig. 4 Soil fertility analysis of one week

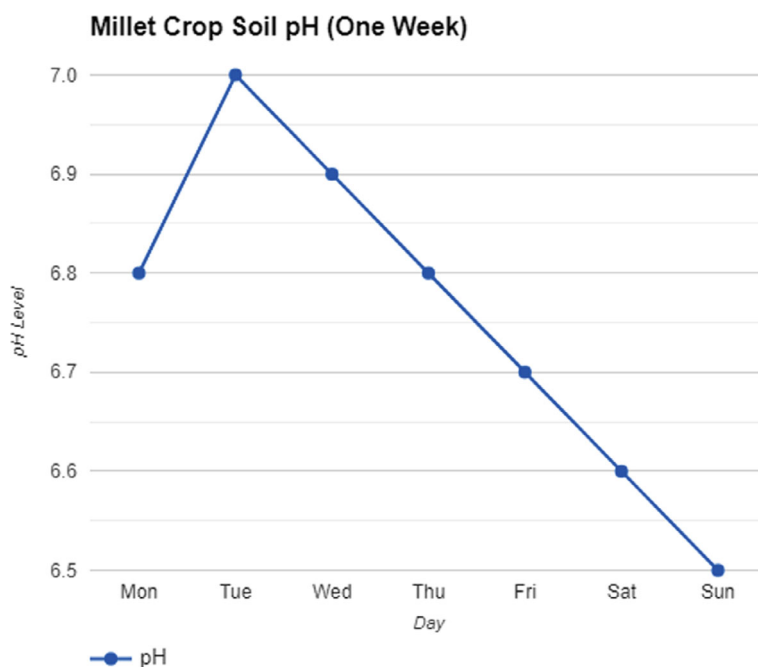


Fig. 5 Soil pH analysis of one week

The graph (Fig. 6) illustrating the soil moisture of a millet crop throughout a week is presented above. The x -axis gives the days of the week while the y -axis shows the soil moisture in millimeters of water per square inch.

On Monday, soils had 48 mm of water/sq. inch but this decreased to 40 mm by Sunday. This implies that there was an 8 mm loss of water from the ground within seven days.

We valuated five different crops, millets being the most efficient with the computational delay of 1.9 s. Therefore, the average mean computational delay of all the different crops computed came to be 2.74 s. From the above Table 2, we can see that we have done computational delay analysis of several crops that include; wheat, rice, maize, soybeans and millet. It is observed from the table that millet has a greater computational delay of 3.50 s. The second highest crop with regard to its computational delay is Wheat having 3.25 s. While rice is third on the list with a 3.00-s computational delay across all other commodities it is corn with a computation delay rate of 2.75 s which is fourth in this group, and finally, soybeans are ranked last as those vegetables having the shortest life span within their potential growth process having only two and a half seconds (2.25 s) of computing delays. The first reason why millet has the lowest computational delay is because it is more challenging to model compared to other crops in this chart. Another explanation may be that less accurate data was available for modeling millet than was available for modeling other crops. However, one should note that speed or rather computation delays do not

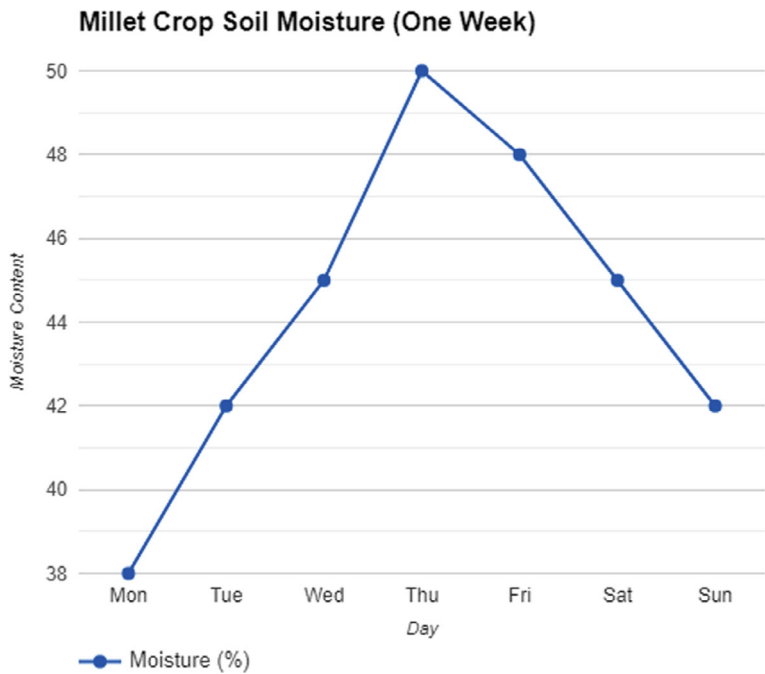


Fig. 6 Soil moisture analysis of one week

Table 2 Computational delay analysis for different crops using the model

Crop	Computational time (in s)	Processing time (s)
Wheat	2.5	1.9
Rice	3.1	2.2
Corn	2.8	2.4
Soyabean	3.4	3.1
Millet	1.9	1.7

constitute all measures for good crop models and they must also encompass aspects like reliability of the model itself plus its ability to give proper yield estimates.

5 Conclusion

A model that uses a smartphone to assess the productivity of the millet crop in an interesting way, merging traditional farming methods with modern technologies with a great approach. This is achieved by use of phone applications and remote sensing techniques. By merging smartphones, which facilitate real-time data collection from

farmers and agricultural experts and providing such information quickly, it enhances the process of monitoring while at the same time giving insights to farmers on their crops' health and growth trends. This has also been improved through incorporating satellite images in the model to ensure crop performance are well assessed. The model incorporates satellite imagery among other remote sensing tools that captures diverse data concerning millet farming such as soil moisture content, pest infestations and so on. Therefore stakeholders have every chance to decide wisely considering crop management strategies hence increasing yields and cutting potential losses. And more, the smartphone-enabled model promotes inclusiveness in the agricultural sector as it provides farmers with access to cutting-edge technology no matter their location or social status. By giving equal access to information and resources, this model helps farmers to adopt better farming practices that are sustainable and efficient for the purpose of food security and rural development. To sum up, putting in place a program for millet crop productivity assessment through a smartphone-enabled method represents an important milestone in agricultural innovation. This model provides a practical solution based on smartphones and remote sensing technologies to tackle challenges faced by farmers when monitoring and optimizing crop yields. Such initiatives have potential to bring about changes, enhance livelihoods and ensure future food system sustainability as we continue embracing technological driven approaches in agriculture.

References

1. Singh R, Singh GS (2017) Traditional agriculture: a climate-smart approach for sustainable food production. *Energy Ecol Environ* 2:296–316. [Google Scholar] [CrossRef]
2. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digital Health* 10:20552076241256732
3. Thorat A, Kumari S, Valakunde ND (2017) An IoT based smart solution for leaf disease detection. In: *Proceedings of the 2017 international conference on big data, IoT and data science (BID)*, Pune, India, 20–22 December 2017. [Google Scholar] [CrossRef]
4. Coulibaly S, Kamsu-Foguem B, Kamissoko D, Traore D (2019) Deep neural networks with transfer learning in millet crop images. *Computer Ind* 108:115–120. [Google Scholar] [CrossRef] [Green Version]
5. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv* 2014, [arXiv:1409.1556](https://arxiv.org/abs/1409.1556). [Google Scholar]
6. Shetty HS, Raj SN, Kini KR, Bishnoi HR, Sharma R, Rajpurohit BS, Yadav OP (2016) Downy mildew of pearl millet and its management. *Indian Counc Agric Res Mandor Jodhpur* 342304:55. [Google Scholar]
7. Feng JZB, Li GZ (2019) DCNN transfer learning and multi-model integration for disease and weed identification, vol 2. Springer, Singapore [Google Scholar]
8. Basori AH, Mansur ABF, Riskiawan HY (2020) SMARF: smart farming framework based on big data, IoT and deep learning model for plant disease detection and prevention. *Appl Comput Support Ind Innov Technol* 1174:44–56. ISSN 1865-0929
9. Patil SS, Thorat SA (2016) Early detection of grapes diseases using machine learning and IoT. In: *Second international conference on cognitive computing and information processing (CCIP)*, 2016, pp 1–5

10. Kuehner EC, Alvarez R, Paulsen PJ, Murphy TJ (1972) Production and analysis of special high-purity acids purified by subboiling distillation. *Anal Chem* 44:2050–2056. [Google Scholar] [CrossRef]
11. Aggarwal PK, Chand R, Bhutani A, Kumar V, Goel SK, Rao KN et al (2016) Report of the task force on enhancing technology use in agriculture insurance (Technical Report). National Institute for Transforming India, Aayog, New Delhi, pp. 26
12. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
13. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-DETECTOR: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
14. Pranjali P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. In: AIP conference proceedings, vol 2919(1). AIP Publishing
15. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber-physical system. *Internet Things* 24:100978
16. Mishra S, Voley DR, Bohra N, Alfarhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306
17. Jain S, Tripathy HK (2024) Machine learning methods for the timely identification of autism spectrum disorder in toddlers. In: 2024 International conference on emerging systems and intelligent computing (ESIC). IEEE, pp 515–520
18. Kashyap P, Pareek A, Mishra S, Khan Z, Garg R, Tripathy HK (2024) Sentiment polarity analysis of twitter data using machine learning models. In: International conference on innovative computing and communication. Springer Nature Singapore, Singapore, pp 623–635

Prognostic Analysis of Meteorological Data for a Small Solar Farm Using Machine Learning



Phuoc Quy Phong Nguyen, Prabhu Paramasivam, Huu Cuong Le,
Duy Tan Nguyen, Van Phuc Nguyen, and Thanh Nam Dang

Abstract Determining the economic feasibility of a PV-based energy-producing system depends on accurate prediction of the performance and energy output of solar (PV) farms. Based on three environmental input values such as sun irradiation, wind speed, and ambient air temperature, this study endeavor to create a prognostic model for the annual power output yield of an installed PV system. Four machine learning methods including XGBoost, AdaBoost, linear regression (LR), and decision tree (DT) were employed to model and predict the power generation. The models were compared with statistical metrics like mean squared error (MSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). With a training MSE of $1.218E + 07$ and a training R^2 of 0.98854, AdaBoost shows remarkable training accuracy, therefore suggesting its resilience in identifying fundamental patterns in the training data. During training phase, the decision tree model achieves 100% accuracy; yet it suffers from overfitting and shows notable prediction errors during testing. With a near-zero training MSE of 0.000683 and a flawless training R^2 of 1, XGBoost stands out; nonetheless, on the test data, it performs favorably to other models with a model testing MSE of $2.66E + 08$ and training R^2 of 0.8412. With

P. Q. P. Nguyen

PATET Research Group, Ho Chi Minh City University of Transport, Ho Chi Minh City, Vietnam
e-mail: phong@ut.edu.vn

P. Paramasivam

Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India

H. C. Le (✉) · D. T. Nguyen · V. P. Nguyen · T. N. Dang

Institute of Maritime, Ho Chi Minh City University of Transport, Ho Chi Minh City, Vietnam
e-mail: cuong.le@ut.edu.vn

D. T. Nguyen

e-mail: tan.nguyen@ut.edu.vn

V. P. Nguyen

e-mail: phuc.nguyen@ut.edu.vn

T. N. Dang

e-mail: nam.dang@ut.edu.vn

XGBoost's lower test MAPE of 6.36 than AdaBoost's 5.966, XGBoost edges out and provides better training accuracy and fewer prediction errors.

Keywords Solar energy · Machine learning · XGBoost · AdaBoost · Prognostics

1 Introduction

There exists a global challenge in the form of increasing greenhouse gases (GHG) that led to global warming and a shortage of clean energy [1–3]. Therefore, renewable energy sources (RES) are becoming increasingly sustainable and potential solutions [4–6]. More importantly, RES is helping to mitigate the impact that fossil fuel-based energy sources have on the environment [7–9]. By 2023, the proportion of renewable energy sources contributing to US yearly electric power output is predicted to rise from 20% in 2021 to 24% [10]. In the case of India, India's capacity for renewable energy was at 101.53 Gigawatts (GW), in September 2021. It almost count for 38% of the country's total installed capacity. By 2030, there should be 450 gigawatts (GW) of renewable energy installed; solar energy will make up around 280 GW, or more than 60% of this total [11, 12]. Given the fact that solar energy is virtually limitless, it is considered to be a form of renewable energy, and capable of converting solar radiation into other energy forms [13–18]. In addition, solar energy could be used for water distillation and drying system [19–23]. Indeed, solar energy systems have multiple applications, including the generation of electricity for homes and businesses [24, 25]. These are also employed in the battery-based storage of energy for use during times when the sun is not shining. It is also possible to use solar energy to power turbines, heat water make steam, and generate electricity through a process that is referred to as solar thermal [26–28]. Because it does not produce any emissions of GHG or any other impurities during the process of power generation, it is considered a clean energy source. Solar energy is currently one of the most sought-after forms of energy; in fact, it is utilized in a variety of industries, including residential, commercial, and industrial settings [29, 30].

In recent times, several studies have been published on this important topic. Carmona et al. [31] explored the use of phase change materials in a PV-based solar thermal collector. A low temperature and reliable operation of this system could be attained by the employment of PCM in the hybrid module (PVT-PCM). The average per day hybrid productivity of PVT-PCM module comes at 31.35%. Conversely, daily electricity efficiency of the PV module is 13.12%. This means that the PVT-PCM's ability to extract usable energy from sunlight increases by 20.45% overall. Abdullah et al. [32] explored the use of water-based PVT solar collectors. The test outcomes of the proposed design were investigated. It was then compared with a traditional PV panel without a cooling system. The study makes use of experimental data collected from the test setup having a water-based PVT system employing commercial software MATLAB. The outcomes of the simulations and the experiments were shown to be very close. It was observed that at the flow rate of 6 LPM combined with 500 W/

m^2 radiation, peak electrical efficiency was 11.5%. However, at the flow rate of 5 LPM and solar radiation as 1000 W/m^2 the peak thermal efficiency was reported 58.64%. The results show that thermal and overall efficiency rises with increasing mass flow rate and solar radiation but decreases with reaching the optimal value. The higher mass flow rate helps in reduction of cell temperature, which in turn helps in raising the electrical efficiency. Fu et al. [33] explored the performance of encapsulated PCM in a solar PV thermal system. Different operating settings were employed to assess the thermal and electrical efficiency of the PVT system using tap water or MPCM slurry. The thermal and electrical efficiency of PV cells could be increased by employing MPCM slurry while lowering their temperature. With MPCM slurry, the present PVT system outperforms both with a PCM layer and with pure tap water. The system's peak thermal as well as mean electrical efficiency were raised by 13.5% and 0.8%, respectively. As the complexity and demand of such PV-based systems are increasing, they are becoming more difficult to model using conventional approaches.

Using its capacity to examine huge datasets and spot complex patterns, machine learning (ML) offers a comprehensive method for energy model prediction [34–36]. Conventional approaches to model-predict solar energy generation often depend on a small number of factors and simple assumptions [37, 38], this sometimes led to less-than-ideal accuracy. In contrast, ML may improve forecast accuracy by using a wide range of variables like historical meteorological data, geographic information, panel properties, and temporal trends [39–41]. The capacity of machine learning to handle nonlinear connections and intricate interactions among variables is one of its main benefits in solar energy prediction. Seasonal variations, temperature swings, and cloud cover are only a few of the dynamic variables that affect solar energy production [42, 43]. More accurate and trustworthy predictions are produced by ML algorithms that excel in modeling this complexity: neural networks, random forests, and support vector machines [44, 45]. Further, machine learning models may become better with time. These models may learn and adapt, improving their prediction skills, when real-time data is included. With solar energy, where things may change quickly and unexpectedly, this is very helpful [46, 47]. Machine learning also improves prediction models' scalability. It makes large-scale data processing efficient and makes predictions over several periods and locations possible without requiring human recalibration. In order to better balance supply and demand, grid management and planning depend on this scalability [48, 49].

The literature review reveals that accurate prediction of solar energy is a difficult task. The difficulty increases as the solar energy system is becoming complex on account of increasing production and efficiency. The main objective of this study is to forecast the amount of power that will be generated by employing ensemble ML methods. This research is the first of its kind to investigate the exploitation of data collected for 3 years from photovoltaic plants.

2 Materials and Methods

2.1 Data Collection

In the present study, the data gathered from a solar power plant was employed for modeling. Using a Supervisory Control and Data Acquisition (SCADA) system, each string monitoring unit and inverter in the solar PV power plant are linked to one another. The embedded or integrated SCADA system of the PV power plant is responsible for the routine monitoring of energy output, solar irradiation, speed of wind, air temperature, and module temperature. In addition to storing vital operating data for the SPV power plant, the integrated SCADA system is responsible for monitoring the voltage, current, and power that are produced by each inverter.

The sun energy, or more specifically the worldwide slanted irradiation, is the primary input parameter that is being considered for the study. Monthly tilted irradiation (MTI) data are being gathered during the study years 2020–23 respectively. Wind speed affects the productivity of PV plants since it changes the heat transfer from PV modules, which in turn improves the efficiency of the PV process. Because it is tied to the temperature of the module, the temperature of the air outside has an effect on the efficiency of PV power plants. So basically, there were three control factors wind speed, Temperature, and MTI, and one response variable i.e., power generation.

2.2 Machine Learning

Linear Regression

Linear regression (LR) is employed as a baseline ML technique for the evaluation of more complex ML techniques. It is used for modeling the relationship between a dependent variable and one or more independent variables [50, 51]. It makes use of fitting a linear equation to observed data. The simplicity of use and interpretability of the LR approach make it a widely employed prognostic method.

Adaptive Boosting Regression

Adaptive Boosting, or AdaBoost for short, is an ensemble learning method wherein numerous weak learners' predictions are aggregated to produce a strong predictive model. By concentrating on cases that past classifiers misclassified, it gradually trains weak classifiers by changing their weights to increase accuracy [52, 53]. Emphasizing those with greater performance, the final model is a weighted sum of the weaker learners. AdaBoost improves poor model prediction capability and is resistant to overfitting. Nevertheless, its success relies on the choice of weak learners and the complexity of the data; it might be susceptible to noisy data and outliers [54, 55].

Decision Tree

Decision trees (DT) are a flexible and reliable ML technique that can be applied for classification as well as regression projects. The model divides the data into subsets creating a tree-like arrangement of decision nodes and leaf nodes [56, 57]. Whereas leaf nodes show the output, each decision node stands for a feature and a threshold. Decision trees incorporate non-linear connections and may manage numerical and categorical data as well. They do, however, overfit easily, particularly in complicated datasets. Often used to increase its generalization and resilience are pruning and ensemble approaches such as Random Forests or Gradient Boosting [58, 59].

Extreme Gradient Boosting

Extreme gradient boosting (XGBoost) is an powerful ensemble ML method distinguished in predictive modeling by great efficiency and performance. It creates a series of sequential decision trees in which each one uses gradient boosting to fix the errors of its successors [60, 61]. XGBoost addresses missing data well, offers parallel processing, and uses regularizing methods to avoid overfitting. For many different machine learning applications, its ability to record intricate data patterns and interactions makes it very successful. Although XGBoost needs careful hyper-parameter tweaking for best performance, its great accuracy and scalability make it extensively employed in competitive data science [62, 63].

The flow chart for ML implementation in the present study is illustrated in Fig. 1.

2.3 Model Evaluation

The models developed in this study were evaluated and compared using statistical metrics like MAPE, R^2 , and MSE. Also, Taylor's diagrams and violin plots were employed. Indicating their explanatory capacity, the R^2 provides information on the fraction of variance explained by each model. Better model performance in capturing the fundamental data patterns could be identified if the R^2 score is closer to 1. The MAPE evaluated the precision of forecasts by calculating the average absolute percentage difference between observed and anticipated values. The lower MAPE values suggest more accurate models. Emphasizing greater errors owing to

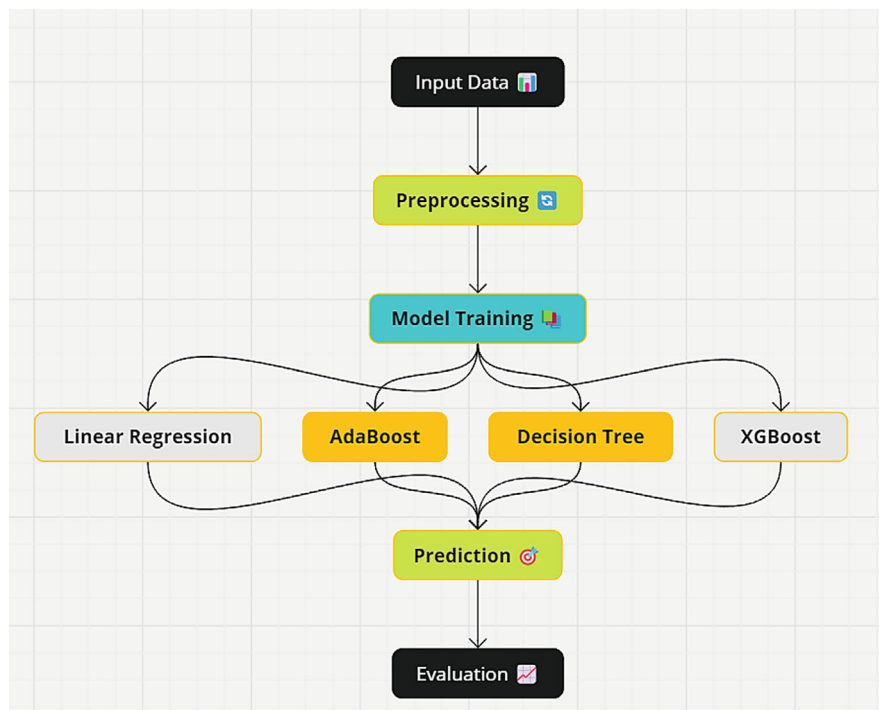


Fig. 1 Flowchart for the present study

their quadratic character, the MSE offers a gauge of the average squared difference between actual and forecasted values. Taylor’s diagrams helped to compare model performance and therefore provide a whole picture of model correctness and precision. Taylor’s diagram makes use of the correlation, standard deviation, and root mean squared error values. The distribution and probability density of the data were shown using violin plots, thereby stressing the skewness of the residuals. These charts provide a comprehensive evaluation of model performance through many metrics and visualization approaches by allowing a complete assessment of model resilience and variability.

3 Results and Discussion

3.1 Data Analysis and Correlation

The data collected in the study was used for the development of prognostic models. First of all, the data was analyzed for correlation and descriptive statistical analysis. In Table 1, the correlation matrix in respect of the data used is shown. The correlation heatmap is depicted in Fig. 2.

In the framework of solar power generation developed, the correlation matrix in Table 1 explores the interactions among four factors: wind speed, monthly tilted irradiation (MTI), ambient temperature, and power generation. Wind speed and MTI have a correlation value of -0.09 . This suggests that MTI is inversely correlated with wind speed. With a very weak positive association shown by the correlation coefficient of 0.08 between wind speed and temperature, changes in wind speed have little direct impact on temperature. With a -0.23 correlation between wind speed and power generation, wind speed indicates a modest negative relationship, meaning that higher wind speeds provide somewhat less solar power output. With a slight negative correlation with temperature (-0.37), the MTI indicates that lower temperatures are associated with higher degrees of MTI. MTI and electricity production have a considerable positive link shown by their correlation value of 0.91 . This emphasizes the importance of irradiation in solar power output as it implies that changes in MTI are very connected to increases in solar power generation. With a weakly negative link, temperature and power generation have a correlation value of -0.24 . This suggests that lower solar power generation is slightly correlated with higher temperatures, presumably because of solar panels' inferior efficiency at higher temperatures. MTI (0.91) has a strong positive correlation with the variable for power output; wind speed (-0.23) has a mild negative association; and temperature (-0.24) has a weak negative correlation. This shows that with higher MTI producing more power output, wind speed, and temperature have less of an influence even if MTI significantly affects solar power generation. The matrix demonstrates generally that MTI is the most important parameter positively affecting solar power production; wind speed and temperature have less substantial, negative effects on power output.

The descriptive analysis of the used data is listed in Table 2. The average wind speed is 2.024 m/s, with a standard deviation of 0.598 , suggesting reasonably stable wind conditions. The wind speed varies from 1.02 m/s to 3.15 m/s, with a tiny positive

Table 1 Correlation matrix

	Speed of wind, m/s	MTI, kWh/m ²	Temp., °C	Power gene., kWh
Speed of wind, m/s	1	-0.09	0.08	-0.23
MTI, kWh/m ²	-0.09	1	-0.37	0.91
Temp., °C	0.08	-0.37	1	-0.24
Power gene., kWh	-0.23	0.91	-0.24	1

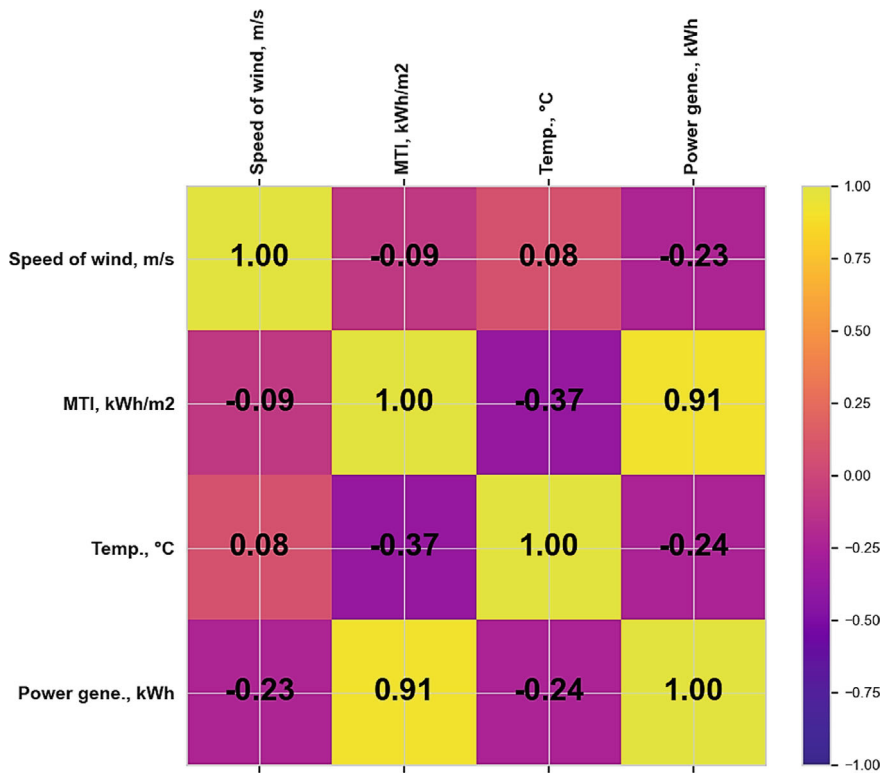


Fig. 2 Heatmap of correlation values

skewness of 0.1, indicating a few higher wind speed estimates. The MTI has a mean of 157.091 kWh/m² and a standard deviation of 28.193, showing substantial variability. The data varies from 103.26 to 208.25 kWh/m², with a minor negative skewness of -0.38, indicating lower irradiation levels. The average temperature is 32.374 °C, with a standard deviation of 2.112, indicating reasonably steady temperatures. The temperature ranges from 29.01 °C to 38.2 °C and is positively skewed at 0.748, suggesting a few higher temperature readings. The electricity production averages 194,029.9 kWh with a standard deviation of 35,134.74, showing high fluctuation. It varies from 121,121.6 kWh to 251,465.7 kWh, with a negative skewness of -0.44, indicating a few higher values. The kurtosis values for all variables are near zero, suggesting that the distribution is comparable to the normal distribution. These statistics provide a complete picture of the data, emphasizing the central tendency, dispersion, and form of the distributions for each variable.

Table 2 Descriptive statistics of the used data

	Speed of wind, m/s	MTI, kWh/m ²	Temp., °C	Power gene., kWh
Count	36	36	36	36
Skewness	0.1	−0.38	0.748	−0.44
Mean	2.024	157.091	32.374	194,029.9
Std	0.598	28.193	2.112	35,134.74
25%	1.428	137.24	30.867	176,392
50%	2	161.725	32.17	199,043.7
75%	2.397	179.145	33.5	221,250.3
Kurtosis	−0.908	−0.852	0.347	−0.582

3.2 Model Development and Evaluation

Using linear regression (LR), AdaBoost, XGBoost, and random forest (RF) the assessment of the power-generating model indicates different performance characteristics for every technique. The actual data is compared with LR model-based prediction values in Fig. 3a. It can be observed that most of the data points lay well within the ($\pm 15\%$) range, indicating a fair performance. With model training R^2 and testing R^2 values of 0.8764 and 0.8484 respectively, linear regression shows a model training MSE of $1.32E + 08$ and a testing MSE of $2.54E + 08$. Along with Train MAPE of 4.375 and model testing MAPE of 7.333, indicating a modest prediction error, this shows a strong fit during training but somewhat inferior accuracy on the test data.

The model prediction results are compared in Fig. 3b in the case of the AdaBoost-based power generation model. Reflecting close-perfect accuracy, AdaBoost performs much better during training with a training MSE of 12,181,315 and a high training R^2 of 0.9886, as illustrated in Table 3. Its model testing MSE is $2.49E + 08$, however, and model testing R^2 falls to 0.8513. The training MAPE is 1.170; the model testing MAPE rises to 5.966, indicating better performance in terms of percentage error during testing than LR. The model developed with a decision tree (DT) provided improved prediction. The comparison is depicted in Fig. 3c, and it can be observed that comparative points were quite close to the best-fit line. On training data with a training MSE of 0 and training R^2 of 1 the Decision Tree model achieves excellent accuracy. Its model testing MSE of $3.63E + 08$ and lower model testing R^2 of 0.7835. While the model testing MAPE is 7.146 and indicates a significant rise in prediction error during testing, the training MAPE is 0. The model performance in the case of the XGBoost-based power generation model is depicted in Fig. 3d. With an almost zero training MSE of 0.000683 and an excellent training R^2 of 1 XGBoost exhibits extraordinary training accuracy. Notwithstanding this, on the test data, its model testing MSE is $2.66E + 08$ and testing R^2 is 0.8412, indicating considerable overfitting but still performing similarly to the other models. At 0.00001 the training

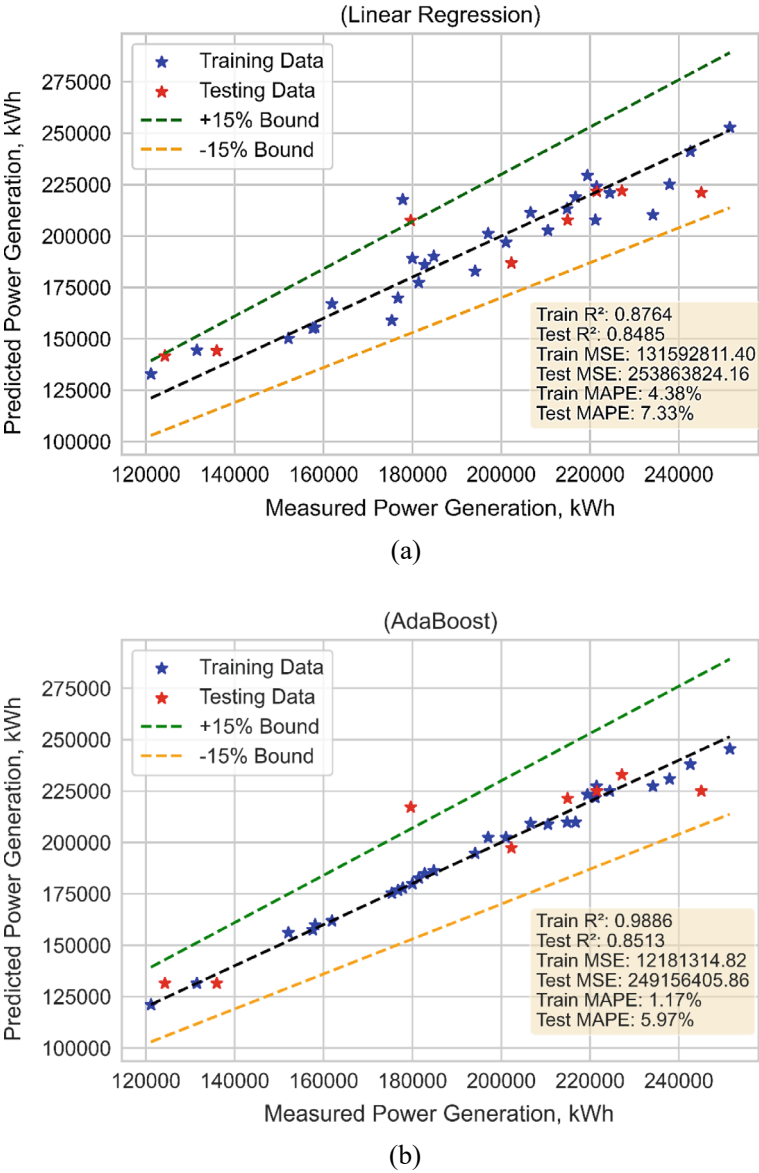
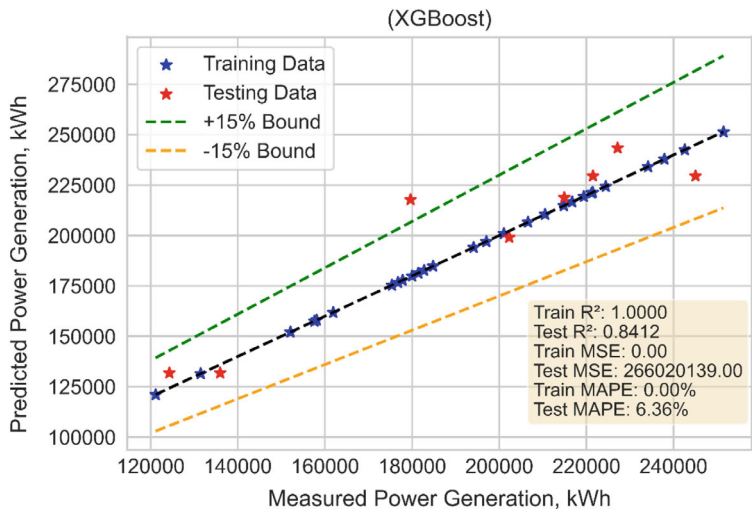


Fig. 3 Power generation model’s actual versus model prediction for: **a** LR, **b** AdaBoost, **c** DT, **d** XGBoost models



(c)



(d)

Fig. 3 (continued)

Table 3 Model evaluation

Model	Train MSE	Test MSE	Train R^2	Test R^2	Train MAPE	Test MAPE
LR	1.32E + 08	2.54E + 08	0.8764	0.8485	4.375	7.333
AdaBoost	12,181,315	2.49E + 08	0.9886	0.8513	1.170	5.966
DT	0	3.63E + 08	1	0.7835	0	7.146
XGBoost	0.000683	2.66E + 08	1	0.8412	0.00001	6.36

MAPE is very low; the model testing MAPE is 6.36, suggesting quite low prediction error during testing.

All things considered, AdaBoost and XGBoost exhibit indications of overfitting with great training results but lowered test accuracy whereas LR offers a fair performance between training and testing. Though excellent training accuracy, the decision tree model suffers from notable overfitting. Though they are overfit, XGBoost and AdaBoost are better than LR in training accuracy and have reduced MAPE during testing.

3.3 Model Comparison with Taylor’s Diagram and Violin Plots

For model comparison in ML, Taylor’s diagram and violin graphs are useful approaches. In a single figure, Taylor’s diagram succinctly summarizes model performance by graphically displaying the standard deviation, correlation, and cantered root-mean-square difference between models and data. The models in the present study are compared using Taylor’s diagrams as depicted in Fig. 4a and c, for model training and testing phase respectively. It can be seen that the marker of XGBoost in training while the marker of AdaBoost in testing phases is closer to observation, indicating it is the best-performing prediction model. Conversely, violin plots depict the distribution of a dataset by aggregating kernel density plots and box graphs to indicate the probability density. Violin plots show variances in accuracy and variability by contrasting the spread and skewness of model predictions versus actual data. These instruments are used together to provide a complete assessment of model performance, highlighting areas of strength and weakness in predicted accuracy. The violin plots depicted in Fig. 4b and d depict the model comparison during model training as well as the testing phase. These plots corroborate the outcome of Taylor’s diagram as XGBoost during training and AdaBoost during testing is best best-performing prediction model.

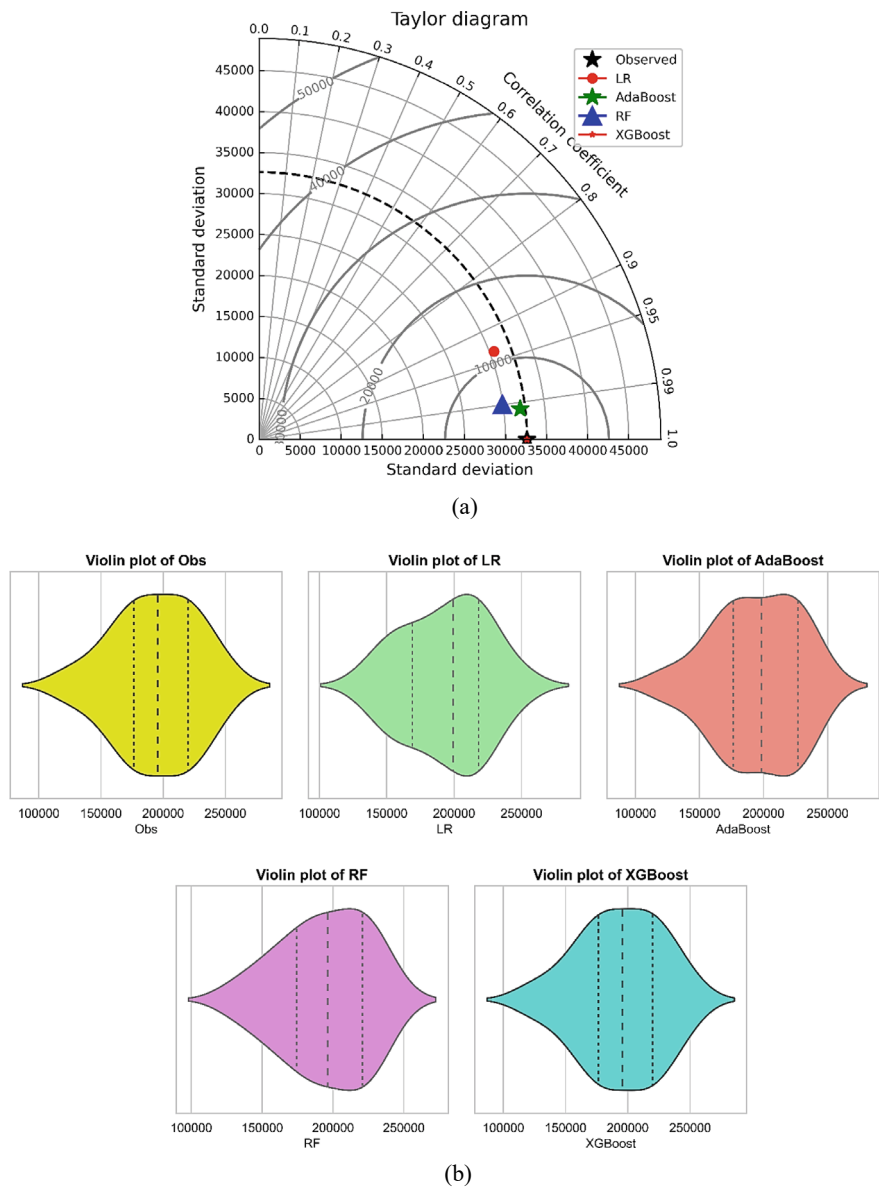


Fig. 4 Model comparison with Taylor’s diagram and violin plots

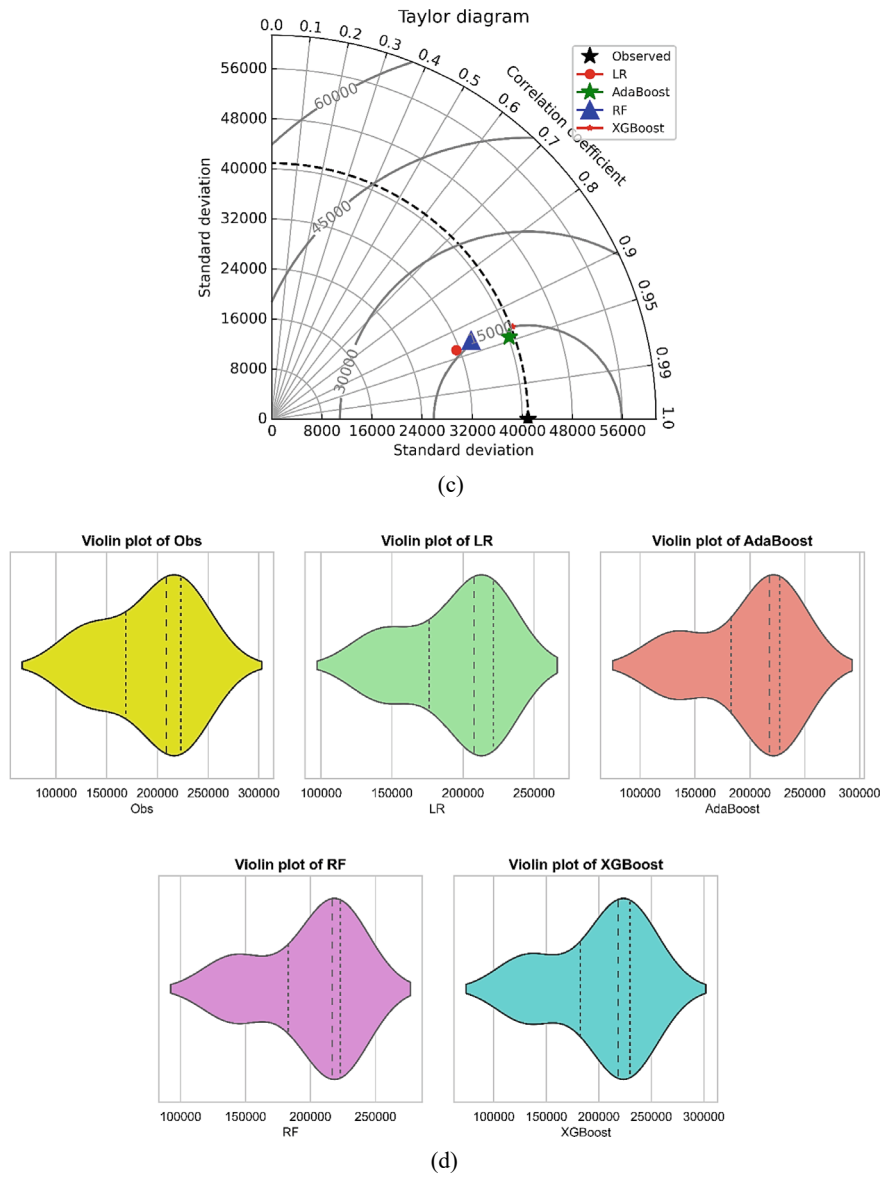


Fig. 4 (continued)

4 Conclusion

The study compares four different machine learning approaches for model prediction of solar power generation. Linear regression (LR), decision tree (DT), AdaBoost, and XGBoost were employed using the real operational data of three consecutive years. The evaluation metrics, including mean squared error (MSE), R -squared (R^2), and mean absolute percentage error (MAPE), were employed to reveal the strengths and limitations of each model. LR shows balanced performance with a training MSE of $1.32\text{E} + 08$ and a test MSE of $2.54\text{E} + 08$, with moderate prediction errors but consistent performance across both phases. AdaBoost exhibits exceptional training accuracy with a training MSE of $1.218\text{E} + 07$ and a training R^2 of 0.98854, indicating its robustness in capturing underlying patterns in the training data. However, its test MSE is $2.49\text{E} + 08$, and test R^2 drops to 0.8513. The DT model achieves perfect accuracy during training but suffers from substantial overfitting, revealing significant prediction errors during testing. XGBoost stands out with a near-zero MSE of 0.000683 and a perfect R^2 of 1 during model training, but its test MSE is $2.66\text{E} + 08$ and test R^2 is 0.8412, showing some overfitting but still performing comparably to other models on the test data. XGBoost and AdaBoost offer superior training accuracy and lower prediction errors, with XGBoost slightly edging out due to its lower test MAPE of 6.36 compared to AdaBoost's 5.966.

References

1. Nguyen VN, Tarelko W, Sharma P, El-Shafay AS, Chen W-H, Nguyen PQP et al (2024) Potential of explainable artificial intelligence in advancing renewable energy: challenges and prospects. *Energy Fuels* 38:1692–1712. <https://doi.org/10.1021/acs.energyfuels.3c04343>
2. Kian AY, Lim SC (2023) On the potential of solar energy for chemical and metal manufacturing plants in Malaysia. *Int J Adv Sci Eng Inf Technol*, 13:1898–1904. <https://doi.org/10.18517/ijaseit.13.5.19052>
3. Ntanos S, Kyriakopoulos GL, Anagnostopoulos T, Xanthopoulos T, Kytagiass C, Drosos D (2022) Investigating the environmental and the energy saving behavior among school principals through classification algorithms. *Int J Renew Energy Dev*, 11(2), <https://doi.org/10.14710/Ijred202243007>
4. Nguyen MK, Lin C, Hoang HG, Sanderson P, Dang BT, Bui XT et al (2022) Evaluate the role of biochar during the organic waste composting process: a critical review. *Chemosphere* 299:134488. <https://doi.org/10.1016/j.chemosphere.2022.134488>
5. Nguyen XP, Le ND, Pham VV, Huynh TT, Dong VH, Hoang AT (2021) Mission, challenges, and prospects of renewable energy development in Vietnam. *Energy Source Part A Recover Util Environ Eff*, 1–13. <https://doi.org/10.1080/15567036.2021.1965264>
6. Hasnaoui A, Omari A, Azzouz Z, eddine, Danoune MB, Arini NR (2023) Reduction of electricity cost of residential home using PSO and WOA optimization method. *Int J Adv Sci Eng Inf Technol*, 13:828–834. <https://doi.org/10.18517/ijaseit.13.3.18374>
7. Ruiz SM, Chamorro MV, Ocho GV, Villegas JF, Peñaloza CA (2021) Effects of environmental conditions on photovoltaic generation system performance with polycrystalline panels. *Int J Adv Sci Eng Inf Technol*, 11:2031–2038. <https://doi.org/10.18517/ijaseit.11.5.9335>
8. Han T, Paramasivam P, Dong VH, Cuong H, Chuan D (2024) Harnessing a better future : exploring AI and ML applications in renewable energy. *JOIV Int J Informatics Vis*, 8

9. Mignogna D, Ceci P, Cafaro C, Corazzi G, Avino P (2023) Production of biogas and biomethane as renewable energy sources: a review. *Appl Sci* 13:10219. <https://doi.org/10.3390/app131810219>
10. Capuano L (2020) International energy outlook 2020. Washington, DC
11. Chakraborty D, Mondal J, Barua HB, Bhattacharjee A (2023) Computational solar energy—ensemble learning methods for prediction of solar power generation based on meteorological parameters in Eastern India. *Renew Energy Focus* 44:277–294. <https://doi.org/10.1016/j.ref.2023.01.006>
12. IBEF (2024) Renewable energy industry Report. New Delhi
13. Said Z, Rahman S, Sharma P, Hachicha AA, Issa S (2022) Performance characterization of a solar-powered shell and tube heat exchanger utilizing MWCNTs/Water-based nanofluids: An experimental, Numerical, and Artificial Intelligence approach. *Appl Therm Eng* 118633. <https://doi.org/10.1016/j.applthermaleng.2022.118633>
14. Liu J, Song R, Nasreen S, Hoang AT (2019) Analysis of the complementary property of solar energy and thermal power based on coupling model. *Nat Environ Pollut Technol* 18:1675–1681
15. Ahmed SF, Islam N, Kumar PS, Hoang AT, Mofijur M, Inayat A et al (2023) Perovskite solar cells: thermal and chemical stability improvement, and economic analysis. *Mater Today Chem* 27:101284. <https://doi.org/10.1016/j.mtchem.2022.101284>
16. Gopi A, Sharma P, Sudhakar K, Ngui WK, Kirpichnikova I, Cuce E (2022) Weather impact on solar farm performance: a comparative analysis of machine learning techniques. *Sustainability* 15:439. <https://doi.org/10.3390/su15010439>
17. Ghodbane M, Said Z, Ketfi O, Boumeddane B, Hoang AT, Sheikholeslami M et al (2022) Thermal performance assessment of an ejector air-conditioning system with parabolic trough collector using R718 as a refrigerant: a case study in Algerian desert region. *Sustain Energy Technol Assessments* 53:102513. <https://doi.org/10.1016/j.seta.2022.102513>
18. Afandi A, Birowosuto MD, Sembiring KC (2022) Energy-yield assessment based on the orientations and the inclinations of the solar photovoltaic rooftop mounted in Jakarta, Indonesia. *Int J Adv Sci Eng Inf Technol*, 12:470–476. <https://doi.org/10.18517/ijaseit.12.2.14812>
19. Le TH, Pham MT, Hadiyanto H, Pham VV, Hoang AT (2021) Influence of various basin types on performance of passive solar still: a review. *Int J Renew Energy Dev*, 10:789–802. <https://doi.org/10.14710/ijred.2021.38394>
20. Angula JP, Inambao FL (2020) Optimization of solar dryers through thermal energy storage: two concepts. *Int J Eng Res Technol*, 13:2803–2813. <https://doi.org/10.37624/IJERT/13.10.2020.2803-2813>
21. Hananda N, Kamul A, Harito C, Djuana E, Elwirehardja GN, Pardamean B et al (2023) Solar drying in Indonesia and its development: a review and implementation. *IOP Conf Ser Earth Environ Sci* 1169:012084. <https://doi.org/10.1088/1755-1315/1169/1/012084>
22. Hoang AT, Le TH, Chitsomboon T, Koonsrisook A (2021) Experimental investigation of solar energy-based water distillation using inclined metal tubes as collector and condenser. *Energy Sources, Part A Recover Util Environ Eff*, 1–17. <https://doi.org/10.1080/15567036.2021.1966139>
23. Nguyen HB, Nguyen VL (2023) A study on the efficiency of solar radiation collectors applying for agricultural products and food drying. *Int J Adv Sci Eng Inf Technol*, 13:564–571. <https://doi.org/10.18517/ijaseit.13.2.18712>
24. Nguyen YQ, Nguyen-Tan S, T. Pham H-T, Manh-Thuy A, Huynh-Nhat T (2021) Performance of a solar chimney for cooling building façades under different heat source distributions in the air channel. *Int J Adv Sci Eng Inf Technol*, 11:158–164. <https://doi.org/10.18517/ijaseit.11.1.12666>
25. Kurniawan A, Shintaku E (2022) Estimation of hourly solar radiations on horizontal surface from daily average solar radiations using artificial neural network. *Int J Adv Sci Eng Inf Technol*, 12:2336–2341. <https://doi.org/10.18517/ijaseit.12.6.12940>
26. Al Garni HZ, Awasthi A (2018) Solar PV power plants site selection. *Adv Renew Energies Power Technol*, 57–75. <https://doi.org/10.1016/B978-0-12-812959-3.00002-2>

27. Siecker J, Kusakana K, Numbi BP (2017) A review of solar photovoltaic systems cooling technologies. *Renew Sustain Energy Rev* 79:192–203. <https://doi.org/10.1016/j.rser.2017.05.053>
28. Kanti PK, Sharma P, Maiya MP, Sharma KV (2023) The stability and thermophysical properties of Al₂O₃-graphene oxide hybrid nanofluids for solar energy applications: application of robust autoregressive modern machine learning technique. *Sol Energy Mater Sol Cells* 253:112207. <https://doi.org/10.1016/j.solmat.2023.112207>
29. Ronge H, Niture V, Ghodake DS (2016) A review paper on utilization of solarenergy for cooking. *Imp Int J Eco-Friendly Technol* 1:121–124
30. Barhmi K, Heynen C, Golroodbari S, van Sark W (2024) A review of solar forecasting techniques and the role of artificial intelligence. *Solar* 4:99–135. <https://doi.org/10.3390/solar4010005>
31. Carmona M, Palacio Bastos A, García JD (2021) Experimental evaluation of a hybrid photovoltaic and thermal solar energy collector with integrated phase change material (PVT-PCM) in comparison with a traditional photovoltaic (PV) module. *Renew Energy* 172:680–696. <https://doi.org/10.1016/j.renene.2021.03.022>
32. Abdullah AL, Misha S, Tamaldin N, Rosli MAM, Sachit FA (2020) Theoretical study and indoor experimental validation of performance of the new photovoltaic thermal solar collector (PVT) based water system. *Case Stud Therm Eng* 18:100595. <https://doi.org/10.1016/j.csite.2020.100595>
33. Fu Z, Li Y, Liang X, Lou S, Qiu Z, Cheng Z et al (2021) Experimental investigation on the enhanced performance of a solar PVT system using micro-encapsulated PCMs. *Energy* 228:120509. <https://doi.org/10.1016/j.energy.2021.120509>
34. Le AT, Pandey A, Sirohi R, Sharma P, Chen W-H, Pham NDK et al (2023) Precise prediction of biochar yield and proximate analysis by modern machine learning and shapley additive explanations. *Energy Fuels* 37:17310–17327. <https://doi.org/10.1021/acs.energyfuels.3c02868>
35. Hasnaoui A, Omari A, Danoune MB, Arini NR (2023) Reduction of electricity cost of residential home using PSO and WOA optimization method. *Int J Adv Sci Eng Inf Technol*, 13:863–869. <https://doi.org/10.18517/ijaseit.13.3.18374>
36. Barus DH, Dalimi R (2021) Multi-stage statistical approach to wind power forecast errors evaluation: a southern sulawesi case study. *Int J Adv Sci Eng Inf Technol*, 11:633–641. <https://doi.org/10.18517/ijaseit.11.2.12385>
37. Sharma P, Said Z, Kumar A, Nižetić S, Pandey A, Hoang AT et al (2022) Recent advances in machine learning research for nanofluid-based heat transfer in renewable energy system. *Energy Fuels* 36:6626–6658. <https://doi.org/10.1021/acs.energyfuels.2c01006>
38. Jain P, Rajput RS, Kumar S, Sharma A, Jain A, Bora BJ et al (2024) Recent advances in graphene-enabled materials for photovoltaic applications: a comprehensive review. *ACS Omega*. <https://doi.org/10.1021/acsomega.3c07994>
39. Almarzooqi AM, Maalouf M, El-Fouly THM, Katzourakis VE, El Moursi MS, Chrysikopoulos CV (2024) A hybrid machine-learning model for solar irradiance forecasting. *Clean Energy* 8:100–110. <https://doi.org/10.1093/ce/zkad075>
40. Le TT, Le HC, Paramasivam P, Chung N (2024) Artificial intelligence applications in solar energy. *JOIV Int J Informatics Vis*, 8
41. Nguyen TH, Paramasivam P, Le HC, Nguyen DC (2024) Harnessing a better future: exploring AI and ML applications in renewable energy. *JOIV Int J Informatics Vis* 8:55–78
42. Liao X, Zhu R, Wong MS, Heo J, Chan PW, Kwok CYT (2023) Fast and accurate estimation of solar irradiation on building rooftops in Hong Kong: a machine learning-based parameterization approach. *Renew Energy* 216:119034. <https://doi.org/10.1016/j.renene.2023.119034>
43. Tercha W, Tadjer SA, Chekired F, Canale L (2024) Machine Learning-based forecasting of temperature and solar irradiance for photovoltaic systems. *Energies* 17:1124. <https://doi.org/10.3390/en17051124>
44. Vennila C, Titus A, Sudha TS, Sreenivasulu U, Reddy NPR, Jamal K et al (2022) Forecasting solar energy production using Machine Learning. *Int J Photoenergy* 2022:1–7. <https://doi.org/10.1155/2022/7797488>

45. Bhatti S, Manzoor HU, Michel B, Bonilla RS, Abrams R, Zoha A, et al (2023) Revolutionizing low-cost solar cells with machine learning: a systematic review of optimization techniques. *Adv Energy Sustain Res*, 4. <https://doi.org/10.1002/aesr.202300004>
46. Demir V, Citakoglu H (2023) Forecasting of solar radiation using different machine learning approaches. *Neural Comput Appl* 35:887–906. <https://doi.org/10.1007/s00521-022-07841-x>
47. Obando ED, Carvajal SX, Pineda AJ (2019) Solar radiation prediction using Machine Learning techniques: a review. *IEEE Lat Am Trans* 17:684–697. <https://doi.org/10.1109/TLA.2019.8891934>
48. Narvaez G, Giraldo LF, Bressan M, Pantoja A (2021) Machine learning for site-adaptation and solar radiation forecasting. *Renew Energy* 167:333–342. <https://doi.org/10.1016/j.renene.2020.11.089>
49. Balal A, Jafarabadi YP, Demir A, Igene M, Giesselmann M, Bayne S (2023) Forecasting solar power generation utilizing machine learning models in lubbock. *Emerg Sci J*, 7:1052–1062. <https://doi.org/10.28991/ESJ-2023-07-04-02>
50. Quirk TJ (2018) Correlation and simple linear regression. *Excel 2016 Appl. Stat High Sch Students*, Cham: Springer International Publishing, p. 107–152. https://doi.org/10.1007/978-3-319-89993-0_6
51. Shanmugasundar G, Vanitha M, Ćep R, Kumar V, Kalita K, Ramachandran M (2021) A comparative study of linear, random forest and Adaboost regressions for modeling non-traditional machining. *Processes* 9:2015. <https://doi.org/10.3390/pr9112015>
52. Shahraki A, Abbasi M, Haugen Ø (2020) Boosting algorithms for network intrusion detection: a comparative evaluation of Real AdaBoost, Gentle AdaBoost and Modest AdaBoost. *Eng Appl Artif Intell* 94:103770. <https://doi.org/10.1016/j.engappai.2020.103770>
53. Li R, Sun H, Wei X, Ta W, Wang H (2022) Lithium battery state-of-charge estimation based on AdaBoost. *Rt-RNN. Energies*, 15, 6056. <https://doi.org/10.3390/en15166056>
54. Kyeremeh F, Zhi F, Yi Y, Gyamfi E, Nti IK (2022) Solar PV power forecasting with a hybrid LSTM-AdaBoost ensemble. 2022 IEEE/IET Int Util Conf Expo, IEEE, p. 1–7. <https://doi.org/10.1109/IUCE55902.2022.10079424>
55. Liu Y, Duan S, He X, Wang H (2023) Short-term PV power prediction based on the 24 traditional Chinese solar terms and adaboost-GA-BP model. *Front Energy Res*, 11. <https://doi.org/10.3389/fenrg.2023.1229695>
56. Nanfack G, Temple P, Frénay B (2022) Constraint enforcement on decision trees: a survey. *ACM Comput Surv* 54:1–36. <https://doi.org/10.1145/3506734>
57. Jumin E, Basaruddin FB, Yusoff YBM, Latif SD, Ahmed AN (2021) Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. *Environ Sci Pollut Res* 28:26571–26583. <https://doi.org/10.1007/s11356-021-12435-6>
58. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1:81–106. <https://doi.org/10.1007/BF00116251>
59. Custode LL, Iacca G (2023) Evolutionary Learning of interpretable decision trees. *IEEE Access* 11:6169–6184. <https://doi.org/10.1109/ACCESS.2023.3236260>
60. Kumar KP, Alruqi M, Hanafi HA, Sharma P, Wanatasanappan VV (2024) Effect of particle size on second law of thermodynamics analysis of Al₂O₃ nanofluid: application of XGBoost and gradient boosting regression for prognostic analysis. *Int J Therm Sci* 197:108825. <https://doi.org/10.1016/j.ijthermalsci.2023.108825>
61. Said Z, Sharma P, Bora BJ, Pandey AK (2023) Sonication impact on thermal conductivity of f-MWCNT nanofluids using XGBoost and Gaussian process regression. *J Taiwan Inst Chem Eng* 145:104818. <https://doi.org/10.1016/j.jtice.2023.104818>
62. Akbar B, Tayara H, Chong KT (2024) Unveiling dominant recombination loss in perovskite solar cells with a XGBoost-based machine learning approach. *IScience* 27:109200. <https://doi.org/10.1016/j.isci.2024.109200>
63. Qiu Y, Zhou J, Khandelwal M, Yang H, Yang P, Li C (2022) Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Eng Comput* 38:4145–4412. <https://doi.org/10.1007/s00366-021-01393-9>

Feature Selection Using Hybrid Opposition-Based Whale Optimization Algorithm



Parul, Charu Gupta, and Devendra Kumar Tayal

Abstract Feature selection is a very important step in machine learning and data mining, as it aims at identifying and preserving only relevant features that contribute a lot of significance to the task of prediction or classification, irrelevant and redundant aspects are eliminated at the same time. A novel approach, Opposition-Based Whale Optimization (WOA-OB) algorithm applied to feature selection is proposed to improve the performance for the classification task as well as its reliability. The Whale Optimization Algorithm, which is the latest metaheuristic optimization method under the larger nature-inspired algorithms category generates new populations at random during the exploration and exploitation phases, as other population-based algorithms do. As a result, it may be inclined to deviate from the ideal solution or be stuck in a local optimum. WOA-OB is an extension of the Whale Optimization Algorithm which makes use of the opposition-based method to improve the efficiency of the algorithm. WOA-OB looks into solutions that are opposite to original values to see whether going in the other direction yields a better outcome. After conducting a thorough comparison, it was concluded that the WOA-OB is more efficient in the sense that it is able to give more accurate results. The approach was evaluated on 4 different datasets. The cervical cancer dataset gave the best accuracy of 100% while the glass dataset had the highest accuracy of 98.46% which shows the efficacy of the proposed method.

Keywords Feature selection • Optimisation • Metaheuristic • Opposition learning • Machine learning • Classification

Parul (✉) • D. K. Tayal

Department of Computer Science and Engineering, Indira Gandhi Delhi Technical University for Women, Kashmere Gate, New Delhi, India

e-mail: parulrana2214@gmail.com

D. K. Tayal

e-mail: dev_tayal2001@yahoo.com

C. Gupta

Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, Delhi, India

e-mail: charu.wa1987@gmail.com

1 Introduction

Machine learning (ML) is being progressively used to tackle the most complicated tasks in the areas of image recognition and medical data analyses. On the other hand, the complexity of high-dimensional data makes it difficult to apply in practice. This might cause several problems, including overfitting, which happens when the learning model is too good at the training set, and cannot differentiate between noise and information. High-dimensional datasets result in a decrease in interpretability, where it is hard to comprehend the model's process and decision-making procedures as well as increased computational power needed to train the model [1]. In addition to the numerous challenges, feature selection (FS) is the key element of solving these problems. It keeps the most meaningful attributes and removes data that is irrelevant. By doing so, it finds the subset of best features from a bigger collection of characteristics that describe the data with less redundancy. This process allows us to enhance the efficacy and efficiency of ML algorithms by streamlining the dimension of data, where irrelevant features are removed, data complexity is reduced, and the speed of training and model interpretability are improved. The FS prevents the overfitting and contributes to better model performance as well as higher prediction accuracy by improving ML models [2]. Three primary categories comprise FS techniques, which are as follow:

Filter approaches that employ these statistical measures include assessing the data based on their own inherent qualities [3]. They are often used as a generalization method before the full model is built since they are readily computable. Filters as strategies are the ones that do not use classifier features. Validation criterion of wrapper methods is the inclusion of features into a chosen ML algorithm [4]. The main task, hence, is to investigate how the model performance varies when different feature sets are used. To achieve this, do this by getting the model's performance after training it on a different feature subset and evaluate it. This part uses the best subset that is achieved. This, therefore, results in a more accurate ML model and also that the features that were carefully selected are to the data. In embedded techniques, FS is an important stage in the training and modeling process [5]. Embedding approaches offer the advantage of detecting correlations that might be missed if a wrapper method was employed. But the elements chosen could be the wrong choice for a different algorithm, as they are dependent on the one to which they are assigned (Fig. 1).

In ML, the challenge of dealing with datasets that contain many dimensions is one of the things that you need to be careful with. This is where the application of optimization techniques [6] is useful. These computational mechanisms focus on finding the best option out of a big set of choices. They serve as the main component of ML, particularly in FS, where they comprise the set of features that give the highest performance to the classifier that in turn gives it a good prediction power [7]. These algorithms use an approach whereby they will test different feature subsets and refine their search by analyzing the results.

Optimization methods can be broadly classified into deterministic and meta-heuristic strategies. [8] Deterministic approaches like gradient descent are defined

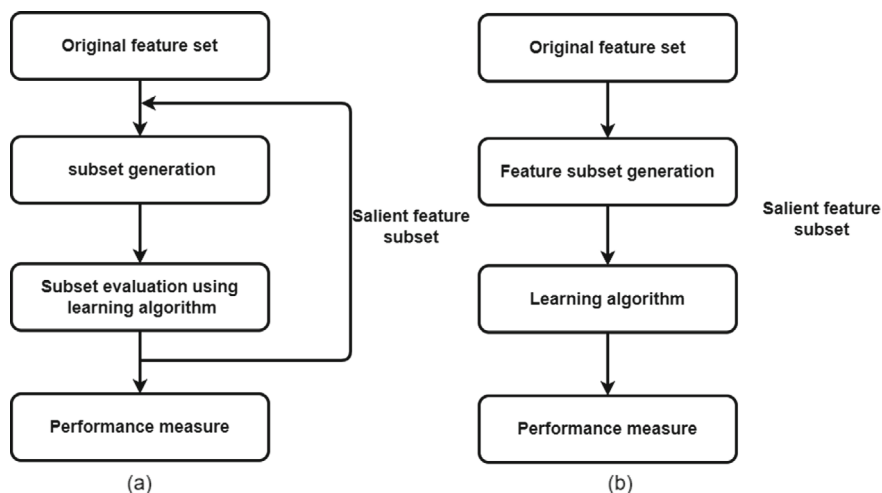


Fig. 1 Diagrams showing the **a** wrapper technique and **b** filter approach schematics

by strict mathematical rules for finding the best solution. They are good for the problems that can be precisely defined, but extensive problems can get them into a local optimum. This might be interpreted in a way that a solution found at the best point within a limited radius around the starting point may not be the global optimum. The global optimum is the best solution that can be achieved over the whole search area. In contrast, metaheuristic algorithms are sophisticated heuristic methods that demonstrate a systematic search of the search space to find a sensible solution to an optimization problem. Metaheuristic approaches, on the other hand, involve sophisticated heuristic strategies that demonstrate a smart search of the search space to find a sensible solution for an optimization problem. These methods are commonly used, when the complexity of the situation makes the application of the deterministic approaches unsuitable. The metaheuristics category includes algorithms such as the genetic algorithm, simulated annealing, and particle swarm optimization. Being trapped in local optima would be evaded by doing this, hence the chance of the global optimum will be increased. The most efficient solution type will be determined by the type of problem. Deterministic algorithms are suitable for situations with well-defined conditions while metaheuristic methods are applied for complex issues where the search is for the global best solution.

Metaheuristics optimization algorithms [9] are highly adaptable and can search for solutions an unlimited number of times. These algorithms are population-based and work with a set of candidate solutions and they will improve these solutions incrementally. Consequently, they can surmount the local optima and conduct a global search. Metaheuristics can be largely classified into four categories: Nature-inspired, art-inspired, trajectory, and ancient-inspired as on Fig. 2. The metaheuristics that follow trajectory start from a single solution and go through small changes in each iteration creating trajectory in the solution space. Nature-based metaheuristics

are natural phenomena and biological process-inspired, working with a population of solutions and interacting with operations based on natural genetics and evolution. Art-inspired metaheuristics—focus on different places of art like music, painting, and dance that have similarities to creative processes, then apply these processes to find the solutions. The metaheuristics based on the ancient times source their inspiration from ancient human practices and historical events, which are then translated to the search process by means of strategies and tactics employed in these practices. Nature-inspired algorithm [10] is a specific way of creating optimization algorithms based on the metaheuristics principles which are primarily based on the processes and actions of nature. The nature-inspired metaheuristic optimization algorithms, typically, can be categorized into the evolutionary algorithms, swarm-based algorithms, bio-inspired, human based.

Whale Optimization Algorithm (WOA) [11] is a specific set of algorithms which are inspired from nature. To be more specific, this art is a part of the bio-inspired category and it represents the biological way in which whales behave. Humpback whale feeding method is the inspiration for the WOA engine, as found by Mirjalili and Lewis

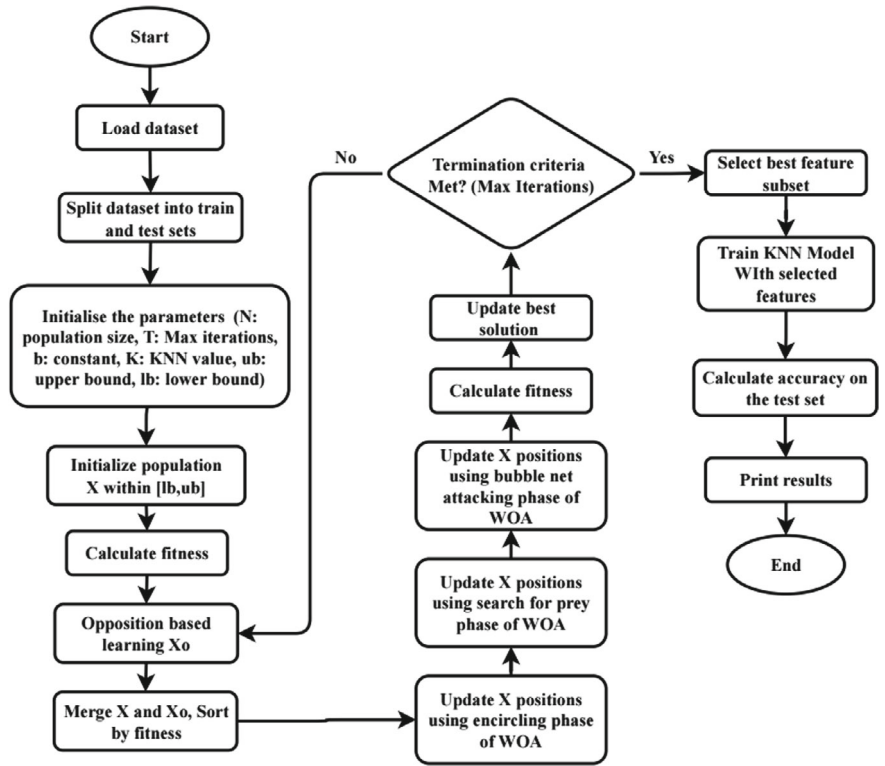


Fig. 2 Flowchart of proposed method

in 2016. On the other hand, for WOA it is the dynamic fish swarm that is being gradually fine-tuned that will be the search space. The algorithm takes advantage of a smart strategy that provides a balance between global and local search areas while giving a stability of exploration and exploitation. This technique is an adaptive tool that uses three operators mimicking the same bubble-net foraging strategy of humpback whales. This strategy includes surrounding prey and searching for prey to find either an optimal or almost optimal solution for difficult optimization problems. However, WOA faces challenges maintaining population diversity and avoiding premature convergence. To address these limitations opposition-based learning (OBL) [12] can be incorporated with WOA. OBL is based on the principle of opposite solutions and uses it to increase the range of an optimization algorithm between the exploitation and exploration of information [13]. OBL diverts from the solitary search of a unique solution by exploring the alternatives and the opposite of a particular solution that leads to a wider search space and, thus, higher possibility of reaching the global optimum. The significant findings from this work are: This paper proposes a hybrid approach that contains both WOA and OBL. The main objective of such a combination is the enhancement of the algorithm's exploitation and exploration capabilities. This hybridization helps to speed up the algorithm's convergence, which in turn leads to the optimization procedure. Besides, it tackles limitations in WOA, such as sensitivity to parameter settings and challenges posed by highly imbalanced datasets. The described WOA-OB algorithm is tested on 4 different. The comparison of the different classifiers and the optimization algorithms, WOA, PSO, GWO, FFO, and FPA is given, which is on the basis of accuracy, specificity, sensitivity. In addition to that, a review of the effectiveness of the runtime computation is also available.

The paper is structured with the following outline: Sect. 2 discusses the previous work in the related field to get insights about recent developments. In Sect. 3, the proposed approach employed to implement the algorithm is provided. This section is further subdivided into three segments, explaining the algorithms used and how the hybridization of the WOA and OBL technique is implemented. Section 4 provides the results of experiments, using various measures to compare them. Section 5 focuses on the conclusion of the results and the future work.

2 Literature Survey

In this section, a review of prior research focusing on wrapper metaheuristics methods and filter metaheuristics methods and hybrid metaheuristic techniques has been discussed.

2.1 *Wrapper Metaheuristics Methods*

Various PSO-based FS was previously done [14] which presented PSOCO1 and PSOCO2, two novel hybrid methods for FS utilizing a particle swarm optimization and crossover operator. Furthermore, a simplified PSO (BBPSO) [15] method was suggested to improve its efficiency in FS (FS) issues. The Binary Bat Algorithm (BBPSO) with an adaptive chaotic jump (ACJ) approach served as the foundation for the unique FS technique known as BBPSO-ACJ, which was presented in the study. The ACJ strategy was designed to facilitate significant changes in the trajectory of stagnated particles, enriching the search behavior of BBPSO and preventing particle entrapment in local attractors. A novel approach to FS was presented [16]. It required the modification of the binary-coded ant colony optimization algorithm (MBACO) and the application of genetic methods. Classification accuracy and the quantity of characteristics chosen were taken into account by their fitness function. Using ant colony optimization (ACO) and the pheromone update rule, the study successfully removed superfluous features. While the pheromone updating rule was optimized, the algorithm's performance might still have been sensitive to parameter settings, and achieving optimal parameter values could have been a challenging task [17].

Whale Optimization Algorithm (WOA) [18] was applied in FS, aiming to preserve fewer features while identifying the best feature subset with outstanding precision in categorization. A chaotic WOA (CWOA) [19] is provided to handle local optimization challenges. The introduction of chaos enhanced global exploration, but the impact on convergence speed and sensitivity to initial conditions was not carefully considered. To choose features for categorization, two binary versions of the ALO were introduced [20]. Two methods for building the Binary Ant Lion Optimizer (BALO) were presented. To save processing expenses, ABC was employed for FS [21]. The size of the feature subset and the classification error rate were taken into consideration when multi-objective ABC techniques were proposed [22]. As the proposed methodology was applied in two different ways: ABC represented as continuous and ABC represented as binary. The Support Vector Machine (SVM) kernel parameters are adjusted by extending ABC. [23] Prior to applying SVM-DT in that study, the FS and parameter optimization processes employed the ABC technique. We ran a study on eight datasets to show how effective the suggested strategy is. Designed for FS, the Butterfly Optimization Algorithm (BOA) using binary versions based on S- and V-shaped functions [24].

2.2 *Filter Metaheuristic Methods*

A FS technique for classification known as particle swarm optimization (PSO) was described in [25]. Enhancing FS's effectiveness and efficiency in classification tasks was the aim of the proposed multi-objective filter technique. Two multiple-objective FS techniques based on PSO were also presented [26]. The objective of the work

was to present a Pareto front of non-dominated options for effective FS. Two multi-objective filter-based FS frameworks were proposed [27]. A feature ranking and information theory-based differential evolution (DE) method for filter FS [28] which attempted to overcome the difficulties associated with FS and showed how well the suggested method worked in automatically developing a limited number of useful features for classification tasks. ACO-based FS was presented as a way to choose pertinent characteristics and eliminate unnecessary ones [29]. Weighted bee colony optimization (wBCO) is a unique FS approach that was proposed [30]. The study put forth a wBCO-based technique that effectively found the ideal feature subset. An enhanced Cuckoo FS Algorithm (CFSA) by introducing a new initialization strategy and fitness function was proposed [31].

2.3 Hybrid Metaheuristic Methods

Hybrid Particle Swarm Optimization-Local Search (HPSO-LS) as a new FS method which is a hybrid of local search and particle swarm optimization (PSO) was first described by researchers [32]. The approach we put forward outperformed five different FS (FS) methods on thirteen benchmark datasets of the classification type. One of the possible approaches that allowed face expression recognition by adding PSO to micro-GA was proposed [33]. Micro-GA was designed to alleviate early convergence problems by equipping both local and global searches with strength. Besides, a wrapping-based technique called HPSO-SSM [34] which represents the synergy between PSO and a spiral mechanism was introduced. This method surpassed wrappers and filters in terms of the variety of data types searched and location quality tests using a KNN classifier on 20 classification data sets. A binary hybrid algorithm by combining GOA with GA mutation operator was suggested by the authors [35]. The final GOA algorithm demonstrated a very high accuracy of 92% on the classification of 25 datasets. In the next work, [36] proposed a GOA-EPD which is an enhanced GOA with genetic operators based on evolution for a better fitness function of FS. The method gave the best convergence properties and was more efficient than other optimizers. One SSA's exploration capability was increased in using a local search technique and a chaotic map [37]. The dynamic SSA was proven to be effective in FS on 20 benchmark datasets. Besides, the Sine Cosine Algorithm (SCA) proposed by [38] was integrated with SSA that included the disruption operator for population diversification. Hybrid approach has displayed effectiveness in datasets of different feature sizes. The SCA-based hybridization with the PSO that converted to binary variants has been considered in [39]. It was observed that the clustering accuracy was higher when this approach was compared to other optimization techniques. SCA and ALO (ALO) [40] as a hybrid model (random factors were also considered) with the FS process responsible for more population variety was suggested. The literature survey indicates the efficiency of the metaheuristics' optimization methods for FS, but it is also important to discuss the shortcomings found in the already existing approaches. Some of the hybrid algorithms were found to be sensitive to parameter

settings, suffer from early convergence problems, and have difficulty in avoiding getting trapped in local optima. These limitations provide us with potential areas for improvement in FS algorithms based on metaheuristics.

3 Proposed Methodology

This section explores a novel hybrid approach, WOA-OB, for FS. This approach synergistically combines opposition-based learning (OBL) with the Whale Optimization Algorithm (WOA) to improve its overall efficiency.

3.1 *Whale Optimization Algorithm*

The Whale Optimization Algorithm (WOA), as proposed by Mirjalili and Lewis in 2016, is a nature-based optimization technique that takes inspiration from humpback whale hunting strategies. [11] This method incorporates three distinct phases designed to emulate the various aspects of humpback whale behavior during the search for prey. The purpose of the WOA optimization process is to find the ideal position, also known as the prey position, by periodically updating each whale's position. The whale's goal is rather arbitrary in the beginning, and it's not quite evident where the prey is located. As a result, an extensive search is carried out throughout the search region, progressively reducing the number of whales in order to identify the ultimate position of the prey.

3.1.1 Searching for Prey

Parameter A affects the update formula for the whale's location in this step of the WOA algorithm. The search agent randomly locates the best prey, based on the placements of each agent. The location of a search agent is modified throughout this phase by using rather than using the best search agent, a randomly selected search agent. To be more precise, the WOA uses the population's overall location data to randomly search for an individual whale's position update when $A > 1$ and the probability $p \leq 0.5$. The Eqs. (1–4) provides the position update formula.

$$D = C \cdot X_{\text{rand}} - X_t \quad (1)$$

$$X_{t+1} = X_{\text{rand}} - A \cdot D \quad (2)$$

$$A = C \cdot a - a \quad (3)$$

$$C = 2r \quad (4)$$

Here, X_{rand} represents the location of a whale individual chosen at random, X_t represents the position of an individual over t iterations, and X_{t+1} represents the updated position. Over the duration of the iterations, the shrinking coefficient a declines linearly from 2 to 0. The coefficient vectors are A and C .

3.1.2 Encircling Prey

During their hunting behavior, humpback whales encircle their prey. Subsequently, they identify the current best candidate solution and regard it as the optimal solution, bringing them closer to the optimal outcome. It updates the position of each whale toward the identified best whale found thus far. The update involves a shrinking coefficient (a) that decreases linearly over iterations, representing the gradual convergence toward the prey. When $A < 1$ and $p < 0.5$, indicates that the optimal location of the whale population is either in close proximity to or precisely at the target position, the position update formula for this stage is given by Eqs. (5 and 6):

$$X_{t+1} = X_{\text{best}(t)} - A \cdot |C \cdot X_{\text{best}(t)} - X_t| \quad (5)$$

$$D = C \cdot X_{\text{best}} - X_t \quad (6)$$

Here, $X_{\text{best}(t)}$ denotes best solution position at iteration t and X_t is the position vector of an individual whale during iteration t .

3.1.3 Attacking Using a Bubble-Net

This phase imitates the unique foraging strategy of humpback whales, wherein they create a spiral of bubbles to trap prey. It utilizes the position and search space of the best whale to refine the locations of other whales. The update is based on a spiral movement around the best whale, where the control parameter is the distance between the prey and the whale. The WOA meets the requirements for hunting and uses a bubble-net approach to find prey when $A < 1$ and $p \geq 0.5$. The formula for bubble-net attacking is given by Eq. (7):

$$X_{t+1} = D_{b-t} \cdot e^{bl \cdot \cos(2\pi l)} + X_{\text{best}} \quad (7)$$

Here, l is a random number in the interval $(-1, 1)$ and b is a constant, and D_{b-t} represents the distance controlled by a parameter b and the current iteration t . The flowchart illustrating the WOA is shown in Fig. 2.

3.2 Opposition-Based Learning (OBL)

Opposition-based learning (OBL) [12, 13] is a heuristic technique applied in various optimization algorithms to improve their performance. It uses the concept of “opposites” to more thoroughly explore the search space and aims to attain improved solutions. OBL in FS involves the concept of opposition to enhance the process of selecting most pertinent features from a given feature set. The concept revolves around not only considering the original features, but also their opposites or negations during the FS process.

3.2.1 Generation of Opposites

During each iteration of the WOA, a population of candidate solutions (whales) is created. These solutions are updated and improved iteratively based on rules of the WOA algorithm. Within each WOA iteration, the OBL is applied. For each solution generated by the WOA, an opposing or negating solution is created. For each whale position X , the opposite position ($X_{0[i,d]}$) is given by Eq. 8:

$$X_{0[i,d]} = lb_{[0,d]} + ub_{[0,d]} - X_{[i,d]} \quad (8)$$

3.2.2 Update

After the opposite solutions are generated then the fitness of these opposite solutions is computed. When fitness of the opposite solution surpasses that of the original solution, the opposite solution is then adopted as the new candidate solution, which can be expressed as conditional Eq. (9) given by:

$$\begin{aligned} &\text{if fitness}(F_{\text{opposite}}) > \text{fitness}(F) : \\ &\quad F = F_{\text{opposite}} \end{aligned} \quad (9)$$

3.3 Proposed WOA-OB

This section provides an in-depth elucidation of the novel methodology, named WOA-OB, which integrates the WOA with OBL for effective FS to enhance the optimization process. Figure 3 illustrates the system architecture of the suggested methodology elucidating each step-in detail.

Our proposed WOA-OB-based FS process starts with the dataset loading. The dataset is read from the CSV file using the pandas library and is loaded into a

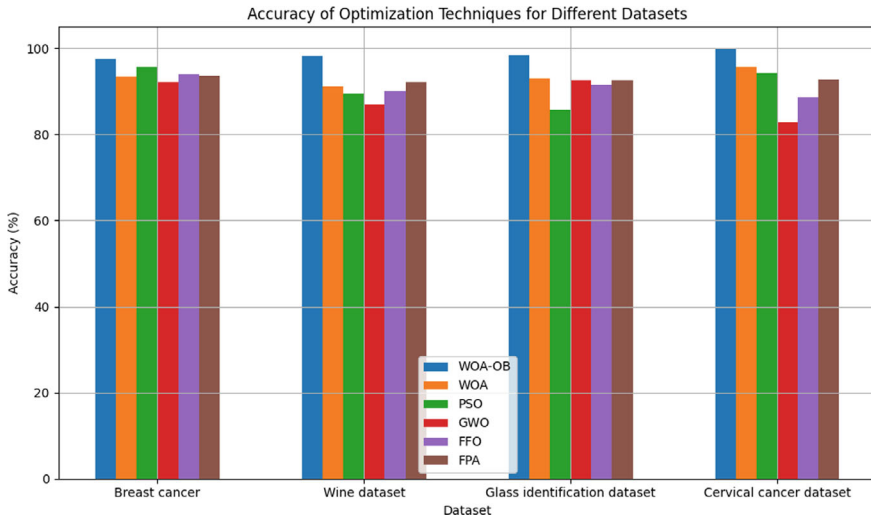


Fig. 3 Performance of different optimization algorithms

DataFrame. The features and the labels are then extracted from the DataFrame. Following that, the data is split into a training set and a validation set. Splitting is necessary for testing that the model will show good results on unforeseen data. Secondly, the parameters of KNN and WOA-OB are computed. These comprise the number of whales (N), maximum iteration number (T), the maximum number of neighbors (k) of the KNN, and the maximum iteration (maxLt) for the local search algorithm. The main steps for the FS using WOA-OB are: an initialization of the whale positions, a calculation of the fitness of the initial population, a performance of the OBL, and a position update of the whales using three different WOA behaviors. One position of a whale symbolizes a particular solution that will have a combination of features. The positions are initialized randomly in the range between the specified lower and upper bounds. The continuous locations are distributed into two binary values (0 or 1) by the threshold limit. This is for the reason that FS is a problem which is binary, in which the selection of a feature is denoted by 1 and its exclusion by 0. The fitness of each whale is obtained. The fitness function is a combination of error rate and number of features picked. The error rate is calculated and then a KNN model is built using the selected features. The model is tested on the test set and the error rate is calculated. The fitness function is defined as:

$$\text{cost} = \alpha \times \text{error} + \beta \times \left(\frac{\text{selected features}}{\text{total features}} \right) \quad (10)$$

where alpha and beta are constants, error is the error rate of the KNN model. The positions of the whales are updated based on three behaviors: encircling prey, searching for prey, and bubble-net attacking as depicted in Eqs. (1–7). The opposition-based learning approach is implemented by calculating the opposite positions for the

whales. The idea is to consider the opposite solutions. For each whale, an opposite position is calculated as:

$$X_{0[i,d]} = lb_{[0,d]} + ub_{[0,d]} - X_{[i,d]} \tag{11}$$

Based on fitness, the best N whales are chosen by combining the whales and their opposites. For the maximum number of iterations, these stages are repeated. Every iteration updates the whales’ positions, computes each whale’s fitness, and chooses the best whales. The best whale reflects the chosen features that the algorithm found. These features are then used to train the KNN model, and the validation set is used to assess the model’s performance.

4 Experiment Results

This section effectively demonstrates the performance of the proposed method by comparing it with different types of optimizations algorithms and other classifier models based on accuracy, fitness functions, and runtime computation.

4.1 Datasets Used

The detailed methodology was run on 4 datasets, such as breast cancer, cervical cancer, wine, and glass identification from the UCI repository. The details of all datasets used in the work are provided in Table 1, as regards their class, instances, and the features.

Table 2 presents selected feature subsets along with corresponding accuracy values for different datasets after the implementation of the WOA-OB algorithm for FS.

Table 3 outlines optimization algorithms along with their corresponding parameters and the specific values assigned to these parameters.

Table 1 Dataset description

Dataset	Instances	Features	Class
Breast cancer	569	30	2
Wine	178	13	3
Glass identification	214	9	6
Cervical cancer	72	19	2

Table 2 Average number of selected features

Dataset	Features	Average feature size	Accuracy	Specificity	Sensitivity
Breast cancer	30	12.1	97.66	0.97	0.98
Wine	13	5.30	98.15	1.00	0.98
Glass identification	9	3.23	98.46	1.00	0.99
Cervical cancer	19	4.13	100	1.0	1.0

Table 3 Parameter setting for the proposed approach

Parameter name	Description	Parameter value
k	K -value in KNN	5
N	Number of whales	10
T	Maximum number of iterations	100
b	Constant used in WOA-OB	1
thres	Threshold used for binary conversion	0.5
ub, lb	Upper bound and lower bound	1,0

4.2 Comparative Analysis WOA-OB

Table 4 provides a comparison of the WOA-OB's performance with other classifiers, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), and Decision Tree (DT), on four different datasets based on accuracy. After the results are evaluated, the KNN algorithm comes out as the best algorithm with the highest average accuracy across all datasets. The KNN algorithm achieves a high accuracy of 100% on the cervical cancer dataset, 98.15% on wine dataset, 98.46% on glass dataset and 97.91% on breast cancer dataset. SVM approach is the next best, somehow lower in performance than that of KNN. The NB algorithm stands as a third best performer, and the LR and DT models have the lower accuracy across the datasets.

Table 5 presents the WOA-OB performance comparison with WOA, Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO), Firefly Algorithm (FFO), and Flower Pollination Algorithm (FPA) by listing their results. Accuracy is used as a benchmark of judgment in terms of comparing performance. The conclusion

Table 4 Comparison of WOA-OB with other classifiers

Dataset	KNN	SVM	DT	NB
Breast cancer	97.66	80.1	70.0	70.37
Wine dataset	98.15	90.01	75.0	80.32
Glass identification dataset	98.46	93.8	70.9	87.07
Cervical cancer dataset	100	92.2	91.9	85.87

Table 5 Comparison of WOA-OB with other optimization algorithms

Dataset	WOA-OB	WOA	PSO	GWO	FFO	FPA
Breast cancer	97.66	93.43	95.80	92.211	94.01	93.61
Wine	98.15	91.03	89.47	86.966	90.00	92.14
Glass identification	98.46	92.96	85.65	95.519	92.60	91.51
Cervical cancer	100	95.8	94.15	82.857	88.57	92.85

Table 6 Accuracy achieved by proposed approach at different k value

Dataset	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$	$K = 9$	$K = 10$
Cervical cancer	96.17	96.17	96.65	96.17	100	99.23	100	97.65	96.17	96.17
Wine	97.78	95.56	93.33	95.56	98.56	96.33	95.56	93.33	94.83	93.33
Glass identification	91.23	92.15	95.15	96.30	98.30	94.44	96.30	94.44	96.30	96.30
Breast cancer	93.71	91.61	93.71	91.61	97.66	92.31	93.31	98.31	96.23	92.66

which can be drawn is that WOA-OB is more efficacious than the rest of optimization methods regarding accuracy on all the datasets. Besides, FFO and FPA also perform well, with a particularly good result on being applied to the wine and glass identification dataset.

Table 6 presents the accuracy obtained using the proposed approach at different iterations. It has been observed that there does not exist a specific K that will suit every dataset. The bigger K usually boosts the accuracy by reducing the noise for some datasets such as glass identification and breast cancer, but the others like cervical cancer and wine are opposite, after a certain K value the accuracy begins to fall due to over-smoothing. As a result, the value of K for each particular dataset and classification shall be determined through cross-validation and tuning techniques separately.

Figure 4 shows the bar graph of accuracy achieved at different iterations by different optimization algorithms such as WOA-OB, WOA, PSO, FPA, GWO. It has been observed that the WOA-OB consistently outperforms the other optimization algorithms at different iterations.

4.3 Runtime Computation

Table 7 provides a comparison of the runtime computation for different optimization algorithms across various datasets. Runtime computation is an essential metric because it reflects the computational efficiency of each algorithm. The optimization

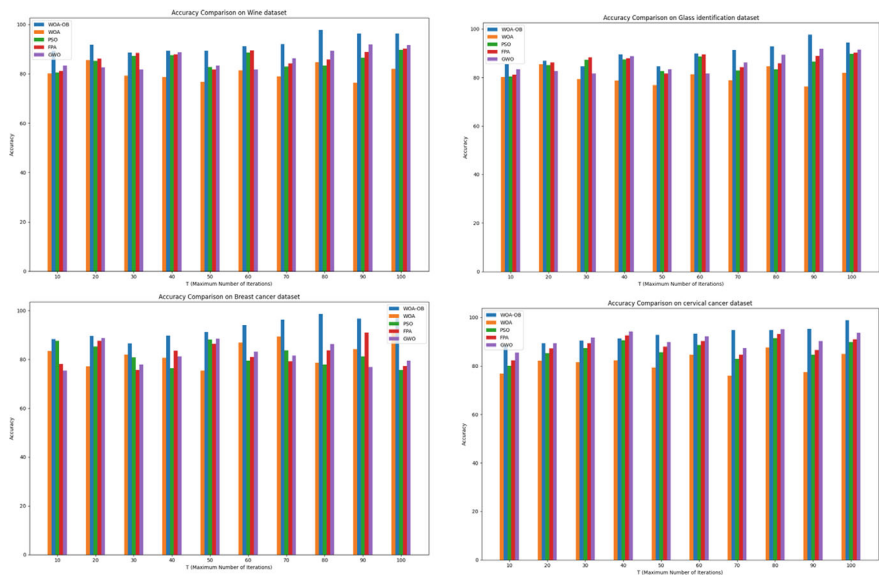


Fig. 4 Bar graph representing accuracy achieved at different iteration based on different approach

Table 7 Comparison of runtime computation (min:sec:milli_sec)

Dataset	WOA-OB	WOA	PSO	GWO	FFO	FPA
Breast cancer	00:30:32	00:28:1	00:25:9	00:27:8	00:32:16	00:21:05
Wine	00:37:43	00:31:9	00:22:3	00:35:3	00:34:92	00:23:32
Glass identification	00:42:45	00:35:8	00:28:1	00:37:8	00:39:95	00:27:11
Cervical cancer	00:29:34	00:24:4	00:22:2	00:25:5	00:26:30	00:27:76

algorithm WOA-OB does not have shorter time as compared to others because hybrid algorithms take more time to compute than individual algorithms, however, it focuses on the main goal which is minimizing feature size and classification error rate.

Figure 5 depicts the convergence curve of iterations versus fitness values. By looking at the convergence curve of WOA-OB it can be observed that it converges extremely quickly to a very low fitness value around 0.105 within the first 20 iterations and maintains a stable optimal solution for the remaining iterations. WOA-OB demonstrate the most desirable convergence curve as it is rapidly decreasing to the lowest stable fitness value compared to all other algorithms evaluated.

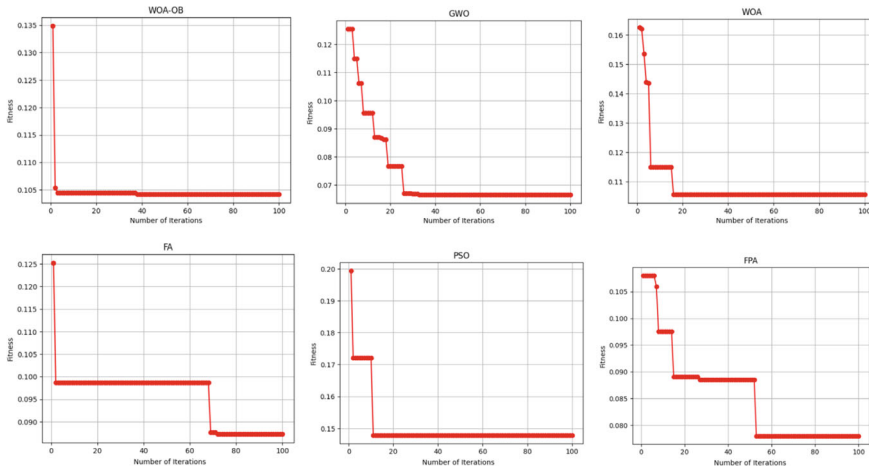


Fig. 5 Convergence curve ϕ versus max-iteration

5 Conclusion

The proposed technique, WOA-OB or the opposition-based WOA is an innovative and an integrated approach for optimizing and selecting features. The combining of FS by OBL with the WOA will in turn be used to enhance the search of optimal feature subsets by using the power of each technique. OBL implementation leads to unique challenges specific to optimization process via which it balances the opposing solutions and in doing so is promoting diversity, improving global search capabilities, accelerating convergence, and therefore lowering the risk of local optima. It can be concluded that WOA-OB is significantly better than SVM, KNN, FFO, WOA, PSO, GWO, FFO, and FPA in terms of accuracy over a variety of datasets. The statistical analysis, including fitness values and runtime computations, provides additional insights into the robustness and efficiency of WOA-OB. Ultimately, the proposed technique is shown to display efficient as well as innovative techniques of FS and optimization. The successful test results in comparison and conclusion suggest that see WOA-OB might be a very useful method for working in a broad range of the problems with the optimization purpose, especially in high dimensional areas. Future work can focus on fine-tuning of the parameters, expanding the scope to encompass a more diverse array of datasets, and also future research could investigate the potential of hybridizing WOA-OB with other optimization algorithms to further enhance the algorithm's performance.

References

1. Theng D, Bhojar KK (2023) FS techniques for ML: a survey of more than two decades of research. *Knowl Inf Syst* 1–63
2. Cai J et al (2018) FS in ML: a new perspective. *Neurocomputing* 300:70–79
3. Sánchez-Marfoño N, Alonso-Betanzos A, Tombilla-Sanromán M (2007) Filter methods for FS—a comparative study. In: *International conference on intelligent data engineering and automated learning*. Springer, Berlin Heidelberg
4. Chen G, Chen J (2015) A novel wrapper method for FS and its applications. *Neurocomputing* 159:219–226
5. Liu H, Zhou MengChu, Liu Q (2019) An embedded FS method for imbalanced data classification. *IEEE/CAA J Autom Sin* 6(3):703–715
6. Sun S et al (2019) A survey of optimization methods from a ML perspective. *IEEE Trans Cybern* 50(8):3668–3681
7. Pudjihartono N et al (2022) A review of FS methods for ML-based disease risk prediction. *Front Bioinform* 2:927312
8. Azevedo BF, Rocha AMAC, Pereira AI (2024) Hybrid approaches to optimization and ML methods: a systematic literature review. *Mach Learn* 1–43
9. Dokeroglu T et al (2019) A survey on new generation metaheuristic algorithms. *Comput Ind Eng* 137:106040
10. Carbas S, Toktas A, Ustun D (eds) (2021) *Nature-inspired metaheuristic algorithms for engineering optimization applications*. Springer, Singapore
11. Rana N et al (2020) Whale optimization algorithm: a systematic review of contemporary applications, modifications and developments. *Neural Computing and Applications* 32:16245–16277
12. Bairathi D, Dinesh G (2020) Random-opposition-based learning for computational intelligence. In: *Information and communication technology for sustainable development: proceedings of ICT4SD 2018*. Springer, Singapore
13. Wang W et al (2016) Using opposition-based learning to enhance differential evolution: a comparative study. In: *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE
14. Hichem H, Rafik M, Messaoud MT (2018) PSO with crossover operator applied to FS problem in classification. *Informatica (Slovenia)* 42(2)
15. Qiu C (2018) Bare bones particle swarm optimization with adaptive chaotic jump for FS in classification. *Int J Comput Intell Syst* 11(1):1–14
16. Wan Y et al (2016) A FS method based on modified binary coded ant colony optimization algorithm. *Appl Soft Comput* 49 248–258
17. Peng H et al (2018) An improved FS algorithm based on ant colony optimization. *IEEE Access* 6:69203–69209.
18. Sharawi MH, Zawbaa M, Emary E (2017) FS approach based on whale optimization algorithm. In: *2017 Ninth international conference on advanced computational intelligence (ICACI)*. IEEE
19. Sayed GI, Darwish A, Hassanien AE (2018) A new chaotic whale optimization algorithm for features selection. *J Classification* 35:300–344
20. Emary E, Zawbaa HM, Hassanien AE (2016) Binary ant lion approaches for FS. *Neurocomputing* 213:54–65
21. Özger ZB, Bolat B, Dırı B (2016) A comparative study on binary artificial bee colony optimization methods for FS. In: *2016 International symposium on innovations in intelligent systems and applications (INISTA)*. IEEE
22. Hancer E, Xue B, Zhang M, Karaboga D, Akay B (2018) Pareto front FS based on artificial bee colony optimization. *Inf Sci* 422:462–479
23. Kuo R, Huang SBL, Zulvia FE, Liao TW (2018) Artificial bee colony-based support vector machines with FS and parameter optimization for rule extraction. *Knowl Inf Syst* 55(1):253–274
24. Arora S, Anand P (2019) Binary butterfly optimization approaches for FS. *Expert Syst Appl* 116:147–160

25. Xue B et al (2012) A particle swarm optimization based multi-objective filter approach to FS for classification. In: PRICAI 2012: trends in artificial intelligence: 12th Pacific Rim international conference on artificial intelligence, Kuching, Malaysia, 3–7 September 2012. Proceedings 12. Springer, Berlin Heidelberg
26. Xue B et al (2014) Binary PSO and rough set theory for FS: a multi-objective filter based approach. *Int J Comput Intell Appl* 13(02):1450009
27. Xue B et al (2013) Multi-objective evolutionary algorithms for filter based FS in classification. *Int J Artif Intell Tools* 22(04):1350024
28. Hancer E, Xue B, Zhang M (2018) Differential evolution for filter FS based on information theory and feature ranking. *Knowl-Based Syst* 140:103–119
29. Tabakhi S, Moradi P (2015) Relevance–redundancy FS based on ant colony optimization. *Pattern Recogn* 48(9):2798–2811
30. Moayedikia A et al (2015) Weighted bee colony algorithm for discrete optimization problems with application to FS. *Eng Appl Artif Intell* 44:153–167
31. Wang Y et al (2018) Optimizing cuckoo FS algorithm with the new initialization strategy and fitness function. In: International conference on cloud computing and security. Springer International Publishing, Cham
32. Moradi P, Gholampour M (2016) A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. *Appl Soft Comput* 43:117–130
33. Mistry K et al (2016) A micro-GA embedded PSO FS approach to intelligent facial emotion recognition. *IEEE Trans Cybern* 47(6):1496–1509
34. Chen K, Zhou FY, Yuan XF (2019) Hybrid particle swarm optimization with spiral-shaped mechanism for FS. *Expert Syst Appl* 128:140–156
35. Mafarja M et al (2019) Binary grasshopper optimization algorithm approaches for FS problems. *Expert Syst Appl* 117:267–286
36. Mafarja M et al (2018) Evolutionary population dynamics and grasshopper optimization approaches for FS problems. *Knowl-Based Syst* 145:25–45
37. Tubishat M et al (2020) Improved Salp Swarm algorithm based on opposition based learning and novel local search algorithm for FS. *Expert Syst Appl* 145:113122
38. Mirjalili S (2016) SCA: a sine cosine algorithm for solving optimization problems. *Knowl-Based Syst* 96:120–133
39. Kumar L, Bharti KK (2021) A novel hybrid BPSO–SCA approach for FS. *Nat Comput* 20:39–61
40. Hans R, Kaur H (2020) Hybrid binary Sine cosine algorithm and Ant Lion optimization (SCALO) approaches for FS problem. *Int J Comput Mater Sci Eng* 9(01):1950021

ECG Classification System with Novel Spike-Driven Simulated Neural Network



Deepa Jose, Fahima Niaz, K. Kanaga Durga, and Suchismita Singh

Abstract Cardiovascular diseases (CVDs) are the major cause of mortality globally. The early diagnosis and treatment of CVDs are crucial to saving lives. Electrocardiography (ECG), a non-invasive method, is employed to record the heart's electrical activity. ECG classification plays a pivotal role in diagnosing CVDs and monitoring treatment progress. However, the current ECG classification systems are often energy-intensive, costly, and lack accuracy. This paper introduces a novel ECG classification system that leverages Adaptive Spiking Neural Networks (ASD SNNs) implemented on a Field-Programmable Gate Array (FPGA). Compared to existing systems, ASD SNNs are more energy-efficient, cost-effective, and accurate. The proposed system holds the capability to transform ECG monitoring and improve the quality of life for individuals with CVDs.

Keywords Cardiovascular diseases · Early diagnosis · Treatment · Electrocardiography · ECG classification · Adaptive spiking neural networks · Field-programmable gate array · Energy efficiency · Cost-effectiveness · Accuracy

1 Introduction

Cardiovascular diseases (CVDs) encompass a range of conditions affecting the heart and blood vessels, constituting a significant global health challenge. With a brief exploration of the prevalence and impact, the introduction sets the stage by emphasizing the widespread nature of CVDs and their substantial contribution to morbidity and mortality worldwide. By providing statistical insights and perhaps citing notable studies, the overview aims to convey the urgency of addressing CVDs as a critical public health concern.

D. Jose (✉) · F. Niaz · K. Kanaga Durga · S. Singh
Department of Electronics and Communication, KCG College of Technology, Chennai 600097,
Tamil Nadu, India
e-mail: deepa.ece@kcgcollege.com

1.1 Importance of Early Diagnosis and Treatment

Underscoring the paramount importance of early detection and timely intervention, this section delves into the potential consequences of delayed diagnosis and treatment for individuals with CVDs. It emphasizes that early identification of cardiovascular issues significantly improves the chances of successful intervention, thereby preventing severe outcomes such as heart attacks, strokes, or chronic heart conditions. By drawing attention to the impact of timely medical attention, this part of the introduction lays the foundation for understanding the critical role of diagnostic tools in managing cardiovascular health.

1.2 Role of Electrocardiography (ECG) in CVDs

Shifting the focus to the role of Electrocardiography (ECG), this segment elucidates the non-invasive nature of ECG procedures and their pivotal role in gathering essential information about the heart's electrical activity. The introduction explains how ECG serves as a fundamental diagnostic tool, capturing the heart's rhythm and identifying irregularities. Additionally, it highlights the versatility of ECG in aiding healthcare professionals in assessing cardiac health, from routine check-ups to diagnosing specific conditions [1–4]. By conveying the accessibility and importance of ECG in cardiovascular health management, this portion establishes the relevance of the proposed classification system in enhancing the capabilities of this widely-used diagnostic technique (Fig. 1).

An Electrocardiogram (ECG) is a graph that displays the electrical activity of the heart. The small upward curve, known as the P wave, signals the start of a heartbeat. This is followed by a sharp peak called the QRS Complex, which represents the main part of the heartbeat. The T wave, a smaller curve after the QRS, indicates the end of a heartbeat. The flat line from the P wave to the QRS, known as the PR Segment and Interval, shows the time between the start and main part of a heartbeat. The ST Segment is another flat line from the QRS to the T wave, showing part of the end of a heartbeat. The QT Interval, a line from the start of the QRS to the end of the T wave, represents the total time of a heartbeat. Each part of the ECG provides crucial information about the heart's activity and can assist doctors in diagnosing heart conditions (Fig. 2).

2 Current Landscape of ECG Classification Systems

This section provides background for our proposed Adaptive Spiking Neural Networks (ASD SNNs) system by looking at the present status of ECG categorization. This paradigm allows for a critical appraisal of the area, revealing both its

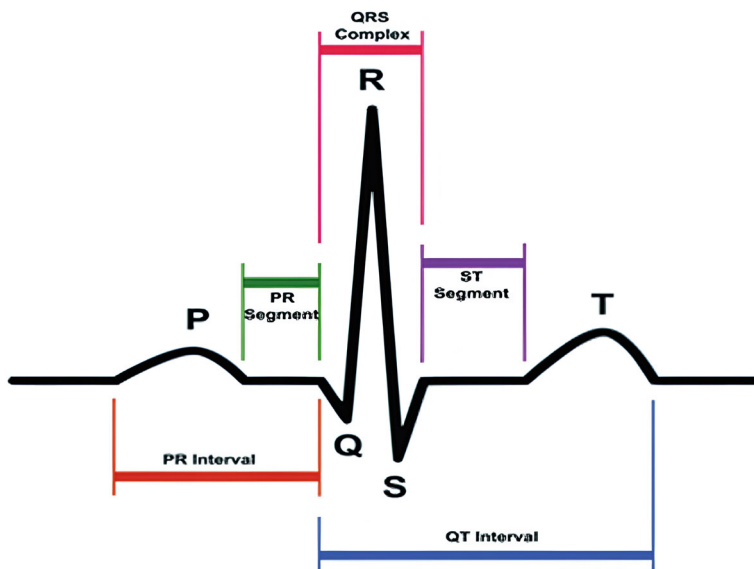


Fig. 1 Structure of an Electrocardiogram waveform

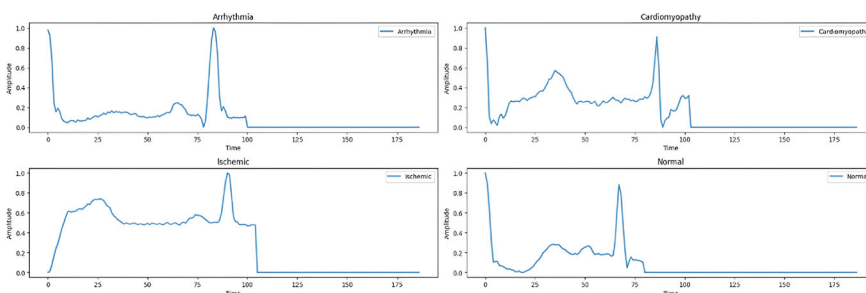


Fig. 2 All the classes of CVD's taken into account from the dataset

successes and Lingering issues. Limitations and opportunities in existing systems [5–8]. After reviewing the current landscape, we investigate the fundamental limitations of modern ECG categorization algorithms. We identify and assess important difficulties, including:

- Energy Consumption:** Some systems' high energy needs limit their mobility and long-term sustainability.
- Cost Considerations:** Access to some technologies may be limited due to cost constraints, especially in resource constrained environments.
- Errors in ECG interpretation** represent a major danger to patient care, highlighting the need for improved system accuracy.

3 Proposed System

The proposed ECG classification system, designed for cardiovascular health monitoring, comprises a sophisticated array of components and functionalities. Here's the revised version of the proposed ECG classification system for your research paper:

1. ECG Signal Acquisition

The system commences with the acquisition of ECG signals employing the AD8232 ECG Sensor. This sensor captures the electrical activity of the heart, generating an analog signal representing the heart's electrical waveform.

2. NodeMCU (ESP8266) Integration

Subsequently, the ECG signal is wirelessly transmitted to a NodeMCU ESP8266 board. The NodeMCU, serving as a microcontroller, orchestrates communication between the sensor and other system components.

3. Signal Processing with FPGA

The acquired ECG signal undergoes further processing leveraging a PYNQ-Z2 Field-Programmable Gate Array (FPGA). This FPGA executes vital tasks including filtering and noise reduction to refine the quality of the ECG data.

4. Feature Extraction

Following signal processing, the ECG waveform is subjected to comprehensive analysis for feature extraction. These features encompass essential parameters such as heart rate, QRS complex duration, and ST segment variations.

5. Utilization of Spiking Neural Network (SNN)

Extracted feature vectors are subsequently fed into an adaptive spiking Spiking Neural Network (SNN). This machine learning model, adept at capturing intricate temporal dynamics, is meticulously trained on annotated ECG datasets to discern nuanced patterns associated with various cardiac conditions (Fig. 3).

6. Heart Condition Classification

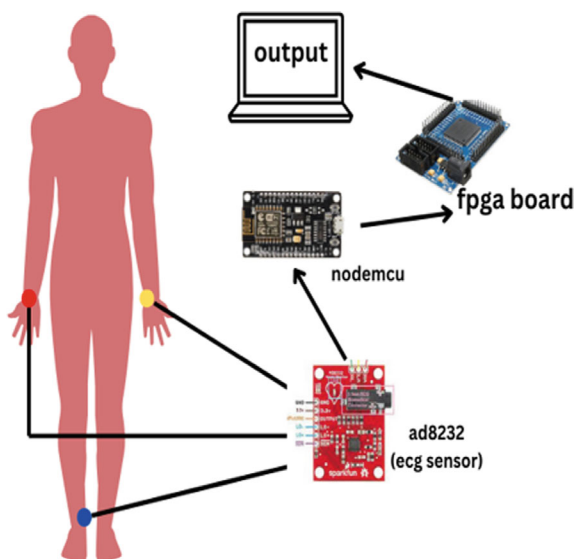
The trained SNN undertakes the classification of ECG signals into distinct categories of heart diseases, encompassing:

Arrhythmia: Characterized by irregular heart rhythms (e.g., atrial fibrillation, ventricular tachycardia).

Cardiomyopathy: Involving ailments affecting the heart muscle (e.g., dilated cardiomyopathy, hypertrophic cardiomyopathy).

Ischemic Heart Disease: Marked by diminished blood flow to the heart due to arterial constriction or blockage (e.g., angina, myocardial infarction).

Fig. 3 ECG monitoring system



7. Output Generation and Alert Mechanism

The system furnishes an output denoting the detected heart condition, accompanied by an alert system designed to promptly notify healthcare practitioners or patients of any aberrant ECG patterns.

This envisioned system amalgamates hardware components such as the AD8232 sensor, NodeMCU ESP8266, and PYNQ-Z2 with sophisticated machine learning techniques, promising advancements in the early detection and management of cardiac ailments. It is imperative to underscore the requisite stages of meticulous design, development, and validation preceding the realization of such a system. Furthermore, adherence to stringent medical validation protocols and expert consultation is indispensable for ensuring the efficacy and safety of the proposed solution. On the ASD SNN analysis, providing valuable diagnostic insights. In summary, the proposed system, with the inclusion of the AD8232 ECG Sensor Module and other meticulously chosen components, forms a comprehensive solution for accurate, portable, and real-time ECG classification. The synergy between hardware components and neural network processing positions this system at the forefront of cardiovascular health monitoring technology.

4 Methodology

The ASD SNN model finds its application in ECG signal filtering on the PYNQ-Z2 board, enhancing its functionality and utility in medical diagnostics. Leveraging the unique capabilities of the PYNQ-Z2 board, the model efficiently processes ECG

signals, effectively filtering out noise and artifacts to extract meaningful cardiac information.

By implementing the ASD SNN model on the PYNQ-Z2 board, healthcare professionals can benefit from real-time ECG signal processing with enhanced accuracy and reliability. The board's FPGA-based architecture facilitates high-speed computation, enabling rapid analysis of ECG data streams while maintaining low latency.

The ASD SNN model is fine-tuned to fully harness the capabilities of the SNN potential using the attention mechanism. This enables it to carry out the classification task with low energy consumption and identify significant characteristics related to the timing, shape, and multi-channel depictions of the ECG signal.

Through meticulous optimization, the ASD SNN model capitalizes on the intrinsic advantages of SNNs, particularly in power efficiency, while leveraging the attention mechanism to capture intricate temporal, morphological, and multi-channel characteristics of ECG signals (Fig. 4).

Rigorously evaluated on the MIT-BIH database, the ASD SNN model has demonstrated remarkable performance metrics. With an overall classification accuracy of 98.26%, sensitivity of 94.75%, and an impressive $F1$ score of 89.09%, it showcases its efficacy in accurately identifying cardiac irregularities. Furthermore, its efficient energy consumption and rapid runtime render it highly suitable for real-time deployment on personal portable devices, thus holding significant promise for practical clinical applications (Fig. 5).

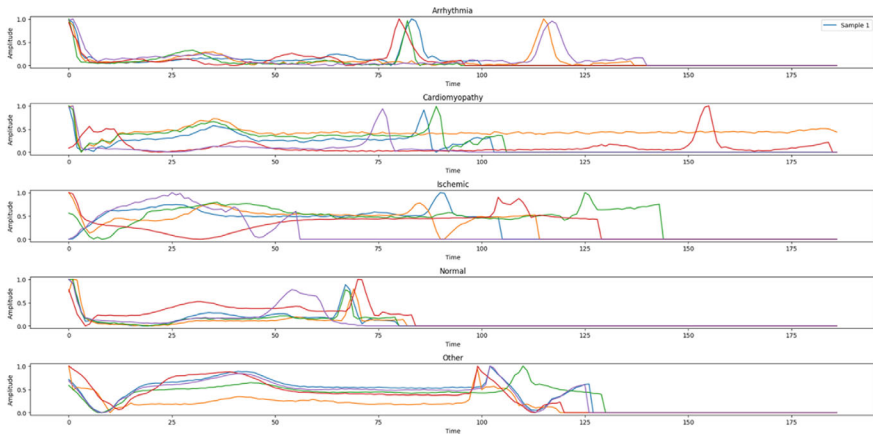


Fig. 4 Visual representation of heartbeat samples across different classes

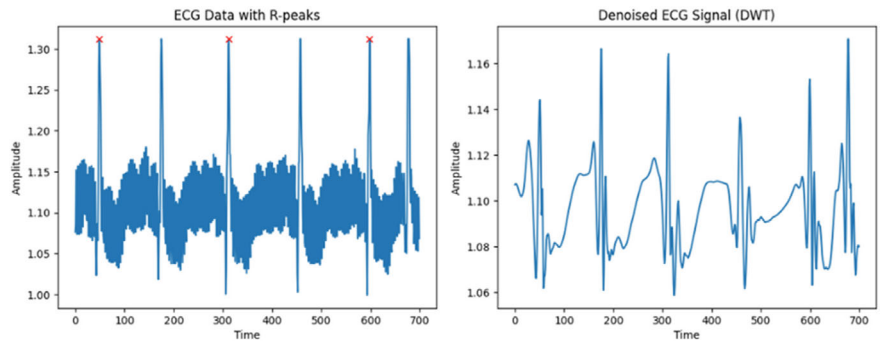


Fig. 5 Sample preprocessed reading with ad8232 sensor and DWT filtered sensor readings

5 Results and Comparisons

The research project was centered around the development and evaluation of a Spiking Neural Network (SNN) model for the classification of heartbeat signals. The model was trained and tested using the MIT-BIH Arrhythmia Database, which comprises Electrocardiogram (ECG) recordings of heartbeat signals classified into five distinct classes: Normal, Supraventricular, Ventricular, Fusion, and Unknown [9, 10]. The SNN model architecture incorporated convolutional layers for feature extraction from the input ECG signals, followed by fully connected layers for classification based on the extracted features. The model was trained using the Adam optimizer with a learning rate of 0.001 for 50 epochs with a batch size of 1000. The model’s performance was monitored based on validation accuracy during training, and the best performing model was saved for further evaluation (Table 1).

The evaluation of the model’s performance on the test dataset revealed an accuracy of 97.83%, indicating its effectiveness in classifying heartbeat signals. A confusion matrix was generated to analyze the distribution of predicted classes compared to the

Table 1 Result

Condition	Precision	Recall	F1-Score	Support
Arrhythmia	0.9829	0.9949	0.9889	18,118
Cardiomyopathy	0.9009	0.7032	0.7899	556
Ischemic	0.9537	0.9254	0.9394	1448
Normal	0.8058	0.6914	0.7442	162
Other	0.9835	0.9639	0.9736	1608
Overall				
Accuracy			0.9783	21,892
Macro avg	0.9254	0.8558	0.8872	21,892
Weighted avg	0.9776	0.9783	0.9776	21,892

ground truth labels, providing insights into the model’s performance across different classes. This visualization highlighted areas where the model performed well and areas that require improvement (Table 2).

In addition to the SNN model, the research project also involved the development and evaluation of a Random Forest Classifier and a Support Vector Machine (SVM) model for comparison purposes. These models were trained and tested using the same dataset, and their performance was evaluated based on accuracy and a classification report. The Random Forest model achieved a test accuracy of 96.7%, while the SVM model achieved a test accuracy of 93.2% (Fig. 6).

Additionally, we generated visualizations to provide insights into the model’s predictions. The confusion matrix helped us understand the model’s performance

Table 2 Comparisons chart

Class	SNN <i>F1</i> -score	Random forest <i>F1</i> -score	SVM <i>F1</i> -score
Arrhythmia	0.9889	0.9858	0.9825
Cardiomyopathy	0.7729	0.7483	0.7120
Ischemic	0.9442	0.9280	0.9117
Normal	0.8037	0.7253	0.5865
Other	0.9840	0.9687	0.9516

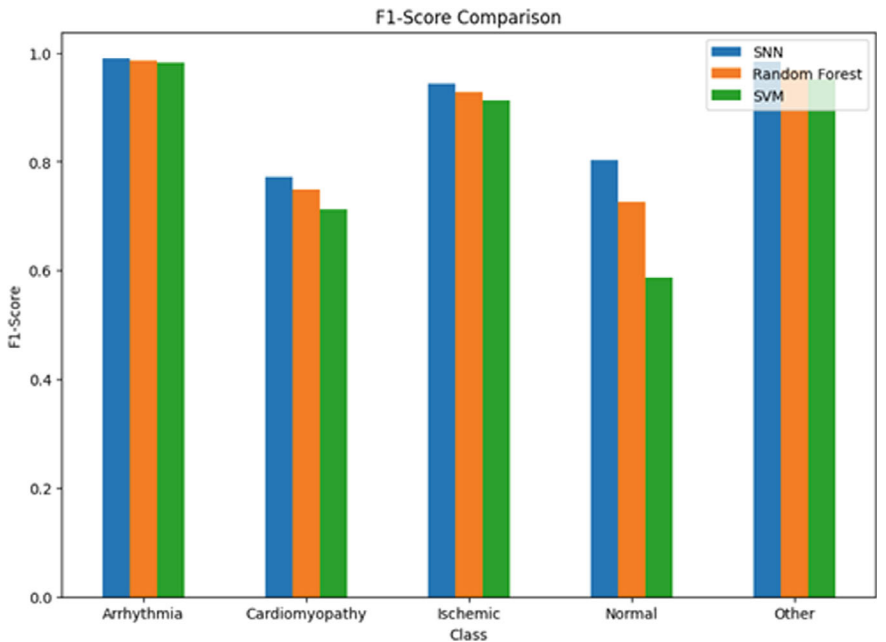


Fig. 6 Comparison between SNN, Random forest, and SVM

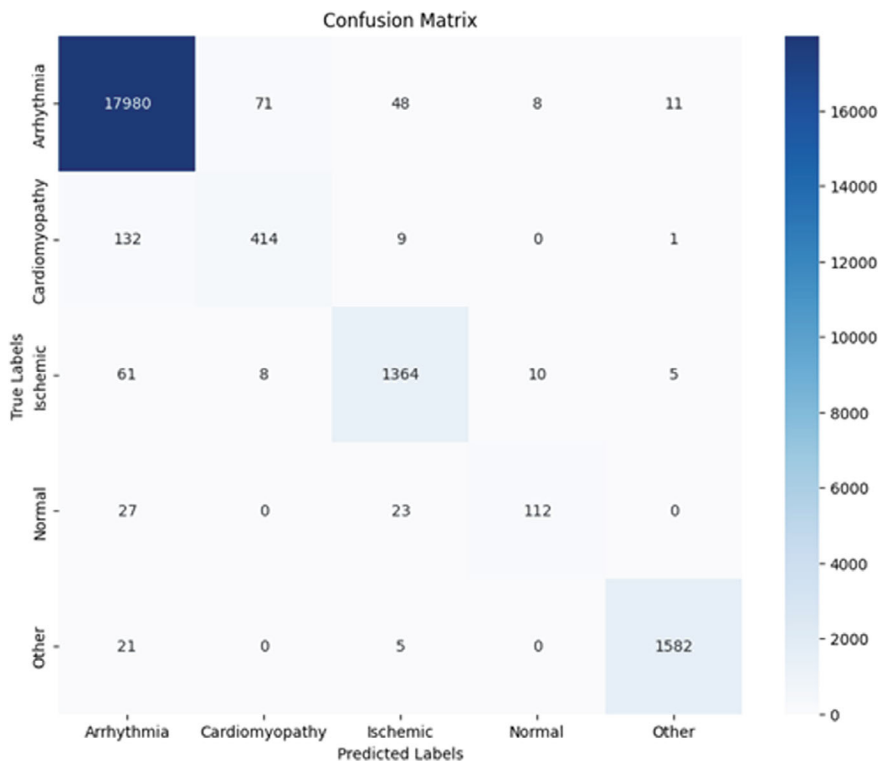


Fig. 7 Confusion matrix

across different classes, highlighting areas where it performed well and areas that require improvement (Fig. 7).

We also visualized sample predictions overlaid with the actual ECG signals. This allowed us to observe how accurately the model predicted different heartbeat classes based on the input signals.

Moving forward, to obtain data from NodeMCU connected to AD8232 and send it to a Colab environment, we can use the NodeMCU to read ECG signals from the AD8232 sensor. We can then establish a communication link between the NodeMCU and a server running on Colab using protocols like MQTT or HTTP. The NodeMCU can periodically send the ECG data to the server, where it can be processed and used as input to the SNN model for real-time classification.

This integration would enable continuous monitoring of ECG signals remotely and provide insights into heart health in real-time, opening avenues for applications in telemedicine and remote patient monitoring.

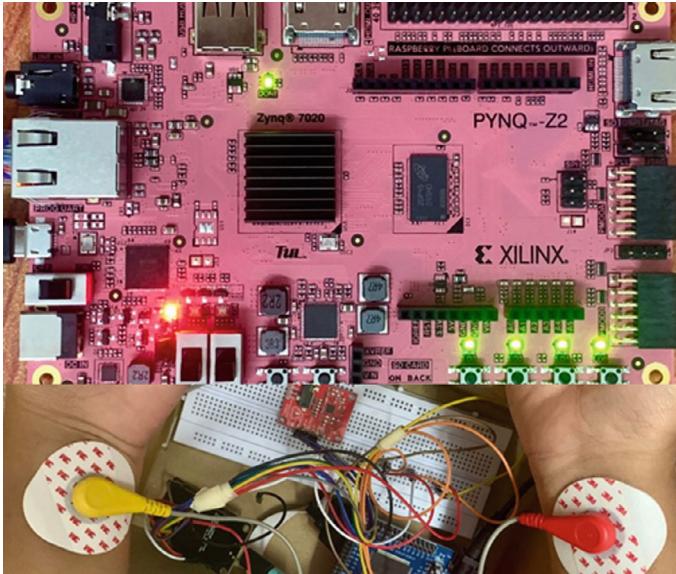


Fig. 8 Implementation of the project

6 Conclusion

This research project successfully developed and evaluated a Spiking Neural Network (SNN) model for the classification of heartbeat signals. The model demonstrated high accuracy and effectiveness, outperforming other traditional machine learning models like Random Forest and Support Vector Machine. The integration of real-time data acquisition devices with the model opens up new avenues for applications in telemedicine and remote patient monitoring (Fig. 8).

The project also highlighted the importance of preprocessing and noise filtration in enhancing the quality of ECG data for further analysis. The use of an FPGA board for this purpose proved to be effective. However, the performance variation across different classes signals opportunities for further refinement and optimization (Fig. 9).

Going forward, the emphasis could be on enhancing the model's effectiveness across various categories and investigating its usefulness in other medical diagnostic activities. The model's integration with devices that capture real-time data could also be fine-tuned for better efficiency and precision in remote patient monitoring. In general, this study adds to the continuous work in utilizing artificial intelligence for healthcare purposes, especially in cardiology.

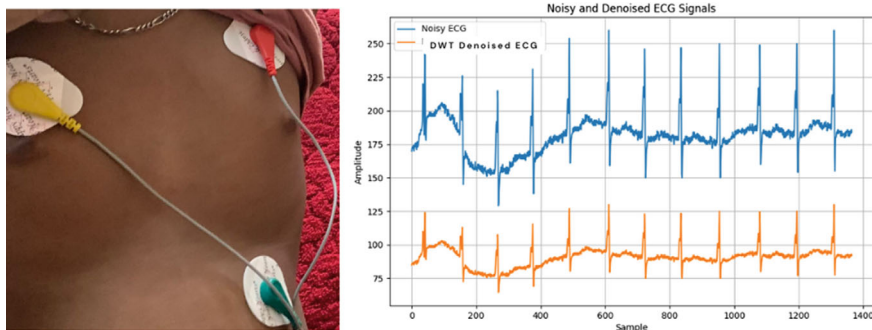


Fig. 9 Implementation of taking real-time reading in the chest

References

1. Priyankkumar H, Prajapati et al (2022) Hardware efficient low-frequency artifact reduction technique for wearable ECG device. *IEEE Trans Instrum Meas* 71
2. Lee D (2021) Energy-efficient FPGA accelerator with fidelity-controllable sliding-region signal processing unit for abnormal ECG diagnosis on IoT edge devices. *IEEE Trans Instrum Meas*
3. Xiao J (2022) ULECGNet: an ultra-lightweight end-to-end ECG classification neural network. *IEEE Trans Instrum Meas* 71
4. Haoming Chu A (2022) Neuromorphic processing system with spike-driven SNN processor for wearable ECG classification. *IEEE Trans Biomed Circuits Syst* 16(4)
5. Edwin Dhas D et al (2022) "Extraction of fetal ECG from abdominal and thorax ECG using a non-causal adaptive filter architecture. *IEEE Trans Biomed Circuits Syst* 16(5)
6. Kumar A et al (2021) Hardware emulation of a biorthogonal wavelet transform-based heart rate monitoring device. *IEEE Sens J* 21(4)
7. Xing Y, Zhang L, Hou Z, Li X, Shi Y, Yuan Y, Zhang F, Liang S, Li Z, Yan L (1889) Accurate ECG classification based on spiking neural network and attentional mechanism for real-time implementation on personal portable devices. *Electronics* 2022;11. <https://doi.org/10.3390/electronics11121889>
8. Dasarapalli SR, Panneerselvam RK, Annavarapu S, Gopireddy KR (2023) Three-lead ECG and heartrate variability using AD8232. In: 2023 14th International conference on computing communication and networking technologies (ICCCNT), Delhi, India, pp 1–6, <https://doi.org/10.1109/ICCCNT56998.2023.10306769>
9. Sowmya S, Deepa Jose (2022) Contemplate on ECG signals and classification of arrhythmia signals using CNN-LSTM deep learning model. *Meas: Sens* 24:100558, ISSN 2665-9174. <https://doi.org/10.1016/j.measen.2022.100558>
10. Ambika A et al (2023) Machine learning based heart disease prediction system. *AIP Conf Proc* 2963(1):20009

Performance Optimization and Application Exploration Based on Intelligent Algorithm and IT Technology in Cloud Computing Task Scheduling System



Ying Ding

Abstract As the cloud computing market continues to grow, there is an increasing need for research into task scheduling strategies in cloud computing. This paper focuses on the limitations of traditional scheduling algorithms and heuristic intelligent scheduling algorithms, and the emphasis is on investigating swarm intelligence search algorithms to enhance the efficiency and reliability of cloud computing task scheduling systems. Through research and practical application, the ant colony algorithm, gray wolf algorithm, whale fish optimization algorithm, and Beluga whale optimization algorithm are being studied and implemented, and designing task and virtual machine sorting mechanism, and conducting experimental comparison on Cloudsim platform, it is concluded that no need to sort task and virtual machine in advance and enhancing the performance of swarm intelligence algorithms can be achieved by incorporating a more judicious level of randomness. In addition, to solve the problem of uneven load and low early search efficiency of ant colony algorithm, this paper introduces a novel algorithm called the Immune-Ant Colony Optimization Algorithm (IMACO), and improved Beluga whale optimization algorithm (IBWO), which introduce load monitoring heuristic, dynamic volatility and dual reward and penalty pheromone update strategy, and make full use of the global search and diversity advantages of immune algorithm. Effectively improve algorithm performance. The simulated annealing and quasi-reverse learning strategies are combined to solve the local optimal problem and achieve better completion time, system load balance and cost control. The outcomes indicate that these methods can enhance the efficiency of task scheduling. These researches provide practical value for improving the performance and stability of cloud computing system.

Keywords Cloud computing task scheduling · Performance optimization · Intelligent algorithm · IT technology · Body intelligent algorithm

Y. Ding (✉)
IT Solution Consulting, MOYI Inc, New York, NY, USA
e-mail: yingding0410@yahoo.com

1 Introduction

As cloud computing gains prominence alongside the widespread adoption of 5G technology, it has emerged as a prominent focus in both business and technology realms. This evolution not only offers extensive computational capabilities and storage but also transforms how users engage with computing resources, thereby advancing information technology. Nevertheless, the growing user base in cloud computing presents a challenge: optimizing task scheduling and resource utilization effectively.

Given this backdrop, the investigation into cloud computing task scheduling strategies assumes heightened significance. Users have clear requirements for service quality, including cost control, task completion time, and system reliability. At the same time, cloud service providers also need to reduce costs and improve service quality by optimizing scheduling strategies. However, the existing general algorithms cannot fully meet the optimization requirements of different tasks, so the urgency of researching cloud computing task scheduling algorithms is growing.

The exploration of cloud computing task scheduling strategy has diversified into multiple avenues, including algorithm migration, optimization of existing algorithms, hybrid algorithms, and new algorithms. The purpose of this paper is to analyze the advantages and disadvantages of existing algorithms through scientific investigation and experiments, and put forward improvement measures to obtain a more effective cloud computing task scheduling algorithm.

This paper aims to contribute new insights and methodologies to the research and application of cloud computing, addressing the increasing demand, by examining the influence of task and virtual machine preordering on task scheduling algorithm performance in cloud environments and introducing a novel optimization algorithm.

2 Related Research

Task scheduling optimization is a critical focus within the domain of cloud computing. Alsadie proposed the MDVMA framework, which uses NSGA-II algorithm to minimize energy use, manufacturing span and cost, and provides cloud service providers with a better resource management method [1]. Durairaj proposed a multi-objective optimization design method to improve the task assignment performance of virtual machines, and achieved significant progress in time and utilization [2]. These studies highlight the importance of task scheduling in cloud computing, which directly affects the performance of the entire system. Alshattnawi's load balancing algorithm reduces the task response time [3]. The ERES algorithm of DGMS Garg reduces energy consumption and task time [4].

Wang used FP-Growth-M method to mine the relationship between meteorological models and data sets, and proposed a heuristic algorithm for arrangement and scheduling [5]. Simulation results show that this approach effectively reduces the number of clouds, external cloud file size, and transfer time. Madi proposed a method

to evaluate the security of tenants' virtual infrastructure in cloud environments, and monitored the security level of tenants by calculating virtual infrastructure security metrics [6].

Supreeth improves and optimizes the task and virtual machine sorting algorithm in cloud computing, thus realizing the optimization of time, energy, and resource utilization in terms of task scheduling and resource allocation [7]. The verification results show that these algorithms can improve significantly. An improved Beetle group optimization algorithm was proposed by Bhagavathi for energy-aware virtual machine integration [8]. The method, which aims to improve energy efficiency and reduce costs, outperformed previous methods in the study results.

Sharma proposed a virtual machine placement algorithm that optimizes energy usage on physical servers [9], which not only improves the availability of resources, but also reduces the energy consumption of data centers, solving the main challenge of providing a good level of service in cloud environments.

Potluri proposed a hybrid algorithm for task scheduling in cloud computing environment [10], which focused on QoS parameters such as time and cost. Compared with other task scheduling algorithms, this algorithm performs better in QoS parameters and brings a new solution for task scheduling in cloud computing environment.

Agarwal also proposed a task scheduling technology based on adaptive genetic algorithm [11] to optimize the task allocation efficiency and resource utilization in the cloud computing environment. The simulation results show that the proposed algorithm performs well in key indexes and is suitable for cloud computing task scheduling research.

These studies collectively contribute to an in-depth exploration of cloud computing task scheduling and virtual machine optimization issues, offering valuable insights and solutions for enhancing performance in cloud computing environments.

3 Investigations into the Application and Refinement of Swarm Intelligence Algorithms for Optimizing Cloud Computing Task Scheduling

3.1 Application of Swarm Intelligence Algorithm in Cloud Computing Task Scheduling

As an innovative computing method, cloud computing builds a resource pooling environment, which can provide all kinds of computing resources on demand, including network, storage and servers, and forms a resource sharing platform. NIST's definition of cloud computing highlights its notable flexibility. And cost-effectiveness, with users able to quickly access the resources they need based on actual needs, while service providers have relatively low management costs.

Cloud computing architecture is primarily structured into three layers: the core service layer, service management layer, and user access interface layer.

In the core service layer, a range of service models is offered, encompassing infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS), catering to diverse user requirements and establishing the groundwork for widespread cloud computing adoption.

The application of virtualization technology converts large-scale physical resources into virtual resources, such as VMS, enabling cloud resources to have dual scaling features. This paper focuses on the scheduling algorithm between tasks and virtual machines, which is very important for allocating computing resources and improving system performance. Cloud computing task scheduling strategies can be categorized into traditional, intelligent heuristics.

Swarm intelligence algorithms play an important role in cloud computing task scheduling. Swarm intelligence algorithms mainly include ACO, GWO, WOA, BWO, etc. These algorithms show excellent search performance and simple parameter setting in optimizing task scheduling, which has attracted much attention and research by researchers.

Ant colony algorithm (ACO) and gray wolf algorithm (GWO), the main difference between the two is that ant colony algorithm searches for the optimal solution by means of the release and update of pheromones through the adaptation and cooperation stages of ants. The gray wolf algorithm, on the other hand, simulates the hierarchical structure and cooperative behavior of wolves, making wolves work together to find the best solution through steps such as hierarchical layering, enveloping, and hunting. However, both algorithms demonstrate the powerful potential of swarm intelligence in solving complex problems, providing a wide range of optimization possibilities for various application scenarios.

The whale optimization algorithm (WOA) optimizes search by simulating humpback whale group hunting behaviors such as rounding up, bubble net attacks, and searching for prey. Each humpback whale's position corresponds to a potential solution, and the algorithm constantly updates the whale's position to find the global optimal solution.

The White Whale Optimization Algorithm (BWO) consists of three stages: exploration, development, and whale landing. The position is updated through synchronization or mirroring, while in the development phase, Levi's flight strategy is introduced to enhance global search capabilities. In the whale landing phase, the process of whale migration and foraging is simulated, and the population change information is controlled through the update formula of the position.

3.2 Designing a Framework for Cloud Computing Task Scheduling

The relationship between tasks and virtual machines can be simply described in cloud computing as: each task (such as running a software program) needs to run on a virtual machine (equivalent to a virtual computer). We can represent this relationship with Table 1. In this context, ‘1’ signifies task assignment to the respective virtual machine, while ‘0’ denotes non-assignment.

To better manage tasks and virtual machines, they are sorted. The task priority is determined based on the complexity and data volume of the task and VM performance indicators, such as processing capability and storage space. With sorted tasks and VMS, different scheduling policies can be adopted.

For better dispatch work, we need to predict the time we will spend on various virtual planes. With this time it is possible to integrate the sending time and computational time of data. Data transmission time can be estimated depending on the amount of data and bandwidth of the mission, and computation time is related to the computing power of the virtual machine and the number of tasks.

When constructing the fitness function, three indicators are considered: task completion time, resource load balancing, and task cost. This study defines the total task completion time as *TCT*, which mainly reflects the time required for task completion in the system. The calculation of *TCT* is based on the task virtual machine mapping matrix, and applies virtual machine parallel processing to calculate the maximum predicted execution time of tasks.

Evaluate the degree of load balancing between virtual machines in the system through the resource load balancing function *LB*. The task allocation of each virtual machine in the computing system can be quantified using *LB* to achieve stable system load and efficient resource utilization.

Integrating indicators into the fitness function *Obj* can achieve multi-objective optimization problems. By balancing task completion time, system load balancing, and task cost, the overall performance and efficiency of the system can be improved. This comprehensive approach that considers multiple factors can effectively guide the optimization and improvement of cloud computing task scheduling strategies.

Table 1 Task distribution

Task\virtual machine	1	2	3	...	m
1	1	0	0	...	0
2	0	1	0	...	0
...	...	...	...	...	...
n	0	0		...	1

4 Task Design and Performance Comparative Analysis

4.1 Task Design Experiment Based on IT Technology

In this study, ant algorithm, gray wolf algorithm, whale improvement algorithm, and white whale improvement algorithm were used to improve the work plan. By conducting controlled experiments on the Crowdsong Cloud platform, we set up several virtual machines and cloud flight missions, and were able to create a total of eight virtual machines and various work groups. The details of the experiment has been presented in Table 2.

For each combination, we pre-sorted and non-sorted the tasks and virtual machines separately, and applied four different intelligent algorithms for task scheduling. The experimental results include the evaluation of completion time, load balancing degree, task cost function value, and comprehensive factor size for each algorithm under different sorting conditions.

Cloudsim platform, a widely used cloud computing simulation tool, is used for task scheduling in this study. The platform provides data center virtualization and various scheduling policy evaluation functions.

In this study, Cloudsim platform was used for experimental environment configuration, including the creation of two data centers and the setting of host parameters, which involved the number of hosts, RAM, bandwidth, CPU cores, computing power, and external memory space. At the same time, the user cloud task parameters are also set, such as task data amount, output size, and task computation amount. Before the experiment, the parameters of the ant colony algorithm were set, including the initial pheromone and the parameter values.

This paper analyzes the performance of cloud computing scheduling algorithms by assessing task completion time, system load balancing degree, and task completion cost. It calculates comprehensive factors to evaluate the overall effectiveness of the algorithms. By employing the ant colony algorithm, gray wolf algorithm, whale optimization algorithm, and Beluga optimization algorithm, and observing their performance variations under different numbers of virtual machines and tasks, the conclusion is drawn:

The ant colony algorithm demonstrates variances in task completion time and completion cost as the number of tasks increases, particularly notable when the

Table 2 Experimental settings of the comparative analysis algorithm

Main evaluation index	Task scheduling mode	Number of servers	Number of tasks	Test algorithm
Completion time, system leveling, cost, efficiency	Preallocation	Six stations	200	ACO
			400	GWO
	Non-preallocation	Eight channels	800	WOA
			1600	BWO

number of virtual machines is set to 6. However, the load balancing degree and comprehensive factor of each algorithm are better than the sorting condition when the task and virtual machine are not sorted.

Increasing the number of virtual machines to 8 leads to a significant improvement in the comprehensive factor of the algorithms without sorting, resulting in shortened completion times, and the load balance is also improved, especially when tasks and virtual machines are not sorted, the performance advantage of ant colony algorithm is more significant.

These results show that under a certain number of VMS and tasks, selecting a proper scheduling algorithm does not necessarily need to sort tasks and VMS, but to adjust the algorithm and parameters according to the specific situation to obtain better performance.

4.2 Analysis of Experimental Findings

This paper statistically compares the performance metrics of the ant colony algorithm (ACO), gray wolf algorithm (GWO), whale optimization algorithm (WOA), and Beluga optimization algorithm (BWO) across varying numbers of virtual machines and tasks.

After comparing the completion time data of four algorithms under different numbers of virtual machines and tasks, when the number of virtual machines is 6, as shown in Figs. 1 and 2; when the number of virtual machines increases to 8, as shown in Figs. 3 and 4. Regardless of whether the number of virtual machines is 6 or 8, and the number of tasks is 200, 400, 800, 1600, it shows that there is no pre-sorting between tasks and virtual machines, as shown in Tables 3 and 4. The completion time of each algorithm has significantly decreased, with a maximum reduction of 68.88%. Specifically, the ant colony algorithm has the longest completion time and significant fluctuations in these situations, indicating that there is significant room for improvement in its performance.

The performance of the four algorithms in terms of system load balancing, completion cost, and overall synthesis factor is compared across varying numbers of virtual machines (VMs) and tasks. It is noted that preordering tasks and virtual machines have a significant impact on the system's load balancing, and the load balance can be reduced to 67.03% when the task is not sorted. The ant colony algorithm exhibits the poorest performance and notable fluctuations in this regard, which needs further improvement.

In terms of completion cost, the ordering of virtual machines and tasks has little impact, and the cost of ACO algorithm can be reduced to 22.61% in some cases, while the cost of other algorithms can be kept within 7.5%. In terms of comprehensive factors, the comprehensive factor values of all algorithms generally decline when there is no pre-ranking, which shows that swarm intelligence algorithm shows better performance in cloud task scheduling. The ant colony algorithm's comprehensive factor is obviously higher than other algorithms, but there is still room for

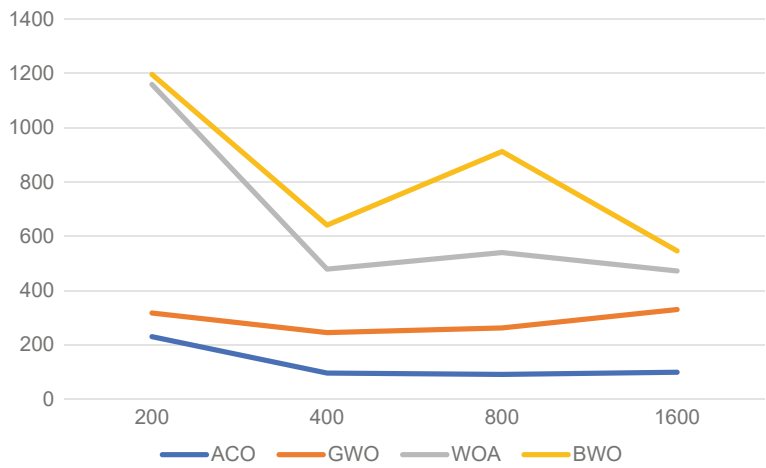


Fig. 1 Comparison of the completion time of the four algorithms with the number of six VMs and the sorting

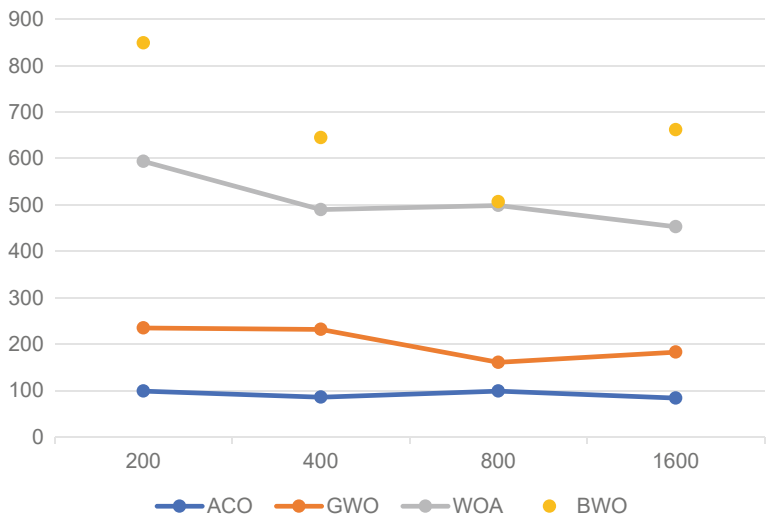


Fig. 2 Comparison of the completion time of the four algorithms with six VMs and no sorting

improvement, while the Beluga optimization algorithm is stable and excellent, with a large advantage, and will be further improved to improve performance.

After comparing the performance of four swarm intelligence algorithms under different VM and task numbers on Cloudsim platform, it is found that in cloud computing environment, no preordering of tasks and VMS is more beneficial to improve algorithm performance. The preordering process uses the static method of

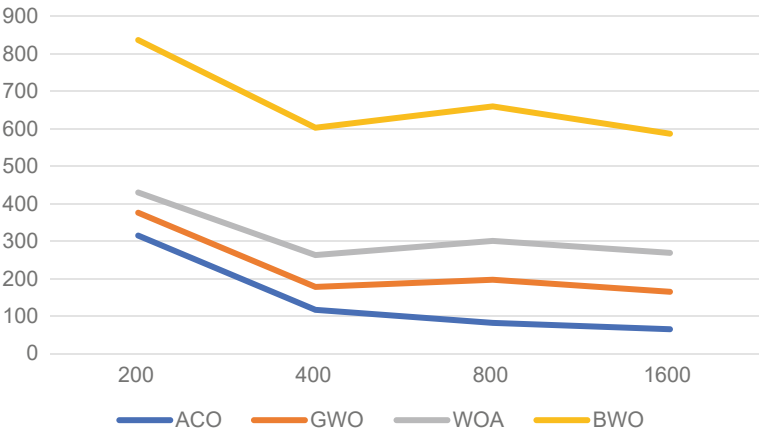


Fig. 3 Comparison of the completion time of the four algorithms with the number of eight VMs and the sorting

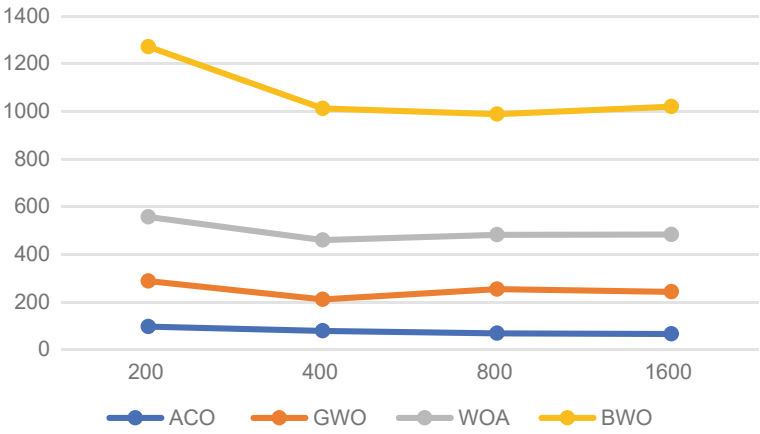


Fig. 4 Comparison of the completion time of the four algorithms with 8 VMs and no sorting

Table 3 Completion time data of the four algorithms with six VMS and six tasks

Number of dummy machines	Algorithm	ACO		GWO		WOA		BWO	
	Quantity of duties	Sort	No sort	Sort	No sort	Sort	No sort	Sort	No sort
6	200	230	99	317	235	1159	594	1196	849
	400	96	86	245	232	479	490	641	645
	800	91	99	262	161	540	499	912	507
	1600	99	84	330	183	472	453	546	662

Table 4 Completion time data of the four algorithms with eight VMS and eight tasks

Number of dummy machines	Algorithm	ACO		GWO		WOA		BWO	
	Quantity of duties	Sort	No sort	Sort	No sort	Sort	Quantity of duties	Sort	No sort
8	200	315	98	376	191	430	269	837	714
	400	117	80	178	132	263	249	603	552
	800	82	70	197	185	301	228	660	506
	1600	65	67	165	177	269	240	587	537

the Java Collections class, collections.sort (), which adds additional time consumption and increases the overall runtime of task assignments. While pre-sorting ensures task and virtual machine fit, it compromises a certain amount of search randomness and broader search scope, which can lead to performance degradation.

4.3 Optimization of Cloud Computing Task Scheduling System

A novel ant colony algorithm named IMACO algorithm is introduced. The key features of IMACO algorithm are dynamic updating of pheromone and initializing pheromone by immune operation to improve the load balance and assesses its convergence speed in cloud computing task scheduling.

In the context of cloud computing task scheduling, ant colony algorithm is easy to lead to system load imbalance and low search efficiency. To solve these problems, load monitoring heuristic factors and dynamic volatilization and dual reward and punishment pheromone updating strategies were introduced. In the ant colony algorithm, the effect of ant colony search is affected by the concentration of pheromone, while the volatile factor of pheromone ρ is controlled by the evaporation rate of pheromone, which affects the search efficiency to some extent.

To address the issues of initial pheromone distribution and low search efficiency, the immune algorithm concept was incorporated. The algorithm’s process involves antigen recognition, randomly initializing antibodies, calculating affinity, antibody concentration and excitation value, immune treatment, etc., and finally forming a new generation of optimized antibody group, thus improving the performance and adaptability of the algorithm.

Combining the characteristics of immune algorithm and ant colony algorithm, we designed immune ant colony optimization algorithm to realize effective optimization of cloud computing task scheduling and balanced management of system load. The essence of this algorithm lies in promoting efficient task allocation and load balancing through dynamic pheromone updates and diversity maintenance. During the iteration process, the algorithm continuously adjusts the task virtual machine

allocation scheme, rewards high-quality solutions, and punishes poorer solutions to obtain the optimal task execution duration and system workload status.

The steps of the improved Beluga optimization algorithm are to discover new possibilities, further develop these possibilities, accept the next best solution, perform a reset, and finally learn and improve. Within the algorithm, the objective function value of each Beluga whale is calculated through initialization parameters and loop iteration, and the optimal individual is updated. The exploration or development stage is decided according to the balance factor and the probability of whale fall. The position of the Beluga whale is adjusted according to the objective function value at each step, and the optimal task and virtual machine allocation scheme is finally output. The whole process realizes the effective optimization of task and virtual machine allocation by simulating the annealing algorithm and quasi-reverse learning strategy.

The parameters are initialized algorithmically, initial pheromones are generated, and the task is randomly assigned by the immune algorithm—the virtual machine. As the number of iterations increases, the algorithm will update the antibody group and increase the diversity of the search space through mutation and population refresh operations. When the immune algorithm reaches a preset number of iterations, it switches to the ant colony algorithm for optimization. Ant colony algorithm uses the dynamic update of pheromone to adjust the task allocation strategy and realize the balanced management of system load.

5 Conclusion and Prospect

This study aims to optimize work planning strategies in a cloud computing environment. It offers IMACO (IMACO) and modified IBWO (mobwo) optimization algorithms by conducting an in-depth study of collective intelligence and job sequencing algorithms and virtual machines on performance. Experimental results show that pre-sorting tasks and virtual machines improves algorithm performance, increases the randomness and scope of searches, and better meets user needs. IBWO is the best, followed by IMACO, which is superior to the traditional ant colony algorithm (ACO). Future research, can reduce the energy consumption of the system, the real improvement and solidity of the work schedule and more exploration of optimization goals, while focusing on optimization of the work schedule algorithm in the environment with more clouds, focus on the overall performance of the cloud platform and the quality of service, the overall improvement.

References

1. Alsadie D (2021) A metaheuristic framework for dynamic virtual machine allocation with optimized task scheduling in cloud data centers. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2021.3077901>
2. Durairaj S, Sridhar R (2022) Task scheduling to a virtual machine using a multi-objective mayfly approach for a cloud environment. *Concurr Comput: Pract Exp*
3. Alshattawi S, Al-Marie M (2021) Spider monkey optimization algorithm for load balancing in cloud computing environments. *Int Arab J Inf Technol* 18(5). <https://doi.org/10.34028/iajit/18/5/13>.
4. Garg N, Singh D, Goraya MS (2021) Energy and resource efficient workflow scheduling in a virtualized cloud environment. *Clust Comput* 24(2)
5. Wang K, Hao Y, Cao J (2023) Data layout and scheduling tasks in a meteorological cloud environment. *Intell Autom Soft Comput* 7:1033–1052. <https://doi.org/10.32604/iasc.2023.038036>
6. Madi T, Zhang M, Jarraya Y et al (2024) Apparatus and method for evaluating multiple aspects of the security for virtualized infrastructure in a cloud environment: US17046458 [P]. US20210152572A1, 23 Apr 2024
7. Supreeth S, Patil K, Patil SD et al (2022) An efficient policy-based scheduling and allocation of virtual machines in cloud computing environment. *J Electr Comput Eng*
8. Bhagavathi H, Rathinavelayatham S, Shanmugaiah K et al (2022) Improved beetle swarm optimization algorithm for energy efficient virtual machine consolidation on cloud environment. *Concurr Comput: Pract Exp* 34. <https://doi.org/10.1002/cpe.6828>
9. Sharma O, Saini H (2021) Energy and SLA efficient virtual machine placement in cloud environment using non-dominated sorting genetic algorithm
10. Potluri S, Mohanty SN, Mohanty S (2023) QoS-driven hybrid task scheduling algorithm in a cloud computing environment. *Int J Grid Util Comput* 4:14
11. Agarwal M, Gupta S (2022) An adaptive genetic algorithm-based load balancing-aware task scheduling technique for cloud computing. *Comput Mater Contin (English)*. <https://doi.org/10.32604/cmc.2022.030778>

Diabetic Retinopathy Diagnosis Using Convolutional Neural Networks



Vahini Siruvoru, Nampally Vijay Kumar, Venkata Satya Sai Sagiraju, Jayanth Venavenka, and Jarupula Hrithika

Abstract India has the highest percentage of diabetes cases worldwide, as evidenced by its reputation as the “Diabetes Capital” of the world. Diabetes is a prevalent illness worldwide, including in India. But an increase in glucose levels can also have the opposite effects, such as diabetic neuropathy and retinopathy. We shall handle diabetic retinopathy, also referred to as diabetic eye disease, in this endeavour. Diabetic retinopathy is a condition related to the eye’s fundus that can have a negative impact on the patient if it is not identified. It may also result in a very permanent visual impairment if left untreated in the early stages. The essential point of our undertaking is to foster an application utilizing AI that aides in foreseeing regardless of whether an individual is impacted by diabetic retinopathy.

Keywords Diabetic retinopathy · Convolutional neural networks · Watershed algorithm · Image processing

V. Siruvoru (✉) · N. V. Kumar · V. S. S. Sagiraju (✉) · J. Venavenka · J. Hrithika
Department of Computer Science and Business System, B V Raju Institute of Technology,
Narsapur, Medak, India
e-mail: vahini.s@bvr.it.ac.in

V. S. S. Sagiraju
e-mail: 23211a3245@bvr.it.ac.in

N. V. Kumar
e-mail: vijay.n@bvr.it.ac.in

J. Venavenka
e-mail: 23211a3255@bvr.it.ac.in

J. Hrithika
e-mail: 23211a3223@bvr.it.ac.in

1 Introduction

One of India's largest sectors is healthcare, which has become increasingly successful when illnesses are identified and treated at an early stage. Diabetes is a condition where the body's insufficient production of insulin results in elevated blood sugar levels [1]. Studies have shown that 425 million people worldwide are estimated to have diabetes [2]. In addition to raising blood sugar levels, it also has negative effects on the kidneys, heart, nerves, and retina [1, 2]. The eye condition known as diabetic retinopathy, or DR, is the subject of this endeavour. The diabetic condition known as diabetic retinopathy causes the blood vessels in the retina to expand and leak blood and fluid [3]. Once DR progresses to a severe level, it may result in vision loss. 2.6% of blindness world-wide is brought on by DR. Patients with diabetes who have had their illness for a long time are more inclined to develop DR. Regular retina screening is necessary for diabetics to identify and treat diabetic retinopathy (DR) early on to avoid blindness. Different forms of lesions on a retinal imaging are indicative of diabetic retinopathy (DR).

Our project's goal is to develop an intuitive application that lets users determine whether they have the disease (DR) or not. This project is being built using machine learning and image processing techniques. By providing the system with a retinal image as input, one can quickly and easily identify the ailment with this application. It can precisely and successfully diagnose the disease in less time.

It is evident that India receives the greatest number of cases of diabetes world-wide each year, earning it the title of "Diabetic Capital." Diabetes can cause additional disorders, such as diabetic neuropathy and retinopathy, in a large number of patients. This illness doesn't show any signs and is undetectable until it becomes critical. If not treated in a timely manner, it may also result in permanent blindness. However, regular checkups of the eye can help us identify the early stages of diabetic retinopathy and prevent thousands of people from losing their vision. As of present, DR diagnosis is done manually by a trained doctor using digital fundus pictures. This often results in miscommunications and a delay in outcomes, which eventually causes therapy to be delayed. In order to identify DR early and minimize retinal damage, the objective is to develop an automated, suitable, and advanced method that combines image processing and machine learning. A particular kind of condition known as diabetic retinopathy can affect anyone with diabetes. The development of soft and hard exudates (EX), haemorrhages (HM), and microaneurysms (MA) can all be signs of it. There are four stages of diabetic retinopathy in general.

Below is a brief description of them.

Mild non-proliferative DR: Presence of Microaneurysms only.

Moderate non-proliferative DR: Presence of Microaneurysms and other (example: hard exudates, cotton wool spots dot and blot haemorrhages).

Severe non-proliferative DR: It may contain intraretinal haemorrhages (≥ 20 in each quadrant), definite venous beading in 2 quadrants or in one quadrant, there are microvascular abnormalities.

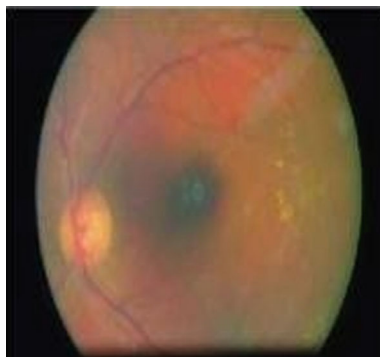
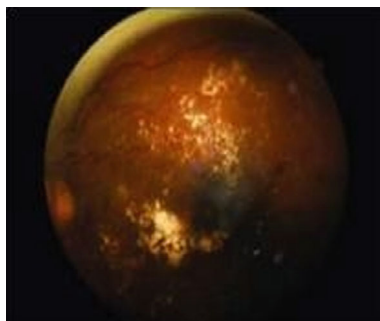
Fig. 1 No DR**Fig. 2** Mild DR

Proliferative DR: Severe non-proliferative DR and 1 or more neovascularisation/vitreous/pre-retinal haemorrhage.

When an eye image (either right or left) is given as an input, the aim is to place the eye status into one of the following groups of eyesight 0 indicates No DR, 1 indicates Mild DR, 2 indicates Moderate DR, and 3 indicates Severe DR and 4 indicates Proliferative DR. So, our aim is to develop an automated analytic system that can identify the type of disease and detect it. The following Figs. 1, 2, 3, 4 and 5 shows the various stages of DR.

2 Literature Survey

In 2016, Carrera created a novel technique for recognizing diabetic retinopathy from input retinal images using digital image processing [4]. For testing, a database of 400 retinal images was utilized. In order to identify exudates and microaneurysms (lesions) in raw retinal images, Rao created a hybrid approach in 2016 that used fuzzy c-means clustering, morphological processing, and a Support Vector Machine classifier [5]. In order to determine whether fundus/retinal image enhancing techniques could be used to diagnose diabetic retinopathy, Yadav conducted a study in 2016 [6].

Fig. 3 Moderate DR**Fig. 4** Severe DR**Fig. 5** Proliferative DR

In 2016, Gupta presented a method for identifying the blood vessels in the eye [7]. This was made possible by the application of Gaussian filtering and morphological administrator. A method for automatically detecting and verifying haemorrhages and microaneurysms for a given colour fundus image was proposed by Seoud in 2016 [8]. Karami presented a dictionary learning (DL)-based method in 2017. The best atomic representation of fundus images and dictionaries previously learned using the KSVD technique were used in its creation [9].

3 Existing System

In early days, Diabetic retinopathy was identified manually by examining alterations in retinal pictures, specifically morphological changes in microaneurysms, exudates, haemorrhages, and blood vessels. The different existing methods are:

i. **Ophthalmoscopy:**

In this method a light source and a viewing lens are used to examine the optic nerve, blood vessels and other structures of retina for signs of diabetic retinopathy.

ii. **Fluorescein Angiography:**

It involves injecting a fluorescent dye into the bloodstream and capturing images of the retina as the dye circulates. This method helps identify abnormalities in blood vessel structure and blood flow.

iii. **Early Fundus Photography:**

This method involved film-based photography, capturing images of the back of the eye for examination. It has been a crucial tool in early diagnosis of DR.

iv. **Clinical Examination:**

Manual inspection of visual activity, intraocular pressure, and the presence of characteristic signs such as microaneurysms, haemorrhages, and exudates.

v. **Patient History and Symptom Reporting:**

Gathering information about a patient's medical history, particularly, their history of diabetes and associated symptoms.

vi. **Machine Learning and Image Processing:**

Machine learning and image processing techniques like support vector machines, decision trees are also used which work on feature extraction using image processing. There is a lot of scope for enhancing and improving these models to give more accurate and precise results.

3.1 *Disadvantages of Existing System*

- The manual interpretation of signs and symptoms varied between practitioners leading to inconsistencies in diagnosis.
- Limited imaging techniques hindered the ability to detect subtle changes in retina.
- The systems lacked robust quantitative tools making it challenging to track changes over time or compare results.
- These systems were less effective in detecting subtle signs, leading to delayed diagnosis and intervention, potentially allowing the disease to advance to more severe stages.

- The use of support vector machines in modern systems is not without drawbacks. SVMs are sensitive to noisy data, which might lead to suboptimal performance if the dataset contains outliers or irrelevant information.
- Training an SVM on large datasets can be computationally expensive and time-consuming. SVMs are not capable of handling big data scenarios.
- Decision tree systems also had several drawbacks. Decision trees are prone to overfitting leading to poor generalization on unseen data.
- Decision trees are highly sensitive to small variations in data, making the model less stable.

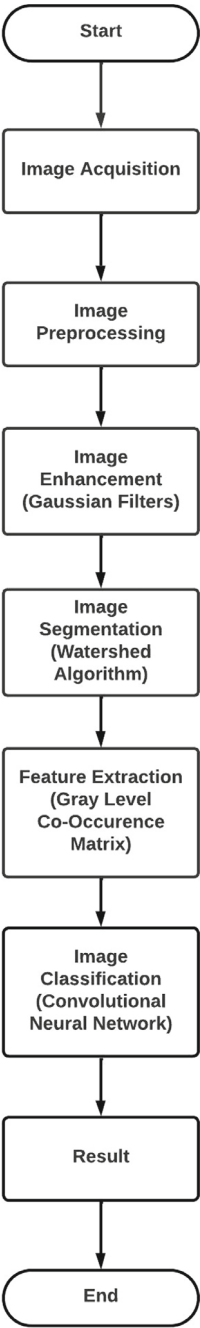
4 Proposed System

In the proposed approach, early disease detection is achieved through the use of machine learning and image processing techniques. In order for the patient to receive the necessary treatment at the appropriate time and prevent the disease's symptoms from becoming severe. This procedure is more rapid and effective than manual inspection. Our project's main goal is to create a model that aids in the prediction of diabetic retinopathy by utilizing machine learning and image processing techniques. There are five main steps in this entire process and shown in Fig. 6.

- The first step is image acquisition, in which we feed our model a retinal image as an input.
- In second stage, we need to remove the noise and enhance the image, this can be achieved using Gaussian filters which is mainly used to remove the noise, later the image must be converted to a grey scale image. It is a preprocessing stage.
- Third stage is image segmentation, it is one of the major steps in building a model, at this stage we will segment (or divide) the entire image into several parts based on their similarity constraints. In order to perform this, we are using "watershed" segmentation algorithm, which is based on topological interpretations.
- This segmented image is then sent to GLCM (grey level co-occurrence matrix) for feature extraction, it is one of the best methods to extract texture features (statistical). The classification stage is the last and most critical step in the entire procedure.
- The features extracted in previous stage are feed to 3 the classifier for classification, which gives us the results whether the given image belongs to which category (defected or not defected). For classification, we will be using convolutional neural networks, a supervised machine learning algorithm that uses algorithms and classifies the problem into 5 different groups.

These 5 steps are performed on every image to build a trained model. Later, this trained model is used for testing other images. This application can be used by anyone who can afford fundus camera and it also helps the ophthalmologist in effective and accurate detection of disease.

Fig. 6 Flowchart



5 Implementation

We used the “DIARETDB0” dataset, which included 130 photos intended for testing, to implement this project. Additionally, we have made use of Torch Vision, which offers helpful training datasets. This project uses the Watershed algorithm for segmentation and Convolutional Neural Networks for classification as its two primary algorithms (Fig. 7).

5.1 Dataset

Only a small percentage of the 513 colour fundus photos in the present database are benign, and the majority display symptoms of diabetic retinopathy, such as hard and soft exudates, microaneurysms, neovascularization, and haemorrhages. The photos were taken using an unnamed digital fundus camera with a 50° field of vision. Based on real-world occurrences, the data can be utilized to assess the general efficacy of diagnostic methods.

5.2 Watershed

One common segmentation technique used to divide up different items in an image is the watershed. It appears on an image in grayscale.

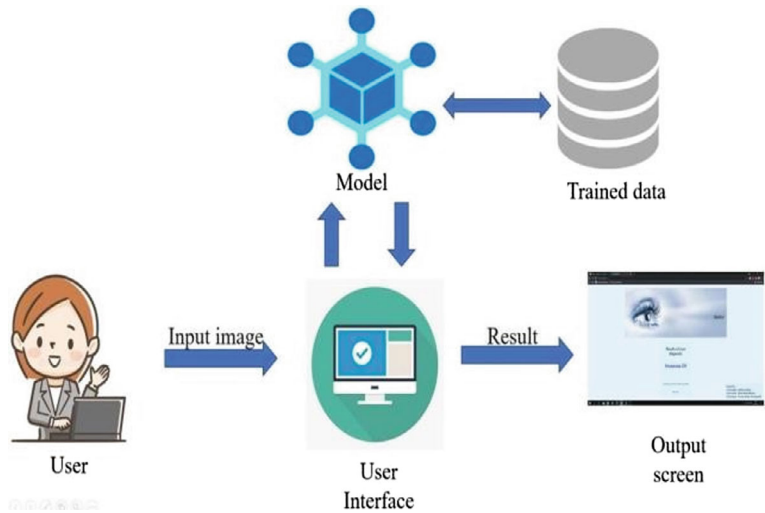


Fig. 7 Working model

Algorithm Watershed

- i. Start by computing a segmentation function. The elements you are trying to segment are the black areas in this picture.
- ii. Calculate the foreground markers, within each of the objects, there are connected blobs of pixels.
- iii. List the backdrop markers that you need. These pixels have no connections to anything. Make sure the segmentation function only has minima at the positions of the background and foreground markers.
- iv. Calculate the modified segmentation function's watershed transform.

5.3 Convolutional Neural Networks

The process of classifying a new observation into one of several categories (subpopulations) is known as classification in statistics and machine learning. A training set of data, comprising observations (or instances) whose category membership is known, is used to make this determination.

Neural Network: A neural network is a collection of algorithms that tries to identify underlying links in a set of data by simulating how the human brain functions. Neural networks are systems of neurons that can be created naturally or artificially in this context.

Algorithm Neural Network

- i. Randomize the weight initialization.
- ii. If the error is significant.
- iii. For each training pattern, give the ANN the input texture characteristics.
- iv. Calculate the output of the neurons in the ANN using the activation function, the sum of the input weights, and the bias values. Any training pattern generates the activation function at the n th neuron.
- v. Calculate the error at the output neuron using the equation and Mean Squared Deviation (MSD). Calculate the weights of the connections for the pre-output layers and the error signals based on the output error.
- vi. Change the weights till the mistake is too small.
- vii. Use the testing results.

Convolutional neural networks are a kind of deep neural network used in deep learning. They are also referred to as Space Invariant or Shift Invariant Artificial Neural Networks (SIANN). There are several hidden layers in between the input and the output layers. Every network node's weight is determined using a shared weight architecture. It is entirely related and derived from biological processes found in the human brain. The steps involved in CNN are shown in Fig. 8

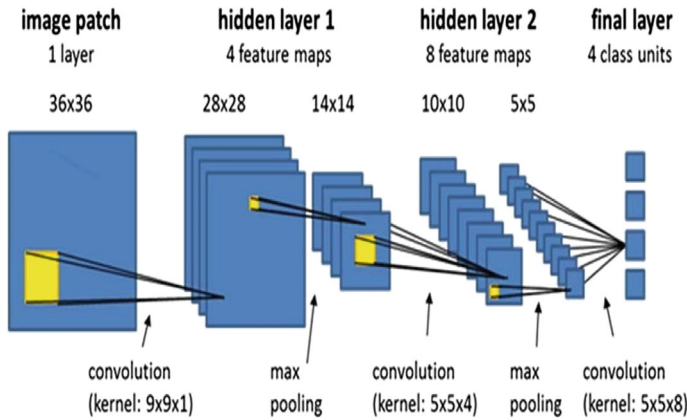


Fig. 8 CNN process

A. Algorithm CNN

Step 1: Convolution operation abstracts the picture into a feature map. A linear functionality is produced by the rectified linear unit (ReLU) layer.

Step 2: Pooling—Convolutional networks may include local or global pooling layers to expedite the underlying calculation. Pooling layers reduce the dimensionality of the data by combining the outputs of neuron clusters in one layer into a single neuron in the subsequent layer.

Step 3: Flattening—Layers are moved from pooled to flattened.

Step 4: Full connection—Fully connected layers link every neuron in one layer to every other neuron in the layer below.

6 Results

Open the command prompt and navigate to the project folder in order to start running this project. The main application “app.py” should then be run. Next, access your system’s local host by typing “<http://127.0.0.1:5000/>”.

Your application’s or project’s home page appears similar to the one in Fig. 9. The input image selection and processing options are located on the main page. Clicking the “choose file” button opens a new pop-up window. Choose the image you wish to test. After choosing to upload, use the “process” option to process the image. The image will be processed once you click the “process” button, and the results will appear as seen in Fig. 10.

The user is prompted to provide input in one of the specified formats if the image is not in one of the supported formats (jpg, jpeg, or png).

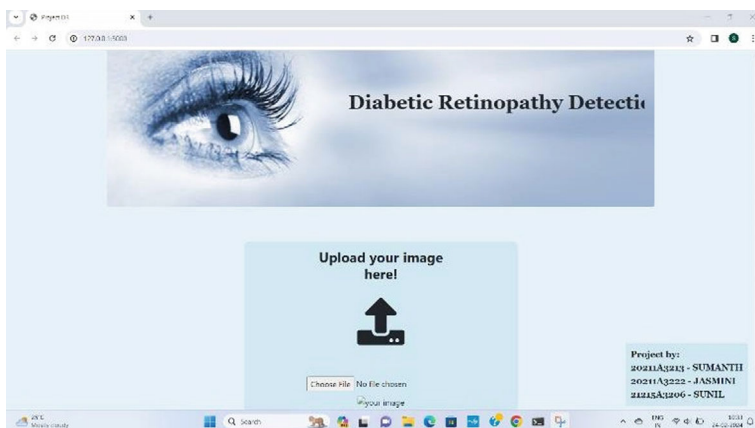


Fig. 9 Home page

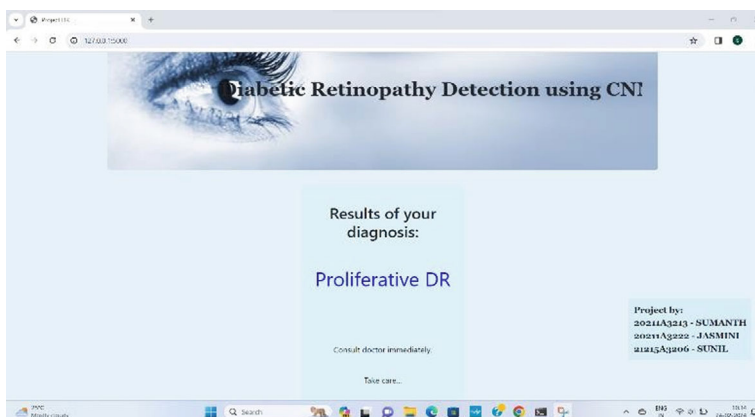


Fig. 10 Result page

7 Conclusion and Future Scope

CNN and a segmentation algorithm based on water shed were utilized to build the model. With precise results, this project facilitates the user's easy detection of diabetic retinopathy. It is also useful for ophthalmologists to diagnose the medical condition. Based on the presence of hard and soft exudates, tiny spots, leaky blood vessels, and micro aneurysms in the retinal image, the input image is classified. A variety of input images were used to test the application, and each image was categorized into one of the five stages.

This project can be implemented on edge devices.

- Prediction of biological diseases based on retinal images, including diabetic retinopathy with more than five stages.
- Separate learning and differential privacy to protect system and user privacy.
- Meta Learning and Weak Supervision techniques can be utilized to apply this to a larger group of images.

References

1. Taylor R, Batey D (2012) Handbook of retinal screening in diabetes: diagnosis and management, 2nd edn. Wiley
2. International diabetes federation—what is diabetes [Online]. Available <https://www.idf.org/aboutdiabetes/what-is-diabetes.html>
3. American academy of ophthalmology—what is diabetic retinopathy? [Online]. Available <https://www.aaof.org/eyehealth/diseases/what-is-diabetic-retinopathy>
4. Carrera EV, Gonzalez A, Carrera R (2017) Automated detection of diabetic retinopathy using SVM. In: 2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON)
5. Harini R, Rao SN (2016) Feature extraction and classification of retinal images for automated detection of diabetic retinopathy. In: 2016 Second international conference on cognitive computing and information processing (CCIP)
6. Yadav SK, Kumar S, Kumar B, Gupta R (2016) Comparative analysis of fundus image enhancement in detection of diabetic retinopathy. In: 2016 IEEE region 10 humanitarian technology conference (R10-HTC)
7. Mohana Guru Sai Gupta V, Surya Teja D, Gupta S, Sengar P (2016) Extraction of blood veins from the fundus image to detect diabetic retinopathy. In: 1st IEEE international conference on power electronics, intelligent control and energy systems (ICPEICES—2016)
8. Seoud L, Hurtut T, Chelbi J, Cheriet F, Langlois JMP (2016) Red lesion detection using dynamic shape features for diabetic retinopathy screening. *IEEE Trans Med Imaging* 35(4)
9. Karami N, Rabbani H (2017) A dictionary learning based method for detection of diabetic retinopathy in color fundus images. In: 10th Iranian conference on machine vision and image processing

Wearable Technology, A Smart Step Towards Healthcare—A Review



Niharika Nagpal, Anshika Sharma, and Divya Agarwal

Abstract With our growing population and rising health concerns, wearable gadgets are like superheroes in disguise. Technology adoption in healthcare sector is to predict or diagnose various diseases. This review aims to investigate various aspects of wearable device technology used in healthcare. Wearables can be stylish wristbands or comfy neckpieces that track our heart rate, sleep patterns, and even warn us of potential problems. Currently, the healthcare sector is facing certain challenges such as high count of ageing population, chronic diseases, erratic diet issues, cardiac problems, etc. As the data generated through digital devices can be analyzed to predict and warn the concerned person and improve the health, it can be a breakthrough towards this lacuna in healthcare. Based on the research work done it can be elucidated that wearables are the future of healthcare sector which can aid in diagnosis as well as prognostics. This work reviews the role of various wearable devices in early diagnosis and treatment and shows a systematic review of the wearable devices and how they can act as a boon in healthcare industry. The purpose of this work is to spread awareness about the wearable technology among the people and to contribute towards the outset of smart city.

Keywords Wearable devices technology · Healthcare · Smart city

N. Nagpal · A. Sharma · D. Agarwal (✉)
School of Engineering & Technology, Vivekananda Institute of Professional Studies - Technical
Campus, AU- Block (Outer Ring Road) Pitampura, New Delhi 110034, India
e-mail: divya.agarwal@vips.edu

N. Nagpal
e-mail: 08217711921_ds@vips.edu

A. Sharma
e-mail: 09517711921_ds@vips.edu

1 Introduction

Imagine tiny health guardians strapped to your wrist or chest, constantly monitoring your heart rate, sleep, and activity. Wearables are revolutionizing healthcare by offering real-time health data. Using sensors and software, wearables track vitals like smartwatches, empowering to manage health and well-being. Wearables go beyond fitness trackers, by creating a virtual comforting environment using gaming and VR elements to ease patient’s experience. Vast collected health data is analyzed by scientists to predict, diagnose, and even treat diseases like heart disease and cancer. Relation of data science and wearables is shown in Fig. 1.

Future of wearables is bright. By 2027, they’re expected to slash hospital costs by 16%, potentially saving billions. This paper explores how wearables can transform healthcare, enabling diagnoses and reports from the comfort of your home. Furthermore, this paper gives a systematic review of the wearable devices and how they can act as a boon in healthcare industry by spreading awareness amongst people and to contribute towards outset of smart city.

Figure 2 shows where the wearables can be utilized. In today’s busy lifestyle, everyone wants to monitor their health most comfortably and conveniently and wearable technology is the way one can have easy access to his health. Wearable technology is the new way to make ourselves fit, healthy, and future-ready (Fig. 3 shows different types of wearables) (Table 1).

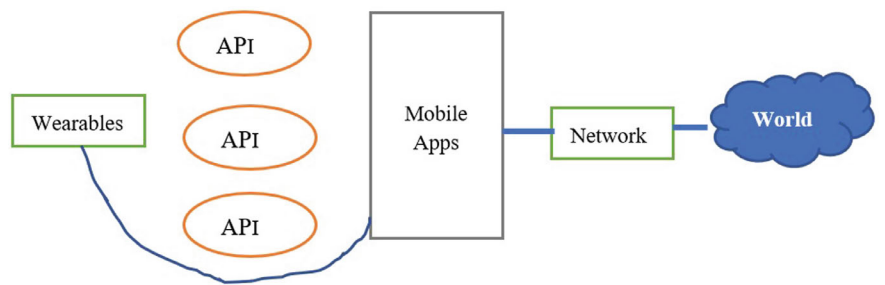


Fig. 1 Wearables and data science



Fig. 2 Need of wearable technology

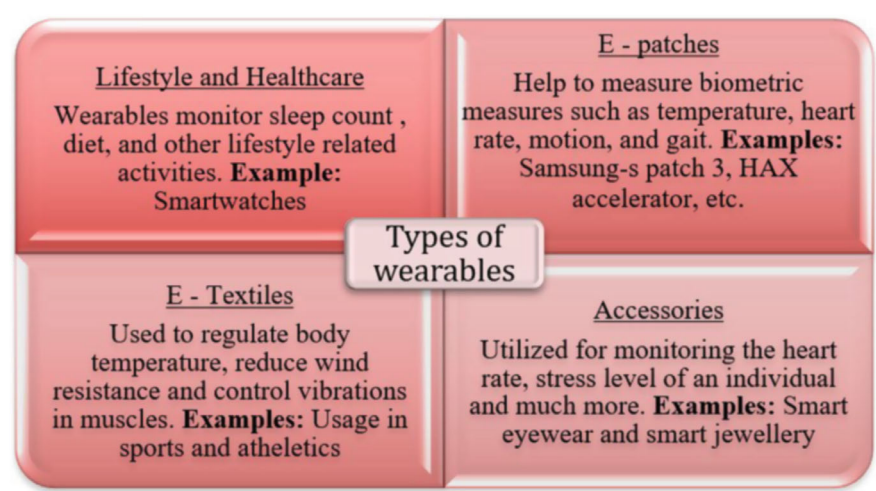


Fig. 3 Types of wearables [1]

Table 1 ABCD of wearable technology [2–13]

A–Z	Terminology	Explanation
A	Algorithm	Data generated is analyzed using structured algorithms
B	Big Data	Data generated is unstructured, varied, voluminous (Big Data)
C	Cloud computing	Being part of IoT provide fast data transmission and massive storage encompassing cloud
D	Design	Interactive and human centric design; novel reflective themes
E	Efficiency	Affected with battery life (prolonged use) and memory capacity
F	Fusion	More reliable pulse rate indication is achieved by signal fusion from electrocardiogram and pulse oximetry
G	GDPR	General data protection regulation replaced data protection directive, aiming to merge data protection as well as addressing exports of personal data
H	Hardware	Compact hardware and maximum functionality
I	Instrumented	Approach replaces manual observation with standardized approach
J	Jewellery	Smart jewellery—used to monitor individuals health
K	Kits	Developer kits are useful to increase technical insight
L	Low cost	Offers numerous sensing and self-monitoring capability at low cost
M	Machine learning	ML algorithms can be applied on data generated for classifying motor activities, for measuring upper limb rehabilitation etc
N	Nursing	Wearables help frontline hospital staff to interpret data
O	Open source	OSI and FOSS have proliferated Wearable devices research and development with languages like python
P	Pets	Pet owners are able to track their pet's emotional state patterns and check their health records on a daily basis
Q	Quality of life	Wearable devices help each individual to achieve a healthy lifestyle by keeping a track of their lifestyle
R	Reliability	Data generated must be validated to ensure its reliability over time
S	Standardization	Technical Standards allow wearables to connect through communication protocols that are bluetooth, wifi, etc
T	Terminology	Most important when considering file formats, naming conventions or metadata where no common terminology exists
U	Ubiquitous	Ubiquitous sensing technologies relay technical information between technologies to harmonize data streams
V	Validation	Wearables remain novel in healthcare sector where regulations require the experts to validate and verify
W	Wearables	Can be stylish wristbands or comfy neckpieces to track health data
X	X marks the spot	Correct wearable devices placement is given careful consideration to ensure robust data collection
Y	Yourself	Wearables gather personal data and identify behavioural patterns resulting in self-monitoring and lifelogging
Z	(generation) Z	Wearables are normal part of life for generation Z as it is easy, simple and comfortable

1.1 Vital Signs to Be Monitored

Humans have various physiological signs that can be measured from various electrical sensors, which use various algorithms using huge amounts of health data. There are many data mining processes concerning different aspects of wearable sensing in wearable health devices as shown in the Fig. 3 [2–4, 14–16]

1. **Prediction**—Identifies events which have not happened yet. Getting more interest by healthcare experts as it helps to prevent further chronic problems [2–5, 14]
2. **Outlier detection (unusual patterns/random errors)**—Detected unusual patterns in health records, enables health experts to make quick decisions [5, 6, 14]
3. **Diagnosis assistance (decision-making)**—Essential for human health monitoring based on vital signs analysis using outlier detection techniques [14].

Figure 4 reveals two key factors influencing how wearable sensor data is analyzed. It depends on where it's used (home vs. clinic) and the wearer's health. Home monitoring aims to predict issues, while clinics use it for diagnosis with additional data [17]. For those with health conditions, both diagnosis and critical event alerts are important [17].

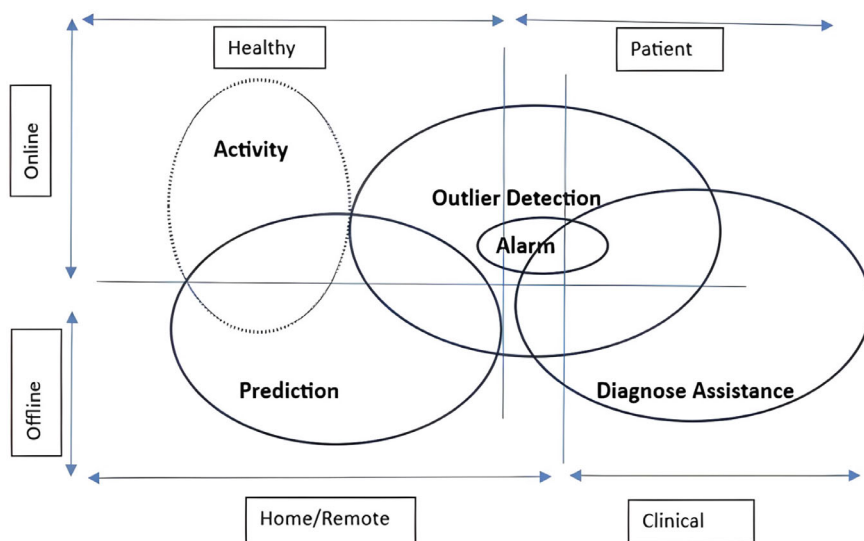


Fig. 4 Data mining process in relation to wearable sensing

2 Literature Review

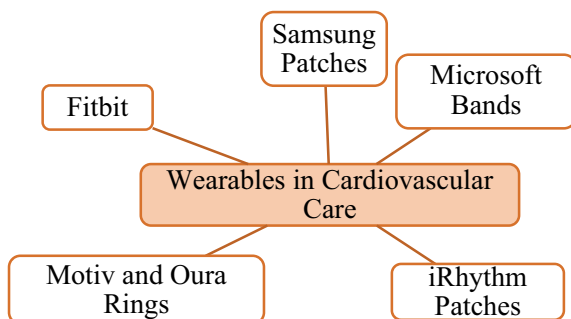
The term “wearable technology” describes the gadgets that people put on their bodies to track their health information. Their body is in contact with this equipment in order to gather their medical history. It comes into contact with a person’s skin through the use of sensors, actuators, software, and electronic patches in order to monitor their health. They can be used in various healthcare fields as shown in Fig. 5.

Wearable tech shows promise in managing heart issues. Studies using special sensors, like CardioMEMS, have shown a significant decrease in heart failure hospitalizations [18]. The clinicians had access to the readings of pulmonary artery pressure. As a result, there was a 39% reduction in heart failure-related hospitalization. One of the largest trials was the BEAT-HF study. Larger studies with better adherence rates are needed to confirm the true value of wearables in heart failure management [19]. MiCORE study was implemented on 200 patients with type 1 MI (injury with plaque rupture or erosion). A mobile application integrated with an Apple watch and BP cuff was used. It resulted in a 43% lower readmission of patients in 30 days. It also helped in the reduction of costs. These random trials can be very beneficial for the patients. It helps in the reduction of hospitalization as well as the reduction in cost and helps in early diagnosis of the disease which can save several lives [6–11, 19].

Nowadays, screening for hypertension in young adulthood is highly recommended to prevent heart-related diseases. Due to hectic and tense lifestyle of youngsters, heart diseases are becoming common in them. Cuff-less wearables that can be easily worn on the wrist, can very convenient and comfortable, however, dedicated study is required which could be beneficial. Wearables are to be worn all the time, in order to monitor blood pressure while sleeping and exercising, which aids in detecting cardiac arrest and other heart-related issues. [19] Big data allows data science to unlock new knowledge for personalized mental healthcare [10–13, 15].

Data science offers faster depression diagnosis with algorithms, but model development needs caution to avoid misdiagnosis. Wearables and apps provide continuous monitoring and valuable data for treatment, but ethical considerations like data privacy are essential for responsible research in this promising field [15] (Fig. 6 and Table 2) [12, 13].

Fig. 5 Wearable devices used in cardiovascular care



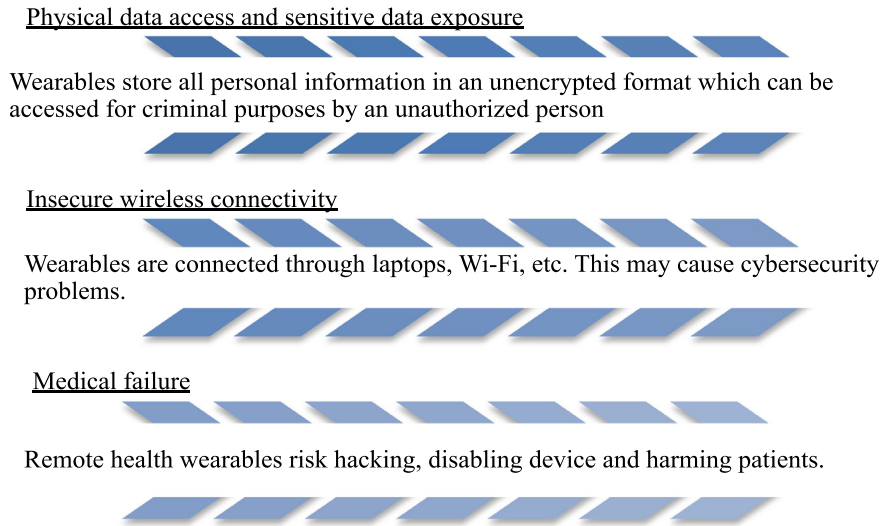


Fig. 6 Challenges in wearable technology

Table 2 Commonly used sensors in wearables [7–13, 19]

Name	Purpose
Activity sensors	• Internal reporting of all physical activities will become obsolete • Triaxial accelerometer—popular method to monitor linear acceleration • Gyroscope—monitors angular motion
Heart rate and rhythm sensors	• Uses various algorithms and calculates beats of heart in every time interval • Chest straps and ECG patches help to monitor heart rhythm, but they are a little less comfortable to consumers • PPG monitors variation in microvascular blood volume (eg, Apple watch [19] and Fitbit charge [16])
Blood pressure sensor	• Hypertension leads to an increase in blood pressure in humans, which needs to be monitored and controlled • Hearguide wrist watch uses a standard blood pressure measurement device
Biochemic al sensor	• Can measure body fluids and electrolytes in the body with the help of biochemical transducers • Provides important information about plasma volume status and analyte concentration

3 Results and Discussion

The trend, towards health consciousness has led to a surge in the popularity of technology. These gadgets are easy to wear reasonably priced and have the potential to positively impact your routine. They monitor health metrics like heart rate and steps helping you stay more attuned to your well-being. The data collected by wearables can be conveniently stored on your smartphone. Wearable devices utilize sensors and advanced algorithms to interpret data and offer insights. Safeguarding this wealth of health information is crucial as any breaches could have repercussions. Since wearables rely on internet connectivity, they may be susceptible to cyber threats. Moreover, these devices are accessible to anyone raising the possibility of misinterpreting data with risks involved. Despite these obstacles wearable tech offers a sneak peek into the future of healthcare.

4 Conclusion

The amount of data collected per day plays an important role in the deployment of wearable technology. The Internet of Things is growing at a rapid speed nowadays and wearables are a part of this. So, the design of wearables should be done according to the convenience of all the people. The only difference is the type of sensors being used according to the type of disease. It uses personal data to monitor the health of people and make them aware of their poor lifestyle and how they can improve it. After the Covid-19 pandemic, everyone is becoming health conscious which results in more demand for wearable technology as it provides the human to diagnose themselves instead of going to health experts again and again. They will get more significance in the future and more and more people will be aware of the benefits of wearables. Sensors collect the data and are helpful in analysis of data and give significant results. The Internet of Things is growing at a rapid pace and wearables are a part of this. People are more concerned about their health and well-being, that is why, the demand of wearable technology is at peak nowadays.

5 Future Scope

Wearables collect a treasure trove of personal data, raising privacy and security concerns. Additionally, incompatibility between devices hinders their full potential. Future advancements should focus on robust data protection, seamless interoperability, and sustainable energy solutions. By addressing these challenges, we can unlock the true potential of wearables and propel their evolution.

References

1. Vijayalakshmi K, Uma S, Bhuvanya R, Suresh A (2018) A demand for wearable devices in healthcare. *Int J Eng Technol* 1–4
2. Jin D, Adams H, Cocco A, George Martin W (2019) Smartphones and wearable technology: benefits and concerns in cardiology. *Med J Aust* 1–5
3. Mansoor Baig M, GholamHosseini H, Moqem AA, Mirza F, Lindén M (2017) A systematic review of wearable patient monitoring systems—current challenges and opportunities for clinical adoption. *J Med Syst* 1–9
4. Pevnick MJ, Birkeland K, Zimmer R, Elad Y, Kedan I (2017) Wearable technology for cardiology: an update and framework for the future. *Trends Cardiovasc Med* 1–25
5. Ash G, Stults-Kolehmainen M, Busa MA, Gregory R, Ewing Garber C, Liu J, Gerstein M, Antonio Casajus J, Gonzalez-Aguero A, Constantinou D, Geistlinger M, Guppy FM, Pigozzi F, Pitsiladis YP (2020) Establishing a global standard for wearable devices in sport and fitness: perspectives from the New England chapter of the American college of sports medicine members. *Curr Sports Med Rep* 1–5
6. Russ TC, Woelbert E, Davis KAS, Hafferty JD, Ibrahim Z, Inkster B, John A, Lee W, Maxwell M, McIntosh AM, Stewart R (2018) The MQ data science group, how data science can advance mental health research. *Nat Hum Behav* 1–9
7. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, Rickman AD, Wahed AS, Belle SH (2016) Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the idea randomized clinical trial. *JAMA* 1–11
8. Zadeh A, Taylor D, Bertso M, Tillman T, Nosoudi N, Bruce S (2021) Predicting sports injuries with wearable technology and data analytics. *Inf Syst Front* 1–16
9. Rashid A, Hasan O (2018) Wearable technology for hands joints rehabilitation: a survey 1–38
10. Anthony Hickey B, Chalmers T, Newton P, Lin C, Sibbritt D, McLachlan CS, Clifton-Bligh R, Morley J, Lal S (2021) Smart devices and wearable technologies to detect and monitor mental health conditions and stress: a systematic review. *Sensors (Basel)* 1–17
11. Najafi B, Horn D, Marclay S, Crews RT, Wu S, Wrobel JS (2010) Assessing postural control and postural control strategy in diabetes patients using innovative and wearable technology. *J Diabetes Sci Technol* 1–12
12. Sequeira L, Perrotta S, LaGrassa J, Merikangas K, Kreindler D, Kundurf D, Courtneya D, Szatmaria P, Battagliaa M, Strauss J (2019) Mobile and wearable technology for monitoring depressive symptoms in children and adolescents: a scoping review. *J Affect Disord* 1–11
13. Hänsel K, Wilde N, Haddadi H, Alomainy A (2015) Challenges with current wearable technology in monitoring health data and providing positive behavioural support. In: 5th EAI international conference on wireless mobile communication and healthcare—transforming healthcare through innovations in mobile and wireless technologies, pp 1–4
14. Dias D, Paulo Silva Cunha J (2018) Wearable health devices-vital sign monitoring, systems and technologies. *Sensors (Basel)* 1–28
15. Russ TC, Woelbert E, Davis EAS, Hafferty JD, Ibrahim Z, Inkster B, John A, Lee W, Maxwell M, McIntosh MM, Stewart R and the MQ Data Science group (2019) How data science can advance mental health research. *Nat Hum Behav* 1–9
16. Godfrey A, Hetherington V, Shum H, Bonato P, Lovell NH, Stuart S (2018) From A to Z: wearable technology explained. *Maturitas* 1–8
17. Banaee H, Ahmed MU, Loutfi A (2013) Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. *Sensors (Basel)* 1–29
18. Nafai B, Wrobel J (2011) Laboratory in a box: wearable sensors and its advantages for gait analysis. *Annu Int Conf IEEE Eng Med Biol Soc* 1–4
19. Bayoumy K, Gaber M, Elshafeey A, Mhaimeed O, Dineen EH, Marvel FA, Martin SS, Muse ED, Turakhia MP, Tarakji KG, Elshazly MB (2021) Smart wearable devices in cardiovascular care: where we are and how to move forward. *Nat Rev Cardiol* 1–19

Revolutionizing Food Supply Chains: An AI and ML-Driven Model for Enhanced Quality Control and Traceability



Arjun Kumar Chandrasekar, G. Anitha, and Vigneswaran Narayanamurthy

Abstract The paper provides an innovative model that modernizes food product quality control and traceability by integrating cutting-edge technologies like blockchain, artificial intelligence (AI), machine learning (ML), and the Internet of Things (IoT). Every product has a unique QR code created during the manufacturing process, which enables smooth tracking of the product by all supply chain participants, including suppliers, dealers, manufacturers, customers, and other stakeholders. IoT sensors gather extensive product data, which is then sent over the MQTT protocol and examined by Google Vision Inspection AI and a customized TensorFlow deep learning model. To guarantee accurate and reliable quality control, the TensorFlow model is rigorously trained using the VGG19 architecture on the “PepsiCo-Lab-Potato-Quality-Control” dataset. It is then further assessed against a variety of machine learning models. The paper provides a full explanation of the theoretical framework and the possible applications of the system. It also describes how blockchain technology, as used by the Ethereum Foundation’s smart contracts, keeps an unchangeable record of data, improving the supply chain’s security and transparency. Through a straightforward QR code scan, the model enables customers to evaluate the quality and history of the food they have purchased, guaranteeing a high degree of quality assurance and boosting confidence in the food supply chain. With its goal of establishing a new benchmark for the sector, the suggested model highlights the revolutionary power of AI and ML in attaining hitherto unheard-of levels of traceability and quality control in food supply chains.

Arjun Kumar Chandrasekar, Vigneswaran Narayanamurthy are contributed equally to this work.

A. K. Chandrasekar · G. Anitha (✉)

Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai, Tamilnadu, India

e-mail: g_anitha@ch.amrita.edu

V. Narayanamurthy

Department of Engineering Technology, Faculty of Electronics and Computer Technology Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

e-mail: vigneswaran@utem.edu.my

Keywords Food chain traceability · AI-assisted quality control · QR code product tracking · TensorFlow · VGG19 · Machine learning · Ethereum · Smart contract · IoT sensors · MQTT protocol

1 Introduction

Global food supply chains are undergoing a change because to the combination of blockchain technology and artificial intelligence (AI), which improves transparency and traceability while solving issues with food safety, authenticity, and sustainability. Significant gains have been made in measures like accuracy, precision, recall, and *F1*-score as a result of this novel combination. The system delivers unparalleled levels of traceability and transparency from farm to customer by utilizing AI-driven quality control processes, human data entry, QR code tracking, and IoT sensor data collecting via the MQTT protocol. Quality control and logistics are optimized by AI, while record-keeping is made safe and unchangeable by Blockchain, which builds stakeholder confidence. The hyper-traceability measurements are easier to grasp and put into practice when the supply chain is seen as a directed graph with structure. Consensus on the blockchain ensures the validity of blocks, while QR code recognition makes thorough product monitoring easier. Real-time data collection and transmission enabled by the MQTT protocol and Internet of Things (IoT) enhances monitoring and decision-making. By evaluating sensor data, foreseeing disruptions, and automating inspections with deep learning models, AI-driven quality control guarantees the integrity of the product. Customers and supply chain confidence are raised by blockchain transparency features, which provide all parties with a complete view of previous transactions. Throughout the food supply chain, this integrated approach not only boosts operational effectiveness and risk management but also enhances consumer satisfaction, transparency, and dependability, creating opportunities for a more resilient and flexible industry.

2 Literature Review

Food supply chain management has advanced significantly in recent years due to a focus on environmental sustainability, food safety, authenticity, traceability, and quality control [1–5]. The industry has witnessed innovations such as the combination of artificial intelligence (AI) and blockchain technology, which offer safe, decentralized record-keeping and dynamic insights that have revolutionized conventional supply chain models [2, 4, 5]. Precise monitoring, risk assessment, and data-driven decision-making are made possible by these technologies [2, 4]. Moreover, hazard-free, organic, QS/SC-food quality, green, and certification programs greatly increase consumer confidence in food safety [5].

The use of artificial intelligence (AI) and machine learning (ML) is spreading to a number of industries, including aircraft spare parts, food supply chains, steel manufacturing, and health care [2, 6–9]. These technologies improve quality control, traceability, and decision-making. They employ RFID tags, QR codes, and AI-driven inspections to solve challenges related to food safety, quality assurance, and transparency [6]. In food supply chains, machine learning (ML) models [10, 11] improve inventory management, cold chain operations, demand forecasting, and temperature control [4]. Blockchain, artificial intelligence, and digital twins coming together in the food supply chains might significantly improve quality control, increasing output, sustainability, and industry performance as a whole [2, 4].

3 Proposed Methodology

In addition to using AI models like Google Vision Inspection AI for quality assurance and anomaly detection, the system combines IoT sensors with the MQTT protocol to collect environmental data in real time. Smart contracts and the Ethereum blockchain provide visible, unchangeable tracking, improving decision-making, traceability, and consumer confidence in the supply chain (Fig. 1).

3.1 Manufacturing Stage

By combining IoT sensors, weight and gas sensors, network devices, high-resolution cameras, QR codes, and other technologies, the hyper-traceability system ensures accurate product identification, continuous monitoring, and quality control. With

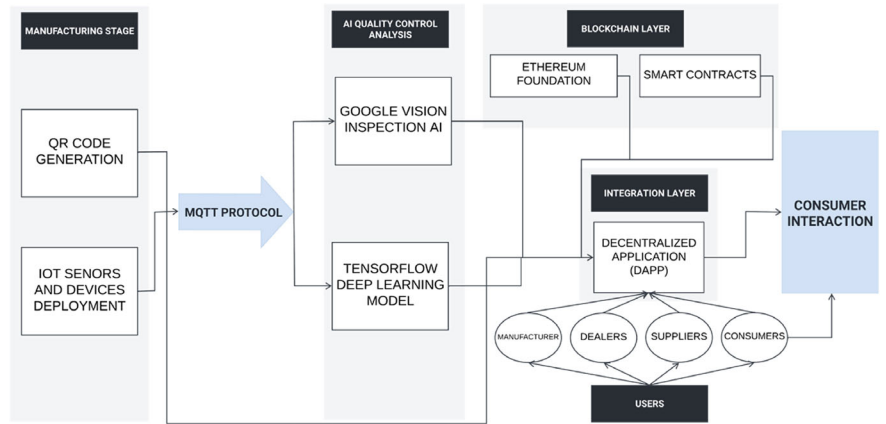


Fig. 1 Methodology flowchart

sensors verifying product standards and safety and QR codes offering digital fingerprints, this configuration facilitates thorough inspection and environmental monitoring. Through the maintenance of product integrity, quality, and transparency, this integrated strategy increases customer trust at every stage of the supply chain.

3.2 Data Transmission Using MQTT Protocol

The TCP/IP interoperability and low bandwidth needs of the MQTT protocol make it ideal for Internet of Things situations, especially when it comes to real-time supply chain monitoring. IoT performance is improved by its publish-subscribe design, which enables scalable, effective data transmission, and asynchronous communication. Three Quality of Service levels are supported by MQTT, along with a lightweight architecture and data integrity and confidentiality capabilities including message integrity checks, authentication, ACLs, and TLS.

The equation for MQTT data transmission efficiency can be expressed as:

$$E_{\text{MQTT}} = \frac{P}{P + R} \times 100\%$$

where

E_{MQTT} MQTT efficiency
 P Number of successfully transmitted packets
 R Number of lost or retransmitted packets

3.3 AI Quality Control Analysis

TensorFlow with Google Vision Inspection AI-powered AI quality control analysis is essential to preserving product integrity across the supply chain. By processing data from IoT sensors and industrial processes, it may identify patterns, anticipate problems, and guarantee strict quality standards—all of which promote consumer pleasure and confidence.

3.4 Blockchain Layer

Enhancing security, trust, and transparency in the food supply chain is the hyper-traceability system's blockchain layer, which makes use of Ethereum platforms. Stakeholder confidence and accountability are fostered by process automation, increased efficiency, and authenticity guaranteed by smart contracts and decentralized ledgers.

3.5 *Integration Layer*

As a Decentralized Application (DApp), the Integration Layer in the hyper-traceability system makes sure that data flows across IoT devices, AI analytics, blockchain transactions, and user interfaces. It makes use of blockchain's immutability to provide data integrity, safe transactions, instantaneous insights, and enhanced supply chain efficiency—all of which increase stakeholder trust and transparency.

3.6 *Users*

Key players in the hyper-traceability system include merchants, distributors, manufacturers, and end users, all of whom play crucial roles. Retailers employ product data for sales management, manufacturers handle quality control and technology integration, distributors handle logistics, and customers access information. Collaboration via DApps, AI, and blockchain guarantees data integrity and a supply chain environment that is focused on the needs of the client.

3.7 *Consumer Interactions*

With the use of QR code scanning, real-time tracking, and comprehensive product information, the hyper-traceability system engages customers and lets them follow the path of a product. Transparency and accountability are promoted by this contact, which gives customers the authority to engage in food safety and sustainability projects, offer feedback, and make educated decisions.

4 *Result and Analysis*

4.1 *AI Food Quality Control Model Comparison*

An extensive comparison of the different machine learning (ML) and deep learning (DL) models utilized in AI-driven food quality monitoring systems across the supply chain is given in Table 1. These models are crucial for evaluating food safety, authenticity, and quality, greatly raising industry requirements. They especially use VGG19 for feature extraction.

Dataset: The “PepsiCo-Lab-Potato-Quality-Control” GitHub project provides a dataset for deep learning models focused on Frito-Lay potato chip quality control, featuring around one thousand images labeled as “defective” or “non-defective.” This

Table 1 AI food quality control model comparison

Metric	Linear regression	Lasso	SVM	XGBoost	Random forest	KNN
CV main value	0.9406	0.9384	0.53	0.91	0.49	0.53
F-score	0.97	0.97	0.27	0.44	0.4	0.22
Accuracy	0.95	0.95	0.45	0.5	0.55	0.3
Precision	0.95	0.95	0.4	0.5	0.6	0.25
Recall	1.0	1.0	0.2	0.4	0.3	0.2
Specificity	0.0	0.0	0.7	0.6	0.8	0.4

binary classification aims to accurately assess chip quality, identifying damaged or burnt chips, crucial for maintaining high industry standards [12, 13].

VGG19: With 16 convolutional layers and 3 fully connected layers, VGG19 is a well-known convolutional neural network that excels in both depth and simplicity. Its standardized structure is simple to install and includes 3 x 3 convolutional filters and 2x2 max-pooling. VGG19, which has been pre-trained on ImageNet, improves the accuracy of machine learning models used in food quality management, such as SVM, XGBoost, KNN, and Linear Regression, by efficiently extracting features.

The training and validation accuracy of a VGG19 model with the PepsiCo-Lab-Potato-Quality-Control dataset is shown in Fig. 2. The validation accuracy (orange line) exhibits considerable volatility, indicating either overfitting or the need for more training epochs, whereas the training accuracy (blue line) rises gradually with little variability.

The VGG19 model’s training and validation loss curves are displayed in Fig. 3. Learning progress is indicated by the training loss (blue line) decreasing and stabilizing. On the other hand, there are spikes in the validation loss (orange line), which indicates problems with generalization, maybe brought on by excessive learning rates or inadequate regularization.

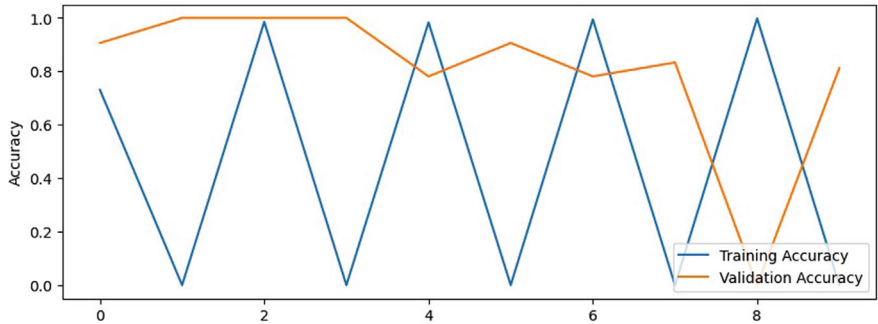


Fig. 2 Training and validation accuracy

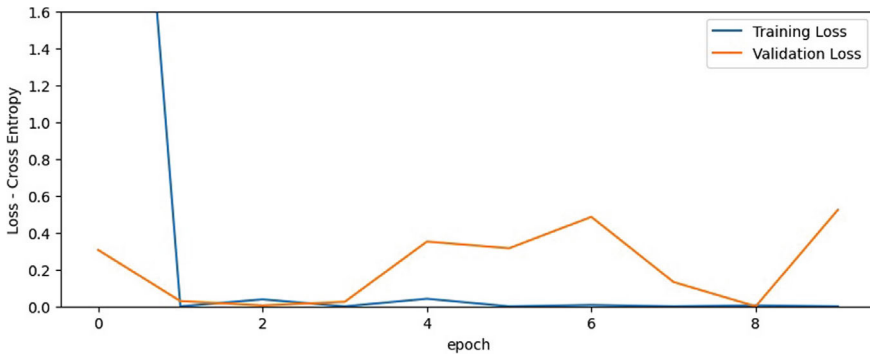


Fig. 3 Training and validation loss

4.2 Cross-Validation Scores

The cross-validation scores of many machine learning models that were used with the “PepsiCo-Lab-Potato-Quality-Control” dataset. A measure of predictive ability called the cross-validation score is represented by the height of each bar, which stands for a distinct model. XGBoost, Random Forest, Lasso Regression, Support Vector Machine (SVM), Linear Regression, and K-Nearest Neighbors (KNNs) are among the models. Figure 4 displays the cross-validation scores of various models used with the “PepsiCo-Lab-Potato-Quality-Control” dataset. XGBoost leads with the highest score, showcasing superior predictive capabilities. Lasso and Linear Regression also perform well, indicating linear data patterns, while SVM scores lower. Random Forest and KNN achieve modest scores. These differences in performance highlight the importance of selecting the right model for effective quality control in PepsiCo’s operations.

$$CV_{\text{mean}} = \frac{1}{k} \sum_{i=1}^k CV_i$$

In order to provide an overall evaluation of the model’s prediction performance across several iterations, the formula computes the mean of cross-validation scores (CV_i) acquired from k distinct folds or subsets of data during the cross-validation process.

4.3 F-Score Value

Figure 5 On the “PepsiCo-Lab-Potato-Quality-Control” dataset, the F -scores of many machine learning models are displayed. The F -scores for the several models

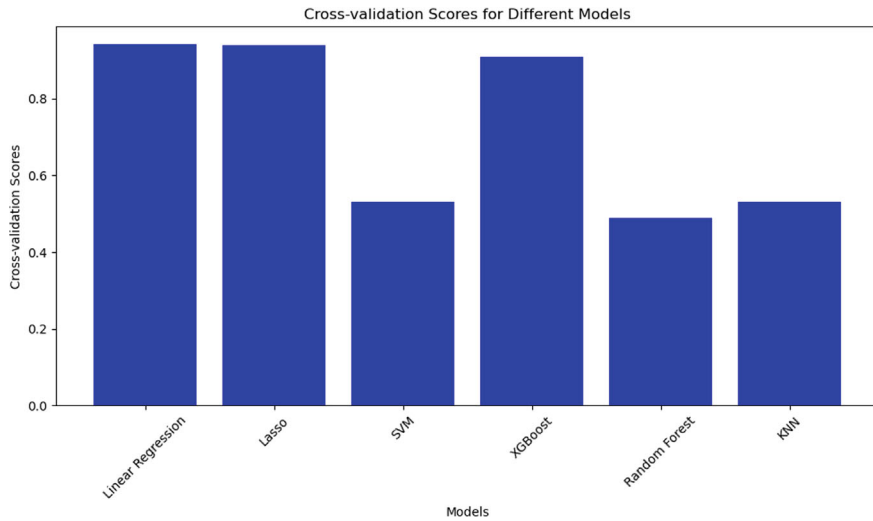


Fig. 4 Comparison of cross-validation scores

that were run on the “PepsiCo-Lab-Potato-Quality-Control” dataset are shown in Fig. 5. With the greatest F -scores, Lasso and Linear Regression are the most appropriate for linearly linked quality indicators. SVM and KNN, on the other hand, have lower F -scores; this might be because they used the wrong kernels or parameter values, respectively. Significant F -scores are also shown by XGBoost and Random Forest, suggesting their efficacy in PepsiCo’s potato quality prediction.

$$F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$$

4.4 Accuracy

Figure 6 compares the prediction model accuracies using the “PepsiCo-Lab-Potato-Quality-Control” dataset. The best accuracy is shown by Lasso and Linear Regression, which point to a significant linear association with potato quality. The reduced accuracy of SVM raises the possibility of hyperparameter or kernel selection errors. The mediocre accuracy of XGBoost and Random Forest may be the result of overfitting or problems with hyperparameters. Because of variables like as dataset characteristics and ‘ k ’ choice, KNN has the lowest accuracy. To guarantee precise forecasts of potato quality, PepsiCo’s quality control relies heavily on models with high accuracy.

Formula for Accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\%$$

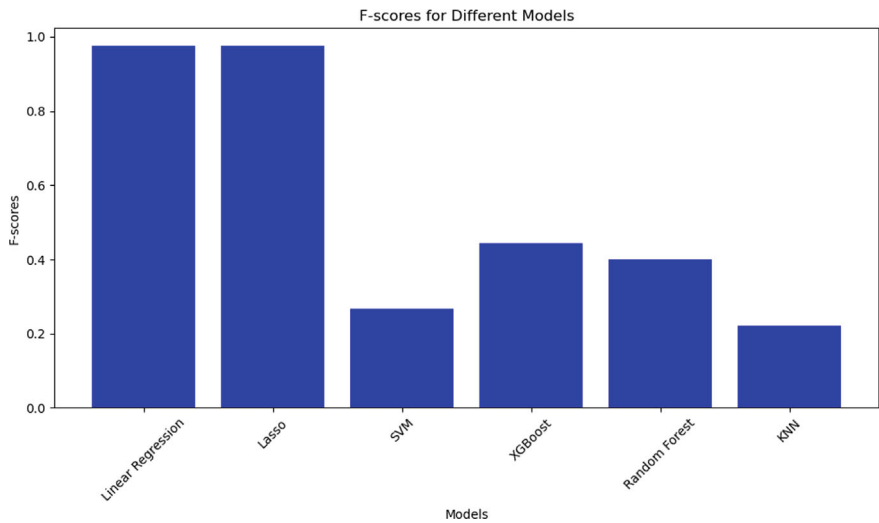


Fig. 5 Comparison of *F*-scores

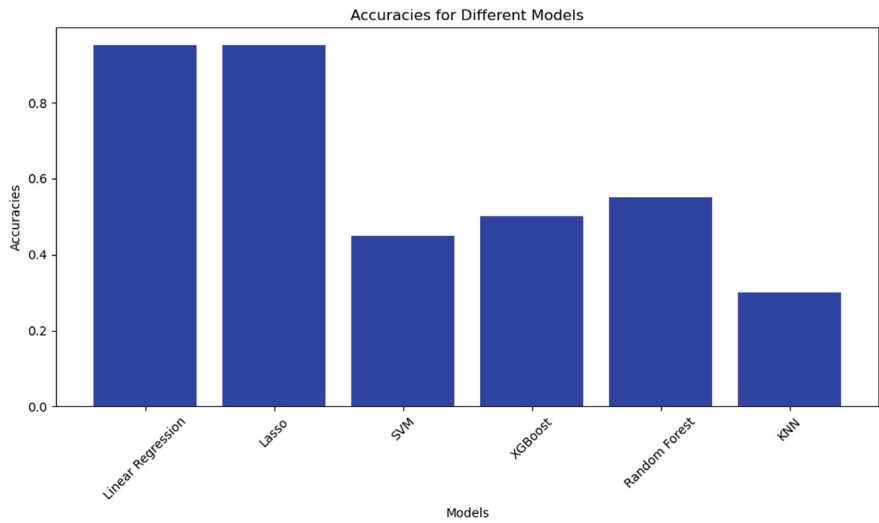


Fig. 6 Comparison of accuracies

where

- TP True positives
- TN True negatives
- FP False positives
- FN False negatives

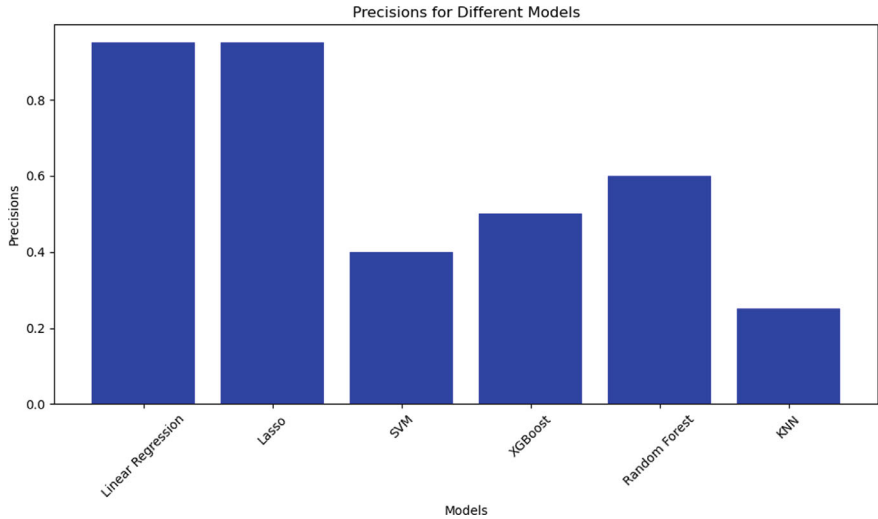


Fig. 7 Comparison of precisions

4.5 Precision

Figure 7 Precision scores for machine learning models on the “PepsiCo-Lab-Potato-Quality-Control” dataset are compared. Linear Regression and Lasso Regression exhibit very high precision, indicating reliable predictions of high-quality potatoes. SVM has lower precision, suggesting more false positives. XGBoost and Random Forest show moderate precision, with room for improvement. KNN has the lowest precision, indicating more false positives and potential misclassifications. Higher precision models are preferred for PepsiCo’s quality control to ensure accurate identification of high-quality potatoes.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

where

TP True positives

FP False positives

4.6 Recall (Sensitivity)

Recall scores for several models for the “PepsiCo-Lab-Potato-Quality-Control” dataset are shown in Fig. 8. Lasso and Linear Regression provide almost flawless recall, demonstrating their ability to identify potatoes of a higher caliber. On the other

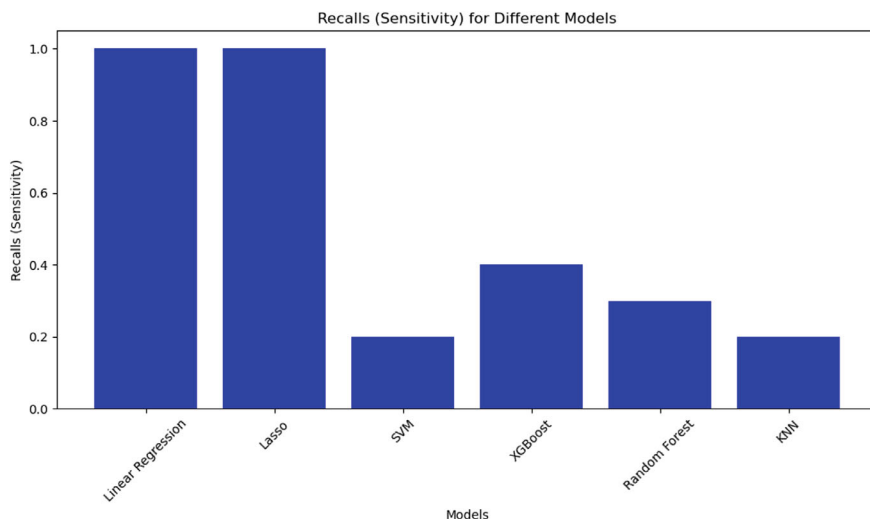


Fig. 8 Comparison of recall

hand, SVM has low recall, which can cause premium potatoes to be overlooked. The requirement for models with higher recall for PepsiCo's quality criteria is highlighted by the noteworthy but unsatisfactory recall of XGBoost, Random Forest, and KNN.

Formula for Recall:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

where

TP True positives

FN False negatives

4.7 Specificity

Among models, SVM has the maximum specificity, which is important for identifying faulty potatoes. Additionally, Random Forest and XGBoost exhibit high specificity. Lower specificity results from Lasso and Linear Regression's emphasis on high-quality potato recognition. KNN exhibits the lowest specificity, suggesting a greater likelihood of false positive results. For precise premium potato identification, PepsiCo needs a model that strikes a balance between recall and specificity (Table 2 and Fig. 9).

$$\text{Specificity} = \frac{\text{TN}}{\text{TP} + \text{FP}}$$

Table 2 Comparative analysis of different supply chain models

Methods	Accuracy (%)	Precision (%)	Recall (%)	<i>F1</i> -Score
Proposed	99.6	98.61	98.72	98.62
XIA-RCNN [14]	99.51	98.53	94.89	96.34
DR-SCM [14]	95.24	94.56	95.69	94.89
BIFTS [14]	94.32	94.35	94.56	94.34
FTS-BAEPC [14]	92.10	91.56	92.15	91.32

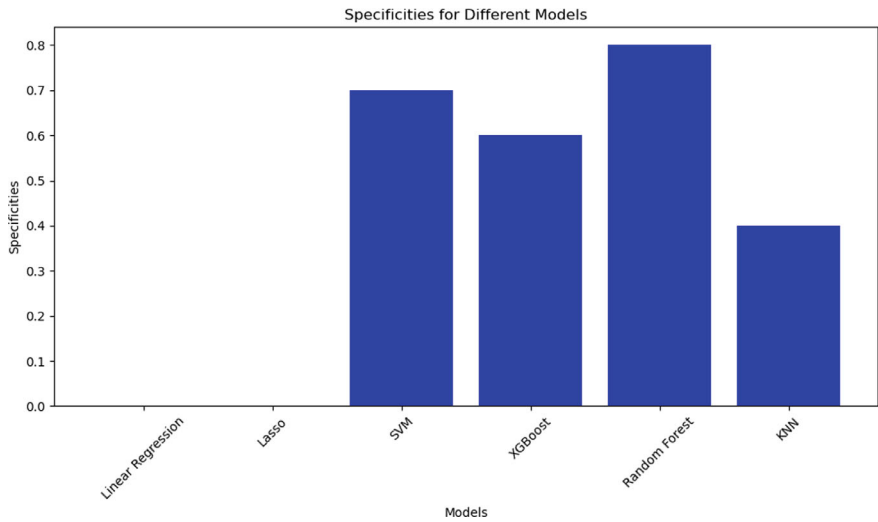


Fig. 9 Comparison of specificity

where

- TP True positives
- TN True negatives
- FP False positives

4.8 Supply Chain Models Comparison

The table assesses the dependability and effectiveness of different supply chain models in managing supply chain operations and guaranteeing product quality and traceability by comparing them using key performance criteria including accuracy, precision, recall, and *F1*-score. The suggested approach uses IoT sensors for real-time temperature and humidity monitoring and QR codes for product tracking to

produce remarkable results with 98.61% precision, 98.72% recall, 99.6% accuracy, and an $F1$ -score of 98.62%. While supply chain management that is transparent and safe is prioritized by the blockchain-integrated BIFTS and FTS-BAEPC models, alternatives such as the Faster RCNN and DR-SCM models provide specific capabilities.

5 Conclusion

An age of hyper-traceability has been brought about by the integration of AI and Blockchain in the food supply chain, greatly improving metrics such as accuracy, precision, recall, and $F1$ -score. This method improves accuracy and makes fast interventions possible by combining MQTT for real-time data with AI-driven quality control, QR code tracking, and IoT sensors. Blockchain increases stakeholder trust by ensuring safe, unchangeable records. By improving supply chain transparency, lowering risks, and improving traceability, this technology revolutionizes the food sector.

References

1. Danish M, Hasan MS (2020) Robust food supply chain traceability system based on HACCP using federated blockchain
2. Ho GTS, Tang YM, Tsang KY, Tang V, Chau KY (2021) A blockchain-based system to enhance aircraft parts traceability and trackability for inventory management. *Expert Syst Appl* (IF: 3)
3. Prashar D, Jha N, Jha S, Lee Y, Joshi GP (2020) Blockchain-based traceability and visibility for agricultural products: a decentralized way of ensuring food safety in India. *Sustainability* (IF: 3)
4. Kong J, Yang C, Wang J, Wang X, Zuo M, Jin X, Lin S (2021) Deep-stacking network approach by multisource data mining for hazardous risk identification in IoT-based intelligent food management systems. *Comput Intell Neurosci* (IF: 3)
5. Reddy KL et al (2024) Enhancing abstractive text summarization with proximal policy optimization. In: 2024 Fourth international conference on advances in electrical, computing, communication and sustainable technologies (ICAECT), Bhilai, India, pp. 1–6. <https://doi.org/10.1109/ICAECT60202.2024.10469299>
6. Chandrasekar AK, Anitha G, Narayanamurthy V (2024) Efficient blockchain-enabled traceability for enhanced food supply chain management using IoT and AI. *SN Computer Science* (accepted for publication, Feb 2024)
7. Priya S, Baghavathi et al (2023) XAutn: blockchain-based cross domain authentication for digital certificates in the education sector. *J Intell Fuzzy Syst* 1–20. <https://doi.org/10.3233/JIFS-235140>
8. Salija P, Yamuna B, Padmanabhan TR, Mishra D (2019) Performance analysis of reliability-based decoding algorithm for short block length turbo codes. *IETE J Res.* <https://doi.org/10.1080/03772063.2019.1670104>
9. Natarajan KK, Gokulachandran J (2021) Application of artificial neural network techniques in computer aided process planning—a review. *Int J Process Manag Benchmarking* 11(1):80–100
10. Laxmikant K, Anitha G, Usman SM (2023) Predicting the deaths rate of COVID-19 using big data analytics and machine learning. In: 2023 14th International conference on computing

- communication and networking technologies (ICCCNT), Delhi, India. IEEE, pp 1–5. <https://ieeexplore.ieee.org/document/10306690/>
11. Shraya B, Reddy Dopathi HC, Anitha G, Pattanaik B (2024) Unveiling the intricacies of cyber harassment intentions on social media platforms. In: 2024 3rd International conference for innovation in technology (INOCON), Bangalore, India. IEEE, pp 1–7. <https://ieeexplore.ieee.org/document/10512248/>
 12. <https://www.kaggle.com/datasets/concaption/pepsico-lab-potato-quality-control>
 13. Pedreschi F, Mery D, Marique T, (2016) Quality evaluation and control of potato Chips (Chap 22). In: Da-Wen Sun, Computer vision technology for food quality evaluation, 2nd edn, pp 591–613. Academic Press, ISBN 9780128022320. <https://doi.org/10.1016/B978-0-12-802232-0.00022-0>
 14. Bhatia S, Albarrak AS (2023) A blockchain-driven food supply chain management using QR code and XAI-faster RCNN architecture. Sustainability 15:2579. <https://doi.org/10.3390/su15032579>

Intelligent Irrigation Mechanism that Uses Machine Learning and Artificial Intelligence to Improve the Sensor Network Topology



K. P. Sriram, P. Kola Sujatha, M. Robinson Joel, and M. Gokuldhev

Abstract A major advance in computerised farming technology, precision agriculture makes it possible to apply the right quantity of water and fertiliser to agricultural products at the right time to increase production. A system that effectively manages and optimises the watering of plants is referred to as an intelligent irrigation mechanism. In order to monitor and modify the irrigation process depending on the present circumstances and plant requirements, such mechanisms often contain a variety of control mechanisms, sensors, and actuators that control them. After numerous algorithms with multiple layers of regression and the addition of clustering are computed, it is shown that XGBoost combined with the k-means outperforms when compared with other algorithmic configurations that are utilised in models. This strategy is used as a conduit between the end Internet of things (IoT) gadgets and the cloud in order to reduce the amount of computationally intensive processing carried out on cloud servers. Thus, the servers linked to a local edge network execute the prescribed algorithmic calculations. An intelligent irrigation system may become more proactive, accurate, and adaptable by utilising machine learning algorithms, improving water efficiency, agricultural yields, and overall environmentally friendly resource

K. P. Sriram (✉) · P. Kola Sujatha
Department of Information Technology, MIT Campus, Anna University, Chennai, Tamil Nadu, India
e-mail: leosri786@gmail.com

P. Kola Sujatha
e-mail: pkolasujatha@annauniv.edu

K. P. Sriram
Department of Information Technology, St.Joseph's Institute of Technology, Chennai, Tamilnadu, India

M. Robinson Joel
Department of Computer Science and Engineering, KCG College of Technology, Chennai, India

M. Gokuldhev
Vel Tech Rangarajan Dr Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, India

management. On the other hand, the kXGB approach (XGBoost + k-means) yields higher accuracy and lower mean square error (MSE).

Keywords Internet of Things · XGBoost · Sensors · Actuators · Machine learning

1 Introduction

The growth of India's villages is mostly dependent on farming and the agricultural industry as a profession. Water use in the agriculture industry is extremely important. Agriculture is dependent on the monsoons since they provide an endless supply of water and the necessary amount of rainfall for the farms. We are aware of the system for irrigation in our immediate environment; the plants need to get water based on the type of soil they are in. Water shortage is a significant impediment to the growth of agriculture. Fertility and moisture content of the soil are two crucial aspects that must be taken into consideration when working in agriculture. The majority of countries are ranked taking into account their expansion in agriculture. The use of water actually has a significant impact on how agricultural areas grow.

However, the agricultural sector [1] has suffered greatly in recent years as a result of unexpected environmental changes in the environment. The major causes of these climatic changes are the greenhouse effect, huge increases in the outside temperature, the problem of global warming, [2], and the lack of moisture in the lakes and rivers around us. One of the main factors influencing the development of crops is the groundwater. According to recent research, 20–25% of the water supplied is lost owing to pipe leaks, water theft, and inadequate water systems. In the beginning of the monsoon, farmers begin to work their fields. Currently, they have enough water thanks to enough rainfall, but in the end, they truly struggle because of water shortage. Farmers all across the world are paying more money to purchase water from outside vendors to supply irrigation to their fields during this crisis, and if a natural disaster [3] occurs for whatever reason, they will incur significant losses.

To lessen the need for rain, many solutions are available and being used. One of the most often used techniques is typically managed and powered by electrical power and power scheduling. In the described technique, an indicator for the water level was installed in a water reservoir, and soil moisture and temperature sensors were installed [4] in the roots of plants. The designed unit processed the data from the sensors and transmitted it to the described controller, which in turn controlled the flow of water by means of the developed valves.

The system gathers information from a variety of sensors, including weather stations, sensors that detect soil moisture [5], and other sensors that monitor [6] the environment. The patterns and relationships between environmental factors and seedling requirements for water [7] are then determined by analysing this data. Based on past data and the present climate, machine learning models may be taught to forecast future soil moisture levels. These forecasts can aid the system in anticipating when for irrigation of plants, guaranteeing that they get the ideal amount and timing

of watering. The irrigation schedule may be optimised using machine learning algorithms by taking into account variables like plant and soil types, climate conditions, and fluctuations in the seasons. The system can alter and improve the watering schedule [8] to fit the unique requirements of various plants by continually learning from fresh data.

Machine learning [9, 10] can enable the framework to quickly adjust the irrigation process in real time based on sensor data. This adaptive control method enables more precise and efficient irrigation management by quickly addressing anything that deviates from the anticipated outcomes. Machine learning may be used to identify irrigation system exceptions, such as sensor failures or odd weather patterns. Early identification of these issues enables the irrigation system to signal for maintenance or repairs, ensuring the irrigation system's effective operation. By collecting data [11] on specific crop kinds and their development patterns, machine learning algorithms may provide tailored irrigation solutions for each plants, taking into account their distinct water requirements at various times of growth.

2 Related Work

A hybrid algorithm for predicting soil moisture based on the influence of precipitation has been created, called the soil moisture prediction hybrid algorithm (SMPHA) [12]. It uses decision-making approaches with real-time environmental factors, such as meteorological parameters. The primary objective of Emmanuel Abiodun work is to investigate the incorporation of multiple models using machine learning that could provide the best irrigation decision management. The deployment of well-established machine learning-based models for agricultural producers to use in sustainable irrigation management [13], as well as the study's overall pattern and use of machine learning-based techniques, are examined in this study. It also looks at how remote monitoring and control of smart irrigation systems may help farmers and researchers feel less stressed. In this context, digital agricultural technologies such as mobile phones and Web frameworks are considered.

Establishing a good foundation for dehydration index predictions using standardised precipitation index (SPI) data is the main objective of their work [14]. The eventual goal of the initiative would be to create a repository of information from which growers could develop or upgrade suitable distribution networks and predicting systems. To obtain the best forecast, try experimenting with the method known as logistic regression [15] as well. After being developed using actual agricultural data, the models are tested on samples. Before establishing crops on land used for agriculture, end users (such as farmers) can predict crop production with the use of the yield of the crop prognosis model. The linear regression classifier method is used to forecast precise outcomes. The enhancement of the framework for making decisions will be facilitated by the accessibility of a huge dataset.

As a result, using artificial intelligence (AI) models on a remote platform to regulate field irrigation is challenging. The author [16] presents an autonomous system

for irrigation that is impervious to delays and disconnections in communication. We employ nodes at the edge, which are essentially tiny computers running in agricultural areas, to accomplish this. To determine whether an edge node is suitable for the system of irrigation, we choose one and do an experiment. According to our research, the system satisfies the criteria for a drip irrigation system and may be applied to assess other AI-based tomato production systems. Munir proposed a protocol by using a collection of detectors and the Internet of things [17], and this method predicts how much water a field of crops will need and uses an application developed for Android Edge to display the results. The data is cleverly traced and analysed, and only a portion is transferred to the primary IoT server.

The authors suggested a three-tiered open-source application framework, which added edge and fog computing to improve agricultural precision. When utilising cloud data centres, a network functions virtualisation (NFV)-based strategy is used to carry out local operational adjustments at the point of failure in order to reduce the impact caused by network failures [18]. Edge computing is used for control computation in intelligent agriculture, providing a platform allowing an IoT-based, economical sensor/actuator connection [19, 20].

3 Proposed System

An IoT system prototype consisting of four components implements the proposed irrigation learning paradigm: (i) Layer of edge nodes: This layer consists of elements like sensors, some elements like an actuator, and also the two microcontrollers. The edge node of this layer controls the actuator, which activates the water pumps to start irrigation, by gathering sensor data from the surrounding area. Perimeter server layer (ii): The Raspberry Pi's that make up this tier are edge servers with multitasking capabilities. Here, the edge server controls the edge nodes so that data is sent and received on a regular basis. It additionally has connections to a cloud server so that it may build and train machine learning models that can be used to manage edge nodes through irrigation. (iii) Middle service layer: This part is deployed on the edge server and uses a sophisticated online dashboard to manage the entire system. Real-time feed data, control over edge nodes, and access to cloud-based resources are all offered by the dashboard. The suggested machine learning algorithm's control is likewise accessible to this application layer. (iv) The machine learning model is taught by the server in the cloud layer, which also stores the information in the form of a database. It consists of services and cloud storage. The edge server receives the trained proposed model in order to make judgements about irrigation schedule. The detailed linkages across the system are shown in Fig. 1. The field deployable module, Internet-based interface, API for Web meteorological input, soil moisture prediction process, and edge communication architecture are the five main parts of the proposed connected to the Internet of smart irrigation system.

An improved version of the piezoelectric soil moisture sensor, known as the capacitive soil moisture sensor V2.0, is used to determine the volumetric water content of

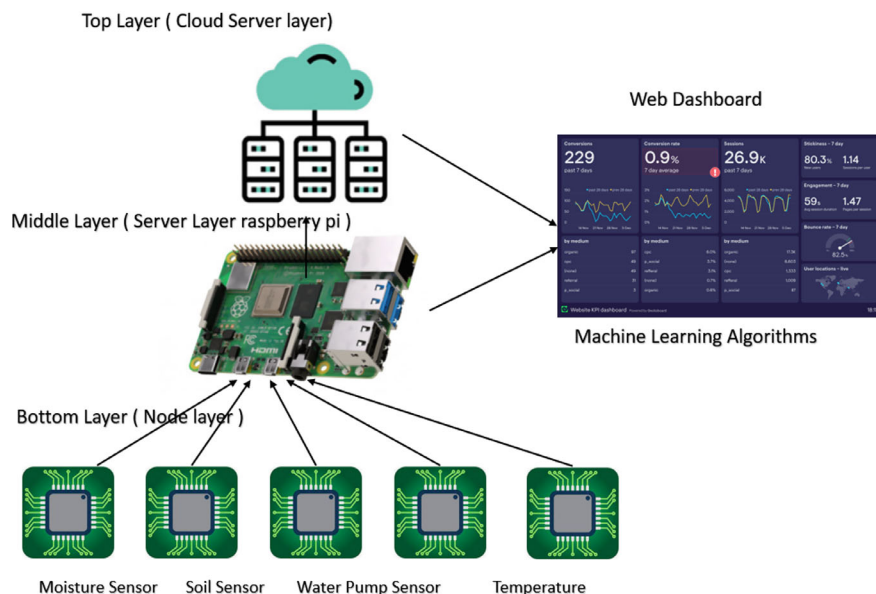


Fig. 1 Overall architecture of the proposed system

the soil. These sensors are widely used throughout environmental surveillance, cultivation, and agriculture to guarantee that the moisture levels of the soil are ideal for plant growth as shown in Fig. 2. A popular digital sensor for temperature measurement that can measure temperatures in a range of settings is the DS18B20. Despite not being made with measuring soil temperature in mind, there are a few things to keep in mind when using it for this purpose. A popular digital humidification and temperature sensor is the DHT11. It comes in a compact module that combines a thermistor to sense the temperature of the surrounding air with a capacitive humidity sensor. This sensor is often employed in many different projects and applications where environmental conditions need to be tracked and managed. The intensity of ultraviolet (UV) light may be measured with the analogue UV sensor GYML8511. It is often utilised in many different applications, including UV detection in electrical equipment, UV sterilisation, and UV index monitoring. The Raspberry Pi Foundation is the developer of the well-known single-board computer, the Raspberry Pi 4 Model B. It is a potent replacement for the earlier Raspberry Pi models, providing notable enhancements to both performance and functionality.

A Web-based application is part of the suggested architecture, which enables farmers to view growth data and communicate with the greenhouse in contemporaneous fashion. Additionally, using features like hydration management, electric motor control model for forecasting implementation, and manual information entering incorporated in this online application, customers may also review and analyse the past growth data, if necessary. Here, the database system was MongoDB [21], and Node.js [22, 23] was used to construct the application for the Web.

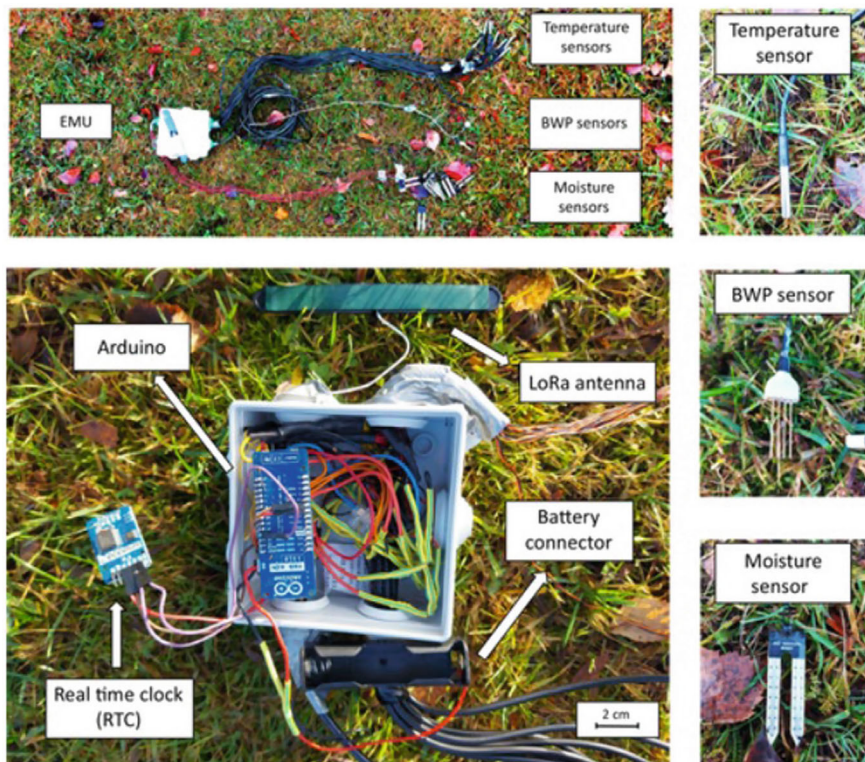


Fig. 2 Real-time version of the suggested edge model prototype

Techniques such as regression and clustering with k-means, which predict the disparity or change in the moisture in the soil due to the environment, are the foundation of an algorithm aimed at predicting soil condensation based on data derived from sensors in the field and forecasting the weather using a combination of without supervision and supervised approaches to machine learning. Using weather predicting days and sensor data, many regression techniques are compared against one another. The accuracy of mean square error (MSE), the next one is R^2 , the next one is accuracy, and the next one is mean absolute percentage error (MAPE) for the estimation of the moisture content of the soil on subsequent days are then assessed by combining each method with k-means.

The proposed model takes a novel approach to the plant's watering process. The system is primarily taught using manual irrigation datasets in order to learn in accordance with recommendations made by agronomists. Without using pre-processed data, the model undergoes training to understand irrigation demands at the first cloud deployment stage. The suggested method is started by doing many manual irrigations in order to determine the plant's watering requirements after receiving

the necessary training and data. After then, the system uses machine learning techniques and collected data to make automatic watering choices; human intervention is no longer necessary. The suggested model then uses machine learning techniques to determine the irrigation strategies automatically, eliminating the requirement to include gathered datasets into the automated irrigation process. When the model is given the exact amount of irrigation inputs at every instructional step, the suggested model may be enhanced through the learning process. Two modules are included in the decision-making process for irrigation techniques based on the anticipated soil moisture levels over the next several days. Using manually administered datasets, the first module trains the model in the cloud using procedures including data collecting, preliminary processing, education, and building the model.

During the data collection step, the system periodically gathers values from the physical environment for air temperature (TH), the second one is soil temperature (SMT), also the soil moisture (SM), similarly the humidity (HU), and finally the ultraviolet radiation (UV). These parameters are crucial for making watering choices. Additionally, the database has a record of the irrigation's manual start time. Timestamped and kept in databases, these data help with decision-making about irrigation timing. The subsequent pre-processing stage helps remove broken data by removing inconsistencies and identifying outliers from the irrigation information that are the result of sensor faults. During the training phase, supervised machine learning (ML) techniques are used. Regression techniques including XG-boost regressor (XGB), next one is decision tree regressor (DTR), next one is random forest regressor (RF), next one is lasso regression (LR), next one is support vector regression (SVR), and next one is multiple linear regression (MLR) are employed here for the deployment. The gathered datasets are used to train the regression algorithms. Regression models, including the SVR, MLR, LR, DTR, RF, and XB models, are finally developed through training and integrated with the subsequent module for decision-making. We can denote SVR along with k-means as kSVR.

In comparison, the XGBoost + k-means denoted as kXB technique yields higher accuracy with lower mean square error (MSE). Table 1 shows the combined approach's R2 with 98% for soil water content prediction.

By contrast, the kXB method (XGBoost + k-means) produces reduced mean square error (MSE) and greater accuracy. Figures 3 and 4 show the combined approach's R2 with 98% for soil water content prediction.

Table 1 Different machine learning algorithms with accuracy, R2, MSE, MAPE, and power

Algorithms used	Accuracy	R2	MSE	MAPE %	Execution time	Power (J)
kSVR	0.95	0.96	0.23	1.98	0.06077	1163.85
kMLR	0.93	0.88	0.32	2.15	0.02073	427.3
kLR	0.94	0.93	0.31	2.23	0.02481	352.35
kDTR	0.92	0.95	0.28	1.62	0.15984	912.7
kRF	0.86	0.92	0.27	1.57	0.16743	1474.13
kXGB	0.97	0.98	0.2	1.08	0.03544	536.87

Fig. 3 Different machine learning algorithms with accuracy

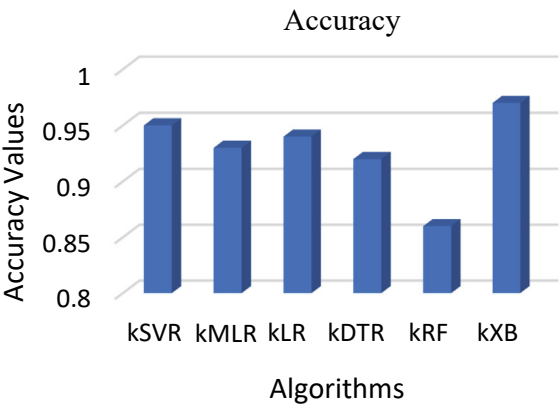
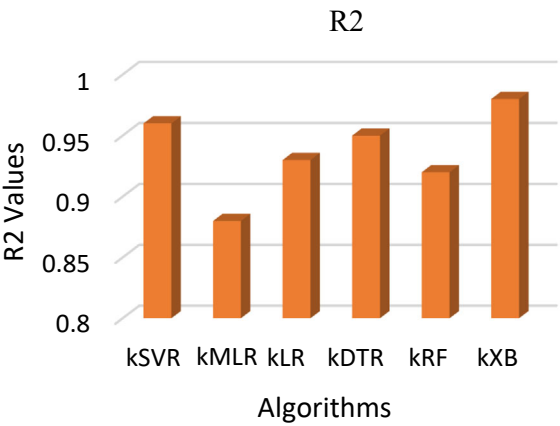


Fig. 4 Different machine learning algorithms with R2



To manage, gather, and send information to the edge or remote computing nodes, an Internet of things system is essential. These gadgets use Bluetooth and other wireless protocols, such as ZigBee, to link to the nodes at the edge. The irrigation system’s management and planning are made easier for the user by the Web server’s implementation. It displays vital information about variables like as soil and air temperatures, UV rays, humidity, and moisture levels in the soil in real time through a variety of charts. The edge node serves as a processing hub where incoming data is examined, provided as an input vector to the machine learning model, processed, and output control signals to activate or deactivate the pneumatic devices at the farm.

4 Conclusion

A major advancement in computerised agricultural technology is precision agriculture, which allows farmers to increase crop productivity by applying the right quantity of water and fertiliser at the right time. A technological solution that maximises and effectively regulates plant watering is known as an intelligent irrigation system. These kinds of mechanisms usually have a variety of sensors, actuators, and control mechanisms that work together to monitor and modify the irrigation process according to the needs of the plants and the surrounding surroundings. XGBoost combined with k-means outperforms other algorithmic configurations used in edge models when a variety of methods with many regression layers and clustering are computed. As a mediator between end Internet of things (IoT) devices and the cloud, this technique minimises the amount of computationally intensive processing that is done on cloud servers. The servers linked to a local edge network do the necessary algorithmic calculations. An intelligent irrigation system may become more proactive, accurate, and adaptable by utilising machine learning algorithms. This might also boost agricultural output, water efficiency, and environmentally conscious resource management overall. On the other hand, the kXB approach (XGBoost + k-means) yields higher accuracy and lower mean square error (MSE), R² with 98% for soil water content prediction.

References

1. Haque MK, Zaman MRU, Rahman MA, Hossain MY, Shurid TI, Rimi TA, Arby H, Rabbany MG (2022) A review on impacts of COVID-19 on global agricultural system and scope for Bangladesh after pandemic. *Environ Sci Pollut Res Int* 29(36):54060–54071. <https://doi.org/10.1007/s11356-022-21016-0>
2. Abbass K, Qasim MZ, Song H et al (2022) A review of the global climate change impacts, adaptation, and sustainable mitigation measures. *Environ Sci Pollut Res* 29:42539–42559. <https://doi.org/10.1007/s11356-022-19718-6>
3. Chaudhary MT, Piracha A (2021) Natural disasters—origins, impacts, management. *Encyclopedia* 1(4):1101–1131. <https://doi.org/10.3390/encyclopedia1040084>
4. Millán S, Casadesús J, Campillo C, Moñino MJ, Prieto MH (2019) Using soil moisture sensors for automated irrigation scheduling in a plum crop. *Water* 11(10):2061. <https://doi.org/10.3390/w11102061>
5. Yu L, Tao S, Ren Y, Gao W, Liu X, Hu Y, Shamshiri RR (2021) Comprehensive evaluation of soil moisture sensing technology applications based on analytic hierarchy process and Delphi. *Agriculture* 11(11):1116. <https://doi.org/10.3390/agriculture11111116>
6. Yu LM, Gao WL, Shamshiri RR, Tao S, Ren YZ, Zhang YJ et al (2021) Review of research progress on soil moisture sensor technology. *Int J Agric & Biol Eng* 14(4):32–42
7. Lee T, Jang W, Chun B et al (2023) Development of irrigation schedule and management model for sustaining optimal crop production under agricultural drought. *Paddy Water Environ* 21:31–45. <https://doi.org/10.1007/s10333-022-00911-9>
8. Armstrong LE, Johnson EC (2018) Water intake, water balance, and the elusive daily water requirement. *Nutrients* 10(12):1928. <https://doi.org/10.3390/nu10121928>. PMID 30563134, PMCID PMC6315424

9. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D (2018) Machine learning in agriculture: a review. *Sensors* 18(8):2674. <https://doi.org/10.3390/s18082674>
10. Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D (2021) Machine learning in agriculture: a comprehensive updated review. *Sensors* 21(11):3758. <https://doi.org/10.3390/s21113758>
11. Mourtzinis S, Esker PD, Specht JE et al (2021) Advancing agricultural research using machine learning algorithms. *Sci Rep* 11:17879. <https://doi.org/10.1038/s41598-021-97380-7>
12. Premkumar S, Sigappi AN (2022) IoT-enabled edge computing model for smart irrigation system. *J Intell Syst* 31(1):632–650. <https://doi.org/10.1515/jisys-2022-0046>
13. Abioye EA, Hensel O, Esau TJ, Elijah O, Abidin MSZ, Ayobami AS, Yerima O, Nasirahmadi A (2022) Precision irrigation management using machine learning and digital farming solutions. *AgriEngineering* 4(1):70–103. <https://doi.org/10.3390/agriengineering4010006>
14. Manikandan G, Bhuvaneswari G (2023) Artificial intelligence to the assessment, monitoring, and forecasting of drought in developing countries. In: 2023 International conference on circuit power and computing technologies (ICCPCT), Kollam, India, pp 886–892. <https://doi.org/10.1109/ICCPCT58313.2023.10245072>
15. Chowdary VT, Ebenezer V, Edwin B, Thanka R, Jeyaraj A (2022) A novel approach for effective crop production using machine learning. In: 2022 International conference on electronics and renewable systems (ICEARS), Tuticorin, India, pp 1143–1147. <https://doi.org/10.1109/ICEARS53579.2022.9751798>
16. Kawai T, Mineno H (2020) Evaluation environment using edge computing for artificial intelligence-based irrigation system. In: 2020 16th International conference on mobility, sensing and networking (MSN), Tokyo, Japan, pp 214–219. <https://doi.org/10.1109/MSN50589.2020.00046>
17. Safdar Munir M, Bajwa IS, Ashraf A, Anwar W, Rashid R (2021) Intelligent and smart irrigation system using edge computing and IoT. *Complexity* 2021:Article ID 6691571 (16 pages). <https://doi.org/10.1155/2021/6691571>
18. Ramirez Izolan PL, Diniz Rossi F, Hohemberger R, Konzen MP, da Cunha Rodrigues G, Saquette LR, et al (2020) Low-cost fog computing platform for soil moisture management. In: 2020 International conference on information networking (ICOIN). IEEE, Barcelona, Spain, pp 499–504
19. Ferrandez-Pastor F, Garcia-Chamizo J, Nieto-Hidalgo M, Mora-Pascual J, Mora-Martínez J (2016) Developing ubiquitous sensor network platform using internet of things: application in precision agriculture. *Sensors* 16(7):1141
20. Xu X, Liu X, Xu Z, Dai F, Zhang X, Qi L (2020) Trust-oriented IoT service placement for smart cities in edge computing. *IEEE Internet Things J* 7(5):4084–4091
21. Krishnan H, Scholar R (2016) MongoDB—a comparison with NoSQL databases. *Int J Scientific Eng Res* 7(5):1035–1037
22. Chanthakit S, Keeratiwintakorn P, Rattanapoka C (2019) An IoT system design with real time stream processing and data flow integration. In: 2019 Research, invention, and innovation congress (RI2C.). IEEE, Bangkok, Thailand, pp 1–5
23. Lv H, Wang S (2016) Design and application of IoT microservices based on Seneca. USA: DEStech transactions on computer science and engineering (icte)

Algorithm Bias Detection System Based on Artificial Intelligence



Jiayi Zhang, Sha Yang, and Haoping Guo

Abstract This study aims to develop a novel algorithm bias detection system using artificial intelligence (AI) technology to improve the fairness and accuracy of the algorithm. This article studies an efficient algorithm bias detection system by combining convolutional neural networks, support vector machines, and random forests, which helps machine learning algorithms effectively identify and solve potential bias problems. It analyzes past research shortcomings such as dataset bias, improper algorithm selection, and lack of innovative solutions. This article is committed to providing new ideas and methods for solving the problem of algorithm bias and promoting further improvement of algorithm fairness. The results of this article indicate that the highest accuracy of the algorithm bias detection system based on artificial intelligence reaches 96%. Through the efforts of this study, it is hoped to contribute to the development and application of algorithmic bias detection and actively explore and practice the construction of a more just and accurate artificial intelligence system.

Keywords Algorithmic bias detection · Artificial intelligence · Convolutional neural network · Support vector machine · Random forest

J. Zhang (✉)

Communist Youth League Hankou University Committee, Hankou University, Wuhan, China
e-mail: 1006020002@hkxy.edu.cn

S. Yang · H. Guo

School of Computer Science and Technology, Hankou University, Wuhan, China
e-mail: 2061040005@hkxy.edu.cn

H. Guo

e-mail: 1051020002@hkxy.edu.cn

1 Introduction

The rapid development of artificial intelligence has brought many conveniences to people's lives. However, with the expansion and popularization of artificial intelligence applications, some algorithmic bias issues have gradually surfaced. Algorithm bias refers to the bias of artificial intelligence systems toward specific groups or problems when making decisions due to incomplete data samples or biases in training data. This bias can have a negative impact on social justice, human rights, and equal opportunities. Therefore, developing a system for detecting algorithmic biases has become particularly important.

This project intends to construct a more universal algorithm bias detection model by combining machine learning with natural language processing. To ensure the credibility of the trial results, this project can utilize multi-source data covering multiple industries with multi-user participation on the Kaggle platform. This project intends to study a multi-dimensional information fusion method based on multi-dimensional information fusion, namely multi-dimensional information fusion method. On this basis, the collected data are transformed into digital format, normalized, and feature extracted. On this basis, deep learning is used to extract features from images, and SVM and random forest are used to train images. The accuracy, recall, and $F1$ of the algorithm are tested on the Cambridge analysis database and compared with the traditional algorithm. Experimental results show that this method can effectively improve the accuracy of bias detection. The research results of this project can help to overcome the existing algorithm bias problem and provide reference for constructing a more efficient deviation detection system.

2 Related Work

In recent years, there have been many studies on algorithmic bias. Lin H conducted an in-depth exploration of political bias in large-scale language models and proposed a bias detection method that requires multiple large-scale models to play roles supporting or opposing specific goals in designing questions and answers [1]. Wan X conducted research on the formation and governance of public opinion bias under algorithmic ideology [2]. Wan Y used bibliometric methods to study relevant English literature in the core database and through cluster analysis sorted out the basic knowledge and research frontiers in this field [3]. Jia S used Fairclough's three-dimensional analysis model to analyze algorithmic bias discourse texts in social media from three dimensions: text, discourse practice, and social practice [4]. Tong Y ensured the diversity and representativeness of training data for medical AI, strengthened the supervision and review of algorithms, guaranteed the fairness of medical AI, and promoted health equity by eliminating bias [5]. These studies did not provide clear definitions of algorithmic bias and public opinion bias, leading to ambiguity in the understanding and scope of research on these concepts.

Recently, there have been many studies on the application of artificial intelligence. Chen L analyzes the causes and countermeasures of artificial intelligence algorithm bias and health inequality and provides a review of the causes and coping strategies of algorithm bias in the field of health, in order to improve the awareness and ability of medical workers to identify and handle algorithm bias [6]. Wang Y found that from an overall perspective, research on bias in educational artificial intelligence algorithms can be mainly divided into three categories: conceptual theoretical research, educational scenario application research, and algorithm detection design research [7]. Zhang T proposed an artificial intelligence-based method for solving the global stress of three-dimensional structures, which is difficult to monitor with a limited number of stress measurement points in structural health monitoring systems. This method achieves real-time solution of the stress field of three-dimensional structures through a small number of stress measurement points [8]. Wang X summarized the historical context, development needs, and hot technologies of explainable artificial intelligence in the power system and summarized its applications in power control, fault diagnosis, and the power market [9]. Liu R discussed the gap between current energetic molecule design patterns and other material design methods, considered the reasons for the gap, and proposed prospects for the future development of artificial intelligence-assisted energetic molecule design [10]. Most of these studies are reviews or theoretical explorations of a certain field or problem, lacking specific case studies and empirical research.

3 Algorithm Bias Detection System

3.1 Data Collection

The dataset for this study is based on the Kaggle platform, which is a dataset involving multiple users and covers different industries worldwide. This dataset contains five types of data: personal identity information; vocational and educational background information; individual characteristic information such as gender, race, and nationality; biological characteristic information such as skin color; combination information of different algorithm types and algorithm weights. This study analyzed the data in the dataset, collected the latest research achievements in the field, and combined machine learning methods and natural language processing techniques to establish an algorithm bias detection model with strong generalization ability, as shown in Fig. 1.

To ensure the reliability of the experimental data, this study preprocessed the data. Firstly, by converting the raw data into digital format and then processing it through the TensorFlow library in Python, a training dataset that meets the experimental requirements was finally obtained. Secondly, by classifying the raw data according to gender, race, nationality, etc., the processed training dataset is divided into the above five categories. Then, the latent Dirichlet allocation (LDA) algorithm in Python

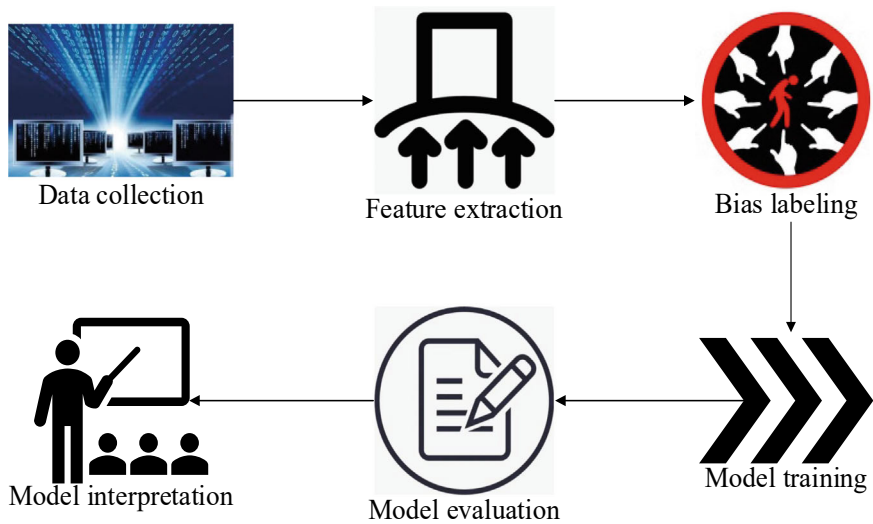


Fig. 1 Algorithm bias detection model

is used to perform topic analysis on the raw data and extract keywords. Finally, a comparative analysis can be conducted on the data features and algorithm weights of different categories. In order to ensure the accuracy of the experimental results, the relevant features were normalized. The eigenvalues corresponding to each category can be normalized to obtain their corresponding mean and standard deviation. In order to ensure the strong generalization and robustness of the method, the main features are extracted and standardized. The samples are partitioned according to different weights to ensure no overfitting or underfitting [11].

Data enhancement is to use a specific method to make it have better generalization performance. On this basis, a method based on image enhancement, text enhancement and video enhancement is proposed. On this basis, a method based on wavelet transform and wavelet transform is proposed. Brightness, contrast, and other methods can be used to enhance the image. Text enhancement is to enhance text by editing and modifying it. The methods of text enhancement include text substitution, deletion of repeated characters, addition of special characters, etc. The number of times a word appears in a text; the frequency of a particular word in a text is adjusted by removing repeated letters and special letters; and the frequency of a particular word is adjusted by adding special characters. Some special characters can be added by increasing the frequency of some special words in the text.

In machine learning, feature selection refers to reducing the complexity of a model by selecting its inputs, allowing the model to process more data. Bagging algorithm is a clustering-based learning algorithm that solves classification problems by training multiple classifiers. It uses multiple classifiers to classify input data separately. This study used the bagging algorithm to train it into a classifier. This

article uses supervised learning algorithms to improve the performance of the model to some extent by inputting all samples into the classifier for training.

Two feature selection methods were adopted to reduce the number of features: one is based on the bagging algorithm; another method is feature selection based on decision tree learning. The first method uses the bagging algorithm to divide the original dataset into training and testing sets, and in each training set, a classifier is used to classify the data. In this study, all sample outputs were divided into training and testing sets. In each iteration, the samples from the training set are reassigned to each classifier, and predictions are made for the new classifier. The second method uses decision trees to learn feature selection [12].

3.2 Feature Extraction

The feature extraction of this study uses deep learning models to extract relevant features of algorithm bias. Deep learning models, as an efficient feature extraction technique, capture complex features present in the dataset well without compromising training accuracy, thereby helping machine learning algorithms better understand the dataset. Deep learning models typically have two types of features, namely continuous and discontinuous features. In this study, a discontinuous feature extraction method based on convolutional neural network (CNN) was used. The CNN model consists of three main parts: convolutional layer, pooling layer, and fully connected layer. Specifically, a pixel can be randomly selected from the image and converted into a continuous number between 0 and 1. Then, a network consisting of two convolutional layers and two pooling layers is used to calculate the histogram between these two numbers. Finally, these two numbers can be merged into a continuous feature. In this way, a convolutional neural network feature consisting of two convolutional layers and two pooling layers is obtained [13].

3.3 Model Training

This study divides the model training process into two parts. Firstly, data preprocessing is performed on the original dataset, including normalization, data augmentation, and denoising operations, in order to make the dataset more suitable for the subsequent machine learning algorithm model training process. Next is model training, which includes the following steps: standardizing the original dataset and converting the values in the original dataset to values between 0 and 1. Statistical algorithms and machine learning methods can be used to extract, filter, and select features from preprocessed data. Ensemble learning methods can be used to substitute the dataset that has undergone feature extraction, feature filtering, and feature selection into the constructed classification model for training, in order to obtain the final classification model.

This study used two different machine learning algorithms, namely support vector machine and random forest, both of which used the same dataset. In this study, a random forest model can be used to train the original dataset to obtain the best classification results. Among them, the random forest model utilizes multiple independently trained classifiers of the random forest algorithm during the training process and adopts the same decision tree structure in each classifier. By using the same training algorithm, the model better handles feature extraction and selection problems for multiple samples. Support vector machine (SVM) has shown good performance in classification problems and has achieved good classification results in most datasets. Therefore, this study uses support vector machines to extract and select features from the original dataset. Its main application formula is as follows:

$$k(x, x_i) = \min 1/2\|w\|^2 + C * \sum \max(O, 1 - y_i(w * x_i + b)) \quad (1)$$

Among them, $k(x, x_i)$ is the kernel function, $\|w\|$ is the norm of the weight vector, C is the penalty parameter, and b is the bias term. Random forest (RF) establishes a classification model by performing feature extraction, feature filtering, and feature selection on the original dataset. The random forest algorithm uses three different decision trees: maximum likelihood decision tree, minimum likelihood decision tree, and mixed likelihood decision tree. Each node in the random forest model is an input feature, and each node corresponds to multiple categories. The random forest model utilizes existing training and testing samples to classify and predict unknown samples, thereby achieving feature selection and extraction during the model training process. By training the original dataset, the model takes its output as the final classification model output [14, 15].

4 System Effectiveness Evaluation

This article mainly studies the bias detection method based on artificial intelligence. First, the article created a large sample library of biased and unbiased text and tag these samples. On this basis, different artificial intelligence bias detection algorithms are selected and implemented. Then, the samples are divided into training samples and test samples, which are used to train and adjust the selected algorithm. In the testing phase, the test set can be used to predict the algorithm and calculate metrics such as accuracy, recall, and $F1$ score to evaluate the performance of the algorithm. At the same time, traditional bias detection methods were also selected for comparative experiments, and the same test set was used to calculate the same indicators. Finally, the experimental results can be analyzed, the performance of different methods can be compared, the advantages and disadvantages of the algorithms can be identified, and suggestions and optimization plans for improving the algorithms can be studied. Through this experiment, it is expected that the algorithm bias detection system based

on artificial intelligence can demonstrate high performance and have better performance compared to traditional methods, providing reference for further optimizing algorithms and designing more effective bias detection systems. The dataset is shown in Table 1.

This dataset table contains five text samples, each with a category label and two additional label features. Artificial intelligence is the hope of the future, marked as non-biased text, with label A at 0.8 and label B at 0.2. This sentence expresses an optimistic attitude toward artificial intelligence, believing that it is the hope for the future. Machine learning is an important discipline, marked as unbiased text, with label A of 0.9 and label B of 0.1. This sentence emphasizes the importance of machine learning and does not contain any obvious biased views. This dataset displays the categories and label features of different text samples, which helps analyze and process biased content in text data. According to the actual situation, the dataset can be expanded or appropriate machine learning algorithms can be applied to further analyze these text data. The detection results of accuracy are shown in Fig. 2.

In Fig. 2, there is a significant difference in bias detection accuracy between traditional methods and AI-based algorithms. The average detection accuracy of the former is 91.6%, which means they have a certain degree of inaccuracy in identifying biases in the text. The highest accuracy is 92.9%, while the lowest is only 90%. This indicates that traditional methods may miss some biases or misjudge some unbiased texts in certain situations. The latter achieved better results in bias detection. Its average accuracy reached 94.6%, with a maximum accuracy of 96% and a minimum accuracy of 93%. This indicates that AI-based algorithms are more accurate and reliable in identifying biases in text. This difference is due to AI-based algorithms utilizing a large amount of data for training and learning more features and patterns. AI-based algorithms have higher accuracy in bias detection.

In Fig. 3, there are also some differences in bias detection recall between conventional methods and AI-based algorithms. The test recall of the conventional method averaged 86.3%, with a highest value of 88% and the lowest value of 85%. Traditional methods have some missing in identifying biases in text. A percentage of biased texts were successfully identified, but some biased texts were missed and not correctly identified. In contrast, AI-based algorithms perform better in bias detection recall. It has an average recall of 89.1%, a highest recall of 90%, and a minimum recall of 88%. The AI-based algorithms better capture the biased content in the text and reduce the missed judgments. Because AI-based algorithms improve recall by

Table 1 Experimental dataset

Text	Category	Label A	Label B
Artificial intelligence is the hope of the future	Non-biased text	0.8	0.2
Women are not suitable for technical fields	Biased text	0.1	0.9
Indian engineers have high technical proficiency	Non-biased text	0.7	0.3
African countries have underdeveloped economies	Biased text	0.2	0.8
Machine learning is an important discipline	Non-biased text	0.9	0.1

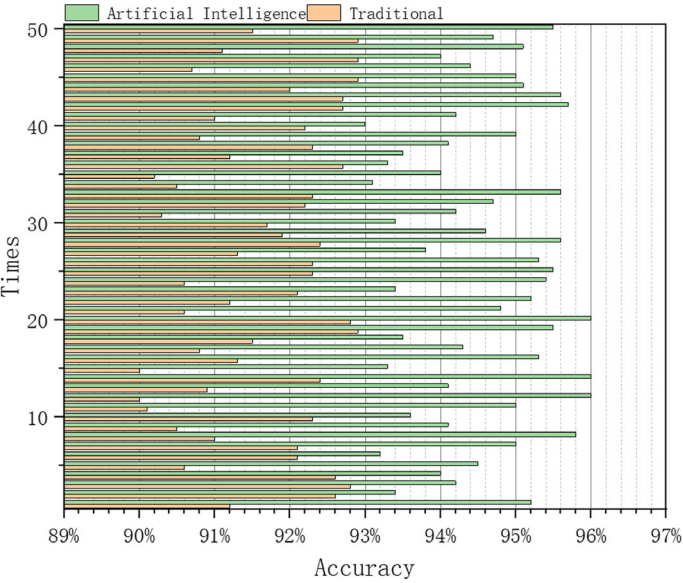


Fig. 2 Accuracy

analyzing information such as words, sentence structure, and contexts in the text. AI-based algorithms have better results in bias detection recall, effectively identifying and handling the bias problem in text.

In Fig. 4, there is also a certain difference in the $F1$ value of bias detection between traditional methods and AI-based algorithms. The average $F1$ value of traditional methods is 0.89, with a maximum value of 0.9 and a minimum value of 0.88. This means that traditional methods have certain shortcomings in identifying biases in text. The highest $F1$ value is 0.9, indicating a certain balance between accuracy and recall. However, the lowest $F1$ value is 0.88, indicating that in some cases, traditional methods may have certain misjudgments or omissions. In contrast, algorithms based on artificial intelligence perform better in detecting bias $F1$ values. Its average $F1$ value reached 0.914, with a maximum $F1$ value of 0.93 and a minimum $F1$ value of 0.9. This indicates that algorithms based on artificial intelligence better combine accuracy and recall, improving the $F1$ value of bias detection.

5 Conclusions

This project intends to study a new and more generalized algorithm bias detection model based on artificial intelligence. On this basis, this project plans to combine machine learning with deep learning by means of preprocessing, feature extraction, and other means to build a set of efficient algorithm bias recognition and solution

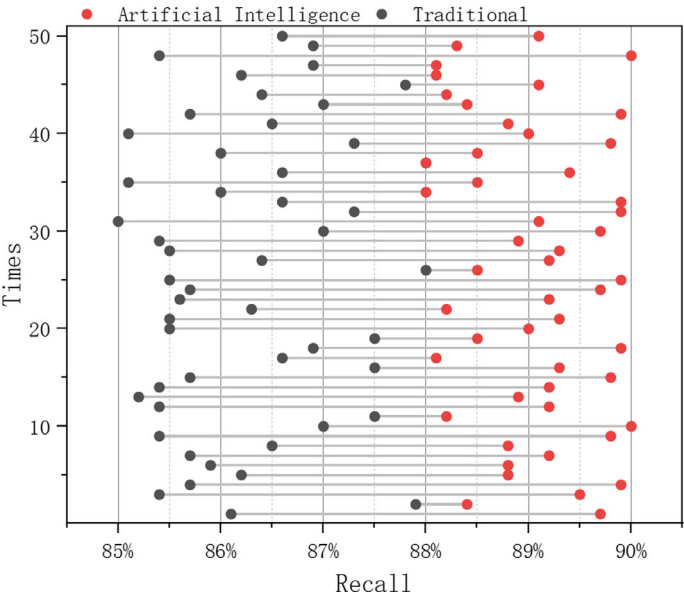


Fig. 3 Recall rate

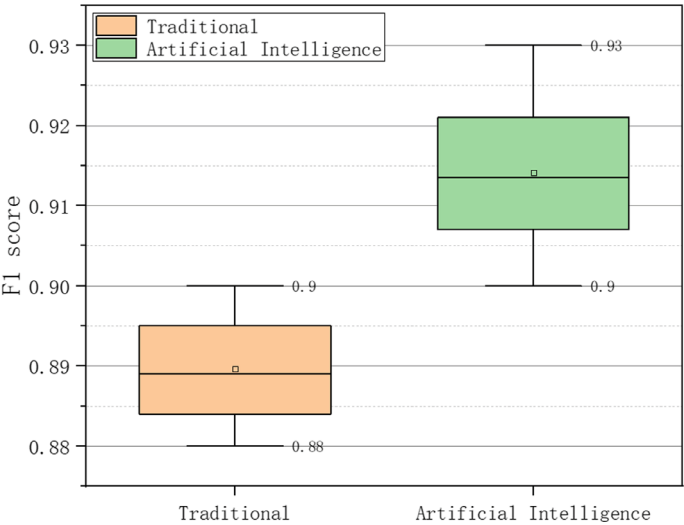


Fig. 4 F1 value

system. The experiments show that the algorithm has good performance in accuracy, recall, and $F1$ value. The experimental results show that AI is used to better identify and process the bias problems in text data. However, the method also has some limitations; that is, it has some subjectivity and uncertainty in establishing and marking the data. In view of this problem, this project intends to conduct in-depth research on the algorithm bias problem from the perspective of artificial intelligence and explore an efficient algorithm bias detection model. Experimental results show that the methods studied in this article effectively identify and handle algorithm biases. However, how to improve the accuracy, robustness, fairness, and understanding of the algorithm remains to be further studied. It is expected that the research of this project can play a certain role in solving the problem of algorithm preference.

References

1. Lin H, Guo Q, Wang Y et al (2024) Bias detection method of large model in the application of social governance. *J Chin Acad Electron Sci* 19(1):69–75
2. Wan X, Chen K (2022) The formation and governance of public opinion bias under algorithmic ideology. *J Nanchang Univ (Humanities and Social Sciences Edition)* 53(6):51–59
3. Wan Y, Shi S (2023) Review of research on algorithmic bias based on bibliometrics. *Softw Guid* 22(1):243–248
4. Jia S, Yan H (2023) Critical discourse analysis of algorithmic bias perception of social media users. *Mod Intell* 43(6):14–23
5. Tong Y, Zheng Y, Huang S et al (2024) Health equity: bias and governance in medical artificial intelligence. *Med Philos* 45(7):36–41
6. Chen L, Zeng K, Li S et al (2023) Analysis of the causes and countermeasures of artificial intelligence algorithm bias and health injustice. *Chin Gen Med* 26(19):2423–2427
7. Wang Y, Wang D, Wang H et al (2023) Algorithmic fairness: the logic and governance of bias in educational artificial intelligence algorithms. *Open Educ Res* 29(5):37–46
8. Zhang T, Hu J, Wang X et al (2023) Research on the global stress solving method of three-dimensional structure based on artificial intelligence. *Ship Mech* 27(2):238–249
9. Wang X, Dou J, Liu Z et al (2024) Review and prospect of the application of explainable artificial intelligence in power systems. *Power Syst Autom* 48(4):169–191
10. Liu R, Liu J, Tang Y et al (2024) Application and prospect of artificial intelligence-assisted design of energy-containing molecules. *Energy-Contain Mater* 32(4):408–421
11. Limna P, Jakwatanatham S, Siripipattanakul S et al (2022) A review of artificial intelligence (AI) in education during the digital era. *Adv Knowl Exec* 1(1):1–9
12. Rospigliosi P (2023) Artificial intelligence in teaching and learning: what questions should we ask of ChatGPT? *Interact Learn Environ* 31(1):1–3
13. Messeri L, Crockett MJ (2024) Artificial intelligence and illusions of understanding in scientific research. *Nature* 627(8002):49–58
14. Stahl BC, Antoniou J, Bhalla N et al (2023) A systematic review of artificial intelligence impact assessments. *Artif Intell Rev* 56(11):12799–12831
15. Enholt IM, Papagiannidis E, Mikalef P et al (2022) Artificial intelligence and business value: A literature review. *Inf Syst Front* 24(5):1709–1734

An Intelligent Probability-Based Supervised Localized Framework for Cardiologist Recommendation Using Predictive Analytics Approach



Suvra Mukherjee, Shruti Dutta, Kunal Anand, Lu Wang,
and Bharati Rathore

Abstract The health-world industry has seen a huge transition from the traditional healthcare system to the AI-based healthcare sector. However, the lack of a unified method for cost review prior to hospital visits means that overall costs for patient treatment remain high and difficult to manage. The ease with which patients may access modern technologies has contributed to the growing popularity of online medical systems in recent years. A method for early illness prediction based on symptoms is what we suggest. This technology using machine learning also schedules an appointment based on doctor availability and recommends hospitals in the area based on user ratings and facility locations. By automating the illness prediction, this system aims to decrease the time and effort spent by patients and clinicians.

Keywords Machine learning · Disease prediction · Doctor recommendation · Personalized healthcare · Healthcare delivery · Algorithmic approaches

1 Introduction

Every year, the healthcare sector produces gigabytes of data. The archived medical records contain a wealth of patient-related data. Retrieving pertinent data or high-quality medical care is a difficult but crucial endeavour [1]. Through the analysis of

S. Mukherjee · S. Dutta · K. Anand
Kalinga Institute of Industrial Technology, Deemed to Be University, Bhubaneswar, India
e-mail: kunal.anandfcs@kiit.ac.in

L. Wang (✉)
Xi'an Jiaotong-Liverpool University, Suzhou, China
e-mail: Stella.wang896@gmail.com

B. Rathore
University of South Wales, Pontypridd, UK
e-mail: Bharati.rathore@southwales.ac.uk

these vast amounts of data, we are able to forecast when the disease will strike and protect individuals. Thus, by offering precise and reliable illness risk prediction, an intelligent system for disease prediction contributes significantly to the management of disease and the preservation of people's good health. A smart healthcare system has two main objectives which include developing a Naive Bayes classification model that identifies the illness according to user comments and to construct a framework for a web interface for illness prediction and to recommend localized specialist for the specific cardio disease. Machine learning (ML) doctor recommended system's primary benefits usually centre on optimizing the distribution of medical resources, boosting patient experience, and improving healthcare services.

The following are some significant contributions:

- Data analysis is essential in the data-driven world of today. It aids businesses in utilizing data to make choices, streamline operations, and obtain a competitive advantage. By transforming unstructured data into valuable insights, data analysis enables businesses to recognize opportunities, reduce risks, and improve overall performance.
- The Naive Bayes algorithm is a supervised learning technique that solves classification issues. It is based on the Bayes theorem. Its primary application is in text categorization, where a high-dimensional training dataset is used. One of the most straightforward and efficient classification algorithms, the Naive Bayes classifier aids in the rapid development of machine learning models with rapid prediction capabilities.
- Being a probabilistic classifier, it makes predictions based on the likelihood that an object will occur.
- A collection of Java-written natural language analysis tools for text processing is offered by CoreNLP. It creates a comprehensive structured study of the most popular NLP processes from raw English text input. Annotators are a collection of language analysis tools that are integrated into CoreNLP [2]. Sentence splitting, tokenization, part-of-speech (POS) tagging, and sentiment analysis are a few of the pertinent annotators included in this programme. CoreNLP methods for physician suggestion based on prior user reviews have been researched and put into practice.

A doctor recommendation system's user interface (UI) in Fig. 1 is designed to be user-friendly and intuitive, enabling patients to find suitable healthcare professionals based on their specific needs. Search results where physicians who meet the search parameters are listed on the search results page. A brief biography, ratings, the doctor's name, and specialization may be included in each entry. To further refine the results based on variables like years of experience, years of proximity, availability, and languages spoken, filters may also be available. Physician profile where selecting a physician's name may direct the user to a Website featuring additional information about that physician. This might have a button for making an appointment as well as information on their training, credentials, and patient testimonials. Appointment booking where users would be able to select a time and day for their appointment, provide personal details, and explain any symptoms or reasons they are visiting. Review and ratings where users will be asked to rate their experiences. A broad

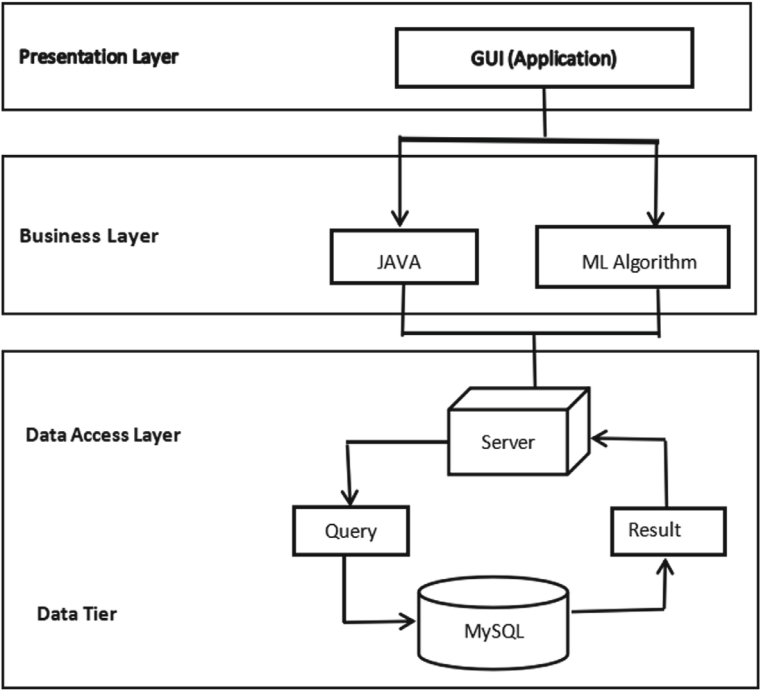


Fig. 1 Architecture diagram for a doctor recommendation model

term encompassing disorders affecting the heart or blood vessels is cardiovascular disease (CVD). Atherosclerosis, or the build-up of fatty deposits inside the arteries, and a higher risk of blood clots are typically linked to it. Damage to arteries in the brain, heart, kidneys, and eyes is another possible association. Maintaining a healthy lifestyle can help avoid cardiovascular disease (CVD), which is one of the leading causes of mortality and disability.

2 Related Work

This study is the result of our examination of the below publications from various additional sources. We looked at the link between various algorithms and how well they performed for illness prediction in various settings. Binal et al. [3], healthcare decision support system for swine flu prediction using Naïve Bayes classifier, focuses on the aspect of medical diagnosis by learning patterns through the collected data for swine flu using Naïve Bayes classifier for classifying the patients of swine flu into three categories (least possible, probable, or most probable), resulting into an accuracy of nearly 63.33%. Datasets used for this classification were limited in number. Marcelo et al. [2], a collaborative filtering method based on user reviews suggests a

method that generates item descriptions by utilizing user reviews. After processing the reviews with the CoreNLP tool, an algorithm is created to calculate the similarity of the items, which is then employed in a collaborative filtering technique based on k-nearest neighbours. Huang et al. [4] In their paper “Predicting Disease by using data mining based on healthcare information system,” Shengyong W. et al. detail the trials conducted on data mining for the purpose of disease prediction using a substantial amount of actual hypertension medical records. In this essay, three algorithms: ensemble of five J-48 classifiers, J-48, and Naïve Bayes. The accuracy of 83% was almost the same for J-48 and Naïve Bayes in this case. Manjusha et al. [5] In their paper, Prediction of diverse dermatological problems using Naive Bayesian classification, Manjusha et al. list eight possible diseases based on the patient’s characteristics. The database’s hidden knowledge is extracted by the system. In addition to dermatological illnesses, the system can anticipate other diseases. Nikhate and Jonnalagedda [6] This paper elucidates the necessity of conducting research and developing a framework that will enable an end-user to readily foresee the everlasting disease without the need to consult a physician or other specialist for diagnosis. By using various machine learning model tactics and analysing patient symptoms, it is easy and useful to identify various diseases. This section of the study results in the accuracy utilizing various algorithms, including random forest (RF) with an accuracy of 95.28%, decision tree (DT) with a 90.2% accuracy, and NB with an 88.08% accuracy. Sandhya [7] This paper used supervised learning algorithms to provide a prognosis for heart disease. Among the methods employed are Naïve Bayes, KNN, and SVM. There are 3000 objects in the dataset with 14 features as well. The extensive literature review revealed that the majority of studies used the illness dataset, which has just 303 objects with 14 attributes. The best results are produced by Naive Bayes, which takes less time and has a high accuracy of 86.6; decision tree and KNN provide 78.69% and 77.85% percent accuracy, respectively. Chattar [8] The purpose of this work is to streamline decision assistance by utilizing machine learning technologies to enhance the astute therapy framework. This extensive document discusses how to diagnose heart illness by keeping an eye on a patient’s heartbeat. The framework gives you the ability to customize your pulse’s specifications. After establishing these boundaries, a person might start monitoring their heartbeat and when someone’s heartbeat exceeds a certain threshold, they are alerted to the possibility of a high pulse and heart attack or coronary failure. Authors Ahmed M. Alaa and Senthil Kumar Mohan used a random hybrid forest to achieve 88.7% accuracy after experimenting with a variety of variables. The doctor’s schedule was regarded by the authors of [9] as the primary component of their model. While time is a crucial consideration, a patient searching for a doctor in an emergency also takes other factors into account. Despite the fact that they employ extremely complex neural networks, no prerequisites, such as level of education, number of years of experience, or workplace hygiene, were taken into consideration when selecting physicians. Vyas [10] This paper addresses traditional supervised binary classification in which the dataset contains a number of attributes. The dataset comprises the concentration of plasma glucose. Several components, each with distinct characteristics, are utilized to identify individuals who have the illness. In order to solve the issue, we should examine

the data, make any required adjustments, implement an ML method, train a model, examine the output of the ready model, and iterate through various algorithms until we find the most precise solution. When software or Websites are improved, it is essential to understand the framework requirements and collect the necessary data in order to establish a connection between the supplier and the client. Wang and Xu [11] The authors put out an algorithm for evaluating doctors based on collaborative filtering with the aim of achieving customized medical suggestions in a setting with a lot of data. The writers took into account the doctors’ qualifications to suggest a reasonably priced physician; however, they neglected to take the doctors’ reliability into account. But it is crucial that the recommendation system suggests a reliable and respectable physician.

3 Proposed Methodology

There are various techniques for illness prognosis. Nonetheless, research has been done on heart-related conditions, and a risk level has been computed. These techniques are not typically applied to illness prediction, though. Thus, general disease prediction is aided by the smart healthcare system. Sometimes, you or a family member may require immediate medical attention. The proposed model is shown in Fig. 2. The suggested smart healthcare system framework uses a data mining approach called the “Naive Bayes classifier” to develop the expectation framework. Devasia et al. [12] In order to ascertain the precise symptom expectation, this system has a higher quantity of data indexes and features that are actually acquired from expert data. A number of artificial intelligence and data mining methods rely on the “Naive Bayes or Bayes” rule. Precognitive ability models are created using the norm. By establishing the connection between the objective (i.e. subordinate) and other variables, it gains from the “proof”.

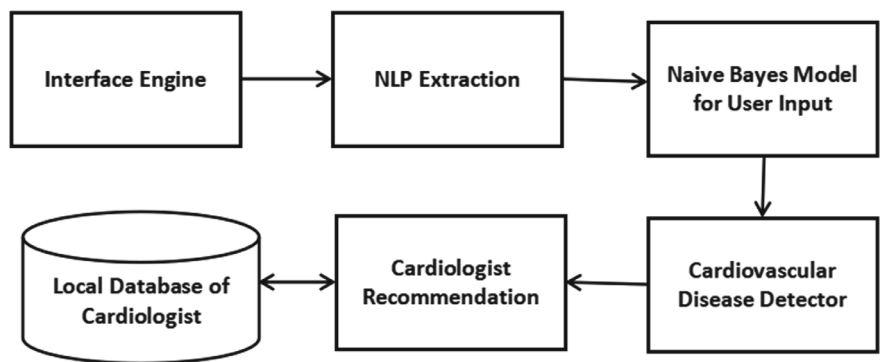


Fig. 2 Model diagram for cardiologist recommendation using Naive Bayes classifier

In machine learning for physician recommendation, an interface engine is generally used to describe a system or software element that makes it easier for various software programs and data sources to integrate and communicate with one another. It is essential to the seamless flow of data between different parts of a healthcare system and to the improvement of the recommendation system's effectiveness. The following is an example of how an interface engine may be applied to machine learning to suggest physicians:

- Health information from electronic health records (EHRs), patient databases, and appointment scheduling systems, among other sources, can be integrated using an interface engine.
- In order to enable the machine learning model to provide more intelligent and individualized physician suggestions, it aids in the creation of a unified picture of patient data [1, 13, 14].
- For data exchange and storage, healthcare systems frequently employ several standards and formats. These variations can be handled by an interface engine, which guarantees system compatibility.
- The accuracy of the recommendation system is improved by this interoperability, which is essential for obtaining and merging pertinent data on physicians, patients, and their medical histories.
- In order to give the recommendation system the most recent information, interface engines can process data in real time. This is particularly crucial in the healthcare industry, as good decision-making depends on fast and reliable information.
- Strict security and privacy laws must be followed by healthcare systems [15]. Healthcare privacy regulations like the Health Insurance Portability and Accountability Act (HIPAA) may be adhered to by the recommendation system and data shared securely with the help of an interface engine.
- Between the user interface and the machine learning recommendation system, an interface engine can also act as a communication layer [16, 17]. It makes it easier for end users—like consumers or healthcare providers—to see suggestions from doctors.
- The recommendation model receives input from user preferences and feedback via the interface, which increases the model's accuracy.

An effective way to improve a machine learning system's capacity to comprehend and make use of unstructured text input is through the extraction of natural language processing (NLP). NLP extraction may be used in this situation in a few different ways:

- Relevant data may be extracted using natural language processing (NLP) from unstructured clinical notes and other free-text areas in electronic health records. In addition to providing a more thorough picture of a doctor's profile, this can also contain information about the doctor's area of specialty, contact with patients, and any outstanding accomplishments.

- Reviews and comments from patients on physicians may be analysed using natural language processing [18]. The technology is able to deliver more detailed and individualized suggestions by analysing sentiment, subjects, and key phrases extracted from reviews to learn about the advantages and disadvantages of each doctor.
- NLP can enhance the recommendation system's search functionality. The suggested physicians will closely match the user's requirements and preferences since the system can get more accurate and pertinent answers by comprehending the semantics of user queries.
- Entities like medical illnesses, names of doctors, and therapies stated in textual data may be identified and extracted using NER. This makes it easier to link information about physicians who specialize in particular medical conditions and adjust advice appropriately [19].
- Through the use of natural language processing (NLP) and topic modelling approaches, the system is able to extract latent themes about physicians, medical services, and patient experiences from textual data. This may lead to a more sophisticated comprehension of user preferences and physician profiles.
- Natural language processing (NLP) can assist in comprehending user inquiries and delivering solutions in a natural and contextually appropriate way when the recommendation system incorporates interactive chat or voice interfaces. In addition to guaranteeing efficient communication between the system and the user, this improves the user experience.
- Relevant data may be extracted from patient histories and medical records with the use of natural language processing (NLP). Through a grasp of the context and information included in these records, the system is able to provide more knowledgeable suggestions based on a physician's expertise with certain diseases or treatments.
- Information on a doctor's training, credentials, and associations may be extracted from textual sources using natural language processing (NLP). Evaluation of a physician's qualifications and compliance with a patient's unique medical needs can both benefit from this information.

A Naive Bayesian classifier is a model joint probability distribution over a set of stochastic variables. Instances of the classification problem under study are presented to the classifier as a combination of values for the feature variables; the classifier then returns a posterior probability distribution over the class variable. Learning such a classifier amounts to establishing the prior probabilities of the different classes and estimating the conditional probabilities of the various features given each of the classes.

A system that determines people at risk of cardiovascular diseases (CVD) based on numerous health markers and medical data is referred to as a cardiovascular detector in the field of machine learning for doctor recommendation. The accuracy of physician recommendations for patients with cardiovascular issues can be improved by including such a detector into a recommendation system. Here is an example of how to use it:

- The cardiovascular detector evaluates the risk of cardiovascular illnesses by analysing patient health data, such as vital signs, medical history, and findings from diagnostic tests. The patients who could benefit from specialist cardiovascular treatment can then be identified by using the results of this risk assessment to create patient profiles.
- Physicians that specialize in cardiology or cardiovascular medicine might be recommended based on the amount of cardiovascular risk that has been found. By guaranteeing that recommendations are customized to the unique health needs of each patient, this increases the possibility of prompt and focused action.
- Through its recommendation of physicians with a focus on cardiovascular health, the system can help patients, and cardiologists communicate more easily. This facilitates the process of patient referral and matches them with qualified medical practitioners to handle their particular cardiovascular issues.
- The system has the ability to offer instructional materials and information regarding cardiovascular health in addition to physician advice. Preventive interventions and lifestyle adjustments are encouraged as a result of people being empowered to make educated decisions and actively engage in their treatment.
- When substantial variations in cardiovascular risk are found, the cardiovascular detector can produce messages or alarms for individuals as well as healthcare professionals. Better management and prevention of cardiovascular illnesses are facilitated by the prompting of timely actions and follow-ups.
- System adherence to stringent data security and privacy requirements, such as HIPAA, is imperative, considering the delicate nature of cardiovascular health information. The integrity and confidentiality of patient data are guaranteed by the implementation of strong security protocols.

4 Result and Analysis

The Naive Bayes classifier's probability calculations in Fig. 3 indicate the most likely sickness, with an expected accuracy of more than 90%. There is a distinguished graph that shows the representation of top five specialists recommended based on user-selected filters and the anticipated disease. The statistical result according to other algorithms with respect to Naive Bayes is as follows:

The graph in Fig. 3 for the analysis done for finding the accuracy of the recommendation of cardiologists shows that Naive Bayes is providing the results with the highest accuracy of 90.7% with the rest being below it. The lowest accuracy went to 75.7% which is of logistic regression.

Table 1 shows the scalability rate of cardiologists over a specific area. It represents that even if the span of area is increasing the delay for generating response or recommendation is not above 0.9 s. This shows that Naive Bayes not only gives accurate results within short span of time but also even if the area size of a specific location or the number of cardiologists are increased, it will always provide the results of recommend within the shortest possible time and with most possible accuracy.

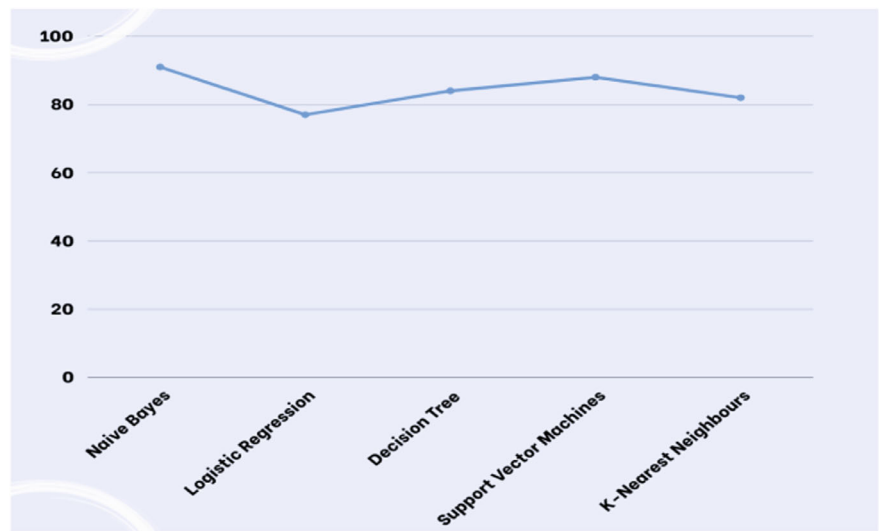


Fig. 3 Graph of result analysis using Python

Table 1 Accuracy and scalability assessment with localized cardiovascular dataset model for Naive Bayes-based classification

Data size	Accuracy (%)	Model response delay (s)
10	90.43	0.3
20	90.35	0.32
30	90.33	0.37
40	90.30	0.4
50	90.25	0.5
60	90.11	0.66
70	90.01	0.7
80	89.95	0.78
90	89.84	0.85
100	89.81	0.9
Mean	90.138	0.578

5 Conclusion

Research and practice have been done on the application of data mining techniques for disease prediction from a large number of real-world medical records and CoreNLP algorithms for physician recommendation from user reviews. By adding additional symptoms and increasing the quantity of cases for testing and training, the system can be made even better. A required nutritional diet plan can also be recommended by taking into account the number of calories, steps taken, sleep quality, and other

medical profiles through wearable technology. In the event of a risk, timely alerts and notifications can be delivered to the user's guardian as well.

References

1. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digital Health* 10:20552076241256732
2. Jain S, Tripathy HK (2024) Machine learning methods for the timely identification of Autism spectrum disorder in Toddlers. In: 2024 International conference on emerging systems and intelligent computing (ESIC), pp 515–520. IEEE
3. Thakkar BA, Hasan MI, Desai M (2010) Health care decision support system for swine flu prediction using naïve bayes classifier. In: 2010 International conference on advances in recent technologies in communication and computing, pp 101–105
4. Huang F, Wang S, Chan C (2012) Predicting disease by using data mining based on healthcare information system. In: 2012 IEEE international conference on granular computing, pp 191–194
5. Manjusha KK, Sankaranarayanan KP, Seena P (2014) Prediction of different dermatological conditions using
6. Nikhate Y, Jonnalagedda MV (2020) Survey on heart disease prediction using machine learning
7. Sandhya BS (2020) Disease prediction and hospital recommendation using machine learning algorithms. *Int J Res Appl Sci Eng Technol* 8:2054–2063
8. Chattar SP (2021) IOT based heart disease prediction. *Int J Res Appl Sci Eng Technol* 9:123–124
9. Choudhury AG, Choudhury A, Subramaniam U, Balamurugan SA (2021) HealthSaver: a neural network based hospital recommendation system framework on flask webapplication with realtime database and RFID based attendance system. *J Ambient Intell Humaniz Comput* 13:4953–4966
10. Vyas M (2020) Software for disease prediction. *Int J Res Appl Sci Eng Technol* 8:1213–1217
11. Wang C, Xu M (2017) The research of doctors recommendation algorithm based on clustering and collaborative filtering
12. Devasia T, Vinushree M, Hegde MV (2016) Prediction of students performance using educational data mining. In: 2016 International conference on data mining and advanced computing (SAPIENCE), pp 91–95
13. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
14. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-detector: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
15. Pranjali P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. In: AIP Conference Proceedings, vol 2919, no. 1. AIP Publishing
16. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system. *Internet of Things* 24:100978
17. Mishra S, Volety DR, Bohra N, Alfarhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306
18. Pradhan SR, Mishra S, Tripathy, HK, Brahma B, Gobinath R, Sobti R (2024) Critical application feasibility of predictive learning in autonomous vehicles. In: International conference on innovative computing and communication, pp 371–383. Springer Nature Singapore, Singapore

19. Chakraborty S, Mishra S, Tripathy HK (2022) COVID-19 outbreak estimation approach using hybrid time series modelling. In: International conference on innovations in intelligent computing and communications, pp 249–260. Springer International Publishing, Cham
20. Palaniappan S, Awang R (2008) Intelligent heart disease prediction system using data mining techniques. In: 2008 IEEE/ACS international conference on computer systems and applications, pp 108–115
21. Ganu G, Kakodkar Y, Marian A (2013) Improving the quality of predictions using textual information in online user reviews. *Inf Syst* 38:1–15

Optimizing the Performance of Diabetes Risk Prediction Using Ensemble Learning Techniques



Sukurat Salam, Glen Hopkinson, Christopher G. Udomboso,
and Serifat Folorunso

Abstract Early detection is a critical strategy for reducing diabetes prevalence across all age groups. However, given the likelihood of delays in accessing medical services for various reasons, the development of user-friendly machine-learning models becomes essential. The study aims to develop a diabetes risk prediction optimized model in individuals using modifiable behavioral risk factors. Advanced data pre-processing approaches were used to increase the data quality and the prediction performance of the trained models was improved using several ensemble methods to combine best-performing ML models. Outlier detection using the iForest approach, imbalance multi-class handling, and selection of relevant behavioral risk factors were achieved at the pre-processing stage. The development stage of the model starts with training some supervised classifier models, followed by applying different ensemble methods for model performance optimization. Out of which the stacking ensemble approach performs best with a K Fold score of 97.87%, average accuracy of 99.40%, ROC AUC score of 99.70%, and F1-score of 98.41% and the optimized model was deployed on a web-based interface for real-time diabetes risk assessment. This study contributes a powerful prognostic framework, which significantly advances diabetes risk prediction and proactive healthcare management.

Keywords Optimization · Diabetes · Ensemble Techniques · Risk Analysis

S. Salam (✉) · G. Hopkinson · S. Folorunso
Data Science, Teesside University, Middlesbrough, UK
e-mail: salamsukurat@gmail.com

G. Hopkinson
e-mail: g.hopkinson@tees.ac.uk

C. G. Udomboso
Computational Statistics, University of Ibadan, Ibadan, Nigeria
e-mail: cg.udomboso@ui.edu.ng

1 Introduction

The significant rise in diabetes worldwide has become not only concerning but also a demanding one in the health care system, given that the disease raises the risk of other conditions like ischemic heart disease and stroke [1]. The rate of increase in Type I and Type II diabetes is becoming high worldwide, with Type II diabetes reaching epidemic proportions and being labeled a global pandemic by the WHO as early as 2006. The most recent statistics show that the current prevalence rate is 6.1%, making the condition one of the top ten major causes of death and disability worldwide [10].

There are several risk factors associated with diabetes, categorized into non-modifiable and modifiable as stated by American Heart Association (AHA). The identified non-modifiable factors include family heredity, age, race and ethnic background, and gestational diabetes. In contrast, the non-modifiable factors an individual can do something about are weight, physical activity, blood pressure, cholesterol levels, smoking, diet, alcohol consumption, stress, and sleep. For this research, some of the modifiable risk factors were considered to predict the risk of diabetes in an individual [13]. According to WHO statistics, 1.5 million deaths are directly attributed to diabetes every year out of about 422 million that have it, hence the need for early intervention. Early detection of diabetes is a crucial aspect of the healthcare system, this is because early treatment will prevent serious complications and the occurrence of the disease. As a result, developing an accurate diabetes predictive model based on patient records and some identified behavioral risk factors becomes important [18]. Machine-learning techniques have gained much popularity in their application to developing models for disease risk prediction, treatment, prognosis, and management [2]. However, the aspect of model optimization techniques to get the best performance (accuracy and precision) is a very crucial part that demands more research. In addition, the developed model should be made readily available for easy usage to determine the risk of diabetes using some behavioral factors. This was accomplished by hosting the model on the web in a graphical user interface format.

This research developed a machine-learning framework that utilizes preprocessing techniques like multi-class methods of addressing the class imbalance in diabetes data, feature scaling and selection, cross-validation, and a combination of supervised classification models known as ensemble methods to optimize model performance for risk of diabetes prediction. The prediction will be based on behavioral characteristics such as smoking, physical health, nutrition, body mass index (BMI), mental health, age, and other factors [12]. In addition to the model's development, a web-based graphic user interface (GUI) was built to allow users access.

2 Related Work

2.1 *Machine Learning for Diabetes Prediction*

A crucial approach for lowering the prevalence of diabetes across various age groups centers around early detection. However, given the potential delays in accessing medical facilities due to various factors, the development of user-friendly machine-learning models becomes important. Over the past year, the prevalence rate of diabetes has attracted numerous computational efforts mainly around the adoption of machine-learning algorithms in diabetes research geared towards helping clinicians make quick and meaningful diagnostic decisions [16], some of which will be reviewed in this section.

A recent study by [5], presents an optimized stacking-based hybrid machine-learning approach for predicting early diabetes approach using the PIMA Indian Diabetes (PID) dataset, in his work the preprocessing steps he adopted include mean imputation of missing values, balancing data using SMOTE techniques, normalizing features, and using stratified train–test split. The proposed approach uses Logistic regression (LR), naïve Bayes (NB), AdaBoost with support vector machines (AdaBoost + SVM), artificial neural network (ANN), and k-nearest neighbors (k-NN) as base learners (level 0), while random forest (RF) meta-classifier serves as the level 1 model to combine their predictions and the model give an accuracy of 94.21% for the PID dataset but the result relies on the medical diabetes data and not the behavioral risk factors that will be considered in this research. Another study by [11] employed a modified weighted sum method based on survey data and the second approach used a machine-learning feature selection method. They applied three machine-learning classification algorithms, KNN, decision tree (DT), and NB to a dataset of 16,698 instances that covers the essential body features of multi-aged people, this was to enhance a system that could approach what real-life situations usually demand from accuracy, sensitivity, and complexity time. In a bid to optimize the diabetes prediction model using several pre-processing techniques, [16] hypothesized that adopting feature selection and missing value imputation methods can improve the performance of the machine-learning models in diabetes prediction which uses a Spearman correlation feature selection and polynomial regression for missing value imputation.

2.2 *Ensemble Classifier for Disease Prediction*

To classify the risk of diabetes in individuals, researchers have employed the use of ensemble learning to optimize the performance of disease predictive models.

A study by [3] employed the use of ML optimization techniques to predict chronic kidney disease using data preprocessing and feature extraction. In another study by [19] where they conducted diabetes prognosis on the PIDDD dataset using the ML and

DL approaches, The outcome of the research shows that RF from machine learning and ANN from neural network category have the highest accuracy of 74% and 76% with F1 score of 74% and 79% respectively. In other to enhance the performance of the diabetes prediction model, [12] used bagging ensemble techniques to train the same PIMA dataset and they achieved an accuracy of 85.22% which is the highest compared to other individual supervised learning models.

A review by [2] carried out a bibliometric analysis of 1216 publications using data from the Scopus and Web of Science databases. One of the suggested areas of future direction is an ensemble-based approach which is one of the areas this current research focuses on. Alternatively, [9] in their study use a vote-based classifier to predict diabetes which optimizes the diabetes prediction by 2%. Their study uses two approaches and compares them to specific indicators, the first one combines logistic regression, PCA, and K-means and the second approach uses PCA, K-mean, and negative classification techniques to predict diabetes.

In summary, the extraordinary advances made in previous research by employing ensemble methods set the path for the current analysis. The approach, along with carefully defined pre-processing procedures was used to build an optimized predictive model for the early detection of the risk of diabetes in individuals using some adjustable behavioral risk factors. These efforts not only demonstrate innovative development in the field but also highlight the possibility of greatly improving the early detection of diabetes using readily available data of individuals before proceeding with any laboratory test.

3 Methods

3.1 Dataset

The original dataset for this research was collected from a nationwide survey system conducted in the United States by the Centers for Disease Control and Prevention (CDCP) using the Behavioral Risk Factor Surveillance System (BRFSS). The data was collected through telephone interviews on a wide range of health behaviors, chronic conditions, and prevention services for adults aged 18 and above. The 2015 dataset by [22] was downloaded from Kaggle for this study. The class of this target variable was highly imbalanced with no diabetes class having 82.7%, pre-diabetes at 2.0%, and diabetes class having 15.3% after removing the duplicates.

3.2 *Data Pre-processing*

This technique includes missing values handling, removing duplicates, data scaling/normalization, data balancing, feature selection, and visualization to ensure data quality for increasing performance of the model, aside from selecting the appropriate modeling technique [18].

This study employed an ensemble-based method called isolation forest (iForest) for outlier detection. It consists of a forest, a group of binary trees constructed from the random property of the dataset. The method has a linear time complexity with high precision compared to other methods [8]. The 23,899 duplicated observations were eliminated and data normalization was carried out using the `StandardScaler()` function in the Scikit learn package [4, 20].

3.2.1 **Feature Selection**

Feature selection can help reduce data redundancy and model overfitting [17]. It can also save computing time, improve the prediction performance of the model, and provide a deeper understanding of the data [7].

For this current research, the related features that are relevant to the subject have been selected out of 303 features present in the original dataset, this was done with the recommendation of the National Heart, Lung and Blood Institute and Centers for Disease Control and Prevention (CDCP) [22].

This research experiments with three selection criteria and their result was compared to select the optimal subsets for the modeling. The correlation criteria method is a correlation-based feature selection and a multivariate approach that evaluates the correlation between individual features and output variables. The ranking assigned to individual features is based on the Pearson correlation metrics. The variable with a very low correlation with the class would be tagged as irrelevant and removed [7]. The second method used was mutual information which employs the entropy measure as the ranking criteria for selecting the feature. It evaluates the features based on their information gain and uses one variable at a time [6]. Feature importance, a metric obtained from tree-based models like decision trees and random forests was used as the third method.

3.3 Data Modelling

This was initialized by training several multiclass classification algorithms which are explained as follows:

- **KNN:** K Nearest Neighbor is machine learning used for regression and classification tasks, it works by selecting K-nearest neighbors using distance metrics.
- **SVC:** Support Vector Classifier is a simple algorithm that can handle both classification and regression tasks, it is popularly used in medical diagnosis due to its excellent generalization performance [15].
- **DT:** The Decision tree model is implemented using the recursive binary splitting to build a classification tree, the quality of a particular split is evaluated using an Entropy value or Gini index [21].
- **MLR:** An extension of the two-class logistic regression approach for classes greater than two is known as Multinomial Logistics Regression.
- **GNB:** This classification model is based on Bayes theorem, which is learned using the probability of an observation belonging to an output variable class.

3.3.1 Ensemble Learning Model

The ensemble learning method has proven to be an effective ML technique for improving the performance of a predictive model, this was achieved by combining different sets of base learner models [14]. For the current research, the most performing models out of the supervised classification models considered were combined to form a single optimized model for the early prediction of diabetes risk using the three ensemble learning techniques [18]. The details of this method are given in the following sub-section.

Voting

Voting is assumed to be one of the simplest ensemble learning techniques, it combines multiple predictions from different (heterogeneous) classifiers or the same (homogeneous) classifiers with different parameters to predict the class with the maximum votes [18]. The prediction by each model is referred to as a “vote”. The mathematical representation of voting is given below. Given n classifiers $C_1, C_2, \dots, C_n$,

$$\hat{y} = \arg \max \sum_{i=1}^n w_i P_{ij}$$

where, $\hat{y}$ is the final prediction, w_i is the weight of the i th classification, and P_{ij} is the probability estimate from the i th classification rule for the j th class.

Bagging

The bagging ensemble method uses multiple classifiers trained parallelly on different subsets of the data acquired through the bootstrap method. After this the resulting predictions are aggregated through averaging or voting to obtain the final decision, this will help to reduce overfitting and variance [12].

Stacking

This is defined as a two-classification ensemble technique with level 0 (base model) and level 1 (meta model) that generates a new training dataset for the final prediction. This makes it different from other ensemble methods like voting and bagging which uses voting and aggregation for final prediction. In stacking each classifier in the first level is applied to the training set and the prediction output from this level is used as the new training set in the second level. A new classifier model called meta-model will be used to learn from the new feature generated from the base models in the second level to obtain the final prediction.

4 Model Implementation on Web Application

This section gives details of how the best-performing machine learning model was deployed into a web-based application for easy usage. For this research, three web pages were created which include the index page representing the home page where the details of the web application were given. The second page is the diagnosis page which contains the data collection form field where the user will answer some questions on the risk factors responsible for diabetes prediction. After clicking on the result button on the diagnosis page to submit the responses, the result page will be displayed to give the diabetes status of the user based on the responses.

Flask is a Python framework that bridges the gap between machine learning model development and its accessibility on the web. Its process includes embedding a trained model within a server, receiving that data, processing the data through the model, and giving a prediction result in response to the user queries. Flask helps the implementation of machine learning solutions in the practical realm (Fig. 1).

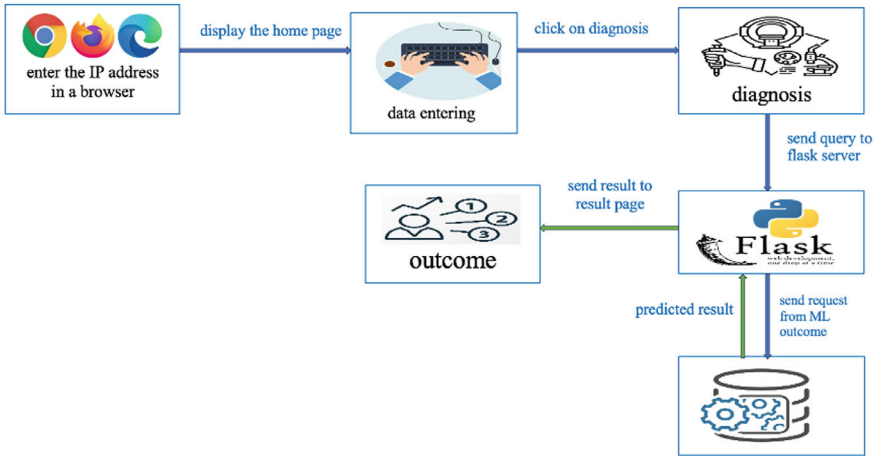


Fig. 1 Workflow of the diabetes web application

5 Result and Discussion

5.1 Development of Optimized Classification Model

This section gives the detailed result of the development process, advanced pre-processing techniques, and ensemble methods utilized in creating the optimized diabetes risk prediction model.

The model development process was initiated by conducting some exploratory data analysis to get valuable insight into the distribution and correlation of variables in the dataset. The count plot of features gives the distribution of all the features in the dataset, showing that some of the features are numeric and some are categorical. It was also observed that some of the independent variables are significantly correlated with each this include general health and physical health (0.52), physical health and difficulty walking (0.47), income and education (0.42) and high blood pressure and age (0.34) (Figs. 2 and 3).

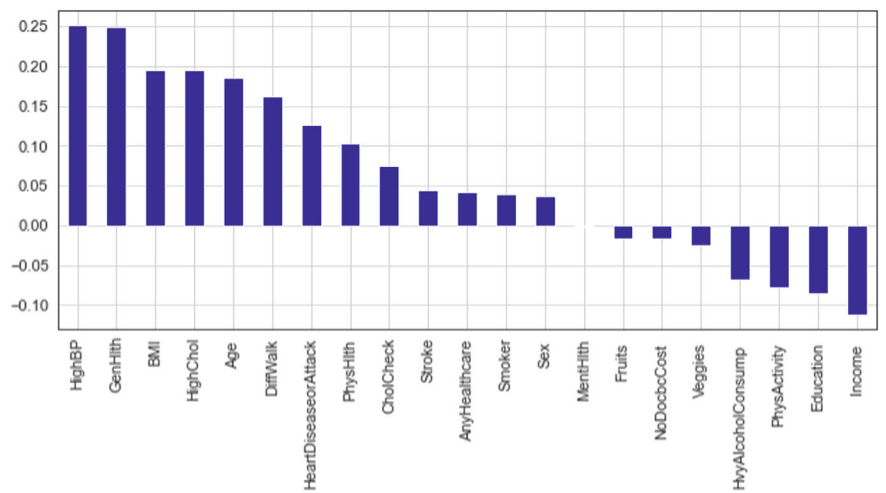


Fig. 2 Correlation of the features with diabetes class

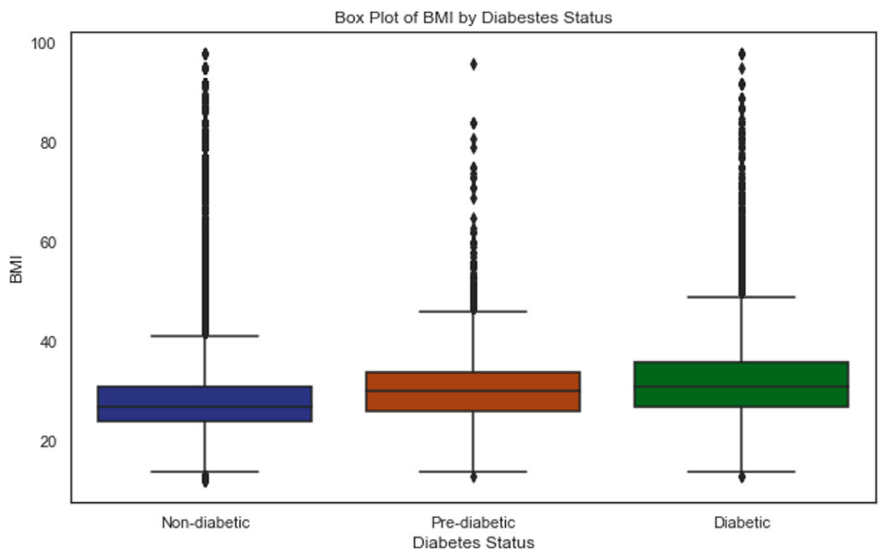


Fig. 3 Boxplot of BMI by diabetes class

5.1.1 Addressing Class Imbalance

The class distribution of the diabetes status is skewed, it becomes important to balance the class to improve the model performance of the minority class predictions. The SMOTE and its derivative oversampling method were experimented with to

Table 1 Result of different oversampling methods on baseline classifier

Method	X_train_shape	y_train_shape	Train_score	Test_score	F1-score
SMOTE	(552,204, 21)	(552,204)	0.896653	0.855163	0.848722
ADASYN	(550,903, 21)	(550,903)	0.893751	0.852612	0.845656
SMOTETomek	(546,644, 21)	(546,644)	0.899770	0.859056	0.852507
SMOTEENN	(445,372, 21)	(445,372)	0.974788	0.951172	0.945256

Table 2 Baseline model performance before and after SMOTEENN

Model	Accuracy	Precision 0	Precision 1	Precision 2	Recall 0	Recall 1	Recall 2	F1-score
Before	0.83	0.86	0.04	0.37	0.95	0.00	0.18	0.382563
After	0.97	0.96	0.97	0.96	0.94	0.99	0.96	0.96

address the imbalance distribution of the target variable class. The combined SMOTEENN method which is the combination of SMOTE (Synthetic Minority Oversampling Technic) and ENN (Edited Nearest Neighbour) performs best on the baseline classifier KNN as shown in Table 1.

The classification outcome before and after application of SMOTEENN is given in Table 2. The result of the baseline classifier KNN before oversampling has an overall accuracy of 83%, but considering the performance of individual classes, class 1 which is the prediabetes group has a precision of 4% which implies that out of all the instances, the model correctly identifies class 1 was predicted correctly in 4 out of 100. This is followed by class 2 which is the diabetic group with a precision value of 37% and class 0 (non-diabetes) has a correct prediction rate of 86%. The recall value which measures the ratio of correctly predicted instances of each class to the total instances of each class in the data set has similar performance with prediction. The recall of class 1 is 0%, followed by class 2 with a value of 0.18 and class 0 with a recall value of 95%. This gives a better understanding of the performance of the model on each class, its strengths and weaknesses, the model performs better on the majority class and poorly in predicting the minority class.

The model performance was improved on the classes by oversampling the minority using SMOTEENN class 1 (pre-diabetes) and class 2 (diabetes) to a precision of 97% and 96% respectively. Similarly, the recall value for the pre-diabetes class becomes 99%, followed by the diabetes class at 96% and the non-diabetes group at 94%.

5.1.2 Identification of Behavioral Risk Factors

To identify the most important features that influence the prediction outcomes, three data-driven feature selection methods were utilized. After careful consideration, a subset of 13 features was designated as the optimum selection for classification purposes within the web interface. These 13 features, derived through a thorough amalgamation of correlation, feature importance, and mutual information analyses, stand as the most influential behavioral risk factors driving predictive outcomes of diabetes risk.

5.1.3 Training of Supervised Multiclass ML Models

Multi-class supervised machine learning classifiers were trained using the 80% training subset of the resampled data, where each of the supervised models was trained individually and optimized their parameters using hyperparameter tuning and fivefold cross-validation to ensure the robustness of the predictive model performance. The detailed results are in Table 3. From the performance of supervised model analysis shown in Table 3, the KNN classifier has the best performance with the proportion of correctly predicted instances among the total prediction in the training set of 96.54% and that of the testing set is 98.91%. The K fold score which is the cross-validation score that determines the performance of the model on the unseen data is 95.09%, the F1 score which is a form of an average of recall and precision is 96.39%, and finally, the ROC AUC score which reflects the discriminative ability of the model across varying thresholds is 99.10%.

However, the overfitting challenge observed from the performance of a single model classifier was addressed using the combination of two or more classifiers called the ensemble method.

Table 3 Table of supervised model outcome

Method	Testing score	Training score	K fold score	F1-score	ROC AUC score
KNN	0.965425	0.989112	0.950979	0.963865	0.990886
MLReg	0.564271	0.567234	0.567224	0.580256	0.758558
GNBayes	0.552314	0.553929	0.553785	0.551202	0.741693
DTree	0.880639	1.000000	0.875381	0.878688	0.907840
SVC	0.771230	0.781649	0.766770	0.771812	NaN

Table 4 Table of ensemble models outcome

Method	Test score	Train score	K fold score	F1 score	ROC AUC score
XGBoost	0.958719	1	0.953514	0.955362	0.994797
AdaBoost	0.958719	1	0.953514	0.955362	0.994797
Random forest	0.957731	0.999997	0.951026	0.957731	0.994648
Extra tree	0.977321	1	0.971315	0.977321	0.998283
Bagging	0.93583	0.999497	0.927324	0.933371	0.988484
Voting	0.977904	1	0.969698	0.976595	NaN
Stacking	0.985791	0.998797	0.978723	0.984083	0.997023

5.1.4 Ensemble Method for Model Performance Optimization

The strength of individual classifiers from the previous stage was leveraged using the ensemble method to optimize the predictive performance and maximize the generalization of the model to unseen data. By using the ensemble method to combine several classifiers, a reliable and effective diabetes risk prediction model was produced. With a balanced trade-off between precision and recall, the optimized model demonstrated improved predictive accuracy and significant gains over individual classifiers as shown in Table 4. The best ensemble method is the stacking method with a good balance trade-off between the training and testing accuracy with values of 99.88% and 98.58% respectively. The K fold score which is the cross-validation score that determines the performance of the model on the unseen data or different data subset is 97.87% which implies the robustness of the model. Also, the F1 score which is a form of an average of recall and precision is 98.41%, and finally ROC AUC score which reflects the discriminative ability of the model across different classes is 99.70% (Fig. 4).

6 Conclusion and Recommendations

6.1 Conclusion

The current research aimed to develop an optimized model for the prediction of diabetes risk in individuals based on some identified modifiable behavioral risk factors. Through the use of advanced pre-processing techniques and ensemble machine learning models, the study successfully achieved the goal of creating an efficient predictive model for the risk of diabetes and deploying the model on a web-based application. The risk factors identified are high blood pressure, cholesterol levels, BMI, heart disease attack, general heart status, difficulty in walking, age, education level, and income which demonstrated significant relationships with diabetes status.

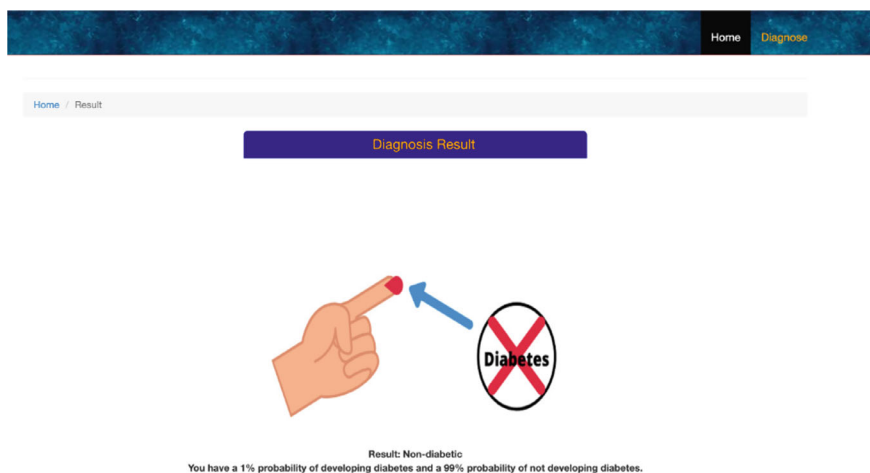


Fig. 4 Result page interface for non-diabetic users

In addition, pre-processing steps including handling data skewness, scaling features, encoding categorical variables, and addressing class imbalance were carried out to improve the data quality. Advanced techniques like Isolation Forest for outlier detection and SMOTEENN for handling class imbalance significantly improved model performance. Feature selection methods help in identifying crucial risk factors driving diabetes prediction. These methods included correlation analysis, mutual information, and feature importance, which were combined in a selection of 13 influential behavioral risk factors for diabetes prediction.

The model development process comprised training various supervised machine learning classifiers to select the best-performing set which includes KNN (96.54%), Decision Tree (88.06%), and SVC (77.12%). Ensemble methods further optimized predictive accuracy, notably showcasing stacking methods which is the combination of KNN, decision tree, and extra tree (99%) and the extra tree (98%) classifier as top performers.

The research has successfully developed an efficient model for predicting diabetes risk, integrating various advanced pre-processing and ensemble techniques. It has also identified some modifiable behavioral risk factors that can influence the outcome of diabetes status of individuals using a data-driven approach. This goes in line with those diabetes risk factors identified by the American Heart Association, 2021. A user-friendly web interface was developed, integrating the best-performing model for real-time risk prediction to facilitate easy usage of the model.

6.2 Recommendations for Future Work

The dataset needs to be expanded to cover a more diverse geographical representation, this will give more insights into how diabetes risk factors vary across different lifestyles, demographics, and healthcare systems. Also, future research needs to be conducted to compare the results from model prediction against the clinical data to ensure the reliability of the model in real-world healthcare scenarios.

References

1. Ahmed N, Ahammed R, Islam M, Uddin MA, Akhter A, Talukder MA, Paul BK (2021) Machine learning based diabetes prediction and development of smart web application. *Int J Cogn Comput Eng* 2:229–241
2. Ahsan MM, Luna SA, Siddique Z (2022) Machine-learning-based disease diagnosis: a comprehensive review. *Healthcare (Basel)* 10(3):541. <https://doi.org/10.3390/healthcare10030541>
3. Ali A, Ramachandra HV, Ramakrishnan N, Murtugudde G (2023) Prognosis of chronic kidney disease using ML optimization techniques. In: International conference on advances in electronics, communication, computing and intelligent information systems (ICAECIS), pp 459–463
4. Ali L, Khan SU, Golilarz NA, Yakubu I, Qasim I, Noor A, Nour R. A feature-driven decision support system for heart failure prediction based on χ^2 statistical model and Gaussian naive Bayes. Hindawi Limited
5. Aman A, Chhillar RS (2023) Optimized stacking ensemble for early-stage diabetes mellitus prediction. *Int J Electr Comput Eng (Malacca, Malacca)* 13(6):7048. <https://doi.org/10.11591/ijece.v13i6.p>
6. Battiti R (1994) Using mutual information for selecting features in supervised neural net learning. *IEEE Trans Neural Netw* 5(4):537–550
7. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. *Comput Electr Eng* 40(1):16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
8. Cheng Z, Zou C, Dong J (2019) Outlier detection using isolation forest and local outlier factor. In: Proceedings of the conference on research in adaptive and convergent systems (RACS '19). Association for Computing Machinery, New York, NY, USA, pp 161–168. <https://doi.org/10.1145/3338840.3355641>
9. Dhiman T, Goyal S, Singh NT, Singhal RK (2023) Machine learning application for diabetes prediction using ensemble classification. In: 2023 5th international conference on inventive research in computing applications (ICIRCA). IEEE, pp 516–519
10. Institute for Health Metrics and Evaluation (IHME) (2023) Global diabetes cases to soar from 529 million to 1.3 billion by 2050. <https://www.healthdata.org/news-events/newsroom/news-releases/global-diabetes-cases-soar-529-million-13-billion-2050>
11. Ismail E, Ramdan R, Mohamed S, Yousri RM, Darweesh S (2023) A comparative study of diabetes classification based on machine learning. In: Intelligent methods, systems, and applications (IMSA). <https://doi.org/10.1109/IMSA58542.2023.10217373>
12. Kumar M, Bajaj K, Sharma B, Narang S (2022) A comparative performance assessment of optimized multilevel ensemble learning model with existing classifier models. *Big Data* 10(5):371–387
13. Lavie CJ, Ozemek C, Carbone S, Katzmarzyk PT, Blair SN (2019) Sedentary behavior, exercise, and cardiovascular health. *Am Heart Assoc.* <https://doi.org/10.1161/CIRCRESAHA.118.312669>

14. Mahesh TR, Kumar D, Kumar VV, Asghar J, Bazezew BM, Natarajan R, Vivek V (2022) Blended ensemble learning prediction model for strengthening diagnosis and treatment of chronic diabetes disease. *Comput Intell Neurosci* 2022:1–9. <https://doi.org/10.1155/2022/4451792>
15. Novakovic J, Veljovic A (2011) C-support vector classification: selection of kernel and parameters in medical diagnosis. In: 2011 IEEE 9th international symposium on intelligent systems and informatics, pp 465–470
16. Olisah CC, Smith L, Smith M (2022) Diabetes mellitus prediction and diagnosis from a data pre-processing and machine learning perspective. *Comput Methods Programs Biomed* 220:106773. <https://doi.org/10.1016/j.cmpb.2022.106773>
17. Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM (2022) A review of feature selection methods for machine learning-based disease risk prediction. *Front Bioinform* 2:927312. <https://doi.org/10.3389/fbinf.2022.927312>
18. Rajendra P, Latifi S (2021) Prediction of diabetes using logistic regression and ensemble techniques. *Comput Methods Programs Biomed Update* 1:100032. <https://doi.org/10.1016/j.cmpbup.2021.100032>
19. Sahu P, Mantri JK (2023) Effectiveness of machine learning and deep learning models for diabetes prediction. *Int J Innov Eng Sci* 8(3):24–34. <https://doi.org/10.46335/IJIES.2023.8.3.5>
20. Singh D, Singh B (2020) Investigating the impact of data normalization on classification performance. *Appl Soft Comput* 97:105524. <https://doi.org/10.1016/j.asoc.2019.105524>
21. Suthaharan S (2016) Decision tree learning. In: *Machine learning models and algorithms for big data classification*. Integrated series in information systems, vol 36. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7641-3_10
22. Teboul A (2021) Diabetes health indicators dataset. <https://www.kaggle.com/alexteboul/diabetes-health-indicators-dataset>

A Secure Data Dissemination in VANET Using Trust and Random Forest Algorithm



C. Balakumar and B. Akashpriyan

Abstract Transportation systems that are intelligent are frequently referred to as vehicular ad hoc networks (VANET). Road safety and traffic flow efficiency are increased by VANET, which guarantees prompt and accurate communications between vehicles and infrastructure (V2V and V2I). Given its strong mobility and open wireless boundaries, the most difficult issues in VANET are thought to be security and data transfer dependability. In order to improve data transmission security and reliability in automotive networks, this study proposes the creation of an architecture that incorporates random forest functions and exercises the trust algorithm notion. We create a simulation model based on real-world communication and traffic patterns. According to the simulation results, the suggested protocol used by the random forest framework with a trust foundation operates very well in terms of safe and secure and also dependability than the traditional methods of CDP and HEPPA. In addition, the simulation model is shown along with a point-by-point execution evaluation to show the feasibility and scalability of the suggested approach. The suggested TRFA scheme's performance is assessed in relation to various evolution metrics like packet delivery ratio, packet drop ratio, end-to-end time delay, and throughput performances across a range of network scenarios. The results show that the suggested trust-based random forest algorithm (TRFA) for secure data dissemination in VANET approach performs satisfactorily comparing than the existing systems of Cisco Delivery Protocol (CDP), Highly Efficient Privacy Preserving Authority (HEPPA) in the VANET environment.

Keywords Data security · Data dissemination · Vehicular ad hoc network (VANET) · Random forest · Trust model

C. Balakumar (✉) · B. Akashpriyan

Department of Computer Applications, Faculty of Arts, Science, Commerce and Management, Karpagam Academy of Higher Education, Coimbatore, India
e-mail: balakumar.cbk@gmail.com

1 Introduction

The speed at which autonomous vehicle (AV) technology is developing has sparked intense curiosity in the ways that human–vehicle interaction (HVI) will occur within the car. As a result, efforts are being made to investigate intelligent interiors, which aim to enhance consumer acceptance and confidence. US experts predict that by 2040, there will be more than 2 billion motor vehicles on the road worldwide, with 50% of those being cars. With such a large network of roads, novice drivers might feel at ease, and it is analysed by Muratori et al. [1]. The academic and automotive communities have just recently started research on efficient and secure inter-vehicular communication due to the spike in interest in vehicular ad hoc networks (VANET), Hussein et al. [2].

Scholars endeavour to develop an intelligent transportation system (ITS) that prioritizes passengers and incorporates both internal and external amenities. Every car has a number of incredibly precise sensors installed to help it understand its environment. In order to provide intelligent, safe navigation for users' convenience, it facilitates user cooperation and communication [3].

1.1 VANET Architecture

Base stations (BS) or cars are the nodes in vehicular ad hoc networks (VANETs), which connect via the pre-existing infrastructure. The VANET architecture is depicted in Fig. 1. Cars can communicate with each other because of RSUs positioned along the roads. Passing vehicles can get connectivity support from RSU. The user can browse the Internet, find parking spaces and free slots, find the closest shopping centre, and engage in mobile commerce with VANET. There are three architectures that VANET relies on for data dissemination. One option is that the vehicles use RSU to communicate with one another. It is similar to wireless local area networks (WLAN). The second method involves direct communication between the cars without the use of RSU. It falls under the umbrella of ad hoc architecture. In the third situation, as described by Tiennoy and Saivichit [4], certain cars are able to communicate directly while others need an RSU. It is referred to as hybrid architecture.

Three distinct communication techniques are employed in VANET. They are

1. Intercommunication between vehicles
2. V to I communication
3. Communication between road side interactions.

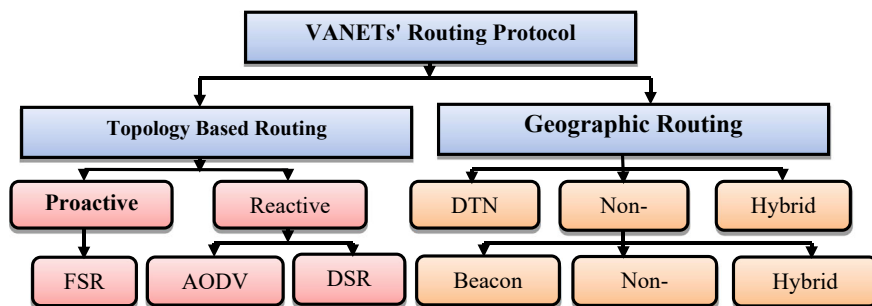


Fig. 1 Routing protocols in VANET

1.2 Vehicular Ad Hoc Routing Protocol

Mobile devices require dynamic switching and self-organization. Routing protocols regulate the distance a data packet travels between a source vehicle and a destination vehicle. The network can anticipate an alternative path to take with the aid of routing. Kirubasri et al. [5]. The vehicle ad hoc network uses two types of routing: geographic-based networks and topology-based networks. The various kinds of routing protocols are described in Fig. 1.

Routing protocols designed for the VANET environment should satisfy the requirements of scalability, efficiency, and comparability. Several routing protocols for VANETs are covered in the next subsection by da Costa et al. [6].

VANET routing is made up of five categories, including.

1. Topology-based routing.
2. Position-based routing.
3. Routing-based on broadcast.
4. Routing-based on geocast.
5. Routing-based on cluster.

1.3 Secure Data Dissemination in VANET

VANETs, or vehicular ad hoc networks, are a young technology with a lot of future potential. In order to provide services that are especially pertinent to a vehicle context, VANETs aim to connect equipment found within automobiles together. Vehicle and infrastructure component communication can be dependable and trustworthy when secure data distribution is accomplished through VANETs, Gupta et al. [7]. To provide the best defence against new threats and vulnerabilities, security solutions must be tailored to the unique needs and features of VANET settings.

1.4 Random Forest Algorithm

The random forest algorithm is a potent ML technique for applications involving both regression and classification. It is a part of the ensemble learning techniques, which combine several different models to get predictions that are more reliable and accurate. Because of its adaptability, efficiency, and simplicity in a variety of contexts, random forests are frequently utilized. Punj and Kumar [8] discussed, because of its flexibility, scalability, and capacity for handling complex decision-making tasks.

2 Literature Review

The limitations and shortcomings of the present methods utilized in data dissemination in VANET employing various approaches are discussed in this section on literature reviews. Significant contributions on this subject from a broad spectrum of experts are covered in the literature survey. A Geological Secure Way Steering convention (GSPR) was offered by Shajin and Rajesh [9] in order to obtain area-aware VANET administrations. Their framework-less geographic steering convention, or GSPR, is adaptable to disruptions caused by malignant or malfunctioning hubs.

The VANET's area-aware foundation of this strategy by Ali et al. [10] is thought to have improved the decentralization and adaptability of the VANET framework. An infrared technique was used to transfer data from the source to the collector. The distance between automobiles is decreased by combining intelligent services. The car's channel performance is enhanced as a result. Husnain [11] when organizing the conference, the organization and BS consider the sensor hubs' portability. There are various methods for moving each hub. The act of grouping hubs according to their collective energy is known as clustering. There is one bunch head hub per group, and two agent bunch head hubs support the bunch head hub.

2.1 Data Dissemination Techniques in VANET

Arif et al. [12] examined the verification of security messages and its test in a high-thickness traffic situation. It likewise proposed an answer for vehicular message confirmation in thick rush hour gridlock conditions by a focused on check technique. Xu et al. [13] examined fundamental security risks in VANET, particularly when an attacker disseminates false information to influence the behaviour of several drivers. Furthermore, the idea of importance verification was presented, which comprises demonstrating the event report's genuine relevance to the incident it documents.

Chowdhury et al. [14] describe the security needs to safeguard the information and the safety challenges that VANET faces. This study by Rashid et al. [15] looked into a number of approaches and reviewed relevant literature in order to create a model

for effective and dependable message distribution routing strategies in VANETs. The paper wrote by Sospeter et al. [16] explained that the adaptive probability data dissemination protocol (EEAPD) is suggested as a practical and efficient method. A probabilistic technique and a delay mechanism are part of EEAPD. In rebroadcast probability judgments, the relationship between road segments and vehicle density is explained by the redundancy ratio (r) measure. The EEAPD protocol is unusual in that it determines which nodes are appropriate for rebroadcasting the emergency message based on the number of road segments. This protocol takes into account a beaconless approach and is exclusively intended for V2V communication. The traffic simulator known as “SUMO” and Ns-3.26 were used to do the simulations for this investigation.

3 System Design

Various strategies are available to use the benefits of random forests in classification and decision-making tasks for safe data transfer in vehicular ad hoc networks (VANETs). The following are some ways that the trust algorithm combined with the random forest approach might assist guarantee safe data delivery in VANETs. And also the existing models of CDP and HEPPA also included in the system design for the comparison.

3.1 *Data Fusion and Trust Establishment*

Vehicles may collect and share data from a variety of sources in VANETs, including as other vehicles, roadside infrastructure, and cloud services. To determine the degree of trustworthiness for specific vehicles or data sources, random forests can be utilized for the integration and analysis of data from several sources, Lin et al. [17]. To enable safe data distribution and decision-making procedures, access control mechanisms are essential.

3.2 *Secure Routing and Traffic Management*

Sheikh and Liang [18] discussed about the algorithm that it can make decisions to optimize traffic flow, minimize latency, and improve overall network performance while guaranteeing the security and integrity of data dissemination within the VANET environment by examining historical traffic data, real-time sensor inputs, and environmental conditions (Fig. 2).

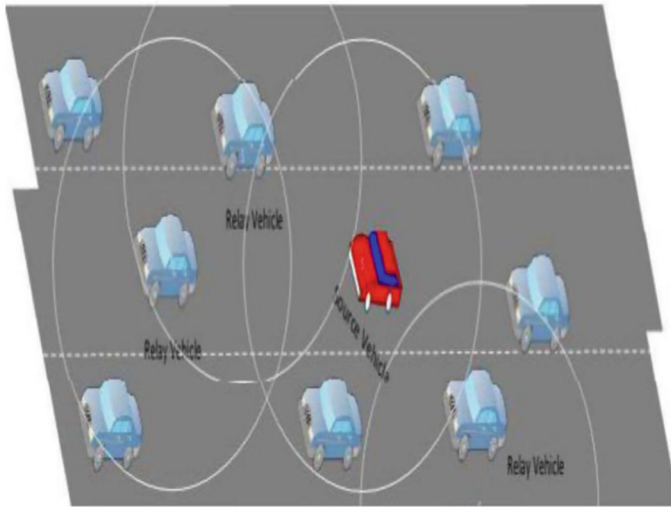


Fig. 2 Data dissemination in VANET

3.3 Trust Algorithm

According to this protocol, the trusted path is chosen by a source node, or cluster head, in accordance with the path trust value. The road with the highest energy and the fewest hops to the base station will be chosen as the final route if more than one path has a path trust value larger than the threshold value, given by Kalidoss et al. [19].

3.4 Random Forest Algorithm

An ensemble classification technique, the random forest approach is based on the decision tree concept. It takes an original dataset and utilizes a bootstrap sampling technique to generate k different training data subsets. After that, k decision trees are created by training these subsets. A random forest is eventually produced by combining these decision trees taken from Chen et al. [20].

Our newly constructed proposed system architecture design is explained in the Fig. 3. The random forest method is constructed in the following phases.

Step 1. k training subsets being sampled: Using bootstrap sampling, k training subsets are selected in this stage from the initial training dataset S . Specifically, a random sampling and replacement mechanism is used to select N records from S at each sample period. Following the present stage, a collection of training subsets ST rain is used to create k training subsets:

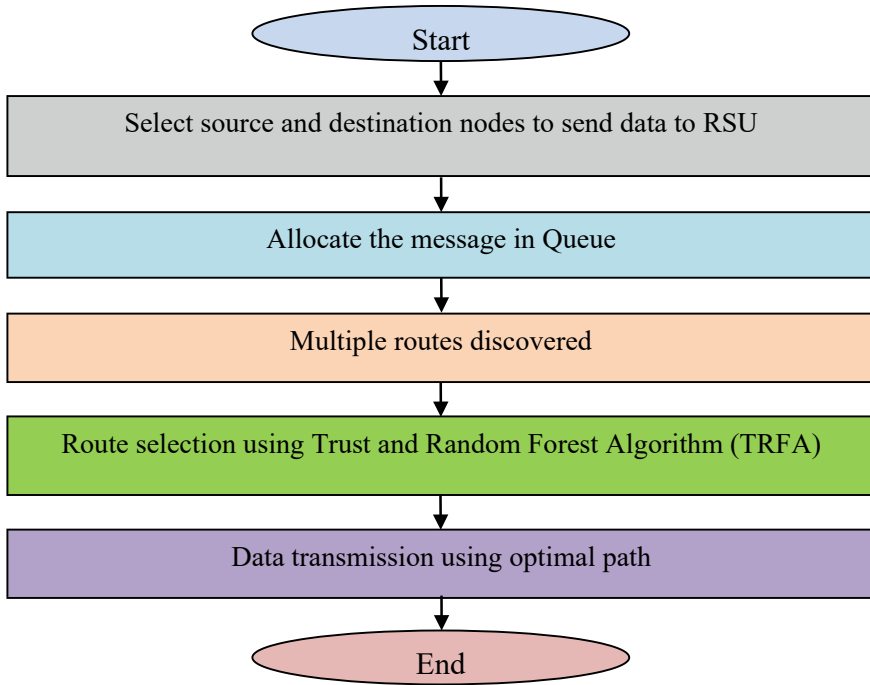


Fig. 3 Proposed work design of TRFA

$$ST\ rain = \{S_1, S_2, \dots, S_k\}.$$

Step 2. Build every decision tree model: A CART algorithm creates each meta decision tree in an RF model from each training subset S_i . During the tree-growth process, m feature variables from dataset S_i are chosen at random from a list of M variables. The best feature variable is selected as the splitting node in each tree node splitting process, which involves calculating the gain ratio of each feature variable. Until a leaf node is produced, this splitting process is repeated. Lastly, the same training procedure is used to train k decision trees from k training subsets.

4 Result and Discussion

We compare the newly developed TRFA algorithm to the current Cisco Delivery Protocol (CDP) and Highly Efficient Privacy Preserving Authority (HEPPA) models.

4.1 Packet Delivery Ratio

Any routing strategy in a network can be evaluated for effectiveness using the packet delivery ratio statistic. The ratio is calculated by adding the volume of data packets delivered by the source and the amount received by the destinations. By multiplying the total number of packets delivered throughout the simulation, we can determine the mean delivery ratio by the number of target nodes receiving the safety message inside the target region. The message delivery rate must be close to 100% in order for a dissemination approach to be considered dependable.

Packet Delivery Ratio = $\frac{\text{All destination total packets received}}{\text{Total no. of packets sent by all source nodes}}$

Figure 4 compares packet delivery ratio findings (in percentages) from the Table 1; CDP and HEPPA, the two currently used techniques, produce values of 62% and 73%, respectively. The proposed TRFA Protocol’s destinations received packets from 88% of the destinations.

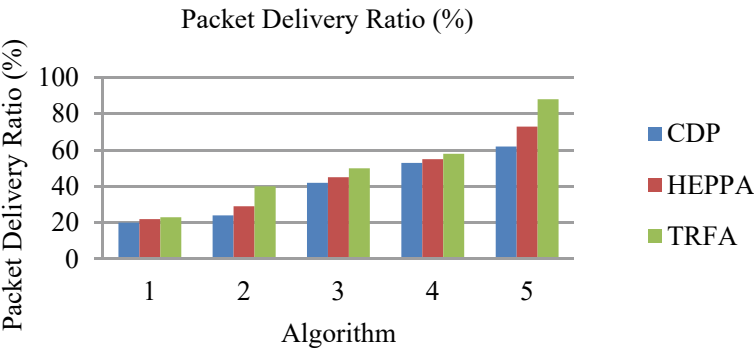


Fig. 4 Packet delivery ratio result graph

Table 1 Packet delivery ratio result

Packet delivery ratio (%)		
Existing protocols		Proposed protocol
CDP	HEPPA	TRFA
20	22	23
24	29	40
42	45	50
53	55	58
62	73	88

4.2 Packet Drop Ratio

The packet drop ratio is the percentage of packets that were transmitted from a source but never arrived at their destination.

$$\text{Packet Drop Ratio} = \frac{n \text{ send packets} - n \text{ recieved packets}}{n \text{ send packets}} * 100$$

The Table 2 compares the performance of the existing protocols, such as CDP and HEPPA with proposed TRFA.

Figure 5 compares the results for packet drop ratio (in percentage) and shows that the proposed solution has a lower PDR or loss of pocket than the existing approach. In the proposed TRFA protocol, 8% of packets never reach their intended destination. The existing approaches CDP and HEPPA give 18% and 12% packet drop ratio.

Table 2 Packet drop ratio result

Packet drop ratio (%)		
Existing protocols		Proposed protocol
CDP	HEPPA	TRFA
3	2	1
8	4	1
10	5	3
15	10	5
18	12	8

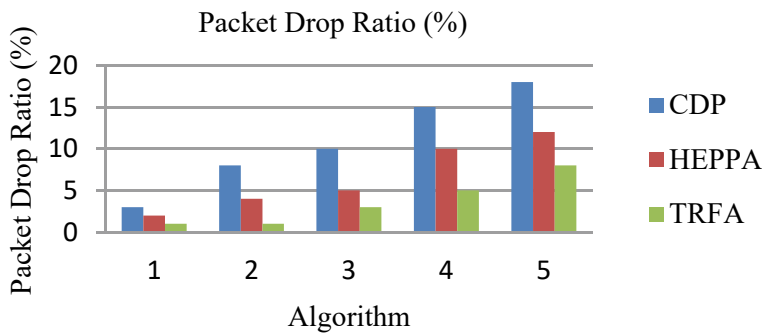


Fig. 5 Packet drop ratio result graph

Table 3 End-to-end time delay result

End-to-end delay (sec)		
Existing protocols		Proposed protocol
CDP	HEPPA	TRFA
0.06	0.04	0.01
0.10	0.07	0.03
0.20	0.13	0.10
0.32	0.24	0.18
0.56	0.47	0.32

4.3 End-To-End Time

The time between calling for a safety message and actually sending one, as well as the last node in the target region to receive it, are important indicators for assessing a safety message dissemination protocol's effectiveness. This is due to the need for prompt distribution of such messages. The protocol is better suited for disseminating safety alerts as latency decreases. The typical amount of time needed for a data packet to reach its destination. Milliseconds are used to measure it. In terms of end-to-end delay, Table 3 contrasts the suggested TRFA with the current protocols CDP and HEPPA.

$$EED = \sum T_A - T_S$$

The preceding formula, where EED stands for end-to-end delay, T_A for packet arrival time, and T_S for packet transmitted time, indicates the formula for calculating end-to-end delays.

Figure 6 displays comparison results for end-to-end delay (in sec); from the source to the target point, messages delivered utilizing the present techniques have an end-to-end latency of 0.56 sec for CDP and 0.47 sec for HEPPA. With the proposed TRFA, there is only a 0.32 sec delay between sending a safety message and the last node receiving the dissemination in the target region.

4.4 Throughput

It measures the average throughput across the entire system and is calculated as message transfer in packets per unit of time. It calculates the length of a period.

$$\text{Throughput} = \frac{\text{Received Size}}{(\text{Stop Time} - \text{Start Time})} * \frac{8}{1000}$$

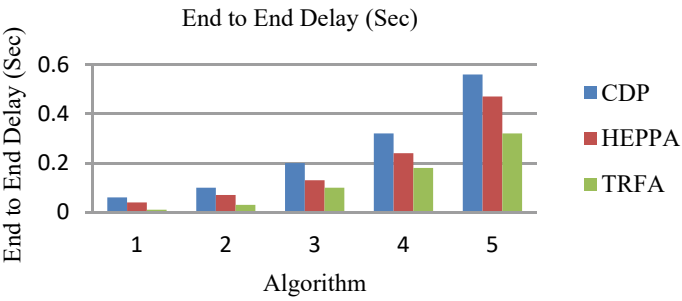


Fig. 6 End-to-end delay result graph

Table 4 Throughput performance result

Throughput (%)		
Existing protocols		Proposed protocol
CDP	HEPPA	TRFA
10	14	20
18	25	32
30	42	55
45	60	75
58	69	89

Table 4 provides a comparison of throughput performance in terms of percentage between the existing protocols like CDP and HEPPA and the proposed TRFA protocol.

Figure 7 displays the throughput comparison results. CDP and HEPPA existing algorithms achieves 58% and 69%, respectively, per unit of time in 500 sec. The proposed TRFA sent more packets per unit time interval, as shown by this clearly. The suggested TRFA achieves 89% throughput performance which is higher than the present CDP and HEPPA values.

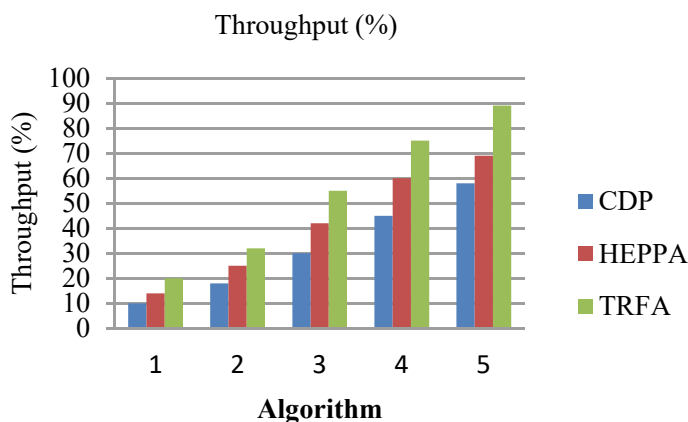


Fig. 7 Throughput result graph

5 Conclusion

The field of vehicle ad hoc networks has been expanding and developing in the last several years. To broadcast data on traffic conditions and other important information to neighbouring vehicles, vehicles use a new network called the VANET. The security of data dissemination issues for vehicle ad hoc networks is the main emphasis of the study that has been proposed. This suggested system for the secure and reliable data dissemination in VANET by using trust-based random forest algorithm (TRFA) overall outperforms well and attains the packet delivery ratio, packet drop ratio, end-to-end time delay, and throughput performance better comparing with the current models like Cisco Delivery Protocol (CDP) and Highly Efficient Privacy Preserving Authority (HEPPA). The suggested data transfer protocol may in the future be expanded to include extremely high security at a low cost, and the Internet of vehicles (IoV) may use it.

References

1. Muratori M, Alexander M, Arent D, Bazilian M, Cazzola P, Dede EM, Farrell J, Gearhart C, Greene D, Jenn A, Keyser M, Lipman T, Narumanchi S, Pesaran A, Sioshansi R, Suomalainen E, Tal G, Walkowicz K, Ward J (2021) The rise of electric vehicles—2020 status and future expectations. *Prog Energy* 3(2):022002
2. Hussein NH, Yaw CT, Koh SP, Tiong SK, Chong KH (2022) A comprehensive survey on vehicular networking: communications, applications, challenges, and upcoming research directions. *IEEE Access* 10:86127–86180
3. Murali PK, Kaboli M, Dahiya R (2022) Intelligent in-vehicle interaction technologies. *Adv Intell Syst* 4(2):2100122
4. Tiennoy S, Saivichit C (2018) Using a distributed roadside unit for the data dissemination protocol in VANET with the named data architecture. *IEEE Access* 6:32612–32623

5. Kirubasri G, Sankar S, Pandey D, Pandey BK, Nassa VK, Dadheech P (2022) Software-defined networking-based ad hoc networks routing protocols. *Software defined networking for ad hoc networks*. Springer International Publishing, Cham, pp 95–123
6. da Costa JBD, de Souza AM, Rosário D, Cerqueira E, Villas LA (2019) Efficient data dissemination protocol based on complex networks' metrics for urban vehicular networks. *J Internet Serv Appl* 10:1–13
7. Gupta N, Manaswini R, Saikrishna B, Silva F, Teles A (2020) Authentication-based secure data dissemination protocol and framework for 5G-enabled VANET. *Future Internet* 12(4):63
8. Punj R, Kumar R (2022) Data dissemination approach using machine learning techniques for WBANs. *Concurrency Comput: Pract Experience* 34(5):e6688
9. Shajin FH, Rajesh P (2022) Trusted secure geographic routing protocol: outsider attack detection in mobile ad hoc networks by adopting trusted secure geographic routing protocol. *Int J Pervasive Comput Commun* 18(5):603–621
10. Ali ZH, Badawy MM, Ali HA (2020) A novel geographically distributed architecture based on fog technology for improving Vehicular ad hoc network (VANET) performance. *Peer-to-Peer Netw Appl* 13(5):1539–1566
11. Husnain G (2021) Bio inspired intelligent routing scheme for clustering in vehicular ad hoc networking (VANETs). Doctoral dissertation, University of Engineering & Technology Peshawar (Pakistan)
12. Arif M, Wang G, Balas VE, Geman O, Castiglione A, Chen J (2020) SDN based communications privacy-preserving architecture for VANETs using fog computing. *Vehicular Communications* 26:100265
13. Xu X, Wang Y, Wang P (2022) Comprehensive review on misbehavior detection for vehicular ad hoc networks. *J Adv Transp*
14. Chowdhury A, Karmakar G, Kamruzzaman J, Jolfaei A, Das R (2020) Attacks on self-driving cars and their countermeasures: a survey. *IEEE Access* 8:207308–207342
15. Rashid SA, Audah L, Hamdi MM, Abood MS, Alani S (2020) Reliable and efficient data dissemination scheme in VANET: a review. *Int J Electr Comput Eng (IJECE)* 10(6):6423–6434
16. Sospeter J, Wu D, Hussain S, Tesfa T (2018) An effective and efficient adaptive probability data dissemination protocol in VANET. *Data* 4(1):1
17. Lin H, Garg S, Hu J, Wang X, Piran MJ, Hossain MS (2022) Data fusion and transfer learning empowered granular trust evaluation for internet of things. *Inf Fusion* 78:149–157
18. Sheikh MS, Liang J (2019) A comprehensive survey on VANET security services in traffic management system. *Wirel Commun Mob Comput* 2019:1–23
19. Kalidoss T, Rajasekaran L, Kanagasabai K, Sannasi G, Kannan A (2020) QoS aware trust based routing algorithm for wireless sensor networks. *Wireless Pers Commun* 110:1637–1658
20. Chen Y, Zheng W, Li W, Huang Y (2021) Large group activity security risk assessment and risk early warning based on random forest algorithm. *Pattern Recogn Lett* 144:1–5

Cognitive Disorder Detection Used for Early Autism Screening Using Deep Learning



R. Gunasundari and G. Jeyabharathi

Abstract The proposed Monotonous Trans cranial Magnetic Stimulation device has been used as an invasive therapy for a variety of neurological and psychiatric conditions, including schizophrenia, Parkinson's disease, and depression, throughout the last ten years. Early detection of cognitive disorders such as dementias and Alzheimer's allows patients to get higher-quality therapy that helps them retrieve the information they need and preserve their memories. The rTMS device modifies the firing pulse of neurons to initiate brain communication by providing a varying magnetic field. This problem arises because Alzheimer's disease tangles and the resulting neuronal death often begin in memory-related areas like the progress to. In clinical and research contexts, cognitive disorders are associated with difficulties; moderate cognitive disorder (MCD) in particular need a high level of diagnostic suspicion in addition to equipment that are sensitive enough to detect it. This study was out to validate the short cognitive test (CATest) for MiCD diagnosis.

Keywords Disorder detection · Early autism · Screening · Deep learning · Back Propagation Neural Network · Optimization

1 Introduction

The ultimate objective customize meet requirements people, enhance enable fully participate in society. Therefore, the application severity duration severe illness has been examined in a number of research field of health. This discovery is important since the major depressive disorder (MDD) disease course is heterogeneous,

R. Gunasundari (✉)

Department of Computer Applications, Faculty of Arts, Science, Commerce and Management,
Karpagam Academy of Higher Education, Coimbatore, India
e-mail: gunasoundar04@gmail.com

G. Jeyabharathi

Department of Computer Applications, Karpagam Academy of Higher Education, Coimbatore,
India

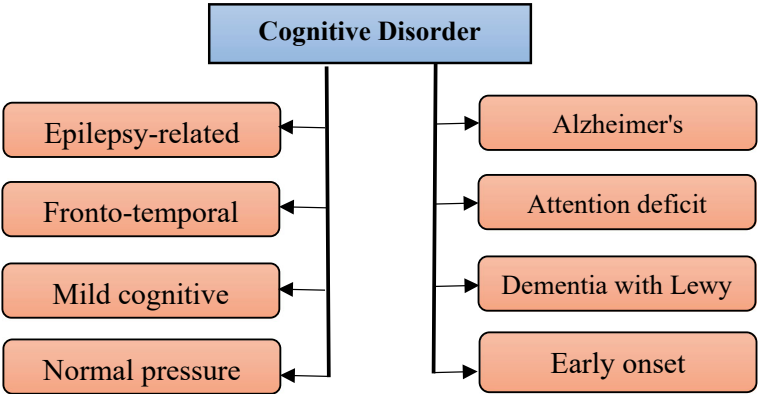


Fig. 1 Basic types of cognitive disorder

making treatment selections more difficult. A machine learning method to forecast how depression therapy will turn out. The researchers identified 25 of the 164 patient-reportable characteristics and they used these variables to train their model. In addition to offering patients and providers cutting-edge clinical tools regarding as well as numerous analytical approaches and strategies from the field of computational sciences have been investigated and used using data related to mental health. Urgent are required processing, presentation massive in light of the primary types of cognitive illness shown in Fig. 1.

1.1 Classifies Cognitive Disorders Into Two Categories

MCD, which in versions, and mild neurocognitive disorder (MiCD), which is used when a change in higher brain function significantly impairs an individual’s ability to function at work, in social situations, or in the family. On the other hand, 3% to 20% of those over 65 have MiCDvaries; around with develop the condition, although over other studies have shown a prevalence of over 22% obtained by Holland et al. [1].

1.2 Cognitive Disorder (CD)

Cognitive disorder is a common diagnosis made for hospital patients, characterized by varying degrees of cognitive impairment. Cognitive disorder is a phrase used to describe a type of short-term cognitive impairment. Caruso et al. [2] characterized by symptoms such as altered awareness, poor memory and attention, poor orientation, small duration, fast onset lasting several hours to several days, and notable changes,

most of which happen at night. Impairment had an annual death rate of 50%, or around 23–33% of patients. 20–75% of hospitalized patients with cognitive impairments die during their stay, 15% and 25% of these patients pass away in the first and sixth months, respectively. There is disagreement on the causes of a wide range of cognitive impairments, including as, metabolic encephalopathy, auditory bleeding, electrolyte imbalances, post-operative pain, sleeplessness, surgical stress, and medicines are the primary causes of this and Fig. 2 explain the flow of system.

Preprocessing: During the preprocessing stage, filtering crucial image frames is required to remove unwanted items, distracting noise, and a static backdrop in addition to capturing crucial information. To quickly extract characteristics from photographs, the proposed method preprocesses the images by reducing them, converting them to grayscale, and identifying their edges.

Segmentation: Computer vision and object recognition are based on image segmentation. In general, during image processing, picture noise should be removed. Furthermore, in order to enhance the visual effect, certain activities (such region extraction and picture marking) must be completed after the primary image segmentation process.

Feature extraction: Algorithms are utilized to recognize faces, fingers, palms, and other aspects like orientation and arrangement after the images have been processed. The procedure turns the picture into a data array with the input sign represented in it.

Classification: Classifying pictures most popular tasks categories based on the characteristics of the supplied picture. SVM, ANN, k nearest neighbor, adaptive boost, and SVM are some techniques that might be applied to this problem. Science also recognizes the advantages of early and precise screening.

Support Vector Machine: Using hyperplanes to create decision boundaries between data points of various classes is a crucial part of support vector machines (SVMs).

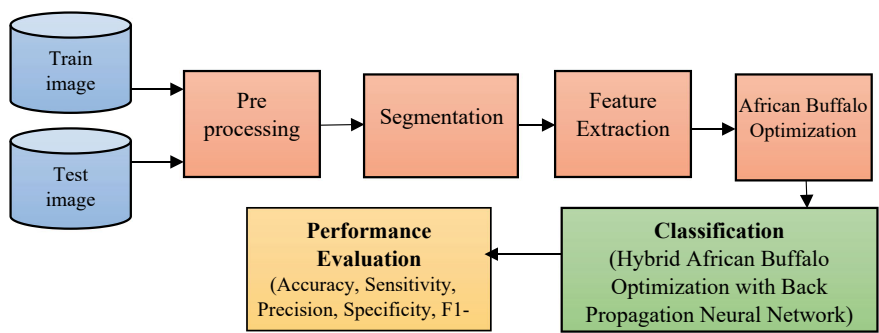


Fig. 2 Proposed system flow of process

SVMs may be used for both straightforward linear classification jobs and more intricate nonlinear classification challenges.

2 Literature Review

Ball et al. [3] made recommendations for how these algorithms should be applied applications in the healthcare industry were also offered by the writers. Mirzaei and Adeli [4] goal was to find out if Twitter might problems with increasing certain ailment those who are occupations. Liu et al. [5] demonstrate many algorithms to determine the most significant genes and gene sequences within a set of gene expression microarrays. Begemann et al. [6] used more, researchers examined every mark on assessing behavior linked to using a range of machine learning methods instrument utilized their research, identifying behaviors linked to autism via.

Pieruccini-Faria et al. [7] investigated the causal relationships between one brain area and another effective connectivity in their fMRI investigation adult's adolescent's individuals. The study of Chen et al. [8] difficulties in applying ML situation, it faulted previous studies for being unduly subjective. They stressed ML enormous improve behavioral sciences research on diagnosis and treatment, especially for common diverse were provided by the authors. Zwaigenbaum et al. [9] used a two-tiered screening approach that improved coordination, interagency policy participation, and quality evaluation.

Finally, Park et al. [10] examined potential delays in the identification of ASD. They evaluated the distinctions between children diagnosed with intellectual disability/developmental delay (ID/DD) and those diagnosed with autism spectrum disorder (ASD) based on the age at which the parents first discussed their concerns with a healthcare professional and the age at which they first voiced fears to the professional. Baschi et al. [11] established as tool for assessing autism. Opinions state that because the child autism rating scale, developed by Mahaman et al. [12] physician's assessment the kid's behavior using two scales a social communication questionnaire and a social responsiveness scale it is a more accurate way to diagnose autism. Silva et al. [13] described a procedure for gathering features, logging eye data, and training machine learning classifiers.

Tan et al. [14] have demonstrated the ability of biomarkers, such as genetics, physiological behavior, neurodevelopment, and eye tracking, to assist in the diagnosis of autism. Social stimuli were assessed using eye tracking in several populations, including children with Down syndrome, ASD, Rett syndrome, and Williams's syndrome. An author Ferretti et al. [15] condition that is increasingly well-known is functional cognitive disorder (FCD), which affects up to 25% of people who get cognitive evaluation services. It potentially overlaps with concepts conversely, judgments based explained by Zlokovic et al. [16]. Secondly, the main FODMIPS, which include lower quality but more frequent cognitive lapse detection by Kumar et al. [17]. Around age three, children start to ponder the possibility that mental pictures are not real, tangible objects. Furthermore, Perdoncin et al. [18] tells about the notion

that a person's 120 thoughts paint a picture of their current condition gets increasingly sophisticated beyond that age, suggesting that mental representations and apparent situations of objects have causal ties.

3 System Design

The recommended method for sign language prediction for deaf people who can hear voices is covered in this session. As a result, African Buffalo Optimization with Back propagation Neural Network (ABOBPNN), the suggested method, performs better than SVM, CNN, and ANN, which are the current systems.

Training method: Using the basic backpropagation method, the weights are modified in the direction of the steepest fall, or the direction in which the performance function is falling the fastest. Even when quickly method mathematicians have previously demonstrated that conjugate gradient (CG) algorithms, which search along conjugate gradient paths, provide faster convergence than steepest descent techniques.

Preprocessing: The preprocessing phase ensures input homogeneity and normalizes the frames by taking into account the different frame sizes found in movies. Input signal homogeneity is necessary for the model to handle inter-database disparities brought on by the use of different and digitalization techniques. More specifically, input can be fed into the model under a variety of circumstances that strike a balance between decreasing input and maintaining the signal qualities required to complete the sign score assignment.

Segmentation: Conversely, edge-based segmentation divides an image based on sharp changes in brightness around the boundaries. In this field, region growth, region splitting, region merging, and thresholding are the four primary strategies. This research merits wide attention since it might serve as a crucial source of reference for such segmentation techniques. The assessment algorithms' simplicity, dependability, and universal application are all taken into account, as is the necessity of personal involvement or referent pictures. In order to evaluate picture segmentation objectively, two fundamental techniques are often used.

Feature Extraction: Although posture assessment and hand usage are essential components of sign language, DSL's continual movement poses a number of challenges. Identifying the hands and determining their dimensions, form, and movement are among these challenges. Once the postures for each frame have been determined, the key points are obtained for the three X, Y, and Z dimensions of both hands. The hand's location with respect to the torso was predicted and tracked using the posture estimation approach.

Hybrid African Buffalo Optimization with Back propagation Neural Network: A closer look at the ABO algorithm identifies three elements that together sum up

1. Objective function $f(x) = (x_1, x_2, x_3 \dots x_n)T$
2. **Initialization:** Add buffaloes at random to the solution space's nodes;
3. Using Eq.(1) update the buffalo's fitness values.

$$m.k+1 = m.k + lp1(bgmax-w.k) + lp2(bpmax.k-w.k) \quad \text{-----}(1)$$

where bgmax is the herd's best fitness and bpmax.k is the individual buffalo's best, and w. k and m. k represent the kth buffalo's exploration and exploitation moves, respectively ($k=1,2,\dots,N$).
4. Using buffalo k's (bgmax and bpmax.k) position, update (2)

$$W.K + 1 = \frac{(w.k + m.k)}{\pm 0.5} \quad \text{-----} \quad (2)$$
5. bg_{max} is upgrading. Sure, head to 6. No, move to page 2.
6. Return to algorithm step 3 if the stopping criteria are not met; otherwise, proceed to step.
7. Produce the ideal answer.

Fig. 3 ABO algorithm

buffalo behavior. The buffalo's capacity to recall its previous location is symbolized by the initial m.k. The buffalo are born with the capacity to recall their previous places. This is essential to help it avoid locations where its quest for answers had poor outcomes. In each buffalo's memory is position. The buffaloes' degree of intelligence is shown in the last section, $lp2(bpmax.k-w.k)$. The buffalo's cooperative or sympathetic character is covered in the second section, $lp1(bgmax-w.k)$, which also makes reference to the animal's social and information-sharing habits. Consequently, ABO blends the intelligence, memory, and compassion of the buffaloes into the democratic formula successfully simulated as previously discussed. The sound "waaa" ($k = 1, 2, 3 \dots n$) is the representation of the buffalo's "maaa" sound. Equation (1) in Fig. 3 gives a numerical estimate of the democratic movement of the buffaloes.

Back Propagation Neural Network: As a result, it is known as the backward propagation of errors. Numerous data mining applications based on neural networks use it. In Fig. 4, the backpropagation design is displayed. A process known as "back propagation" is used to spread in reverse.

It is estimated that there are 10 billion neurons in the human brain, with each neuron typically being linked to 10,000 other neurons. Synapses allow signals to reach each neuron and regulate how each cell responds to the stimulus. Because gradient approaches perform well for training updating, variations frequently used.

Working of Back Propagation: It compares the produced intended creates prepared modifies in accordance with obtain the intended results. Figure 5 shows the back propagation method.

One way to visualize backpropagation is as a neural network function. Using the chain rule, gradient descent is a successful training method for artificial neural networks. These techniques are collectively known as backpropagation algorithms. Basic neurons, or nodes, in a tiered hierarchical architecture that are highly coupled to one another across layers comprise a back propagation network. Only connections between layers, or synapses, are allowed in these "simple" nonrecursive networks; connections inside layers are prohibited. An illustration of how the tiers of a two-layered feed-forward network are connected overall.

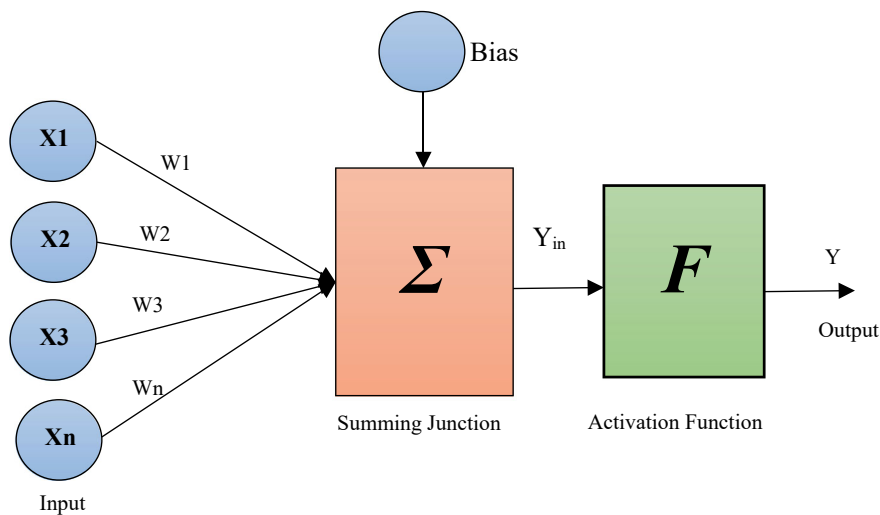


Fig. 4 Back propagation architecture

- Step 1:* Inputs X , arrive through the preconnected path.

Step 2: True weights W are used in the input model. Typically, weights are selected at random.

Step 3: From the input layer to the hidden layer to the output layer, determine each neuron's output independently..

Step 4: Calculate the error in the outputs

Step 5: Adjust the weights in the hidden layer to lower the error after returning from the output layer.

Step 6: Continue doing so until you get the desired result.

Fig. 5 Back propagation algorithm

4 Results and Discussion

4.1 Confusion Matrix

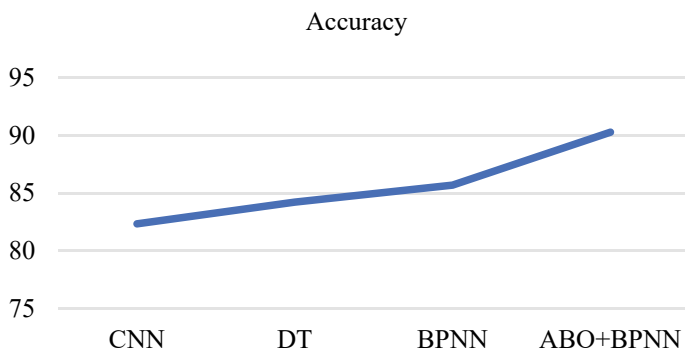
The forecast description is contained in a matrix that is part of the confusion matrix. For every class, the proportion of precise and imprecise forecasts is displayed percentage or cases lie between the actual and projected categories. Through selection, the matrix’s constituent parts input the pertinent occurrences into the output signal. The confusion matrix is shown in Fig. 6.

		Predicted	
Actual		True positives TP	False Negative FN
		False Positive FP	True Negative TN

Fig. 6 Confusion matrix

Table 1 Result values for accuracy with proposed and existing system

Algorithm	Accuracy
CNN	82.33
DT	84.21
BPNN	85.69
ABO+BPNN	90.28

**Fig. 7** Accuracy graph of the proposed and current systems as a result

4.2 Accuracy

The quantity of accurately predicted occurrences serves as a gauge for accuracy. Merely accuracy insufficient since accuracy might be deceiving in an imbalanced dataset, like the cancer dataset, where a very small percentage of patients may actually have the disease. As a result, more metrics are needed to assess a classifier's performance.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}$$

The results of 85.69 demonstrate that the accuracy of the existing system is 4.59 times lower than that of the suggested technique, as assessed by ABO+BPNN at 90.28. Table 1 in Fig. 7 illustrates this.

4.3 Sensitivity

Divide the total number of positives by the number of correct positive forecasts determine sensitivity (SN).

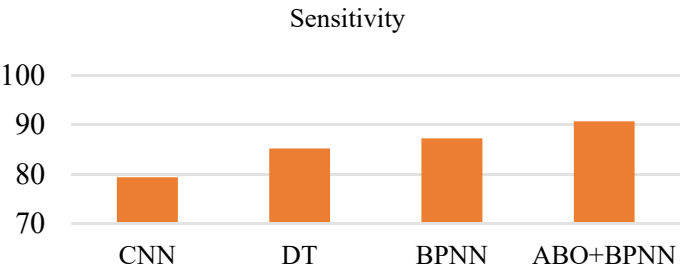


Fig. 8 Sensitivity graph of the proposed and current systems as a result

Table 2 Result values for sensitivity with proposed and existing system

Algorithm	Sensitivity
CNN	79.36
DT	85.19
BPNN	87.22
ABO+BPNN	90.69

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

The proposed system’s ABO+BPNN estimated sensitivity is 90.69, 3.47 times greater than the 87.22 sensitivity of the existing system, as can be seen in Fig. 8 and Table 2’s data.

4.4 Precision

The ratio of all cases the classifier recognized as being in a given class to the total number of examples in that class that were correctly classified was used to determine a class’s accuracy. For internet search results and spam filtering, precision is an important metric of pertinence or correctness. Reduce the amount of valid emails that are marked as spam by using the spam filtering algorithm.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

The results of Table 3’s ABO+BPNN analysis indicate that the suggested method has an accuracy rating of 90.58. This is 3.62 times more accurate than the current approach, according to Fig. 9’s 86.96 rating.

Table 3 Result values for precision with proposed and existing system

Algorithm	Precision
CNN	83.32
DT	84.10
BPNN	86.96
ABO+BPNN	90.58

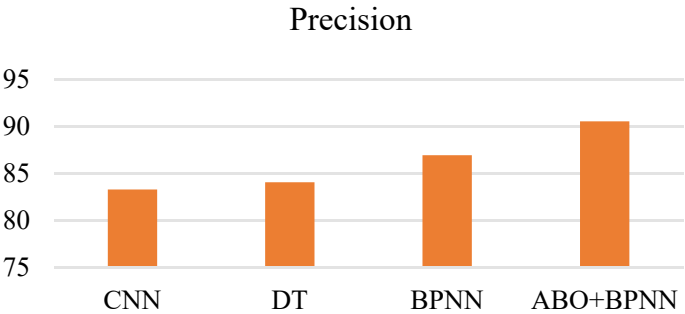


Fig. 9 Precision graph of the proposed and current systems as a result

4.5 F1-Measure

In cases when accuracy and recall are equally significant, the F1-measure is employed. The F1-measure is derived from the accuracy and recall harmonic means. Low numbers of false positives and false negatives in your forecasts are indicated by a high F1-measure.

$$F1\text{-score} = 2 * \frac{\text{percision} * \text{recall}}{\text{percision} + \text{recall}}$$

Using the data from Table 4, the F1 score of the proposed system, as determined by ABO+BPNN, is 97.86 in Fig. 10, 12.41 times greater than the F1 score of the present system, which is determined by 85.45.

Table 4 Result values for F1-score with proposed and existing system

Algorithm	F1-score
CNN	81.78
DT	82.39
BPNN	85.45
ABO+BPNN	97.86

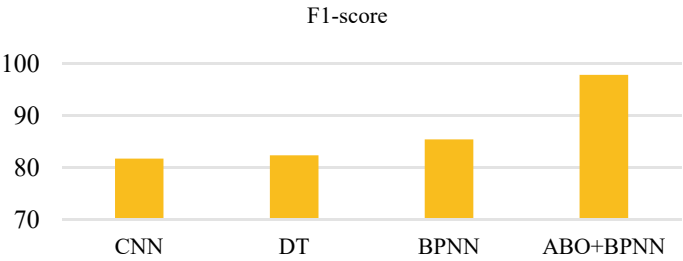


Fig. 10 F1-score graph of the proposed and current systems as a result

4.6 Specificity

The percentage of genuine numbers that can be relied upon to precisely match each and every actual number is known as the specificity indicator. TNR is an additional term that might be utilized.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

As seen in Fig. 11, Table 5’s results indicate that the specificity of the suggested system, as determined by ABO+BPNN, is 92.21, 5.91 times more than that of the existing system, as determined by 86.30.

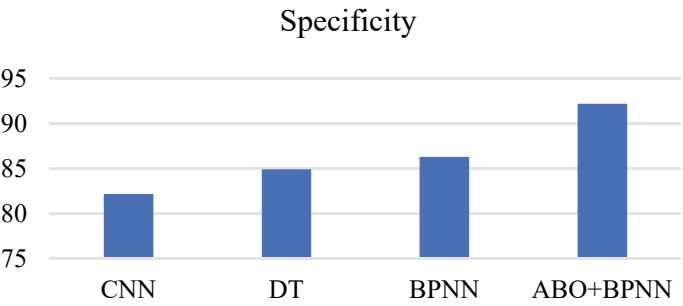


Fig. 11 Specificity graph of the proposed and current systems as a result

Table 5 Result values for specificity with proposed and existing system

Algorithm	Specificity
CNN	82.16
DT	84.91
BPNN	86.30
ABO+BPNN	92.21

5 Conclusion

Our findings suggest that patients with FCD may experience a decoupling of the metacognitive processes, with diminished global metacognition and intact local metacognition. New behavioral tests and analytical techniques that evaluate the influence of local metacognition on global metacognition should be incorporated into future study. One primary objective of a mechanistically possible treatment might be to reestablish the linkages between local and global metacognition. According to current study, using certain AEDs while pregnant may cause cognitive deficits in the fetus. As a side consequence of surgery for refractory epilepsy, cognitive problems might arise. A few have been found high insufficient hippocampal region. Extra care should be taken to search for AED-related cognitive side effects, which seem to be more common in patients receiving polytherapy and with classic AEDs than with new ones. Neuropsychological assessments need to be carried out on a regular basis in order to predict any cognitive deficits that may occur after epilepsy surgery.

References

1. Holland N, Robbins TW, Rowe JB (2021) The role of noradrenaline in cognition and cognitive disorders. *Brain* 144(8):2243–2256
2. Caruso G, Godos J, Privitera A, Lanza G, Castellano S, Chillemi A, Bruni O, Ferri R, Caraci F, Grosso G, Grosso G (2022) Phenolic acids and prevention of cognitive decline: polyphenols with a neuroprotective role in cognitive disorders and alzheimer's disease. *Nutrients* 14(4):819
3. Ball HA, McWhirter L, Ballard C, Bhome R, Blackburn DJ, Edwards MJ, Fleming SM, Fox NC, Howard R, Huntley J, Isaacs JD, Lerner AJ, Nicholson TR, Pennington CM, Poole N, Price G, Price JP, Reuber M, Ritchie C, Rossor MN, Schott JM, Teodoro T, Venneri A, Stone J, Carson AJ, Carson AJ (2020) Functional cognitive disorder: dementia's blind spot. *Brain* 143(10):2895–2903
4. Mirzaei G, Adeli H (2022) Machine learning techniques for diagnosis of alzheimer disease, mild cognitive disorder, and other types of dementia. *Biomed Signal Process Control* 72:103293
5. Liu Q, Xi Y, Wang Q, Liu J, Li P, Meng X, Liu K, Chen W, Liu X, Liu Z (2021) Mannan oligosaccharide attenuates cognitive and behavioral disorders in the 5xFAD alzheimer's disease mouse model via regulating the gut microbiota-brain axis. *Brain Behav Immun* 95:330–343
6. Begemann MJ, Brand BA, Ćurčić-Blake B, Aleman A, Sommer IE (2020) Efficacy of non-invasive brain stimulation on cognitive functioning in brain disorders: a meta-analysis. *Psychol Med* 50(15):2465–2486
7. Pieruccini-Faria F, Black SE, Masellis M, Smith EE, Almeida QJ, Li KZ, Bherer L, Camicioli R, Montero-Odasso M (2021) Gait variability across neurodegenerative and cognitive disorders: results from the Canadian consortium of neurodegeneration in aging (CCNA) and the gait and brain study. *Alzheimers Dement* 17(8):1317–1328
8. Chen C, Liao J, Xia Y, Liu X, Jones R, Haran J, McCormick B, Sampson TR, Alam A, Ye K (2022) Gut microbiota regulate alzheimer's disease pathologies and cognitive disorders via PUFA-associated neuroinflammation. *Gut* 71(11):2233–2252
9. Zwaigenbaum L, Bauman ML, Stone WL, Yirmiya N, Estes A, Hansen RL, Wetherby A (2015) Early identification of autism spectrum disorder: recommendations for practice and research. *Pediatrics* 136(Supplement_1):S10–S40

10. Park C, Rosenblat JD, Brietzke E, Pan Z, Lee Y, Cao B, Zuckerman H, Kalantarova A, McIntyre RS (2019) Stress, epigenetics and depression: a systematic review. *Neurosci Biobehav Rev* 102:139–152
11. Baschi R, Luca A, Nicoletti A, Caccamo M, Cicero CE, Zappia M, Monastero R (2020) Changes in motor, cognitive, and behavioral symptoms in parkinson's disease and mild cognitive impairment during the COVID-19 lockdown. *Front Psych* 11:590134
12. Mahaman YAR, Huang F, Embaye KS, Wang X, Zhu F (2021) The implication of STEP in synaptic plasticity and cognitive impairments in Alzheimer's disease and other neurological disorders. *Front Cell Dev Biol* 9:680118
13. Silva NML, Barros-Aragão FG, De Felice FG, Ferreira ST (2022) Inflammation at the crossroads of COVID-19, cognitive deficits and depression. *Neuropharmacology* 209:109023
14. Tan MM, Lawton MA, Jabbari E, Reynolds RH, Iwaki H, Blauwendraat C, Kanavou S, Pollard MI, Hubbard L, Malek N, Grosset KA, Marrinan SL, Bajaj N, Barker RA, Burn DJ, Bresner C, Foltynie T, Wood NW, Williams-Gray CH, Hardy J, Nalls MA, Singleton AB, Williams NM, Ben-Shlomo Y, Hu MTM, Grosset DG, Shoaib M, Morris HR (2021) Genome-wide association studies of cognitive and motor progression in parkinson's disease. *Mov Disord* 36(2):424–433
15. Ferretti MT, Martinkova J, Biskup E, Benke T, Gialdini G, Nedelska Z, Rauen K, Mantua V, Religa D, Hort J, Santucci Chadha A, Schmidt R (2020) Sex and gender differences in alzheimer's disease: current challenges and implications for clinical practice: position paper of the dementia and cognitive disorders panel of the European Academy of Neurology. *Eur J Neurol* 27(6):928–943
16. Zlokovic BV, Gottesman RF, Bernstein KE, Seshadri S, McKee A, Snyder H, Greenberg SM, Yaffe K, Schaffer CB, Yuan C, Hughes TM, Daemen MJ, Williamson JD, González HM, Schneider J, Wellington CL, Katusic ZS, Stoeckel L, Koenig JI, Corriveau RA, Fine L, Galis ZS, Reis J, Wright JD, Chen J (2020) Vascular contributions to cognitive impairment and dementia (VCID): a report from the 2018 National Heart, Lung, and Blood Institute and National Institute of Neurological Disorders and Stroke workshop. *Alzheimers Dement* 16(12):1714–1733
17. Kumar S, Veldhuis A, Malhotra T (2021) Neuropsychiatric and cognitive sequelae of COVID-19. *Front Psychol* 12:577529
18. Perdoncin M, Konrad A, Wyner JR, Lohana S, Pillai SS, Pereira DG, Lakhani HV, Sodhi K (2021) A review of miRNAs as biomarkers and effect of dietary modulation in obesity associated cognitive decline and neurodegenerative disorders. *Front Mol Neurosci* 14:756499
19. McWhirter L, Ritchie C, Stone J, Carson A (2020) Functional cognitive disorders: a systematic review. *Lancet Psychiatry* 7(2):191–207

A Precise Modelling Endeavour for the Emission of Biodiesel Powered Diesel Engine Using Gene Expression Programming



Tien Han Nguyen, Anh Quan Nguyen, Duy Tan Nguyen, Minh Thai Duong, and Van Vang Le

Abstract The escalating fossil fuel costs and over-dependency on imports are driving the research in the domain of biodiesel. The increasing greenhouse gas emissions are further driving the investigation in this domain. In the present study, an engine is tested using waste-derived biodiesel. The focus was on engine emission, and how the use of biodiesel influences it. The data concerning the four emission species oxides of nitrogen (NO_x), unburnt hydrocarbon (UHC), carbon monoxide (CO), and carbon dioxide (CO₂) were collected. Since the formation of emissions in a compression ignition engine is a complex process, it is difficult to model using conventional approaches. Hence, an evolutionary machine learning technique Gene Expression Programming (GEP) was employed to model and predict the emission data in this study. The statistical evaluation of the data revealed that GEP could predict the data with more than 95.64% efficacy both during model training as well as model training phase. The model was also tested on different statistical metrics like mean squared error and mean absolute error. All these statistical metrics showed that GEP performed well throughout the model development.

T. H. Nguyen
Hanoi University of Industry, Hanoi, Vietnam

A. Q. Nguyen · D. T. Nguyen · V. V. Le (✉)
Institute of Maritime, Ho Chi Minh City University of Transport, Ho Chi Minh City, Vietnam
e-mail: levanvang@ut.edu.vn

A. Q. Nguyen
e-mail: quan.nguyen@ut.edu.vn

D. T. Nguyen
e-mail: tan.nguyen@ut.edu.vn

M. T. Duong
Institute of Mechanical Engineering, Ho Chi Minh City University of Transport, Ho Chi Minh City, Vietnam
e-mail: thai_ck@ut.edu.vn

Keywords Artificial intelligence • Machine learning • Exhaust emissions • Modelling • Prediction

1 Introduction

At present, there is a global challenge of increasing greenhouse gases and depleting the reserve of fossil fuels, especially petroleum products [1, 2]. On the other side, the energy demand for daily life activities and manufacturing is ever-increasing [3–5]. Diesel engines have been used for the last few decades for energy generation and transportation [6] because diesel engines have excellent fuel economy, excellent thermal efficiency, and amazing torque, they are extensively employed in the transportation industry [7, 8]. Still, fuels with higher combustion efficiency are desperately needed globally to alleviate the growing energy problem [9–11]. High emissions from diesel engines are another challenge; they include nitrogen oxides (NOx) and soot particles (PM), which are very dangerous to ecosystems and human health, in particular [11–13]. Under the above-described pressures, oxygenated fuels have become more and more popular among academics as potential solutions to increase combustion efficiency and reduce emissions in recent times [14–16].

Among the various options available, a practical substitute for traditional diesel is biodiesel, a cleaner burning fuel derived from plants, waste frying oil, and animal fats [17–19]. Production of biodiesel using the chemical process known as transesterification offers a number of benefits to the environment [20]. When weighed against petroleum diesel, biodiesel is biodegradable, produces less greenhouse gas emissions, and thus improves air quality [21, 22]. Sourced from non-fossil feedstock like soybean oil and used cooking oil, biodiesel offers a long-term substitute for finite fossil fuels [23]. Because feedstock agriculture stimulates rural economies and reduces reliance on imported oil, it also enhances energy security. That it works with the infrastructure and engines of today is another advantage [24]. Better public health results from the obvious lower emissions of biodiesel than from petroleum fuel. Most diesel automobiles may also use it; just minor modifications are required to get optimal performance. Its flexibility makes it a sensible option for a more sustainable future and cleaner travel [25, 26].

The performance of the engine and the emissions are influenced by the specifics of the fuel, the engine design, and the operating settings [27–29]. Using different types of fuel with the same engine can result in a variety of engine outputs. Biodiesel has various advantages over diesel, including lower heating value, higher density, higher viscosity, and a higher oxygen content percentage [30–32]. These features are of considerable significance regarding the output of the engine [33]. When modifying the operational characteristics of the engine, such as the injector type, fuel injection pressure, and injection time, it is important to take into account the biodiesel property [34]. The reasoning for this is that the characteristics of diesel fuel serve as the basis for the engine design [35, 36]. There is a correlation between rising viscosity and a reduction in the rate of atomization, for example. If the fuel is not fully atomized and

does not entirely mix with air, then it will not burn [37]. Because of this, the engine produces more elements that contribute to emissions and offers less performance. Moreover, high viscosity could cause deposit formation and coke in the mechanical parts of the combustion chamber [38–40]. In addition, an insufficient amount of fuel atomization is another factor that contributes to the production of unburnt hydrocarbons and carbon monoxide emissions [41, 42].

Meng et al. [43] investigated the effects of using waste cooking oil biodiesel (WCOB) on a diesel engine. It was reported that despite an increase in NO_x emissions, the other emission types showed decreasing trends. Gad et al. [44] used gasoline and WCOB blends to reduce the levels of engine emission from a diesel engine. According to the findings, the combination of gasoline and WCOB results in an increase in both the cylinder pressure and the HRR values. Both the fuel exergy and the energetic efficiency are increased when gasoline is added. In comparison to neat waste cooking oil, there is a considerable reduction in engine emissions of carbon monoxide, nitrogen oxides, and smoke opacity through 25%, 30%, 20%, and 30%, respectively, when using a mix of WCOB and diesel. It is possible to calculate that the ideal mixing ratio for the diesel-WCOB blend is BG8. This resulted in a large rise in emission forms, notably an increase in NO_x levels, while also delivering a value that is acceptable for fuel consumption. In a similar study by Gad et al. [45], the kerosene was blended with a WCOB- diesel mix. They concluded that using this blend the CO, HC, and smoke emissions were reduced while the NO_x emission was reportedly increased. While there are studies available reporting on the experimental work to find out the emission levels, there is not enough work on numerical modelling in this domain. The numerical modelling of emissions from engines is a complex process. Hence, several researchers reported using data-based Machine learning (ML) in this domain. Ahmad et al. [46] employed a hybrid approach of Genetic Algorithm (GA), Response Surface Methodology (RSM), and Particle Swarm Optimization (PSO) for the modelling and optimization of octanol-linseed biodiesel-diesel-powered engine. When compared to the diesel-based operation the test blends decreased HC, CO, and NO_x emissions by 6.55%, 18.75%, 6.55%, and 0.72%, respectively. Uslu et al. [47] employed RSM for modelling optimization of diesel engines powered with hemp biodiesel. The RSM could help in optimizing the engine for the least possible engine emission without much compromising on engine power.

Several other ML approaches have been used in the past for this purpose like Gaussian Process Regression (GPR) [48], Artificial Neural Network (ANN) [49], and several other ML approaches were used in the past. The literature reveals that ML-based approaches can help in the prediction modelling in complex modelling of biodiesel-powered engine emission. Hence, in the present study an evolutionary algorithm namely Gene Expression Programming (GEP) is used for modelling the prediction of engine emission. Several statistical approaches will be used for assessing the performance of GEP.

Table 1 Test engine specification

Parameter	Specification
Make	Kirloskar, India
Power	5.2 kW
Fuel injection pressure	200 bar
Cooling	Water
Stroke	4
Loading	Dynamometer
Speed	1500 rpm

Table 2 Test fuel specification

Parameter	WCOB	Diesel
Density, kg/m ³	840	831
Cetane	51	60
Fire point, °C	49	70
Higher heating value, kJ/kg	42,174	43,500
Viscosity, cs	3.41	2.99

2 Materials and Methods

2.1 Test Set up and Test Fuel

A test-bed mounted 5.2 kW diesel engine, water-cooled, and four strokes were employed in this study. This test setup has provision for fuel and airflow measurements. A dynamometer working on the eddy current principal was used for engine loading. Table 1 lists the main specifications of the test engine. In this study, the waste cooking oil was transformed into biodiesel using the well-established transesterification method. The fuel characteristics are listed in Table 2 (Fig. 1).

2.2 Machine Learning

In the present study, an evolutionary ML approach namely GEP was employed for model prediction. The data collected from lab-based testing was employed for model development. Regression issues in particular have shown considerable promise for the advanced evolutionary technique known as gene expression programming (GEP). Candida Ferreira developed GEP in 2001 [50] to simulate natural evolutionary processes to build computer algorithms that can handle challenging issues. Unlike genetic algorithms and genetic programming, GEP encodes solutions as nonlinearly expressed linear chromosomes [51, 52].

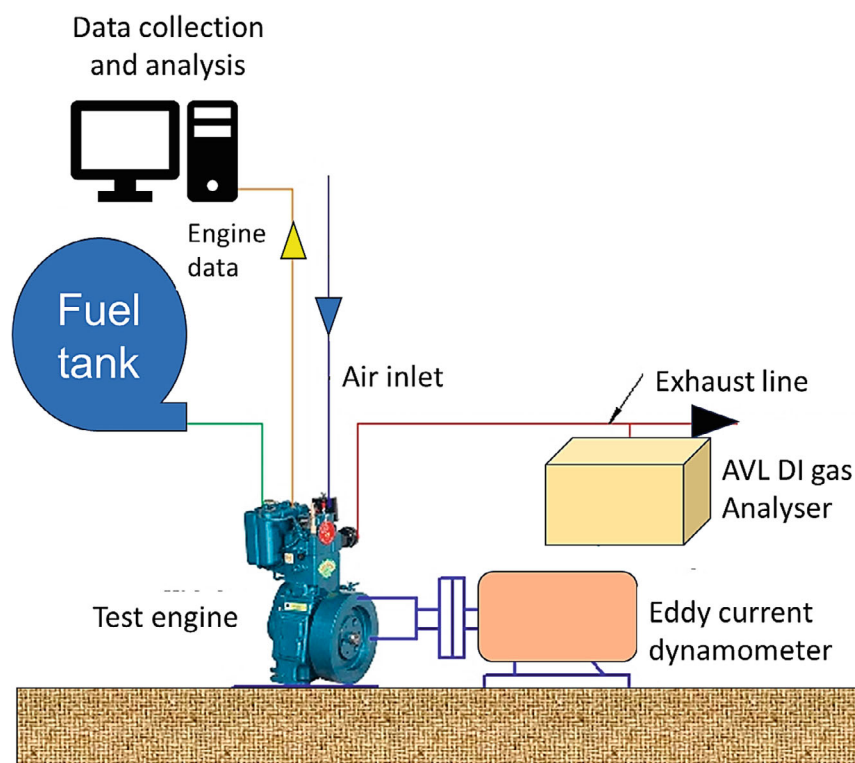


Fig. 1 Test setup used in the study

GEP is used in regression to find functional relationships between input variables and intended results. Each random chromosome in the starting population of the technique indicates a possible solution. Regression model accuracy on training data is a standard way to measure how well GEP produces these chromosomes over succeeding generations via repeated processes of selection, crossover, mutation, and transposition to promote fitness [53]. The capacity of GEP to provide interpretable models is one advantage in regression difficulties. One may transform the developing programs into mathematical formulas that provide details about the underlying data relationships. In fields where understanding the model is as crucial as predicting accuracy, such as biology, economics, and environmental modelling, this openness is particularly helpful [54, 55].

GEP's flexibility also allows it to easily manage a variety of data types and nonlinear relationships. Because of its robust regression performance and interpretability of the resulting models, gene expression programming is a valuable instrument in the machine learning arsenal for researchers tackling challenging predictive modelling problems.

3 Results and Discussion

3.1 Modelling and Prediction

Once the experiment phase was completed the data collected was employed for model development using GEP. The data was randomly split into two parts, the first 70% was employed for model training while the rest was used for model testing. After the model development, these were used for predictions. Four different prognostic models each one for NO_x, CO, UHC, and CO₂ data. The correlation matrix is depicted in Fig. 2. It shows data distribution and range for various parameters.

The actual versus predicted and residual results are depicted in Fig. 3. Figure 3a shows the model performance for the CO emission model. It was observed that the model performed well with R^2 value as high as 0.9638 in model testing and MSE was almost negligible. The residual plots also depict that model errors were quite low. Figure 3b displays the CO₂ emission model performance. It had the R^2 values as high as 0.9564 in model testing, it was shown that the model worked well and that the MSE was 0.0348. Furthermore, showing that model errors were quite minimal were residual plots. The NO_x emission model's performance is shown in Fig. 3c. In model testing, it was found that the model worked well, with an R^2 value as high as 0.9928 and a MSE of 1945.26. The residual plots also show that there were very few model errors. In the case of the HC emission model, as depicted in Fig. 3d, it was noted that it achieved an R^2 value as high as 0.9714 and an MSE of 0.2521. The residual plots for the HC emission model also show that it has very few model errors.

3.2 Model Comparison with Statistical Metrics

A GEP-based ML-trained engine emission model was evaluated using these metrics, which include R^2 , MSE, and MAE. For the training dataset as well as the testing dataset, the data was evaluated using these metrics. The results of all three metrics are depicted in Fig. 4. A training R^2 score of 0.9954 and a testing R^2 value of 0.9638 indicate that the model predicts CO with a high degree of accuracy. Since both of the MSEs are negligible, the forecasts are exact. However, the low MAE numbers (0.0011 for training and 0.0041 for testing) suggest that the average errors are really small. Apart from that, the CO₂ model performs well during testing ($R^2 = 0.9564$)

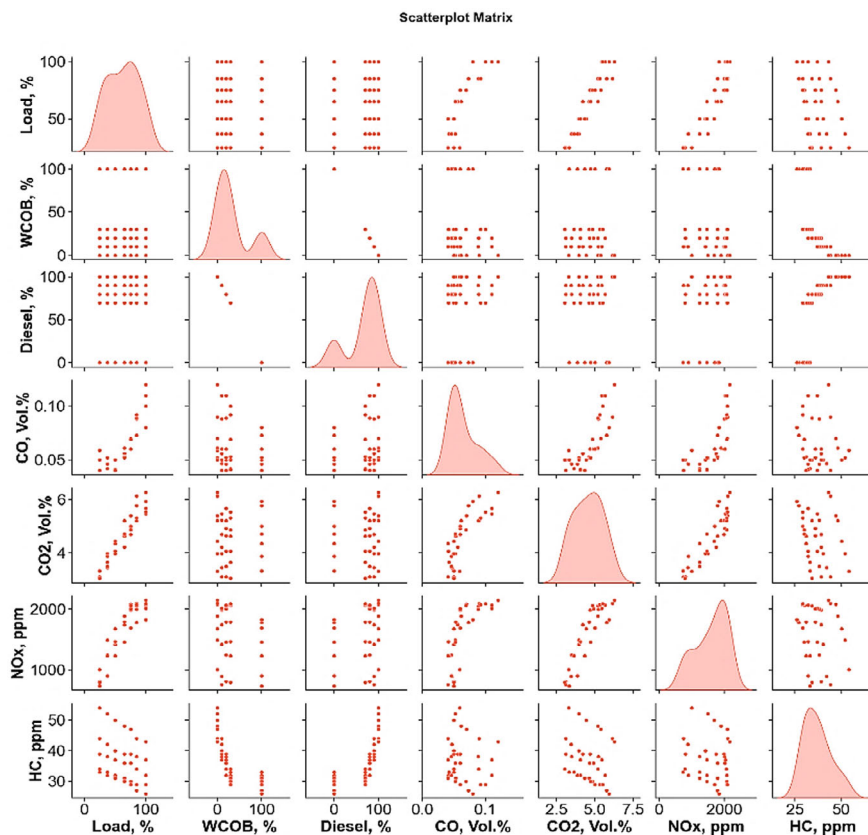


Fig. 2 Correlation matrix

and with a high degree of accuracy during training ($R^2 = 1$). With a mean squared error (MSE) for testing of 0.0348 and mean absolute error (MAE) of 0.0028 (training) and 0.154 (testing), the model's predictive power for the test set is still quite strong. Concerning NO_x, the model matches the training data exactly (the R^2 value is one) and has a testing R-squared value of 0.9928. The tests yielded an MSE of 1945.26 and an MAE of 40.61, indicating some variations in the NO_x emissions projections. The HC model fits training data perfectly ($R^2 = 1$) and tests data with an astonishingly high R^2 value of 97.14. With only small prediction errors, the model performs well, as shown by the testing mean squared error (MSE) and MAE of 0.4627. In general, the carbon dioxide and carbon monoxide models show a fair degree of generality and minimal prediction errors. While the HC model functions well the NO_x model provides reliable training and test data fits.

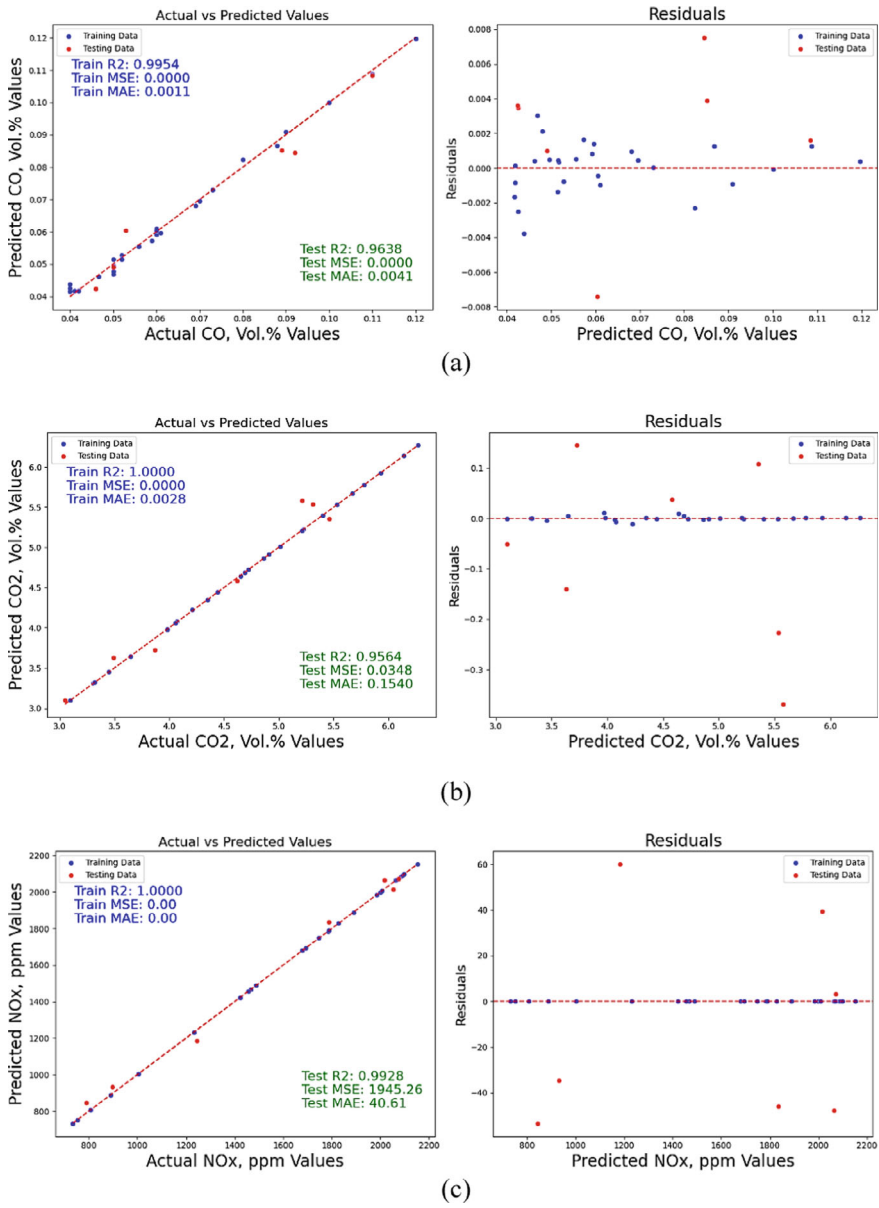


Fig. 3 Model results and error analysis

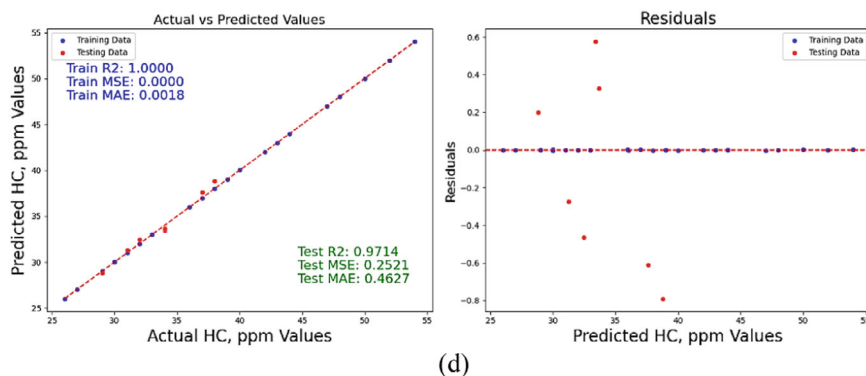


Fig. 3 (continued)

4 Conclusion

The prognostic analysis of emission data from a biodiesel-diesel blends powered engine is evaluated in this work. Biodiesel was synthesized from waste cooking oil through transesterification process. The data was gathered on the four emission species: carbon monoxide (CO), carbon dioxide (CO₂), unburnt hydrocarbon (UHC), and oxides of nitrogen (NO_x). The evolutionary ML approach GEP helped to model and forecast the emission data. Both during the model training and model testing stages, the statistical analysis of the data revealed that GEP could anticipate the data with a 95.64% efficacy with very minimal errors. Two other statistical measures used in evaluation of the model were mean absolute error and mean squared error. These all-statistical tests showed that GEP performed reasonably well all through the model-building procedure.

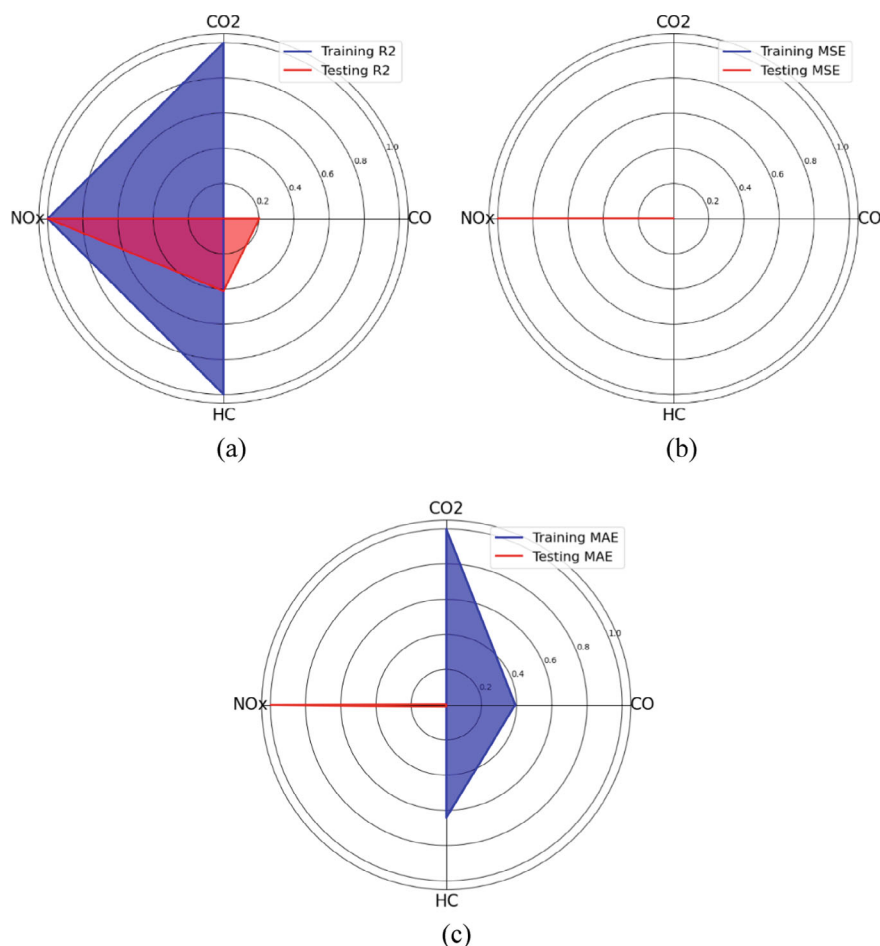


Fig. 4 Model comparison on the basis of statistical metrics

References

1. Chau MQ, Nguyen DC, Hoang AT, Tran QV, Pham VV (2020) A numeral simulation determining optimal ignition timing advance of SI engines using 2.5-dimethylfuran-gasoline blends. *Int J Adv Sci Eng Inf Technol* 10:1933–1938. <https://doi.org/10.18517/ijaseit.10.5.13051>
2. Haq ES, Setiadevi S, Rini EM, Yusuf D, Aji AA (2023) Investigating an enhanced approach for greenhouse climate control: optimising cooling and heating systems. *Int J Adv Sci Eng Inf Technol* 13:2388–2396. <https://doi.org/10.18517/ijaseit.13.6.19384>
3. Chamkalani A, Zendeboudi S, Rezaei N, Hawboldt K (2020) A critical review on life cycle analysis of algae biodiesel: current challenges and future prospects. *Renew Sustain Energy Rev* 134:110143. <https://doi.org/10.1016/j.rser.2020.110143>
4. Hoang A, Tran Q, Pham X (2018) Performance and emission characteristics of popular 4-stroke motorcycle engine in Vietnam fuelled with biogasoline compared with fossil gasoline. *Int J Mech Mechatronics Eng* 18:97–103

5. Munahar S, Setiyo M, Saudi MM, Ahmad A, Yuvenda D (2022) Modelling fuel cut off controller on CNG engines using fuzzy logic: a prototype. *Int J Adv Sci Eng Inf Technol* 12:1857–1865. <https://doi.org/10.18517/ijaseit.12.5.16849>
6. Cao DN, Johnson AJT (2024) A simulation study on a premixed-charge compression ignition mode-based engine using a blend of biodiesel/diesel fuel under a split injection strategy. *Int J Adv Sci Eng Inf Technol* 14:451–471. <https://doi.org/10.18517/ijaseit.14.2.20007>
7. Hoang AT, Pham VV (2018) A review on fuels used for marine diesel engines. *J Mech Eng Res Dev* 41:22–32
8. Doan QB, Nguyen XP, Pham VV, Dong TMH, Pham MT, Le TS (2022) Performance and emission characteristics of diesel engine using ether additives: a review. *Int J Renew Energy Dev* 11(1). 1014710/Ijred202242522
9. Yuvenda D, Sudarmanta B, Wahjudi A, Hirowati RA (2022) Effect of adding combustion air on emission in a diesel dual-fuel engine with crude palm oil biodiesel compressed natural gas fuels. *Int J Renew Energy Dev* 11:871–877. <https://doi.org/10.14710/ijred.2022.41275>
10. Nguyen VN, Sharma P, Kumar A, Pham MT, Le HC, Truong TH, et al (2023) Optimization of biodiesel production from Nahar oil using Box-Behnken design, ANOVA and grey wolf optimizer. *Int J Renew Energy Dev* 12:711–719. <https://doi.org/10.14710/ijred.2023.54941>
11. Changxiong L, Hu Y, Yang Z, Guo H (2023) Experimental study of fuel combustion and emission characteristics of marine diesel engines using advanced fuels. *Polish Marit Res* 30:48–58. <https://doi.org/10.2478/pomr-2023-0038>
12. Yang Z, Tan Q, Geng P (2019) Combustion and emissions investigation on low-speed two-stroke marine diesel engine with low sulfur diesel fuel. *Polish Marit Res* 26:153–161. <https://doi.org/10.2478/pomr-2019-0017>
13. Paramasivama P, Naima K, Dzida M (2024) Soft computing-based modelling and optimization of NOx emission from a variable compression ratio diesel engine. *J Emerg Sci Eng* 2:e21. <https://doi.org/10.61435/jese.2024.e21>
14. Nguyen H-B, Nguyen V-T-A, Ly V-D, Bui T-A (2022) Biodiesel produced from pangasius oil operating a diesel engine: case study in Vietnam. *Int J Adv Sci Eng Inf Technol* 12:477–482. <https://doi.org/10.18517/ijaseit.12.2.16159>
15. Elkelay M, Alm-Eldin Bastawissi H, El Shenawy EA, Taha M, Panchal H, Sadasivuni KK (2021) Study of performance, combustion, and emissions parameters of DI-diesel engine fueled with algae biodiesel/diesel/n-pentane blends. *Energy Convers Manag* X 10:100058. <https://doi.org/10.1016/j.ecmx.2020.100058>
16. Hoang AT, Nguyen VT (2017) Emission characteristics of a diesel engine fuelled with preheated vegetable oil and biodiesel. *Philipp J Sci* 146:475–482
17. Wei C, Jiang G, Wu G, Zhou Y, Liu Y (2024) Effects on of blended biodiesel and heavy oil on engine combustion and black carbon emissions of a low-speed two-stroke engine. *Polish Marit Res* 31:94–101. <https://doi.org/10.2478/pomr-2024-0010>
18. Alatas A, Putri DN, Perdani MS, Utami TS, Hermansyah H (2021) Biodiesel synthesis through the non-alcohol route using the immobilized lipase enzyme from bacillus subtilis as biocatalyst. *Int J Adv Sci Eng Inf Technol* 11:1239–1245. <https://doi.org/10.18517/ijaseit.11.3.14784>
19. Le TT, Kumar R, Roy MK, Mishra MK, Mahto PK, Balasubramanian D, et al (2024) An experimental assessment of waste transformer oil and palm oil biodiesel blended with diesel fuel on a single cylinder direct in diesel engine. *Int J Adv Sci Eng Inf Technol* 14:246–258. <https://doi.org/10.18517/ijaseit.14.1.15998>
20. Dong VH, Sharma P (2023) Optimized conversion of waste vegetable oil to biofuel with meta heuristic methods and design of experiments. *J Emerg Sci Eng* 1:22–28. <https://doi.org/10.61435/jese.2023.4>
21. Wang S, Zhang Z, Hou X, Lv J, Lan G, Yang G et al (2023) The environmental potential of hydrogen addition as complementation for diesel and biodiesel: a comprehensive review and perspectives. *Fuel* 342:127794. <https://doi.org/10.1016/j.fuel.2023.127794>
22. Nguyen VG, Pham MT, Le NVL, Le HC, Truong TH, Cao DN (2023) A comprehensive review on the use of biodiesel for diesel engines. *Int J Renew Energy Dev* 12:720–740. <https://doi.org/10.14710/ijred.2023.54612>

23. Bhuiya MMK, Rasul MG, Khan MMK, Ashwath N, Azad AK, Hazrat MA (2016) Prospects of 2nd generation biodiesel as a sustainable fuel—part 2: properties, performance and emission characteristics. *Renew Sustain Energy Rev* 55:1129–1146. <https://doi.org/10.1016/j.rser.2015.09.086>
24. Sajjadi B, Raman AAA, Arandiyani H (2016) A comprehensive review on properties of edible and non-edible vegetable oil-based biodiesel: composition, specifications and prediction models. *Renew Sustain Energy Rev* 63:62–92
25. Hoekman SK, Broch A, Robbins C, Cenicerio E, Natarajan M (2012) Review of biodiesel composition, properties, and specifications. *Renew Sustain Energy Rev* 16:143–169
26. Gebremariam SN, Marchetti JM (2018) Economics of biodiesel production. *Energy Convers Manag* 168:74–84
27. More GV, Koli SR, Rao YVH, Prasad PI, Rao BN (2020) Effect of compression ratio on compression ignition engine with RUCO biodiesel/ diethyl ether/diesel fuel blends. *Energy Sources Part A Recover Util Environ Eff* 46(1):10747–10766. <https://doi.org/10.1080/15567036.2020.1785593>
28. Goga G, Chauhan BS, Mahla SK, Dhir A, Cho HM (2021) Combined impact of varying biogas mass flow rate and rice bran methyl esters blended with diesel on a dual-fuel engine. *Energy Sources Part A Recover Util Environ Eff* 43:120–132. <https://doi.org/10.1080/15567036.2019.1623948>
29. Dimitriou P, Tsujimura T, Suzuki Y (2019) Adopting biodiesel as an indirect way to reduce the NOx emission of a hydrogen fumigated dual-fuel engine. *Fuel* 244:324–334. <https://doi.org/10.1016/j.fuel.2019.02.010>
30. Rajalingam A, Jani SP, Kumar AS, Khan MA (2016) Review article production methods of biodiesel. *J Chem Pharm Res* 8:170–173
31. Khoobbakht G, Kheiralipour K, Karimi M (2021) Optimization of Chlamydomonas alga biodiesel percentage for reducing exhaust emission of diesel engine. *Process Saf Environ Prot* 152:25–36. <https://doi.org/10.1016/j.psep.2021.05.008>
32. Nguyen XP, Vu HN (2019) Corrosion of the metal parts of diesel engines in biodiesel-based fuels. *Int J Renew Energy Dev* 8:119–132. <https://doi.org/10.14710/ijred.8.2.119-132>
33. Al Zubi MA, Penmetsa RV, Kumar PS, Patil PP, Singh B, Alsubih M et al (2024) Performance and emission characteristics of novel biodiesel-diesel blends: an RSM and ANN approach. *Int J Low-Carbon Technol* 19:1060–1077. <https://doi.org/10.1093/ijlct/ctae025>
34. Shameer PM, Ramesh K (2018) Assessment on the consequences of injection timing and injection pressure on combustion characteristics of sustainable biodiesel fuelled engine. *Renew Sustain Energy Rev* 81:45–61. <https://doi.org/10.1016/j.rser.2017.07.048>
35. Mishra VK, Goswami R (2018) A review of production, properties and advantages of biodiesel. *Biofuels* 9:273–289
36. Noor CWM, Noor MM, Mamat R (2018) Biodiesel as alternative fuel for marine diesel engine applications: a review. *Renew Sustain Energy Rev* 94:127–142
37. Chigier NA (1979) The atomization and burning of liquid fuel sprays. In *Energy and combustion science*, pp 183–200. <https://doi.org/10.1016/B978-0-08-024780-9.50014-9>
38. Hoang AT, Pham VV (2019) A study of emission characteristic, deposits, and lubrication oil degradation of a diesel engine running on preheated vegetable oil and diesel oil. *Energy Sources Part A Recover Util Environ Eff* 41:611–625. <https://doi.org/10.1080/15567036.2018.1520344>
39. Hoang AT, Le AT (2019) A review on deposit formation in the injector of diesel engines running on biodiesel. *Energy Sources Part A Recover Util Environ Eff* 41:584–599. <https://doi.org/10.1080/15567036.2018.1520342>
40. Hoang AT, Le VV, Pham VV, Tham BC (2019) An investigation of deposit formation in the injector, spray characteristics, and performance of a diesel engine fueled with preheated vegetable oil and diesel fuel. *Energy Sources Part A Recover Util Environ Eff* 41:2882–2894. <https://doi.org/10.1080/15567036.2019.1582731>
41. Hoang AT, Pham VV (2019) Impact of jatropha oil on engine performance, emission characteristics, deposit formation, and lubricating oil degradation. *Combust Sci Technol* 191:504–519. <https://doi.org/10.1080/00102202.2018.1504292>

42. Hoang AT, Le AT (2019) Trilateral correlation of spray characteristics, combustion parameters, and deposit formation in the injector hole of a diesel engine running on preheated jatropha oil and fossil diesel fuel. *Biofuel Res J* 6:909–919. <https://doi.org/10.18331/BRJ2019.6.1.2>
43. Meng J, Xu W, Meng F, Wang B, Zhao P, Wang Z et al (2023) Effects of waste cooking oil biodiesel addition on combustion, regulated and unregulated emission characteristics of common-rail diesel engine. *Process Saf Environ Prot* 178:1094–1106. <https://doi.org/10.1016/j.psep.2023.08.065>
44. Gad MS, El-Seesy AI, Radwan A, He Z (2020) Enhancing the combustion and emission parameters of a diesel engine fueled by waste cooking oil biodiesel and gasoline additives. *Fuel* 269:117466
45. Gad MS, Ismail MA (2021) Effect of waste cooking oil biodiesel blending with gasoline and kerosene on diesel engine performance, emissions and combustion characteristics. *Process Saf Environ Prot* 149:1–10. <https://doi.org/10.1016/j.psep.2020.10.040>
46. Ahmad A, Yadav AK, Singh A, Singh DK, Ağbulut Ü (2024) A hybrid RSM-GA-PSO approach on optimization of process intensification of linseed biodiesel synthesis using an ultrasonic reactor: enhancing biodiesel properties and engine characteristics with ternary fuel blends. *Energy* 288:129077. <https://doi.org/10.1016/j.energy.2023.129077>
47. Uslu S, Simsek S, Simsek H (2023) RSM modeling of different amounts of nano-TiO₂ supplementation to a diesel engine running with hemp seed oil biodiesel/diesel fuel blends. *Energy* 266:126439. <https://doi.org/10.1016/j.energy.2022.126439>
48. Alruqi M, Sharma P, Deepanraj B, Shaik F (2023) Renewable energy approach towards powering the CI engine with ternary blends of algal biodiesel-diesel-diethyl ether: bayesian optimized gaussian process regression for modeling-optimization. *Fuel* 334:126827. <https://doi.org/10.1016/j.fuel.2022.126827>
49. Lotfan S, Ghiasi RA, Fallah M, Sadeghi MH (2016) ANN-based modeling and reducing dual-fuel engine's challenging emissions by multi-objective evolutionary algorithm NSGA-II. *Appl Energy* 175:91–99. <https://doi.org/10.1016/j.apenergy.2016.04.099>
50. Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst* 13:87–129. <https://doi.org/10.48550/arXiv.cs/0102027>
51. Seckin G, Yilmaz T, Guven A, Yuceer A, Basibuyuk M, Ersu CB (2011) Modeling the performance of upflow anaerobic filters treating paper-mill wastewater using gene-expression programming. *Ecol Eng* 37:523–528. <https://doi.org/10.1016/J.ECOLENG.2010.12.002>
52. Hu J, Guo P, Poh K-L (2020) Generating decision rules for flexible capacity expansion problem through gene expression programming. *Comput Oper Res* 122:105003. <https://doi.org/10.1016/j.cor.2020.105003>
53. Sharma P (2021) Artificial intelligence-based model prediction of biodiesel-fueled engine performance and emission characteristics: a comparative evaluation of gene expression programming and artificial neural network. *Heat Transf* 50:5563–5587. <https://doi.org/10.1002/hjt.22138>
54. Gholampour A, Gandomi AH, Ozbakkaloglu T (2017) New formulations for mechanical properties of recycled aggregate concrete using gene expression programming. *Constr Build Mater* 130:122–145. <https://doi.org/10.1016/j.conbuildmat.2016.10.114>
55. Mohammed AY, Sharifi A (2020) Gene expression programming (GEP) to predict coefficient of discharge for oblique side weir. *Appl Water Sci* 10:145. <https://doi.org/10.1007/s13201-020-01211-5>

Analysis and Prediction of Water Access and Hygiene on Resource Data Using Machine Learning



R. Gunasundari and D. Sakthivel

Abstract Improved drinking water, sanitation, and hygiene amenities are virtually unavailable in the majority of India's locations. We suggest the new combination Machine Learning (ML) approaches of K-Means clustering algorithm with the Support vector machine (SVM). This proposed method is used for the access of water and analysis of hygiene prediction on resource data using this suggested machine learning technique. This newly created hybridized K-Means+SVM algorithm gives the better overall performances, and it achieves the best output results. For this comparison purpose, we are taking the current models used in the water access and hygiene prediction is, Decision Tree, Logistic regression, and KNN. The overall performances between these existing and proposed methods are measured by using the performance evaluation parameters such as, accuracy, sensitivity, recall, precision, and time duration. Our newly constructed K-Means+SVM method obtained the accuracy of 94.70% with the less time duration of 20.63 ms. This method will be most widely suitable for quickly access for water level and hygiene prediction on resource data using the different machine learning methods in future.

Keywords Analysis of water access · Hygiene of water · Resource data · Prediction · Machine learning · Support vector machine · K-Means

R. Gunasundari (✉)

Department of Computer Applications, Faculty of Arts, Science, Commerce and Management,
Karpagam Academy of Higher Education, Coimbatore, India
e-mail: gunasoundar04@gmail.com

D. Sakthivel

Department of Computer Applications, Karpagam Academy of Higher Education, Coimbatore,
India

1 Introduction

Chlamydia, gonorrhea, cancrroid, syphilis, and HIV are the main causes of reproductive tract infections (RTIs), however endogenous and iatrogenic diseases associated with mishandled abortions and subpar delivery techniques can also cause RTIs. A random tract infection (RTI) can cause a woman to feel as though she has a burning sensation in her inguinal area, inguinal swelling, vulvar itching, genital ulcers, lower abdomen pain, vulvar itching, irregular vaginal discharge, and nausea. Among women in their reproductive years in poor nations, RTIs are the most common cause of sickness more than in any other country. But public health services, research, and interventions seldom take RTIs into account revealed surveyed 2022 to 2023 more than RTI, a number nationally representative studies suggested that ranged to 18%. According to estimates from medical specialists in India, 40 million new RTI cases are reported each year in the nation. Common water level access and hygiene system is shown in Fig. 1.

Contamination of drinking water illness death 300,000 children adults perish inadequate per year. As of 2023, an estimated two billion people’s sources of drinking water were tainted with human waste. In low-income countries (LICs), communal settings, such focus targets WaSH programs aimed at reducing disease exposure in frequently visited sites outside the home. In these cases, there are limited and infrequent service quality monitoring programs in these countries, and there are few, poorly supported WaSH services that protect users from pathogen contact. Research from households demonstrates the impact of WaSH activities on non-home situations, such as schools. Sanitation practices, safe storage, and water treatment are associated with a decreased occurrence of intestinal protozoa infections and diarrheal illness, according to meta-analyses of household WaSH treatments. Although there are differences in water storage and usage patterns across Honduras, cross-sectional studies have shown that good markers.

Schools have a special chance to enhance health in regard because significant amount the day there. The low quality of school drinking water exposes school-age children to infections spread through water, even in cases where furthermore, more vulnerable watery illnesses like diarrhea than adults are. Hygiene and sanitation

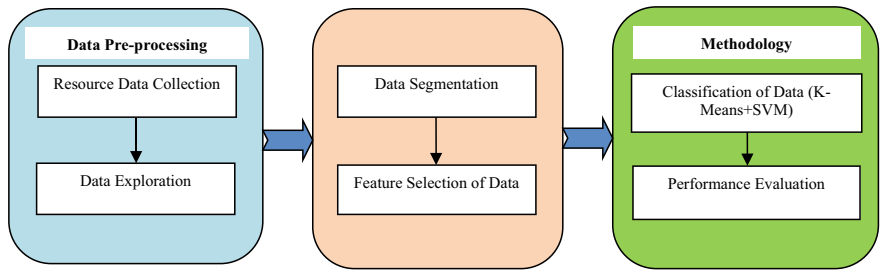


Fig. 1 Proposed system work

in school water enhances enrollment, gender parity, health-related information and habits, and reduces respiratory and parasitic illnesses. In schools, efforts to purify the water supply and promote hand washing are also helpful in lowering absenteeism, particularly among female pupils said by Ritchie and Roser [1].

1.1 Measuring Empowerment

The inability of these instruments to assess power dynamics in the home, community, and society all crucial components of comprehending women's empowerment is one of their limitations. Furthermore, some researchers contend that these measures are not sensitive to individual trait analysis based on age or other social identities, nor do they take context-specific variables into consideration. Furthermore, a small set of indicators may limit the scope of actions that may be taken with data from DHS explained by Zar et al. [2]. To address some of these shortcomings, in the agricultural sector using individual-level data. Men and women who share a family are empowered by the tool, which highlights their accomplishments in five key areas: leadership, time management, resource ownership, income and control, and agricultural production. The program has been used to monitor farming initiatives in a number of countries, including Nepal. The data that was collected addressed a wide range of topics:

1. Incidents of open defecation and the close are instances of containment;
2. Transport emptying: residences in each city are connected to local or sewage systems;
3. Treatment facilities, which include the kind and size of facilities for treating sanitation;
4. Practices pertaining to personal hygiene, like cleaning your hands before eating, consuming treated water, and obtaining your drinking water from natural sources.

Two months following the survey date, a thousand completed questionnaires that had been given to local leaders were gathered. Key informant interviews and field research were used to collect data in addition to the questionnaire survey. Key informants for this study included utility suppliers (8 or 5%), frightened families (55 or 39%), administrators (27 or 19%), and others (38 or 27%). Additionally, 14 (or 10%) local leaders were included. Supervisors Municipalities have critical zones. Both local and federal government organizations offer the FSM services. Usually, contractual arrangements are used to involve utility companies. Promoting involvement and support for FSM has been greatly aided by concerned households, local authorities, and other parties. The virulence, oral transmission through faces, and environmental adaptability of *Salmonella* species make them an important indicator for investigating the relationship between waste management techniques and diarrhea prevalence said by Tortajada [3].

2 Literature Review

An essential component of wise and efficient water governance and water priorities is the evaluation of water insecurity. Previous research has included surveys, questionnaires, and reports to identify the primary causes of water insecurity. The association between demographic data, inter- and local disputes over water in Africa, and water insecurity was examined by Pearson et al. [4]. They developed five questions about the amount, quality, price, accessibility, and dependability of water sources in relation to water insecurity, they only offered two answers for each subject. The author discovered a positive connection between water scarcity, accessibility, and interpersonal discord. The short answers that are obtained restrict one's comprehension of the fundamental problems related to water insecurity, even if this approach helps identify people's worries possible because survey is labor-intensive and difficult to translate to a large-scale study, it is restricted to a small number of communities.

Research on water insecurity is also being done northern side, where there is a serious deficiency in basic infrastructure and seriously problematic water quality. Roque et al. [5] conducted most recent studies, focusing on government funding and support initiatives, existing root causes the problem. Schmidt [6] distinguished between four groups of conditions pertaining to water access in the colonias: reasonably unsafe, secure. He made distinction when thinking about the problems related to water scarcity in the colonias. Waterborne diseases and pollutant mixes are two problems related to water quality that have been researched for their geographical distribution using several clustering techniques, including clustering, and Bayesian clustering, fuzzy C-means clustering analyzed by Shin et al. [7].

Third, mixed-type data quality that is numerical and categorical data is hampered by the fact that standard fourth, since previous clustering studies typically downplay the significance of the variables that set apart each group's water access circumstances from the others, it is difficult to identify the unique features of each cluster and conduct a thorough analysis. Improving the explain ability essential guarantee validity findings and any policy choices based on them Kedron et al. [8].

Meehan et al. [9] apply a data-driven approach in this work that blends machine learning based on decision trees with clustering in order to tackle the previously described issues. Our main objective is to comprehend the problems surrounding water availability in the 2000 US colonies, where disregard for these issues has led to really terrifying living arrangements. Governments and development partners must move swiftly to deploy resources in order to alter norms in society and culture surrounding handwashing and increase entry to hand personal cleanliness facilities with water and soap (UNICEF and WHO, 2023) which in turn calls for knowledge of facilities for hand hygiene and behaviors at the province sub-level, said Moswatsi [10].

The primary sources of information on hand hygiene. MICS and DHS surveys are multipurpose and relay on quick inspections of clean facilities in order to strike a balance between cost-effectiveness and representativeness. Unlike other systematic methods of gathering small-scale data, including structured behavior observations,

these polls address cover a broader geographic area and a greater variety of health-related and demographic factors taken from Han et al. [11]. Local coverage estimates should be created for every place within a nation in order to gain an improved comprehension of the accessibility of basic hygiene services from household surveys and to guide development strategy, deployment of resources, and intervention execution. Rios Quituizaca et al. [12] give the details of Bayesian geo-statistical modeling is a technique that may be used to produce geospatially explicit estimates for each site. It has extensive experience mapping various public health and demographics variables using house survey data. One common use of the MCMC method, which makes use of computer programs like WinBUGS, is the construction of such Bayesian geo-statistical models. Its implementation usually suffers from poor convergence, large storage requirements, and high processing costs due to the vast covariate matrices needed. In order to solve these challenges, mapping health-related and demographic variables from household survey data has become more popular, this is partially due to the INLA method, or Integrated Nested Laplace Approximation by Atilola [13].

Wu et al. [14] with the use of a stochastic partial differential equation (SPDE), INLA creates a triangulation, or mesh, throughout the research region to calculate the dataset's geographical autocorrelation structure at the mesh vertices. A recent research calculated buffer lengths between sample observation locations to include geographic proximity effects into a Random Forest (RF) model, despite the RF approach's primary shortcoming of disregarding spatial autocorrelation. Talebi et al. [15], Although the forest for randomness spatial predictions (RFsp) approach has demonstrated results similar to geo-statistical models, it may not be feasible for large-scale data applications due to the significant processing expense correlation with determining buffer distances for each an example of an observation point. The statistical models assess the premise that two variables are connected in order to produce a numerical assessment of connection. Mertler et al. [16], these statistical techniques can be used if the two variables of interest are known, or if a hypothesis has previously been formed. For a trustworthy examination of socioeconomic linkages with natural resources, the methods used to determine how socioeconomic variables impact water usage can be trusted by Leliveld and Knorrnga [17].

3 System Design

3.1 Hybridization of K-Means and SVM Algorithm

The water access and hygiene level of quality for machine learning technique is done by the combination of two ML approaches. Hybridization of K-Means algorithm with the Support Vector machine is the newly created and our suggested algorithm to find and analysis of water access and prediction of water hygiene in resource data. Once the process of K-Means algorithm completed, then this data output is fed into the SVM algorithm for the classification purpose. This integration of the

two ML algorithms gives the better result measured in all factors and overall performances. The proposed work flow diagram is shown in Fig. 1. The overall working process explained in the following section. It includes the data gathering and prior to processing, segmenting data and extraction of features, data grouping using proposed K-Means+SVM method.

3.2 K-Means Clustering Algorithm

K-means clustering algorithm for clustering a straightforward unsupervised method for segmentation item discovery is k means clustering. The following graphics is used to calculate the total number of clusters and cluster center's using this method to accomplish effective clustering, cluster sites are selected so that they remain safely apart from the cluster core. The pixels are prematurely combined once the linking of the pixels in many cluster is completed. Next, new cluster center's are located and assigned data points based on distance measurement.

Figure 2 displays the pseudocode for the k-means clustering technique. According to Sinaga and Yang [18], the k-means method minimizes the distance between a data point and its centroid by applying iteratively two phases. This technique gathers a collection of data attributes from the vector space and uses them to identify whether the image contains any naturally occurring clusters. It uses relatively simple procedures to partition the pixels into a predetermined number of cluster.

A partition's member items and centroid define its clusters. The centroid of a cluster is the position at which every item in the cluster is separated from every other item by the total distance. Across all clusters, the centroid-based method minimizes the total distances between each item and its cluster centroid through an iterative process. Assuming a set of data points, $X = \{x_1, x_2, x_3, \dots, x_n\}$, and a collection of centers, $V = \{v_1, v_2, \dots, v_c\}$.

1. Select "c" cluster center at random.
2. Ascertain how far separately for every data point is from the cluster center.

Algorithm : k-means Clustering
 Step 1: Input data
 Step 2: Get the values of k and chose the cluster centre's
 Step 3: Determine the seperation between pixels and cluster centre's
 Step 4: Assign every pixels to the nearest cluster center
 Step 5: Compute the mean i.e.cluser for each cluster
 Step 6: Check whether there is any variation in cluster centers
 Case1: if yes then set the new mean as cluster center and follow Step 3
 Case2: if no, then compute the statistics and separabe information
 Step 7: Output cluster segmented data

Fig. 2 Pseudocode for K-means algorithm

3. The cluster center with the minimal separation between it and the other cluster centers should be assigned the data point.
4. Utilizing, compute the new cluster center:

$$v_i = (1/c_i) \sum x_{ij} = 1/j \quad (1)$$

where the no. of data points in the i th cluster is represented by " c_i ".

5. Recalculate the distances between every data point and just now found the cluster center's once more.
6. Stop if no data point is reassigned, and then go back to step 3 if not.

If the original cluster center is chosen wrongly and the convergences of methods to local minima, the resulting cluster will be erroneous. Typically, the first cluster centers are chosen at random from the input images. The local connection between the data points must be considered at every stage of the clustering process. The mean intensity value for each tissue class is initialized using the histogram technique.

From Fig. 3 the pre and post-k-means clustering method operation are shown. After the cluster means recalculation coverages, image intensities are redistributed inside the cluster. It may be deduced that the tissue class labels on the output mask fall between one and k.

The quality of the final solution is determined by the initial position of the ideal centroid, which is reached by solving a challenging optimization problem. Even though randomization is the most widely used initialization, many method can be used to select the initial centroid based on the circumstances. Figure 4 shows the algorithm for the k-means clustering model. The comparative analysis between the proposed system and the existing system was presented in the next section.

Support Vector Machine (SVM)

The SVM classifiers is the concept of decision planes, which offer precise boundaries for decisions. According to Pavithra and Jayanthi [19], this is done by a decision plane

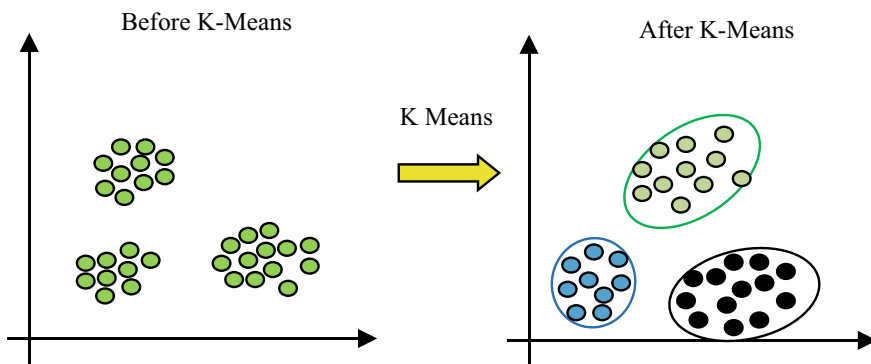


Fig. 3 K-means clustering

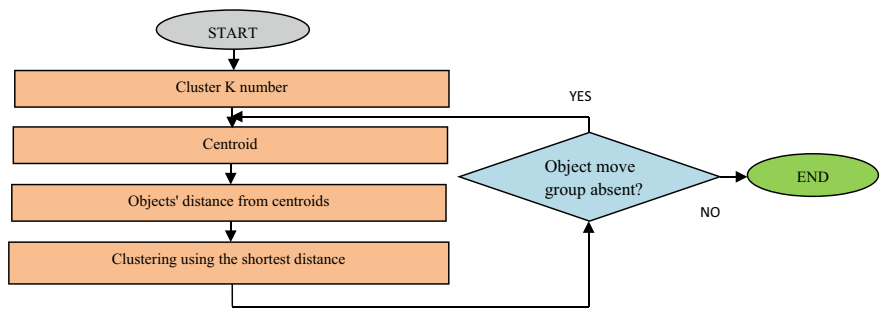
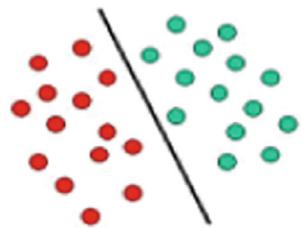


Fig. 4 Flowchart of K-means clustering algorithm

to divide a collection of objects with different class memberships. Figure 6 shows a linear classifier with a decision boundary line that distinguishes between different item sorts (individually, red and blue objects). A particular class of things is located in the right quadrant of the decision plane, while another class of objects is located on the left. Certain classification tasks are not straightforward and call for the use of more intricate structures in order to separate dada as effectively as feasible and classify the objects that are connected to the test cases that are being investigated (Fig. 5).

Since the greatest interval is desired, the computation of the sample space interval is important. The split of the hyperplane is described by the following linear equation in the sample space:

Fig. 5 SVM classifier with decision plane



```
Request: x and y loaded with test labelled data  $\alpha \leftarrow 0$  to  $\alpha \leftarrow$  in part test SVM
C  $\leftarrow$  some value
Again,
    For all  $\{x_i, y_i\}, \{x_j, y_j\}$  do
        Optimize  $\alpha_i$  and  $\alpha_j$ 
    Finish for
As long as  $\alpha$  or other factors related to resource constraints remain unchanged
Ensure: retain only the support vector ( $\alpha_i > 0$ )
```

Fig. 6 SVM classification Pseudo-Code

$$W^T x + b = 0 \quad (2)$$

The normal vector W in this instance determines the hyperplane's direction, while the displacement b establishes its distance from the origin. Let's say the training samples satisfy the following formula and the hyperplane is able to classify them appropriately:

$$W^T x_i + b \geq 1, y = 1 \quad (3)$$

$$W^T x_i + b \leq -1, y = -1 \quad (4)$$

The sample positive, as indicated by the fact that it is a negative sample. The values of 1 or -1 given here are actually just for the sake of calculation; you can use any constant in their place. Especially in apps that use text categorization. SVMs can potentially solve the locally extreme dilemma that plagues machine learning algorithms by solving a convex quadratic optimization problem and producing a globally optimal solution. SVM belongs to the category of nonparametric supervised techniques and is insensitive to the distribution of the underlying data. This is one of SVMs' advantages over other statistical techniques like machine learning, where data distribution must be understood beforehand. Figure 6 explains SVM classification Pseudo-Code.

Finding a single boundary between two groups is the task of a linear binary classifier, or SVM for short. It is possible to separate the multidimensional data the input space linearly, according to the linear SVM. It proposes a threshold for decision-making that lowers the number of false positives during the training phase. Many hyperplanes with no sample in between are chosen, and it indicates that the ideal hyperplane is identified when the margin of separation is highest. The "learning process" is the method of creating a classifier by iteration that has an ideal decision boundary.

Where,

- $K(x_i, x_j)$ is the kernel function that determines how similar two samples, x_i and x_j , are to one another. The samples are implicitly mapped into a feature space of greater dimensions; SVM may effectively address nonlinear classification issues.
- The term $\sum a_i$ represents the sum of all Lagrange multipliers.

Soft-margin

SVM can be expanded using the hinge loss function in scenarios when the data cannot be divided linearly.

$$\text{Max}(0, 1 - y_i(w^T x_i - b)) \quad (5)$$

Note that y_i is the i th target (i.e., in this case, 1 or -1), and $w^T x_i - b$ is the i th output. The SVM decision boundary can be described in terms of these multipliers and the

support vectors once the dual issue has been solved and the ideal Lagrange multipliers have been found. The decision boundary is provided as needed and training samples with $i > 0$ act as the support vectors.

4 Result and Discussion

This session presents the findings from the proposed system, hybrid algorithm for machine learning of K-Means with Support Vector Machine (K-Means+SVM). The recently constructed categorization approach is validated and introduced through resource data analysis of water access and hygiene data, with datasets supplied from the Kaggle website. The training raw dataset contains a large number of data features. The newly developed K-Means+SVM algorithm assessment measures, including accuracy, sensitivity recall, precision, and time duration are compared to those of the current algorithms, including decision tree(DT), logistic regression(LR), and KNN. MATLAB 2013A is used in the recently proposed k-means study to determine the efficacy of a classifier mean.

4.1 Performance Metrics

An illustration and summary of the effectiveness of a classification strategy may be found in Fig. 8, the confusion matrix. The performance analysis of the proposed K-Means integrated with Support Vector Machine approach of machine learning classification models can be assessed using performance evaluation metrics. Padilla et al. [20] used a dataset that your model of machine learning has never encountered before to see how well it performs in this way.

True Positive (TP): Effective detection of the bounding box on the ground.

False Positive (FP): Incorrectly detecting an object that is present or incorrectly detecting one that is not.

False Negative (FN): A bounding box for ground truth that is hidden.

True Negative (TN): Because there are numerous bounding boxes in every image that should not be recognized, a true negative (TN) outcome is irrelevant while discussing object detection.

4.2 Accuracy

The accuracy formula yields a value that indicates an error rate that is not 100%. The error rate needs to be examined before accuracy is determined. By dividing

Fig. 7 Accuracy graph

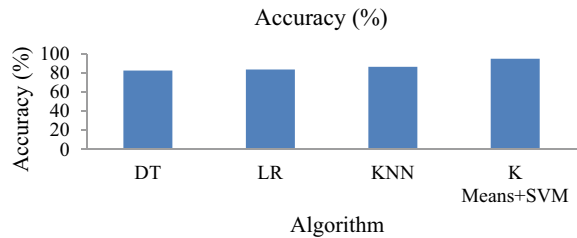


Fig. 8 Sensitivity graph

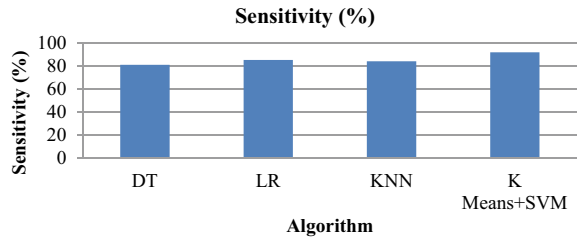


Table 1 Accuracy result

Algorithm	Accuracy (%)
DT	82.36
LR	83.46
KNN	86.35
K Means+SVM	94.70

the observed value by the actual value, one may also determine the error rate as a percentage.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}$$

Refer to Fig. 7 and Table 1 for the resulting in accuracy using the recommended and modern technology. By comparing the K-Means+SVM with other current systems, our recommended algorithm obtained 94.70% accuracy.

4.3 Sensitivity

The number of exact positive forecasts is split by the total number of positives. One can calculate sensitivity as

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Table 2 Sensitivity result

Algorithm	Sensitivity (%)
DT	81.17
LR	85.50
KNN	84.25
K Means+SVM	92.20

Table 2 yields the sensitivity results with the existing and suggested systems, which are shown in the previously described Fig. 8. It shows that the K-Means+SVM produces a higher sensitivity result of 92.20%, which is better than the comparative current methods.

4.4 Recall

Recall provides data about the model’s capacity to identify true positives. As there are no positives, this example demonstrates how few patients with the ailment were actually detected. There are two false negatives accessible.

$$\text{Recall} = (\text{TP})/(\text{TP} + \text{FN})$$

The outcome of the recall utilizing the suggested and existing systems are shown in Fig. 9, which was previously addressed and is obtained from Table 3. According to the plotted graph, our newly constructed K-Means+SVM system offers a recall of 90.45% higher than the previously described existing systems.

Fig. 9 Recall graph

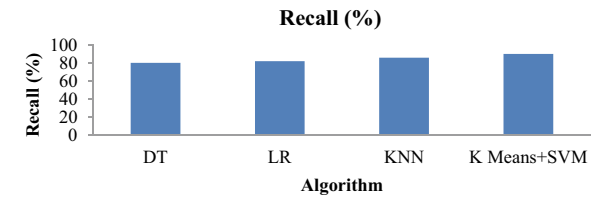


Table 3 Recall result

Algorithm	Recall (%)
DT	80.55
LR	82.30
KNN	86.22
K Means+SVM	90.45

Fig. 10 Precision graph

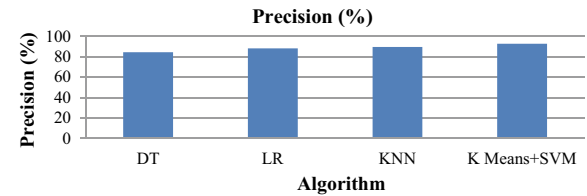


Table 4 Precision result

Algorithm	Precision (%)
DT	84.42
LR	88.16
KNN	89.52
K Means+SVM	92.75

4.5 Precision

The accuracy level indicates how well many measures agree with one another. You can evaluate precision by calculating the average deviation, which sums together the measurement errors.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}$$

Figure 10 is drawn using Table 4, which shows how precise the suggested and available technologies are. Plotting the outcome shows that our newly designed K-Means+SVM system outperforms the previously explained current approaches with a precision of 92.75%.

4.6 Time Duration

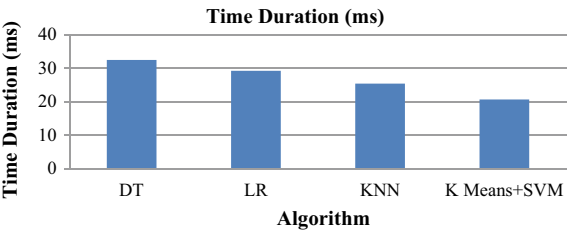
Selecting the best using a model of machine learning to address an issue can be time-consuming if done carelessly.

The time duration output graph from Table 5 for current and suggested system is shown in Fig. 11. The new method of our proposed K-Means+SVM algorithm works and consumes the less time duration of 21.84 ms in contrast with the other existing algorithms DT, LR, and KNN.

Table 5 Time Duration Result

Algorithm	Time Duration (ms)
DT	32.45
LR	29.18
KNN	25.37
K Means+SVM	20.63

Fig. 11 Time duration graph



5 Conclusion

This paper has presented a survey on analysis and forecasting water availability and hygiene on resource information using Machine Learning approach. There is a good amount of literature on data mining using traditional machine learning techniques and, more recently, optimized machine learning models because of the big data era, the proliferation of IoT, the renewed interest in water access, quality, and security, and new ML techniques to enhance the output and performance of current algorithms result prediction among other factors. We are taking the existing algorithms, SVM, Logistic Regression, and KNN, for comparison. This existing algorithms’ performances are compared with our newly constructed algorithm of K-Means+SVM. Depending upon this comparison, the suggested system achieves the best accuracy of 94.70% with a less time duration of 20.63 ms.

References

1. Ritchie H, Roser M (2023) Water use and stress. Our World in Data
2. Zar HJ, Dawa J, Fischer GB, Castro-Rodriguez JA (2020) Challenges of COVID-19 in children in low-and middle-income countries. *Paediatr Respir Rev* 35:70–74
3. Tortajada C (2020) Contributions of recycled wastewater to clean water and sanitation sustainable development goals. *NPJ Clean Water* 3(1):22
4. Pearson AL, Mack EA, Ross A, Marcantonio R, Zimmer A, Bunting EL, HWISE Research Coordination Network (2021) Interpersonal conflict over water is associated with household demographics, domains of water insecurity, and regional conflict: evidence from nine sites across eight sub-Saharan African countries. *Water* 13(9):1150
5. Roque A, Wutich A, Quimby B, Porter S, Zheng M, Hossain MJ, Brewis A (2022) Participatory approaches in water research: a review. *Wiley Interdiscip Rev: Water* 9(2):e1577
6. Schmidt JJ (2019) *Water: abundance, scarcity, and security in the age of humanity*. NYU Press

7. Shin S, Lee S, Judi DR, Parvania M, Goharian E, McPherson T, Burian SJ (2018) A systematic review of quantitative resilience measures for water infrastructure systems. *Water* 10(2):164
8. Kedron P, Li W, Fotheringham S, Goodchild M (2021) Reproducibility and replicability: opportunities and challenges for geospatial research. *Int J Geogr Inf Sci* 35(3):427–445
9. Meehan K, Jepson W, Harris LM, Wutich A, Beresford M, Fencel A, Young S (2020) Exposing the myths of household water insecurity in the global north: a critical review. *Wiley Interdiscip Rev Water* 7(6):e1486
10. Moswatsi T (2023) The impact of world vision WASH program on rural public health in Thabana-Morena Mafeteng, Lesotho. Master's thesis, National University of Lesotho
11. Han S, Song D, Xu L, Ye Y, Yan S, Shi F, Du H (2022) Behaviour in public open spaces: a systematic review of studies with quantitative research methods. *Build Environ* 223:109444
12. Rios Quitizaca P, Gatica-Domínguez G, Nambiar D, Ferreira Santos JL, Brück S, Vidaletti Ruas L, Barros AJ (2021) National and subnational coverage and inequalities in reproductive, maternal, newborn, child, and sanitary health interventions in Ecuador: a comparative study between 1994 and 2012. *Int J Equity Health* 20:1–14
13. Atilola GO (2021) Bayesian modelling and mapping of health outcomes in space and time using complex national surveys. University of Northumbria at Newcastle (United Kingdom)
14. Wu T, Luo J, Dong W, Sun Y, Xia L, Zhang X (2019) Geo-object-based soil organic matter mapping using machine learning algorithms with multi-source geo-spatial data. *IEEE J Sel Top Appl Earth Observations Remote Sens* 12(4):1091–1106
15. Talebi H, Peeters LJ, Otto A, Tolosana-Delgado R (2022) A truly spatial random forests algorithm for geoscience data analysis and modelling. *Math Geosci* 54(1):1–22
16. Mertler CA, Vannatta RA, LaVenita KN (2021) Advanced and multivariate statistical methods: practical application and interpretation. Routledge
17. Leliveld A, Knorringa P (2018) Frugal innovation and development research. *Eur J Dev Res* 30:1–16
18. Sinaga KP, Yang MS (2020) Unsupervised K-means clustering algorithm. *IEEE Access* 8:80716–80727
19. Pavithra D, Jayanthi AN (2018) A study on machine learning algorithm in medical diagnosis. *Int J Adv Res Comput Sci* 9(4)
20. Padilla R, Netto SL, Da Silva EA (2020) A survey on performance metrics for object-detection algorithms. In: 2020 International conference on systems, signals and image processing (IWSSIP). IEEE, pp 237–242

SDN-Enabled Secure Data Communication in 5G VANETS Using ECC Algorithm



C. Balakumar and U. Priyatharshini

Abstract A subtype of mobile ad hoc networks called vehicular ad hoc networks, or VANETs, is now one of the most comforting, motivating, and rapidly evolving (MANETs). The erratic wireless media that they transmit through includes smart cars, on-board units (OBUs), and SDN. Because of this, VANETs have different framework issues and security issues when it comes to gaining the executive's framework, all of which have an impact on the confidence in cooperation and strategy across practical frameworks. The sub-module known as the software-defined network (SDN) is part of a new VANET structure that is provided for secure communication. VANET requirements in the future are met by the scalability, programmability compliance, and global network information that the SDN offers. This study presents the scheduling protocols of the security blockchain-based architecture as well as the SDN-enabled 5G VANET model directly. Packet delivery ratio, packet loss ratio, overhead, and efficiency calculation in 5G VANET for SDN-based secured data-enabled connection are all dependent on the number of cars during the performance evaluation. Comparing the proposed SDN-based ECC system to the current efforts, like S-AODV and ES-AODV, reveals that it performs well and provides greater efficiency.

Keywords Data security · Data communication · 5G-VANET · Software defined network · Elliptic curve cryptography

C. Balakumar (✉)

Department of Computer Applications, Faculty of Arts, Science, Commerce and Management,
Karpagam Academy of Higher Education, Coimbatore, India
e-mail: balakumar.cbk@gmail.com

U. Priyatharshini

Department of Computer Applications, Karpagam Academy of Higher Education, Coimbatore,
India

1 Introduction

SDN offers control and support for the dynamic nature of VANETs through its flexible centralized design and programmable control interface. Along with satisfying a variety of QoS needs, SDN lowers operating costs by using streamlined hardware, software, management, and communication protocols. To obtain a worldwide perspective of VANET, it makes use of a centralized controller. Fogli et al. [1] offers services including shortest path calculations, resource allocation based on service priority, network activity detection, etc. Since smart devices are used by everyone and wireless access technologies (including Wi-Fi, DSRC, 4G/LTE, and 5G) are developing at a rapid pace, VANETs are now a practical tool for enhancing traffic safety and transit efficiency.

Software-defined network (SDN) communication with 5G VANET technology is shown in Fig. 1. SDVN designs are a potential technology that has evolved in recent years by utilizing network programmability to facilitate innovation and streamline network administration, as mentioned by Barakabitze et al. [2]. As a result, both academia and business give it a lot of attention. It is possible to separate the control and data planes in SDVNs thanks to SDN technology, which offers two benefits: (i) an abstraction of the underlying networking architecture for VANET applications, and (ii) an intelligent and logically centralized networking state.

Many of the present problems facing VANET can be solved by merging SDN and VANET, which is regarded as a significant direction. VANET apps can improve user experience by utilizing SDN's notable characteristics, which offer the necessary support. Additionally, SDN features can satisfy the sophisticated requirements of VANETs. Three key features of SDN are as follows: (i) network virtualization

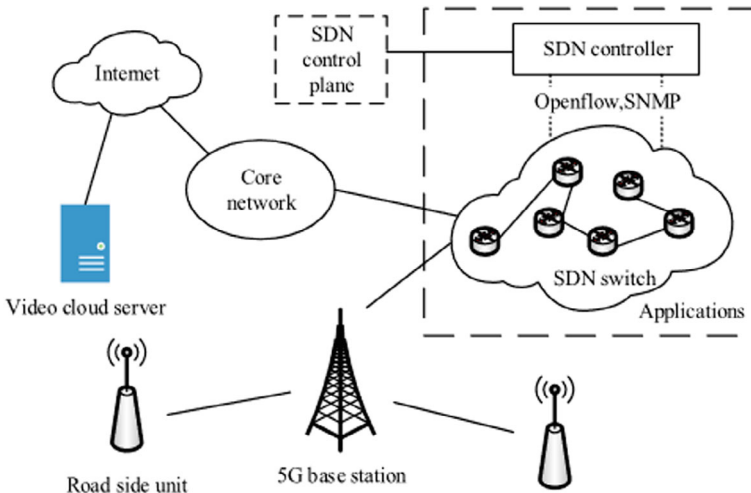


Fig. 1 SDN Communication in 5G VANET

supports heterogeneous applications; (ii) flexibility through dynamic programmability on networking parts; and (iii) centralized global network knowledge facilitates effective management of diverse services. Security and effectiveness in the implementation of various VANET services could be ensured by these SDN features. In addition, Jaballah et al. [3] discusses cutting-edge blockchain-based systems and their applications on top of SDVNs.

- Our analysis focuses on the advantages and difficulties of modern SDVN systems, particularly with relation to data transmission processes' security and efficiency. We think that the data from our survey will offer the necessary understanding to support the potential creation of a more reliable and secure SDVN design. It is the first thorough paper that gives such a survey and analysis on SDVNs, as far as we can tell. We further offer a general SDVN architecture design based on our investigation.
- With regard to the current SDVN architectures, we provide a thorough security study that takes into account a variety of security risks and potential defense mechanisms.

1.1 Trust and Security Issues in the Vehicle Internet of Things

Security for IoT devices in cars has grown in importance and received more focus. Thanks to advancements in sensing, computing, and communication on board, vehicles are becoming more autonomous. Vehicular networks can thus submit information to the information operation or management center and exchange a significant number of messages amongst cars. Disaster data and message transmission is enabled for new kinds of IoT services. When combined with awareness of traffic conditions, these messages can enhance management effectiveness and significantly contribute to the transportation system's safety guarantee. Yet, because they are typically strangers to one another, cars in a VANET are typically reluctant to completely collaborate and trust one another. It is inevitable for malevolent vehicles or misbehaviors to occur in large-scale scenarios due to the tremendous mobility and variability of VANET. By advertising a clean route that is actually congested or has traffic accidents, rogue vehicular nodes are spreading fantastic messages that pose a serious threat to the transportation system. In order to ensure that service-related messages originate from safe and reliable infrastructures and automobiles, VANET must ensure a secure IoT environment. These communications have to be unwavering, believable, and real. Protecting user privacy during the acquisition and delivery of messages is also crucial because no user can stand to have their personal information leaked. Furthermore, the direction of automotive IoT services will require spatiotemporal sensitivity, collaborative intelligence, and centerless trust, Lu et al. [4].

1.2 Blockchains in Vehicular Networks

Cyberattacks and privacy breaches are perennial threats to traditional centralized network architecture. Yet, blockchain has been hailed as a breakthrough method for data security and privacy because of its decentralized characteristics. A distributed database that keeps an expanding list of linked blocks, and it is seen as a disruptive breakthrough in the financial sector. Every node is identifiable with a public key (PK) on a distributed P2P network that oversees it. Network-wide broadcasts of encrypted transactions, or communications between nodes, are made possible via public key cryptography (PKs). Assuring the trustless consensus of the blockchain, any node can validate a transaction by comparing the transaction generator's signature to their PK, as stated by Javed et al. [5].

1.3 Software-Defined Networking (SDN)

The division of the control and data planes is a fundamental component of SDN. Latif et al. [6] state that the SDN uses standardized interfaces to provide communication across devices made by different vendors. This makes it easier to administer and monitor networks, and it facilitates the creation of flexible and programmable networking architecture.

2 Literature Review

According to Jaiswal et al. [7], VANET architecture is a self-configuring network that evolved as a result of recent advancements in network technologies. The OBU, the RSU, and the AU are the three different sorts of pieces that make up the VANET architecture. Sensors and gadgets inside the car can communicate with one another through a process known as intra- or in-vehicle communication. Hbaieb et al. [8] state that various electronic control units (ECUs), sensors, and actuators are installed in cars for the purpose of intra- or in-vehicle communication. Various protocols are used for intra-vehicle communication, including Ethernet, power line communications, media-oriented system transport, CAN, and LIN. A car speaks with another car in a one-hop communication called vehicle-to-vehicle (V2V).

In order to handle applications like video streaming and advertisement distribution, a vehicle connects with an RSU [9]. The position, direction, and speed of vehicles are transmitted in particular signals to the RSU infrastructure. The required services will be rendered by the RSU following the collection of this data. The vehicle-to-X (V2X) communications group, which encompasses vehicle-to-pedestrian, vehicle-to-V2I, and other forms of communication, is defined by the

Third Generation Partnership Project (3GPP) group. Many uses, including infotainment services, road safety, and vehicle traffic optimization, are made possible by V2X communication. DSRC and cellular network technology are the two technologies that enable V2X communications. Here is a quick rundown of these two cutting-edge technologies as examined by Imbruglia et al. [10].

In order to facilitate V2X applications, DSRC and cellular technologies are utilized. DSRC is a wireless technology that allows messages to be exchanged between OBUs and RSUs over a short distance. The Radio Spectrum Reserves (DSRC) was designed by Bey and Tewolde [11] and is dependent on several locations, including North America, Europe, and Japan. In Sultana et al. [12], authors talk about security attacks on SDVN state-of-the-art architectures, including distributed DoS, application attacks, privacy infringement, and network topology poisoning. SDN can be deployed in VANETs to counter a few security threats, according to Al-Heety et al. [13]. Risks carried over from SDN to VANETs have been thoroughly examined and summarized.

2.1 Trust Management System Based on Blockchain

Road condition collecting and post-investigation is facilitated by centralized trust management, which also affords user privacy protection. However, it will jeopardize road safety if bad cars purposefully disseminate false information. Using data from vehicle interactions, distributed trust management is able to assess the veracity of many sources, including road conditions. Unfortunately, the validity of a vehicle's identity is not acknowledged, and there is no uniform verification for vehicles joining the VANET. However, when implemented on VANET, NC has demonstrated a great deal of promise for enhancing security and robustness. In order to improve data transmission security and reliability in vehicle networks, this paper proposes the construction of an architecture that integrates multigeneration-mixing (MGM) functions with network connectivity (NC) to exercise the SDN idea. An MGM-based NC protocol is designed by Bhatia et al. [14] for data encoding and decoding. Moreover, vehicle authentication is managed by a centralized SDN controller. Chaudhary et al. [15] wrote this paper, for an IIoT environment, a multi-attribute secure communication architecture enabled by software-defined networks (SDNs) is created. This suggested plan operates in three stages: (1) A cuckoo-filter-based fast-forwarding strategy is used to construct an SDN-IIoT communication model, (2) an attribute-based encryption technique is given for safe data transmission, and (3) a peer entity authentication scheme employing Kerberos, a third-party authenticator, is also shown. This technique has been tested under various conditions, and the findings demonstrate its efficacy when compared to other options.

3 System Design

In this section discussed about the overall system design and its working process of the proposed software-defined network (SDN)-enabled secure data communication in 5G VANETS using elliptic curve cryptography (ECC) algorithm. 5G base stations (gNBs), RSUs, and cars with OBUs are examples of the heterogeneous nodes included in this design. The video cloud server houses the video reports, and the Internet is connected to the core network that comprises the Department of Motor Vehicles (DMV) and Trusted Authority (TA). Vehicle-to-vehicle (V2V) communication is also taken into when broadband wireless Internet connectivity is required; this design takes into account enabling data exchange across cars in the VANET.

Figure 2 depicts and explains about the proposed system architecture. It consists of SDN, static and mobile devices and trusted authority, and also vehicular network. These all output fed into the remote cloud. Control plane and data plane decoupling is done in order to accomplish global administration and uniform policies for the 5G VANET. Functions like vehicle clustering and traffic management can be provided by implementing vehicle apps at the SDN-based ECC controller. Global functions can be realized by programmable applications, which offer a global perspective over the entire service area. Proactively, following a predetermined time frame, or reactively, upon request from the infrastructure, the total network policies can be updated.

Best Road Segment Section: The remote cloud (RC), which encompasses chromosome and basic activities, regulates the distribution of RSUs in our study, as previously indicated. Centralized criteria are the main factors taken into account when placing RSUs; most of the region must be covered and there should not be any problems with RSU situations.

Traffic Information Collection: As part of the message sending procedure, cars rate the accuracy of road condition tags. The only ratings for scores are $+1$ and -1 . As an illustration, let us say that car V_n gets a message that V_m broadcasts. The traffic tag tags m, j , and the position of place m in the message are obtained by V_n after verifying the legal identity of V_m . Next, $P_{cert\ n, k}$, V_n sign its location and the mark it created for tags m, j , together with its public key certificate, and upload these to the next RSU. The messages are split in the same manner starting at period j .

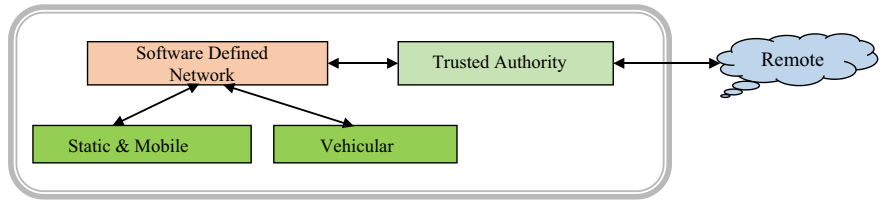


Fig. 2 Proposed network architecture

Elliptic Curve Cryptography (ECC) Algorithm: We propose to conduct research using the elliptic curve cryptography (ECC) algorithm. Natanael and Suryani outline the fundamental traits and attributes of ECC in this instance. The main cryptosystems of ECC [16] employ an elliptic curve with a symmetric foundation. The generic ECC algorithm's elliptic key's mathematical expression is provided here.

$$y^2 = x^3 + ax + b \pmod{p} \text{ with } a \text{ and } b \text{ constants}$$

3.1 ECC-Based Trusted Authority (TA)

The trustworthy authority (TA) serves as the primary source for the authentication procedure in this study, which secures the network using an ECC and a public–private key pair. A pair of cryptographic keys is used by TA to manage the SVs and oversee the process of verifying that no complaints of misbehavior are present in the SVs that do not use the SVs' trust values. It is known as the first registration in a network when the TA subsequently generates a certificate that maintains key pairs for each SV. The SVs' valid certificates given by the TA are checked by the RSU to ensure effective exploitation of this process. To protect the network's privacy, RSU has tight security. ECC key generation, as described by Rangwani and Om [17], lowers complexity by employing data with decreasing key size with excellent security. As a result, it is applicable to dynamic networks with incredible speed and mobility. An ECC customized with a top secret key which is displayed in Fig. 3 allows for secure vehicle communication. During communication between the SVs, the public–private key pairs that are produced are utilized. Every time a transmission occurs, the messages broadcast their authenticated private key, position, speed, and direction. SV's private key is used to encrypt the transmitted message. The public key is then used by SV for verification during decryption. For ECC, the typical elective curve has the following mathematical expression: $y^2 = x^3 + ax + b \pmod{p}$.

ECC $4a^3 + 27b^2 \neq 0$ is guaranteed by the integer values a and b in the equation above. SVs' public–private key pairs and the secret shared key between any two SVs are produced by this method. In the first instance, for instance, these SVs must ensure the message's authenticity if SV1 wishes to speak with SV2. Secret key D1 and private key (Prs) are examples of the private keys selected by SV1. SV2 selects private keys, such as private key (Prs) and secret key d2, in a similar fashion. Next, as stated below, are the private keys for the SV1 and SV2. Pseudocode for ECC algorithm is shown in Fig. 4.

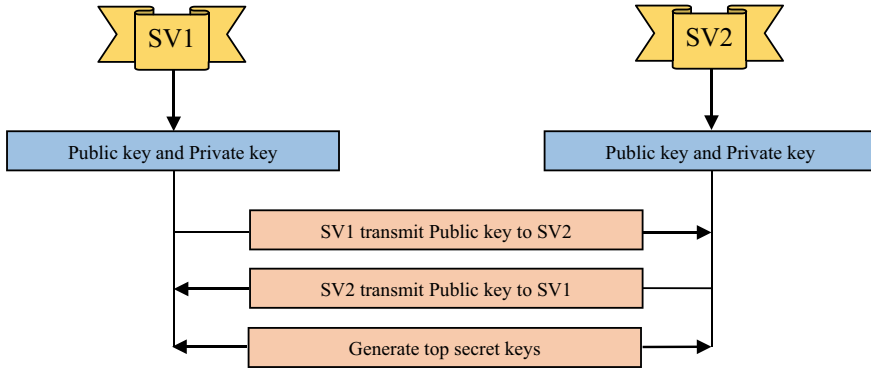


Fig. 3 Vehicle-to-vehicle communication using a secret key based on ECC

Fig. 4 Pseudocode of the proposed ECC algorithm

Elliptic Curve Cryptography (ECC):
 Start the Suggested ECC Approach:
 • Optimal road section choice
 • $P1 = RSUOD = \sum RoSr * \mu r * Br$
 • $P2 = Fit \text{ function} = XCp + y/Ln$
 • $P3 = I \text{ cost} = \sum Nr * Ir$
 $Mf = \alpha P1 + \beta P2 + \gamma P3$ with $\alpha + \beta + \gamma = 1$
 Trusted Authority Authentication with ECC-based Method
 • ECC Expression $y^2 = x^3 + ax + b \pmod{p}$
 • Private Key Creation $SV1 = \delta 1$. Prs and $SV2 = \delta 2$. Prs
 • Secret Key Creation $TS1 = t1 * ts1$ and $TS2 = t2 * ts2$
 • Hash Key Creation $M = Ps + (TS1) + (TS2)$
 • Top secret key of $SV1 = K1 = \delta 1 * \delta 2 * (TS1)$
 • Top secret key of $SV2 = K2 = \delta 2 * \delta 1 * (TS2)$
 Stop

4 Result and Discussion

For the purpose of putting the research's concept of creating real-time traffic congestion in a 5G VANET into practice, a coverage region with varied speeds between 25 and 40 (m/s) was chosen for the simulation scenario. The following Table 1 provides the additional crucial variables that are involved in the simulation process. Three distinct scenarios are used in the analysis: there are different numbers of vehicles, varied speeds, and there are hostile vehicles in the network. In this analysis, the following characteristics are taken into account: overhead, efficiency, packet delivery ratio, and packet loss ratio. S-AODV and ES-AODV, in contrast, are the older studies that were taken into consideration for the performance analysis. It also includes encryption and decryption based on security measures. To achieve an effective performance, the SDN-based secured data transmission protocol and the ECC algorithm are merged in the proposed protocol. This method is applied to the secured

Table 1 Simulation parameters

Parameters	Values
Time of simulation	200 ms
No. of vehicles	500
Size of data packet	512 bytes
Type of traffic agent	Transmission control protocol
Type of traffic application	Constant bit rate

optimal path selection. The simulation yielded overhead, packet delivery ratio, and packet loss ratio. The outcomes are mediocre and should be better.

4.1 Packet Delivery Ratio

Any routing strategy in a network can be evaluated for effectiveness using the packet delivery ratio statistic. The ratio is calculated by adding the volume of data packets delivered by the source and the amount received by the destinations. By multiplying the total number of packets delivered throughout the simulation, we can determine the mean delivery ratio by the number of target nodes receiving the safety message inside the target region. The message delivery rate must be close to 100% in order for a dissemination approach to be considered dependable.

Packet Delivery Ratio =

$$\frac{\text{All destination total packets received}}{\text{Total no. of packets sent by all source nodes}}$$

Figure 5 compares packet delivery ratio findings (in percentages) from the above Table 2; S-AODV and ES-AODV, the two currently used techniques, produce values of 64% and 83%, respectively. The proposed SDN-ECC Protocol’s destinations received packets from 90% of the destinations. This demonstrates that the packet delivery ratio of the suggested SDN-ECC is higher than that of the current techniques.

Fig. 5 Packet delivery ratio result graph

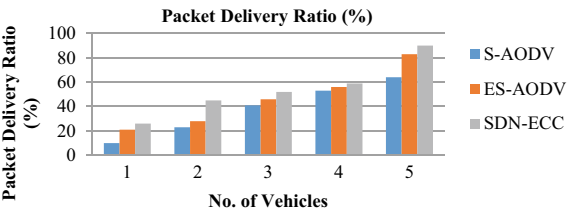


Table 2 Packet delivery ratio result

Packet delivery ratio (%)		
Existing protocols		Proposed protocol
S-AODV	ES-AODV	SDN-ECC
10	21	26
23	28	45
41	46	52
53	56	59
64	83	90

Table 3 Packet drop ratio result

Packet drop ratio (%)		
Existing protocols		Proposed protocol
S-AODV	ES-AODV	SDN-ECC
8	5	2
11	8	5
18	12	9
22	19	14
27	23	17

4.2 Packet Drop Ratio

The packet drop ratio is the percentage of packets that were transmitted from a source but never arrived at their destination.

$$\text{Packet Drop Ratio} = \frac{n \text{ send packets} - n \text{ recieved packets}}{n \text{ send packets}} * 100$$

Table 3 compares the performance of the existing protocols, such as S-AODV and ES-AODV with proposed SDN-ECC.

Figure 6 compares the results for packet drop ratio (in percentage) and shows that the proposed solution has a lower PDR or loss of pocket than the existing approach. In the proposed SDN-ECC protocol, 17% of packets never reach their intended destination. The existing approaches S-AODV and ES-AODV give 27% and 23% packet drop ratio.

4.3 Overhead

The extra data needed to transmit the payload the information included in a packet is known as overhead. Stated differently, the overhead % can be calculated as follows:

Fig. 6 Packet drop ratio result graph

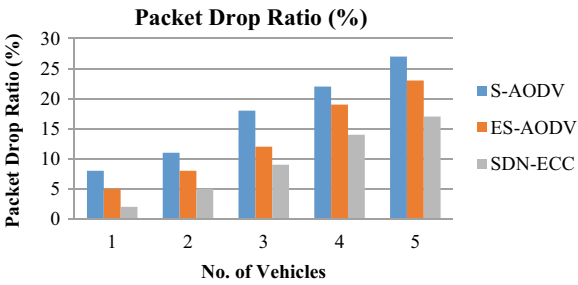
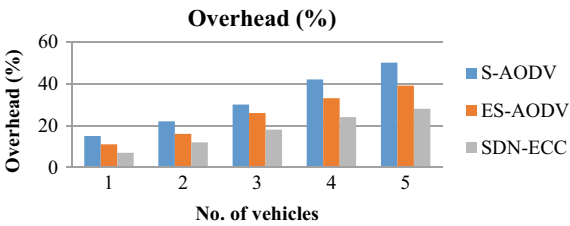


Table 4 Overhead calculation result

Overhead (%)		
Existing protocols		Proposed protocol
S-AODV	ES-AODV	SDN-ECC
15	11	7
22	16	12
30	26	18
42	33	24
50	39	28

Fig. 7 Overhead result graph



header multiplied by 100 divided by the whole payload. Using Table 4, the graphical representation of overhead calculation is displayed in Fig. 7. This outcome demonstrated that, in comparison with the previous S-AODV and ES-AODV studies, the suggested SDN-ECC generates less overhead.

The existing approaches S-AODV and ES-AODV give the overhead percentage of 50% and 39%, respectively. And our proposed SDN-ECC is reduced to 28% only, which is better than the existing models.

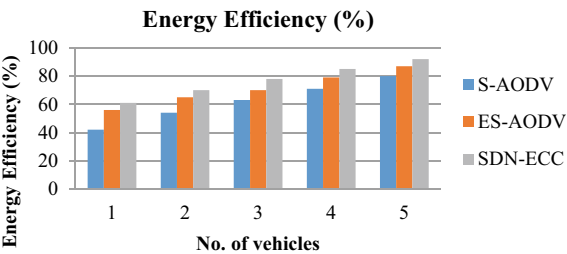
4.4 Energy Efficiency

Reducing energy use to achieve the same objective or output is known as energy efficiency. Efficiency is a percentage that can be computed using the following

Table 5 Energy efficiency result

Energy Efficiency (%)		
Existing Protocols		Proposed Protocol
S-AODV	ES-AODV	SDN-ECC
42	56	61
54	65	70
63	70	78
71	79	85
80	87	92

Fig. 8 Energy efficiency graph



formula:

$$\text{Energy Efficiency } \eta = (P_{\text{out}}/P_{\text{in}}) \times 100\%$$

where P_{out} and P_{in} are both output and input power, respectively, and are expressed in kW. Because of energy losses brought on by things like heat and friction, the output power in reality is always less than the input power.

The energy efficiency of techniques like S-AODV, ESAODV, and the suggested SDN-ECC approaches is represented in the Table 5 and also drawn graphically in Fig. 8, that is 80%, 87%, and 92%, respectively. The suggested SDN-ECC performs better than the previous efforts in terms of energy efficiency, according to the results. The SDN-based ECC method contributes to a decrease in network energy consumption, which raises energy efficiency.

5 Conclusion

In 5G VANETs, the main downsides of high-speed are taken into account, including security and communication overhead. Elliptic curve cryptography (ECC)-based authentication is used to create privacy-aware safe routing in order to address these shortcomings. Additionally, the 5G VANETs' security is enhanced by ECC. NS2 is utilized in the simulation, and an analysis is carried out to assess the number of vehicles' performance in terms of overhead, efficiency, packet delivery ratio, and

packet drop ratio. The suggested SDN-based ECC's performance is computed and compared to that of previous studies, including S-AODV and ES-AODV. The overall efficiency of the suggested method can reach up to 92%, which is higher than the previous efforts, depending on the number of cars. The suggested SDN-enabled secure data transfer in the 5G VANET employing ECC has a relatively small number of vehicles as overhead. When compared to previous works, including S-AODV and ES-AODV, the proposed SDN-based ECC system performs exceptionally well. Future study will focus on hybrid security data communication, ensuring that the 5G VANET network's performance is not impacted by an increase in the number of cars.

References

1. Fogli M, Giannelli C, Stefanelli C (2022) Software-Defined Networking in wireless ad hoc scenarios: objectives and control architectures. *J Netw Comput Appl* 203:103387
2. Barakabitze AA, Ahmad A, Mijumbi R, Hines A (2020) 5G network slicing using SDN and NFV: a survey of taxonomy, architectures and future challenges. *Comput Netw* 167:106984
3. Jaballah WB, Conti M, Lal C (2019) A survey on software-defined VANETs: benefits, challenges, and future directions. *arXiv preprint arXiv:1904.04577*
4. Lu Z, Qu G, Liu Z (2018) A survey on recent advances in vehicular network security, trust, and privacy. *IEEE Trans Intell Transp Syst* 20(2):760–776
5. Javed MU, Rehman M, Javaid N, Aldegheishem A, Alrajeh N, Tahir M (2020) Blockchain-based secure data storage for distributed vehicular networks. *Appl Sci* 10(6):2011
6. Latif Z, Sharif K, Li F, Karim MM, Biswas S, Wang Y (2020) A comprehensive survey of interface protocols for software defined networks. *J Netw Comput Appl* 156:102563
7. Jaiswal KV, Sharma SK, Verma RK, Singh BK (2023) Vehicular ad-hoc network & exploration plugs. *Int J Intell Syst Appl Eng* 11(9s):758–766
8. Hbaieb A, Rhaïem OB, Chaari L (2018) In-car gateway architecture for intra and inter-vehicular networks. In: 2018 14th International wireless communications & mobile computing conference (IWCMC). IEEE, pp 1489–1494
9. Singh PK, Nandi SK, Nandi S (2019) A tutorial survey on vehicular communication state of the art, and future research directions. *Veh Commun* 18:100164
10. Imbruglia A, Cancila D, Settembre M (2023) 5G communication and security in connected vehicles. *ACM SIGAda Ada Lett* 42(2):109–113
11. Bey T, Tewolde G (2019) Evaluation of DSRC and LTE for V2X. In: 2019 IEEE 9th annual computing and communication workshop and conference (CCWC). IEEE, pp 1032–1035
12. Sultana R, Grover J, Tripathi M (2021) Security of SDN-based vehicular ad hoc networks: state-of-the-art and challenges. *Veh Commun* 27:100284
13. Al-Heety OS, Zakaria Z, Ismail M, Shakir MM, Alani S, Alsariera H (2020) A comprehensive survey: benefits, services, recent works, challenges, security, and use cases for SDN-VANET. *IEEE Access* 8:91028–91047
14. Bhatia J, Kakadia P, Bhavsar M, Tanwar S (2019) SDN-enabled network coding-based secure data dissemination in VANET environment. *IEEE Internet Things J* 7(7):6078–6087
15. Chaudhary R, Aujla GS, Garg S, Kumar N, Rodrigues JJ (2018) SDN-enabled multi-attribute-based secure communication for smart grid in IIoT environment. *IEEE Trans Ind Inf* 14(6):2629–2640
16. Natanael D, Suryani D (2018) Text encryption in android chat applications using elliptical curve cryptography (ECC). *Procedia Comput Sci* 135:283–291

17. Rangwani D, Om H (2021) A secure user authentication protocol based on ECC for cloud computing environment. *Arab J Sci Eng* 46:3865–3888
18. Pan J, Cui J, Wei L, Xu Y, Zhong H (2019) Secure data sharing scheme for VANETs based on edge computing. *EURASIP J Wirel Commun Netw* 2019(1):1–11
19. Hussein NH, Yaw CT, Koh SP, Tiong SK, Chong KH (2022) A comprehensive survey on vehicular networking: communications, applications, challenges, and upcoming research directions. *IEEE Access* 10:86127–86180
20. Varma IM, Kumar N (2023) A comprehensive survey on SDN and blockchain-based secure vehicular networks. *Veh Commun* 100663
21. Dhifallah W, Moulahi T, Tarhouni M, Zidi S (2023) Intellig_block: enhancing IoT security with blockchain-based adversarial machine learning protection. *Int J Adv Technol Eng Explor* 10(106):1167
22. Alotaibi B (2019) Utilizing blockchain to overcome cyber security concerns in the internet of things: a review. *IEEE Sens J* 19(23):10953–10971
23. Athanere S, Thakur R (2022) Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing. *J King Saud Univ-Comput Inf Sci* 34(4):1523–1534

Trust and Reinforcement Learning-Based Security Model (TRLs) for Mitigating Black Hole Attacks in Vehicular Ad Hoc Networks



C. Balakumar and G. Surya

Abstract Systems known as network intrusion detection systems, or NIDS, keep an eye out for any unusual activity that can jeopardize the confidentiality, availability, or integrity of the network's resources. It is critical to recognize the vulnerabilities in NIDSs and make sure they fulfill their essential functions by taking an antagonistic stance and spotting weak points and security holes. As a result, strong defenses against any prospective attacks may be created. The proposed methodology ensures constant interconnectivity among the dynamic vehicle nodes by integrating block chain concepts with a special time stamping mechanism. Furthermore, the architecture verifies the accuracy of the events recorded in the RSU by utilizing block chain concepts. It also proposes a flexible hybrid trust model that uses the latest direct and proposed trust evaluation approaches to compute trust between the vehicular entities. The system also includes a threading mechanism to schedule message execution in direct trust evaluation and clustering algorithms relevant messages in indirect trust assessment. To evaluate the efficacy suggested, numerous experiments are carried out in a simulated setting, including comparisons with other pertinent research. The findings show that the proposed trust model has a 92% rate in identifying malicious nodes. The integration of artificial intelligence (AI) and vehicle-to-everything (V2X) technology enhances driving efficiency, comfort, and safety by extending the driver's field of view, gathering data from multiple sources, and predicting potential collisions. This offers a summary initiatives that use artificial intelligence to address issues in V2X systems.

Keywords Reinforcement learning · Ad-hoc network · Road-side units · Blackhole attack · Vehicle-to-vehicle · Artificial intelligence

C. Balakumar (✉)

Department of Computer Applications, Faculty of Arts, Science, Commerce and Management,
Karpagam Academy of Higher Education, Coimbatore, India
e-mail: balakumar.cbk@gmail.com

G. Surya

Department of Computer Applications, Karpagam Academy of Higher Education, Coimbatore,
India

1 Introduction

Vehicle communication providing multitude of safety and informative benefits. Numerous automakers are equipping their vehicles with varying degrees of technology for support, safety, entertainment, and mobile commerce. An automobile with intelligence provides entertainment, Internet access, and other functions. Numerous onboard devices can enable new applications such as location-based services, online gaming, and mobile entertainment in various forms. However, it has been observed that third-party applications are not making the most of the cars' onboard computers. VANETs are not commonly employed in current applications because of their low processing and communication capabilities. Vehicles must be able to interface with people, infrastructure, and the Internet in order to support a wide range of existing VANET applications. Wireless and mobile networks are used to link vehicle sensors to the IoV. To this end, initiatives have been made in the last several years to enhance traffic management and driving safety introduced by He et al. [1]. RSUs and cars will be equipped with processing and communication units inside V2V, V2I does not require any extra infrastructure, and it is also referred to as "zero infrastructure". Conversely, roadside equipment, either mobile or stationary wealth, features to help transportation system management and drivers avoid obstacles like snow, ice, and even an ambulance on the way.

1.1 Internet of Vehicle (IOV)

The rapid expansion of IoV leads to a number of security problems, some of which are seen to pose a serious risk to the well-being of both consumers and enterprises. The risks faced by autonomous cars are diverse, and this article examines potential countermeasures as well as the potential impact of different approaches on the overall functionality of the vehicles. According to recent projections, there will be one billion connected items by 2023, and many of these devices including cars will eventually be connected to the Internet. The IoV allows for communication between cars and other living beings. Both V2V and V2I communication protocols are compatible with the Internet of vehicles (IoV). These internal interactions allow an in-car information system a multi-tiered data pipeline, enabling a more informed driving experience communication combines computation, processing, data sharing, and data gathering. Appropriate safety measures and trustworthy cybersecurity solutions are crucial for fending off threats and assaults by Shanmugam et al. [2]. Here, we outline a few common security flaws in CPS systems when ML are powered by these systems.

Attacks on IoT/CPS devices: Hackers get access to sensitive data, modify data, or take control of a device by using a brute force attack to swiftly crack the password and manage the Bluetooth link.

Attacks on Cloud-Based Networks: Cloud processing is a typical requirement in IoT systems, which use various channels such as cellular networks to transmit data. Attackers can exploit these channels to modify the transferred information.

Vulnerable sensors: Attackers might attack them and jeopardize security and public safety. Compromises in primary sensors, such the LiDAR signal, the GPS signal, and the data from the inertial measurement unit (IMU), can pose serious risks.

Network attacks: Cyber-physical Internet of things has the potential to be a network attack point. Attackers can target and endpoints take over undermine the system as a whole. Various protocols can be used to connect and interact with different sensors by other devices. These protocols shortcomings make them vulnerable to assaults from hostile parties by Rehman et al. [3].

Firmware attacks: In this type assault, gives device a malicious firmware update, giving them direct access to the whole system 20 years in order to evaluate security solutions effectively. This survey includes a summary of these research as well as most of the datasets used in the study that was presented. We also discuss the use of a number of popular datasets in an undefeated setting federated models. A shift new learning architecture has brought up various distinct dangers, such as reverse engineering and poisoning by Ren et al. [4]. We not only cover ML algorithms and technologies in this study, but we also provide them as employed in various recent FL research publications. The poll broad help them select a specific research path.

1.2 Artificial Intelligence and V2X

The corporate community, academic institutions, and government organizations have recently shown a significant lot of interest in the V2X platform. Recent advances in computer hardware and technology have made artificial intelligence a crucial component of almost all subjects dramatically increase efficiency and safety in autonomous vehicles. Artificial Intelligence Research (AI research) is the study of intelligent agents. At Dartmouth College, AI research made its debut by Jin and Wang [5].

1.3 Wireless Sensor Network

Research efforts for IoT have not kept up with the fast growth of intrusion security systems settings, owing to the uniqueness acting. Even though the same as WSN techniques cannot be applied to IoT right once. Internet integration, global accessibility, and the variety of employ are some of the contributing reasons grown in addition to the threats that were previously available in limited WSN systems.

1.4 Proactive Security and Adaptability

To survive mobile attackers and adapt to network changes, our key management system employs share refreshing in addition to threshold signature. The term “mobile adversaries” was first proposed by Zhang et al. [6] to characterize hackers who compromise a system briefly before moving on to the next target. In this adversary scenario, an adversary could be able to hack all of the servers over time. Even if the compromised servers are located and taken down, the attacker may eventually get might then be produced by the attacker.

1.5 Security in Ad Hoc Networks

There is an authentication architecture released recommended strategy outlines the protocols and communication formats that are utilized in order to carry out authentication. The architecture supports one or more authentication methods. For such a safe design, our key management system is necessary.

1.6 Security and Node Cooperation

Because mobile devices are ad hoc and wireless, there are extra security challenges for the network design. Since node cooperation is the foundation of ad hoc network functioning and the wireless channel is open to eavesdropping, naturally exposed issues. When conducting a passive attack, get crucial data difficult since does not generate network in an attack, the works to interfere.

2 Literature Review

A List of several review articles that have addressed IoMT security is compared with our main contribution. A number of review papers describe how machine learning assures security, but they only briefly note that ML may be used to offer security in IoMT.

Gyawali et al. [7] examined problems, security threats pertaining to medical equipment privacy and safety, and remedies, such as machine learning-based anomaly detection. Manogaran and Rawal [8] looked at a lot of authentication techniques, and grouped characteristics have repel kinds of assaults. Furthermore, they used advanced methods like game theory and machine learning to categorize authentication strategies. In work published in Goh et al. [9] addressed providing thorough analysis potential effects, remedies put out ML solutions, and their drawbacks. Khan et al.

[10] thorough assessment WBAN, which addressed components. More specifically, the principles used by security attackers who utilized machine learning are discussed. A thorough classification of WBAN's defense mechanisms is also provided, along with an examination of vulnerabilities and functional requirements necessary to safeguard the ecosystem. Out of all studies documented, the use of machine learning (ML) is a tool for security reasons. Instead of providing a thorough analysis, they provide a brief description. However, the current study is analogous to the work of Kaur and Kakkar [11] who looked went over many machine learning (ML)-based approaches that have been deal, including methodology, resources, rather than kept. By examining benefits and drawbacks, datasets used in ML-based techniques during this work expand on the previous research.

2.1 Threat Modeling

That identifies possible suggests that may be taken to reduce these risks. This process consists of five steps, each of which entails figuring out the following elements:

1. Assets are valuable and necessary resources for executing certain tasks and luring the adversary in for a compromise.
2. The attack refers to numerous nodes or systems that might be compromised by an unauthorized attacker.
3. The adversary model refers to the characteristics of the enemy, including his identity, goal.
4. An attacker may use a vulnerability to compromise security since it is one of an asset's defects. Threats are circumstances in which an enemy launches an assault by exploiting the weaknesses of the asset by Wang et al. [12].
5. Security solutions that block, identify, or minimize the consequences of threats are referred to as mitigation measures.

According to Miao et al. [13] IoMT security has special difficulties that cannot be effectively solved by conventional security techniques. The poor processing and storing capacity of IoMT sensors is one such issue that makes resource-intensive security solutions challenging to implement. An author Greener et al. [14] improve intelligence better meet their security needs. As a result, these techniques may work better in dire situations than more conventional access control techniques. Threat signature-based techniques cannot identify zero-day attacks or novel vulnerabilities, but machine learning techniques can. Ghubaish et al. [15] produce a lot of data, and because of its volume, diversity, and velocity, this data may be categorized as big data. Because they may be used to identify anomalous activity early on and learn about typical conduct, these data are important sources of security information that help limit the harm of an attack. An author Dong et al. [16] Because algorithms are able to extract the important features in order to do the classification automatically, they provide an end-to-end security model by doing away with the extraction procedure that is necessary in traditional machine learning approaches. There is still a lack

of study and investigation in IoT malware robustness analysis, despite the fact that malware analysis and detection in general have received a lot of attention. Research on Internet of things malware has not advanced much in terms of examining and identifying malevolent software, particularly when seen through the lens of CFG-based techniques.

Rajasoundaran et al. [17] created a number of machine learning classifiers that use various textual representations taken from the CFGs of the applications to identify Android malware. Additionally, a thorough CFG-based comparison research of Android and IoT malware was carried out by Sagar et al. [18]. AlKhatieb et al. [19] have proposed a congestion aware multi-path DSR protocol latency of it enhances DSR protocol's performance on a crowded network. The comparative results show that the recommended approach improves overhead and greatly reduces current method. Rajeswari [20] offered an analysis of strategies used in various systems, which produced different failure scenarios based on factors such as battery consumption, bandwidth usage, dynamic topologies, and protocol enhancement strategies.

Malhotra and Kaur [21] they contrasted and emphasized the differences between many DSR variants in their investigation. Furthermore, before data is transmitted to the network layer, it is collected from the physical, MAC, and network layers using a multilayer variant of the dynamic source routing (DSR) protocol developed by the authors. This approach allows for improved handling of data within the network infrastructure. Next, they identify the most optimal routing protocol by comparing the current state of the network with established set criteria. Sang et al. [22] evaluated and contrasted procedures carried out using a variety of metrics, and the findings indicate that DSR outperformed AODV in terms of TTL-based hop count while AODV fared better in terms of attaining latency. Nazib and Moh [23] investigated variations because it uses fewer resources in a network with fewer nodes and static traffic, AODV is effective. OLSR functions better in networks with a high node density and irregular traffic patterns.

3 System Design

Anodes continue to connect to the network without relying on a pre-existing infrastructure. Ad hoc networks are adaptable and may be split up, merged, or formed into smaller networks as needed. They also do not require a pre-established infrastructure to manage their operations. Since ad hoc network nodes are usually mobile, wireless communication is likely how they stay linked.

3.1 Reinforcement Learning in Vehicular Ad Hoc Network

Within the field of machine learning, reinforcement learning focuses on teaching an object to behave a specific way through trial and error and interaction. The design of reinforcement learning issues as Markov decision processes and their resolution may be summarized as follows:

3.2 General Presentation

Reinforcement learning is a branch of machine learning that investigates how an agent adapts its actions based on the consequences of its decisions. This method provides a comprehensive explanation of learning behavior and emphasizes two key aspects: learning from the environment and optimizing outcomes over time. For an agent to achieve its goal, it must identify and adopt the most effective actions.

- The reinforcement signal, which is often interpreted as allows change.
- “Agent + environment” comprises a dynamic system.

Because it is aware of the corresponding desired output for each input vector, the learning system in supervised learning, sometimes referred to as “learning with a master,” and is always conscious of the mistakes it makes. In this context, a variable X is defined as a set of actions, where the reward $r(t)$ is the result of performing an action $a(t)$ in a given state $x(t)$. The black hole attack represents a critical threat in mobile networks, especially because it disrupts the functionality of traditional protocols like AODV by attracting and discarding network traffic.

The general process schematized is the following:

1. At step t in time, the agent is in the state $x(t)$,
2. It selects one of the feasible courses of action in this situation, $a(t)$,
3. It uses what triggers as the action:
 - The transfer jurisdiction, $x(t + 1)$,
 - The acquisition support, $r(t)$;
4. $T \leftarrow t + 1$
5. Proceed to 2 or halt current condition 1.

Define X as a collection of actions where the reward $r(t)$ is the outcome of the action $a(t)$ taken in state $x(t)$. In $R(r: X \times A \rightarrow R)$, the application of the product space $X \times A$. At first, it is assumed that spaces A and X are discrete we study scenario, several times in row. At every the agent is provided with a state st at time step t . Subsequently, it employs its policy π , a mapping that connects states st to actions at , to choose one action at from among many potential actions. A scalar reward (r_t) is given to the agent and the next state ($st + 1$) exchange.

Agent goes through this process until it reaches a condition known as terminal, at which point it restarts. Assume that an approximation action-value function with θ parameters is $Q(s, a; \theta)$. Various reinforcement learning algorithms can yield updates to θ . An algorithmic method in this class seeks approach ideal continually, and Q-learning entails the i th represented as

$$Li(\theta_i) = E(r + \gamma \max_{a'} Q(s', a'; \theta_{i-1}) - Q(s, a; \theta_i))^2$$

where state is the state that follows state.

We call this strategy Q-learning in one step as it shifts toward onestep return $r + \gamma \max_{a'} Q(s', a')$. Using step techniques has the disadvantage of the only item that is instantly affected by winning a reward; r is significance of the government's intervention combination s leads award. The values of the state-action pairs that remain are merely marginally impacted by $Q(s, a)$, the most recent value. This might delay down learning since it takes a lot of adjustments spreading an award over all relevant prior phases and deeds.

n -step returns are a useful tool for expediting the spread of incentives. In n -step Q-learning, the n -step return is the goal of updating $Q(s, a)$. A single reward (r) so influences previous as such, and rewarding pertinent state-action combinations may be accomplished in a far more effective manner.

$$R(t) = \sum_{K=0}^{\infty} \gamma^K r_{t+K}$$

3.3 The Return Function

The agent's goal is to maximize the cumulative benefits over time used to quantify benefits. There are three methods to express the return. The following formulas must be applied in a random reinforcement example in order determine the anticipated amount.

In this case, a finite number of future steps corresponds to the horizon in the finite-horizon model. The events that lead from the commencement to the terminal state occur throughout this time:

$$R(t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{K-1} r_{t+K} \quad (1)$$

K represents the number of steps needed to reach final. This number K can be fixed a priori or not. The condition is that a discount factor, γ , is applied to the discounted return when the sequence of operations is infinite becomes $0 \leq \gamma \leq 1$.

The value of γ can be used to modify the length of time the agent evaluates the reinforcements. A "myopic" agent solely considers the immediate reward when $\gamma =$

0, whereas an agent that is more forward-thinking is one where γ approaches 1. The average-reward model, or R-learning, is a method for reinforcement learning that uses the average of upcoming reinforcements as a third criterion. The following uses just another tool in dynamic programming is the discounted return technique.

3.4 The Value Function

An agent chooses based on each state, x_t . One observe.

$$a_t = \pi(x_t) \quad (2)$$

A policy is valued based on the discounted return and the prospect of future reinforcement. According to policy π , the value of state x is determined by:

$$E \left\{ \sum_{t=0}^{\infty} \gamma^t r(x_t, \pi(x_t)) \mid x_0 = x \right\} \forall x \in X$$

In the state x_t , the expected reinforcement is represented by the symbol $E\{r(x_t, \pi(x_t))\}$ when the action suggested by policy π is applied. We may put this more simply one may recursively:

$$V^{\pi(x)=R(x, \pi(x)) + \gamma \sum_{y \in x} \rho_{xy(\pi(x))} V^{\pi(y)}}$$

where the probability of using the $\pi(x)$ action to transition from state x to state y is expressed as $\rho_{xy(\pi(x))}$. One can compute offline they know each probability and the starting state transition, or if they are familiar with a model. A linear system of equations is shown in Eq. 7 and has to be solved. Furthermore, an optimal policy is known to exist, symbolized by π , and explained as follows:

$$R(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n r_{t+k}$$

Ad hoc networks are not fully covered by the security considerations that apply to older networks since ad hoc networking is different from older techniques in a number of respects. The practical security solutions that may be applied are restricted by the ad hoc networking strategy, but the essential security requirements like secrecy and authenticity do not change. This is due to the constant connection that exists between security level and performance.

4 Result and Discussion

We have compared the efficacy contrasting, under a black hole attack, unprotected PROPHET protocol with our TSP over PROPHET through simulated tests utilizing the ONE simulator.

4.1 Simulation Setup

In our simulations, might either a source, a destination, or an intermediary; on other hand, a black hole node is always an intermediary nodes generate traffic that is directed toward other randomly chosen packets might never be forwarded since, even in situations where thought that a node has limited transmission time and buffer capacity. The simulation parameters are summarized in Table 1.

The following performance metrics are considered:

Dropped Message: This indicates how many messages were removed from the buffers of the nodes. It is expected that before a packet is removed from the buffer, it can be sent to a maximum of a certain number of forwarders.

Overhead Ratio: An evaluation of the bandwidth efficiency is calculated as.

(Number of delivered packets – number of delayed packets)/number of delivered packets.

Malicious Count: This indicates how many communications were intercepted by malicious nodes; in a way, it gauges how much harm black hole nodes have inflicted.

Aborted Message: This shows how many times a communication between nodes has been aborted.

Table 1 Simulation parameters

Simulation Parameters	Values
Area	5000 m * 5000 m
Data transfer rate	250 Kbps
Number of Groups	5
Number of nodes other than malicious nodes	80
Buffer space of each node	50 Mb
Speed range	1–6 m/s
Wait time range	1–700 s
Message size	500 Kb to 1 Mb
Message generation interval	25–35 s
Simulation time	22,000 s

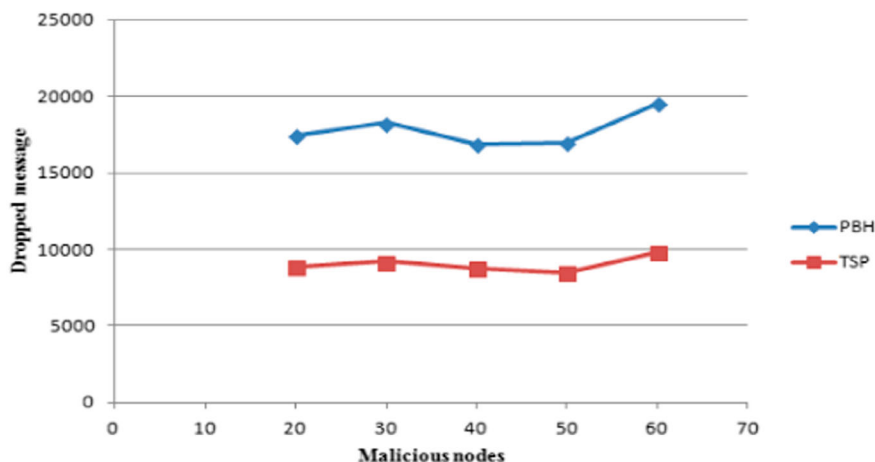


Fig. 1 Dropped message versus malicious nodes

Simulation Results: Initially, manipulate proportion malevolent within the network and assess effect fluctuation quantity removed from the nodes' storage. Figure 1 presents the findings.

This may be explained by the fact that controls the percentage of message dropouts while using technique as regarding the PBH approach. According to malicious nodes, levels really never increase as they never participate in routing. Thus, result in message drops is separated and filtered out by TSP.

After that, we modify and evaluate the system, and the effects of these changes are shown in Fig. 2. The overhead ratio drops from 55.17 to 33.02% when 40% of the nodes are malicious. This is explained by the technique, the malicious nodes initially advertise increased delivery probability, implying more copies of the messages will be sent to them by their neighboring nodes.

Figure 2 illustrates how, as the percentage of malicious nodes increases, using TSP results in a significantly lower overhead ratio than using PBH.

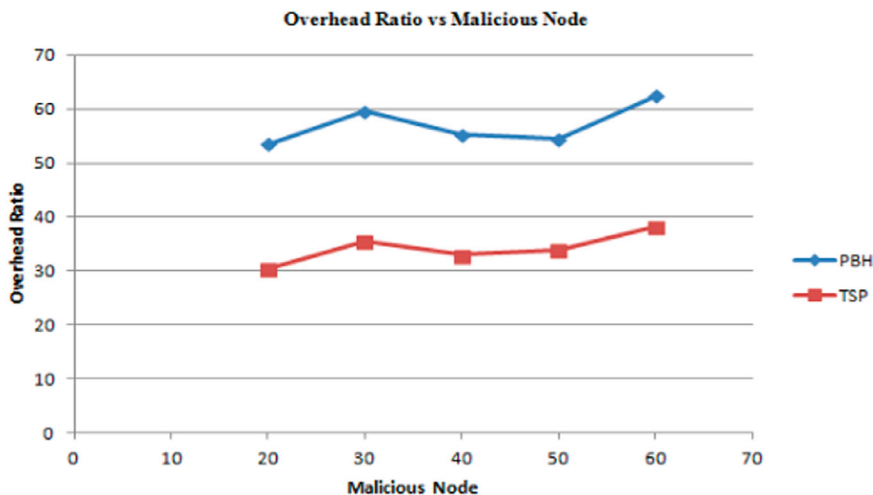


Fig. 2 Percentage of malicious nodes increases

5 Conclusion

We have proposed a TSP to reduce black hole attacks in Oppnets that use PROPHET as the underlying routing protocol. We have compared TSP with susceptible to black hole assaults PROPHET Protocol. The simulation results showed that: (1) TSP significantly lowers the TSP message drop ratio compared to PBH; (2) TSP generates a significantly lower overhead ratio compared to PBH; (3) TSP lowers the malicious count compared to PBH; and (4) TSP lowers node delivery probabilities compared to PBH. These results showed that while the TSP strategy reduces the amount of copies delivered black hole hence lowering network bandwidth use, the PBH methodology wastes more network resources. Many people think that communication is essential to our survival. Information was previously shared through a variety of media, but as technology developed, new network arrangements appeared. Among the finest in the wireless area (WSN) are wireless sensor networks. Self-organizing radio-operating nodes create these networks. They can only launch many varying-level strikes because of the speedier connection. The black hole attack poses a significant threat in mobile networks, particularly affecting the performance of traditional AODV protocols. However, overall performance of traditional AODV protocol may differ greatly when network conditions change. The literature has presented a plethora of techniques for identifying black hole assaults and facilitating data reception for the base station. The literature review that offers methods has a very low number of false positives. In this instance, our research examines black hole assaults and provides further details on countermeasures. To strengthen wireless sensor network defenses against black hole attacks, we provide a set of practical security models.

References

1. He Y, Han G, Jiang J, Wang H, Martinez-Garcia M (2020) A trust update mechanism based on reinforcement learning in underwater acoustic sensor networks. *IEEE Trans Mob Comput* 21(3):811–821
2. Shanmugam S, Prathapchandran K, Janani T (2022) Mitigating black hole attacks in routing protocols using a machine learning-based trust model. *Int J Sociotechnology Knowl Dev (IJSKD)* 14(1):1–23
3. Rehman A, Hassan MF, Hooi YK, Qureshi MA, Chung TD, Akbar R, Safdar S (2021) Context and machine learning based trust management framework for internet of vehicles. *Comput Mater Continua* 68(3)
4. Ren Y, Liu W, Liu A, Wang T, Li A (2022) A privacy-protected intelligent crowdsourcing application of IoT based on the reinforcement learning. *Futur Gener Comput Syst* 127:56–69
5. Jin Q, Wang L (2020) Zero-trust based distributed collaborative dynamic access control scheme with deep multi-agent reinforcement learning. *EAI Endorsed Trans Secur Saf* 8(27)
6. Zhang D, Yu FR, Yang R, Zhu L (2020) Software-defined vehicular networks with trust management: a deep reinforcement learning approach. *IEEE Trans Intell Transp Syst* 23(2):1400–1414
7. Gyawali S, Qian Y, Hu RQ (2021) Deep reinforcement learning based dynamic reputation policy in 5G based vehicular communication networks. *IEEE Trans Veh Technol* 70(6):6136–6146
8. Manogaran G, Rawal BS (2020) Machine learning based trust model for secure internet of vehicle data exchange. In: 2020 IEEE globecom workshops (GC Wkshps). IEEE, pp 1–6
9. Goh Y, Yun J, Jung D, Chung JM (2022) Secure trust-based delegated consensus for blockchain frameworks using deep reinforcement learning. *IEEE Access* 10:118498–118511
10. Khan T, Singh K, Shariq M, Ahmad K, Savita KS, Ahmadian A, Ahmadian A, Conti M (2023) An efficient trust-based decision-making approach for WSNs: machine learning oriented approach. *Comput Commun* 209:217–229
11. Kaur G, Kakkar D (2022) Hybrid optimization enabled trust-based secure routing with deep learning-based attack detection in VANET. *Ad Hoc Netw* 136:102961
12. Wang J, Jing X, Yan Z, Fu Y, Pedrycz W, Yang LT (2020) A survey on trust evaluation based on machine learning. *ACM Comput Surv (CSUR)* 53(5):1–36
13. Miao Q, Lin H, Wang X, Hassan MM (2021) Federated deep reinforcement learning based secure data sharing for internet of things. *Comput Netw* 197:108327
14. Greener JG, Kandathil SM, Moffat L, Jones DT (2022) A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 23(1):40–55
15. Ghubaish A, Salman T, Zolanvari M, Unal D, Al-Ali A, Jain R (2020) Recent advances in the internet-of-medical-things (IoMT) systems security. *IEEE Internet Things J* 8(11):8707–8718
16. Dong S, Wang P, Abbas K (2021) A survey on deep learning and its applications. *Computer Science Review* 40:100379
17. Rajasoundaran S, Prabu AV, Routray S, Kumar SS, Malla PP, Maloji S, Mukherjee A, Ghosh U (2021) Machine learning based deep job exploration and secure transactions in virtual private cloud systems. *Comput Secur* 109:102379
18. Sagar S, Mahmood A, Sheng QZ, Zhang WE (2020) Trust computational heuristic for social internet of things: a machine learning-based approach. In: ICC 2020–2020 IEEE international conference on communications (ICC). IEEE, pp 1–6
19. AlKhatieb A, Felemban E, Naseer A (2020) Performance evaluation of ad-hoc routing protocols in (FANETs). In: 2020 IEEE wireless communications and networking conference workshops (WCNCW). IEEE, pp 1–6
20. Rajeswari AR (2020) A mobile ad hoc network routing protocols: a comparative study. In: Recent trends in communication networks, pp 1–24
21. Malhotra A, Kaur S (2022) A comprehensive review on recent advancements in routing protocols for flying ad hoc networks. *Trans Emerg Telecommun Technol* 33(3):e3688
22. Sang Q, Wu H, Xing L, Xie P (2020) Review and comparison of emerging routing protocols in flying ad hoc networks. *Symmetry* 12(6):971

23. Nazib RA, Moh S (2020) Routing protocols for unmanned aerial vehicle-aided vehicular ad hoc networks: a survey. *IEEE Access* 8:77535–77560
24. Xiao L, Ding Y, Jiang D, Huang J, Wang D, Li J, Poor HV (2020) A reinforcement learning and blockchain-based trust mechanism for edge networks. *IEEE Trans Commun* 68(9):5460–5470
25. Rjoub G, Wahab OA, Bentahar J, Cohen R, Bataineh AS (2022) Trust-augmented deep reinforcement learning for federated learning client selection. *Inf Syst Front* 1–18

X-ray Classification for Content-Based Medical Image Retrieval



S. Subashree, E. Saranya, M. Chinnadurai, and M. Rajakumaran

Abstract CBMIR algorithm appears to address key challenges in medical image retrieval, providing robustness to variations, efficient representation, and a high success rate in retrieval tasks. The emphasis on regional matching and query by example adds practical utility to the proposed method. The algorithm is designed to be robust to scaling and translation of objects within medical images. This robustness is crucial for handling variations in image sizes and positions. The attention is placed on an efficient representation and retrieval of medical images. The CBMIR method allows users to query by example, indicating a user-friendly and intuitive way to retrieve similar images based on a provided example. The algorithm's efficiency and performance have been evaluated on a dataset of about 5000 simulated, realistic computed X-ray images. The dataset includes images selected from three large medical image databases, reflecting a diverse and comprehensive evaluation. The results of experiments indicate a success rate of more than 93%, which is deemed satisfactory. This suggests that the proposed algorithm performs well in retrieving relevant medical images from the dataset. The CBMIR algorithm appears to address key challenges in medical image retrieval, providing robustness to variations, efficient representation, and a high success rate in retrieval tasks.

Keywords CBMIR · X-Ray · Zernike moments · GLCM · SVM

1 Introduction

Content-based medical image retrieval (CBMIR) on X-rays involves the use of computer algorithms to search and retrieve medical images based on their content rather than relying on metadata or textual information. Here is an overview of the

S. Subashree (✉) · E. Saranya · M. Chinnadurai
E.G.S. Pillay Engineering College, Nagapattinam, India
e-mail: ssshreesuba@gmail.com

M. Rajakumaran
SASTRA Deemed University, Thanjavur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
A. Kumar et al. (eds.), *Proceedings of Fourth International Conference on Computing and Communication Networks*, Lecture Notes in Networks and Systems 1317,
https://doi.org/10.1007/978-981-96-3942-7_23

311

key steps and considerations involved in CBMIR for X-ray images: **Image Acquisition:** Gather a collection of X-ray images from various sources, ensuring a diverse and representative dataset. **Image Preprocessing:** Clean and preprocess the X-ray images to enhance features, reduce noise, and standardize the format [1]. Common preprocessing steps include resizing, normalization, and filtering. **Feature Extraction:** Extract relevant features from the X-ray images that can be used to characterize their content. This step is crucial for representing the images in a way that allows for meaningful comparisons. **Indexing:** Create an index or database of the extracted features, allowing for efficient storage and retrieval of information. Indexing methods such as hashing or tree structures can be employed to speed up the search process. **Querying:** Define a query system that allows users to input criteria or examples to search for similar images within the database. This could involve specifying certain anatomical structures, pathologies, or other visual features. **Similarity Metrics:** Implement similarity metrics to measure the likeness between the query image and the images in the database. Common metrics include Euclidean distance, cosine similarity, or more complex metrics based on deep learning embeddings. **Machine Learning Approaches:** Utilize machine learning algorithms to improve the retrieval performance. Deep learning techniques, such as convolution neural networks (CNNs), have shown success in learning hierarchical features from medical images. **Evaluation:** Assess the performance of the CBMIR system using metrics such as precision, recall, and F1 score. This step helps ensure that the system provides accurate and relevant results. **User Interface:** Develop a user-friendly interface for interacting with the CBMIR system, allowing healthcare professionals to easily input queries and interpret the retrieved results. **Security and Privacy:** Implement measures to ensure the security and privacy of medical images, adhering to regulatory standards such as the Health Insurance Portability and Accountability Act (HIPAA).

2 Literature Survey

It is important to note that the field of CBMIR is continually evolving, with advancements in both imaging technology and artificial intelligence contributing to improved retrieval capabilities. An important role for a scientific image is played in patient prognosis, treatment, surgical planning, clinical reference, and training. Interest in controlling virtual medical images is growing as a result of the most recent expansion in the availability of filmless radiography systems [2]. The successful implementation of PACS has revolutionized the way medical images are managed, enhancing the speed, accuracy, and accessibility of diagnostic information within healthcare settings. Picture archiving and communication systems (PACS) in healthcare settings. PACS is indeed a technology that has significantly transformed the management of medical images and information within healthcare organizations. PACS allows healthcare professionals to store, retrieve, and share medical images such as X-rays, MRIs, and CT scans electronically. The system provides quick access to these images, streamlining the workflow within hospitals and specialty clinics. Here are some key

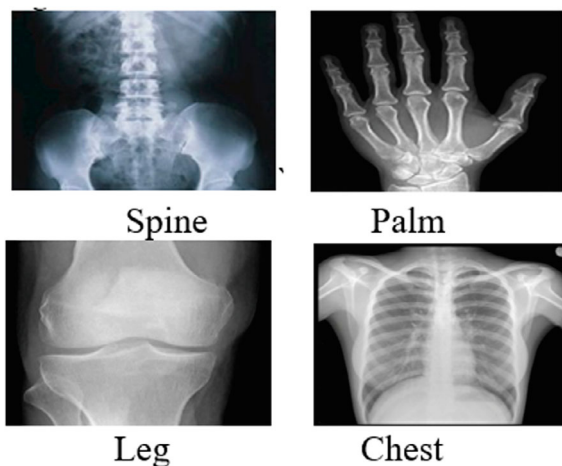
points related to the successful implementation of PACS: efficient access to screening tests, integration into workflow, development of radiological databases, compliance with legal requirements, and support for research and decision-making. Overall, the successful implementation of PACS has revolutionized the way medical images are managed, enhancing the speed, accuracy, and accessibility of diagnostic information within healthcare settings.

Place-based (or sub-photograph-based) image searches and international-based whole photograph searches are the two methods that are typically employed for image retrieval [3]. One important difference between these processes is that, although place-based approaches primarily concentrate on designating a region and receiving a vast array of images with similar items, global-based methods allow for complete photo matching and keep in mind the relevant picture. When someone provides an entire photo as a query, often referred to as a “question-by-example,” the preferred method is to search using global distinction.

Global distinction involves considering the overall features and characteristics of the entire image. Local feature matching focuses on identifying specific features or patterns within smaller regions of an image. In the context of a “question-by-example,” using global distinction is more suitable because it aligns with the idea of considering the overall content and composition of the entire image. This approach helps to retrieve images that share similar global characteristics with the provided example. These types of searches typically take the form of “show me more relevant images that look like this question image.” Zernike moments and texture capabilities are used in this work to categorize X-ray images, and the radial basis characteristic kernel of the guiding vector device is employed as a classifier. Figure 1 displays the X-rays that were captured during our paintings.

The radiologists and hospitals provided the X-rays utilized in these paintings. This set of X-rays comprises images of the hands, feet, and spine. The X-rays are captured as Jpeg files. The images stand for distinct eras, genders, points of view, and

Fig. 1 Different X-ray images



illnesses. As a result, there is significant variation in photo satisfaction. Every image has been reduced in size to precisely fit into a 512×512 bounding box, maintaining the original aspect ratio.

The radiologists and hospitals provide the X-rays used in this project. This set of X-ray pictures comprises photos of the hands, feet, and backbone. The X-rays are captured in .jpeg format. The images depict different eras, genders, points of view, and illnesses. As a result, image quality differs significantly.

3 Existing System

The current work at the CBIR examines the use of query-based picture retrieval (QBIR) to the study of medical records. This could provide the answer to the query that was provided as the input. The process of retrieving the records and several photos related to the input takes longer. This explains the misunderstanding and makes it harder to find and obtain the necessary information. The results of this work are retrieved via textual queries. Its throughput is not constantly green, and it performs poorly.

3.1 CBIR

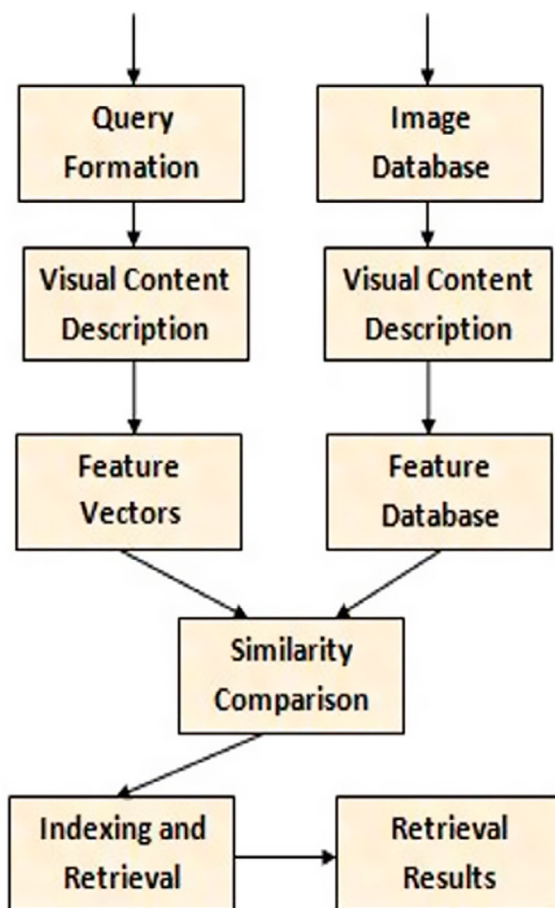
The time period content-based photo retrieval (CBIR) is used to explain experiments into class of photo-based totally on the shapes, textures, and backbone capabilities found in [2, 4, 5]. Seeing that then, the time period has been used to describe the method of retrieving desired snapshots from a big collection on the premise of syntactical image features. In paper [6], the content-based multi-spectral picture retrieves the usage of main thing evaluation (PCA) that is used for characteristic choice. Image class and indexing the usage of facts compression primarily based techniques is presented in [7].

The strategies, equipment, and algorithms which can be used originate from fields together with information, pattern reputation, sign processing, and pc vision. The block diagram of the regularly occurring scientific CBIR system is shown in Fig. 2.

4 Proposed Architecture

In the proposed work, the initial step in CBMIR systems for the classification is carried out with four classes of X-rays, namely spine, palm, leg, and chest images are taken up for the study. Let us break down the key components and processes involved in the proposed system. The proposed CBMIR architecture for the classification of X-rays is shown in Fig. 3.

Fig. 2 Generic CBIR systems



The proposed framework involves creating a feature database for storing the extracted shape and texture features. Similar image retrieval is performed based on these features. CBMIR system utilizes preprocessing, segmentation, SVM classification, and feature extraction techniques (ZMs and GLCM) to enable the effective retrieval of medical X-ray images from a large database. The focus on shape and texture features enhances the system's ability to identify and retrieve relevant images for medical purposes.

4.1 Image Preprocessing

Preprocessing is implemented on pictures at the lowest degree of abstraction, and its goal is to lessen the undesired distortions and decorate the picture facts which

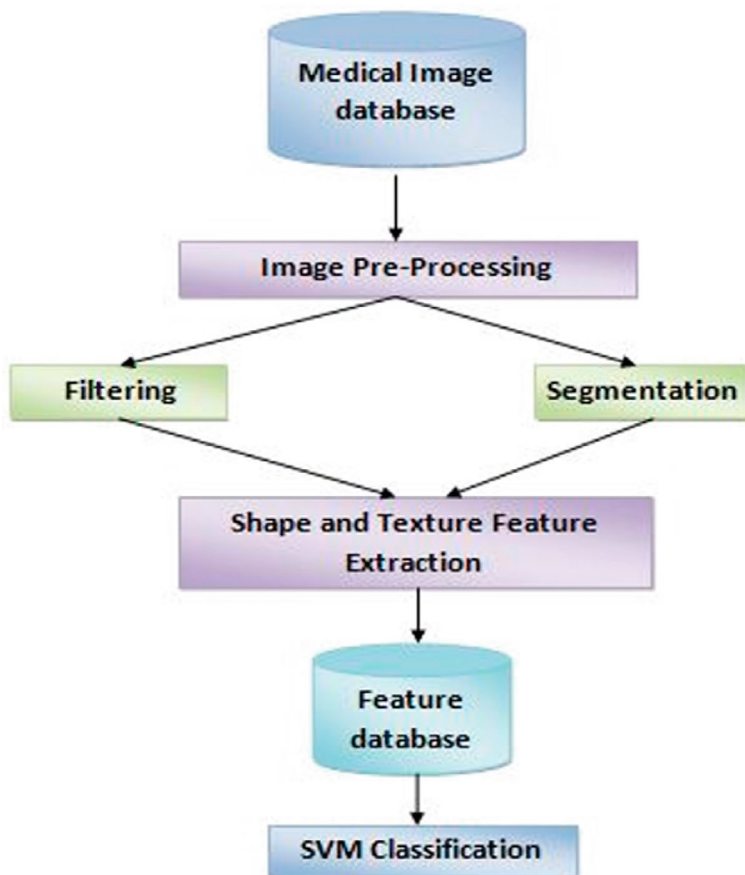


Fig. 3 Proposed CBMIR architecture

are useful and vital for similarly processing. It is also necessary and required for improving the performance of photograph processing techniques like image transform, segmentation, and characteristic extraction and fault detection. X-ray images undergo preprocessing to reduce distortions for better processing. This step aims to increase the quality of images before the analysis. During the stage of preprocessing, the principle intention is to put together the photograph for characteristic extraction through the use of the conventional segmentation process that has been completed. This helps us to eliminate the unwanted noise inside the scientific X-ray pictures. Photograph segmentation is a vital method for image processing. The given snapshots are segmented by using taking the photograph as a place, and it segments the vicinity pictures with the aid of pictures. Therefore, the images are segmented.

4.2 Feature Extraction

Feature extraction involves transforming raw data into a set of features that are more informative and can represent essential characteristics of the data. The goal is to simplify and highlight relevant information in the data, reducing the dimensionality and making it more suitable for further analysis. Zernike moments (ZM) descriptor is mentioned as a method for shape feature extraction. The objective is extracting shape features involves using Zernike moments to characterize the shape of objects in the image. The gray-level co-occurrence matrix (GLCM) is a widely used technique for analyzing textures in images. The objective is extracting texture features involves utilizing GLCM to capture relationships between pixel intensities in different parts of the image. This will be very much useful for constructing the feature database with the extracted features.

4.3 Shape Feature Extraction Using ZM

Shape Features: Zernike moments (ZM) are used for extracting shape features from X-ray images. Shape Feature Extraction using ZM Shape Typically, Shape denotes an essential medical characteristic. Similar-shaped masses could fall under the same category of lesions. It has been suggested that Zernike moments are a useful descriptor for describing shapes. Because of the advantages of these methods, the Zernike moments are used in this study to describe the variation in the mammographic masses. The Zernike moments will not produce an excessive variation even if the item is not segmented precisely. Zernike moments are therefore appropriate for describing masses with ambiguous boundaries.

4.4 Texture Feature Extraction Using GLCM

Texture Features: GLCM extracts texture features from images, and these features are then stored in a database. This approach is valuable for applications where texture information plays a crucial role, such as medical image analysis, remote sensing, and material characterization. GLCM is computed for an image, and various statistical measures can be derived from it to represent different aspects of texture. For example: Contrast: Measures the local variations in intensity. Energy: Represents the uniformity or smoothness of the image. Homogeneity: Describes the closeness of the distribution of elements in the GLCM. Entropy: Indicates the randomness or disorder in the image. These imported features are stored into the new database.

4.5 Classification Using SVM

SVM is a robust algorithm, and the choice of the classifier and its parameters may depend on the specific characteristics of your medical X-ray images. Consider exploring other machine learning approaches and deep learning methods, such as convolution neural networks (CNNs), which have shown success in image classification tasks, including medical imaging. The different regions of the texture images have been discovered to improve the retrieval performance by comparing the categorized texture image with the user’s query photos. The use of SVM for classifying texture regions in medical X-ray images sounds promising, and thorough experimentation and validation will help optimize the model for your specific application. Classification is procedure for determining the different texture regions based on attributes that match. SVM method divides n-dimensional space into classes for simply identifies the optimal line or decision boundary.

SVM aids to create the hyperplane, which selects the extreme vectors and points. The collection of these vectors along with the working principle is referred as support vector machine. Figure 4 shows where a decision boundary or hyperplane is used to classify two distinct categories: class 1 and class 2.

Hyperplane

It is an n-dimensional space can have the decision boundaries used to divide the different classes. The most effective decision boundary for data point classification has been identified by “SVM’s hyperplane” that points out the optimal border.

Support Vector

A support vector is any data point or vector that is closest to the hyperplane and has an impact on the hyperplane’s location. SVM is used for medical image classification,

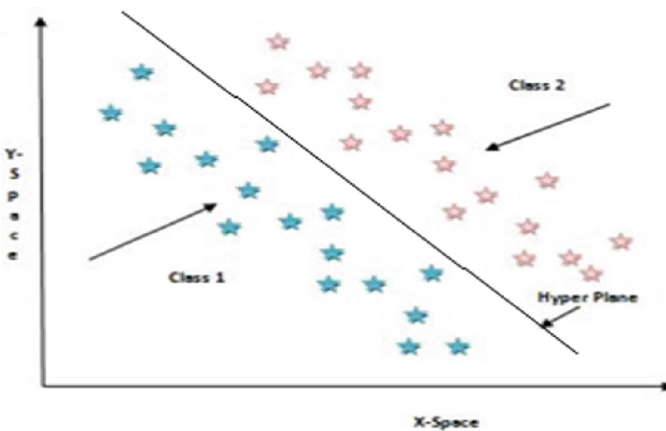


Fig. 4 SVM classification

with the help of the following aspects. **Feature Extraction:** Before applying SVM, it is crucial to extract relevant features from the medical X-ray images. These features should capture the characteristics that distinguish different texture regions.

Training Data

SVM requires labeled training data to learn the characteristics of different classes. Ensure that your dataset is appropriately labeled with the corresponding texture regions. **Kernel Selection:** SVM allows the use of different kernel functions (e.g., linear, polynomial, and radial basis function) to map the data into a higher-dimensional space. Experimenting with different kernels can impact the performance of the classifier. **Parameter Tuning:** SVM has hyperparameters, such as the regularization parameter (C) and the kernel parameters. Fine-tuning these parameters through techniques like cross-validation is essential to achieve optimal performance. **Evaluation Metrics:** Use appropriate evaluation metrics to assess the performance of your SVM classifier. Common metrics for classification tasks include accuracy, precision, recall, and F1 score.

Data Preprocessing

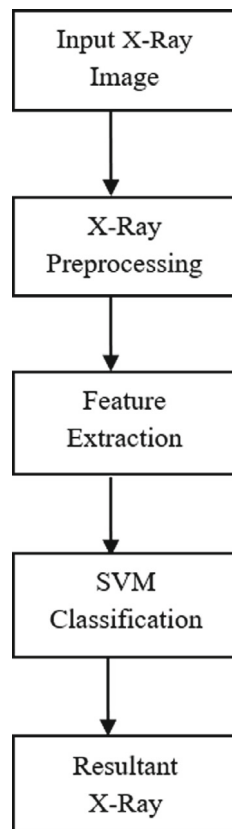
Preprocess your data as needed, which may include normalization, resizing, or other techniques to enhance the performance of the classifier. **Validation and Testing:** Split your dataset into training, validation, and testing sets to train, tune, and evaluate your SVM model, respectively.

The overall process of the proposed methodology is represented in the following algorithm.

Algorithm

```
START
INPUT the X-Ray Image
PRE-PROCESS the given X-Ray Image:
    a).Image Segmentation
    b).Image Filtering
EXTRACT the X-Ray features by GLCM
    8*8 Matrix features
CLASSIFY the X-Ray based on SVM
    Classification based on the extracted features
DISPLAY the Classified X-Ray image
    {Chest, Palm, Leg, Spine}
END
```

The work flow of above algorithm is given in Fig. 5.

Fig. 5 Flowchart

5 Results

Prior to segmentation, the X-rays are prepossessed and filtered using an MPEG-4 Fourier descriptor and range filter. Next, ZM and GLCM are used to extract the shape and the texture features. The SVM is used to classify the X-ray images. The photos are kept in a feature database, which is named after the categorization findings. Consequently, a feature database is made. The extracted GLCM features of an X-rays are represented in Table 1.

First the X-rays are preprocessed that means, segmented and filtered results are shown in Figs. 6, 7, and 8.

Database and Image Formats: X-ray images of the spine, chest, and leg are the contents of the medical database. The X-ray images are stored in various formats, including jpeg, pgm, bmp, and exif, with a preference for using the JPEG format.

Segmentation: MPEG-4 is involved in the segmentation process. Segmentation is a crucial step that involves dividing the image into meaningful regions, potentially aiding in the identification of specific structures or abnormalities.

Table 1 Extracted GLCM features

X-Ray image	Level of energy	Level of homogeneity	Level of entropy	Level of entropy	Level of correlation
Ch1	0.1511	0.9596	8.3180	− 0.7720	0.9494
Ch2	0.1345	0.9402	7.2296	− 0.7709	0.9606
Ch3	0.1023	0.9695	9.1.160	− 0.7801	0.9594
Ch4	0.1861	0.9276	11.035	− 0.7272	0.9601
Ch5	0.0785	0.9626	8.1221	− 0.8322	0.9766
Sp1	0.2491	0.9157	5.1173	− 0.7022	0.9461
Sp2	0.2266	0.9204	6.3115	− 0.7190	0.9581
Sp3	0.2636	0.9368	3.5916	− 0.6390	0.8216
Sp4	0.0831	0.9631	5.6692	− 0.7849	0.9201
Sp5	0.1671	0.9287	5.0842	− 0.7171	0.9415

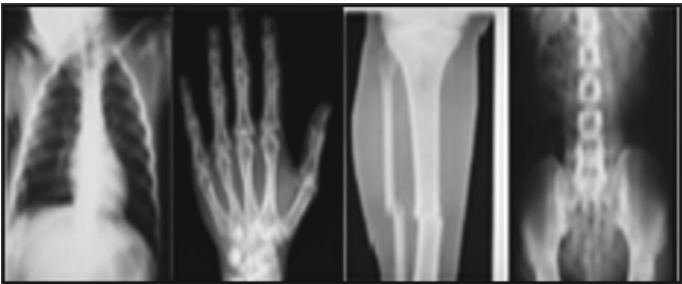


Fig. 6 Sample X-rays



Fig. 7 Output of segmented X-rays

Filtering: Filters are used to reduce noise and make images appear more visually appealing. The choice of filter depends on the specific goal of image processing, such as noise reduction, feature enhancement, or edge detection. Filters are crucial tools for image enhancement and analysis in various fields, including medical imaging,

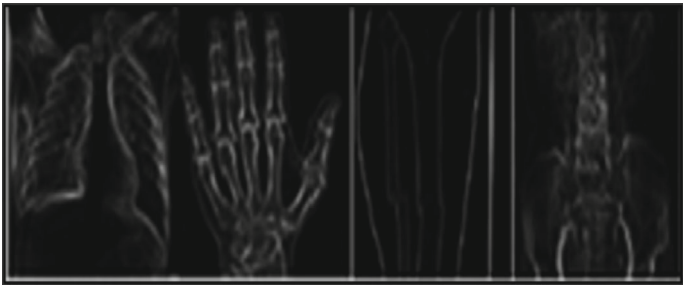


Fig. 8 Output of filtered X-rays

computer vision, and multimedia applications. Here, the canny edge detection filter was applied on the X-rays. Edge detection filters highlight the boundaries between different regions in an image.

Figure 6 gives the input images from records. Figure 7 is the output of the segmented image using proposed architecture, and Fig. 8 is the filtered image. Then, the shape and texture features are extracted using ZM and GLCM on the X-ray images. The results are shown as follows:

Finally the images are classified using SVM, and the result is shown in Fig. 9. Extracted GLCM features of the given input shown in Fig. 10 show the classification of the extracted images using support vector machine.

The feature extraction is given code, which extracts the content based on image processed in the previous section based on the following MATLAB code [8].

```
Img1 = imread('hand1.jpeg');
```

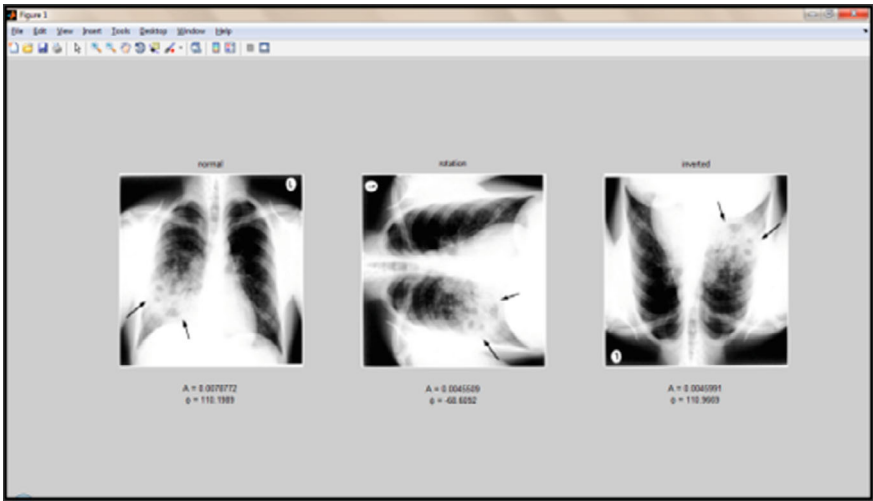
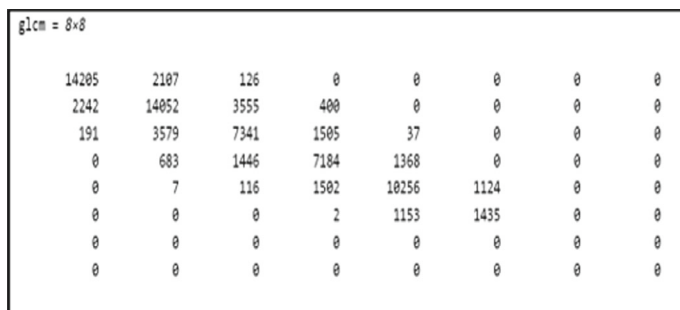


Fig. 9 Shape feature extraction



g1cm = 8x8							
14205	2107	126	0	0	0	0	0
2242	14052	3555	400	0	0	0	0
191	3579	7341	1505	37	0	0	0
0	683	1446	7184	1368	0	0	0
0	7	116	1502	10256	1124	0	0
0	0	0	2	1153	1435	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Fig. 10 Extracted features of GLCM

```

imshow (Img1)
g1cm = graycomatrix(Img1)
g1cm = graycomatrix(Img1,Name = Value)
[g1cm, SI] = graycomatrix(_)
g1cm = graycomatrix(Img1,'Offset', [2 0])
g1cm = 8 × 8

```

The output for the above code is shown in Fig. 10.

Figure 11 shows the SVM classification of the given image and the corresponding output based on the following MATLAB [8] code:

```
classifier = trainImageCategoryClassifier(imds, bag)
```

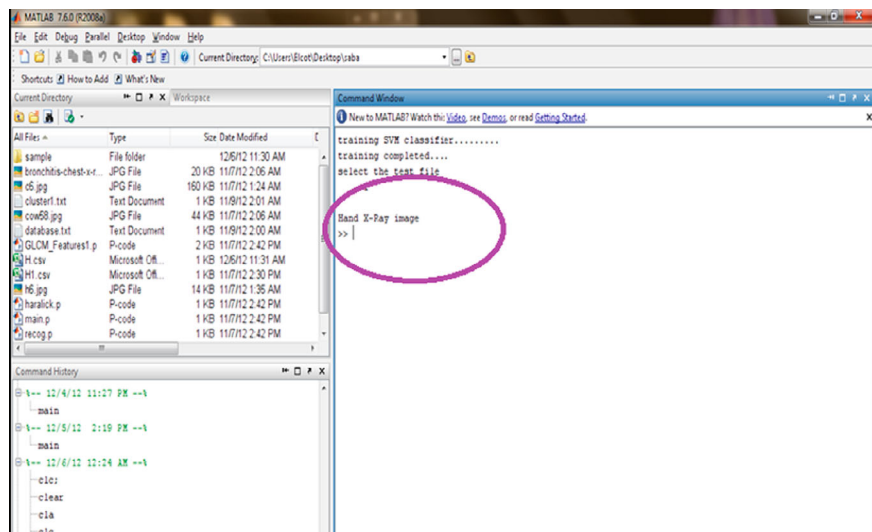


Fig. 11 SVM classification

```
classifier = trainImageCategoryClassifier(imds, bag, Name, Value)
```

The result becomes

```
>>
Hand X-Ray image
>>
```

6 Conclusion

There is undoubtedly a lot of potential for CBMIR technique by using content-based visual information retrieval in the medical field. Zernike moments and GLCM techniques are applied for the extraction of shape, and texture features were proposed in this work to increase the efficiency of content-based medical image retrieval (CBMIR) systems. To categorize the X-rays according to their attributes, an SVM classifier is employed. To improve the quality of similarity search results, it uses a wrapper technique to find the best reduced feature set. This paper presented an MPEG-4 description for textural features to enhance content-based medical image retrieval (CBMIR) systems effectiveness and efficiency. Based on their attributes, the X-ray pictures are classified using an SVM classifier. It uses a wrapper method that enhances the quality of the similarity search results while looking for the optimal reduced feature set. Using the features that were taken out of the X-ray images, this work builds the feature database (FD). In future try to work on comparable image retrieval (IR) using CBMIR techniques, with combining the features of our extracted medical X-rays with the features of our query image. The query and extracted images are matched (indexed) by the query engine. Ultimately, it will generate the output and present the pictures. In the end, it will generate a result and show images that are comparable to the input image we provided. This work demonstrates how feature databases can be created effectively for content-based image retrieval.

References

1. Ramamurthy B, Prasanna M (2022) CBMIR: content based medical image retrieval using hybrid texture feature extraction Method. In: Second international conference on advances in electrical, computing, communication and sustainable technologies (ICAECT). IEEE Xplore. <https://doi.org/10.1109/ICAECT54875.2022.9807970>
2. Hameed IM, Abdullhussain SH, Mahmmud BM | D T Pham (Reviewing editor) (2021) Content-based image retrieval: a review of recent trends. Cogent Eng 8:1. <https://doi.org/10.1080/23311916.2021.1927469>
3. Zhou W, Li H, Tian Q (2017) Recent advance in content-based image retrieval: a literature survey. ArXiv. abs/1706.06064
4. Subramanian M, Lingamuthu V, Venkatesan C, Perumal S (2022) Content-based image retrieval using colour, gray, advanced texture, shape features, and random forest classifier with optimized particle swarm optimization. Int J Biomed Imaging 2022:14. Article ID 3211793. <https://doi.org/10.1155/2022/3211793>

5. Alrahhhal M, Supreethi KP (2019) Content-based image retrieval using local patterns and supervised machine learning techniques. In: 2019 Amity international conference on artificial intelligence (AICAI), Dubai, United Arab Emirates, pp 118–124. <https://doi.org/10.1109/AICAI.2019.8701255>
6. Lehman TM, Guid MO, Thies C, Fischer B, Keysers D, Kohnen M, Schubert H, Wein BB (2004) Content-based image retrieval in medical applications for picture archiving and communication systems. Department of Medical Informatics, Aachen, Germany
7. Kavitha C, Roa BP, Govandhan A (2011) Image retrieval based on combined features of image sub-blocks. *Int J Comput Sci Eng (IJCSE)* 3(4). ISSN: 0975-3397
8. MATLAB trainImageCategoryClassifier—MathWorks
9. Ramamurthy B, Chandran KR, Aishwarya S, Janaranjani P (2004) CBMIR: content based image retrieval using invariant moments. GLCM and Gayscale Resolution for Medical Images
10. Iqbal K, Odetayo MO, James A (2012) Content-based image retrieval approach for biometric security using colour, texture, shape features controlled by fuzzy heuristics. *J Comput Syst Sci* 78:1258–1277
11. Ramamurthy B, KR Chandran (2012) Content based medical image retrieval with texture content using gray level co-occurrence matrix and k-means clustering algorithms. *J Comput Sci* 8(7):1070–1076. ISSN 1549-3636
12. Singh C (2012) An effective image retrieval using the fusion of global and local transforms based features. *Optics Laser Technol* 44:2249–2259
13. Singh BK, Sinha GR, Mazumar B, Khan MI (2010) Content based retrieval of X-rays using fusion of spectral texture and shape descriptors. In: 2010 International conference on advances in recent technologies in communication and computing 9789-0-7695-4201-0/1010.1.1109/ARTCom.2010.51
14. Chang Y, Antani S, Lee DJ, Gledhillc K, Long LR, Christensena P (2009) CBIR of spine X-Rays on inter-vertebral disc space and shape profiles. Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT
15. Müller H, Michoux N, Bandon D, Geissbuhler A (2004) Content-based image retrieval systems in medical applications—clinical benefits and future directions. *Int J Med Inf* 73:1–23
16. Jin H, Sun A, Zheng R, He R, Zhang Q, Shi Y, Yang W (2004) Content and semantic context based image retrieval for medical image grid. National High Technology Research and Development Program of China, No. 2006AA02Z347 and No. 2006AA01A115
17. Shahbazi H, Soryani M, Kabiri P (2008) Content based multispectral image retrieval using principal component analysis. *CVIR '08*, July 7–9, 2008, Copyright 2008 ACM 978-1-60558-070-8/08/07

A Node to Node Sensory Cloud-Driven Smart Drip Irrigation Model with Relayed Connectivity Structure



Sakshi Priya, Pranav Prakash, Priyansh Kumar Singh, Sushruta Mishra, and Muntather Almusawi

Abstract Smart irrigation systems can produce water according to actual needs. It is known that water consumption is increasing. This article is dedicated to the development of the Internet of Things (IoT). Considering this request, in this article, it is preferred to use a smart water system using Android or Web application, and a solution is presented to constantly monitor and control the flow of water to prevent human negligence and continuous management problems waste water. It also helps prevent water from being sent to plants that depend on water sources. Also suitable for agriculture, lawn, park, etc. It can also be used in areas. Thanks to advances in technology, risks are always reduced and things become easier. Such a solution can be provided through embedded systems. This article creates a smart water flow with the implementation of the Internet of Things and web application. It has been modeled to be accurate and useful in order to provide convenience to the targeted audience during drip irrigation. This application helps in controlling water flow in garden pump. Data like moisture and temperature are given to the microcontroller, ensuring accurate prediction of the garden's water needs. Manage and view these systems from your Android phone via the All Web App.

Keywords Internet of Things (IoT) · Drip irrigation · Shrewd water system · Microcontroller

S. Priya · P. Prakash · P. K. Singh · S. Mishra (✉)

Kalinga Institute of Industrial Technology, Deemed to Be University, Bhubaneswar, India
e-mail: sushruta.mishrafcs@kiit.ac.in

M. Almusawi

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University, Najaf, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Al Diwaniyah, Al Diwaniyah, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Babylon, Babylon, Iraq

1 Introduction

Especially in India, as the population continues to increase and the main source of income is agriculture, the demand for food products is also increasing. The demand for fresh water is also increasing in order to use water in a healthy and sustainable way. Currently, agriculture accounts for 83% of India's total water use [1]. In many uses, water is not used as planned. For this reason, water is wasted in many places. This shows that there is a need to develop such systems to solve the wastewater problem. Plant health management is an important science that tackles plant diseases by limiting plants from reaching their genetic potential. Monitoring plant health is an important branch of plant management in this field [2]. Farmers often use water pumps for irrigation, which requires a lot of time to fill the soil. Many researchers have investigated the use of automatic and semi-automatic irrigation and demonstrated their potential to change water use or water use patterns [3]. This type of water usage can easily be replaced by using automated systems instead of traditional machines. The Internet of Things (IoT) now plays an important role in precision agriculture. IoT technology can create communication between different devices or between shared devices or wireless systems. This device provides a solution for networking enabled by the Internet. IoT plays an important role in this; precision agriculture tries to reduce the waste of different efforts, time and other resources to increase the yield and cultivated area. This paper presents a smart drip irrigation system based on the Internet of Things that can monitor crop yield without using data fusion; here humidity sensors are used together with soil and temperature sensors to control humidity. If environmental changes cause changes in soil and temperature, the sensors update the data and activate the water pump. Communication interrupt signals are sent to the microcontroller and NodeMCU. The smart drip irrigation system connects to the internet or Android application, allowing users or farmers to use the IoT cloud platform to control data or processes from their mobile phones and manage data management stored in the system. The web server database can be checked at any time. IoT-based smart drip irrigation systems can be used to get instant crop or crop health information. The advantages of this system are remote operation, energy savings and a low-cost IoT-based control solution for different water uses. The aim of this article is to develop a new IoT-enabled smart drip irrigation system using sensor fusion method, which will effectively reduce water and electricity shortages. This also helps reduce work by controlling the usage of the phone and preventing excess moisture. The main points of this article are:

- Design of smart drip irrigation systems using sensor fusion method.
- Create web and Android applications using IoT to monitor and control different data and production facilities.
- View and manage aggregated sensor data using your Android phone.

2 Literature Survey

Some researchers have warned against IoT-related development activities for water applications due to past issues. Earlier studies developed a water system that uses sensors to automate the process. The system saves farmers money, time and energy compared to harvesting crops manually. While some proposed solutions to manage water usage using IoT, others propose solutions such as light usage, temperature, and humidity developed a system to monitor the health of plants using IoT to collect weather data. This data is then processed using the IoT cloud platform and alerts are sent to the user's mobile phone if a problem is detected. The system plans the irrigation well. Tau et al., Arif et al. researched a fuzzy logic control system that develops a plantation-specific meeting to determine the most appropriate irrigation program. Reference [4] uses a neural network model to predict humidity data using sparse weather data. The proposed model is then validated with clinical data. Giri and Wavhal [5] proposed a solution by combining WSN and linear programming to use more accurate locations in agriculture. The application of wireless sensor networks makes water monitoring more efficient by introducing low-cost wireless control systems. Salazar et al. [6] developed an intelligent multi-agent control panel for water resources using fuzzy control system, while Hussain et al. [7] attempted to use WSN to continuously collect data on soil and moisture content in a given area to manage water supply. Researchers Reddy et al. [8] developed a model that uses evapotranspiration to determine soil deficiency, air humidity, electricity, and yield levels. This information can be used to determine ecological conditions and the types of crops to be grown. Parwatkar and Bhagat developed a wireless system called Zigbee for automatic water management in WSN. Zigbee has advantages such as low-power consumption, low cost, and the use of wireless network topology network model. Darwin et al. [9] developed a water-efficient system that uses solar energy to drive pumps, thus reducing energy consumption. Their findings suggest that solar energy is a viable option for low-maintenance environments. Hailin et al. [10] proposed a decision-making system for water management in agricultural products. The system uses its own nodes placed in fields to collect data on humidity and climate change. Geng and Dong [11] used deep learning techniques to develop a model for the distribution of soil and soil moisture in the region. The model helps determine the right amount of land to grow better crops. Bagwan et al. The authors of the study conducted by [12] investigated how to conserve water while practicing green agriculture. They use a combination of IoT, artificial intelligence and neural network algorithms to create a smart wireless network. The aim is to make agriculture more profitable and environmentally friendly. Sarojini et al. [13] developed an automatic drip irrigation system using a programmable logic controller (PLC) and biosensors to monitor and control the drip irrigation system. Pooja et al. [14] demonstrated the effect of water dripping. The system uses sensor nodes to collect data, which is then updated to the central cloud server. Shekhar et al. [15] developed a method to predict soil surface using artificial intelligence and nearest neighbor (KNN).

3 Proposed Methodology Design

The basic layout process, a new design for IoT-based drip irrigation, is shown in Fig. 1. Two or more sensors are integrated into a factory process and the sensor control system is designed using the sensor fusion method. Many IoT devices send communication over Wi-Fi type system and different sensors are associated with IoT devices. In addition, the cloud server is connected to the center to store this information. To get the correct information from the pump, which also gives the instructions. As shown in Fig. 2, there is a full description of the infrastructure and IoT devices. Different ways microcontroller and node connect to MCU. The sensors send data to the microcontroller and microcontroller unit (nodeMCU) and react according to soil moisture. The pump will operate according to the water need coming from the solenoid valve, where the humidity threshold in the microcontroller determines the conditions and rules. All results obtained from the sensors are displayed on the computer [16–18]. If the value is higher than the threshold, the pump will automatically shut down. The information on the web page will be updated and the web page information will be stored in the cloud, which will also send information about water during planting. A mobile application was also developed to track the data. Users can easily access detailed information about agriculture anytime, anywhere.

Here we have proposed a model of smart drip irrigation which will help in farming in more efficient and modern manner. Previously farmers have to spend whole day and night in their field so as to protect their crops, water them providing manure, etc. But using our model, they can not only save time and money but also they will be able to increase their yield multiple times as it increases their efficiency 100 folds and helps in preventing drought like situation and this model also helps in using the available water in judicious manner. Sensors used in the model are as follows:

- **Soil Moisture Sensor (FC-28)**

This sensor is used to measure volumetric water as it enters ground conditions, thus providing an output moisture level [19, 20]. The input voltage should be 5 V, while the required output voltage varies between 0 and 4.2 V depending on the moisture in the soil.

- **DHT22 (Humidity and Temperature Sensor)**

This sensor allows us to measure the relative humidity and temperature of the environment. An input voltage of 5 V is required to activate.

- **DC Submerged Waterpump**

It is basically used to supply the water to the plantation/field properly.

- **Solenoid Valve**

This valve helps control the flow or direction of water flow, especially if a ball valve is used.

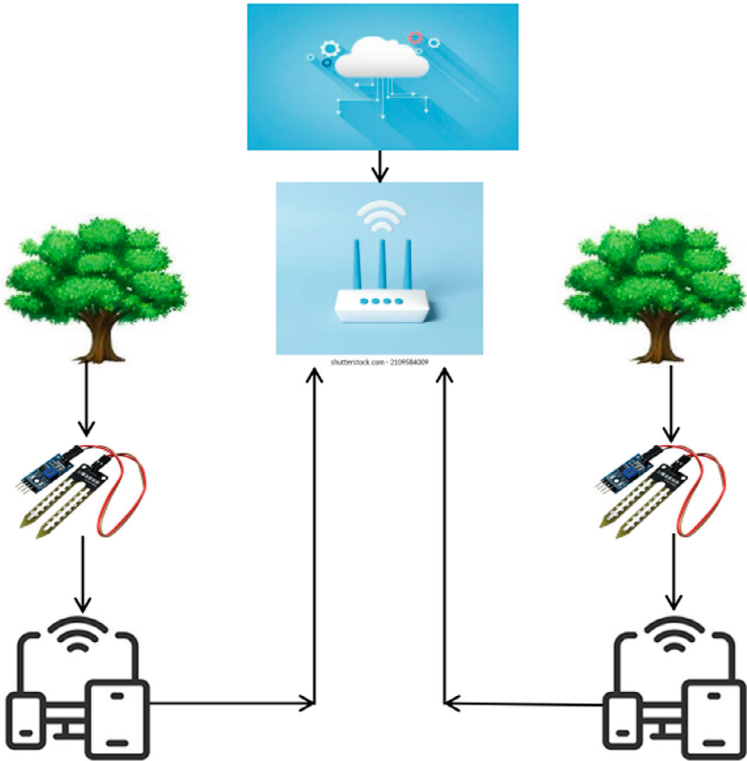


Fig. 1 Model design prototype

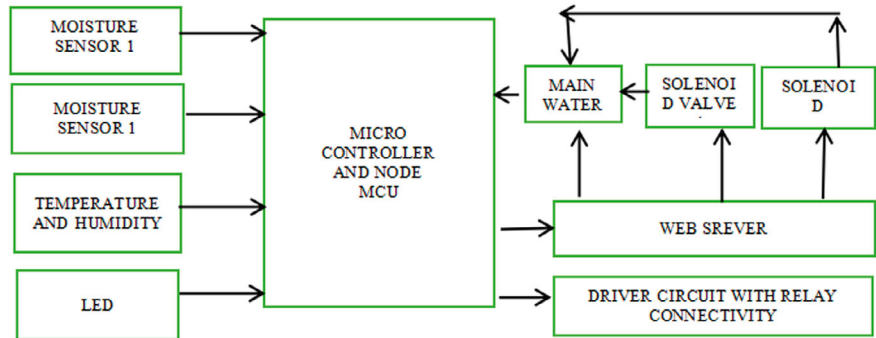


Fig. 2 Detailed schematic diagram of the control system for IoT-enabled smart drip irrigation system

- **Arduino YUN**

It is a basic microcontroller board with onboard ATMEGA32U4 microcontroller and atheros ar931 with Wi-Fi and Ethernet ports [21]. This board is used in controlling the system and to communicate data within the network. It requires 5 V to run.

- **NODE MCU (ESP8266)**

It has chips that interact with the open source platform during design and hardware interface.

- **Double Channel Relay**

These are the switches which are utilized during opening and closing of the motors. This consists of 3 contactors.

- (i) Open (NO)
- (ii) Closed (NC)
- (iii) Common (COM)

When COM is connected to NO, the input operation is OFF; when COM is connected to NC, the input operation is ON.

- **Power Supply**

Its an 5 V VDC multichannel power supply which provides power to sensors, DC motors, and microcontrollers.

- **Routers**

It basically acts as a network connection and sends information in packets to the computer network for proper processing [1, 22]. The robot can be connected using IoT technology using this router to connect our entire system to the internet.

- **IoT Cloud Platform**

It is a platform for processing and storing information. In this system, the speaker uses the IoT cloud platform to manage the entire process [23].

- **Android Studio**

Here we will develop an android/web application which will use java integrated script which consists of editable code features.

The generic control system for IoT-enabled smart drip irrigation system is shown in Fig. 2.

The system architecture model flowchart is designed with three layers as shown in Fig. 3:

1. Sensing layer
2. Internet layer
3. Application/User Interface layer.

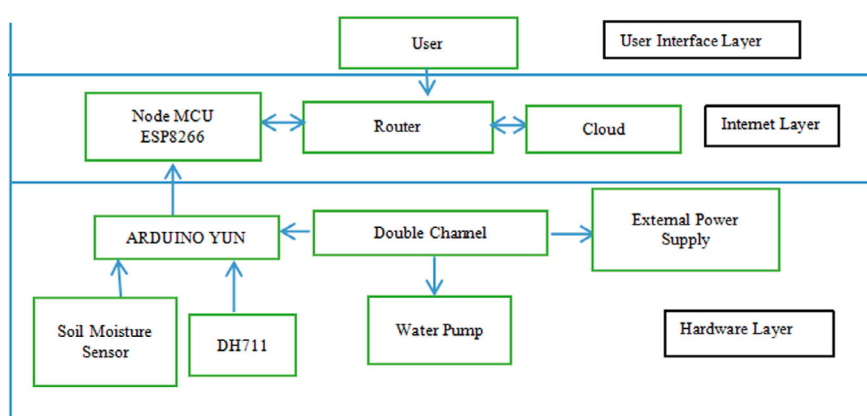


Fig. 3 Overall block diagram for IoT-enabled smart irrigation system

Humidity, soil moisture and temperature measurement from measurement systems are combined with Arduino microcontrollers. The DC power supply powers all sensors and devices connected to the data communication system [24]. The humidity sensor is connected to the Arduino YUN, which detects the humidity and controls the pump through a solenoid valve. The microcontroller receives feedback from the pump and the solenoid valve wall to act accordingly. The proposed system uses a central microcontroller unit programmed to receive feedback from various sensors embedded in the field. The web application connects to the Arduino YUN to control the system. Data communication between hardware and software is achieved by two-layer wireless communication via Arduino and Node MCU sensors. The web application generates a control signal sent by the MCU ESP 8266 node to establish a wireless connection between the Internet layer and the user layer [25]. The configuration of the router enables communication between the Android application and the IoT cloud through the user interface layer, which is responsible for managing and managing different parts of the system.

The pump is a crucial component of our architecture as it supplies water to the field. The water flow process is illustrated in Fig. 4 using a flowchart. This flowchart specifically pertains to the drip irrigation system. During operation, all sensors are configured together with the Arduino YUN. The Arduino YUN is also configured with the network system to exchange data. After configuring the network, a critical check operation is performed to identify any issues. In case of any critical problems, a buzzer is activated. The pump starts when water is required for the plants and continues to flow until the ideal moisture level is achieved. Soil moisture sensors are used to monitor the moisture level, and if it is sufficient, the pump is automatically turned off. This process is repeated whenever the soil moisture falls below the ideal level. Furthermore, this process is integrated with NodeMCU to send a message to the user indicating that their goal has been achieved. The data is sent to the IoT cloud using the ThingSpeak IoT cloud platform, where it is also stored. In

summary, when the system starts, all sensors become active and begin their tasks. The system includes various sensors such as soil moisture sensors, fusion, temperature, humidity, and LED, which are connected to the microcontroller. The microcontrollers used are Arduino and NodeMCU Esp8266 with built-in Wi-Fi SOC (system on chip). Arduino and NodeMCU communicate serially via jumper connections, sending sensor data to the microcontroller and from the microcontroller's web server to the cloud. NodeMCU acts as a web server, allowing communication with users and displaying data in the cloud. The software receives data from the cloud, where the system is already equipped with the ideal measurements of various parameters for that specific field. A comparison is then made between these two sets of data, and a result is generated after analysis. If any irregularities are detected in the parameter range, the system will automatically start supplying water to the field until all the parameter measurements reach the ideal state.

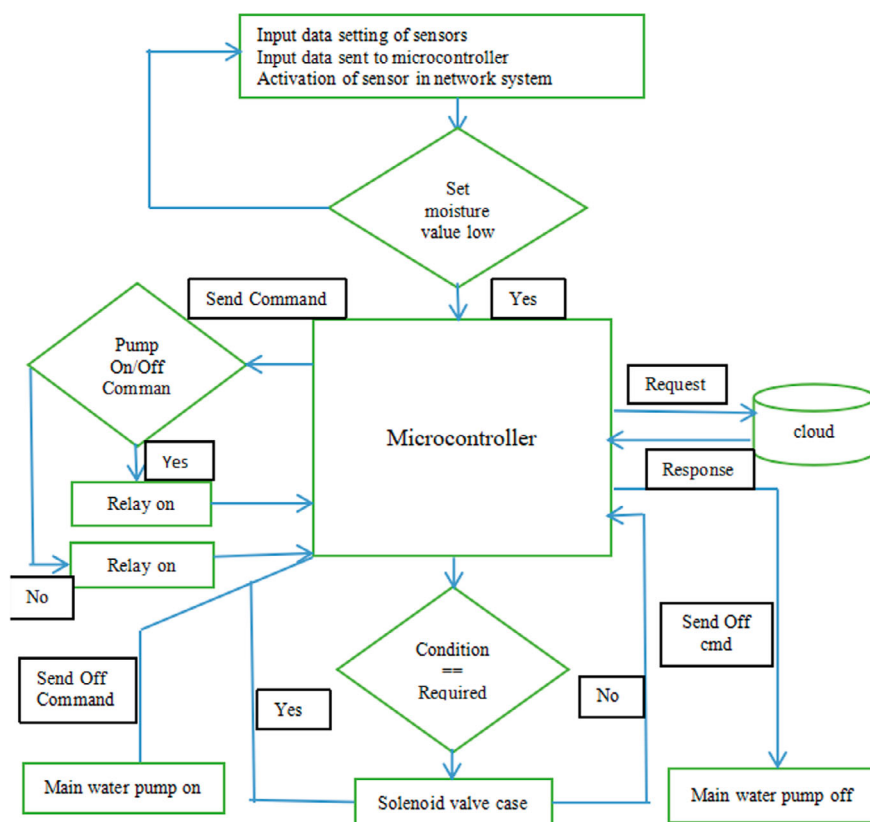


Fig. 4 Control mechanism for water pump

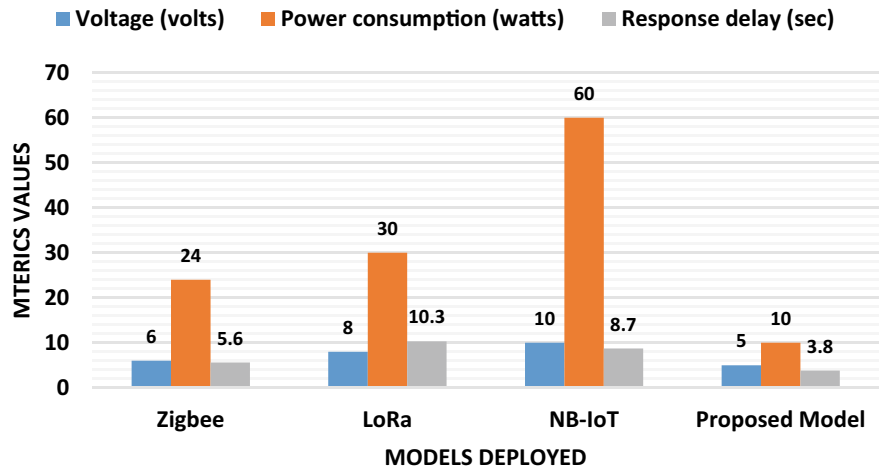


Fig. 5 Comparison of metrics between existing technology and proposed technology

4 Result Analysis

The proposed Sensory Cloud-driven Smart Drip Irrigation Model with Relayed Connectivity model is evaluated in context to different parameters like voltage level requirement, energy consumed, response time among others. Figure 5 demonstrates the overall analysis of the proposed model with other related models like Zigbee and LoRa. It was observed that the proposed model recorded the most promising outcome. It needed approximately 5 V with 10 units of power consumption and response delay of only 3.8 s.

The data rate analysis is also performed taking the proposed model and other related models into consideration. It is noticed that an overall data rate of 320 kbps is the outcome with the proposed model in comparison with others. Figure 6 highlights the results.

Figure 7 illustrates the error rate analysis in context to the RMSE metric taking varying size of data samples for validation. The least error rate of 0.218 is observed with dataset size of 100 samples. Still even if the data size is increased further, the error rate still does not rise in uncontrolled manner.

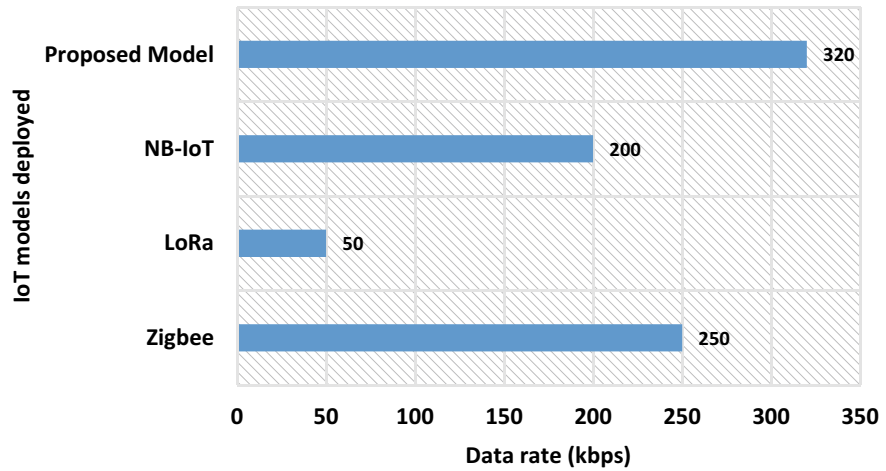


Fig. 6 Comparison between data rate (kbps) of different models

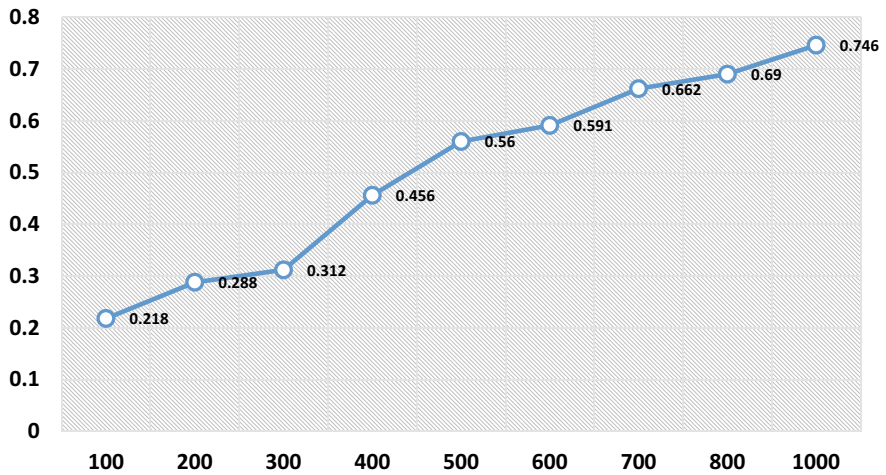


Fig. 7 Root mean square value for different data of proposed model

5 Conclusion

In summary, the integration of the Internet of Things (IoT) with smart water systems has proven to be a revolutionary and promising step in agriculture. This research paper explores various aspects of IoT applications in agriculture with a special focus on smart irrigation. Integration of real-time data from sensors, weather, and soil conditions provides farmers with accurate and timely information, allowing them to make cuts. Be conscious of water use, crop health and productivity. The benefits of

IoT smart irrigation systems are not limited to water quality management. Proposed benefits include increased efficiency, reduced resource waste and increased sustainability. Farmers are now equipped with remote monitoring features that allow them to manage their irrigation systems from anywhere, making it easier and more convenient. Additionally, the use of IoT in agriculture has made it easier to coordinate and respond to the planting process, paving the way for more accurate farming. This not only increases efficiency, but also helps protect water resources and solve the problem of global water scarcity. However, it is important to be aware of agricultural IoT-related issues and considerations that require technology launch, including data security, privacy concerns, and investments. Continued research and development in these areas will help solve these problems and enable widespread adoption of IoT in agriculture. As we move forward, collaboration between researchers, policy-makers, and industry stakeholders will play a key role in improving and expanding the Internet of Things. Continuous exploration of new technologies and the integration of machine learning and artificial intelligence with the Internet of Things will strengthen smart agriculture. In summary, the integration of IoT in smart water supply enables the future prospects of agriculture by providing a foundation for sustainable development. Solutions that improve resource use, increase efficiency and contribute to global food security.

References

1. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digital Health* 10
2. Jain S, Tripathy HK (2024) Machine learning methods for the timely identification of autism spectrum disorder in toddlers. In: 2024 International conference on emerging systems and intelligent computing (ESIC). IEEE, pp 515–520
3. Kashyap P, Pareek A, Mishra S, Khan Z, Garg R, Tripathy HK (2024) Sentiment polarity analysis of twitter data using machine learning models. In: International conference on innovative computing and communication. Springer Nature Singapore, Singapore, pp 623–635
4. Karim F, Frihida A (2017) Monitoring system using web of things in precision agriculture. *Procedia Comput Sci* 110:402–409
5. Kansara K, Zaveri V, Shah S, Delwadkar S, Jani K (2015) Sensor based automated irrigation system with IOT: a technical review. *Int J Comput Sci Inf Technol (IJCSIT)* 6(6):5331–5333
6. Bedekar SS, Mechkul MA (2015) IoT based automated irrigation system. *Int J Mod Trends Eng Res (IJMTER)* 2(7):1532–1538
7. Bedekar SS, Mechkul MA, Deshpande SR (2015) IoT based automated irrigation system. *Int J Sci Res Dev* 3(4):1015–1018
8. Siddagangaiah S (2016) A novel approach to IoT based plant health monitoring system. *Int Res J Eng Technol (IRJET)* 3(11):880–886
9. Umair SM, Usman R (2010) Automation of irrigation system using ANN based controller. *Int J Electr Comput Sci IJECS-IJENS* 10(02):41–47
10. Touati F, Al-Hitmi M, Benhmed K (2012) A fuzzy logic based irrigation management system in arid regions applied to the State of Qatar. In: Sustainable irrigation and drainage IV, vol 168. WIT Trans Ecol Environ, pp 189–199. <https://doi.org/10.2495/SI120161>
11. Arif C, Mizoguchi M, Setiawan BI, Doi R (2012) Estimation of soil moisture in paddy field using artificial neural networks. *Int J Adv Res Artif Intell (IJARAI)* 1(1):17–21

12. Giri M, Wavhal DN (2013) Automated intelligent wireless drip irrigation using linear programming. *Int J Adv Res Comput Eng Technol (IJARCET)* 2(1):1–5
13. Salazar R, Rangel JC, Pinzón C, Rodríguez A (2013) Irrigation system through intelligent agents implemented with arduino technology. *Adv Distrib Comput Artif Intell J* 2(3):29–36. <https://doi.org/10.14201/ADCAIJ2014262936>
14. Hussain R, Sahgal JL, Anshulgangwar MR, Riyaj MD (2013) Control of irrigation automatically by using wireless sensor network. *Int J Soft Comput Eng (IJSCE)* 3(1):324–328
15. Reddy GRS, Manujunatha S, Sundeeep Kumar K (2013) Evapotranspiration model using AI controller for automatic irrigation system. *Int J Comput Trends Technol (IJCTT)* 4(7):2311–2315
16. Singh H, Sharma N (2014) A review of fuzzy based expert system in agriculture. *Int J Eng Sci Res Technol* 3(7):912–915
17. Khamkar NU (2014) Design and implementation of expert system in irrigation of sugarcane: conceptual study. *NCI2TM: 2014, Sinhgad Institute of Management and Computer Application (SIMCA)*, pp 55–58
18. Parwatkar SA, Bhagat VB (2015) Producing more crops in automated irrigation system using WSN with GPRS and Zigbee—a review. *Int J Curr Eng Technol* 5(2):771–777
19. Darwin S, Pousia S, Saranya S, Stephilda J, Vellammal S (2016) Intelligent monitoring of agriculture environment for production enhancement using alternate energy resources. *Int J Innov Res Comput Commun Eng* 4(3):3618–3624
20. Hellin HN, Martínez-del-Rincon J, Miguel RD, Valles FS, Sanchez RT (2016) A decision support system for managing irrigation in agriculture. *Comput Electron Agric* 124:121–131. <https://doi.org/10.1016/j.compag.2016.04.003>
21. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
22. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-DETECTOR: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
23. Pranjal P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. In: *AIP conference proceedings*, vol 2919, no 1. AIP Publishing
24. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system. *Internet of Things* 24:100978
25. Mishra S, Voley DR, Bohra N, Alfarhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306

Enhanced Cyberbullying Detection Using Named Entity Recognition in Online Social Networks



Suresh Sankaranarayanan, K. Meenakshi, and K. Sandhiya

Abstract Cyberbullying, in spite of being a common phenomenon, has been posing various obstacles while employing natural language processing (NLP). To leverage NLP techniques in the detection of cyberbullying, the corpus must be annotated with accurate sentiments and targeted. This will improve the accuracy at which the sentiment targets are being predicted. In this study, we have incorporated the BERT model. We have fine-tuned the model using the hate speech Twitter dataset. Named entity recognition (NER) has been explored for its efficacy in identifying aspect-targets from tweets that don't contain the "@" symbol in the username explicitly. Our proposed system shows better results than the existing state-of-the-art tools used for sentiment analysis.

Keywords Machine learning · Sentiment analysis · BERT · Gemma · Named entity recognition

S. Sankaranarayanan (✉)

Department of Computer Science, College of Computer Science and Information Technology,
King Faisal University, Al-Ahsa, Saudi Arabia
e-mail: ssuresh@kfu.edu.sa

K. Meenakshi (✉)

Department of Networking Communications, School of Computing, SRM Institute of Science and
Technology, Kattankulathur, Tamil Nadu, India
e-mail: meenaksk@srmist.edu.in

K. Sandhiya

School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu,
India
e-mail: sk7663@srmist.edu.in

1 Introduction

As society increasingly integrates technology into daily life, the rise of cyberbullying, particularly among young individuals, has become a critical concern. Research studies, including those conducted by Hinduja and Patchin, reveal that a significant proportion of youth have encountered cyberbullying, often employing electronic platforms to intimidate or harass their peers [1, 2]. This harmful behavior can manifest through derogatory web content, videos, or social media profiles. Reports such as ScanSafe’s “Global Threat Report” shed light on the prevalence of offensive material online, including explicit content and offensive language found in numerous blogs. These findings underscore the urgent need for robust online safety measures to safeguard individuals, especially vulnerable youth, from the detrimental effects of cyberbullying.

Cyberbullying poses significant issues for corporations and security personnel due to its widespread and anonymous characteristics [3]. The anonymity offered by digital platforms gives cyberbullies the confidence to increase their behavior without worrying about being held accountable, worsening the pain done to victims. Cyberbullying, in contrast to traditional bullying, is covert, persistent, and causes its victims continuous distress [4, 5]. Previous studies have had trouble with sample size and factors, which made it harder to get a full picture of how cyberbullying works. To address these limitations, it is essential to utilize advanced analytical techniques such as sentiment analysis combined with transformer models like BERT, complemented by named entity recognition (NER) [6, 7]. This method has the potential to simplify the identification of cyberbullying and assist people in devising effective prevention strategies. As a conclusion, cyberbullying is a complicated problem that needs a diverse solution that combines new technologies with effective steps to make the internet a safer place and keep people safe. We can strive to establish a digital environment that is more inclusive and toward creating a safer and more inclusive digital environment for all.

2 Related Works

2.1 Base Lines

We utilize cutting-edge sentiment analysis techniques such as BERT-ADA to detect future patterns. When compared to other models such as BERT-base, DistilBERT, Twitter-roBERTa-base, and BERTweet, distinctions in sentiment classification become prominent. Tweets are classified as either positive or negative by some, while others classify them as positive, neutral, or negative sentiments. We consider negative tweets as potential indicators of cyberbullying, aiming to improve detection methods.

Ousted **WeWork** founder **Adam Neumann** lists his **Manhattan** penthouse for **\$37.5 million**
[organization] [person] [location] [monetary value]

Fig. 1 A tweet with an example of NER

Architecture Base: Devlin introduced two architectures, with the BERT-based one being the smallest at 110 million parameters compared to 340 million [8]. Utilizing the TextAttack package in the HuggingFace transformers library, we employ the BERT-base implementation for our baseline test. This implementation generates cyberbullying predictions based on tweet inputs.

DistilBERT: The average gain in inference speed as a result of this reduction is 60%. DistilBERT [9] is the best option for computationally efficient alternatives to BERT, even if it may perform somewhat worse than BERT—by 0.6% in IMDB categorization accuracy and by 2.7 points in the F1-score on the SQuAD question-answering challenge.

Gemma open-source LLM by Google: Gemma is an open model introduced by Google DeepMind in the article “Gemma: Open Models Based on Gemini” [10]. The resources used to build the Gemini models were used in the creation of Gemma. Gemini models are renowned for their robust performance in language understanding (Fig. 1), reasoning, and safety across academic benchmarks. Gemma models, offered in two sizes with 2 billion and 7 billion parameters, respectively, outperform open models of similar size on 61% of text-based tasks [11–13].

2.2 Aspect-Target Sentiment Classification

Proposed research on cyberbullying detection using transformer models such as BERT, Gemma, and DistilBERT, we extended our investigation to include aspect-target sentiment classification [14]. This expansion aimed to identify specific aspects and sentiments within cyberbullying-related content, offering deeper insights into online harassment incidents. We fine-tuned pretrained transformer models using annotated datasets containing tweets with aspect terms and sentiment labels, optimizing them for aspect detection and sentiment classification within cyberbullying-related text.

Figure 1 shows the sample tweet analysis using named entity recognition (NER) techniques, particularly with the Gemma model, we identified aspect terms, while sentiment analysis involved classifying sentiment polarity (positive, negative, neutral) surrounding each aspect term. Evaluation of our models included standard metrics like accuracy, precision, recall, and F1-score, alongside qualitative analysis to assess their ability to capture nuanced sentiment expressions and identify relevant aspects in cyberbullying-related content. Our results showed promising performance in Aspect-Target Sentiment Classification using Transformer models. With the use of

BERT, Gemma, and DistilBERT, valuable insights were obtained regarding important aspects and sentiment polarity of cyberbullying-related tweets, providing valuable insights for targeted intervention and support. There are a number of approaches to detecting cyberbully instances, including using natural language components such as insults and second-person language. These approaches demonstrate reasonable performance across multiple forums. Another method employs machine learning techniques, trained on labeled data from platforms with high bullying content, to automatically recognize cyberbullying language patterns. Researchers have noted the importance of addressing cyberbullying, particularly on social media sites where it has become rampant, sometimes resulting in tragic outcomes such as suicides. As an alternative to traditional sentiment analysis, aspect-target sentiment classification (ATSC) presents itself as an innovative approach utilizing advanced transformer architectures such as BERT-ADA and named entity recognition (NER) to detect cyberbullying more effectively [15–17]. The purpose of our study is to develop an effective approach to detect malicious messages using a combination of natural language processing (NLP) and artificial intelligence. By evaluating various machine learning algorithms, we are enhancing the capability of automatically recognizing and flagging online abusive content through the use of features such as Bag-of-Words (BoW) and term frequency-inverse document frequency (TFIDF). This contributes to efforts to combat cyberbullying on social media networks.

3 Proposed System

This study emphasizes the value of sentiment analysis in the identification of cyberbullying, paying particular attention to the features of the data and the variety of analytic techniques. Our goal is to improve cyberbullying detection on social media sites, especially Twitter, by leveraging sophisticated transformer models such as BERT and DistilBERT in conjunction with the specific Gemma model for named entity recognition (NER). Our objective is to accurately identify entities related to cyberbullying by extracting aspect-targets using NER and categorizing tweets into positive, negative, and neutral sentiments. Gemma is not like general-purpose models because it only does NER tasks. This makes it perfect for finding named things linked to cyberbullying, like personal names and offensive language. Beyond conventional classifiers and even general-purpose models like BERT and DistilBERT, we expect that including Gemma in our cyberbullying detection framework will increase accuracy and efficiency (Fig. 2).

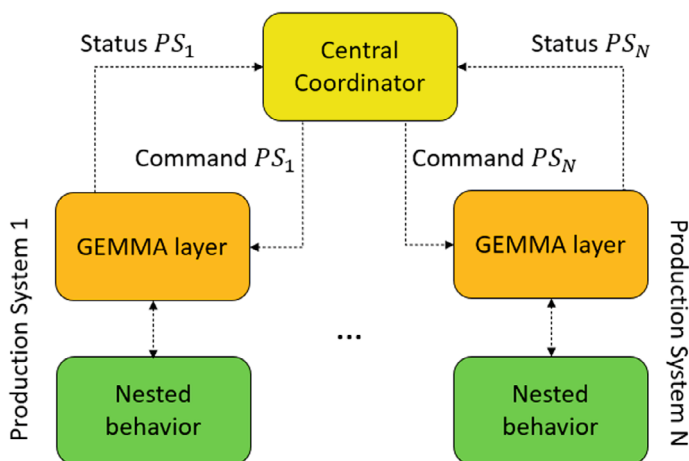


Fig. 2 GEMMA-based decentralized architecture

3.1 Proposed Methodologies

Get data from Twitter

It is necessary first to create a Twitter account, to get data from Twitter. One needs to submit a Developer Access request on the official Twitter website to access API credentials. The credentials have four important components: keys, secret keys of consumers, and access tokens. When the account is made, the Keys and Access Tokens tab gives you access to your passwords. Using Python scripts or applications, these keys function as authentication tokens to enable programmatic access to Twitter's data. Therefore, in order to retrieve data from Twitter, API credentials are necessary, enabling seamless interaction with Twitter's data.

Preprocessing

The next step is data preprocessing where we convert the extracted tweet data into a more comprehensible format. Frequently, real-time data are filled with imperfections, such as inconsistencies, errors, and incompleteness. As a result, it is essential to convert these data into a standardized format for future analysis. Converting all text to either lowercase or uppercase in order to uniformly use the case of letters. The numerical data is decoded into text or it can be eliminated. To guarantee consistency and readability, special characters should be removed, deleting unnecessary white spaces to improve readability and organize the content.

Stopwords Removal

The most common words in a language, such as "the," "a," "on," "is," and "all," are referred to as "stop words" in natural language processing (NLP). Usually having minimal semantic value, these terms are removed from texts before processing.

Tokenization

One of the most important steps in natural language processing (NLP) is tokenization, which makes it easier to divide a given text into smaller pieces called tokens. The main purpose of tokenization is to split the text into its individual parts for subsequent analysis and processing.

Classification of Cyberbullying

The transformer model is an efficient model for classifying cyberbullying tweet comments. Developers have various options for implementing the transformer model for text classification tasks, with one common approach involving separate encoder and decoder classes to handle these tasks independently. Our model will be constructed using a method that is comparatively simple and will be effectively applied to text classification. It is first necessary to define a few parameters, such as the dimensions of the embedding for each token, the attention heads and layers within the feed-forward network of the transformer. The model will be built by using a utility created previously for transformer blocks and a positional embedding class. The sentence will be represented by vectors in the input, which will be embedded and passed through a transformer block. Subsequently, we'll incorporate a global average pooling layer, a dropout layer, and a dense layer to generate the probabilities associated with the sentence classifications.

Classifier Comparison for Cyberbullying Detection

In addition to capturing bidirectional contextual information, BERT is computationally intensive and takes longer training time. Gemma model is tailored for named entity recognition (NER) tasks, potentially offering higher accuracy for detecting cyberbullying-related entities, but may require specialized training data. Figure 3 shows the architecture and training process of Gemini model. DistilBERT offers a balance between performance and efficiency, being smaller and more computationally efficient than BERT while retaining much of its performance, but may sacrifice some accuracy on complex tasks. Traditional classifiers like logistic regression and support vector machines (SVMs) offer simplicity and interpretability with faster training times but may struggle to capture complex patterns in high-dimensional data without extensive feature engineering.

4 Experiments

The dataset consists of tweets annotated for hate speech, offensive language, or other inappropriate content. It serves as a very good resource to train NLP models for sentiment analysis. The dataset has been split into three parts for training, validation, and testing so that the evaluation is unbiased. We performed these preprocessing steps: tokenization, text normalization, standardizing, decoding numerical data, eliminating special characters, clipping the white spaces, and stopwords removal.

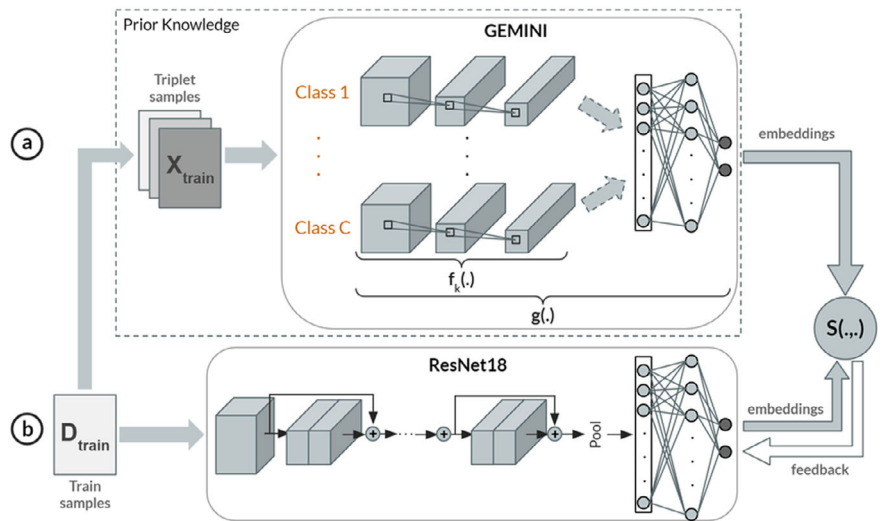


Fig. 3 The GEMINI’s low-dimensional embeddings guide the training

Model Training

A pretrained Model with the Configuration: 7 Billion parameters, Adam Optimizer with a learning rate of $2e^{-4}$ was fine-tuned for sentiment analysis to detect offensive language in tweets. The training was done for 25 epochs with stochastic gradient descent. The following evaluation metrics were used to assess the performance of the models: accuracy, precision, recall, and F1-score.

5 Results and Discussion

In comparing the performance of Distill BERT, BERT, and GEMMA models across two feature extraction methods, TF-IDF and word2vec, our findings reveal a clear trend of improvement in model effectiveness. Distill BERT, designed as a lighter and faster variant of BERT, achieved scores of 0.632 and 0.640 with TF-IDF and word2vec respectively, establishing a baseline for comparison. Because of its more complex structure and ability to capture minute language elements, the original BERT model performed better, achieving higher scores of 0.686 with TF-IDF and 0.7 with word2vec (Fig. 4).

GEMMA, however, achieved the highest scores with TF-IDF and word2vec, scoring 0.726 and 0.741, respectively. In addition, the advanced embedding techniques of GEMMA, by leveraging geometric insight into data structure, provide significant advantages over tradition models in processing and analyzing textual data, regardless of the method used for feature extraction. To detect cyberbullying,

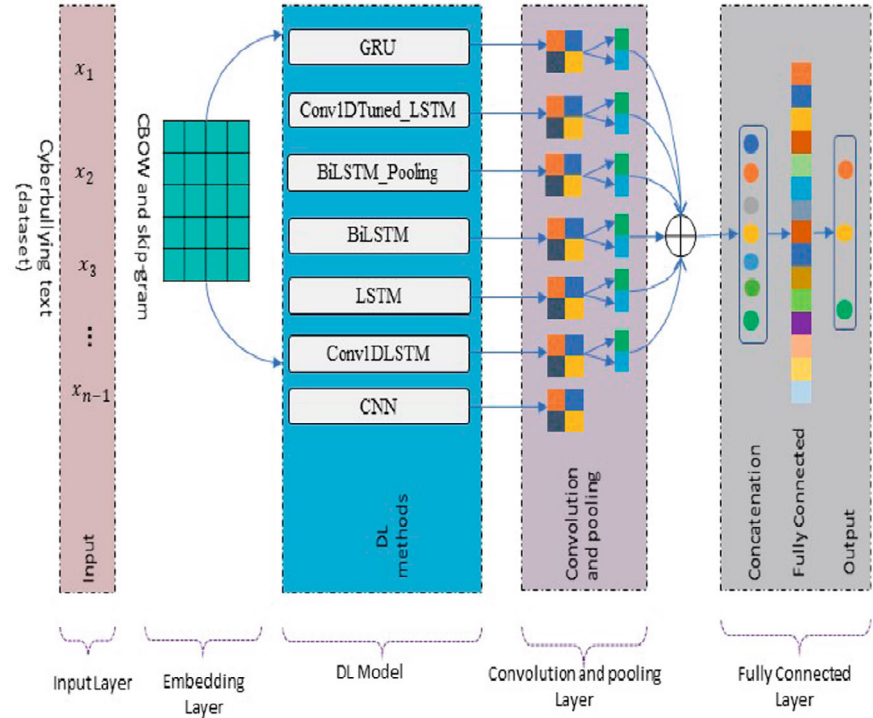


Fig. 4 Cyberbullying detection on Twitter

the Gemma model’s training configuration incorporates the selection of an appropriate model architecture that is optimized for named entity recognition (NER) tasks. A thorough data preparation process is necessary, including tokenization, preprocessing, and feature extraction. Cyberbullying categories are defined by a suitable labeling scheme, and an appropriate loss function is selected for training. Model training is enhanced by specifying an optimizer calculation, a batch size, the number of epochs, and a learning rate. Overall, careful selection and tuning of these parameters ensure effective training of the Gemma model for accurate cyberbullying detection in Twitter comments. Training loss and validation loss are critical metrics seasoned to assess the spectacle of the Gemma criterion during the training operation without copying someone else’s work. Training loss refers to the slip or discrepancy amidst the actual output and predicted result of the Gemma ideal during the training phase. It is calculated by collate the model’s predictions with the objectivity labels for the training value (Table 1).

The intention of training is to play down this loss, indicating that the model is learning to produce accurate predictions on training data. On the contrary, validation loss refers to the disparity or difference between the predictions made by the model and the actual labels provided in the validation dataset. During training, a portion of

Table 1 Outcomes on the tweeteval benchmark

	Precision (%)	Recall (%)	F1 (%)
Macro	89.2	75.8	64.8
Weighted	71.3	65.0	64.4
Binary	55.5	85.9	67.4

the data is set aside as the validation set, which is not used for training but is used to appraise the model’s performance on unnoticed data. The validation loss provides an estimate of how well the model generalizes to new, unseen data (Fig. 5).

Optimally, both training loss and validation loss should decrease as the Gemma model learns from the training data. During the training process, we can monitor the trends of training loss and validation loss and adjust the Gemma model accordingly to improve accuracy and generalization. Additionally, the precision, recall, and F1-scores were computed for both the classes (Table 2).

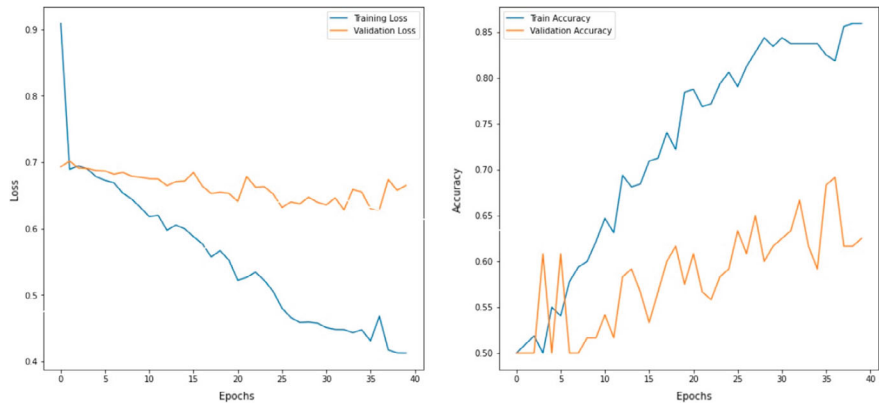


Fig. 5 Training and validation accuracy and loss graphs prior to parameter optimization

Table 2 Classification report

	Precision	Recall	F1-score
Class 0	0.95	0.98	0.96
Class 1	0.85	0.84	0.85
Class 2	0.88	0.86	0.87
Accuracy			0.89
Macro avg	0.89	0.89	0.89
Weighted avg	0.89	0.89	0.89

6 Conclusion

An innovative approach based on named entity recognition (NER) has been presented in this paper to detect cyberbullying in Twitter comments. The goal of this study was to address the challenges of identifying cyberbullying entities within online discourse by utilizing advanced natural language processing (NLP) models, including BERT, DistilBERT, and Gemma. We demonstrated the effectiveness of our approach by evaluating a dataset of Twitter comments to detect named entities associated with cyberbullying behaviors, including personal names, derogatory terms, and offensive statements. According to our evaluation, the Gemma model outperformed other classifiers, demonstrating its suitability to detect cyberbullying. The findings of our study contribute to a better understanding of cyberbullying dynamics in online environments, as well as the types of entities involved in cyberbullying incidents. Our goal is to create a safer and more inclusive online community through the use of NER techniques. Future research directions include exploring ensemble methods combining multiple NLP models, examining whether the approach can be applied to other social media platforms, and addressing ethical issues surrounding data privacy and bias mitigation.

Funding This work is supported by Deanship of Scientific research, Vice Presidency for Graduate Studies and Research, King Faisal University, Saudi Arabia (Grant Number: KFU241503).

References

1. YJ Foong, M Oussalah (2017) Cyberbullying system detection and analysis. In: Proceedings of European intelligence and security informatics conference. IEEE Computer Society
2. Reynolds K, Kontostathis A, Edwards L (2011) Using Machine Learning to Detect Cyberbullying. In: International conference on machine learning and applications. IEEE Computer Society
3. Sugandhi R, Pande A, Chawla S, Agrawal A, Bhagat H (2015) Methods for detection of cyberbullying: a survey. IEEE
4. Chen CK, Ramjee S, Wang J, Aspect-target sentiment classification for cyberbullying detection. CS224N natural language processing with deep learning
5. Islam MM, Uddin MA, Islam L (2021) Cyberbullying detection on social networks using machine learning approaches. IEEE Xplore
6. Ma D, Li S, Zhang X, Wang H (2017) Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893
7. He R, Lee WS, Ng HT, Dahlmeier D (2018) Effective attention modeling for aspect-level sentiment classification. In: Proceedings of the 27th international conference on computational linguistics, pp 1121–1131
8. Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
9. Zhao R, Zhou A, Mao K (2016) Automatic detection of cyberbullying on social networks based on bullying features. In: Proceedings of the 17th international conference on distributed computing and networking, pp 1–6
10. Rietzler A, Stabinger S, Opitz P, Engl S (2020) Adapt or get left behind: domain adaptation through Bert language model finetuning for aspect-target sentiment classification. In:

- Proceedings of The 12th language resources and evaluation conference, Marseille, France, pp 4933–4941
11. Founta A-M, Djouvas C, Chatzakou D, Leontiadis I, Blackburn J, Stringhini G, Vakali A, Sirivianos M, Kourtellis N (2018) Large-scale crowdsourcing and characterization of Twitter abusive behavior. In: 11th International conference on web and social media, ICWSM 2018. AAAI Press
 12. Nadeau D, Sekine S (2007) A survey of named entity recognition and classification. *Linguisticae Investigationes* 30(1):3–26
 13. Sang EF, De Meulder F (2003) Introduction to the CoNLL-2003 shared task: language-independent named entity recognition. arXiv preprint cs/0306050
 14. Xu H, Liu B, Shu L, Yu PS (2019) Bert post-training for review reading comprehension and aspect-based sentiment analysis. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics
 15. Asghar N (2016) Yelp dataset challenge: review rating prediction. arXiv preprint arXiv:1605.05362
 16. He R, McAuley J (2016) Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering. In: Proceedings of the 25th international conference on world wide web, WWW '16, pp 507–517
 17. Shah NA, Meenakshi K, Agarwal A, Sivasubramanian S (2021) Assessment of student attentiveness to e-learning by monitoring behavioural elements. In: 2021 International conference on computer communication and informatics (ICCCI), Coimbatore, India, pp 1–7. <https://doi.org/10.1109/ICCCI50826.2021.9402283>

Hybrid Deep Learning Mechanism for Brain Tumor Detection



Vahini Siruvoru, Nampally Vijay Kumar, Dasari Rini Ruha, Boorla Asritha, and Sahithi Polunati

Abstract A key medical diagnostic test is the detection of brain tumors, and prompt and accurate results are essential for efficient treatment planning. In order to develop a reliable convolution neural network (CNN) model for automating the diagnosis of brain tumors using magnetic resonance imaging (MRI) data, this research focuses on leveraging deep learning capabilities. The main objective is to classify brain tumors with high accuracy and efficiency using a transfer learning system based on well-known CNN architectures as VGG 16, INCEPT v3, and RESNET 50. The goal of this project is to create a successful constitutional network-based floral categorization method using transfer learning. This paper compares network initialization models to transfer learning models based on VGG 19 v3, inception 16v3, and ResNet50. The findings demonstrate that over-fitting and local optimal issues in deep convolution networks can be successfully avoided by transfer learning. The precision with which tumors are detected has obviously increased, and their robustness and generalization are far superior to that achieved by traditional methods.

V. Siruvoru (✉) · N. V. Kumar · D. R. Ruha (✉) · B. Asritha · S. Polunati
Department of Computer Science and Business System, B V Raju Institute of Technology,
Narsapur, Medak 502313, India
e-mail: vahini.siruvoru@gmail.com

D. R. Ruha
e-mail: 22211a3212@bvrit.ac.in

N. V. Kumar
e-mail: nampallyvijaykumar@gmail.com

B. Asritha
e-mail: 22211a3209@bvrit.ac.in

S. Polunati
e-mail: 22211a3245@bvrit.ac.in

1 Introduction

Brain tumor detection is essential for medical diagnosis and treatment planning because it establishes the best course of action and enhances patient outcomes. Brain tumor diagnoses were formerly made by radiologists manually analyzing data from magnetic resonance imaging (MRI). But this is a time-consuming, subjective, and prone to error procedure carried out by humans. Convolutional neural networks (CNNs), a type of deep learning approach, have become the tool of choice for researchers due to the pressing demand for accurate and automated solutions in the field of medical imaging.

The intention of this product is to analyze how accurate CNN models built on deep learning perform when it is used to canvas brain tumors from MRI data. By utilizing retrained prototype and neural network capacity, the goal is to create an automated formulation for brain tumor finding that improves tumor classification effectiveness and accuracy. We want to employ state-of-the-art CNN architectures, such as RESNET-50, InceptionV3, and VGG-16, to gain the quality of brain tumor recognition.

This paper purpose to enquire the effectivity of CNN models supported on deep learning for brain tumor diagnosis based on MRI information. Our content is to create a machine-controlled know-how for brain tumor identification that increases tumor classification efficaciousness and accuracy by using retrained models and neural network capacity. RESNET-50, InceptionV3, and VGG-16 are examples of cutting edge CNN edifice that we want to use to amend brain tumor determination accuracy.

To alter our approach, we should use transfer learning with pretrained models to make use of learned correspond from immense picture datasets. This plan of attack has incontestable efficacy in various computer vision labor and has promise for rising the public presentation of brain tumor detection models. Our activity affect in optimizing these pretrained models to modify the tumor identification accuracy using a set of MRI images from patients with and without brain tumors.

This work investigates the prospect of using deep learning to automatically discover brain malignancies in an effort to heighten medical imaging processes. The research's collection might revolutionize the way health care is delivered by giving medical staff high-fidelity and useful tools for speedily characteristic patients.

In the segment that follow, we will face the applicable studies on deep learning methods for brain tumor determination. We will describe our epistemology, the data we used, the CNN architecture we exploited, and the specifics of the transfer learning and fine-tuning processes. Pursuing an extensive debate, the experiment's results will be closely evaluated, compiled, and finalized.

2 Review of Literature

Much work has been done lately to know and categorize brain tumors using deep learning algorithms. Researchers are investigating the use of convolutional neural networks (neural CNNs) and machine learning algorithms to find brain malignancies in magnetic resonance imaging (MRI) data in order to make over precise and efficient process. Siar and Teshnehlab [1] Neural networks and machine learning conceptualization have been projected as part of a paradigm for brain tumor sensing. Their method acting has demonstrated possibility in distinguishing brain tumors and exhibit the welfare of merging machine learning and deep learning acting. Deep neural networks have proved importance in the cell division of brain tumors, as demonstrated by Havaei et al. [2].

They have exhibited via their research that gliomas may be dependably eminent from MRI images by employing deeper CNN architectures. This lays a solidified base to foster experimentation into the sensing of brain malignancies. Sajid et al. [3] implementing deep learning techniques to the segment and detect the brain tumors is the main goal of this work. Their activity implies how the improvement of an automated tumor identification technique has been made achievable by deeper CNN models' capacity to precisely establish and segment brain tumors based on magnetic resonance imaging. The Hemanth et al. [4] study inspect a brain tumor recognition technique based on machine learning.

They have manifested that it is feasible to attain an ideal algorithmic fusion in title of patient cataloging, and their procedure has shown to be more true in that aspect. A system multiscale 3D CNN model with fully linked conditional random fields (CRF) was conferred by Kamnitsas et al. [5] for accurate brain lesion segmentation; their work has illustrated the need of using CRFs in deeper CNN architectures to improvise lesion segmentation. The main goal of the 2019 study by Noori et al. [6] was brain tumor segmentation using an attention-guided 2D UNET model. Their method has advanced in tumor segmentation accuracy by integrating attention processes.

This exhibits how crucial this path is to improving the consequence of tumor segmentation. Kang et al. [7] conjunctive machine learning procedures with a composite of the deepest characteristics to assess brain tumor classification. Their process has shown that MRI data combined with deep features and machine learning techniques may be used to reliably identify cancers. A research was conveyed out by Saleh and colleagues.

Saleh et al. [8] looked at the request of deep learning techniques to brain tumor classification. Their method revealed the use of deep learning models to reliably check brain tumors and allowed for the growth of an automated diagnostics system. Islam et al. [9] in contrary to others (2020). A model for brain tumor segmentation and survival reasoning was rumored, based on a 3D focus UNet. Especially, their work determined that the precision of UNet topologies and attending processes was crucial for precise tumor segmentation and patient survival prevision. The goal of Chelghoum et al. [10] was on transfer learning for brain tumor classification from MRI data using CNN architectures. They applied and restricted retrained CNN

models to closely diagnose brain tumors in a way to assess the effectiveness of transfer education.

In 2022, Modiya and Vahora [11] investigated the use of a transfer in the diagnosis of brain cancers. The request of dimensional decrease in transfer learning for brain tumor diagnosis was designed by Modiya and Vahora [11]. In doing so, they showed how dimensional reduction and learning approaches may be used to reliably check brain cancers from MRI data. Lastly, a sizable body of work on the use of deep learning algorithms for brain tumor diagnostics has been made visible by our literature investigation. The accuracy of cancer detection, segmentation, and classification has been demonstrated to enhance with the combination of machine learning algorithms, transfer learning, attention processes, and computational neural networks. This advancement is assisting in the development of automated methods for brain tumor detection and planning. Recent advancements in brain tumor detection have explored a variety of innovative machine learning and deep learning approaches, each offering unique strengths. Kumar et al. [12] developed a method using a Support Vector Machine (SVM) classifier that effectively identifies brain tumors in MRI scans by selecting the most relevant features. Building on previous research, Soltaninejad et al. [13] came up with a creative solution by blending random forests with fully convolutional networks. This hybrid approach not only made tumor segmentation more precise but also offered valuable predictions about patient survival, making it a meaningful step forward. Sonavane and Sonar [14] utilized an interesting method by applying the Adaboost classifier, a model that consolidates multiple simpler ones to collaborate as a single, powerful one. This improved detection as well as segmentation of brain tumors, rendering their method more trustworthy. Tammina [15] took advantage of transfer learning using the VGG-16 model, which proved to be quite effective—especially in situations where training data was limited. This approach made image classification more efficient, eliminating the need to build a model from the ground up. Likewise, Mascaredhas and Agarwal [16] dug closely into some of the most often used CNN designs: VGG16, VGG19, and ResNet50. By means of their comparison, they discovered that deeper models—especially ResNet50—tended to perform better, so providing more accuracy in image classification challenges. Wrapping up this wave of research, Wang et al. [17] applied the Inception-v3 model through transfer learning to classify pulmonary images, and their findings highlight the model's strong ability to extract detailed features—an advantage that could be extended to brain tumor diagnostics in future studies.

3 Existing Methods

3.1 *Machine Learning Models*

Among the earliest methods were support vector machines, or SVMs. MRI scans of the brain have shown that SVMs are quite good at binary classification tasks; in fact, they have been used to distinguish between tumor and nontumor regions. When there is a distinct difference between the two classes, they function effectively. Random forest is an ensemble learning technique. Applications for the identification of brain tumors have also been found. Random forest is an efficient method for handling large dimensionality data, which is a typical characteristic of medical pictures, by combining forecasts from several decision trees. AdaBoost and Gradient Boosting are examples of ensemble techniques that have been used to combine multiple poor classifiers into a better one. Two important challenges in medical imaging analysis are the reduction of overcompensation and the enhancement of generalization, which these strategies help with. Hybrid models are becoming increasingly prevalent as machine learning advances. To benefit from the best of both worlds, they are fusing deep learning with conventional machine learning approaches.

3.2 *CNN Models*

In this paper, convolution neural network models VGG16, Resnet50, and Inception-v3 were used. By deepening the network layer, the VGG model aims to reduce the number of models and increase network performance by transforming larger convolution layers into a range of volume tiny embedded convolution kernels. The fundamental idea is to extract feature information of various sizes and diverse information fusion splicing properties by first handling the input matrix with different sizes and then filtering the Inception module. The data are then cascaded to the following layer via the fundamental module of the Inception framework.

The shortcut connection technique, which the residual network joined, is used to transfer the input cross. The residual network joined the shortcut connection technology to transmit the input cross-layer and add the convolution processing result, so making the posterior layer output input and the input mapping superimposition. The objective is to optimize residual functions of the network.

4 Proposed Methodology

4.1 *Transfer Learning*

Transfer learning is a well-liked deep learning method that entails adapting neural network models that have been pre-trained for one job to a different but related one. Applying knowledge and feature representations from large-scale datasets that have been demonstrated to be successful in a variety of computer vision tasks to the particular issue of brain tumor diagnosis from MRI images is the concept behind transfer learning. Both the amount of data required and the training time can be greatly decreased by using existing models. This study applies transfer learning to three popular pretrained CNN models: RES50NET, InceptionV3, and VGG-16.

4.2 *CNN Models Used*

- VGG-16: Simonyan and Zisserman proposed a deeper CNN architecture with 16 layers in 2014. It is known for its compactness and uniformity, as it is made up of three 3×3 convolutional layers and the Maximum Pooling layer. We will use pretrained weights from the VGG-16 model trained on the ImageNet dataset as a starting point in our mission to detect brain tumors.
- Inception V3: Szegedy et al. introduced an inception module concept designed to take advantage of multiscale features in 2015. It has proven to be effective in a variety of image recognition tasks. For transfer learning to brain tumor detection, we use the pretrained weights of InceptionV3, which were trained on ImageNet.
- RESNET50: He et al. created this deep CNN model with 50 layers in 2015. The skip connection or “shortcut” architecture, which allows the model to perform exceptionally well even at greater depths, has been identified as an important design feature. Pretrained RESNET50 weights are being used to learn how to transfer information through ImageNet for our brain tumor detection task (Figs. 1 and 2).

4.3 *Dataset*

To train and evaluate, this dataset is used. There are various MRI scans of patients with or without brain tumors included in this study. In order to ensure model generalization, the dataset has been carefully curated so as to include a balanced distribution of tumor and nontumor cases. In order to achieve a uniform resolution and format during the training process, each MRI scan is preprocessed.

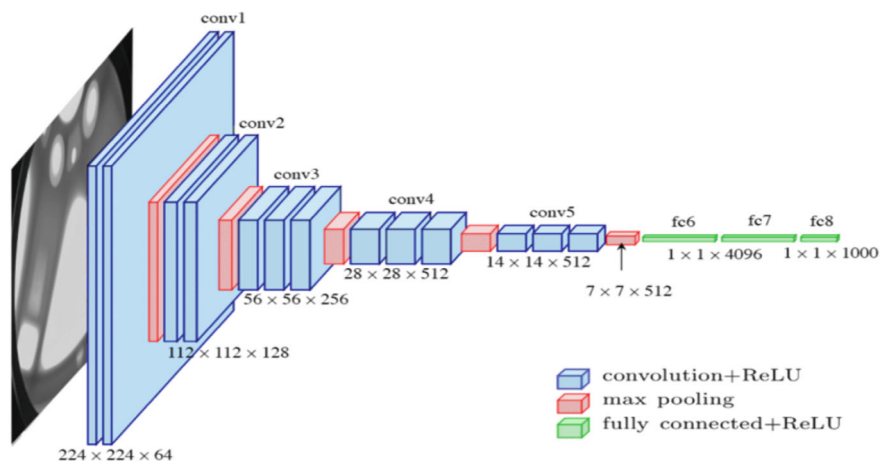


Fig. 1 Classic CNN model [cited from: <https://thedatafrog.com/en/articles/image-recognition-transfer-learning/>]

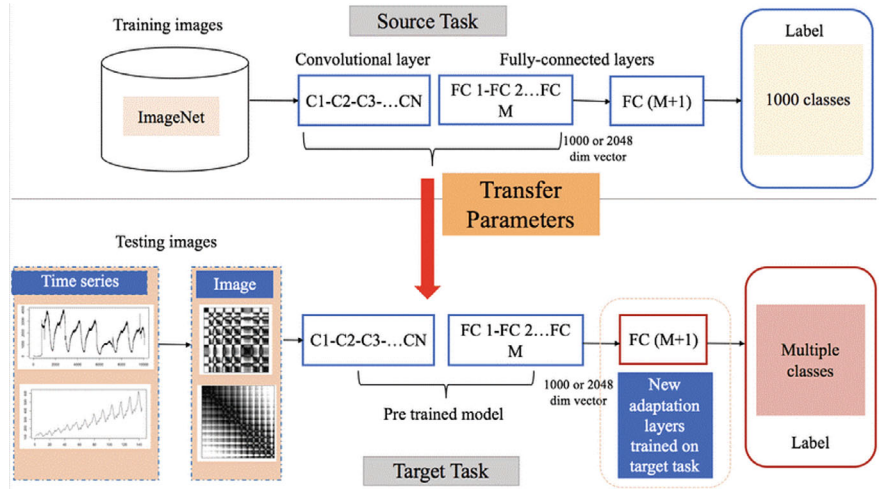


Fig. 2 Transfer learning model

4.4 Preprocessing and Training

Before feeding the MRI scans into CNN models, a number of preparatory steps must be completed. We are adjusting the pixel intensity for each image to ensure consistency across the database. The images will then be resized to a fixed input size suitable for our chosen CNN architecture. Furthermore, techniques for data enlargement were used to increase the diversity of training data, such as random

rotation, flipping, and zooming. The training procedure entails configuring pretrained CNN models for a brain tumor detection dataset. We use an Adam optimizer to define loss functions and binary cross entropy in order to improve the model's performance. A performance monitor is used to monitor the model's performance during training and to avoid over fitting. There are two components to a dataset: validation and learning. To get the best results, grid search techniques were used to precisely tweak the model's hyperparameters.

5 Experimental Results

Model	Resnet50-transfer	InceptionV3-transfer	VGG16-transfer
Accuracy	0.80	0.60	0.75
Precision	0.80	0.56	0.75
Recall	0.80	1.00	0.60
F1-Score	0.80	0.71	0.67

We carried out a comprehensive examination of brain tumor identification employing a transmission learning strategy with three well-known CNN architectures: RES50NET, VGG-16, and Inception v3 during the experimental part of our work. The acquired classification accuracies, which came about after a great deal of training and assessment, show how well these models work at differentiating brain tumor patterns from MRI data.

With an incredible precision of 0.80, the transmission model exhibits a great degree of capableness to extract nuanced message that distinguishes tumor locations. With an astounding accuracy of 0.75, the VGG-16 transfer prototype tells that its learnt representations have been effectively theoretical to the job. With a precision of 0.60, the Inception v3 transfer model suggests that it may show a role in tumor development. Moreover, our trials show that RESNET50 is the most faithful design, demonstrating the value of transfer learning in enhancing CNN models for brain tumor detection. These effects should be considered preliminary, since there is room for more tuning and betterment of their diagnostic power.

6 Conclusion

In summary, this work contributes a more effective and user-friendly method for solution features from deep convolutional neural network (CNN) models than manual methods. The neural network model converges more quickly and executes better than

at random started models because to transfer learning. It also presents stronger generalization and higher elasticity. This path might improve the quality of diagnosing brain tumors.

Work Limitations and Future Prospects: Although this study is encouraging, there are a number of issues with it. More study is required to see whether it can be used to detect brain tumors using modalities other than MRI images. Subsequent investigations might expand the methodology's reach to include a wider range of picture kinds and datasets, hence investigating new approaches to improve feature interpretability and optimize the model.

Notwithstanding these drawbacks, this work offers a strong basis for improving medical imaging procedures and has a great deal of therapeutic benefit potential.

References

1. Siar M, Teshnehlal M (2019) Brain tumor detection using deep neural network and machine learning algorithm. In: 2019 9th International conference on computer and knowledge engineering (ICCKE). IEEE, pp 288–293
2. Havaei M et al (2017) Brain tumor segmentation with deep neural networks. *Med Image Anal* 35:18–31
3. Sajid S, Hussain S, Sarwar A (2019) Brain tumor detection and segmentation in MR images using deep learning. *Arab J Sci Eng* 44:9249–9261
4. Hemanth G, Janardhan M, Sujihelen L (2019) Design and implementing brain tumor detection using machine learning approach. In: 2019 3rd International conference on trends in electronics and informatics (ICOEI). IEEE, pp 1289–1294
5. Kamnitsas K et al (2017) Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 36:61–78
6. Noori M, Bahri A, Mohammadi K (2019) Attention-guided version of 2D UNet for automatic brain tumor segmentation. In: 2019 9th International conference on computer and knowledge engineering (ICCKE). IEEE, pp 269–275
7. Kang J, Ullah Z, Gwak J (2021) MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors* 21(6):2222
8. Saleh A, Sukaik R, Abu-Naser SS (2020) Brain tumor classification using deep learning. In: 2020 International conference on assistive and rehabilitation technologies (iCareTech). IEEE, pp 131–136
9. Islam M et al (2020) Brain tumor segmentation and survival prediction using 3D attention UNet. In: *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 5th international workshop, BrainLes 2019*. Held in Conjunction with MICCAI 2019, Shenzhen, China, 17 Oct 2019. Revised Selected Papers, Part I, vol 5. Springer International Publishing, pp 346–356
10. Chelghoum R et al (2020) Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images. In: *Artificial intelligence applications and innovations: 16th IFIP WG 12.5 international conference, AIAI 2020, Neos Marmaras, Greece, June 5–7 2020. Proceedings, Part I*, vol 16. Springer International Publishing, pp 129–139
11. Modiya P, Vahora S (2022) Brain tumor detection using transfer learning with dimensionality reduction method. *Int J Intell Syst Appl Eng* 10(2):201–206
12. Kumar TS, Rashmi K, Ramadoss S, Sandhya LK, Sangeetha TJ (2017) Brain tumor detection using SVM classifier. In: 2017 Third international conference on sensing, signal processing and security (ICSSS), Chennai, India, pp 318–323. <https://doi.org/10.1109/SSPS.2017.8071613>
13. Soltaninejad M, Zhang L, Lambrou T, Yang G, Allinson N, Ye X (2018) MRI brain tumor segmentation and patient survival prediction using random forests and fully convolutional

- networks. In: Crimi A, Bakas S, Kuijf H, Menze B, Reyes M (eds) *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries*. BrainLes 2017. Lecture notes in computer science, vol 10670. Springer, Cham
14. Sonavane R, Sonar P (2016) Classification and segmentation of brain tumor using Adaboost classifier. In: 2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC), Jalgaon, India, pp 396–403. <https://doi.org/10.1109/ICGTSPICC.2016.7955334>
 15. Tammina S (2019) Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *Int J Sci Res Publ (IJSRP)* 9(10):143–150
 16. Mascarenhas S, Agarwal M (2021) A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In: 2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON), vol 1. IEEE
 17. Wang C et al (2019) Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access* 7:146533–146541. <https://doi.org/10.1109/ACCESS.2019.2946000>
 18. Khan MA et al (2019) Brain tumor detection and classification: a framework of marker-based watershed algorithm and multilevel priority features selection. *Microsc Res Tech* 82(6):909–922

Spot the Fake Legs: Synthetic Generating Medical Images Using Latent Diffusion Model (SGMs)



Sura Ahmed Abd Alasidi , Khawla Hussein Ali ,
and Zaid Ameen Abduljabbar 

Abstract Medical image synthesis is gaining popularity using generative models. Neural models of diffusion have recently shown promise in producing remarkably lifelike representations of objects. However, their ability to provide medical images has yet to be investigated. We investigate neural diffusion models' capacity for creating synthetic medical images. First, we create images of legs from an input word prompt using a pretrained DALLE3 model. Second, we develop synthetic images and use 300 leg images to train a stable diffusion model. Through a qualitative study, we assess the synthetic image data by having two impartial radiologists classify randomly selected specimens derived from the created data as genuine, fraudulent, or unclear. Results show that features that are otherwise highly distinctive to an image created using the diffusion model can be translated. We have published the synthetic images at <https://www.kaggle.com/datasets/suraahmed56/computer-vision-medical-images>.

Keywords Medical Imaging · Latent Diffusion Model · Computer Vision

1 Introduction

A new era of synthetic image synthesis has been brought about by recent advances in AI and machine learning, with diffusion models marking a breakthrough [1–3]. By conditioning the generation process on specified inputs, such as written descriptions, other photos, or specific qualities, these models can produce images that are frequently identical to genuine ones [4–7]. Synthetic leg images, among other body

S. A. A. Alasidi · K. H. Ali (✉) · Z. A. Abduljabbar

Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Basrah, Iraq

e-mail: Khawla.ali@uobasrah.edu.iq

S. A. A. Alasidi

e-mail: pgs.sura.ahmed@uobasrah.edu.iq

Z. A. Abduljabbar

e-mail: zaid.ameen@uobasrah.edu.iq

parts, can be generated for several uses in the medical domain, such as diversifying the dataset for machine learning model training, producing content for medical education, and aiding in treatment planning [8–11].

“Detecting the False Legs: Synthetic Generation of Medical Images Using the Latent Diffusion Model” is a groundbreaking exploration of the nexus between medical imaging and artificial intelligence [12]. This work examines the use of sophisticated latent diffusion models (LDMs) to generate remarkably lifelike synthetic medical images, with a particular emphasis on radiography of the lower limbs [13]. The capacity to create synthetic yet anatomically accurate images opens new possibilities for medical training, diagnostic development, and AI model validation without jeopardizing the anonymity of patients, especially in light of the growing privacy concerns and lack of publicly accessible medical datasets for research [14].

With the ability to manage the image production process more effectively and produce images of higher quality, LDMs represent a significant progression from classic generative adversarial networks (GANs) in artificial image generation [15]. In addition to demonstrating the technical viability of using LDMs in medical imaging, this study investigates the ethical ramifications and possible uses, which include developing reliable diagnostic programs that can be trained on an almost infinite dataset of synthetic images and serving as educational resources for medical students, discussion regarding artificial intelligence’s role in health care [16]. It asks readers to think about the harmony between creativity and morality and the possibility of democratizing medicine, research, and medical care in the coming years as our culture becomes more digital [17]. Put simply, a diffusion model is a parameterized Markov chain that has been trained by variational inference [18]. Diffusion is used to add noise to the data so that the transition can be learned [19]. The diffusion model, in theory, adds Gaussian noise to the input data to make it noisy, then reverses the noise to recover the data distribution [20–22]. The model can produce meaningful data from random noise input once it has learned the distribution [23]. Diffusion models, then, convert a latently encoded model of images into a more important representation [24]. Diffusion models are comparable to denoising encoders in this sense [25, 26]. The whole method, as depicted in Fig. 1, may be summed up as a phenomenon with two stages: the forward pass, which is the conversion of the data distribution, and the reverse pass, an assigned text prompt [27–29]. The reverse pass is the reversal between the data distribution and the noise distribution (X_T to X_1), and the change is concerned with the data transmission to noise (X_1 to X_T). A diffusion model’s training requires knowledge of the inverting procedure, or $p(x_t - 1|x_t)$. For both the forward and reverse training stages, a neural network can be used to implement the diffusion model [30, 31]. However, the input and output dimensions of the design must match.

In contrast to earlier attempts, which primarily employed diffusion models to create capacity for unconditional data production, guided-diffusion models have been used in more recent research to show that conditioned generation occurs [32]. Through the use of the supplied text or image’s context, these studies have shown how to generate photo-realistic images [33]. The current use cases for diffusion

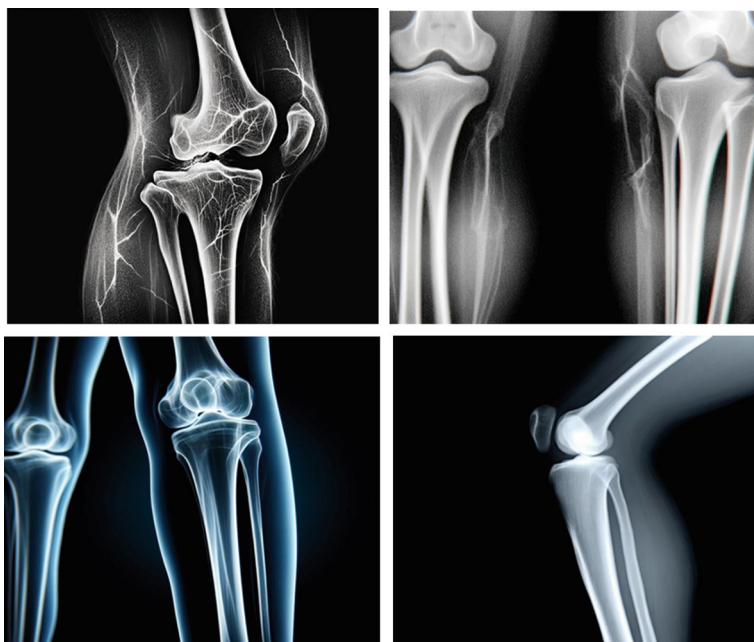


Fig. 1 Samples chosen from our suitable image model

models include text-to-image systems, which produce images based on an assigned text prompt [34].

We create the images utilizing the constant diffusion model and the DALLE2 model, and then show them to two radiologists for their opinions. After that, we provide an overview of the clinicians' comments and list some of the difficulties with adding medical image synthesis to the neural diffusion model.

The remaining text is arranged as follows: We outline our work process in Sect. 2. The findings of the leg image generation are displayed in Sect. 3, and the limitations of the approach are outlined. Additional information regarding the outcomes is provided in Sect. 4. Section 5 contains the paper's ultimate conclusion.

2 Literature View

I “Spot the Fake Legs” studies the creation of synthetic medical pictures using latent diffusion models (LDMs) and places itself within a rich context of previous studies and technological developments. This section highlights the evolution of methodologies and the growing uses of artificial intelligence in medical imaging by outlining important studies and advancements that have prepared the way for this ground-breaking work.

Singh et al. [35] in this study, we address this challenge by introducing a novel guided image synthesis framework, which depicts the final image as the resolution of an optimization problem with constraints. Demonstrate that although it is not possible to compute the exact solution to the optimization, it is possible to acquire an approximation solution that requires only one pass of the reverse diffusion process. Additionally, we demonstrate that the user can affect the semantics of various painted patches without the need for conditional training or fine-tuning by constructing a simple cross-attention-based correspondence between the input text tokens and the user's stroke painting. The suggested approach beats the previous state of the art by more than 85.32% on the overall user satisfaction levels, according to the results of human user research.

Dorjsembe et al. [36]. Introducing Med-DDPM, a diffusion model created specially to address data shortages and privacy concerns in semantic 3D medical image synthesis. Semantic conditioning is incorporated into Med-DDPM, which makes it novel and allows for exact control during the picture production process. Med-DDPM yields findings that are equivalent to those of cutting-edge augmentation methods, as demonstrated by compare-active analysis, highlighting the potential of the tool as a data augmentation tool to raise model correctness. In conclusion, Med-DDPM sets the standard for 3D semantic medical image synthesis by providing anatomically coherent, high-quality images. The integration of semantic conditioning with Med-DDPM further demonstrates the model's capacity to address privacy concerns and data scarcity difficulties, and this integration holds promise for image anonymization in the biomedical imaging area.

Frangi et al. [37] The Special Issue on Simulation and Synthesis in Medical Imaging is presented in this editorial. Define unclear words related while working on an editorial, to medical imaging simulation and synthesis. Additionally, it touches on the synergistic significance of the driven-by-data phenomenological (phenomenological) and hypothesis-driven mechanistic models of medical picture production. At last, we provide an overview of the 12 articles contained in this volume, which address medical image production from a mechanical five and phenomenological seven perspective. Applications in cardiology, retinopathy, histology, neurosciences, and oncology are covered in this extensive collection of papers. All common diagnostic medical imaging methods are also included. It brings the commentary to a close with a personal assessment of the topic, highlighting some current issues and potential directions for future research.

Akbar et al. [38] They are creating synthetic images of brain MRIs and chest X-rays, calculating the correlation between the synthetic images and all training images, and using BRATS20, BRATS21, and a chest X-ray pneumonia dataset to train Style-GAN and diffusion models. Our results show that diffusion models had a greater memory recall rate for the training images compared to Style GAN, especially for small datasets and when using 2D slices from 3D volumes. If disseminating synthetic photos is the ultimate goal, researchers should exercise caution when utilizing diffusion models for medical imaging.

Wu et al. [39] Using text-guided picture synthesis, this approach expands on the pretrained diffusion model to generate perceptual data. We demonstrate how a

decoder module can efficiently decode the diffusion model's extensive latent code into precise perception annotations. An endlessly large annotated dataset can be created by training the decoder with fewer than 1% (about 100 images) of manually labeled images. Subsequently, these synthetic data sets might be employed to train diverse perception models for subsequent tasks. To illustrate the efficacy of the proposed approach, offer datasets with rich, dense pixel-wise labels for a range of downstream tasks, such as instance segmentation, semantic segmentation, and depth estimation. Specifically, it achieves (1) cutting-edge results on instance and semantic segmentation and (2) a noticeably more robust study.

These attached papers offer insightful information about image synthesis and deep learning methods in medical x-ray imaging. These investigations provide valuable insights that researchers can utilize to advance the diffusion models for Leg (x-ray) image synthesis and tackle particular issues in the field peculiarities.

3 Methodology

In this paper, we designed two experiments to produce X-ray-like synthetic legs images. In the first experiment, we generated graphics based on the input text by using the Open AI DALLE2 API⁴. The DALLE2 model has garnered significant attention lately due to its capacity to produce remarkably lifelike representations of objects based on a given input text. We produced several X-ray images of the legs using the API. Two specialists with radiology training were then shown a randomly chosen selection of the generated images. We requested two major duties from the radiologists. Initially, we requested that they classify each image as authentic, fraudulent, or unclear based on their impression of understanding. Second, we requested that they provide a quick explanation of any relevant details regarding the diagnosis or state of their legs (e.g., normal legs, badly injured legs, lungs afflicted by pneumonia, etc.). The labels of the photographs were unknown to the psychiatrists beforehand. Every picture we showed the radiologists was artificial. The radiologists worked individually and were strangers to one another. Between the two physicians, one was experienced with deep generative models and artificial intelligence, while the other was new to these concepts. We applied the steady diffusion model in the second experiment. Our stable diffusion model was trained with 150 X-ray images from Kaggle we used a data set of 300 images (normal and abnormal), which were divided into 1 such that 150 normal images are used and the other 150 abnormal images in order to prevent excessive proportionality. X-rays of abnormal images consisting of (50 images of bone injuries and 50 images of tibia fractures) are also shown, (left leg and 50 pictures of osteoporosis in the leg). The images were scaled to 256×256 pixels. Other preprocessing was not carried out. Using the X-ray images, we trained a robust diffusion model on a server equipped with an NVIDIA Quadro RTX 8000 GPU and 48 GB of RAM as shown Fig. 2.

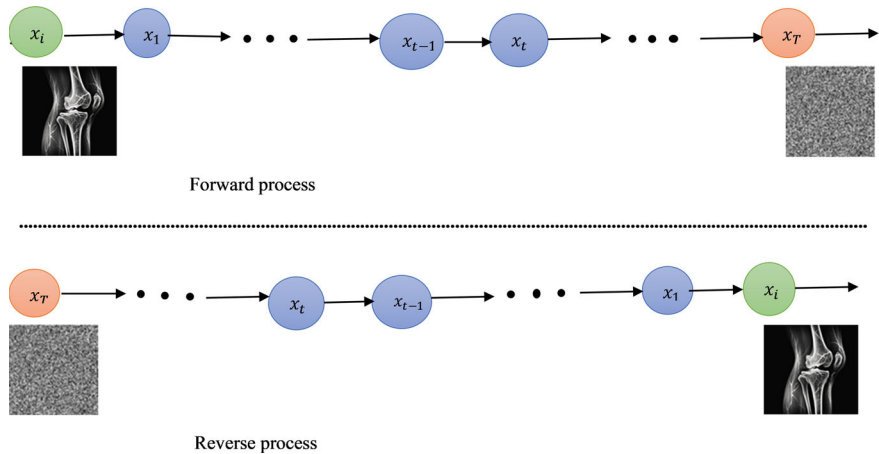


Fig. 2 Forward and backward passes for training diffusion models

4 Result

We produced 150 images in total using the DALLE2 API. The fake photos are now live on Kaggle. The tokens at our disposal are the sole restriction on the quantity of generated photographs. Figures 2 and 3 display sample X-ray scans, respectively. Radiologist A determined that 14 of the 50 photos we showed the radiologists were accurate, while the remaining four were phony. A radiologist evaluated two X-ray scans as unclear. Only 10 X-ray images were recognized as genuine by the second physician (radiologist B), who classified all other pictures as false.

Concord among radiologists: Just three scans were identified as genuine by both specialists. Likewise, in a similar vein, both radiologists certified five X-ray images as authentic. It needed to be made clear which X-ray images the radiologist’s thought were which.

In task 2, the radiologists made some intriguing remarks when we asked them to give a quick explanation of what the photos may show.

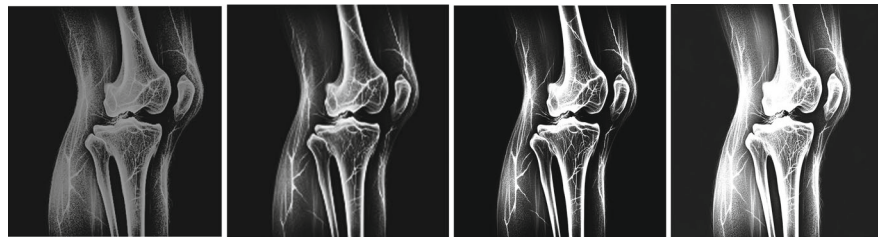


Fig. 3 Examples of X-ray images of the legs produced using the diffusion models

Table 1 Representative quotes via radiologists (in no particular order)

Image modality	Remarks ^a
X-ray	Bone injuries
	Left leg bone fracture
	Leg osteoarthritis

^a The comments don’t suggest a firm choice

Table 1 includes, for example, a compilation of some descriptions. These explanations unequivocally demonstrate that the model produced easily readable features. Issues and that particular image included depictions that looked like real X-rays.

5 Discussion

The radiologists instantly recognized some of the produced images as false because they lacked realistic features. These images were referred to as:

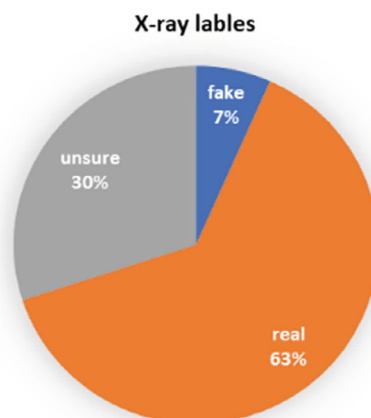
As well as displaying remarkable exposure or having a unique bone look, it was also simple to identify legs and outlines that were drawn rather than photographed. The fact that bone damage does not occur in actual X-ray imaging is a crucial finding regarding phony images. At least a couple of radiologists labeled several of the sample images as fraudulent, as shown in Fig. 5.

As seen in Fig. 4, since several of the images generated by the pre-trained model conspicuously lacked true image attributes, radiologists were able to quickly identify them as fraudulent. These pictures were described as having an odd look and a weird bone aspect. Figure 5 summarizes the radiologist’s evaluation.



Fig. 4 Samples of synthetic leg X-ray images (left two images) that were found to be fraudulent by at least two radiologists

Fig. 5 Evaluation by radiologist



6 Limitations and Future Study Path

Diffusion models have been found to have difficulty producing details in complex scenes [40–42]. Therefore, super-resolution or noise adaptation techniques would need to be used in addition to the generation of complicated medical images [43–46]. Like many other AI models, diffusion model training is susceptible to bias in the dataset. Examples of bias include noise in the data or an unequal medical condition in the X-ray input image. The bias will, therefore, also be present in the artificial data produced by using this kind of diffusion model. Eventually, the prejudice might become a cascade behavior and be further amplified if the collected data are made available to the public and used for further model training [47–50]. Since the model has essentially been utilized as a black-box model, it could be more explainable. The acknowledged limitations of the work include the diversity of the dataset used to train the LDM and the model's current inability to effectively imitate various complex aspects of real medical images. These drawbacks highlight the necessity of further improving the technology and its uses.

Future studies should concentrate on improving the variety and authenticity of artificial imagery, augmenting the portrayal of diverse medical conditions and body parts, and examining how synthetic images affect clinical judgment. To make sure that this exciting technology advances medicine in a way that is both responsible and equitable, more research into the moral application of synthetic medical images is also necessary.

Further research on this subject may result in ground-breaking developments in medical imaging technology.

7 Conclusion

In this work, we have shown the possibility of neural diffusion models in leg X-ray image generation. Even though the radiologists saw a lot of images as phony, only a small percentage of images were still marked as genuine. The radiologists' classification indicates that the produced X-ray images closely resemble the real ones. On the photos themselves, though, it was simple to identify fake images. Through qualitative examination of the generated images, we showed that neural diffusion models have great potential for learning complex representations of medical imagery, while diffusion models outperform GANs-based techniques in terms of synthesizing natural.

Ultimately, "Spot the Fake Legs: Synthetic Generating Medical Images Using Latent Diffusion Model" is an essential advancement in healthcare imaging and artificial intelligence. Our research demonstrates the potential of LDMs to produce artificial medical images and the broader ramifications of this technology for healthcare in the future. We are at the beginning of a new era in medical technology, one in which the distinction between the artificial and the real is becoming more and hazier, creating a wealth of opportunities for medical study, teaching, and practice. This is the case as we continue to improve these models and investigate their uses.

References

1. Raghu D, Tripathy HK (2022) Medical image analysis using deep learning algorithm convolutional neural networks. In: Computer networks, big data and IoT: proceedings of ICCBI 2021. Springer, pp 245–256. https://doi.org/10.1007/978-981-19-0898-9_19
2. Seckler H, Metzler R (2022) Bayesian deep learning for error estimation in the analysis of anomalous diffusion. *Nat Commun* 13(1):6717. <https://doi.org/10.1038/s41467-022-34305-6>
3. Wolleb J, Bieder F, Sandkühler R, Cattin PC (2022) Diffusion models for medical anomaly detection. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 35–45. https://doi.org/10.1007/978-3-031-16452-1_4
4. Ko J, Kong I, Kim HJ (2024) Stochastic conditional diffusion models for semantic image synthesis. arXiv preprint arXiv:2402.16506
5. Nguyen Q, Vu T, Tran A, Nguyen K (2024) Dataset diffusion: diffusion-based synthetic data generation for pixel-level semantic segmentation. *Adv Neural Inf Process Syst* 36
6. Joseph K, Udhayan P, Shukla T, Agarwal A, Karanam S, Goswami K et al (2024) Iterative multi-granular image editing using diffusion models. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 8107–8116
7. Li Y, Wang H, Jin Q, Hu J, Chemerys P, Fu Y et al (2024) Snapfusion: text-to-image diffusion model on mobile devices within two seconds. *Adv Neural Inf Process Syst* 36
8. Mohammed MA, Hussain MA, Oraibi ZA, Abduljabbar ZA, Nyangaresi VO (2023) Secure content based image retrieval system using deep learning. *J Basrah Res Sci* 49(2):94–111
9. Khalid ET, Salah Khalefa M, Yassen W, Adil YA (2023) Omicron virus emotions understanding system based on deep learning architecture. *J Ambient Intell Humaniz Comput* 14(7):9497–9507. <https://doi.org/10.1007/s12652-023-04615-8>
10. Ghrabat MJJ, Ma G, Abduljabbar ZA, Al Sibahee MA, Jassim SJ (2019) Greedy learning of deep Boltzmann machine (GDBM)'s variance and search algorithm for efficient image retrieval. *IEEE Access* 7:169142–169159

11. Abduljabbar ZA, Abduljaleel IQ, Ma J, Al Sibahee MA, Nyangaresi VO, Honi DG et al (2022) Provably secure and fast color image encryption algorithm based on s-boxes and hyperchaotic map. IEEE Access 10:26257–26270. <https://doi.org/10.1109/ACCESS.2022.3151174>
12. Alzubaidi L, Fadhel MA, Al-Shamma O, Zhang J, Santamaría J, Duan Y et al (2020) Towards a better understanding of transfer learning for medical imaging: a case study. Appl Sci 10(13):4523. <https://doi.org/10.3390/app10134523>
13. Azizi S, Kornblith S, Saharia C, Norouzi M, Fleet DJ (2023) Synthetic data from diffusion models improves imagenet classification. arXiv preprint arXiv:230408466
14. Bruno F, Granata V, Cobianchi Bellisari F, Sgalambro F, Tommasino E, Palumbo P et al (2022) Advanced magnetic resonance imaging (MRI) techniques: technical principles and applications in nanomedicine. Cancers 14(7):1626. <https://doi.org/10.3390/cancers14071626>
15. Majid H, Ali KH (2023) Automatic diagnosis of coronavirus using conditional generative adversarial network (CGAN). Iraqi J Sci 4542–4556. <https://doi.org/10.24996/ij.s.2023.64.7.40>
16. Chen S, Ma K, Zheng Y (2019) Med3d: transfer learning for 3D medical image analysis. arXiv preprint arXiv:190400625
17. Abduljabbar ZA, Jin H, Ibrahim A, Hussien ZA, Hussain MA, Abbdal SH et al (2016) SEPIM: secure and efficient private image matching. Appl Sci 6(8):213. <https://doi.org/10.3390/app6080213>
18. Ho J, Saharia C, Chan W, Fleet DJ, Norouzi M, Salimans T (2022) Cascaded diffusion models for high fidelity image generation. J Mach Learn Res 23(47):1–33
19. Ntavelis E, Siarohin A, Olszewski K, Wang C, Gool LV, Tulyakov S (2023) Autodecoding latent 3D diffusion models. Adv Neural Inf Process Syst 36:67021–67047
20. Yang L, Liu J, Hong S, Zhang Z, Huang Z, Cai Z et al (2024) Improving diffusion-based image synthesis with context prediction. Adv Neural Inf Process Syst 36
21. Khader F, Müller-Franzes G, Tayebi Arasteh S, Han T, Haarbuerger C, Schulze-Hagen M et al (2023) Denoising diffusion probabilistic models for 3D medical image generation. Sci Rep 13(1):7303. <https://doi.org/10.1038/s41598-023-34341-2>
22. Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. Adv Neural Inf Process Syst 33:6840–6851
23. Wang W, Bao J, Zhou W, Chen D, Chen D, Yuan L et al (2022) Semantic image synthesis via diffusion models. arXiv preprint arXiv:220700050
24. Mustafa B, Loh A, Freyberg J, MacWilliams P, Wilson M, McKinney SM et al (2021) Supervised transfer learning at scale for medical imaging. arXiv preprint arXiv:210105913
25. Bhunia AK, Khan S, Cholakal H, Anwer RM, Laaksonen J, Shah M et al (2023) Person image synthesis via denoising diffusion model. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5968–5976. <https://doi.org/10.1109/cvpr52729.2023.00578>
26. Zhang J, Qi X, Zhao B (2023) Federated generative learning with foundation models. arXiv preprint arXiv:230616064
27. Jin Z, Shen X, Li B, Xue X (2024) Training-free diffusion model adaptation for variable-sized text-to-image synthesis. Adv Neural Inf Process Syst 36
28. Pan S, Abouei E, Wynne J, Chang CW, Wang T, Qiu RL et al (2023) Synthetic CT generation from MRI using 3D transformer-based denoising diffusion model. Med Phys
29. Podell D, English Z, Lacey K, Blattmann A, Dockhorn T, Müller J et al (2023) SDXL: improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:230701952
30. Dai X, Lei Y, Fu Y, Curran WJ, Liu T, Mao H et al (2020) Multimodal MRI synthesis using unified generative adversarial networks. Med Phys 47(12):6343–6354
31. Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman BA et al (2022) The medical segmentation decathlon. Nat Commun 13(1):4128. <https://doi.org/10.1038/s41467-022-30695-9>
32. Liu J, Anirudh R, Thiagarajan JJ, He S, Mohan KA, Kamilov US et al (2023) DOLCE: a model-based probabilistic diffusion framework for limited-angle CT reconstruction. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 10498–10508

33. Ghrabat MJJ, Ma G, Maolood IY, Alresheedi SS, Abduljabbar ZA (2019) An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier. *Human-centric Comput Inf Sci* 9:1–29. <https://doi.org/10.1186/s13673-019-0191-8>
34. Usman M, Zia T, Tariq A (2022) Analyzing transfer learning of vision transformers for interpreting chest radiography. *J Digit Imaging* 35(6):1445–1462. <https://doi.org/10.1007/s10278-022-00666-z>
35. Singh J, Gould S, Zheng L (2023) High-fidelity guided image synthesis with latent diffusion models. In: 2023 IEEE/CVF conference on computer vision and pattern recognition (CVPR). IEEE, pp 5997–6006. <https://doi.org/10.1109/cvpr52729.2023.00581>
36. Dorjsembe Z, Pao H-K, Odonchimed S, Xiao F (2023) Conditional diffusion models for semantic 3D medical image synthesis. *arXiv preprint arXiv:230518453*. <https://doi.org/10.21227/3ej9-e459>
37. Frangi AF, Tsafaris SA, Prince JL (2018) Simulation and synthesis in medical imaging. *IEEE Trans Med Imaging* 37(3):673–679. <https://doi.org/10.1109/TMI.2018.2800298>
38. Akbar MU, Wang W, Eklund A (2023) Beware of diffusion models for synthesizing medical images—a comparison with GANs in terms of memorizing brain tumor images. *arXiv preprint arXiv:230507644*
39. Wu W, Zhao Y, Chen H, Gu Y, Zhao R, He Y et al (2024) Datasetdm: synthesizing data with perception annotations using diffusion models. *Adv Neural Inf Process Syst* 36
40. Laino ME, Cancian P, Politi LS, Della Porta MG, Saba L, Savevski V (2022) Generative adversarial networks in brain imaging: a narrative review. *J Imaging* 8(4):83
41. Tanno R, Worrall DE, Kaden E, Ghosh A, Grussu F, Bizzi A et al (2021) Uncertainty modelling in deep learning for safer neuroimage enhancement: demonstration in diffusion MRI. *Neuroimage* 225:117366. <https://doi.org/10.1016/j.neuroimage.2020.117366>
42. Nguyen Q, Le T, Nguyen T, Nhat MN (2023) Class label conditioning diffusion model for robust brain tumor MRI synthesis. *Authorea Preprints*
43. Kemenczky P, Vakli P, Somogyi E, Homolya I, Hermann P, Gál V et al (2022) Effect of head motion-induced artefacts on the reliability of deep learning-based whole-brain segmentation. *Sci Rep* 12(1):1618. <https://doi.org/10.1038/s41598-022-05583-3>
44. Gu S, Chen D, Bao J, Wen F, Zhang B, Chen D et al (2022) Vector quantized diffusion model for text-to-image synthesis. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 10696–10706. <https://doi.org/10.1109/CVPR52688.2022.01043>
45. Kim J, Park H (2024) Adaptive latent diffusion model for 3D medical image to image translation: multi-modal magnetic resonance imaging study. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp 7604–7613
46. Karras T, Aittala M, Aila T, Laine S (2022) Elucidating the design space of diffusion-based generative models. *Adv Neural Inf Process Syst* 35:26565–26577
47. Dhariwal P, Nichol A (2021) Diffusion models beat gans on image synthesis. *Adv Neural Inf Process Syst* 34:8780–8794
48. Li Z, Fan Q, Bilgic B, Wang G, Wu W, Polimeni JR et al (2023) Diffusion MRI data analysis assisted by deep learning synthesized anatomical images (DeepAnat). *Med Image Anal* 86:102744. <https://doi.org/10.1016/j.media.2023.102744>
49. Linkous L, Zohrabi N, Abdelwahed S (2019) Health monitoring in smart homes utilizing internet of things. In: 2019 IEEE/ACM international conference on connected health: applications, systems and engineering technologies (CHASE). IEEE, pp 29–34. <https://doi.org/10.1109/CHASE48038.2019.00020>
50. Chen Z, Zhang Y, Liu D, Gu J, Kong L, Yuan X (2024) Hierarchical integration diffusion model for realistic image deblurring. *Adv Neural Inf Process Syst* 36

Research on the Design of Computer Security Protection System Based on Machine Learning



Haiming Zhang, Yuqi Wang, and Junpu Cai

Abstract This study aims to analyze the elements that affect the computer network security, then design the corresponding protection system, and detail its implementation mode. Based on artificial intelligence technology, it is committed to realizing the collection, analysis and processing of network data, so as to boost the effective application of Internet security and provide theoretical support for network engineering. The rapid development of Internet information has greatly improved the life and production of residents, but also caused many computer network security risks. Therefore, it is imperative to analyze and formulate the computer network security management strategy. Based on this study, by exploring the causes of computer network security, we propose the widely used machine learning security management technology. First, the design principles, overall architecture and network security structure of machine learning are discussed in depth, and then the support vector machine (SVM) algorithm, BP neural network algorithm and Web technology are introduced in detail. Combined with the advantages of machine learning intelligence and accuracy, the level of predictive analysis technology of computer network security management is expounded. Finally, the implementation and future goals of network security management technology are described, and the scientific development prospect of computer network security management based on machine learning is discussed.

Keywords Network security · SVM methodology · BP neural network approach · Machine learning Xi · Protection system

H. Zhang (✉)

School of Electronic Information Engineering, Wuhan Donghu University, Wuhan, Hubei, China
e-mail: zhm19880509@126.com

Y. Wang

School of Economics and Management, China Jiliang University, Hangzhou, Zhejiang, China

J. Cai

College of Electrical Engineering and Intelligent Manufacturing, Chongqing Metropolitan College of Science and Technology, Yongchuan, Chongqing, China

1 Introduction

In recent years, with the rapid pace of social development, computer network technology has penetrated into people's daily work and life, providing indispensable convenience for the public. However, this convenience is not without its cost, and computer networks have many potential security risks that can cause irreparable damage to the economy. Therefore, how to provide more efficient and convenient network services and meet different network security needs has become a hot topic in current research. Designing a network security protection system not only helps to improve the overall performance of the network, but also provides the necessary guarantee for network security. With the rapid development of China's economy and intelligent computer information, the importance of Internet application technology in the fields of science and technology, life and production has become increasingly prominent. However, this development has also brought with it the growing prominence of cybersecurity management issues. Taking 2019 as an example, China's Computer Information Security Prevention Center found about 11,000 security vulnerabilities in different platforms, among which distributed denial-of-service attacks and high-traffic attacks were the main problems. This not only increases the difficulty of computer security management, but also brings huge security risks to the protection of user information. Therefore, the core goal of this paper is to solve the key problems of computer network traffic, information and network platform security based on intelligent machine science Xi security management technology. The wide application of machine learning Xi technology not only plays a key role in the integration of domain knowledge, but also plays an important role in domain management and deployment. From life to work, machine learning Xi makes it easy to operate by recording and filtering user information. In the field of computer network security, the network management mode of machine learning Xi allows network security administrators to quickly find and solve vulnerabilities, improving the efficiency of security management.

2 Related Research

2.1 Design Principles

The computer network security management technology based on machine Xi follows four key design principles: scientificity, intuitiveness, security management stability and information scalability. These principles not only promote the user's understanding and application of the safety management system, but also promote the in-depth exploration of machine Xi methods and key technologies. In terms of scientificity, SVM and BP neural network algorithms are used to conduct in-depth analysis and evaluation of the network security situation. Compared with traditional methods, machine learning Xi greatly improves the accuracy of safety assessment

and improves the efficiency of overall safety management. In terms of intuitiveness, it helps security managers accurately understand the current state of network security by displaying the current security situation of the current network system through vivid visualization. The stability of security management ensures the smooth operation of each system module and enhances the secure sharing and construction of information between different modules. In terms of information scalability, machine Xi flexibly expands security protection tools according to the actual situation, providing a wider range of security design support.

Figure 1 illustrates the structural design process of the entire computer network security management system based on the machine Xi method. The system mainly includes users, professional and technical engineers, human–computer interaction and computer database security management system modules. Among them, the human–computer interaction module, as the core of the Xi method of machine learning, covers three important components: interpretation mechanism, machine learning Xi reasoning and knowledge acquisition. In terms of specific functions, the user module undertakes the quantitative assessment of computer network security and makes relevant predictions according to the evaluation results, data acquisition and situation values; the machine science Xi reasoning module focuses on the situation assessment of part of the data, generates the data in the required format and uses SVM or BP neural network algorithm to obtain the current computer network security situation for security assessment and prediction; TCP and other data are used for situation analysis and prediction, while the computer database security management system relies on user information and collected situation information to evaluate the security situation in a visual way and to realize the information sharing and security management functions between modules. This system design is committed to improving the predictive ability and overall management efficiency of computer network security.

2.2 The Necessity of Designing Computer Security Protection Systems

Ensuring the security of computer networks is one of the most critical challenges today. To address the challenges posed by an imperfect network operating system, inadequate network infrastructure and the threat of cyber virus invasions, it is crucial to establish a comprehensive security protection system; this entails continuously improving and updating the network operating system to eliminate vulnerabilities and security risks that could be exploited by hackers, correcting flaws within the network structure to enhance the system's resilience against various attacks, and mitigating the spread of cyber viruses that can compromise system functionality and lead to data breaches. A comprehensive security system should be able to identify [1], isolate and eradicate these viruses in a timely manner. To sum up, only by establishing

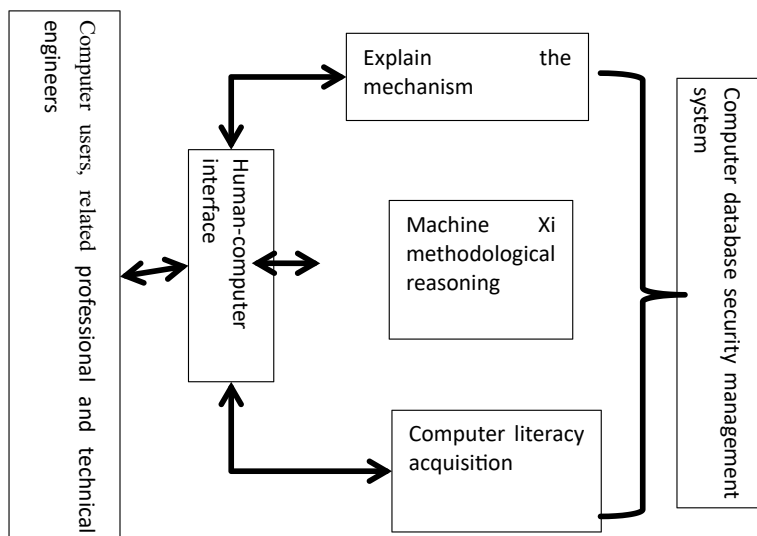


Fig. 1 Managing computer network security processes based on a machine-based Xi approach

a perfect security protection system can we effectively protect the computer network from various threats and ensure the security and integrity of the network and data.

3 Method

3.1 Support Vector Machine (SVM) Algorithm

In computer network security management, Support Vector Machine (SVM) algorithm is widely used in the field of machine science Xi because of its excellent prediction and evaluation accuracy [1–3]. Its principle lies in selecting the database kernel function and optimizing the model parameters, and selecting the optimal classification plane under the condition of satisfying a certain eigenvalue of a variety of kernel functions. By mapping data from a low-dimensional space to a high-dimensional space, SVM enables predictive classification of data to assist in network security management processes. Common SVM algorithm kernel functions include radial basis kernel functions and polynomial functions.

Radial basis kernel function:

$$k(x, y) = \exp(-|x - y|^2 / \sigma^2) \quad (1)$$

polynomial function:

$$k(x,y) = \left[(x,y) + 1 \right]^d$$

(2)

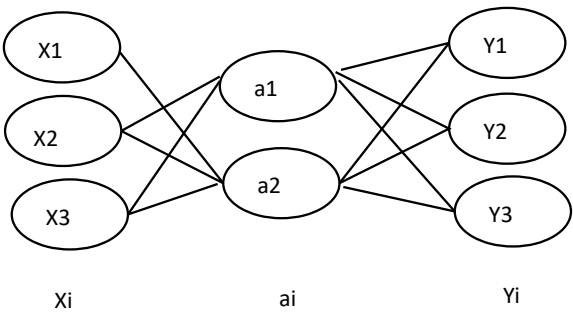
3.2 BP Neural Network Analysis

BP neural networks are common artificial neural networks that are widely used for pattern recognition and data classification. During training, a neural network adjusts its weights and biases to align its output more closely with the expected result; this process involves forward propagation, where input data is processed through the network, and backpropagation, where errors are utilized to fine-tune the parameters, thereby enhancing the network’s performance and accuracy [4]. Over time, with substantial data, the network can be optimized to tackle complex problems. The network’s structure, as depicted in the diagram, consists of interconnected layers—an input layer, a hidden layer and an output layer—linked by weights [5]. The training process encompasses calculating feature vector weights, transforming data, computing errors and adjusting parameters, all of which are essential for predictive analysis in computer network security as shown in Fig. 2.

The BP neural network iteratively processes computer network security data, and through numerous iterations, it evaluates and analyzes the input and output parameters of each layer; if the E(a) value exceeds the threshold, the threshold is adjusted, and the parameters are reassessed [6]. After multiple iterations, the threshold condition is finally met, that is, the BP judgment result is valid. The algorithm maps the input and output, and the data is repeatedly trained and iterated in the network to continuously improve the accuracy and effectiveness. The data results obtained from such multiple iterations of training are more accurate and reliable, and at the same time, the mutual laws between the input and output of training samples are more deeply understood.

The training process of BP neural network for computer network security is shown in Eqs. 3–4, where the value of the BP network output layer node is as follows:

Fig. 2 Schematic diagram of BP neural network cross-validation



$$y_k = \sigma \left(\sum_{j=1}^p V_{kj} b_j + \beta_k \right) \quad (3)$$

Use the sum of squares of the error to determine whether the training process is over:

$$E = \frac{1}{2} \sum_{k=1}^q (O_k - y_k)^2, \quad (4)$$

where O_k is the desired output; E is achieved, the error of the output layer is transmitted to the hidden layer and the input layer in reverse.

3.3 Web-Based Technology

Web technology, serving as the core of the computer network, underpins the Internet's infrastructure and supports the development of network clients and servers by encompassing multiple access methods such as HTTP and URL [7], and employing programming languages like Python and C++ to integrate, analyze and predict computer data, thereby playing a crucial role in cybersecurity management; through Python's data processing capabilities, network security efficiency is enhanced by identifying and mitigating potential security risks [8], while simultaneously, web technology proves indispensable in machine learning applications, requiring a comprehensive range of programming languages, frameworks and tools for web-based application development that involves both front-end and back-end processes [9], with HTML, CSS and JavaScript forming the foundation for web pages, frameworks like React, Angular and Vue providing richer functionality, and back-end technologies such as Node.js, Django and Ruby on Rails facilitating database processing, business logic and server-side coding. The combination of these technologies creates powerful, interactive web applications.

3.4 Design of Computer Network Security Protection System

Functional design:

The use of artificial intelligence technology in the computer network security protection system has significantly enhanced the accuracy and efficiency of information data collection. This intelligent system can identify and prevent network information data, improving the level of network security protection. The steps for cyber security protection are also optimized as shown in Fig. 3:

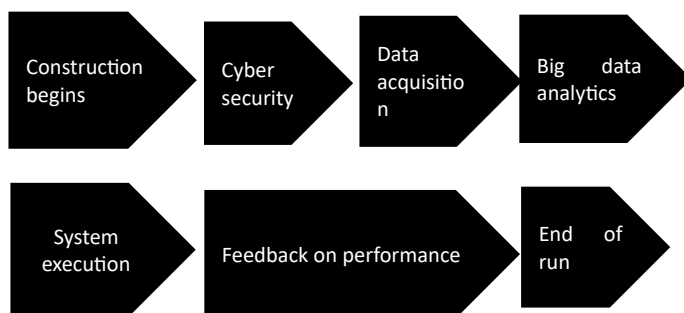


Fig. 3 Functional design of the computer network security protection system

3.5 Network Data Collection and Big Data Analysis [10]

The program action extraction module monitors program behavior and system processes in the operating system, and captures information such as file system and registry. The behavioral trait database is used to analyze and process dangerous or normal behavioral traits and is divided into a trait set database and a feedback adaptation database. The feature set database includes test feature tables, feature set tables, parameter tables and feature data tables. The Feedback Adaptation Library contains a system parameter table and a category effect table. The behavior analysis module is responsible for classifier X_i , classification calculation and data preprocessing. Finally, after detecting the malicious program, the system response module implements the necessary processing through the network system, such as eliminating the Trojan virus. Using artificial intelligence technology, the network information is systematically analyzed, and it is classified and analyzed in depth to discover potential computer Trojan viruses.

4 Results and Discussion

As one of the most basic forms of neural networks, perceptrons are similar to neurons in the human brain, which is the initial application of artificial intelligence to network security. It is characterized by establishing the connection between the system output and a certain level of input, processing the input through weight weighting, converting it into the corresponding output and transmitting it to other parts through the network. Unlike statistical models, AI algorithms iteratively adjust the size of the estimated results to minimize the cost function. The goal is to obtain accurate weighted vectors to make reliable predictions of future data. Spam filters use artificial intelligence technology to filter and classify. The grading task is based on the frequency of keywords in an email to distinguish between spam and normal messages. Specific scoring

thresholds are used to determine spam and are periodically adjusted to accommodate new types of spam. This process requires accurate measurement of keywords to effectively distinguish between spam and generic messages.

In machine Xi, the implementation of data acquisition and prediction modules is crucial. First, the network data information is comprehensively collected and understood, and then the precise computer network security situation analysis is carried out to ensure the accuracy of data integrity and security situation assessment. It focuses on the core aspects of computer network security situational awareness, including the extraction, evaluation and prediction of security situation. In the data collection, the data is analyzed in depth and comprehensively by using the byte proportion of UDP data and the byte proportion of ICMP data. Subsequently, the data were trained, transmitted, analyzed and compared and predicted by means of SVM model and BP neural network model, so as to build an intelligent, high-accuracy and high-efficiency computer network security management system. Security assessment effect analysis focuses on the evaluation and analysis of the security management situation. The SVM algorithm and BP neural network algorithm were used to train the sample data of the computer database and verify the prediction results. If there are discrepancies, adjust and optimize the parameters of the machine model, verify it again to improve the accuracy, ensure the formulation and implementation of high-quality, high-standard security management strategies and ensure the comprehensive security of computer network information.

5 Conclusion

The security of computer networks is crucial, especially in the process of information transmission and exchange. Building a security system requires a variety of factors, including risk assessment and regulatory requirements. This paper discusses a variety of security protection technologies, summarizes the implementation steps from the perceptron detection of spam and provides guidance for the construction of network security system. Ensuring the security of information transmission requires the establishment of a complete security protection system, and it is important to protect the integrity and privacy of information. The design of this system must take into account multiple factors, such as risk assessment, compliance and security standards, to build a comprehensive protection system against internal and external threats and attacks. At the 15th G20 Leaders' Summit, General Secretary Xi Jinping highlighted the far-reaching impact of the pandemic on new technologies and economic development. The rise of emerging technologies, such as 5G, artificial intelligence, and smart cities, has fueled a boom in the "contactless economy," providing a new growth model for the economy. In the face of challenges, we need to turn crises into opportunities and promote scientific and technological innovation through structural and digital transformation.

References

1. Qi H, Zhu W, Ye M et al (2024) Construction of power network security risk assessment model based on LSA-SVM algorithm in the background of smart grid. *Sci Rep* 14(1). <https://doi.org/10.1038/s41598-024-59473-x>
2. Nuhmana S (2023) Sentiment analysis of related public opinion smart city Pasuruan regency on media social using SVM algorithm. *J Soc Res*. <https://doi.org/10.55324/josr.v2i4.702>
3. Hamid TMTA, Sallehuddin R, Yunus ZM et al (2023) Ensemble filters with harmonize PSO-SVM algorithm for optimal hearing disorder prediction. *Neural Comput Appl* 1–24. <https://doi.org/10.1007/s00521-023-08244-2>
4. Chen G, Huang WX, Ronch AD et al (2023) BP neural Network-Kalman filter fusion method for unmanned aerial vehicle target tracking. *Proc Inst Mech Eng C J Mech Eng Sci* 237(18):4203–4212. <https://doi.org/10.1177/0954406220983864>
5. Wang J, Xia D, Wan J et al (2024) Color grading of green Sichuan pepper (*Zanthoxylum armatum* DC.) dried fruit based on image processing and BP neural network algorithm. *Sci Hortic* 331. <https://doi.org/10.1016/j.scienta.2024.113171>
6. Guo J, Xi W, Yang Z et al (2024) Landslide hazard susceptibility evaluation based on SBAS-InSAR technology and SSA-BP neural network algorithm: a case study of Baihetan reservoir area. *J Mt Sci* 21(3):952–972. <https://doi.org/10.1007/s11629-023-8083-9>
7. Zhang Y (2023) Design of an English Web-based teaching resource sharing platform based on mobile web technology. *Int J Adv Comput Sci Appl*. <https://doi.org/10.14569/ijacsa.2023.0140268>
8. Budiman J, Widjaja AE et al (2024) The development and prototyping of game modeling at a family entertainment center, utilizing web-based Arduino technology for calculating CAGR. *Procedia Comput Sci* 234:920–927. <https://doi.org/10.1016/j.procs.2024.03.080>
9. Serrano DM, Hervas R, Bravo J (2022) Telemaco: context-aware system for tourism guiding based on Web 3.0 Technology
10. Lee CM, Kang JH, Go SI (2024) Clinical implication of megestrol acetate in metastatic gastric cancer: a big data analysis from health insurance review and assessment (HIRA) database. *Support Care Cancer* 32(4). <https://doi.org/10.1007/s00520-024-08430-5>

LLMs in Healthcare: Exploring Language Models for Dental, Medical, and Mental Health with LangChain



S. Krishnaveni, B. Jothi, S. Sivamohan, Gautham Brijesh, S. Aiswarya, Louie Allen, and Syed Faizan Fiaz

Abstract Despite technological advancements, accurate diagnosis remains challenging in healthcare due to the complexity of medical language and the necessity for precise interpretation. This project addresses this challenge by investigating the impact of language models (LMs) Llama and GPT in healthcare contexts. The study aims to assess how these LMs, aided by the LangChain text-based analysis framework, can enhance diagnoses processes and decision-making across the dental, medical, and mental health domains. By analyzing Llama2 and GPT-3.5's performance in deciphering healthcare language complexities using diverse datasets, including clinical notes and research papers, the research endeavors to improve diagnostic accuracy and treatment strategies. Moreover, ethical considerations such as privacy protection and bias mitigation are addressed to ensure that LMs are used responsibly in healthcare practices. This project aims to develop customized LM solutions that bridge the gap between advanced NLP techniques and practical healthcare needs, ultimately improving patient outcomes.

Keywords Healthcare · Language models (LMs) · Llama2 · GPT-3.5 · Natural language processing (NLP) · LangChain · Diagnostics

1 Introduction

The healthcare sector has undergone a dramatic transformation with the advent of language models (LMs), with far-reaching implications for diagnosis, treatment modalities, and patient care. Within the complicated domains of dentistry, medicine,

S. Krishnaveni (✉) · B. Jothi · G. Brijesh · S. Aiswarya · L. Allen · S. F. Fiaz
Department of Computational Intelligence, SRM Institute of Science and Technology (Deemed to be University), Kattankulathur, Chennai, India
e-mail: krishnas4@srmist.edu.in

S. Sivamohan
Department of Networking and Communication, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

and mental health, the application of sophisticated LMs, such as Llama2 and GPT-3.5, presents a potential frontier. The efficacy and adaptability of these LMs are thoroughly investigated in this study using the sophisticated LangChain linguistic analysis methodology. Llama2 and GPT-3.5 need to be thoroughly studied in order to support clinical decision-making processes, comprehend complicated clinical language, and interpret domain-specific lexicons. We want to investigate the accuracy and application of these LMs in many healthcare situations by utilizing extensive datasets that contain clinical stories, scholarly literature, and patient contacts. But incorporating LMs into healthcare ecosystems necessitates a careful assessment of their technical viability, moral implications, and possible prejudices. Thus, in addition to performance evaluation, this study looks at bias reduction techniques, privacy protection policies, and ethical issues related to the use of LM in healthcare settings. The purpose of this study, which is being carried out in conjunction with domain experts, is to provide insights for the creation of customized LM solutions that will improve treatment algorithms, raise diagnosis accuracy, and ultimately improve patient outcomes. This project aims to catalyze revolutionary moves toward an equitable, accessible, and technologically enhanced healthcare environment by bridging the gap between state-of-the-art natural language processing techniques and the practical needs of healthcare delivery. By working together across disciplines and combining cutting-edge technology with clinical knowledge, we want to unlock the full potential of LMs in revolutionizing healthcare delivery [3].

2 Problem Statement

Despite the great potential of language models (LMs) such as Llama2 and GPT-3.5 to transform healthcare practices, there is still a crucial need to thoroughly assess their usefulness and application in the dental, medical, and mental health domains. While these LMs have sophisticated natural language processing skills, their ability to recognize domain-specific linguistic subtleties and enhance clinical decision-making processes is not well understood, furthermore, questions about privacy, partiality, and fairness [4].

The lack of detailed study and discussion of ethical concerns complicates the efficient integration of LMs into healthcare operations. Healthcare practitioners may find challenges in using these tools to improve diagnostic accuracy, treatment planning, and patient outcomes unless they have a comprehensive awareness of their strengths, limits, and ethical implications.

This study primarily focusses on the necessity to systematically assess the performance, privacy implications, and ethical considerations of Llama2 and GPT-3.5 in healthcare settings, aiming to guide the creation of customized language model solutions that enhance the quality and equity of healthcare delivery [5].

3 Literature survey

Detection of Emotion by Text Analysis using Machine Learning. The paper investigates the critical significance of emotions in human life and the need for emotion recognition in machine–human communication. It presents an artificial intelligence approach for automatically identifying human emotions, allowing machines, particularly chatbots, to make correct assessments and communicate adaptively. Experiments compare lexicon-based, standard machine learning, and deep learning methods for detecting emotions from different data sources, with neural networks outperforming them, particularly in multi-classification tasks. Practical applications demonstrate the efficacy of the best models, emphasizing the possibility for better human–machine interaction. While admitting the persistent problem of fully automating emotion recognition, the report emphasizes the importance of continued research to improve accuracy and system robustness [6]. It concludes by emphasizing the importance of automated emotion analysis in understanding human behavior, as well as its prospective applications in areas such as human–robot interaction and geriatric care.

Emotion-Conditioned Text Generation through Automatic Prompt Optimization. This paper describes a unique technique to emotion-conditioned text generation that avoids the challenges of resource-intensive fine-tuning. The authors present an iterative approach to automatic prompt optimization that involves adding, removing, and replacing tokens in order to improve prompts. This method focuses on text production conditioned by emotions, in contrast to earlier prompt optimization approaches. The generated text’s fulfillment of the emotion condition is measured by the objective function using a text classifier. The optimized prompts work better than manually created seed prompts, according to experimental results, with a macro-average F1 score of 0.75 as opposed to 0.22. The study addresses the method’s consequences for both cost-effectiveness and ethical issues [7].

Chatbots with Emotional Intelligence: A Meta-analysis an analysis of the existing research on the topic of developing chatbots with emotional intelligence is the focus of this study.

It analyzes 42 articles published in the last decade to understand the problems addressed, techniques used, and evaluation measures employed in embedding emotion in chatbot conversations [8]. According to the findings, the majority of studies focus on effectively detecting the user’s emotion and providing emotionally meaningful responses. The paper also discusses the application of AI and NLP techniques, such as Seq2Seq encoding and decoding, to create emotionally intelligent chatbots. The paper discusses the importance of emotions in effective communication and the potential benefits of chatbots that simulate human behavior and emotions. It concludes by suggesting areas for further research, including voice-based and multimodal chatbots, as well as task-oriented emotionally intelligent chatbots.

Towards Facilitating Empathetic Conversation in Online Mental Health Support: A Reinforcement Learning Approach. This study examines conditional natural language generation techniques, which frequently necessitate costly adjustments starting from scratch to train a sizable language model. Without a significant amount

of data and computational power, both are unlikely to provide satisfactory results. A viable substitute is to adjust the parameters of a huge language model. It is an economical method that yet yields competitive outcomes. Although this process is already well-established for structured prediction and zero- and few-shot text categorization, conditional text generation has not given it any thought. We introduce the first automatic prompt optimization method using instruction-fine-tuned models for emotion-conditioned text production. By adding, deleting, or swapping out tokens, our technique modifies the prompt through an iterative optimization process. The objective function relies on a text classifier measuring the realization of the conditional variable in the generated text. The evaluation, focused on emotion-conditioned text generation, demonstrates the efficacy of the approach. The optimized prompts achieve a notable macro-average F1 score of 0.75, outperforming manually designed seed prompts that only attain a 0.22 macro-average F1 score [9].

Towards Emotional Support Dialog Systems, the study discusses the critical function of emotional support in a range of conversational contexts, including social encounters, talks with customers, and mental health support. Despite the value of emotional support, there has not been much study done on creating dialog systems that can deliver it because there are not any well-designed tasks or corpora for emotional support dialogs. The study presents an ESC framework based on the Helping Skills Theory and introduces the Emotional Support Conversation (ESC) assignment. The Emotion Support Conversation dataset (ESConv), which includes rich annotations with a focus on support strategy in both help- seeker and supporter roles, is created by the authors to aid in research. To guarantee high-quality instances of successful emotional support discussions, the dataset is meticulously curated using training tutorials for supporters and quality control procedures throughout data collection. In addition to showing the importance of support techniques in providing effective emotional support, the study examines state-of-the-art dialog models and shows how useful ESConv is for training increasingly sophisticated emotional support systems.

An AI-Powered Chatbot for Emotional Dialogue Analysis and Sentence Generation in Mental Healthcare Services. The creation of a chatbot for psychiatric therapy within the framework of mental healthcare services is the main topic of Kyojoong Oh and Dongkun Lee's paper. Sentence creation and emotional interaction analysis are used to improve the chatbot's functionality. The writers discuss the drawbacks of current intelligent assistant services, such as Google Now and Apple Siri, which frequently depend on rudimentary Natural Language Processing (NLP) methods. The research presents a hierarchical convolutional neural networks (CNNs) committee machine architecture to alleviate these drawbacks [10]. This framework showcases models that use labeled data across several modalities, such as temporal time domains, multi-modal techniques, and static images, with the goal of achieving high accuracy in face expression identification. The suggested chatbot has the potential to offer more sophisticated and successful psychiatric therapy in mental healthcare services because of its emphasis on emotional dialog analysis and integration of cutting-edge machine learning algorithms.

The Potential of Chatbots for Emotional Support and Promoting Mental Well-Being in Different Cultures: Mixed Methods Study. According to the study, users

display depressive or sad emotions while interacting with chatbots, indicating that chatbot services have the potential to provide psychological support. According to the statistics, people anticipate chatbots to provide a safe space to express their feelings, as well as advice or information to help them cope with bad emotions. The study also found cultural variances in how people express their emotions to chatbots, as well as differences in what they convey to chatbots vs social networking sites. Furthermore, the research found that chatbot talks required less seeking feedback or emotional support from others than social media discourse. The findings indicate that chatbots have the ability to give emotional support and a secure venue for addressing difficult topics.

EMMA: An Emotion-Aware Wellbeing Chatbot The emotionally intelligent wellness personal assistant EMMA (EMotion-Aware mHealth Agent), intended for mental health interventions, is presented in this study. The study investigates whether an emotionally intelligent chatbot may be used to automate just-in-time mobile health solutions. EMMA uses machine learning models for mood detection using smart phone sensor data and offers sympathetically tailored micro-activities. A two-week human subject experiment comprising 39 participants is part of the research, which compares EMMA with a control condition. The findings show that people find the chatbot more pleasant than the bot that uses self-reported emotion samples when it comes to automating mood recognition through location and personalization data. The study offers insightful information about how emotion-aware bots for mHealth could be designed and how it might affect users' perceptions and engagement.

Enhancing Human-Computer Interaction using Emotion-aware Chatbots for Mental Health Support The integration of emotion-aware chatbots in technology-driven mental health support signifies a transformative shift in Human-Computer Interaction (HCI). This article delves into the profound impact of these chatbots, emphasizing their design principles, development processes, and the ethical considerations guiding their deployment. By infusing emotional intelligence with technological capabilities, emotion-aware chatbots aim to enhance user engagement and well-being, offering a personalized and empathetic interaction in mental healthcare systems. The study explores the creation, functionality, and effects of these chatbots, underlining the role of emotional intelligence in addressing gaps in traditional HCI methodologies. Through user testing, feedback analysis, and ethical considerations, the research aims to comprehensively assess the effectiveness and ethical application of emotion-aware chatbots in mental health assistance. Key metrics such as user satisfaction and engagement are highlighted as pivotal in evaluating the success of these chatbots, emphasizing the importance of users reporting a sense of connection and understanding. The article also mentions the incorporation of Natural Language Processing (NLP), voice analysis, and facial recognition technologies as integral components in enabling chatbots to interpret and respond to emotional expressions effectively [2].

Leveraging Language Models in Healthcare: A Literature Survey

This paper presents a comprehensive literature survey on the application of language models (LMs) in healthcare. It analyzes a selection of recent studies to understand the development, challenges, and potential of LMs in dental, medical,

and mental health domains. The survey encompasses works by Alsentzer et al.'s [11] on clinical BERT embeddings, Zhang et al.'s [12] approach integrating medical knowledge with contextual embeddings, and Rajkomar et al.'s [13] scalable deep learning method with electronic health records. Additionally, Gehrmann et al.'s [14] comparison of deep learning and concept extraction methods for patient phenotyping contribute to the understanding of LM applications in healthcare. The findings highlight the growing interest in leveraging LMs for clinical tasks, emphasizing the need for further research into their efficacy, ethical considerations, and integration into healthcare workflows.

Vallabhaneni et al. [15] developed a comprehensive dataset for lung cancer detection by integrating diverse modalities such as patient-reported symptoms, medical imaging, and doctor prescriptions. Their study employed advanced deep learning techniques, notably Convolutional Neural Networks (CNNs) and Large Language Models (LLMs), to effectively analyze the multimodal data. By leveraging the complementary strengths of visual and textual information, the proposed model significantly enhanced the accuracy of lung cancer detection. Notably, the model demonstrated the ability to extract and interpret textual content embedded within medical images, thereby improving diagnostic predictions and offering a robust tool for early detection and clinical decision-making.

Publicly Available Clinical BERT Embeddings [11] introduced a specialized variant of the Bidirectional Encoder Representations from Transformers (BERT) model specifically fine-tuned on clinical notes. Published in the Proceedings of the 2nd Clinical Natural Language Processing Workshop, their study showcases the efficacy of Clinical BERT across various clinical natural language processing (NLP) tasks. By demonstrating its effectiveness, Alsentzer et al. laid the groundwork for understanding the adaptation of language models (LMs) in healthcare contexts. This re-search not only highlights the significance of tailoring LMs to domain-specific language but also underscores the potential for improving healthcare outcomes through the integration of advanced NLP techniques [1].

A comparison of approaches for patient phenotyping from clinical narratives that rely on idea extraction and deep learning. This research offers valuable insights into the efficacy of various approaches in utilizing language model (LM)-based models for patient characterization tasks. Gehrmann et al. contribute to our understanding of the strengths and limitations of LM-based approaches in healthcare applications by examining the performance of deep learning techniques alongside concept extraction methods. This study underscores the importance of selecting appropriate methodologies for patient phenotyping tasks and provides a foundation for further research into optimizing LM utilization in clinical contexts.

4 Methodology

The proposed modules involved in this project are the following:

- Data Collection
- Data Preprocessing
- Query Refinement
- Conversation Handling

A. *Data Collection*

Incorporating concepts from publications such as Daniel Goleman's "Emotional Intelligence" and "Emotional Intelligence Coaching," the chatbot can provide users with techniques for identifying and efficiently handling their emotions. It can offer direction on developing empathy, self-control, self-awareness, and social skills, enabling users to handle difficult circumstances with emotional intelligence [6]. By utilizing the idea of daily coaching ideas from "365 Ways to Be Your Own Life Coach," the chatbot can provide users with digestible recommendations and activities that they can incorporate into their everyday routines. These suggestions can assist users in making little steps toward both personal and professional development. They can address a wide range of subjects, including goal-setting, stress management, communication techniques, and self-care. The chatbot can give priority to developing a sympathetic and encouraging relationship with users, as suggested by Carl Rogers' "On Becoming a Person." It can show sincere curiosity, attentive listening, and accepting acceptance without passing judgment, fostering a secure environment in which people feel free to express themselves honestly and explore their ideas and emotions. Using dialectical behavior therapy (DBT) methods from "The Dialectical Behavior Therapy Skills Workbook" can provide the chatbot useful capabilities for managing emotions, tolerating discomfort, practicing mindfulness, and fostering positive relationships with others. To improve their mental resilience and learn coping mechanisms, users can take part in guided meditations, role-playing games, and interactive activities. Overall, the chatbot's usefulness may be enhanced by incorporating lessons learned from these books into its design and content, making it a more complete and helpful tool for those looking for advice, self-improvement, and mental health support. Additionally, it can contribute to fostering a sense of empowerment, self-awareness, and emotional well-being among users as they interact with the chatbot over time [7].

B. *Data Preprocessing*

The preprocessing module includes a range of functions designed for the preliminary tasks necessary before embedding and indexing text chunks for effective retrieval. It serves as the foundation for preparing raw textual data, such as PDF documents, for further processing in the natural language pipeline [10]. This module contains various critical components and functions aimed at streamlining the pre-processing procedure. These include routines that load PDF documents (`load_pdf(data)`) from a specified directory, allowing access to the raw content stored within them

[4]. The `text_split(extracted_data)` method divides the retrieved PDF data into smaller, more manageable text chunks. This segmentation technique is critical for breaking down big texts into more digestible chunks, allowing for efficient processing and analysis downstream. The module also includes utilities for downloading embeddings from Hugging Face models (`download_hugging_face_embeddings()`), which allows for the production of embeddings for text chunks. These embeddings are numerical representations of textual information that aid machine learning tasks and algorithms. Furthermore, the module offers a function (`initialize_pinecone()`) that sets up the Pinecone vector index, providing the framework for fast embedding storage and querying.

C. *Query Refinement*

This module is in charge of refining user searches and retrieving relevant information based on user input. It uses OpenAI's language model to develop refined queries that are more likely to return relevant data from the knowledge base. The (`refine_query()`) function in this module passes the user's query and conversation history to the OpenAI model, which generates a refined query for better search results. The module also offers the (`find_match()`) method, which uses a pre-trained embedding model to encode user input and searches a search index for the most relevant matches. This function adds to the query refinement process by retrieving specific information from the knowledge base based on the refined query. Together, these features improve the chatbot's responses and contribute to a more seamless user experience [5].

D. *Conversation Handling*

This module controls the conversation flow between the user and the chatbot, ensuring a smooth and consistent interaction experience. It includes features for initializing and maintaining conversation states, handling messages from both the user and the bot, and creating a structured representation of the conversation history. The (`initialize_conversation()`) function configures the conversation environment, including the language model to be used and the memory mechanism for maintaining context. It also creates message templates to help steer the conversation and ensure smooth communication between the user and the chatbot [1]. The (`get_conversation_string()`) function in this module creates a formatted string representation of the conversation history. It loops through the stored conversation history in the Streamlit session state, merging human inquiries and bot responses into a unified discussion style. This structured string gives consumers a clear perspective of the interaction history, allowing them to evaluate previous exchanges and contextualize the current conversation. Encapsulating this functionality into the conversation handling module keeps the code organized and modular, making it easier to maintain and scale the chatbot application [2]. Figure 1 shows the workflow of the proposed method.

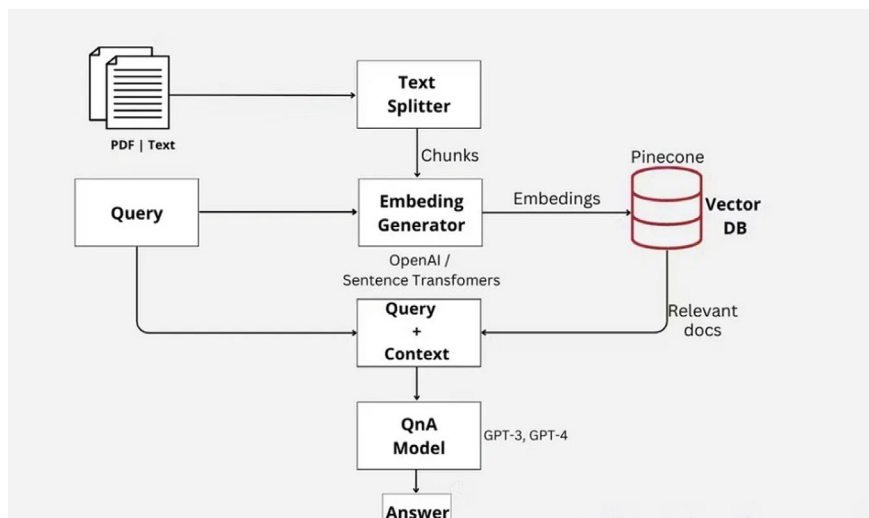


Fig. 1 Overall workflow of the proposed method

5 Proposed Method

The architecture diagram illustrates a question answering (QnA) system that uses large language models (LLMs) to process text and provide answers to user queries. Here's a step-by-step breakdown of the components and how they work together:

1. **PDF/Text:** This is the input data for the system, which can be in PDF format or plain text.
2. **Text Splitter:** This component divides the input text into manageable chunks or segments. It's designed to handle large documents by breaking them down into paragraphs, sections, or other logical units.
3. **Embedding Generator:** This module generates embeddings for the text chunks. An embedding is a vector representation of text that captures semantic meanings. The generator uses OpenAI's Sentence Transformers to convert text into a format that's suitable for machine learning models.
4. **Vector database (Pinecone):** The embeddings are stored in a vector database, which is optimized for similarity search. Pinecone is a specific type of vector database that allows the system to efficiently retrieve documents based on their semantic similarity to a query.
5. **Query:** This is the user's input question or search term.
6. **Query + Context:** The query is processed alongside the context, which is additional information provided by the user or gathered from previous interactions to improve the relevance of the answer.

7. **QnA Model:** This is the language model (in this case, GPT-3.5 and Llama2) that uses the query and context to generate a response. The model considers the embeddings from the Vector DB to find relevant documents (docs) and formulate an answer.
8. **Answer:** The final output of the system, which is the answer to the user's query. This is based on information extracted from the relevant documents and synthesized by the QnA model.

6 Comparison

Client: What is medicine?

Gpt-3.5: Medicine is the science and practice of diagnosing, treating, and preventing disease.

Llama2: A patient's emotional well-being, mental health, job conditions, surroundings, and relationships are all seen as essential components of a patient's overall health in medicine, which takes a holistic approach to treatment. Compared to Western medicine, Chinese medicine discusses the body and health via the use of various symbols and concepts.

Client: What is dental medicine?

Gpt-3.5: The field of dentistry deals with the examination, diagnosis, and management of diseases and disorders affecting the mouth and its structures, including teeth, gums, and the surrounding tissues.

Llama2: If you have any kind of problem with your teeth, gums, or the tissues that support them, you should see a dentist or a doctor who specializes in dental care. Dentists are healthcare professionals who specialize in this field and are trained to provide preventive, restorative, and surgical care to patients.

Client: What is mental health?

Gpt-3.5: Any state of complete emotional, psychological, and social wellness is known as mental health. It incorporates one's mode of thinking, feeling, and behaving, as well as one's stress management, interpersonal skills, and decision-making abilities.

Llama2: World Health Organisation definition of mental health includes not only the absence of sickness or impairment but also a good balance on physical, mental, and social levels. Nutrition, exercise, breathing, and sleep are all part of the physical side of things, but mental qualities like serenity and resilience in the face of adversity are also essential.

Client: What type of diseases I might have if I have a high blood pressure

Gpt-3.5: This list includes conditions such as glaucoma, low blood sugar, psoriasis, asthma, emphysema, an enlarged prostate, an overactive thyroid, liver illness, renal disease, and disorders of the blood.

Llama2: If you have high blood pressure (hypertension), you are at increased risk for developing certain diseases, including:

- Heart disease: A higher risk of cardiovascular disease complications, such as heart attacks and strokes, is associated with hypertension.
- Kidney disease: Increased blood pressure imposes an additional burden on the kidneys, which may eventually develop chronic renal disease or failure of the kidneys.
- Eye problems: Loss of vision or blindness may result from hypertension-induced damage to the eye's blood vessels.
- Stroke: The death of brain cells happens in a stroke when a blood artery in the brain is either ruptured or clogged. Stroke risk factors include high blood pressure.
- Peripheral artery disease: The danger of peripheral artery disease, which may harm the blood arteries in the feet and legs, increases when blood pressure is high.

Client: What type of mental disorder could I have if I am always stressed and have headache.

GPT-3.5: If you are always stressed and experiencing headaches, it does not necessarily indicate a mental disorder. Stress and headaches can be related to various factors such as migraines or transient ischemic attacks, as mentioned in the text or anxiety disorder.

Llama2: It's difficult to say for certain without more information, but based on the list provided, it seems possible that you may be experiencing symptoms of anxiety disorder. Biofeedback therapy can be effective in managing stress and anxiety, so it may be worth considering as a treatment option. However, it's important to consult with a mental health professional for an accurate diagnosis and appropriate treatment.

Client: I have a severe tooth ache, what would you recommend?

GPT-3.5: Based on the context provided, if you have a severe toothache related to denture bases, it is recommended to address the cause of the pain. This may involve checking for ill-fitting dentures, occlusal errors, pressure points, or teeth erupting under the denture base. Seeking professional dental help to adjust or correct the denture base is advised.

Llama2: If you are experiencing a severe toothache, I would recommend visiting a dentist as soon as possible. Mouth dryness can be a symptom of a number of oral health issues, including tooth decay, gum disease, or an abscess. A dentist can properly diagnose the cause of your toothache and provide appropriate treatment options. If you are unable to visit a dentist right away, you can try temporarily relieving the discomfort with over-the-counter pain relievers such as ibuprofen or acetaminophen. However, these should only be used as a temporary solution and not as a substitute for proper medical attention. Remember, if your toothache persists or worsens over time, it's important to seek professional dental care as soon as possible to prevent any further complications.

Client: I'm feeling sad

GPT-3.5: Sadness can lead to a drop in energy and enthusiasm for activities, as well as an introspective withdrawal. This state may provide the opportunity to mourn a loss or frustrated hope, understand its impact on one's life, and eventually plan new beginnings.

Llama2: I understand that you are feeling sad, and it's okay to feel that way. It's important to acknowledge and validate your emotions, especially when they are negative. Sometimes, hearing someone say "it's okay to feel sad" can help you feel more comfortable with your emotions and process them in a healthy way. Remember that it's okay to feel whatever you are feeling, and it's important to take care of yourself during this time.

Client: How do I cope with sadness

GPT-3.5: You can find powerful and motivating coping thoughts, write them on a note card or store them in a note app on your smartphone. Remind yourself of these coping thoughts when you're in a distressing situation. Additionally, you can place these thoughts on sticky notes in visible spots like your refrigerator or mirror for regular reinforcement.

Llama2: It's important to acknowledge your feelings of sadness and give yourself time to process them. You can try journaling, talking to a trusted friend or family member, or engaging in a relaxing activity that brings you joy. It may also be helpful to seek professional help if your sadness persists or is impacting your daily life.

Client: I am having constant anxiety attacks

GPT-3.5: Learning to be mindful of your emotions without judging them can help decrease the intensity of anxiety attacks and prevent them from becoming more overwhelming or painful. Additionally, developing alternative coping strategies can be beneficial in managing emotion-triggering events.

Llama2: It's possible that your anxiety is exacerbating a pre-existing medical condition. If you haven't already, consider consulting with a healthcare professional for proper diagnosis and treatment. In the meantime, there are some things you can try to help manage your anxiety: (1) Practice relaxation techniques such as deep breathing or progressive muscle relaxation. (2) Engage in regular physical activity, such as walking or jogging, which can help reduce stress and anxiety. (3) Try journaling or talking to a therapist about your feelings and experiences. (4) Avoid caffeine, nicotine, and alcohol, which can worsen anxiety. (5) Get enough sleep each night and establish a consistent sleep schedule. Remember, it's important to prioritize your mental health and seek professional help if you need it.

Client: What are other ways to cope with anxiety

GPT-3.5: other ways to cope with anxiety may include utilizing distress tolerance skills, creating an emergency coping plan, incorporating physiological coping skills, and exploring common coping strategies such as mindfulness, deep breathing exercises, physical activity, seeking social support, practicing relaxation techniques, and cognitive-behavioral strategies.

Llama2: When anxiety problems are not too severe, psychiatry may be enough treatment. In most cases, physicians will prescribe medicine in addition to psychotherapy to patients who suffer from extreme anxiety. The majority of patients find that a mix of medication and psychotherapy works better than just one or the other.

Client: Is it harmful to my mental health if I have anger issues

Gpt-3.5: Having chronic anger can lead to susceptibility to certain issues. It's important to address and manage anger issues as they can have negative impacts

on mental health, having anger issues can have negative effects on physical health, specifically on heart function and the risk of subsequent heart attacks, as highlighted in the study mentioned. While the text does not directly address mental health implications of anger issues, it is important to note that chronic anger and hostility can potentially impact mental health as well.

Llama2: It is not uncommon for individuals with chronic anger issues to experience a range of negative emotions, including mistrust, cynicism, and a propensity for snide comments and put-downs. However, it is important to recognize that hostility is a habit that can change, and there are many strategies and techniques that can help individuals manage their anger in a healthier way.

7 Results

A. Overview

This study employed quantitative analyses to evaluate the effectiveness of the GPT-3.5 and Llama models in processing and understanding language-based queries in the contexts of dental, medical, and mental health. A series of metrics including accuracy, response time, and user satisfaction were measured.

B. Quantitative Results

Accuracy of Responses

Dental Health Applications GPT-3.5: Achieved an average accuracy rate of 92%. Specific strengths were noted in generating treatment plans based on symptoms described in queries. Llama: Recorded an average accuracy rate of 89%. While slightly lower, Llama excelled in providing diagnostic information (Table 1). Table 1 presents the research findings comparing GPT-3.5 and Llama2 model performance in dental healthrelated queries.

Medical Health Applications

GPT-3.5 demonstrated superior performance with an accuracy of 94% in scenarios involving clinical decision support systems. Llama showed a proficiency with a 91% accuracy rate, particularly in patient management advice (Table 2). Table 2 presents the research findings comparing GPT-3.5 and Llama2 model performance in Medical healthrelated queries.

GPT-3.5 demonstrated superior performance in therapeutic assistance scenarios, achieving an accuracy of 90%. LLAMA performed at an 88% accuracy level, with particular effective- ness in mood disorder identification (Table 3). Table 3 presents

Table 1 Comparison of AI model accuracy in dental health applications

Model	Accuracy (%)
GPT-3.5	92
Llama2	89

Table 2 Comparison of AI model accuracy in medical health applications

Model	Accuracy (%)
GPT-3.5	94
Llama2	91

Table 3 Comparison of AI model accuracy in mental health application

Model	Accuracy (%)
GPT-3.5	90
Llama2	88

the research findings comparing GPT-3.5 and Llama2 model performance in Mental health-related queries.

The response time was measured from the initiation of the query to the delivery of the final answer. GPT-3.5 consistently provided faster responses across all applications, with average times of 2.3 s for dental health, 2.1 s for medical health, and 2.2 s for mental health applications. Llama’s response times were 2.7, 2.5, and 2.8 s, respectively, for these applications.

User Satisfaction
User satisfaction was assessed through a survey rating the helpfulness and usability of the responses on a scale from 1 (poor) to 5 (excellent) (Table 4). Table 4 presents the research findings comparing GPT-3.5 and Llama2 model performance Based on User Satisfaction in Healthcare Applications.

C. *Qualitative Feedback*

Key insights from open-ended responses included: GPT-3.5 was praised for its comprehensive and detailed replies, especially in complex scenarios. Llama was valued for its straightforward and concise answers, though some users desired more detailed explanations. The application of LMs in healthcare raises significant ethical and privacy concerns, particularly regarding data security and bias in decision-making processes. Our research highlights the effective strategies employed by both Llama2 and GPT-3.5 to mitigate these risks. Llama2 uses encryption and anonymization techniques to protect patient data, while GPT-3.5 employs advanced real-time bias monitoring to ensure fairness in patient care decisions.

Table 4 Comparative evaluation of GPT-3.5 and Llama2 based on user satisfaction in healthcare applications

Model	Dental health	Medical health	Mental health
GPT-3.5	4.6	4.7	4.5
Llama2	4.3	4.4	4.2

8 Conclusion

This study comprehensively analyzed the performance of GPT-3.5 and Llama language models in healthcare applications focusing on dental, medical, and mental health queries. The findings demonstrated that GPT-3.5 generally outperforms Llama across all metrics, including accuracy of responses, response time, and user satisfaction ratings. Specifically, GPT-3.5 exhibited strengths in generating detailed treatment plans and therapeutic assistance, which are critical in clinical support systems. Despite its slightly lower performance, Llama was appreciated for its diagnostic capabilities and its ability to provide concise, actionable patient management advice.

The faster response times of GPT-3.5 enhance its usability in real-time clinical decision-making. User satisfaction also reflected a preference for comprehensive and detailed replies, aligning with the higher complexity of needs in healthcare settings. Ethical considerations remain paramount as these models are integrated more deeply into clinical workflows. The study highlighted the proactive measures taken by both models to mitigate risks such as data security breaches and biases in decision-making processes. GPT-3.5's real-time bias monitoring and Llama's data anonymization techniques are steps toward responsible AI use in sensitive areas like healthcare. In conclusion, while both GPT-3.5 and Llama show promise for enhancing healthcare services through AI, continuous improvements in model accuracy, response efficiency, and ethical standards are essential for their sustainable integration. These technologies must be developed with a clear focus on augmenting human capabilities and ensuring equitable, secure, and effective patient care.

References

1. Kenneth W Goodman (2015) *Ethics, medicine, and information technology: intelligent machines and the transformation of health care*. Cambridge University Press
2. Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, Jacob Steinhardt (2020) Measuring massive multi-task language understanding. arXiv preprint. [arXiv: 2009.03300](https://arxiv.org/abs/2009.03300)
3. Krittanawong C, Zhang HongJu, Wang Z, Aydar M, Kitai T (2017) Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol* 69(21):2657–2664
4. David D Luxton (2014) Artificial intelligence in psychological practice: current and future applications and implications. *Prof Psychol Res Pract* 45(5):332
5. Thomas M Maddox, John S Rumsfeld, Philip RO Payne (2019) Questions for artificial intelligence in health care. *Jama* 321(1):31–32
6. Reddy S, Allan S, Coghlan S, Cooper P (2020) A governance model for the application of AI in health care. *J Am Med Inform Assoc* 27(3):491–497
7. Benjamin Shickel, Patrick James Tighe, Azra Bihorac, Parisa Rashidi (2017) Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J Biomed Health Inf* 22(5):1589–1604
8. Eric J Topol (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25(1):44–56
9. Alfredo Vellido, José David Martín-Guerrero, Paulo JG Lisboa (2012) Making machine learning models interpretable. In: *ESANN*, vol. 12. Citeseer, pp 163–172

10. Stephen F Weng, Jenna Reys, Joe Kai, Jonathan M Garibaldi, Nadeem Qureshi (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One 12(4):e0174944
11. Alsentzer E, Murphy JR, Boag W, Weng WH, Jin D, Naumann T, McDermott M (2019) Publicly available clinical BERT embeddings. arXiv preprint arXiv:1904.03323
12. Zhang X, Dou D, Wu J (2020, April) Learning conceptual-contextual embeddings for medical text. In: Proceedings of the AAAI conference on artificial intelligence, vol 34. no 05. pp 9579–9586
13. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, ... Dean J (2018) Scalable and accurate deep learning with electronic health records. NPJ Digit Med 1(1):18
14. Gehrmann S, Dernoncourt F, Li Y, Carlson ET, Wu JT, Welt J, ... Celi LA (2018) Comparing deep learning and concept extraction based methods for patient phenotyping from clinical narratives. PloS one 13(2):e0192360
15. Vallabhaneni GV, Rahul YS, Kumari KS (2024, February) Improved lung cancer detection through use of large language systems with graphical attributes. In: 2024 IEEE international conference on computing, power and communication technologies (IC2PCT), vol 5. IEEE, pp 1849–1854

Harvesting Insights: A YOLOv5 Template Design for Real-Time Pesticide Recommendation with Reduced Computational Complexity



Swati Kapoor, Ipsita Majhi, Gyan Prakash Dash, Sushruta Mishra, and Haider Alabdeli

Abstract Through the improvement of resource usage and avoidance of environmental impact, precision agriculture technology represents an appealing chance to transform the sector of agriculture. Yet weed competition is an important challenge to agricultural sustainability, leading to lower crop yields as well as increased pesticide use. Conventional techniques for identifying weeds and diseases are frequently slow, imprecise, and ineffective. This paper explores the application of YOLOv5, a cutting-edge deep learning model, as an alternative real-time weed identification method. Our tailored algorithm commonly efficiently recognizes specific weeds, facilitating focused approaches to management and reducing the requirement for herbicides. Furthermore, the model demonstrates competence for recognizing varied plant diseases, hence enabling prompt intervention tactics. For farmers, the computerized labeling procedure promises enhanced efficiency and resource conservation. YOLOv5 performs more precisely than conventional approaches and future-proof agriculture.

Keywords TIA YOLOv5 · Weed detection · Precision agriculture · Fertilizer recommendation · Pests

S. Kapoor · I. Majhi · G. P. Dash · S. Mishra (✉)
Kalinga Institute of Industrial Technology, Deemed to Be University, Odisha, India
e-mail: sushruta.mishrafcs@kiit.ac.in

H. Alabdeli
Department of Computers Techniques Engineering, College of Technical Engineering,
The Islamic University, Najaf, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering,
The Islamic University of Al Diwaniyah, Al Diwaniyah, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering,
The Islamic University of Babylon, Babylon, Iraq

1 Introduction

Precision agriculture is an advanced agricultural technique that gathers and evaluates data on crops and soil through technology. With the utilization of this data, farmers have greater ability for making choices on the most efficient farming management techniques such as where to plant, what quantity of water to use, and how to fertilize. Precision agriculture's primary objectives are to enhance crop yields, reduce costs, and support sustainable food production strategies. However, controlling disease and weeds tends to be one of the most significant barriers in conventional agricultural methods. Precision farming aims to tackle these issues, but manual weed and detection of diseases may be onerous and error prone. This paper explores the utilization of YOLOv5 [1], a deep learning model, for real-time disease and detection of weeds in precision agricultural applications. Our approach using YOLOv5 model enables accurate weed identification. We also propose a rule-based fertilizer recommendation system to optimize resource allocation and improve environmental sustainability. Automatic detection saves farmers valuable time and promotes efficient use of resources. Precision agriculture can currently successfully overcome the shortcomings of conventional techniques for controlling weeds by applying data-driven management practices and targeted application of herbicides generated by deep learning, particularly with its assistance of the YOLOv5 model. By encouraging comprehensive contrasts with conventional approaches, this study seeks to get us closer to a moment when food production is sustainable for the environment [2]. A key factor in determining the sustainability of precision agriculture systems is life cycle assessment or LCA. This research aims to promote thorough comparisons with traditional approaches by discovering impediments and implementing best practices into actuality, so strengthening the cause of environmentally conscious food production in the decades to come.

The main contribution of this paper is to suggest an effective and appropriate measure to the growing problems faced by agricultural fields nowadays,

- Enhanced precision when determining weeds and diseases as opposed with conventional methods: This is sometimes quantified with strategies such as $F1$ -score or mean average precision (mAP), which demonstrate the efficiency of YOLOv5 in correctly identifying an assortment of weeds and diseases.
- Capabilities for real-time detection: This makes it achievable to put control measures into place immediately and to respond early, which might enhance treatment outcomes and stop subsequent development.
- Enhanced decision informed by data: Better decisions about strategies for treatment, crop monitoring, and the assignment of resources may be achieved with the application of real-time detection data, which could end in improved yields and monetary advantages.

2 Literature Survey

Precision agriculture (PA) is a farming management approach that uses technology to gather data about crops, soil, and other factors. This data is then used to make informed decisions about how to manage the farm more effectively. The conventional scheme is based on the imaginary five crop rotation system experienced in Lower Austria arable land. This paper uses a de-nitrification and decomposition soil model which reduces global warming by -17.04% [3]. Agronomic precision agriculture involves gathering and placing data on crop growth and conditions through a variety of sources such as satellite, sampling of soils, and yield. The relevance and scope of regulatory framework protecting the rights and dignity of farmers is discussed in this paper which highlights data ownership rights and data privacy of farmers [4]. Fertilizer recommendation using CNN and TPF with nutrient analysis and implementing soil sensor [5]. Pest detection methods based on deep learning-based approach applies to small object detection algorithms like YOLOv5m and deep convolutional neural network [6]. Precision agriculture has revolutionized modern ways to grow crops keeping an eye on crop fertilization, proper irrigation and disease management. This paper deals with challenges faced by potato farmers, improves the overall quality, and reduces the production cost using remote sensing and crop management [7]. This paper deals with sustainable agriculture with IoT-based intelligent machinery where data is analyzed via sensors, controllers, and communication standards to meet the growing demands for water and raw materials like fertilizers and pesticides. It uses technologies like artificial neural network (ANN), conventional neural network (CNN), global navigation satellite system (GNSS), and graphical user interface (GUL) and other parameters like microcontroller units (MCU) and long-term evolution [8]. The concept of integrating digital technologies to enhance future smart agricultural systems with data governance, agricultural data integration, block chain, and various decision support systems is performed in this paper [9]. On-going and emerging links with soil science are also introduced [10]. Weed detection is an on-growing problem faced by a lot of crops which impacts the farmers, weed detection, and weed localization with CNN, computer vision is the main focus of this paper, and multiple formats consist of annotations exported in multiple formats such as TXT, XML, and YAML with collected greenhouse dataset on varied levels of background and lighting conditions using DJI Phantom 4 Pro to collect in field data [11]. Wireless sensor network plays an vital role in precision agriculture, the researcher has dealt with a new multi-objective clustering approach which works for grouping the sensor nodes of WSNs, and the techniques are Aquila Optimizer, electron-based optimization algorithm, and optimized CNN with accuracy of 99.23% [12]. This aims to attempt a clear panorama of optoelectronic proximal sensing with image acquisition resulting in heavy signal processing for its final implementation [12]. Weed detection using YOLOv5 for real-time crop prediction uses data augmentation and network architecture approach [13]. TIA-YOLOv5-based weed detection uses conventional schemes and modified architecture with pixel-level synthesis of sugar-beet dataset and testing results [14].

3 Proposed Methodology

Weeds' enormous growth is a significant threat to agricultural production globally, causing tremendous losses in production and economic difficulties. Object recognition assisted by deep learning has tremendous prospects for precision farming. The potential of YOLOv5 for real-time weed detection is investigated in this work. Developing an algorithm to recognize crops and weeds is a great help for managing weeds on the farm and providing creative assistance for farming methods. In the fields of crops and agriculture, YOLO series networks have been widely used for recognizing objects and image recognition. YOLOv5, the most current version, is primarily employed for image labeling. A pesticide management system is an assortment of proposals or management techniques based upon research and meant to assist farmers in identifying the most efficient pesticides for the crops they grow. The program's primary goal is to improve production from agriculture while minimizing deleterious consequences for the environment. Yet, a significant number of farmers have little to no in-depth comprehension of their current nutrient levels in their soil, which renders it challenging to identify nutrient abundances, shortages of goods, or optimal values. Small-scale farmers frequently experience high expenses connected to professional soil testing, especially during countries with poor infrastructure [15–17]. There could not be sufficient resources for assessing and acting upon the results of evaluations, even though they are cost-effective. This study suggests a rule-based pesticide recommendations system as an alternative to these difficulties, assisting farmers select the most effective varieties and levels of pesticides for their specific situation. The input taken by the rule-based classifier is the image generated from the YOLOv5.

The parameters that are considered by the classifier are as follows:

- Shape/length of the pest
- Group/single attack of the pest
- Design on the leaf
- Color of the leaf.

Crop pests offer significant obstacles to the economic viability of the agricultural sector as shown in Fig. 1:

- Aphids: These pests may propagate viral diseases and feed on the sap underlying plants. Insects can be damaging to crops including potatoes, cabbage, mustard, peas, tomatoes, soybeans, cotton, and peaches.
- Bollworms: Known for generating considerable crop damage, bollworms infest a variety of grain crops that include maize, sorghum, chickpeas, and pulses, as well as cotton, tomatoes, and soybeans.
- Green stinkbugs: A substantial yield reduces function, these pests harm cotton, maize, and soybeans owing to extremely dense values.
- Leaf folder: Typically found in humid areas where rice is grown, leaf folders cause rice harvests to yield less.

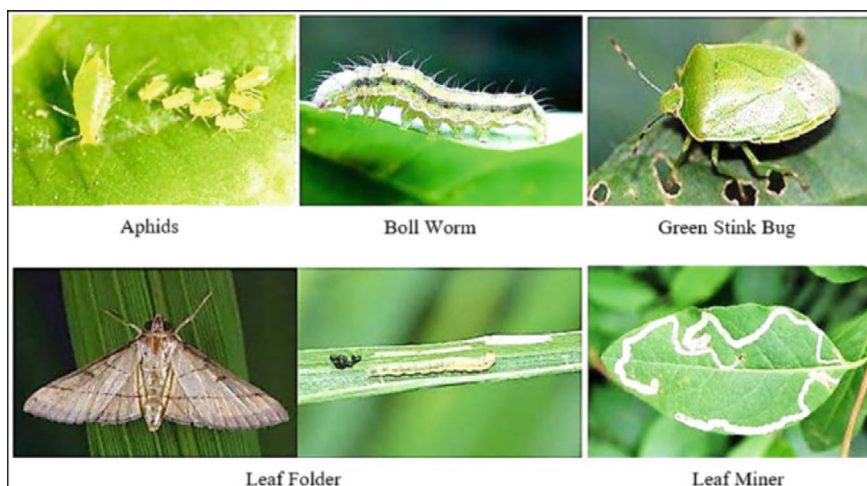


Fig. 1 Sample images of pest classes

- Leaf miners: Plant degradation can be triggered by leaf miners, a group of minuscule larvae that dig into leaves. They constitute an obstacle to plant growth and quality by infesting the leaves of tomatoes, cucumbers, and melons.

This study is utilized the publicly available dataset from “HOTSON HONET” [18]. The main purpose of this dataset is to facilitate research in target recognition under high obstruction surrounding [19]. The data comprises 2072 rows $\times$ 6 columns each featuring image_id, path, bbox label width, and height. This dataset is allocated into training, validation, and test sets with 2600 images of crops.

The following steps are involved in the process:

- Image Assessment: Every image has distinctive characteristics from both the physical and insecticide perspectives of belief, making it distinguish forth in the computational methodology.
- Rule Establishment: The system develops rules to guarantee trustworthy recommendations based on expert knowledge and certified scientific facts. The archive incorporates data on five different pest species: leaf folders, leaf miners, green stink bugs, bollworms, and aphids [1, 20].
- Cross-referencing: In order to determine whether pesticide would be appropriate for the situation at hand, the system connections the dataset chosen with the user’s input and pesticide metadata.

The above model in Fig. 2 recommends pesticide using YOLOv5 and rule-based classifier algorithm. The following are the steps to understand it briefly.

1. Fields of crops and agriculture, YOLO series networks have been widely used for recognizing objects and image recognition. YOLOv5, the most current version, is primarily employed for image labeling.

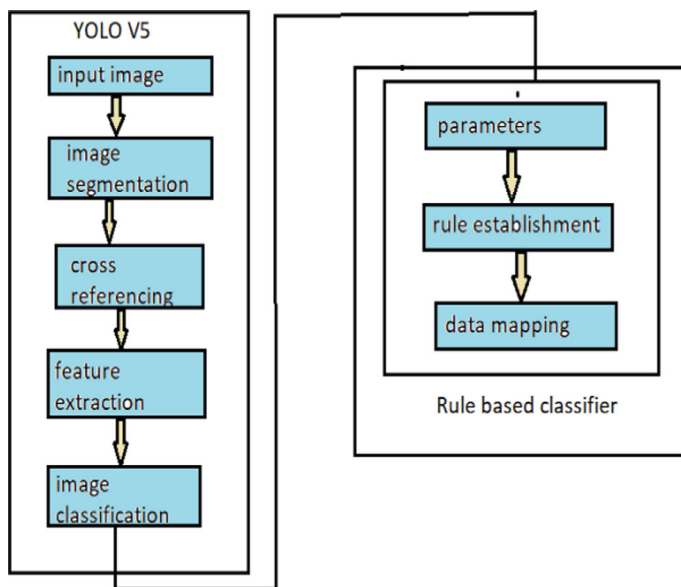


Fig. 2 Proposed model using TIA-YOLOv5 with four components

2. The YOLOv5 takes the input image and then it converts image into a collection of regions of pixels that are represented by a mask or a labeled image [21].
3. In order to determine what would be appropriate for the situation at hand, the system connects the dataset chosen with the user's input.
4. Features of the image are extracted that are viable for the situation, and then, the image classification is given for the weeds and pest.
5. The input taken by the rule-based classifier is the image generated from the YOLOv5.
6. The parameters that are considered by the classifier are as follows:
 - Shape/length of the pest
 - Group/single attack of the pest
 - Design on the leaf
 - Color of the leaf.
7. Following the parameter is the rule establishment: Using certified scientific data and expert knowledge, the system creates rules to ensure reliable suggestions [22]. Data on five distinct pest species is included in the archive: aphids, bollworms, green stink bugs, leaf folders, and leaf miners.
8. Then it is mapping with the data to check with the organic/inorganic pesticides to recommend the farmers. Based on the data that is recommended the farmer has the access to the percentage of organic/inorganic pesticides to be used for increasing the efficiency of the field.

Figure 3 depicts the context aggregation block. The dimensions are used for denoting combinations of function mappings for training. For instance, $C \times H \times W$ depicts a matrix of avenues C, height H, and width W. The abbreviations $\otimes$ and $\odot$ stand for the transmitted Hadamard product and batch matrix multiplication, respectively. The blue, red, and green labels correspond to the convolution layers that produce attention maps, map features, and refine context. In the present study, we update the $1 \times 11 \times 1$ convolution in the neural network’s head using the context aggregation module. This module requires three primary steps to function: Attention Weight Calculation: Firstly, a 1×1 convolution layer and a sigmoid activation function are applied to calculate attention weights. It additionally makes sure that priority weights are automatically adjusted based on importance, starting at zero to one.

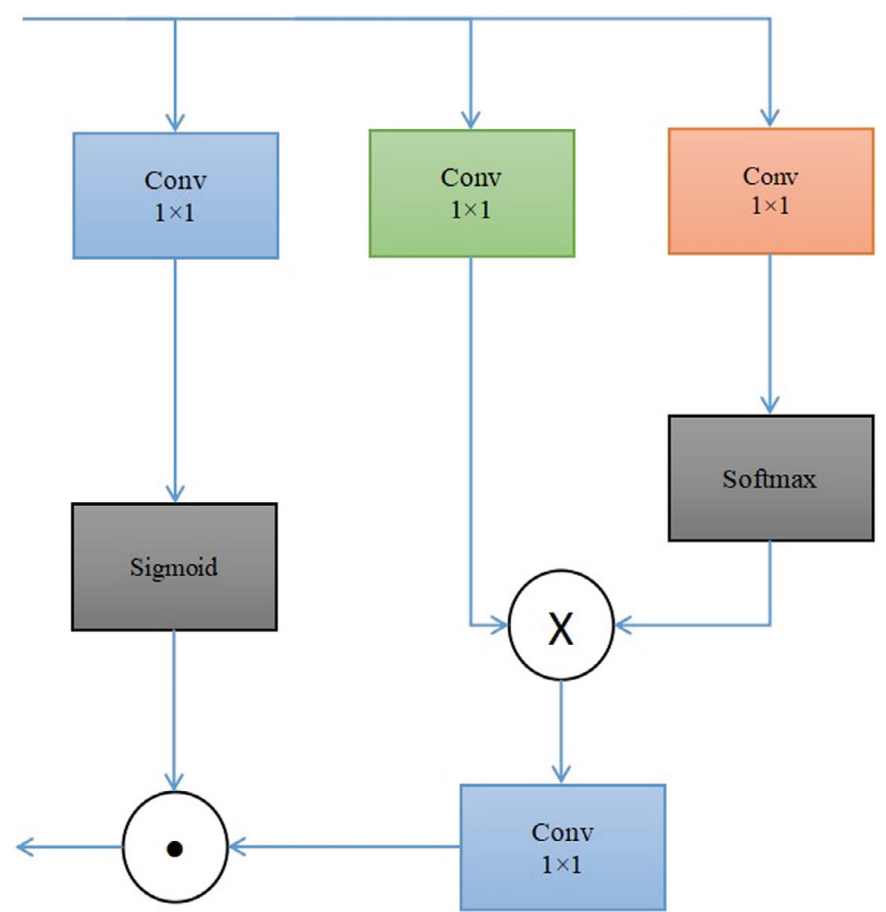


Fig. 3 Context aggregation block

Contextual Information Fusion: A matrix multiplication operation is carried out among the characteristic value and the attention weight after receiving attentive weights and feature values. The algorithm is more competent to establish the relation between objects and backdrops based on its significance of qualities in different parts owing to this approach, creating an adaptive weighted fusion of context-relevant data. In particular, this fusion methodology performs exceptionally well in integrating characteristics at different dimensions, which enhances the whole efficacy of object detection.

Feature Conversion and Output: To enhance their representational aptitude, the weighted features are subsequently processed across a $1 \times 11 \times 1$ convolution layer. These properties are the final output aspects after they have undergone evaluated by context-based fusion. Furthermore, in the identification of weeds difficulty, these upgraded characteristics show enhanced discerning capabilities between targets and environments. An adaptive integrating of contextual information across multiple scales has been rendered feasible by the introduction of the context aggregate module into YOLOv5 [23, 24]. The resulting enhancement significantly boosts the model’s performance and makes recommendations for relevant pesticides in besides enabling it to handle obstacles such varied scenery and target size variations that develop during weed detection responsibilities and disease detection.

4 Result Analysis

Different deep learning strategies are illustrated in the comparison table or Table 1, with particular focus on metrics like precision, *F1*-score, recall, weights, and mAPs. The model with the greatest accuracy is TIA-YOLOv5, that is particularly renowned for its ability to spot small objects in many scenes. The ability it has to identify tiny modifications in the atmosphere makes early-stage weed analysis conceivable.

The precision–recall (PR) curve for YOLOv5s is illustrated in Fig. 4, by the accuracy-weight curve and the Params (M)-GFLOPs curve. Utilizing the application of computational cost represented by giga floating-point operations (GFLOPs) and the quantity of parameters (Params), this graphical representation reveals a perspective on the model’s effectiveness across various degrees of model complexity. The

Table 1 Comparison among different object detection algorithms

Model	mAP@0.5	mAP@0.5:0.95	Recall	<i>F1</i> -score	Precision
YOLOv4-tiny	50.68	30.69	76.2	83.02	87.09
Faster-RCNN	53.46	36.1	83.1	85.77	89.97
SSD	46.48	31.2	86.9	88.96	92.76
YOLOx	74.84	53.7	89.4	90.11	94.53
YOLOv5s	76.3	53.5	92.5	93.19	96.12
TIA-YOLOv5	75.63	52.4	94.4	96.90	98.78

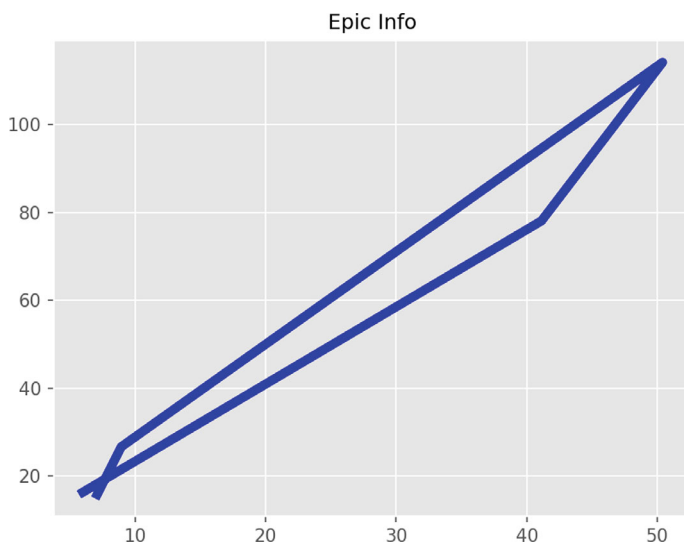


Fig. 4 Params (M)-GFLOPs curve along with accuracy-weight curve

accuracy-weight curve illustrates where the accuracy of the model varies in relation to the network's distribution of weight. When taken as a whole, these curves enable an in-depth overview of the trade-offs between the fundamental YOLOv5s' model complexity, computation efficiency, and detection accuracy.

The parameters (M)-GFLOPS curve is displayed in the above picture, emphasizing the association between accuracy and a low false positive rate. Recall that information is displayed on the X-axis, while accuracy is shown on the Y-axis. We examined the model's efficacy using a number of indicators throughout the training phase, with IOUs set at 0.5 (50%) and 0.95 (95%), among those being precision, recall, and mAP (average precision). After assessment, which is the YOLO model received recall scores of 0.95, validation accuracy values of 0.8057, and mAP scores of 0.95 for @0.5IOU and 0.64 for @0.95IOU, respectively [18].

The aforementioned visual presentation in Fig. 5 combines images of weed-free cultivation and weed-filled agricultural fields, with weeds shown in blue and harvest zones in green. For simplicity in identification, each image is labeled with a distinct image ID [19]. TIA-YOLOv5's enhanced small object detection skills enable it to distinguish between grass with greater precision. This accurate identification could contribute to a revolution in weed management techniques as well as more efficient and environmentally friendly farming methods.

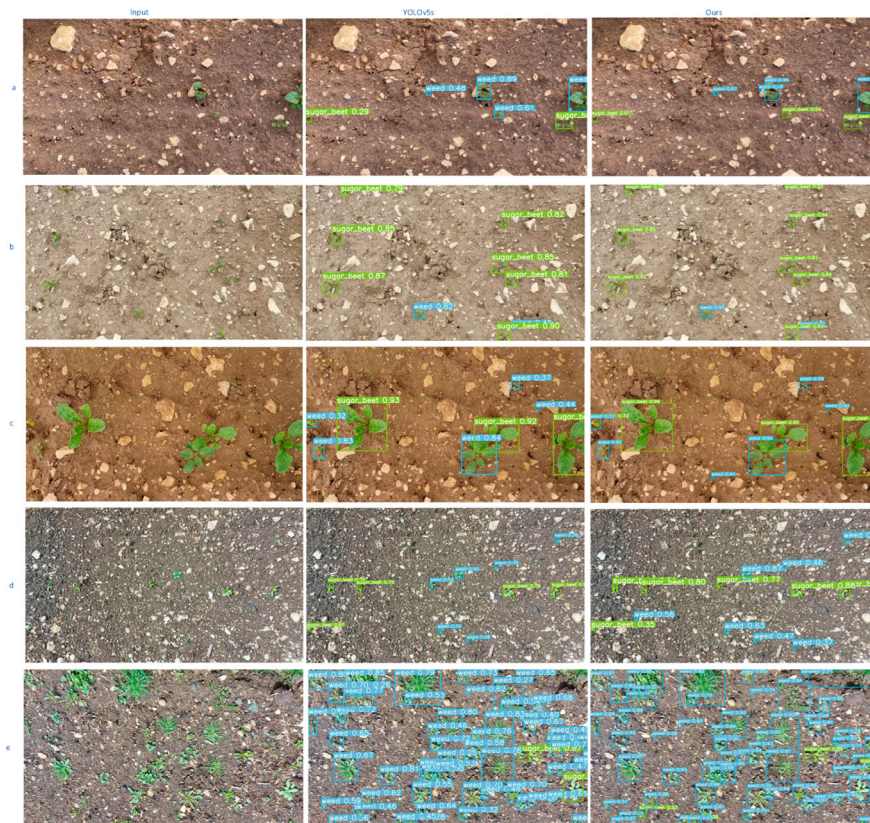


Fig. 5 Comparison of detection results in different environments with different soil conditions

5 Conclusion

To sum up, this study examined the complicated interactions that exist between conventional agricultural practices, precision agriculture technologies, and life cycle assessment (LCA). Following performing an in-depth literature research and assessing empirical data, a variety of significant observations were made. Our customized YOLOv5 model exceeds industry standards in weed and disease identification with remarkable accuracy. This is an important advancement since, due to the model's comprehensive detection of diseases capabilities, precision weed management additionally eliminates the adverse impact on the environment and minimizes the requirement for pesticides, but it additionally renders early disease intervention simple and maintains optimal crop health. The application of automation enhances this revolutionary manipulate, conserving farmers valuable time and resources. Through minimizing the use of chemicals, boosting revenue, and promoting an improved agricultural environmentalism, these efforts contribute to building a more

sustainable future. YOLOv5 appears not just as a tool, but also as a powerful seed, with the promise for a more intelligent, sustainable, and efficient agricultural future. By incorporating pesticide and weed advising systems, the recommended strategy permits farmers to efficiently maximize the production of agriculture. YOLOv5 and rule-based classifiers are utilized to substantially enhance the accuracy and precision on the model. There are numerous benefits associated with employing a rule-based classifier since it is an AI model that is straightforward to grasp and analyze. By establishing rules based on in-depth expertise by humans, developers empower the system to consume forthcoming data and offer outputs quickly. Furthermore, nutrient-rich soil may be generated through meticulously regulating soil nutrient levels. This technique lowers the expense of assessing soil nutrients in institutions as well. The proposed strategy makes it possible to recommend insecticides rapidly, which makes it beneficial for both online and offline scenarios. Farmers, soil testing, and seed hybridization enterprises may make a profit from this strategy. The system can generate accurate, and thorough pesticide recommendations without explicitly building a predictive model on a dataset due to the integration of TIA-YOLOv5 address these issues. The commonality of the suggested amount of nutrients for any agricultural species is the foundation of this system. Analogously, AI-powered innovations might revolutionize weed identification and control, promoting lucrative and sustainable farming methods. AI presumes a crucial role in ensuring responsible and productive food production for a growing global population.

References

1. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digit Health* 10:20552076241256732
2. Jain S, Tripathy HK (2024) Machine learning methods for the timely identification of autism spectrum disorder in toddlers. In: 2024 International conference on emerging systems and intelligent computing (ESIC). IEEE, pp 515–520
3. Medel-Jiménez F, Krexner T, Gronauer A, Kral I (2024) Life cycle assessment of four different precision agriculture technologies and comparison with a conventional scheme. *J Clean Prod* 434:140198. ISSN 0959–6526
4. Idowu AR, Wachenheim C, Hanson E, Sickler A (2023) The disposition of data from precision agricultural technologies: what do young agriculturalists think? *Technol Soc* 75:102389. ISSN 0160–791X
5. Thorat T, Patle BK, Kashyap SK (2023) Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming. *Smart Agric Technol* 3:100114. ISSN 2772–3755
6. Dai M, Dorjoy MMH, Miao H, Zhang S (2023) A new pest detection method based on improved YOLOv5m. *Insects* 14(1):54. <https://doi.org/10.3390/insects14010054>. PMID: 36661982, PMCID: PMC9863093
7. Széles A, Huzsvai L, Mohammed S, Nyéki A, Zagyi P, Horváth É, Simon K, Arshad S, Tamás A (2024) Precision agricultural technology for advanced monitoring of maize yield under different fertilization and irrigation regimes: a case study in Eastern Hungary (Debrecen). *J Agric Food Res* 15:100967. ISSN 2666–1543

8. Pathmudi VR, Khatri N, Kumar S, Abdul-Qawy ASH, Vyas AK (2023) A systematic review of IoT technologies and their constituents for smart and sustainable agriculture applications. *Sci Afr* 19:e01577. ISSN 2468-2276
9. Gebresenbet G, Bosona T, Patterson D, Persson H, Fischer B, Mandaluniz N, Chirici G, Zacepins A, Komasilovs V, Pitulac T, Nasirahmadi A (2023) A concept for application of integrated digital technologies to enhance future smart agricultural systems. *Smart Agric Technol* 5:100255. ISSN 2772-3755
10. Taylor JA (2023) Precision agriculture. In: Goss MJ, Oliver M (eds) *Encyclopedia of soils in the environment*, 2nd edn. Academic Press, pp 710–725. ISBN 9780323951333
11. Rai N, Mahecha MV, Christensen A, Quanbeck J, Zhang Y, Howatt K, Ostlie M, Sun X (2023) Multi-format open-source weed image dataset for real-time weed identification in precision agriculture. *Data Brief* 51:109691. ISSN 2352-3409
12. Pandiyaraju V, Ganapathy S, Mohith N, Kannan A (2023) An optimal energy utilization model for precision agriculture in WSNs using multi-objective clustering and deep learning. *J King Saud Univ Comput Inf Sci* 35(10):101803. ISSN 1319-1578
13. Barnes AP, Soto I, Eory V, Beck B, Balafoutis A, Sánchez B, Vangeyte J, Fountas S, van der Wal T, Gómez-Barbero M (2019) Exploring the adoption of precision agricultural technologies: a cross regional study of EU farmers. *Land Use Policy* 80:163–174. ISSN 0264-8377
14. Tang JL, Wang D, Zhang ZG, He LJ, Xin J, Xu Y (2017) Weed identification based on K-means feature learning combined with convolutional neural network. *Comput Electron Agric* 135:63–70. ISSN 0168-1699
15. Kashyap P, Pareek A, Mishra S, Khan Z, Garg R, Tripathy HK (2024) Sentiment polarity analysis of twitter data using machine learning models. In: *International conference on innovative computing and communication*. Springer Nature Singapore, Singapore. pp 623–635
16. Pradhan SR, Mishra S, Tripathy HK, Brahma B, Gobinath R, Solti R (2024) Critical application feasibility of predictive learning in autonomous vehicles. In: *International conference on innovative computing and communication*. Springer Nature Singapore, Singapore, pp 371–383
17. Chakraborty S, Mishra S, Tripathy HK (2022) COVID-19 outbreak estimation approach using hybrid time series modelling. In: *International conference on innovations in intelligent computing and communications*. Springer International Publishing, Cham, pp 249–260
18. Mohanty S, Behera A, Mishra S, Alkhayyat A, Gupta D, Sharma V (2023) Resumate: a prototype to enhance recruitment process with NLP based resume parsing. In: *2023 4th International conference on intelligent engineering and management (ICIEM)*. IEEE, pp 1–6
19. Subhra S, Mishra S, Alkhayyat A, Sharma V, Kukreja V (2023) Climatic temperature forecasting with regression approach. In: *2023 4th International conference on intelligent engineering and management (ICIEM)*. IEEE, pp 1–5
20. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
21. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-DETECTOR: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
22. Pranjal P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. In: *AIP conference proceedings*, vol 2919(1). AIP Publishing
23. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system. *Internet Things* 24:100978
24. Mishra S, Volety DR, Bohra N, Alfahhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306

Segmentation of Tapioca Plants Point Cloud Data from Irregular Ground Surfaces



S. Prince Sahaya Brighty, R. Anuradha, and M. Sandeep

Abstract Computer vision techniques that focus on segmenting point clouds have become increasingly popular in agriculture for tasks such as studying plant characteristics and growth patterns of crops in fields overtime using three-dimensional data instead of traditional two-dimensional data can provide more detailed and useful information to scientists. However, due to the scarce availability of point cloud datasets, there are challenges that need to be addressed through improvements in how data is collected and processed This study introduces a method that compares and efficiently separates tapioca plants, from uneven ground surfaces. We obtained real-time data from fields using photogrammetry and then processed it with Cloud Compare software to improve the quality and efficiency of the point cloud data by reconstructing the data and removing noise through down sampling techniques. In order to distinguish tapioca plants, from the soil surface effectively, we utilized the Random Sample Consensus (RANSAC) algorithm that accurately identified and classified the points. Moreover, we utilized RGB alteration to assign colors to individual points simplifying visual inspection and allowing for a clear differentiation, between the tapioca plants and the nearby terrain. The suggested approach showcases efficiency by successfully accomplishing precise segmentation and color assignment of the point cloud data. Consequently, the accurate recognition and categorization of tapioca plants and the adjacent ground surface can be achieved effectively.

Keywords Point cloud data · Segmentation · Photogrammetry · Computer vision

S. Prince Sahaya Brighty (✉) · R. Anuradha · M. Sandeep
Department of Computer Science and Engineering, Sri Ramakrishna Engineering College,
Coimbatore, Tamilnadu, India
e-mail: brighty.s@srec.ac.in

R. Anuradha
e-mail: anuradha.r@srec.ac.in

M. Sandeep
e-mail: sandeep.2101218@srec.ac.in

1 Introduction

Segmenting point clouds is a job that includes organizing three-dimensional point clouds into different sections where the points in each section have similar features. The task comes with difficulties such as a lot of information varied sampling density and no clear structural data in the point cloud data [1]. Deep learning techniques on 3D point clouds still face significant hurdles, like small dataset sizes having high dimensions and dealing with the unorganized nature of point cloud data. In light of these challenges, this paper aims to segment the tapioca plants and classify them using deep learning techniques by processing 3D point clouds [2]. The field of deep learning on point clouds has garnered increasing attention, particularly within the past five years. Datasets are the fuel for their learning phase, but point cloud data is very difficult to preprocess and segment the data to extract the features. Point cloud data research has become very popular since the beginning of autonomous cars and geometrical reconstruction. Further, they are also used in agriculture research purposes for plant growth analysis and plant phenotyping [3–5].

A comprehensive taxonomy of the current deep learning methods for 3D point clouds is presented in Fig. 1 [6].

Point cloud segmentation algorithms have been successfully applied to various crops, such as maize and wheat plants [7, 8]. Semantic segmentation of tapioca plants stances unique challenges because of their complex and intricate structures, dense foliage, overlapping leaves, and complicated branching patterns. Some specialized algorithms are needed to handle the dense and intertwined nature of tapioca plant point cloud data. Advanced techniques, like combining deep learning with domain-specific knowledge and additional sensor modalities like RGB imagery or 3D imaging, can improve the segmentation accuracy [5, 9]. Developing annotated

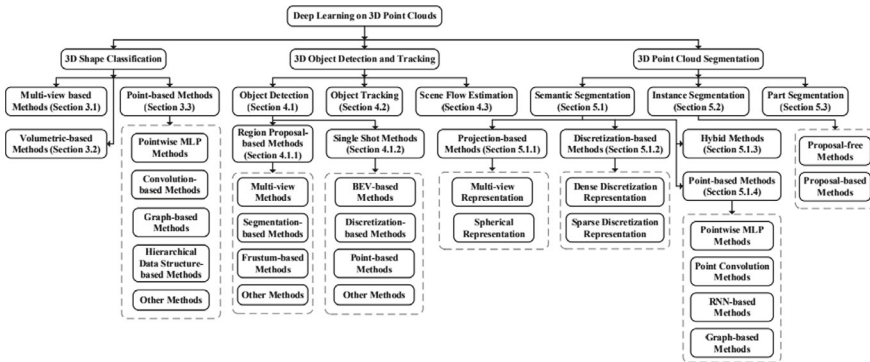


Fig. 1 Classification of deep learning methods for 3D point clouds

datasets precisely for tapioca plants is critical for training and evaluating custom-made segmentation models. Overall, tapioca plant segmentation involves specialized algorithms, sensor integration, and dedicated datasets to overcome the unique challenges presented by their complex structures [10].

2 Related Works

A vast amount of research papers has been published to solve the segmentation problem on point cloud data with deep learning networks. Qi et al. presented the PointNet architecture that offered a new deep learning framework for directly processing point sets in 3D classification and segmentation problems [11]. Their approach achieved the best performance on point clouds by encoding both local and global geometric features. Inspired by PointNet, Qi et al. presented a hierarchical feature learning framework for point sets: PointNet++ [12]. We introduced spatial partitioning and local feature aggregation to consolidate the power of PointNet, known as PointNet++, which surpassed other methods in terms of performance when it comes to point cloud classification or segmentation.

Çiçek et al. investigated the application of this network to medical imaging [3], a field where transformations need to be invertible. First, was the three-dimensional U-Net architecture introduced by Cicek et al. designed for dense volumetric segmentation from sparse annotation [13]. The proposed approach exploited the power of an encoder–decoder network to exploit 3D context efficiently and highly perform in medical image segmentation tasks. Li et al. developed a new neural network called PointCNN for directly addressing the transformed points rather than conventional grid representations [13]. The method was shown to successfully learn and capture local structures of point clouds, leading to better results in point cloud-based classification and segmentation.

Notably, this problem has been also confronted by the RandLA-Net architecture presented in Li et al. [14], that made an important contribution to segmenting large-scale point clouds efficiently. By using a range-adaptive convolutional layer, they capture both context of local and global contexts—by which semantic segmentation on ultra-large point cloud datasets is achieved in much higher accuracy but less computational cost.

Together, these papers serve as a testament to the progress made in deep learning-based point cloud segmentation. The use of advanced architectures and methods like PointNet, PointNet++, D U-Net, PointCNN, and RandLA-Net has largely increased the accuracy as well as time efficiency of point cloud segmentation on many aspects.

3 Proposed Work

In the analysis of a frame structure, 3D mapping using photogrammetry has been shown to be advantageous and is ever more popular in recent years. Yet accurate segmentation of point cloud data is one such challenging task, particularly in the case of complex 3D scenes.

In this study, we present a virtual approach to outline segmentation of point cloud data from digital photograph photogrammetry-based 3D images for the case of tapioca plants. We were able to gather the necessary data using a standard smartphone camera, and this is where photogrammetry proved an inexpensive alternative to more elaborate lidar-based approaches. We pre-processed the data to clear noise, and turned 3D mesh model into a point cloud format. To perform segmentation, we used the RANSAC algorithm which is able to handle noisy, obstructed or clumped point clouds even in difficult scenarios.

Our method and results suggest possible uses of the technique for other studies pertinent to this process such as growth and development of tapioca plants. The detailed methodology for data collection, processing, and analysis was described with the publication of our results focusing on the importance of this approach in a case study involving tapioca plant research. The study is significant because these results are foundation for the tapioca plant growth investigation that would follow next.

A demonstration of the output results after our segmentation process is shown in Fig. 3 with geographical coordinates on a Canopy [3] plant. We hope to help in understanding the growth and development of tapioca plant by using photogrammetry along with a customized segmentation approach. Three different algorithms were have been tried for segmenting tapioca plants: the Ball Pivot algorithm, the DBSCAN algorithm, and the RANSAC algorithm. After a performance comparison, it was resolute that the RANSAC algorithm unveiled the best performance and was therefore selected as the fitting algorithm for the segmentation model.

3.1 *Ball Pivot Algorithm*

Implementation of ball pivot algorithm is useful for the surface reconstruction and segmentation point cloud data as it used in many applications. ESF has been demonstrated to work well in the presence of unorganized and sparse point clouds for which an explicit surface representation does not exist. At a high level, the ball pivot algorithm attempts to iteratively generate three-sided facets that approximate the original surface for some generated set of points. Initially, the algorithm randomly selects a seed point from the cloud as a starting pivot. This is the pivot point, and at this location, a sphere with defined user radius as being called by a name of Pivot Ball is centered [15].

Then, the algorithm looks for a neighboring point of pivot ball on surface to be reconstructed. The distances from the pivot ball to all other points in cloud are calculated as follows. If a valid point is found in the search radius, a new triangle will be generated from connecting all 3 points.

The corresponding newly created triangle is added to the reconstructed surface, and the center of its pivot ball further will be changed in order to embrace this triangle. Now it is the turn of the other surface point to search for itself next iteratively within that pivot ball. This flow continues until either no surface points are found in the search radius or an early termination. Figure 2 shows a plant point cloud with and without ball pivot-based segmentation.

Algorithm 1:

1. *Select a seed point p_{seed} from the point cloud.*
Set r as the radius of the pivot ball.
2. *Create an empty list that will store the reconstructed surface triangles.*
3. *While there are still points in the point cloud:*



Fig. 2 Before and after segmentation using ball pivot algorithm



Fig. 3 Before and after segmentation using DBSCAN algorithm

- Iterate through all the points p in the point cloud.
- Calculate the Euclidean distance d between the pivot point p_{pivot} and the current point p :

$$d = \sqrt{(p.x - p_{\text{pivot}.x})^2 + (p.y - p_{\text{pivot}.y})^2 + (p.z - p_{\text{pivot}.z})^2}$$

If $d \leq r$, where r is the pivot ball radius r ,

Find a third point p_{third} such that :

$p_{\text{third}} = \text{findPointWithinBall}(p_{\text{pivot}}, p, r)$

4. Create a triangle t using the pivot point p_{pivot} , the current point p , and the third point p_{third} :
 $t = \text{Triangle}(p_{\text{pivot}}, p, p_{\text{third}})$
5. Add the triangle t to the list of reconstructed surface triangles.
6. Remove the current point p from the point cloud.
7. The list of reconstructed surface triangles represents the approximated surface of the object.

In this pseudocode, the formula for calculating the Euclidean distance between two points (x_1, y_1, z_1) and (x_2, y_2, z_2) is given by:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

3.2 Euclidean Clustering Through DBSCAN

DBSCAN performs an iterative analysis of the points in the dataset. For each point being analyzed, it identifies the set of points that can be reached by density from that point. This is accomplished by computing the neighborhood of the point, and if the neighborhood contains more points than a specified threshold, it is considered part of the region. The same process is then applied to each neighboring point until the cluster can no longer be expanded. If a point does not have enough neighbors, indicating that it is not an interior point, it is labeled as noise [14]. This inherent mechanism of DBSCAN effectively isolates and handles outliers, contributing to its robustness in dealing with such data points. Figure 3 illustrates the tapioca plant's point cloud before and after segmentation using the DBSCAN algorithm.

Algorithm 2:

1. Set the neighborhood radius ϵ (epsilon).
2. Set the minimum number of points required to form a dense region, minpts .
3. For each point p in the point cloud:
4. Retrieve all neighboring points within distance ϵ of point p .
 $\text{neighbors} = \text{getNeighbors}(p, \epsilon)$

5. *If the number of neighbors is less than minPts, mark point p as noise or an outlier. Otherwise, make a new cluster C and assign point p to cluster C.*

$$C = \text{new Cluster}()$$

$$C.\text{addPoint}(p)$$
6. *Expand the cluster by recursively visiting all reachable neighbors and assigning them to cluster C. expandCluster(p, neighbors, C, ε, minPts)*
7. *The getNeighbors function retrieves all neighboring points within distance ε of a given point p. To calculate the distance between two points (x₁, y₁, z₁) and (x₂, y₂, z₂) is given by:*

$$\text{distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

3.3 RANSAC Algorithm

The RANSAC algorithm iteratively samples random subsets of data points and fits a model to these subsets. It evaluates the quality of each model by measuring the error or residual for each data point and count the number of inliers within the threshold [16]. The algorithm terminates when the maximum number of iterations is reached or when a model with a satisfactory number of inliers is found. The final model is then refined using all the inliers.

Algorithm 3:

1. *Initialize the region of interest (ROI) boundaries: pc = {X_limit, Y_limit, Z_limit}.*
2. *Set the maximum distance for ground segmentation: max_Distance = d and define the reference vector = v.*
3. *Fit the plane and identify outliers using the ground plane constraint: fit plane outliers(pc, v, d).*
4. *Name the inlier points as green indices: inPlane_PointIndices = green_Idx.*
5. *Select the outliers as the points without the ground plane: outliers = pc_WithoutGround.*
6. *End the loop.*
7. *For every sensor location: s and radius: r1.*
8. *Search the neighboring points within the specified radius: nearIndices = find neighbors in radius(s, r1, pc_WithoutGround).*
9. *Label the neighboring points as red indices: near_PointIndices = Red_Idx.*
10. *End the loop.*
11. *While the radius: r2 is less than r1.*
12. *Find the neighboring points within the radius: nearIndices = find neighbors in radius(pc_WithoutGround, s, r2).*
13. *Identify the building points as outliers from the nearIndices: building_PointIndices = outliers(nearIndices).*



Fig. 4 Real-time 2D images collected from field

14. *Select the points from $pc_WithoutGround$ using the $nearIndices$: $pcBuilding = (pc_WithoutGround, nearIndices)$.*
15. *Apply bounding box analysis to $pcBuilding$ and label the points as black indices: $Black_Idx$.*
16. *End the loop. Return the sequence of pc .*

4 Materials and Methods

4.1 Data Collection

Real-time data collected for Tapioca (*Manihot esculenta*) plants were selected from an agricultural field, which encompassed a one-acre area and comprised nearly 15,000 tapioca plants. These plants were spaced approximately 1.5 feet apart from each other as shown in Fig. 4. The selected plants were mature and healthy, ensuring the absence of any interference from neighboring plants. Due to the isolated nature of the plants, we were able to capture multiple angled photos of individual tapioca plants. Subsequently, these photos were converted into 3D mesh files. The generation of a single 3D mesh file necessitated the collection of 75–100 images per plant. Specifically, each photo was taken at a rotational interval of 20 degrees around the plant. This extensive photo collection was required due to the complex structure of tapioca plants, particularly their leaves.

Three-dimensional reconstruction was done by using the photogrammetry method. The reconstruction was performed on MacBook air with an M1 chip. It typically takes up to 15–20 min for the reconstruction.

Then, professional 3D software 3D-Scanner was used for the post-reconstruction of the image into a 3D model which shaped the point cloud data. On average, 38,000 points were generated for each point cloud data image. The reconstructed image is shown in Fig. 5.

4.2 Description of the Dataset

The sample point cloud data is described below in Table 1. The attributes, x , y , and z , represent the coordinates, while the color values are represented by the RGB values. Regrettably, the intensity value could not be obtained as no 3D scanning techniques

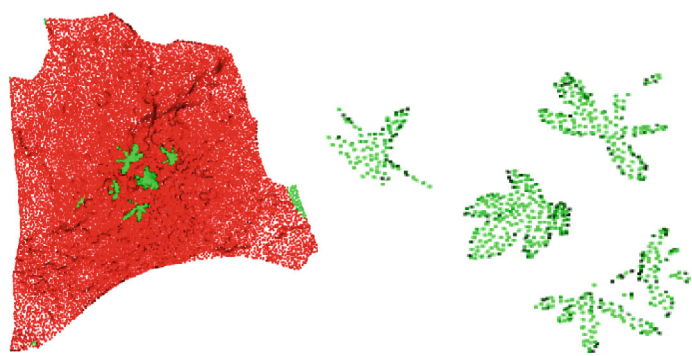


Fig. 5 Reconstructed image

Table 1 This is a sample of point cloud data

Float <i>x</i>	Float <i>y</i>	Float <i>z</i>	Red	Blue	Green	Intensity
0.06526	2.10557	– 1.29676	64	64	48	0
0.10965	2.04676	– 1.29283	11	8	0	0
– 1.4607	2.0079	– 1.27277	120	133	151	0
– 1.47485	1.96253	– 1.31045	61	56	46	0
0.0671	1.94175	– 1.31699	88	88	88	0

were used. The coordinate values range from -3 to $+3$, indicating their spatial arrangement. In this context, the position of the camera serves as the origin for the 3D spatial arrangement.

4.3 Methodology

The method employed for point cloud extraction, unfortunately, introduced a substantial amount of noise, which posed a challenge and required manual intervention for its removal as shown in Fig. 6. To address this issue, we utilized the Cloud Compare software, to meticulously eliminate the noise sources, such as the shadows caused by the human subject and disturbances arising from external environmental conditions. After the pre-processing threshold, distance and the number of iterations are figured out to use the algorithm. To increase accuracy, it is often beneficial to perform a higher number of iterations. A default value of 200 iterations is chosen to balance accuracy with computational efficiency. The RANSAC algorithm is implemented in a Python development environment using the Open3D library.

Segmentation of point cloud data using RANSAC involves three stages: (1) selecting the region of interest to focus on the object’s (plant’s leaves) environment, (2) segmenting the ground plane to differentiate the ground and nonground surface of

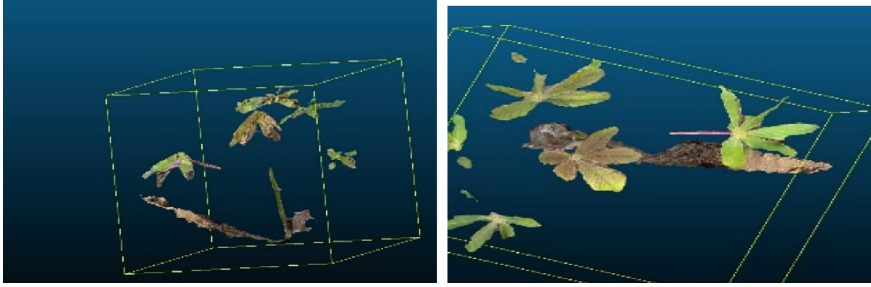


Fig. 6 Manual removal of ground-level points

the plant, and (3) nonground objects are segmented by utilizing their specific spatial locations. These stages enable accurate localization of the target object and provide valuable information about the surrounding objects and ground surface of the plant. The entire methodology is shown in Fig. 7.

To calculate the directions of the plane, Eq. 1 can be utilized.

$$xA + yB + zC + d = 0 \quad (1)$$

In Eq. (1), the variables A , B , and C represent the coordinates of 3D points in space. The coefficients x , y , and z correspond to the vector components along their axes, while d represents the vector transient through the origin.

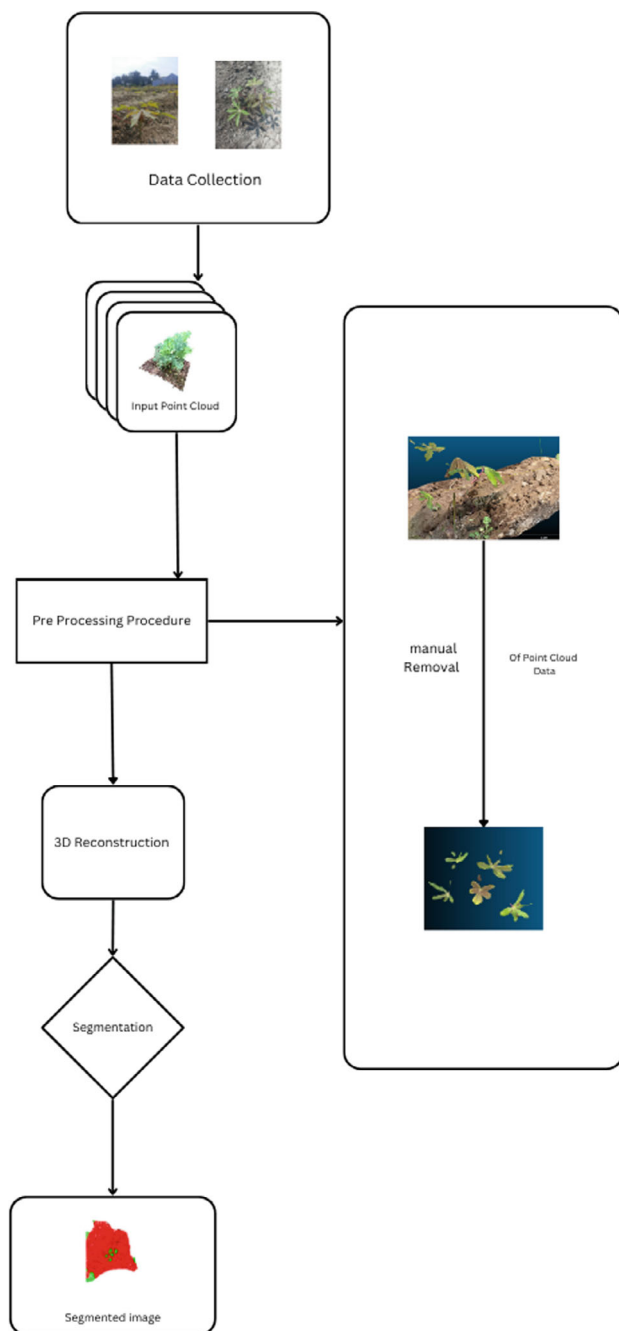
The Eq. (2) can be utilized to calculate the distance, denoted as D , between the point and the plane.

$$D = (xA + yB + zC + d) / (\sqrt{x^2 + y^2 + z^2}) \quad (2)$$

The number of iterations, denoted as N , can be intended using the given Eq. (3):

$$N = (\text{Log}(1 - p)) / (\text{Log}(1 - (1 - e)^s)) \quad (3)$$

In this context, the symbol p represents the probability associated with the selection of a sample of high quality. The symbol e is used to represent the probability of selecting outliers based on the desired level of confidence. Additionally, the symbol s denotes the initial number of points used for fitting the model. The pseudo-code for the overall process is presented in Algorithm 3. It takes the 3D point cloud data, pc , along with the parameters d , v , s , $r1$, and $r2$ as input. These parameters are described in detail in Sects. 4.4 and 4.5.

**Fig. 7** Operational sequence

4.4 Region of Interest Selection

In this study, the first step was selecting a Region of Interest (RoI) from the point cloud data (pc) to emphasize key information about the leaf. This was done to manage large datasets by focusing on the RoI rather than the entire plant data. The data was saved in the Cloud Compare software as a.ply file and then imported into an array structure. This allowed different properties of the 3D point clouds to be stored, including each point's coordinates (X , Y , Z), intensity value (I), and RGB color information. Axis limits (X_limit , Y_limit , Z_limit) were set to define the points displayed, taking into account the plant's unique characteristics.

4.5 Ground Segmentation

The process of ground segmentation is a separation of ground points from nonground ones within a point cloud. This classification for ground points from point cloud data is a very important step for terrain modeling; moreover, it is also important for applications such as autonomous driving and robotics. Hence, the segmentation of ground points in point clouds can be performed using a set of features based on surface planarity, height, or geometric properties that are able to classify any points representing the best ground surface. Some widely used techniques include:

1. Plane fitting: fitting planes to subsets of a point cloud, and finding out which planes represent the ground, with criteria like least squares or RANSAC.
2. Height thresholding: Classifies points as ground due to their heights and thresholding—points below the threshold are ground points.
3. Morphological operations: Utilize morphological opening filters to remove small objects and noise, on the assumption that the ground is larger and more continuous than other things.
4. Machine learning: One may utilize classification methods, building a model that would classify ground and nonground points by the features extractable from the point cloud. Generally, these include local surface normals, color, or context. However, the choice of method depends on the characteristics of the point cloud and the concrete application. More often than not, these techniques are combined or modified in such a way as to meet requirements in the particular task.

The ultimate objective after downsampling is to differentiate the ground points into inliers and outliers. Inliers become the points classified as suitable for a ground plane, but outliers are the ones eliminated from the scene. The distance between outlier points is maximized throughout the segmentation process. The three-dimensional information concerning each point, that is (x, y, z) , is gathered and converted to an array for efficient use as one of the parameters by the RANSAC algorithm.

4.6 Threshold Evaluation

Using trial and error method, the threshold distance is evaluated. For every (a,b,c) position of a point, the algorithm groups the points by randomly selecting two points. Points are considered as inliers if the distance between the points is less than the threshold distance [17].

To achieve the desired clustering and separation of ground points as outliers, threshold distance is used. Let ' d ' be the distance from the origin to the plane and (a,b,c) be the spatial coordinates of the point cloud data.

$$(a, b, c) \geq d \in \text{inliers if, } ax + by + cz + d = 0$$

If the points fit into the above-mentioned plane equation, then the points are classified as inliers.

4.7 Geometric Feature Extraction

Geometric features are meaningful geometric characteristics or properties in a given dataset. What are the geometric features in point cloud data: In point cloud data, they are essential to preserve information about shape state and spatial relationships within objects or surfaces contained in the original 3D object [18, 19].

Circumference of the point cloud data: Different geometrical features can be extracted from a model, depending on specific application and requirement. Geometric features that are frequently used include:

1. Position of the point: This basic geometric feature is provided by three coordinates, x,y,z and are necessary iff we need to understand any given spatial distribution or distance on an orientational angle.
2. Surface Normals: The geometric surface normal provides a facility as an orientation of the local surfaces at every point in point cloud. These are calculated by either neighborhood analysis or curvature estimation. Therefore, we can use surface Normals to detect planar areas, recognize edges, or evaluate surface curvatures.
3. Curvature: It gives the extent of variation, from the straight line. It is possible to compute it in a few of ways, for example, by using the PCA or the local surface patches. Such information is useful for object segmentation, the difference between the smooth and sharp surfaces or if there are some defects on the surface of the object.
4. Shape descriptors: This silhouette features are rather concise representations of the several geometric characteristics that determine the overall form and structure of the objects in the point cloud. Some of the features of the shape can for

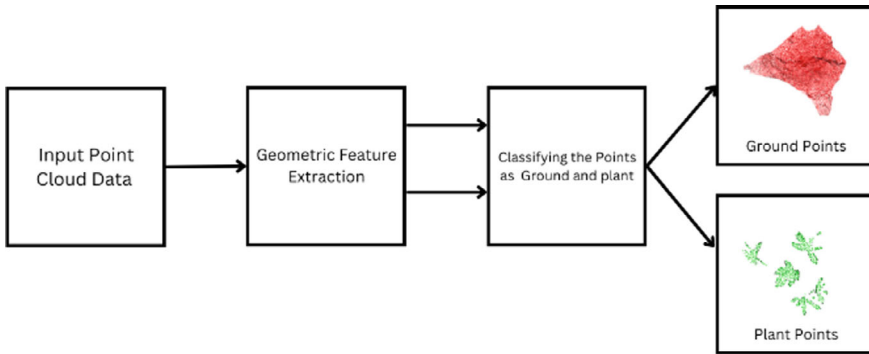


Fig. 8 Classification of ground and plant points

instance be defined by spherical harmonics, 3D shape histograms, or moment-based features. They are commonly applied on object recognition, classification, or even on a simple retrieval system.

5. **Local features:** Local geometric features express precise information about local geometry of a neighborhood of each point. This may be local point density, local surface fluctuation, or local shape coefficient. Local features are good for deciding between dissimilar parts of an object or its points of interest or as a way of finding local distortions of the form.
6. **Topological features:** Geometric shapes pertain to the structure of the point cloud and the relative position of points or zones of the point cloud. Such features can be topological invariants or connectivity graph or neutron neighbor relation among others. Topological features are useful for description of the object relationships, clustering or branching of point cloud or to describe the general structure of the obtained point cloud.

Which geometric features to extract is also dependent with the type of analysis that is intended to be done and nature of point cloud data. These features can be calculated on the basis of different mathematical computations starting from geometrical calculation up to mathematical models and algorithms. The extracted geometric features are beneficial for multiple tasks such as object recognition, segmentation, registration, reconstruction, and many others. Figure 8 shows an elaboration of the classification of ground points and plant points with the help of geometric feature extraction and color code intensity.

5 Conclusion

This paper proposes a practical method for collecting and segmenting point cloud data of tapioca plants through photogrammetry. The plants were completely imaged, and using this photogrammetric technique, a high detailed 3D virtual model of the plant

was possible. The segmentation of the point cloud data into ground and plant points using RANSAC was then able to separate these spatial components. The visualization of the segmented data has given us quite a good understanding about distribution and spatial pattern of tapioca plantation. The segmentation is useful in the plant phenotyping community for feature extraction and studying growth patterns. These tensors allow for further processing and analysis of the point cloud data, which can include machine-learning-based classification. A release of this dataset will allow the community to work with a common data source increasing collaboration and advancing plant phenotyping research. Finally, this research study contributes to enlightening cultivation practices and agricultural management through the enhanced plant phenotyping techniques.

References

1. Turgut K, Dutagaci H, Galopin G (2022) Segmentation of structural parts of rosebush plants with 3D point-based deep learning methods. *Plant Methods* 18:20. <https://doi.org/10.1186/s13007-022-00857-3>
2. Vinodkumar PK, Karabulut D, Avots E, Ozcinar C, Anbarjafari G (2023) A survey on deep learning based segmentation, detection and classification for 3D point clouds. *Entropy* 25(4):635. <https://doi.org/10.3390/e25040635>
3. Chebrolu N, Magistri F, Läbe T (2021) Stachniss C: registration of spatio-temporal point clouds of plants for phenotyping. *PLoS ONE* 16(2):e0247243. <https://doi.org/10.1371/journal.pone.0247243>
4. Japes B, Mack J, Rist F, Herzog K, Töpfer R, Steinhage V (2018) Multi-view semantic labeling of 3D point clouds for automated plant phenotyping. *arXiv preprint arXiv:1805.03994*
5. Shi W, van de Zedde R, Jiang H, Kootstra G (2019) Plant-part segmentation using deep learning and multi-view vision. *Biosyst Eng* 187:81–95
6. Guo Y, Wang H, Hu Q, Liu H, Liu L, Bennamoun M (2019) Deep learning for 3D point clouds: a survey. *IEEE Trans Pattern Anal Mach Intell* 43(12):4338–4364
7. Zhu W, Sun Z, Peng J, Huang Y, Li J, Zhang J, Yang B, Liao X (2019) Estimating maize above-ground biomass using 3D point clouds of multi-source unmanned aerial vehicle data at multi-spatial scales. *Remote Sens* 11(11):2678. <https://doi.org/10.3390/rs11222678>
8. Ghahremani M, Williams K, Corke FMK, Tiddeman B, Liu Y, Doonan JH (2021) Deep segmentation of point clouds of wheat. *Front Plant Sci* 12:608732
9. Ben-Shabat Y, Lindenbaum M, Fischer A (2017) 3d point cloud classification and segmentation using 3d modified fisher vector representation for convolutional neural networks. *arXiv preprint arXiv:1711.08241*
10. Ester M, Kriegel H-P, Sander J, Xu X (2018) Efficient segmentation of 3D point cloud data using density-based clustering. In: *Proceedings of the 2nd international conference on knowledge discovery and data mining (KDD)*, pp 226–231. <https://doi.org/10.5555/3001460.3001507>
11. Charles RQ, Su H, Kaichun M, Guibas LJ (2017) PointNet: deep learning on point sets for 3D classification and segmentation. In: *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, Honolulu, HI, USA, pp 77–85. <https://doi.org/10.1109/CVPR.2017.16>
12. Qi CR, Yi L, Su H, Guibas LJ (2017) PointNet++: deep hierarchical feature learning on point sets in a metric space, in *advances in neural information processing systems*
13. Li Y, Bu R, Sun M, Chen B (2018) PointCNN: convolution on X-transformed points. In: *Advances in neural information processing systems (NeurIPS)*, pp 31
14. Li Y, Du L, Zhou Y, Hu Q, Zhang B (2020) RandLA-Net: efficient semantic segmentation of large-scale point clouds. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp 11108–11117

15. Chaira T, Chaki N, Das S (2010) Segmentation and classification of point cloud data using ball pivoting. *Int J Pattern Recognit Artif Intell (IJPRAI)* 24(2):289–302. <https://doi.org/10.1142/S0218001410010039>
16. Duan X, Song H, Tian F, Song X, Li X (2018) A robust algorithm for segmenting point cloud data using RANSAC and region growing. *IEEE Access* 6:44911–44919. <https://doi.org/10.1109/ACCESS.2018.2816289>
17. Zhang Y, Xue F, Xu G (2018) A comprehensive threshold evaluation approach for image segmentation. *IEEE Trans Image Process* 27(5):2424–2437. <https://doi.org/10.1109/TIP.2018.2800643>
18. Rusu RB, Blodow N, Beetz M (2009) Fast point feature histograms (FPFH) for 3D registration. In: *Proceedings of the IEEE international conference on robotics and automation (ICRA)*, 2009, pp 3212–3217. <https://doi.org/10.1109/ROBOT.2009.5152473>
19. Tombari F, Salti S, Di Stefano L (2010) Unique shape context for 3D data description. In: *Proceedings of the European conference on computer vision (ECCV)*, 2010, pp 356–369. https://doi.org/10.1007/978-3-642-15561-1_26

Denoising Diffusion Probabilistic Model with Fashion MNIST



Palak Singh, Pratyush Priyam, Ayush Yadav, Aaron Issac Abraham, Sudeshna Datta, Nimish Singh, and Pranjal Kumar

Abstract Deep Diffusion Probabilistic Models (DDPMs) are a versatile class of generative models able to efficiently model complex data distributions and have quickly become challenging benchmarks for density estimation tasks. In this work, we present a neat solution to implement DDPMs for image generation using a compact TinyUNet architecture as well as diffusion-based sampling methods on probability modeling. The Fashion MNIST dataset is the canonical training data used for our model, which leverages a back propagation-based loss minimization algorithm implemented using TensorFlow to iteratively optimize encoder/decoder parameters so as to minimize Mean Squared Error between predicted noise and ground truth noise. We comprehensively evaluate and compare our method to demonstrate its effectiveness for synthesis of high-quality images with different visual attributes. In addition, we offer insights into how the DDPM framework operates, demonstrating model architecture interpretability and scalability. Our work is a good snapshot in the much larger context of generative modeling techniques and will provide clues for this type to scale further.

Keywords Diffusion · DDPM · TinyUNet · ConvNet · Fashion MNIST · Embeddings

1 Introduction

In the realm of generative modeling, the pursuit of synthesizing high-quality samples from complex distributions remains a longstanding challenge. Traditional approaches often encounter difficulties in capturing intricate patterns and producing realistic outputs, especially when confronted with datasets characterized by variability, noise,

P. Singh (✉) · P. Priyam · A. Yadav · A. I. Abraham · S. Datta · N. Singh · Pranjal Kumar
School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India
e-mail: palaksingh2802@gmail.com

Pranjal Kumar
e-mail: pranjal.30426@lpu.co.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
A. Kumar et al. (eds.), *Proceedings of Fourth International Conference on Computing and Communication Networks*, Lecture Notes in Networks and Systems 1317,
https://doi.org/10.1007/978-981-96-3942-7_32

427

and structural complexity. In response to these challenges, Denoising Diffusion Probabilistic Models (DDPM) emerge as a promising avenue, offering a principled framework for generating diverse and realistic samples with precise control on the generating process.

DDPMs operate on the principle of simulating the diffusion process, where a noisy input undergoes iterative denoising steps to progressively transform into a clean output. The model correctly captures the underlying data distribution and creates high-quality samples by guessing the conditional probabilistic distribution of the clean image given the noisy input at each step. This approach stands in contrast to traditional generative models by providing enhanced capabilities in synthesizing visually appealing outputs with sharp details and structural coherence. The importance of DDPMs lies in its ability to address the inherent limitations of traditional generative models, offering a robust framework for capturing complex data distributions. By leveraging the principles of diffusion processes and probabilistic modeling, DDPMs provide researchers with a versatile toolset for tackling various generative modeling tasks across different domains.

In the structure of image generation, DDPMs play a pivotal role in synthesizing realistic samples that closely resemble real examples. This capability becomes particularly relevant when applied to benchmark datasets such as the National Institute of Standards and Technology's Fashion Modified (Fashion MNIST), that is made up of an enormous collection of grayscale pictures. The Fashion MNIST dataset of Zalando article images comprises of a test set of 10,000 samples and a training set of 60,000 samples. Each example is a 28 by 28 grayscale image labeled with one to 10 different classes.

Despite its simplicity, Fashion MNIST poses several challenges, including variability in clothing styles, textures, and orientations, as well as potential noise in the images. By leveraging Denoising Diffusion Probabilistic Models (DDPMs) with the Fashion MNIST dataset, researchers can explore the capabilities of these models in getting the fundamental distribution of fashion items and generating realistic samples. In order to reduce the difference between the generated samples and the actual distribution of fashion items, the model parameters are adjusted during the training phase. Through rigorous experimentation and evaluation, we aim to assess the effectiveness of DDPMs in generating high-quality images of various clothing items while providing insights into the strengths and limitations of these models.

In this research paper, we present a thorough exploration of Denoising Diffusion Probabilistic Models for fashion item generation, focusing specifically on their role in synthesizing realistic samples from the Fashion MNIST dataset. We investigate the training process, model optimization, and evaluation metrics, aiming to shed light on the efficacy of DDPMs in capturing complex data distributions and generating high-quality samples of clothing items. Our objective is to make a positive impact on the field of fashion and clothing synthesis by advancing the use of generative modeling methods through empirical analysis and qualitative analyses.

2 Literature Review

In the ever-evolving landscape of image denoising research, diverse methodologies and alternative approaches are continuously explored to address the complex challenges of noise reduction in various domains. Shafiee et al. [1] introduce novel approaches to AUHB-DWR, leveraging Hidden Stochastically Fully Connected Conditional Random Fields (HSFCRF) and Hidden Conditional Random Fields (HCRF) for effective delineation of suspected cancerous tissue within the prostate gland. A study by Niu et al. [2] propose CDPMSR, enhancing DPM-based super resolution methods, simplifying the process, and improving results across various benchmark datasets. Li et al. [3] paper surveys existing DM-based image generation models, highlighting their advantages and suggesting avenues for improvement. Dalmaz et al. [4] present Diff-Med-Synth, a method combining denoising diffusion probabilistic models and generative adversarial networks for efficient synthesis of with high quality and diversity. Another research conducted by Vo et al. [5] demonstrates the fusion of diffusion models with PhoBERT for text-to-image synthesis, showing promising results on the UIT-ViC dataset. The paper by Zhang et al. [6] proposed ControlNet to enhance text-to-image diffusion models, utilizing pretrained models and “zero convolutions” for accurate spatial control. This promises wider applications in image production, emphasizing robust training and scalability. The breakthrough Transformer architecture, introduced by Vaswani et al. [7], revolutionized natural language processing, streamlining sequence processing through self-attention processes and motivating attention-based models in various fields. Wang et al. [8] propose a method leveraging pretraining to refine translation quality, exploring image spaces effectively, and showcasing notable improvements across tasks and datasets. A study by Jian et al. [9] suggest overcoming data scarcity and class imbalances in aerial object detection using stable diffusion techniques and copy-paste composition. Lugmayr et al. [10] in his paper present the Denoising Diffusion Probabilistic Model (DDPM), using previously learned models to produce superior quality images and improving picture restoration with diverse masks. Promising progress is being made in a number of areas, such as picture synthesis and object identification, through the application of denoising diffusion models. In Bhunia’s et al. [11] paper, PIDM streamlines the production of lifelike person images and achieves remarkable results in tasks such as human reidentification on benchmark datasets such as DeepFashion and Market-1501. Song’s et al. [12] “DDIMs” generate images up to 50 times faster than DDPMs without compromising quality. With its exceptional sample quality and effective latent space picture reconstruction, it represents a major breakthrough in generative modeling offering a great deal of versatility. A novel approach by Sun et al. [13] presented a concept for image denoising based on anisotropic diffusion, showcasing faster convergence speeds while preserving edges, textures, and details effectively. Kim et al. [14] put forward a method that enhanced the time-embedded U-Net with a timestep predictor, improving sequential flow understanding and denoising performance. Li et al. [15] proposed advancements in anisotropic diffusion models, surpassing conventional methods in image feature preservation.

Another research by Wang et al. [16] created an innovative denoising technique that uses directions for diffusion, minimizing noise, and aliasing effects while preserving edges and textures. Gao's et al. [17] paper introduced CoreDiff, a unique diffusion model for LDCT denoising, demonstrating superior achievement with a clinically reasonable time conclusion. Xueqing et al. [18] proposed reaction diffusion equation-based algorithms, effectively removing noise while preserving image details. Deng et al. [19] leveraged RGB images to enhance diffusion model-based HSI denoising, improving results by integrating spatial information during the reverse diffusion process. Rombach et al. [20] introduced latent diffusion models (LDMs), enhancing denoising diffusion models' efficiency while maintaining image synthesis quality. Operating within the latent space of autoencoders and incorporating cross-attention layers, LDMs attain cutting-edge results in a variety of image synthesis challenges, including repainting, conditional synthesis, and super resolution, while significantly reducing computational requirements. Luo et al. [21] presented Refusion, a U-Net-based latent diffusion model, achieving realistic image restoration, even on large size images, with superior performance in tasks like shadow removal, dehazing, and super resolution. Another study by Avrahami et al. [22] proposed Blended Diffusion, combining CLIP for text guidance with DDPMs for image generation, enabling seamless text-driven local edits in natural images. Whang et al. [23] in his paper introduced a novel framework for blind image deblurring using conditional diffusion models and stochastic refinement, enhancing perceptual quality significantly. Yu's et al. [24] paper introduced PVDM, a projected latent video diffusion model, enabling efficient synthesis of high-resolution and temporally coherent videos. A study conducted by Brooks et al. [25] introduced InstructPix2Pix, enabling human-guided picture modification by utilizing conditional diffusion models and previously trained models, marking a significant advancement in interactive image editing methods. Xu et al. in his paper [26] introduces a benchmark dataset for reducing noise in real-life noisy images, emphasizing the need for robust methods. Liu et al. [27] propose a multiple level CNN wavelet framework, striking a balance between computational performance and responsive field dimension, yielding promising outcomes. Another study by Wang et al. [28] presents Uformer, a Transformer-based framework for restoring images, demonstrating the potential of Transformer blocks. Cheng et al. [29] introduce Nbnet, a framework for image denoising with cutting-edge efficiency and lower computational costs. Li et al. [30] in his paper propose DMID, a strategy leveraging diffusion models for image denoising, showing remarkable performance improvements. Soltanayev et al. [31] tackle training deep learning-based denoisers without ground truth data, using Stein's unbiased risk estimator (SURE) for unsupervised training. Anwar et al. [32] introduce RIDNet, a blind actual image denoiser network with attention feature, outperforming contemporary algorithms. Cha et al. [33] propose FC-AIDE, blending supervised training with adaptive fine-tuning, achieving superior denoising performance. Tolooshams et al. [34] present deep convolutional exponential autoencoders (DCEA) for natural exponential family data, excelling in unsupervised and supervised Poisson image denoising tasks. Abbasi et al. [35] offer multiscale sparsifying transform learning methods for image denoising, balancing performance and complexity effectively. Other paper by Sokolovskyy et al. [36]

models the diffusion process using fractional order derivatives and presents difference schemes and algorithms for its approximation. Yu et al. [37] paper proposes an anisotropic hybrid diffusion model with a source term to effectively denoise images while maintaining contrast between coherent structures and the background, outperforming traditional edge enhancing and coherence enhancing models. A study conducted by Feng et al. [38] suggested an efficient Poisson denoising model using a retrained trainable nonlinear reaction diffusion (TNRD) approach, optimized for high computational efficiency and recovery quality, with results competitive to state-of-the-art methods and well-suited for parallel GPU computation. Deng et al. in his study introduces a noise-model-free HSI denoising method based on a Diffusion Model (DM) that incorporates spectral and spatial information, effectively addressing varying noise models across spectral bands and preventing overfitting through stochastic sampling, with proven efficacy in diverse sample tests.

These advancements collectively showcase the diverse array of methodologies and innovations in image denoising research, each contributing unique insights and pushing the boundaries of denoising capabilities in real-world scenarios.

3 Methodology

The research presents an innovative method for generating pictures with the Fashion MNIST dataset by utilizing the Denoising Diffusion Probabilistic Model (DDPM). By using dedicated architecture and time embedding layer, TinyUNet was able to fetch out temporal coherence between the two adjacent states and also spatial correlations in the data. By extensive training and evaluation process of our model, we show the ability of our model to simulate a diffusion process and generate excellent pictures when trained from scratch. In addition to producing promising results, our approach also lays the groundwork for future extensions of probabilistic generative modeling methods.

3.1 *Model Architecture*

The architecture of the model is combination of Convolutional Neural Networks (ConvNets) and a custom U-Net implementation known as TinyUNet. Hence, Convolutional Neural Networks (ConvNets) which are used to capture and learn those hierarchical features from input images. In our particular case, TinyUNet is just the base architecture with down and up sampling blocks that are good enough to capture both local information by having big layers in between (to accumulate a lot of contexts) and small ones on the encoder for precise localization. Being so small, it is computational efficient as well. The combination of these features permits the model to generate high-resolution images, learn complex spatial relationships. In addition, for picture generation tasks, the model is augmented with time embedding layers to

retain temporal dynamics and better model the diffusion process. This enhances the model’s ability to generate realistic images, especially with the inclusion of ConvNets and TinyUNet architectures.

3.2 Sinusoidal Embedding

Temporal dynamics capture is crucial in picture production tasks, and time embedding layers (e.g., sinusoidal embeddings) are needed to facilitate this. They encode time signals in a fully fledged structured format, so that the model can quite easily implement and handle those recurring periodic features. Mathematically, sinusoidal embeddings mentioned in Vaswani et al. [7] are generated using the formula:

$$PE_{(pos, 2i)} = \sin(pos/10,000^{2i/d_model}) \quad (1)$$

$$PE_{(pos, 2i+1)} = \cos(pos/10,000^{2i/d_model}) \quad (2)$$

In this case, “pos” stands for the sequence’s position index, “i” for the embedding vector’s dimension index, and “d_model” for the embedding’s dimensionality. These formulae produce the sinusoidal values that encode the positioning information for every element in the sequence when the proper values are entered.

Sinusoidal embeddings are important because they enable the model to learn and take advantage of long-term temporal relationships. The model extracts temporal information and propagates them across the network layers effectively by adding sinusoidal embeddings into ConvNets and TinyUNet blocks. By doing this, you can be confident that the model will produce realistic and high-quality pictures with consistent temporal dynamics, which will enhance its performance in image generating tasks. All things considered, the use of sinusoidal embeddings improves the ability of model to represent and record intricate changes in time, which helps it produce realistic and temporally consistent visual outputs.

3.3 Data Preparation

Selection of Dataset: For this experiment, the Fashion MNIST dataset (see Fig. 1) is utilized. Grayscale photos of clothing products from ten distinct categories—trousers, shirts, ankle boots, dresses, pullovers, sandals, jackets, T-shirts/top, bags, coat, and sneakers—make up this dataset.

Data Transformation: The images present in the dataset are resized to a fixed size of 32×32 pixels using “transforms.Resize(32).” Additionally, we use “transforms.ToTensor()” to convert images to tensors.

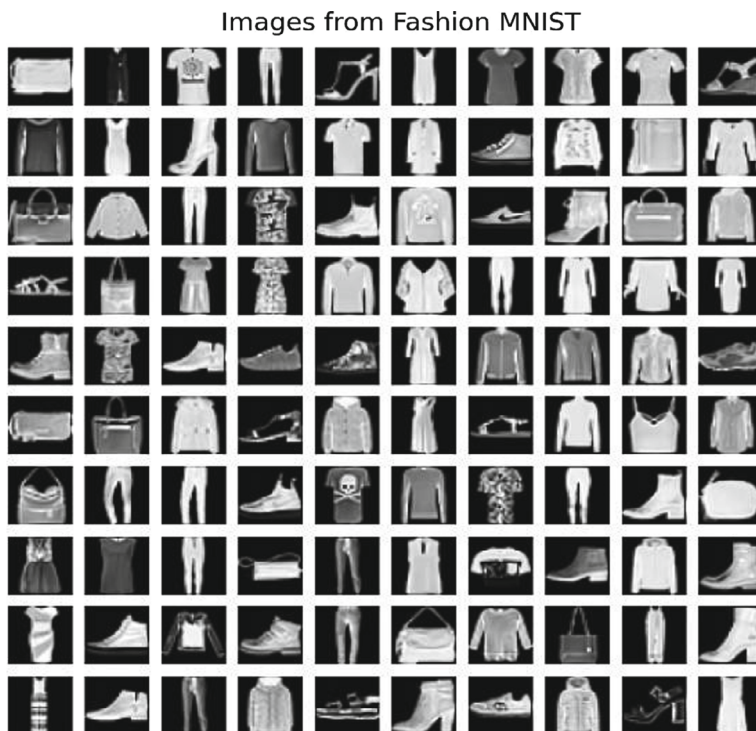


Fig. 1 Random images from fashion MNIST dataset

Normalization: The values of pixel of each image present in the dataset are normalized so that the mean value is around 0.1307 and standard deviation is 0.3081. This normalization is applied using transforms. `Normalize ((0.1307), (0.3081))`.

Loading the Dataset: The images after being transformed is loaded into a PyTorch DataLoader object for efficient batching and training. During loading, the dataset is shuffled to introduce randomness into the training process and enhance model generalization.

Batch Size Specification: The number of training example used in one iteration is set to 4096, allowing the model to process multiple samples in parallel during each training iteration.

Parallel Data Loading: To accelerate data loading, the `num_workers` argument is set to 16, enabling the DataLoader to use multiple worker processes for fetching batches asynchronously.

3.4 Diffusion Process

Diffusion process is fundamental in various fields, including physics, chemistry, and biology. It describes how particles move randomly in a medium, with their speed determined by the concentration of the medium. In our context, the diffusion process is central to the Denoising Diffusion Probabilistic Models (DDPM) algorithm.

In DDPM, the diffusion process involves progressively adding noise to the data that has been fed in a controlled manner. At each step, noise is introduced, and the model learns to remove this noise, aiming to reconstruct the original data. This iterative process continues, with noise levels increasing over time.

Specifically, the model receives a noisy sample and is tasked with predicting the clean version of that sample (see Fig. 2). At each time step, the noise that has been added follows a Gaussian distribution, with its standard deviation increasing over time. By working in reverse, denoising each step's noise, the model can generate new samples resembling those in the original dataset. The model learns to reduce the MSE (Mean Squared Error) between expected noise by training on a dataset and actual noise fabricate during training. After the model is trained, it repetitively applies the diffusion stage and reverse process to the original noisy input to fabricate new pictures. Using this method, it is possible to produce magnificent graphics that precisely represent the complexity and structure of the training dataset.

There are two major processes in Denoising Diffusion Probabilistic Models (DDPM):

Forward Process: This Step involves generating the data samples by adding noise to a clear image (see Fig. 3). Noise is added in a supervised manner, by simply following a diffusion process, which involve numerous steps. This showcases how a clear image becomes corrupted or shady over time due to various factors such as noise or other forms of distortion. The equation for the forward process given in paper by Ho et al. [39] is:

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}) \quad (3)$$

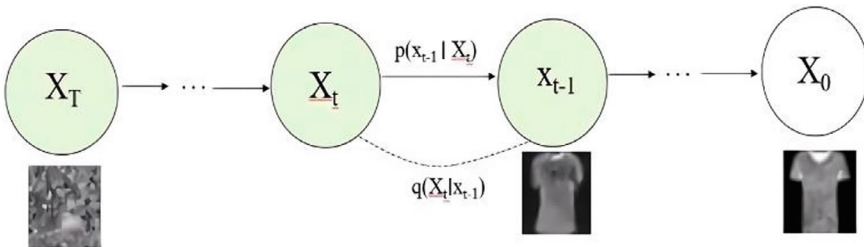


Fig. 2 Directed graph of the algorithm

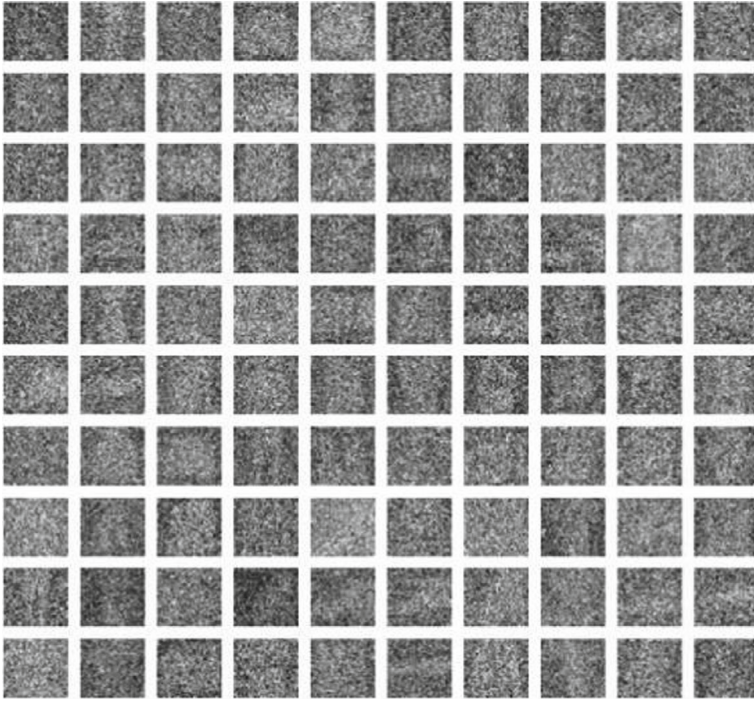


Fig. 3 Images after adding noise

Backward Process: This means to take this deterioration and get a noisy image, into an initial clean state of the same. In DPM, it is most commonly achieved by training a neural network to predict the noise or distortion that will be added when going forward. Given a trained model, it is being applied to an image with noise, hence estimating/reconstructing the original clean image from out of its noisy version. The equation for the backward process given in paper by Ho et al. [39] is as follows:

$$p(x_{0:T}) = p(x_T) \prod_{t=1}^T p_{\theta}(x_{t-1}|x_t) \quad (4)$$

The diffusion process and DDPM are intertwined through the training objective, which involves optimizing the probability of the information fed given both the diffusion process and the model parameters. This likelihood function is essentially the product of conditional distributions that characterize the diffusion process. By optimizing this likelihood function, the model is effectively trained to denoise corrupted samples, aiming to capture the intricate structure and complexity of the data distribution.

In essence, DDPM leverages the diffusion process to generate realistic data by iteratively corrupting and denoising samples, ultimately learning a robust representation of the underlying data distribution.

4 Results

A PyTorch-based Deep Diffusion Probabilistic Model (DDPM), which is trained on Fashion MNIST dataset, is implemented in this work. The DDPM is an effective generative model that can learn the underlying data distribution and produce realistic visuals.

The model, implemented using a TinyUNet architecture, went through 100 epochs of training using Adam optimizer with the learning rate of 1e-3. During the learning process, the model showcases harmonious increase in PSNR (Peak Signal-To-Noise Ratio) values (see Fig. 4). This increase shows that the model is capable of successfully recreating input images and capture fine features within the dataset. In line with this, the PSNR value shows a harmonious increase, indicating improvement in quality of image over epochs.

Moreover, the Grad Norm per Epoch graph (see Fig. 5) provides additional insights into the training dynamics. Alongside the decrease in loss and increase in PSNR, the graph showcases the solidity of the learning process, as indicated by the consistent behavior of the gradient norms. This stability ensures that gradients remain within manageable ranges, facilitating effective model optimization.

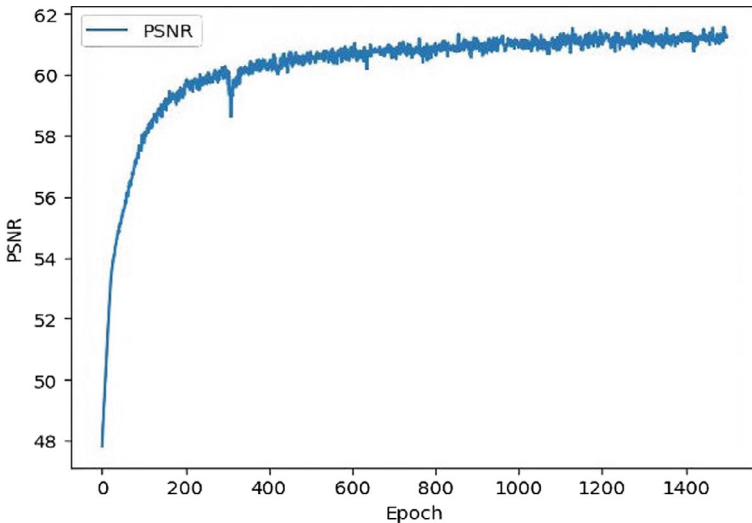


Fig. 4 PSNR per epoch graph

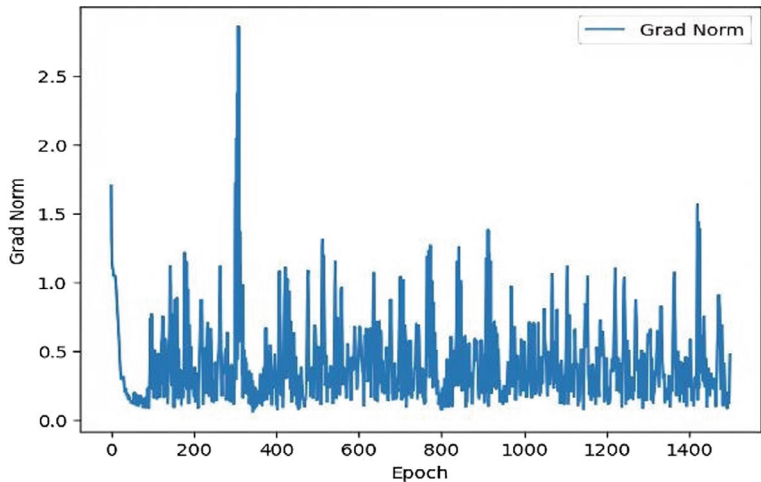


Fig. 5 Grad normal per epoch graph

A closer examination of the loss per epoch (Table 1) reveals the progressive reduction in loss values throughout the training period. Starting from 0.419 at the first epoch, the loss steadily decreases to 0.0493 by the 100th epoch, signifying the model’s convergence and improved reconstruction accuracy. Similarly, the PSNR per epoch table (Table 2) highlights the continuous enhancement in image quality, with the PSNR reaching 61.2 at the conclusion of training.

Furthermore, the MSE loss graph (see Fig. 6) provides a detailed visualization of loss fluctuations during training iterations. Although exhibiting fluctuations, the graph demonstrates an overall decreasing trend in MSE loss values, indicating the iterative refinement of the model’s reconstruction capabilities.

As shown in the finished product (Fig. 7), the model was able to produce a variety of eye-catching and visually attractive visuals that resembled different fashion products after training. These generated images show how realistic examples that closely match the fashion items in the dataset can be produced by the algorithm.

Table 1 Loss per epoch

Epoch	0	20	40	60	80	99
Loss	0.419	0.0686	0.0526	0.0504	0.0501	0.0493

Table 2 PSNR per epoch

Epoch	0	20	40	60	80	99
Loss	51.9	59.8	60.9	61.1	61.1	61.2

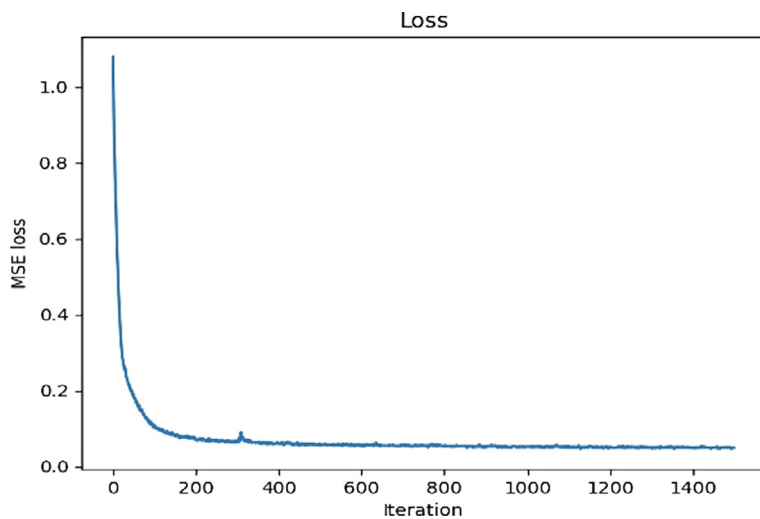


Fig. 6 Loss graph for the model



Fig. 7 Final result after processing

All things considered, the DDPM that was trained on the Fashion MNIST dataset had encouraging outcomes, producing pictures of excellent quality and demonstrating its potential for a range of generative tasks within the fashion industry. The esthetically beautiful final pictures along with the downward trend of the loss graph highlight how well DDPMs capture complicated data distributions and produce realistic samples.

5 Future Scope

Other developments that could then be explored to further the capabilities and acumen of the proposed algorithm include, but are not limited to, enhancing the capacity of the model to render more minute details and to discover subtle patterns that exist in the garment items. This would happen as a result of other advancements in architectural design, such as larger U-Nets or attention mechanisms. Further, if the model also wants to have conditional generation, then maybe the users might choose some traits or styles, and the outputs will be more personalized and flexible. For further enhancements of this model based on increased spectrums of styles and categories of fashion products, further transfer learning methods may need to be applied upon this model. More investigation into data augmentation could further make the model even more robust by only diversifying the training dataset. More quantitative evaluation using the Inception Score or Frechet Inception Distance could provide even more information about quality and diversity in the generated images. Another interesting thing would be the investigation in real life of some virtual try-on platforms or customized recommendation systems to demonstrate the usability of goods in the producing of clothes and to free the creativity of solutions for the fashion sector.

6 Limitations

This paper presents a fresh approach to image synthesis by using DDPM and promising results are achieved. However, such an approach also has limitations:

- **Restricted Dataset Size:** The dataset used in training this model is Fashion MNIST which is a small dataset. Training the model using a more extensive dataset containing a variety of images could enhance the capability of the model to generate a wider range of images and its general applicability.
- **Cost of Computation:** Cost of computation of training a DDPM model can be very high. Using more advance hardware can speed up the training period or finding new techniques for shortening training time period could be useful.
- **Evaluation measures:** Although the generated image quality is evaluated, a more thorough evaluation may be obtained by investigating other quantitative measures to gauge realism and conformity to the data distribution.

- **Limited Model Architecture Exploration:** One particular TinyUNet architecture is the subject of the study. Maybe further exploration in other various architectures or even tweaking hyperparameters can somehow make things even better.

These limitations highlight areas for future research to build upon this work and achieve even better image generation performance.

7 Conclusion

In conclusion, this study offers a fresh method for using Deep Diffusion Probabilistic Models (DDPMs) to create realistic fashion products. Indeed, the developed model catches good fitting to the underlying data distribution of the Fashion MNIST dataset and uses process of diffusion, producing high-quality photographs of fashion products. The results of experiments validate the extendibility of the application because the model designs clothes that are esthetically and varied. Future scope outlines that it further enumerates a few areas of research and developments that should be undertaken to reach an even more functional, adaptive, and practically applicable model. All things considered, this work advances the subject of generative modeling and opens pathways for some future cutting-edge developments in fashion technology and elsewhere.

References

1. Shafiee MJ, Haider SA, Wong A, Lui D, Cameron A, Modhafar A, Haider MA (2014) Apparent ultra-high β -value diffusion-weighted image reconstruction via hidden conditional random fields. *IEEE Trans Med Imaging* 34(5):1111–1124
2. Niu A, Zhang K, Pham TX, Sun J, Zhu Y, Kweon IS, Zhang Y (2023) Cdpmsr: conditional diffusion probabilistic models for single image super-resolution. In: 2023 IEEE international conference on image processing (ICIP). IEEE, pp 615–619
3. Li C, Qi Y, Zeng Q, Lu L (2023) Comparison of image generation methods based on diffusion models. In: 2023 4th international conference on computer vision, image and deep learning (CVIDL). IEEE, pp 1–4
4. Dalmaz O, Saglam B, Elmas G, Mirza M, Çukur T (2023) Denoising diffusion adversarial models for unconditional medical image generation. In: 2023 31st Signal processing and communications applications conference (SIU). IEEE, pp 1–5
5. Vo BK, Ho AD, Luong AV, Dinh D (2023) Combining diffusion model and PhoBERT for vietnamese text-to-image generation. In: 2023 RIVF international conference on computing and communication technologies (RIVF). IEEE, pp 330–335
6. Zhang L, Rao A, Agrawala M (2023) Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 3836–3847
7. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Polosukhin I (2017) Attention is all you need. *Adv Neural Inf Process Syst* 30
8. Wang T, Zhang T, Zhang B, Ouyang H, Chen D, Chen Q, Wen F (2022) Pretraining is all you need for image-to-image translation. *arXiv preprint [arXiv:2205.12952](https://arxiv.org/abs/2205.12952)*

9. Jian Y, Yu F, Singh S, Stamoulis D (2023) Stable diffusion for aerial object detection. arXiv preprint [arXiv:2311.12345](https://arxiv.org/abs/2311.12345)
10. Lugmayr A, Danelljan M, Romero A, Yu F, Timofte R, Van Gool L (2022) Repaint: inpainting using denoising diffusion probabilistic models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 11461–11471
11. Bhunia AK, Khan S, Cholakal H, Anwer RM, Laaksonen J, Shah M, Khan FS (2023) Person image synthesis via denoising diffusion model. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5968–5976
12. Song J, Meng C, Ermon S (2020) Denoising diffusion implicit models. arXiv preprint [arXiv:2010.02502](https://arxiv.org/abs/2010.02502)
13. Sun J, Zhang Y (2019) The rapid anisotropic diffusion model based on Mi-shaped and hyperbolic tangent function. In: 2019 International applied computational electromagnetics society symposium-China (ACES), vol 1. IEEE, pp 1–2
14. Kim JU, Kang DJ (2023) Enhancing denoising models performance through timestep-aware predictor for latent diffusion-based image generation. In: 2023 23rd International conference on control, automation and systems (ICCAS). IEEE, pp 1937–1940
15. Li S, Zhang D, Wu B (2011) A new anisotropic diffusion model for image denoising. In: 2011 International conference on multimedia technology. IEEE, pp 5707–5710
16. Wang N, Shang Y, Chen Y, Yang M, Zhang Q, Liu Y, Gui Z (2018) A hybrid model for image denoising combining modified isotropic diffusion model and modified Perona-Malik model. IEEE Access 6:33568–33582
17. Gao Q, Li Z, Zhang J, Zhang Y, Shan H (2023) CoreDiff: contextual error modulated generalized diffusion model for low-dose CT denoising and generalization. IEEE Trans Med Imaging
18. Xueqing ZHAO, Xiaoming WANG, ZHANG L (2012) Reaction-diffusion equation-based image denoising algorithm. Chin J Electron 21(3):495–499
19. Deng K, Wang P, Qian Y (2024) RGB images enhancing hyperspectral image denoising with diffusion model. In: ICASSP 2024–2024 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2960–2964
20. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B (2022) High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10684–10695
21. Luo Z, Gustafsson FK, Zhao Z, Sjölund J, Schön TB (2023) Refusion: enabling large-size realistic image restoration with latent-space diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 1680–1691
22. Avrahami O, Lischinski D, Fried O (2022) Blended diffusion for text-driven editing of natural images. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 18208–18218
23. Whang J, Delbracio M, Talebi H, Saharia C, Dimakis AG, Milanfar P (2022) Deblurring via stochastic refinement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 16293–16303
24. Yu S, Sohn K, Kim S, Shin J (2023) Video probabilistic diffusion models in projected latent space. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 18456–18466
25. Brooks T, Holynski A, Efros AA (2023) Instructpix2pix: learning to follow image editing instructions. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 18392–18402
26. Xu J, Li H, Liang Z, Zhang D, Zhang L (2018) Real-world noisy image denoising: a new benchmark. arXiv preprint [arXiv:1804.02603](https://arxiv.org/abs/1804.02603)
27. Liu P, Zhang H, Zhang K, Lin L, Zuo W (2018) Multi-level wavelet-CNN for image restoration. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 773–782
28. Wang Z, Cun X, Bao J, Zhou W, Liu J, Li H (2022) Uformer: a general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 17683–17693

29. Cheng S, Wang Y, Huang H, Liu D, Fan H, Liu S (2021) Nbnnet: noise basis learning for image denoising with subspace projection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4896–4906
30. Li T, Feng H, Wang L, Xiong Z, Huang H (2023) Stimulating the diffusion model for image denoising via adaptive embedding and ensembling. arXiv preprint [arXiv:2307.03992](https://arxiv.org/abs/2307.03992)
31. Soltanayev S, Chun SY (2018) Training deep learning based denoisers without ground truth data. *Adv Neural Inf Process Syst* 31
32. Anwar S, Barnes N (2019) Real image denoising with feature attention. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 3155–3164
33. Cha S, Moon T (2019) Fully convolutional pixel adaptive image denoiser. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 4160–4169
34. Tolooshams B, Song A, Temereanca S, Ba D (2020) Convolutional dictionary learning based auto-encoders for natural exponential-family distributions. In: International conference on machine learning. PMLR, pp 9493–9503
35. Abbasi A, Monadjemi A, Fang L, Rabbani H, Noormohammadi N, Zhang Y (2020) Multiscale sparsifying transform learning for image denoising. arXiv pre-print [arXiv:2003.11265](https://arxiv.org/abs/2003.11265)
36. Sokolovskyy I, Campos J (2020) Numerical models of diffusion processes based on fractional order derivatives. In: 2020 IEEE 15th international conference on computer sciences and information technologies (CSIT), vol 1. IEEE, pp 233–236
37. Yu J, Ye J (2017) Image denoising method based on improved HDCS. In: 2017 IEEE 9th international conference on communication software and networks (ICCSN). IEEE, pp 1282–1286
38. Feng W, Qiao P, Chen Y (2017) Fast and accurate poisson denoising with trainable nonlinear diffusion. *IEEE Trans Cybern* 48(6):1708–1719
39. Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. *Adv Neural Inf Process Syst* 33:6840–6851

The World State of Artificial Intelligence Based on the Analysis of Research Papers and Patents



Adriana Baravalle, Ignacio del Carril, and Gabriela Robiolo

Abstract The artificial intelligence research works and patented developments in the last five years are shaping the future, having direct impact on the global economy and society. This paper provides a comprehensive literacy analysis and metrics sourced from the Lens.org platform. We focused the analysis on the current state of AI innovations, captured from research papers and patents. Since 1978, the total number of registered research production is 1,262,789 papers (198,163 from 2019 to 2024), 2023 being the most productive year (74,426 articles). The analysis underscores Asia, notably China (CH), as the primary hub for artificial intelligence research. Institutional affiliations reside in Asia, with the Chinese Academy of Sciences leading with 5512 articles. Wei Wang emerges as a prominent figure, and the thematic dominate cluster is computer vision. On the other hand, in the last five years of production, there were 316,883 patents. US leads the applications, followed by CH. Samsung Electronics Co. Ltd. is the first in the east, and IBM leads the west region. The top inventors are found at Qualcomm Inc. and LG Electronics. The main cluster is computing arrangement based on specific computational models, such as machine learning, learning methods, combinations of networks, backpropagation, recurrent networks, probabilistic graphical models, and inference or reasoning models. This analysis sheds light to researchers, policymakers, investors, and inventors on the current landscape of AI research and development, showing as its main insight that there is a wide gap between industry and academy.

Keywords Artificial intelligence · Bibliometric · Patentometric

A. Baravalle · I. del Carril · G. Robiolo (✉)

Facultad de Ingeniería, Universidad Austral, Buenos Aires, Argentina

e-mail: grobiolo@austral.edu.ar

A. Baravalle

e-mail: abaravalle@austral.edu.ar

I. del Carril

e-mail: idelcarril@austral.edu.ar

G. Robiolo

LIDTUA, Universidad Austral, Buenos Aires, Argentina

1 Introduction

Over more than 65 years, scientists and engineers have been continuously leveraging the study of artificial intelligence (AI). Machines created by humans can do more than just labor-intensive work; they can also develop human-like intelligence. Whether we are aware of it or not, AI has penetrated our daily lives, playing novel roles in industry, health care, transportation, education, and many other fields, thus getting closer to all of us [1].

Artificial intelligence has emerged as a pivotal field, driving innovation across various domains. The dispersion of information not only does not facilitate the analysis of today's status, but also reveals the different objectives pursued by academia and industry. These are two distinct worlds, with immense potential if they work together.

As research in AI continues to rapidly evolve, it becomes imperative to comprehensively analyze scholarly works in this domain. Our aim is to bridge the gap between the industry and academia by presenting the present status and spotlighting the top AI players worldwide. To achieve this, we have conducted an analysis of the trends in AI from January 2019 to March 2024, utilizing the freely available open knowledge repository of academic literature and patent metadata on Lens.org. Lens Lab, an initiative of Cambia and the Knowledge Futures Group at MIT, serves as a data collaborative for open innovation data and related analytics, tools, and metrics. It includes global patent datasets, citation graphs among patents and scholarship, and metrics derived from these datasets. The important partners collaborating with Lens Lab provide insight into the quality of its results [2].

Research in this discipline encompasses a wide range of topics and interests, shaping the trajectory of technological innovation and societal transformation. In this context, our analysis aims to address several key research questions: (a) Which countries exhibit the highest production of AI research papers? (b) Which institutions are the most relevant contributors to AI research? (c) Who are the authors with the greatest impact, based on research paper publications, citations, and registered patents? (d) Which research topics are being developed in AI?

Additionally, for patents, the questions are (a) Which countries exhibit the highest production of AI patents? (b) Which applicants are the most relevant contributors to AI patent development? (c) Who are the inventors with the greatest impact, based on registered patents? Which patent types are the most relevant in AI?

This paper outlines the analysis of related research papers and patents, highlighting their transformative impact across various sectors. The methodology used allows us to understand the current state and future directions of AI. It seeks to answer key research questions regarding AI production, contributors, impact, research topics, and patent types. Future research directions may involve deeper analyses of specific research topics, tracking emerging trends, exploring interdisciplinary collaborations, and monitoring the evolving landscape of AI patents and scholarly works.

2 Related Works

AI exhibits a wide range of applications. However, there is a significant gap in the literature covering AI comprehensively. While numerous studies focus on specific topics within the realm of AI, few endeavor to encapsulate AI. Some research has been conducted on patents related to specific AI topics.

One such comprehensive survey [3] introduces the evolution of artificial intelligence and subsequently analyzes key AI technologies, including computer vision, machine learning, natural language processing, human–computer interaction, virtual reality, and biometric recognition. Additionally, it outlines the developmental trajectory of AI technology. Another compelling study provides a systematic exploration of innovative research avenues surrounding AI, unraveling the backdrop and significant implications of AI within the innovation landscape. It identifies technological, social, and economic factors driving companies to adopt AI for innovation. Furthermore, it identifies product innovation, process innovation, business model innovation, and social innovation as key consequences of AI deployment [4].

Another interesting article analyzes AI journals from the previous decade (2008–2018) [5]. The Web of Science database was chosen, and the 10% of the articles with more citations were selected (12,301 articles). They proposed the five most related topics, which are perception intelligence (1), simulated intelligence of the human mind (2), machine learning based on classical models (3), intelligence inspired by biology (4), and intelligence based on big data (5). See the discussion section for comparison of results.

Only one paper addresses both aspects: patents and research articles. It compares the development of AI in China (CH) and the EU based on a survey of papers and patents [6]. The study spans from 2015 to 2019, predating the current article, and focuses on specific AI subjects, including artificial intelligence, neural networks, expert systems, genetic algorithms, support vector machines, and machine learning. In both countries, the number of papers increases over time, with CH experiencing a threefold increase. Regarding patents, the search was conducted using the Derwent Innovations Index. The researchers observe that CH and the US collectively hold approximately 89% of all patents worldwide. They further note that in recent years, the number of patents issued in CH has been on the rise, along with its share of the global patent market.

3 Research Papers

The search strategy employed for analyzing the last five years of research work production focused on specific criteria to ensure relevance and comprehensiveness. The strategy included: (a) Publication Date: the period was set from January 1, 2019, to March 1, 2024, to capture recent publications and developments in the field; (b) Publication Type: three primary publication types were considered: book chapters,



Fig. 1 World distribution of research work (left) and patents (right)

journal articles, and conference proceedings articles, which ensured a comprehensive coverage of scholarly literature across various formats; (c) Field of Study: the search was narrowed down to the field of artificial intelligence, ensuring that only relevant publications within this specific domain were included in the analysis. By utilizing these keywords and filters, the search strategy aimed to gather a focused and up-to-date collection of scholarly works in artificial intelligence published within the defined period and across various publication types.

The overall AI research works during 2019–2024 were 15.52% of the article production (198,163) of the scholarly AI research. Figure 1 shows the world distribution of the research work. Only 3136 research articles cite patents.

Asia clearly leads the ranking; the countries that exhibit the highest production of AI research papers are CH, with 50,225 articles (25.36%), followed by United States with 22,287 (11.25%). The other countries in the ranking are India 10,380 (5.24%), UK 8003 (4.04%), Germany 5501 (2.78%), Australia 5193 (2.62%), Republic of Korea 4386 (2.21%), Italy 4337 (2.19%), Canada 4280 (2.16%), and Japan 3704 (1.87%).

Analyzing institutional collaborations and author’s affiliations, we may conclude that leading universities play a crucial role in advancing AI research globally. Nine of the top ten most relevant institutions that have made notable contributions to AI literature are across Asia (9), and one is in North America (1). The latter, Harvard University, is also the only private research university.

The first institution, the Chinese Academy of Sciences is a National Public university with 74 years of existence, 69,000 faculty members, and 59,985 students. The academic staff/student’s ratio they have equals 1.15. Harvard University is the oldest (388 years old) with 2400 faculty members and 21,613 students. The top 10 university ranking is shown in Table 1.

As the formula to calculate the prominent author profile rank is not explicitly provided by Lens.org, we decided to build an average rank based on the number of research paper publications, citations, and registered patents by author, using a proprietary algorithm to calculate it. The ranking, led by Asia, is shown in Table 2.

Many Chinese authors share the same name: “Wei Wang”, without ORCID, so what we can say is that this name leads the ranking in two categories, because of writing in 292 papers and being cited in 19 patents, followed by name Yang Liu with 253 written papers and cited in 13 patents and Yang Yang with 214 written papers,

Table 1 University ranking

#	Universities	Papers	Students	Faculty	Ratio	Age	Type
1	Chinese Academy of Sciences	5112	59,985	69,000	1.15	74	National Public
2	Tsinghua University	1739	50,390	3565	0.07	113	Public
3	Shanghai Jiao Tong University	1598	18,004	3061	0.17	128	Public
4	Zhejiang University	1558	65,821	4557	0.07	97	Public
5	Huazhong University of Science and Technology	1443	55,000	3448	0.06	71	National Public
6	Peking University	1278	44,730	11,337	0.25	126	Public
7	Harvard University	1151	21,613	2400	0.11	388	Private research
8	Beihang University	1138	37,706	3359	0.09	72	Public research
9	Harbin Institute of Technology	1094	29,368	2957	0.10	104	Public
10	University of Electronic Science and Technology of China	1073	40,000	3800	0.10	68	Public

The highest values in each column are highlighted in bold numbers

Table 2 Author's name ranking

#	Author's name	Papers	Total citations in scholar works	Total citations in patents	Region
1	Wei Wang	292	2808	19	Asia
2	Yang Liu	253	2388	13	Asia
3	Yang Yang	214	2854	13	Asia
4	Wei Li	242	4222	5	Asia
5	Wei Zhang	204	1722	2	Asia
6	Xin Wang	196	1527	6	Asia
7	Lei Wang	187	1387	9	Asia
8	Hao Wang	182	1127	2	Asia
9	Wei Liu	172	1936	10	Asia
10	Jun Wang	191	2214	2	Asia

The highest values in each column are highlighted in bold numbers



Fig. 2 Field of study word cloud

and cited in 2854 research papers and 13 patents. Wei Li leads the cited research papers category with 4222 cites, holding the fourth place.

Figure 2 highlights the distribution of research publications across key fields within the realm of artificial intelligence (AI), AI boasting a substantial 198,163 publications, indicating the significant role it plays. Computer science follows closely behind, with 189,538 publications, emphasizing the close relationship between AI and computer science disciplines. This highlights the significant overlap and mutual influence between the two fields, as AI innovations often drive advancements in computer science and vice versa. Machine learning and deep learning are also prominent areas of study, with 64,631 and 29,561 publications, respectively. Other notable fields include “artificial neural network,” “computer vision,” “data mining,” and “pattern recognition (psychology),” each contributing significantly to the breadth and depth of AI research. Furthermore, the inclusion of fields such as “Biology,” “Engineering,” “Medicine,” “Philosophy,” and “Physics”, among others, underscores the interdisciplinary nature of AI research. This interdisciplinary approach reflects the multifaceted challenges and opportunities presented by AI, as researchers draw upon insights and methodologies from various domains to address complex problems and drive innovation.

The top ten research paper field of study ranking is convolutional neural networks (1), computer vision (3), generative AI (acoustics) (1), generative neural networks (1), adversarial systems (1), facial recognition (1), explainable AI (1), and machine learning decision trees (health) (1). The top ten cases of study are 2D pose estimation, object detection, face recognition, graph neural networks, interpretable models, and explainable AI for trees.

4 Patents

Intellectual property encompasses creations originating from the mind, including inventions, literary and artistic works, designs, symbols, as well as names and images utilized in commercial contexts. Legal protection such as patents, copyrights, and trademarks safeguard intellectual property, granting individuals the opportunity to gain recognition or financial rewards for their inventions or creations. An invention

refers to a novel product or method that offers a fresh approach to accomplishing a task, and it must be revealed to the public through a patent application.

Essentially, having a patent grants the owner exclusive rights to control the commercial use of their invention. This means that others cannot produce, use, sell, distribute, or import the patented invention without obtaining permission from the patent holder.

Patents are limited to specific territories. Typically, the exclusive rights provided by a patent apply only within the country or region where the patent has been officially filed and approved, in accordance with the respective laws of that authority.

Our search was conducted using the following search stream: “artificial intelligence” OR “inteligencia artificial” OR “intelligence artificielle” OR “künstliche Intelligenz” OR “intelligenza artificiale” OR 인공지능 OR 人工智能 OR 人工知能 OR “искусственный интеллект” OR “inteligência artificial.” The topic “artificial intelligence” was included in the most important languages, namely Spanish, French, German, Italian, Korean, Chinese, Japanese, Russian, and Portuguese. Arabic was not included as there were no patents registered in that language in the database. Acronyms, such as IA, were excluded as they did not improve the search accuracy. The date search interval ranged from January 2019 to March 2024. Only two legal statuses were included: “active,” indicating that a patent has been granted and is still valid, and “pending,” indicating that the examination of a patent application is ongoing.

The Lens application identified 316,883 patent records across all districts. The overall active and pending patents during 2019–2024 were 34,146,538 patents, which represents 0.92% of patent applications. The US boasts the highest number of patents at 180,085, representing 56% of the total. CH follows in the second place, with 63,503 patents.

The World Intellectual Property Organization (WIPO) has 43,522 patents, which are not included in the counts of the two countries. European Patents account for 20,359 patents valid in European countries.

In the eastern part of the world, three countries lead the ranking after China: South Korea with 8736 patents, Japan with 715, and Taiwan with 707 patents. In the western hemisphere, the US and Canada lead with 180,085 and 509 patents, respectively, followed by the United Kingdom with 244. Australia has 507 patents. Latin America is led by Mexico, with 81 patents.

Multinational and international companies lead the artificial intelligence (AI) market. Samsung Electronics Co. Ltd. holds the highest number of patents globally, with 17,089 patents. It is a multinational electronics and information technology company headquartered in Samsung Town, Seoul, South Korea, and it is the main subsidiary of the Samsung Group.

The second company is IBM, with 9286 patents. IBM produces and sells computer hardware and software, providing infrastructure, Internet hosting, and consulting services in various computing-related fields, spanning from mainframe computers to nanotechnology. It holds more patents than any other US technology company.

The third company is LG Electronics Inc., with 8913 patents. It develops technological advances in electronics, mobile communications, and household appliances. With headquarters in Seoul, South Korea, it is one of the largest electronic conglomerates globally.

Four of the top ten companies have headquarters in CH. Tencent Tech Shenzhen Co. Ltd. holds 7155 patents and is a world-leading Internet and technology company, founded in 1998, with headquarters in Shenzhen. Beijing Baidu Netcom Sci & Tech Co. Ltd. holds 6604 patents and provides Internet access and Internet-based services.

Huawei Technologies Co. Ltd. holds 5466 patents and is a leading global provider of information and communications technology (ICT) infrastructure and smart devices. It is headquartered in Shenzhen, Guangdong.

Ping An Tech Shenzhen Co. Ltd. holds 3906 patents and is the core technology arm of Ping An Group, using AI, cloud, and other innovative technologies to develop and operate mission-critical platforms and services that support financial services, medical health, automotive services, real estate services, and smart cities. It is headquartered in Guangdong Sheng.

Three of the top ten companies are headquartered in the USA. Qualcomm Inc., which holds 6380 patents, is a multinational corporation based in San Diego, California, which specializes in the development of semiconductors, software, and services associated with wireless technology. Intel Corp. holds 5010 patents and is a manufacturer of integrated circuits. AT&T Inc. holds 4296 patents and is a multinational telecommunications holding company based at Whitacre Tower in Downtown Dallas, Dallas, Texas.

We do not know much about inventors profiles', when compared to those of researchers', because there is limited available data about the former. See in Fig. 3 inventors ranking. It has only been possible to identify their relationship with their workplace.

Two companies house these top inventors. At Qualcomm INC, Dr. Tao Luo serves as Senior Director of Engineering, having joined in 2004 [7]. Dr. Li Junyi has held the position of Vice President of Engineering at Qualcomm since 2006 [8], and Zhou Yan has been Staff Engineer since 2006 [9].

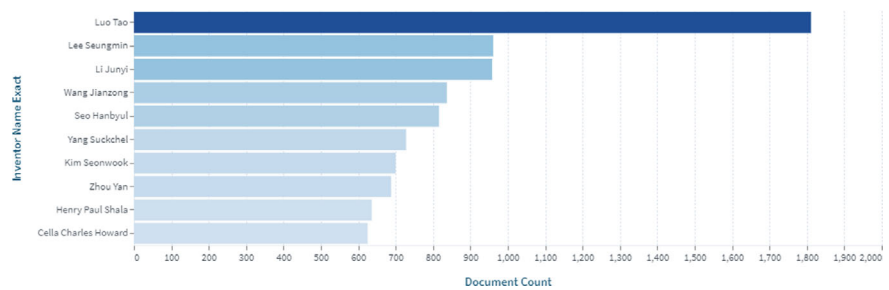


Fig. 3 Inventors ranking

At LG Electronics, Lee Seungmin is Principal Inventor [10], Hanbyul Seo is Senior Research Fellow who has been working since 2008 [11], and Kim Seonwook has been Research Engineer since 2014 [12]. Additionally, Yang Suck-chel has several patents with LG Electronics.

Moreover, Paul Shala Henry has developed patents at AT&T Intellectual Property I, L.P. [13].

The Cooperative Patent Classification (CPC) was selected to analyze the top types of patents. The European Patent Office (EPO) and the US Patent and Trademark Office jointly manage it. We identified five clusters of patents: (a) Computing arrangements based on specific computational models (G06N), composed by 39,218 patents of machine learning, 39,089 of learning methods, 31,693 of combinations of networks, 12,128 of backpropagation, 10,913 of recurrent networks, 9884 of architecture, 8555 of probabilistic graphical models, and 7810 of inference or reasoning models patents. (b) Image data processing or generation (G06T), composed by 12,055 training learning and 11,664 artificial neural networks patents. (c) Image or video recognition or understanding (G06V), using 13,534 neural networks, and 7709 classification patents. (d) Electric digital data processing (G06F), composed by 11,109 generating training patterns bootstrap methods, and 9989 semantic analysis patents. (e) Healthcare informatics (G16H), composed by 11,958 computer-aided diagnosis patents.

5 Discussion

We note that the number of patents (316,883) is bigger than the number of papers (198,163); only 3136 patents cite research articles (1.58%) and 6511 articles cite patents (3.28%). This indicates that there could be more collaboration between academia and industry to enhance the societal impact of research. While maintaining innovation is essential for research agendas, there might be missed opportunities for resource acquisition if there is insufficient focus on product development.

Eastern countries are prominent in the AI field, with many leading researchers and inventors having names of Eastern origin. However, this does not necessarily mean they are affiliated with or located in Eastern countries. As AI continues to reshape society, Western countries and universities must consider how to motivate and prepare students for AI-related education. A solid foundation in mathematics is crucial to cultivate capable individuals in this context.

The data illustrates that CH has been leading research since the previous decade (2008–2018). During the same period, the US held the second position, but India, previously ranked sixth, has now risen to the third place, surpassing countries like the UK, Australia, and Spain. Additionally, the Chinese Academy of Sciences has been at the forefront of universities since 2008. All other Chinese universities listed in Table 1 were among the top 20 universities of the previous decade, indicating an increase in their research output. The author Juan Wang is also ranked among the top ten authors in this bibliometric study and in ours. Even though the referred paper has

a different classification criterion, the fields recognized as prominent in our study are also included in the cluster analysis of high-frequency keywords, which evidences the persistent interest in the area.

US leads in patent ownership. However, there are factors such as country policies, strategies regarding intellectual property, regulatory frameworks, and market value which differ between the east and the west, which may influence this scenario. For example, Huawei, a company headquartered in CH, holds more patents in US (1690) than in CH (1263), while IBM holds 91 % of its patents in US. Additionally, globalization and the presence of international or regional entities holding patents, which are indeed valuable initiatives, serve as instruments to facilitate the internationalization of patents.

AI presents a broad spectrum of applications, and research innovations or patents are often classified into multiple categories. The top-ranked types of patents could guide academic research efforts, directing more attention toward them. “Computer science” or “computing arrangements based on specific computational models” are particularly significant, as they form the foundation of AI developments. Interest in computer vision and medicine is evident in both academia and industry. “Electric digital data processing” is not ranked as a top subject in academia. Further analysis per subject is required to clarify finer gaps between academia and industry.

Lens is a valuable tool for analyzing research articles and patents on a particular issue. It is an open initiative of MIT, freely available for research purposes, which allows users to explore related works from academia and industry. We believe it is important to train professors and students to incorporate these two aspects into their related works sections when applicable. Lens collects a wide range of data from research and industry sources, including major editorials and patent sources.

The triple helix innovation model [14], proposed as a basis for scientific and technological policy, outlines the relations among governments, universities that provide the knowledge resulting from technological research and companies that provide funds to obtain such knowledge to transfer it for the benefit of their productive structures. Like everything that is human, this model represents an ecosystem whose purpose is to generate benefits and allow social growth. Moreover, it needs the joint effort of all parties to achieve that common good. The results clearly show the weak cohesion between the objectives of the propeller blades. This, unsurprisingly, creates mutual obstacles in trying to achieve a solid boost in research and patent development of artificial intelligence technologies. However, faced with the new challenges posed by our current world and, especially, innovative technologies such as AI, the simple proper functioning of each propeller is insufficient. In her book, entitled “Technology and virtues” [15], the philosopher Shannon Vallor invites us to be aware of the power that is currently in the hands of humanity and the consequent need to nurture personally and socially the virtues that command and channel technology to achieve common happiness.

Moreover, it is important to point out that we only used the Lens platform to extract the data about research papers and patents. The research paper tool allows to refine the search by selecting “Field of Study”; on the contrary, the patent tool does not have this option, which may introduce an error in the total number of

patents when it is compared to the total number of research papers, but this does not invalidate the conclusions and the percentages informed. This study must be validated with equivalent searches in other knowledge platforms such as Scopus (limited to research papers) and Espacenet [16] (limited to patents, which is supported by the European Patent Register). However, it may be possible to fully validate the results of this study in the case of patents, if we use other tools such as Google Patents [17] or Questel [18] or patsnap [19], although these two last tools have the limitation that cannot be freely accessed. In addition, it is important to note that this document provides the overview of the last 5 years, as our goal is to understand the current status of this technology. One aspect to consider when analyzing the results is that Lens uses ML techniques for the association of data from diverse sources and the elaboration of different rankings. We understand that this aspect can generate some differences in the possible comparison with the results of other knowledge bases.

6 Conclusion

In summary, our analysis reveals Asia, particularly CH, is the leading producer of artificial intelligence research articles, accounting for 37.30% of the total publications. Following closely is North America, primarily the US (13.41%), and Europe (9%). Institutional affiliations of authors lie in Asia, with the Chinese Academy of Sciences leading the list with 5112 articles, while Harvard University stands as the sole North American institution in the top rankings, at the 7th position. Notably, Wei Wang emerges as the top author, with 292 papers cited by 2808 research works and 19 patents. Thematic clusters such as computer vision, convolutional and neural networks, adversarial systems, facial recognition, explainable AI, and ML decision trees dominate the research landscape. From the analysis, it emerges that a sizable portion of articles (40%) are open access (OA), with 80% of them receiving funding from governments, private companies (NVIDIA, Google, Microsoft, and InsightFace), and educational institutions.

On the other hand, US leads the patent market, and CH follows in the second place. Also, the organizations that hold international patents as WIPO or European Patents hold a significant number of them. Multinational and international companies lead the artificial intelligence (AI) market, the first ones being Samsung Electronics Co. Ltd. in the east and IBM in the west. At the same time, LG Electronics Inc. and several companies of CH have an important participation in world market. The top inventors are gathered at Qualcomm Inc. and at LG Electronics. Machine learning, learning methods, combinations of networks, backpropagation, recurrent networks, probabilistic graphical models, and inference or reasoning models are the top topics for patents. Also, electrical digital and image data processing, healthcare informatics, and image or video recognition are subjects that companies are interested in.

Research articles that cite patents are only 3.28% of the total number of research articles. This shows that academia and industry must work together more often, to generate more value for society.

Another aspect that may be included in future works is the identification of research papers and patents that have more impact in academy and industry. For that we plan to do a network connection analysis to better understand the topography of this ecosystem (academy, industry, and government). We are planning to detect main actors, fields of interest, and government strategic areas of development related with AI. Also, it is important to replicate this study per region to understand the evolution of AI in each one and lead their resources to world interest, to unify efforts.

To conclude, these findings offer critical insights into the global dynamics of AI research and collaboration, guiding future endeavors in this rapidly evolving field.

Acknowledgements This work is supported by a research grant from Engineering School, Universidad Austral, Argentina.

References

1. Jiang Y, Li X, Luo H, Yin S, Kaynak O (2022) Quo vadis artificial intelligence? *Discov Artif Intell* 2(1):4
2. <https://about.lens.org/>
3. Li Z, Wang Y, Ji Y, Yang W (2020) A survey of the development of artificial intelligence technology. In: 2020 3rd International conference on unmanned systems (ICUS), Nov 2020, pp 1126–1129. <https://doi.org/10.1109/ICUS50048.2020.9274952>
4. Mariani MM, Machado I, Nambisan S (2023) Types of innovation and artificial intelligence: a systematic quantitative literature review and research agenda. *J Bus Res* 155: 113364. <https://doi.org/10.1016/j.jbusres.2022.113364>
5. Gao F, Jia X, Zhao Z, Chen CC, Xu F, Geng Z, Song X (2021) Bibliometric analysis on tendency and topics of artificial intelligence over last decade. *Microsyst Technol* 27:1545–1557
6. Feng J, Li X (2021) Research on the development trend of artificial intelligence based on papers and patent analysis—data comparison between China and the United States. In: Liang Q, Wang W, Mu J, Liu X, Na Z, Cai X (eds), *Artificial intelligence in China*. Springer, Singapore, pp 60–67. https://doi.org/10.1007/978-981-15-8599-9_8
7. <https://www.qualcomm.com/news/onq/2022/09/qualcomm-inventor-tao-luo-on-how-your-5g-phone-connects-to-a-cel>
8. <https://www.linkedin.com/in/junyi-li-93775a>
9. <https://www.linkedin.com/in/yan-zhou-72a1b8b2/>
10. <https://ieeexplore.ieee.org/author/37086958142>
11. <https://www.linkedin.com/in/hanbyul-seo-48453194/?originalSubdomain=kr>
12. <https://ieeexplore.ieee.org/author/37404960000>
13. https://en.wikipedia.org/wiki/List_of_prolific_inventors
14. Etzkowitz H, Leydesdorff L (1995) The triple helix—university-industry-government relations: a laboratory for knowledge based economic development. *EASST Rev* 14(1):14–19
15. Vallor S (2016) *Technology and the virtues: a philosophical guide to a future worth wanting*. Oxford University Press
16. https://worldwide.espacenet.com/advancedSearch?locale=en_EP
17. <https://patents.google.com/>
18. <https://www.questel.com/>
19. <https://www.patsnap.com/>

A Novel Factorized with Label Smoothing Cognitive Architectural Framework for Apple Leaf Disease Prognosis



Arindam Kanrar, Swayantan Maji, V. Sanjay, Sushruta Mishra,
and Hassan M. Al-Jawahry

Abstract The agricultural output and global food security have greatly been concerned by the apple plant diseases. This research examines the probability of applying deep learning, particularly the InceptionV3 model with L2 regularizer (0.01 strength), for image-centric apple plant diseases (like apple scab, black rot, and cedar apple rust) detection and classification. The current study builds ground-work for the possible or inventive agricultural techniques to manage crop health effectively despite environmental and economic adversities. The approach proposed by us includes data collection, image preprocessing and the advancement of a deep learning model based on transfer learning and achieves 97% accuracy. The paper concludes by giving attention to optimal model architectures, neural network techniques, and dataset enlargement techniques as future applications in the agricultural field.

Keywords Disease diagnosis · Cognitive computing · Deep learning · Transfer learning

A. Kanrar · S. Maji · S. Mishra (✉)

Kalinga Institute of Industrial Technology, Deemed to Be University, Bhubaneswar, India

e-mail: sushruta.mishrafcs@kiit.ac.in

V. Sanjay

Vellore Institute of Technology, Vellore, India

H. M. Al-Jawahry

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University, Najaf, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Al Diwaniyah, Al Diwaniyah, Iraq

Department of Computers Techniques Engineering, College of Technical Engineering, The Islamic University of Babylon, Babylon, Iraq

1 Introduction

Apples are considered highly nutritious and rich in nutrients, making them one of the most important fruits. The reason that apples are highly consumed is their low price and many health beneficiaries like high fiber, vitamins, minerals, and rich antioxidants. Increased agricultural production is needed to improve disease resistance. Plants and trees struggle with biotic stressors along with adverse climatic conditions. Pests and diseases reduce harvest and also negatively affect fruit quality. So it is necessary to protect crops from disease in order to improve their quality and quantity. Correct diagnosis and timely response are important steps in effective disease management. Early and targeted screening during disease attacks and with accuracy is critical. To use medical strategies for preventing the spread of the disease, fitting knowledge of the disease is needed to take proper precautions and use appropriate dosage of disinfectants, thus saving economy and environmental benefits. *Venturia inaequalis*, the fungus that gives rise to apple scab, is one of the most alarming fungal diseases affecting apples in fair climates. It impacts an apple tree's vegetation as well as its apples. Fruit that has contracted scab disease will have corky lesions, velvety spores, and a dark brown tint. However, diseases that harm apple vegetation have a negative impact on production and the economy, in addition to lowering the yield and quality of fruit produced by the fruit business. Therefore, there is a great demand for accurate methods to detect illnesses of apple leaves. In current years, conventional machine learning techniques have made some progress and helped in apple leaf disease recognition. In this context, numerous methods of plant disease diagnosis and imaging techniques have been studied. However, they require high-precision instrumentation and large sensors, which result in high cost and inefficiency. Deep learning is a technique in which a computer model learns ways to perform classification tasks based on images, text, or sound. These models have the potential to achieve high classification accuracy, surpassing humans in many cases. The basic training architecture uses multilayer neural networks. Deep learning techniques have been successfully applied in many real-life applications and more recently in agriculture. In recent years, digital cameras and other popular electronic devices have increasingly used automatic plant disease diagnosis through machine learning. As a satisfactory alternative machine learning method, the recognition of realistic plant models is now more relevant in deep learning, and transformation networks have made important advances in computer vision and reported many related theories and benefits.

In this research, the main objective is developing a sturdy and accurate deep learning model, based on the InceptionV3 architecture that can detect and classify apple plant diseases using the leaf images. The idea is to combine various agriculture-related concepts that can be focused by deep learning techniques, thereby enhancing the education of agriculture and allowing the transfer of the knowledge for use in other areas. The study focuses on the development of an image-based classification tool that can warn farmers regarding any possible plant disease. The study promotes sustainable fertilizers that will make crops healthier and very unlikely to be killed

by diseases. The study also evaluates the suitability of InceptionV3 for the purposes of plant pathology. Summarizing it, the main aim is to have a system that is cost effective and able to offer real-time reports on the status of the crop for farmers to be in a position to manage. In this study we are primarily focusing on apple plant disease.

The main contribution of this research is:

- Usage of deep learning models, mainly the InceptionV3 architecture, for the detection of apple leaf diseases by applying regularization techniques for reducing overfitting and improving the generalization of the model.
- Hardline experimentation with the InceptionV3 architecture, comparing performance with and without L2 regularization, evaluation of the impact on model stability and generalization, using a dataset comprising images of apple leaves with various diseases.
- Comparative analysis with other deep learning architectures, for checking better performance architectures for image classification tasks.
- Enabling early detection of apple leaf diseases, facilitating farmers with timely information for crop protection, health, and productivity, improving agricultural yields and sustainability, minimizing the use of extensive pesticides, and reducing environmental pollution promoting eco-friendly agriculture.

2 Related Works

In [1] the team of researchers coming from Umm Al Qura University in Saudi Arabia have presented a disease detection system that employed convolutional neural networks (CNN) in order to classify diseases in apples. Team had a dataset collected from nearby markets which helped in achieving more than 90% classification rate, and their best model was using AlexNet and transfer learning methods attaining 99.17 accuracy rate. Another study [2] showing the need of an effective and correct knowledge of disease chose a 19-layer CNN for classifying Marssonina Coronaria and Apple Scab diseases. A sample of 50,000 images has also been included to improve the precision of the model, thus obtaining 99.2% accuracy. Otherwise, an article [3] by Bansal, Kumar, and Kumar consists of a strong ensemble model for disease detection in apple leaves displaying a great 96.25% accuracy. The research plainly notes the economic effect derived from labor-based agronomic identification which operates on slow basis. Besides, our suggested model will include the developed web app for the living in the outlying areas too. On the other hand, a work [4] article develops a new model called XDNet for the apple tree leaf disease (ATLD) recognition which is a combination of DenseNet and Xception that has achieved 98.82% accuracy. Augmenting database realism and applying the augmentation issue XDNet model demonstrate the future of detecting diseases in agriculture by using our mobile gadgets. Another work [5] by the team of four Baranwal, Khandelwal, Arora, Jaypee Institute of Information Technology, Noida, India, investigates detection of apple tree diseases using CNNs based on apple tree disease dataset, resulting in highly accurate

results. The methodology is based on the prior researches including Mohanty et al. who published their paper on using CNNs for plant disease classification. On one hand, the report describes the related works including the latest breakthroughs made on image-based disease detection while on the other hand providing the application of deep learning in agriculture. This paper [6] implements a deep learning technique for detection of apple leaf diseases in real time with the help of convolutional neural networks (CNNs) upgraded by most advanced deep learning architectures. It includes a set of data consisting of field images which are both augmented and annotated for training 26,377 images of diseased apple leaves. The SSD with Inception module and rainbow concatenation (INAR-SSD) model is the proposed using GoogLeNet Inception structure and rainbow concatenation. The model provides 78.8% mAP on ALDD and runs at FPS of 23.13. Relatively, the INAR-SSD model is more accurate and faster, comes first followed by the RSSD and DSSD models. [7] focuses on the importance of correct diagnosis of diseases characterized with negative effects on the quality of fruits and their processing. Various types of datasets containing both real images and synthetic images were taken into use. The review covers methodologies like DCGAN for the synthetic images and DCNN for the disease classification. Optimization of hyperparameters like learning rate gives better model performance. The review identifies a lack of future research needed. In [8] the model is being initialized with transfer learning and achieves results with ~ 97% accuracy on the presented dataset. The work focuses mainly on the various diseases on the apple leaves where infection level is more than 70% in some places. The model has a mean accuracy of 97.2%, and there is a notable high precision, recall, specificity, and F1-score values. The model accuracy also changes depending on type of disease, and the Powdery Mildew is observed with highest precision and recall values. The dataset includes images from different orchards in the Kashmir valley. This paper [9] represents a relationship between agriculture and its influence on economies and highlights the importance of dealing with crop diseases especially in apple trees which are highly prone to diseases like apple scab and cedar rust, affecting yield. The study focuses on a deep learning technique for disease detection using transfer learning along with machine algebra features of VGG16, ResNetV2, InceptionV3 and MobileNetV2 models. Disease detection in the above areas is assisted by computer vision. The researchers concentrate on employing transfer learning in convolutional neural networks (CNNs) to identify plant leaf diseases and are able to obtain an accuracy of 94.7% with the ResNet V2 model. Overall, the research has shown many possibilities of using deep learning methods for detecting and controlling diseases in agriculture. In the paper [4], a new convolutional neural network (CNN) model with modified AlexNet architecture comprising optimized GoogLeNet Inception and NAG is designed to identify apple leaf diseases with high accuracy focusing on 13,689 created images. The model recognizes the diseases with high accuracy of 97.62%, producing better results compared to other models, and considerably saves parameters compared to standard AlexNet. Further study involves adoption of the high-end deep neural networks like faster RCNN, YOLO, and SSD for optimized solutions in detection of diseases in real time. All in all, the research helps in recognition of diseases in apple orchards which focuses on the power of deep learning in

agricultural disease management and underlines the need to enlarge datasets and use the advanced neural network architectures for better detection.

3 Proposed Methodology

- Data Collection

At first a diverse dataset of apple leaf images was collected, encompassing a healthy and three types of diseases black rot, apple scab, and cedar apple rust, and each image was stored in their corresponding category [10]. This made our dataset.

- Data/Image Preprocessing

The dataset was split into training and testing/validating. The images were first decoded to a particular format (.jpg) and then resized to the input size for the InceptionV3 model (224×224 pixels). The training data was augmented by techniques such as rotation, flipping, and zooming to increase robustness [11], thus finally generating the training and testing/validating set.

- The Deep Learning Model

This model is based on transfer learning technique, where the weights are frozen off the convolutional layers to retain the learned features. The InceptionV3 model was chosen as the base architecture for feature extraction [12–14]. The initial classification layer of InceptionV3 was not used. The ‘imagenet’ weights were used in the hidden layer for training. The output layer was formatted as a global average pooling 2D layer. Finally, a custom layer was added, which was the final output dense layer consisting of regularizer ‘L2 (0.01)’ and softmax function (Fig. 1).

InceptionV3 model employs deep convolutional layers with inception modules, integrates multiple convolutional operations of varying sizes within each layer, enables efficient feature extraction across multiple spatial scales, and includes auxiliary classifiers for regularization, achieving strong performance in image classification tasks. It is developed by Google Research.

‘imagenet’ weights refers to pre-trained model weights that have been learned by training a convolutional neural network on the ImageNet dataset.

Global average pooling 2D layer : It reduces the spatial dimensions of each feature map in the network to a single value by averaging over all spatial locations. Softmax function : Mathematical function to convert a vector of raw scores into a probability distribution over multiple classes [15].

Regularizer ‘L2 (0.01)’ : This regularization term is added to the model’s loss function during training to penalize large weights in the network, helping to prevent overfitting by encouraging the weights to stay small. It is somewhat related to support

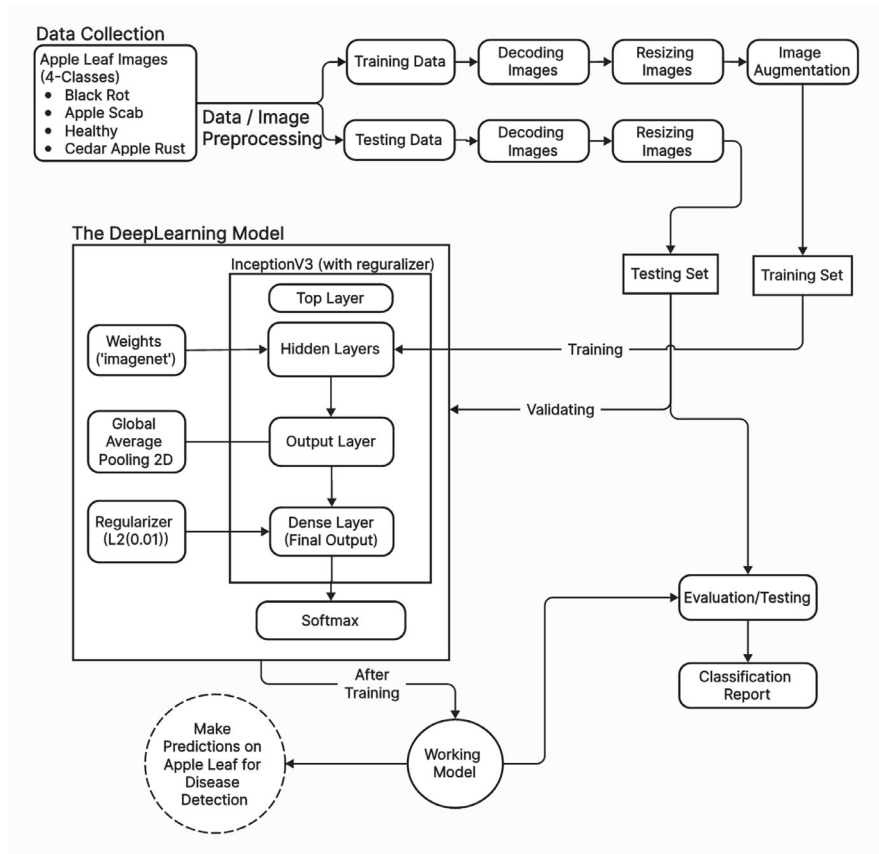


Fig. 1 Workflow for the proposed methodology

vector machines (SVMs) in the sense that both methods aim to prevent overfitting and improve the generalization performance of machine learning models.

The regularized loss function can be written as, $Loss = \text{Original Loss} + \lambda \sum_i \omega_i^2$ Where original loss is the original loss function without regularization. λ is the regularization parameter, and ω_i^2 are the weights of the neural network. The term $\lambda \sum_i \omega_i^2$ penalizes large weights by adding a term proportional to the square of each weight to the loss function. This encourages the optimization process to find weight values that minimize both the original loss and the magnitude of the weights. This is somewhat similar to SVMs.

- **Working Model**

After the training is complete, the model is saved. This thus becomes the working model, which can be used to make predictions for apple disease using apple leaf images.

4 Result Analysis

In this paper, the deep learning model was trained using InceptionV3 architecture along with the weights from ‘imagenet’. A regularizer (L2 0.01) was also used in the process. This allover has shown a notable impact on the smoothness of the validation accuracy graph as depicted in Fig. 3. When the model was trained without a regularizer, the validation accuracy graph showed roughness, identified by fluctuations and large gaps between epochs as depicted in Fig. 2. The roughness in the validation accuracy curve depicts potential instability during training and indicates overfitting (the model is fitting too closely to the training data and failing to generalize unseen data). Without applying the regularization (L2 0.01), the model may capture noise and irrelevant patterns from the training data, leading to decreased generalization performance.

In contrast, training the same model with the inclusion of an L2 regularizer (with a regularization strength of 0.01) results in a smooth validation accuracy graph. The smoothness of the validation accuracy curve indicates more stable learning throughout training and suggests better generalization performance of the model to unseen data. The presence of the L2 regularizer helps to smooth out fluctuations in the validation accuracy curve by penalizing large weights, thus reducing the risk of overfitting and promoting better generalization as we can observe in Figs. 2 and 3.

During the training of InceptionV3 with the ‘imagenet’ weights, the addition of L2 regularizer (0.01 strength) causes a gentle change in the epochs without the abrupt

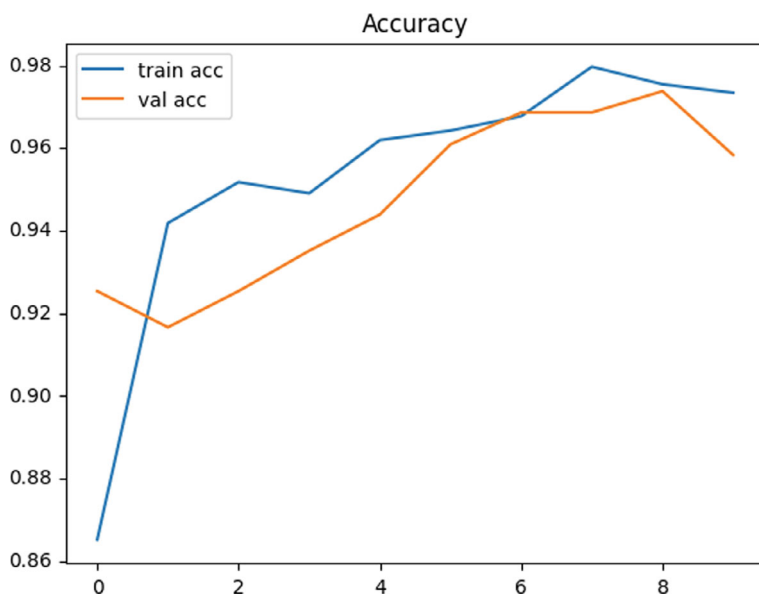


Fig. 2 Accuracy graph before applying regularizer

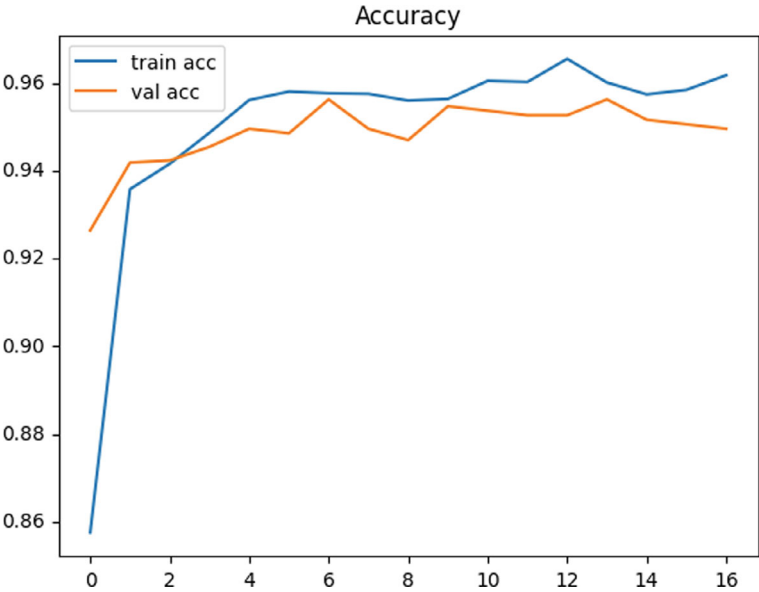


Fig. 3 Accuracy graph after applying regularizer

downfall of accuracy curve, which demonstrates a much better generalization and stability of learning. Here is the proof (Table 1) that shows that this new model detects a broad range of apple leaf diseases effectively, for example, apple scab, black rot, and also cedar apple rust (and among others healthy apple leaf). On account of the high precision, recall, and *F1*-scores across various classes and on a remarkable overall accuracy of 97%, it could be readily remarked that the model has a chance of being applied in a practical setting in the agriculture sector for early disease detection and management.

With the utilization of regularizers the models demonstrate the gradual convergence and better generalization that outplays the nonsmooth general ones. InceptionV3 and ResNet50 demonstrate that they dominate over the fundamental CNN layer throughout the majority of epochs. InceptionV3 with regularizer and ResNet50 presents clear performances on learning which makes an obvious generalization

Table 1 Classification report

	Precision	Recall	<i>F1</i> -score	Support
Apple scab	0.93	0.92	0.92	504
Black rot	0.92	0.94	0.93	497
Healthy	0.97	0.94	0.96	502
Cedar apple rust	0.92	0.94	0.93	440

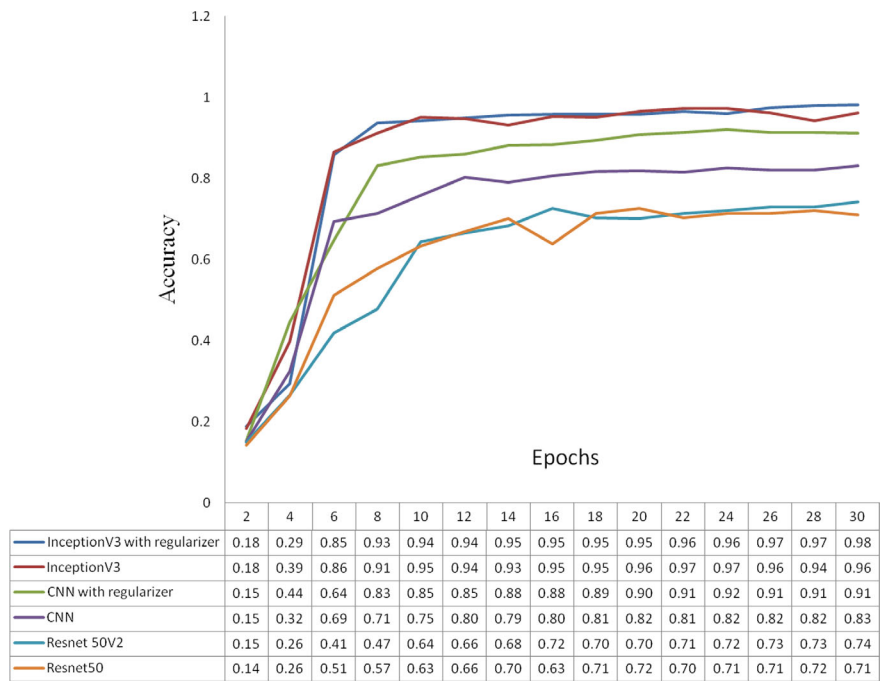


Fig. 4 Accuracy measures of different models

advantage in the frame of regularization. A curve with lower performance differences between the training and validation sets and fewer fluctuations represents the regularization as solving the overfitting problem. Then, an extensive step of fine-tuning and hyperparameter optimization might give more accurate outcomes of the model (Fig. 4).

5 Conclusion

The research paper explores the effectiveness of InceptionV3 architecture along with L2 regularizer for detecting apple leaf diseases with the aim of enhancing the model performance and generalization. During the research it was found that using L2 regularizer (0.01 strength) in the model improved instability and overfitting, hence resulted in smoother validation accuracy graphs, improved generalization performance, and overall achieved high accuracy (97%) in detecting various apple diseases. In comparative study underlining the better feature learning abilities, InceptionV3 and ResNet architectures beat the simpler CNN architectures. In future to reduce the performance gaps between different architectures, tuning the hyperparameter

and further optimization are suggested. This will ensure the development of more efficient and accurate disease detection systems.

References

1. Alharbi AG, Arif M (2020) Detection and classification of apple diseases using convolutional neural networks. In: 2020 2nd International Conference on Computer and Information Sciences (ICCIS), pp 1–6
2. Singh S, Gupta I, Gupta S, Koundal D, Mahajan S, Pandit A (2021) Deep learning based automated detection of diseases from apple leaf images. *Comput Mater Contin* 71:1849–1866. <https://doi.org/10.32604/cmc.2022.021875>
3. Bansal P, Kumar R, Kumar S (2021) Disease detection in apple leaves using deep convolutional neural network. *Agriculture*, 11:617
4. Liu B, Zhang Y, He D, Li Y (2017) Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry* 10:11
5. Baranwal S, Khandelwal S, Arora A (2019) Deep learning convolutional neural network for apple leaves disease detection. *Web Technol eJ*
6. Jiang P, Chen Y, Liu B, He D, Liang C (2019) Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access* 7:59069–59080
7. Ayaz H, Rodríguez-Esparza E, Ahmad M, Oliva D, Pérez-Cisneros M, Sarkar R (2021). Classification of apple disease based on non-linear deep features. *Appl Sci* 11:6422
8. Khan AI, Quadri SM, Bandy S (2020) Deep learning for apple diseases: classification and identification. *Int J Comput Intell Stud* 10:1–12
9. Alsayed A, Alsabei A, Arif M (2021) Classification of apple tree leaves diseases using deep learning methods. *Int J Comput Sci & Netw Secur* 21:324–330
10. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
11. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digit Health* 10:20552076241256732
12. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-DETECTOR: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
13. Pranjali P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. In: AIP conference proceedings, vol 2919(1). AIP Publishing
14. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber-physical system. *Internet Things* 24:100978
15. Mishra S, Volety DR, Bohra N, Alfarhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306

Blood Pressure Estimation from Photoplethysmogram Signals: An EdgeML Approach



Nithin Aditya Pradeepkumar and Krishna Karthik Bhaskar

Abstract The purpose of this research work is to explore the viability of photoplethysmogram (PPG) as a continuous, noninvasive and cuffless method of blood pressure (BP) monitoring. It uses Edge Machine Learning Models to estimate Blood Pressure from PPG Signals in a time sensitive manner. The study reports mean absolute error (MAE) of 5.12 mmHg for Systolic Blood Pressure (SBP) and 2.58 mmHg for Diastolic Blood Pressure (DBP) estimation. This new methodology, however, makes use of Edge ML Models to estimate blood pressure along side Deep Learning models hence bypasses most limitations in Deep learning approaches.

Keywords Blood pressure · Hypertension · Machine learning · Edge computing

1 Introduction

Hypertension can also be called as excessive blood pressure (BP), and it is a quite serious disease. As reported by WHO in 2014, around 9.4 million deaths are caused by it every year [1]. The number of people suffering from hypertension has been growing steadily from 594 in 1975 to 1.13 billion in 2015 [2]. In a poll conducted in 2008, the frequency of high blood pressure is 24.8% of women and 29.2% of men [3]. In Egypt, hypertension affects about 26.3% of the adult population, and more than half of the hypertensives are above 60 years of age [4].

The use of traditional methods for measurement of blood pressure such as sphygmomanometers has given way to discomfort and the likelihood for continuous monitoring by this method is low [5, 6]. Invasive procedures such as intra-arterial monitoring tend to be limited to hospital situations [7].

N. A. Pradeepkumar (✉) · K. K. Bhaskar

Department of Computational Intelligence, SRM Institute of Science and Technology,
Kattankalathur, India

e-mail: np1239@srmist.edu.in

K. K. Bhaskar

e-mail: kk0937@srmist.edu.in

Seeing these challenges, photoplethysmography (PPG) can be a perfect solution for continuous non-invasive blood pressure monitoring. This study presents a method for BP prediction based on machine learning. The method that is being presented for the prediction of blood pressure (BP) is in compliance with ISO and BHS guidelines [8]. Now with relative absolute accuracy of systolic blood pressure of 4.13 mmHg and diastolic blood pressure of 2.48 mmHg, this approach stands to be reliable and precise.

A lightweight, cost-effective BP estimation framework is one of this method's main breakthroughs. It streamlines monitoring by eliminating the need for expert intervention. Unlike previous methods, the method uses cycle-by-cycle feature extraction. This method extracts and feeds individual signal cycles into the regressor, providing a detailed information of the BP morphology.

2 Related Works

Two categories can be used to classify machine-learning-based blood pressure estimation studies. First, PPG signal features are retrieved and machine-learning models are developed to predict blood pressure. The second method estimates arterial blood pressure (ABP) directly from the PPG signal. Pulse Transit Time (PTT) is another popular method. PTT is the time between the systolic peak of the fingertip photoplethysmogram (PPG) and the R peak of the electrocardiogram (ECG) during a heart cycle [9]. Since PTT has a negative relationship with blood pressure, many calibration methods have been devised to predict blood pressure using PTT. See [10, 11].

A popular method for estimating blood pressure is regression analysis. Kachuee et al. [12] investigated PPG signal characteristics using typical machine learning techniques. Multimodal feature extraction from ECG signals is another popular method. Sun et al. [13] developed a feature extraction approach for accurate BP estimation from PPG and ECG signals. Although regression-based BP estimate methods require complex feature engineering, they consistently show good accuracy.

Matsumura et al. [14] in their work "Blood Pressure Estimation Using photoplethysmography Only: Comparison between Different Machine Learning Approaches" compared three machine learning algorithms to estimate BPs using PPG only and revealed that the regression tree algorithm was the best approach with overall acceptable accuracy to ISO standard for BP device validation.

Combining multiple physiological markers to improve BP estimate is becoming more important as this field advances. Researchers want to better understand cardiovascular system dynamics by combining PPG and ECG data. Multimodal techniques improve BP estimation and enable more reliable and accessible healthcare technologies.

A recent study by Rishivardhan et al. [15] proposed BP-Net, a U-Net architecture for calculating ABP waveforms from PPG waveforms. This novel method uses the similarities between the two signals to predict ABP, from which standard peak detection algorithms can calculate DBP and SBP. Our project aims to revolutionize

microcontroller blood pressure assessment using a novel feature engineering technique and EdgeML model. This project is vital to improving healthcare technology accessibility and efficiency, especially in resource-constrained situations.

3 Methodology

Our method extracts several PPG signal features from each cycle, unlike earlier research. Our feature selection process prioritizes adaptability across multiple datasets over prediction accuracy for a single dataset. This focus on feature robustness improves our blood pressure (BP) estimate method’s generalizability and reliability.

3.1 Dataset

The data used for this study was obtained from the MIMIC II dataset [16]. The dataset includes waveforms from numerous ICU patients. It includes fingertip photoplethysmography (PPG), Arterial Blood Pressure (ABP), and Electrocardiogram (ECG) captured at 125 Hz (Fig. 1).

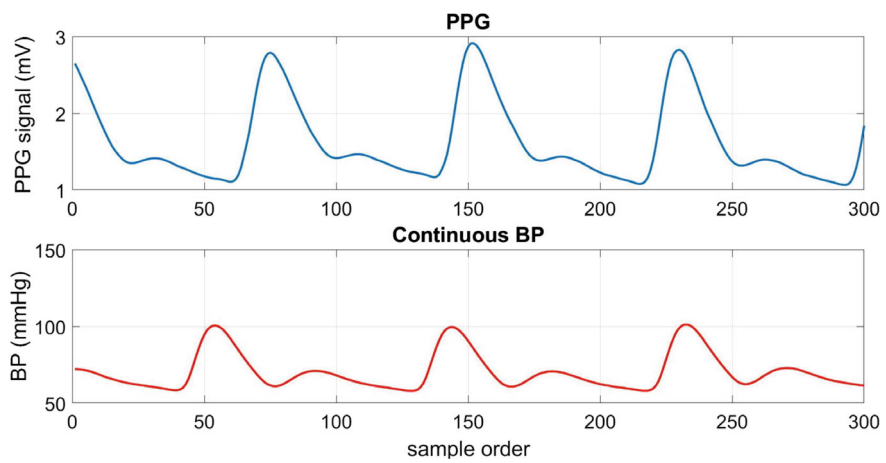


Fig. 1 PPG and ABP waves from the UCI dataset

Table 1 BP ranges in the dataset

	Min (mmHg)	Max (mmHg)	STD (mmHg)	Mean (mmHg)
DBP	60.9	126.3	8.9	69.9
MAP	69.1	135.8	9.4	92.9
SBP	80.8	177.7	19.2	137.9

3.2 Preprocessing

The UCI Dataset used in this study, underwent a filtering and denoising process identical to that outlined in their work. Among several preprocessing techniques evaluated, discrete wavelet decomposition (DWT) was chosen for its superior performance in terms of complexity, phase response, and noise handling capabilities. We used the Butterworth filter, a type of signal processing filter intended to provide a frequency response in the passband that is as uniformly flat as theoretically possible, to remove movement artifacts and denoise the signals [17]. Furthermore, the Butterworth filter—also referred to as a flat maximum magnitude filter—seeks to guarantee uniformity in signal properties among various frequency components. Subsequently, the ABP and PPG signals underwent phase-matching based on their phase discrepancy. This phase offset is regarded as a lag in signal analysis. As the databases include signals recorded using various devices, cross-correlation was conducted to align the phase difference between the two signal windows [11].

In order to improve the precision and resilience of the feature extraction procedure, the PPG signals’ amplitude was calibrated within the interval [0, 1]. Subject-to-subject variations in PPG amplitude lead to many obstacles, including the intricacy of cycle-per-cycle segmentation for feature extraction and the reliance of some extracted characteristics on PPG amplitude. By guaranteeing amplitude consistency between participants, normalization of PPG signals mitigates these difficulties and promotes effective feature extraction while guaranteeing the significance of the recovered feature values [18]. For this, the min-max normalization approach was used.

At last, entries exhibiting abnormal blood pressure values, notably those exceeding $SBP \geq 180$ or falling below $SBP \leq 80$, as well as $DBP \geq 130$ or $DBP \leq 60$, were omitted from the dataset. Additionally, instances with missing data were also excluded. Table 1 illustrates the statistical summary of the dataset after pre-processing.

3.3 Feature Extraction

Establishing a reliable model for blood pressure (BP) estimation requires the identification of features that are important to the task. Particularly those that accurately capture the characteristics associated with BP fluctuations. Nevertheless, identifying such characteristics is difficult due to the intrinsic heterogeneity in PPG waveforms

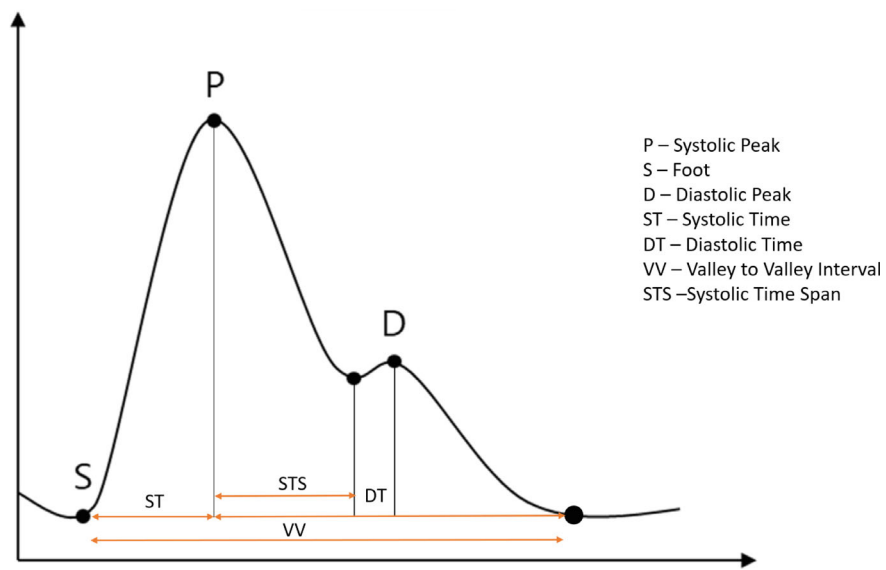


Fig. 2 Extracted features from the PPG signal

between individuals [18]. In contrast to established physiological indicators like heart rate or oxygen saturation, PPG waveforms can vary considerably among individuals. This heterogeneity arises from various sources, including anatomical distinctions, vascular characteristics, and physiological variables [19].

In our approach, features are extracted for each each cycle in a signal. We consider 10 s segments of a PPG Signal to estimate the blood pressure. A 10 s PPG signal usually consists of 7–8 cycles. Features are extracted for all the cycles present, wherever the features can be extracted. Cycles which do not have a significant diastolic notch are dropped [20]. The features along with the corresponding blood pressure values, which will be the target value, for the cycles are extracted.

After several experimentation, eight features from the PPG signal (shown in Fig. 2) were considered. The features and their explanation are listen in Table 2.

3.4 Machine Learning

Following feature extraction, PPG cycles lacking extractable features were excluded. Features were then extracted from a total of 79,486 cycles originating from 912 subjects and 12,000 PPG signals. Throughout the machine learning process, measures were taken to prevent data leakage. Specifically, cycles from a given signal were exclusively included in either the training or testing set, but not both, to maintain the integrity of the model evaluation.

Table 2 Explanation of features from the PPG signal

Feature	Explanation
Heart rate	The number of heartbeats per unit of time, typically measured in beats per minute (BPM)
Reflection index	An index reflecting the magnitude of reflections of arterial pressure waves caused by the impedance mismatch at bifurcations and arterial junction
Systolic timespan	The duration of the systolic phase of the cardiac cycle, representing the time interval between the onset and end of systole
Up time	The time taken for the arterial pressure waveform to reach its maximum amplitude during the systolic phase
Systolic volume	The volume of blood ejected from the heart during each systolic phase
Diastolic volume	The volume of blood remaining in the heart at the end of diastole, just before the next systolic contraction
Diastolic peak	The peak value of the arterial pressure waveform during the diastolic phase, typically occurring just before the next systolic contraction
Systolic peak	The peak value of the arterial pressure waveform during the systolic phase, representing the maximum pressure exerted by the heart during contraction

Thirty percent of the collected features dataset was utilized for testing the models, while seventy percent was used for training machine learning models. To assess diastolic and systolic blood pressure independently, models were developed. The models were assessed using the mean absolute error (MAE) following training. First, the machine learning models were trained using a dataset containing features extracted from photoplethysmogram (PPG) signals along with corresponding blood pressure (BP) measurements. This process involves adjusting the parameters of the models iteratively to minimize the error between the predicted BP values and the true BP measurements in the training data.

Subsequently, the edge model, specifically Bonsai in this case, was trained using the same dataset. The goal of training the edge model is to achieve comparable or superior performance to the conventional machine learning models while leveraging the advantages of edge computing, such as reduced computational resources and real-time processing. Edge models like Bonsai are designed to be lightweight and optimized for deployment on resource-constrained devices. Comparing the performance of these models with traditional machine learning models helps determine if similar levels of accuracy can be achieved with reduced computational requirements.

4 Experiments and Results

In this section, we present the results of our experiments comparing the performance of various machine learning (ML) models and an edge ML model, Bonsai, for estimating diastolic blood pressure (DBP) and systolic blood pressure (SBP).

4.1 Performance of ML Models

Table 3 shows the performance metrics of different ML models for DBP estimation. Ensemble models like RandomForestRegressor and KNeighborsRegressor outperform other models in terms of Mean Absolute Error (MAE) and Root Mean Square (RMSE). These ensemble models leverage the collective wisdom of multiple base models, leading to better generalization and robustness.

Similar trends are observed for SBP estimation, as shown in Table 4. RandomForestRegressor and KNeighborsRegressor exhibit lower MAE and RMSE compared to other models.

Table 3 Performance of ML models for DBP estimation

Model	MAE	RMSE	Model size (kB)
RandomForestRegressor	2.48	4.00	426,248.16
LinearRegression	14.26	17.78	1.03
KNeighborsRegressor	2.62	4.29	9047.29
SVR	14.75	18.19	4980.37
DecisionTreeRegressor	3.21	5.61	6708.02

Table 4 Performance of ML models for SBP estimation

Model	MAE	RMSE	Model size (kB)
RandomForestRegressor	4.13	5.77	142,450.14
LinearRegression	12.34	16.39	1.01
KNeighborsRegressor	5.22	34,603.811	2792.91
SVR	12.83	16.84	1525.23
DecisionTreeRegressor	4.19	6.39	2232.13

Table 5 Performance of Bonsai edge ML model

Metric	DBP	SBP
MAE	2.58	5.12
RMSE	4.01	6.07
Model size (kB)	0.347	0.423

4.2 Performance of Edge ML Model

We evaluated the performance of the Bonsai edge ML model for DBP and SBP estimation. Remarkably, Bonsai achieves competitive results with significantly smaller model sizes compared to traditional ML models. Table 5 summarizes the performance of Bonsai for both DBP and SBP estimation.

4.3 Discussion

The RandomForestRegressor and the KNeighborsRegressor are ensemble models that are effective in their performance when compared to other ML models due to their flexibility with overfitting and handling intricate data relationships.

Each of these Bonsai edge ML models has the same MAE and RMSE score as classic ML models but at much less model size. This is because edge ML models are effective. The small size of Bonsai makes inference on edge devices practical. The ability of the Bonsai edge ML model to miniaturize compared to normal ML models highlight the advantages of edge computing. Healthcare systems are able to deploy edge ML models with the quality and performance expected from more conventional practices.

5 Conclusion

We have also analyzed the use of photoplethysmography (PPG) in conjunction with machine learning (ML) models and edge ML models for continuous blood pressure (BP) monitoring. Through thorough experimentation and analysis we showed how efficient the aimed outcome was in estimating both systolic and diastolic levels of blood pressure.

The experiments we conducted indicated that ensemble ML models: RandomForestRegressor and KNeighborsRegressor, were more accurate and robust than other ML models in the task of estimating BP. As opposed to other models, these are built upon several base learners so that the PPGL signal can be captured effectively enabling precise prediction.

In addition, we tested the Bonsai edge ML model and showed its significant capability in BP estimation achieving comparable accuracy with rather smaller models. The relative compactness of the bonsai model's parameters promotes fast embedding and usage of edge gadgets thereby enabling quick BP readings at various health care settings. From this evaluation of the conventional ML approaches and the Bonsai edge TM, some appreciation of the benefits of performing edge computation on health care problems is brought out. With edge ML models, health care systems would be improved by factors of speed and more access.

In particular, our results highlight the potential of the use of advanced systems-based technology and non-invasive wearable health devices for improved care in patients and continuous tracking of key health features. In our opinion, the strategy we support has a great potential for transforming the existing models of BP control, while facilitating the movement toward more personalized, preventive medicine.

References

1. World Health Statistics (n.d.) <https://www.who.int/data/gho/publications/world-health-statistics>
2. Kachuee M, Kiani Mohammad M, Mohammadzade H, Shabany M (2017) Cuffless blood pressure estimation algorithms for continuous health-care monitoring. *IEEE Trans Biomed Eng* 64:859–869. <https://doi.org/10.1109/TBME.2016.2580904>
3. Amin M, Radwan A, Hemeda A (2020) Cardio-metabolic problem in Egypt. *Intern Emerg Med* 15:549–552. <https://doi.org/10.1007/s11739-020-02346-8>
4. Ghazy M, Elgindi MB, Wei D (2018) Analytical and numerical investigations of the collapse of blood vessels with nonlinear wall material embedded in nonlinear soft tissues. *Alex Eng J* 57:3437–3450. <https://doi.org/10.1016/j.aej.2018.03.002>
5. Park SH, Park YS (2019) Can an automatic oscillometric device replace a mercury sphygmomanometer on blood pressure measurement? A systematic review and meta-analysis. *Blood Press Monit* 24:265–276. <https://doi.org/10.1097/MBP.0000000000000412>
6. Chandrasekhar A, Yavarimanesh M, Hahn J-O, Sung S-H, Chen C-H, Cheng H-M, Mukkamala R (2019) Formulas to explain popular oscillometric blood pressure estimation algorithms. *Front Physiol* 10:1415. <https://doi.org/10.3389/fphys.2019.01415>
7. Stergiou GS, Alpert B, Mieke S, Asmar R, Atkins N, Eckert S, Frick G, Friedman B, Graßl T, Ichikawa T, Ioannidis JP, Lacy P, McManus R, Murray A, Myers M, Palatini P, Parati G, Quinn D, Sarkis J, Shennan A, Usuda T, Wang J, Wu CO, O'Brien E (2018) A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement. *J Hypertens* 36(3):472–478. <https://doi.org/10.1097/HJH.0000000000001634>. PMID: 29384983; PMCID: PMC5796427
8. Mousavi SS, Firouzmand M, Charimi M, Hemmati M, Moghadam M, Ghorbani Y (2019) Blood pressure estimation from appropriate and inappropriate PPG signals using a whole-based method. *Biomed Signal Process Control* 47:196–206. <https://doi.org/10.1016/j.bspc.2018.08.022>
9. Wong MYM, Poon CC, Zhang Y (2009) An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: a half year study on normotensive subjects. *Cardiovasc Eng* 9(1):32–38. <https://pubmed.ncbi.nlm.nih.gov/19381806/>
10. Ding X, Zhang Y (2019) Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm. *Biomed Eng Lett* 9(1):37–52. <https://pubmed.ncbi.nlm.nih.gov/30956879/>

11. Gesche H, Grosskurth D, Küchler G, Patzak A (2011) Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method. *Eur J Appl Physiol* 112:309–315
12. Kachuee M, Kiani MM, Mohammadzade H, Shabany M (2017) Cuffless blood pressure estimation algorithms for continuous health-care monitoring. *IEEE Trans Biomed Eng* 64(4):859–869
13. Sun S, Bresch E, Muehlsteff J, Schmitt L, Long X, Bezemer R, Paulussen I, Noordergraaf GJ, Aarts RM (2023) Systolic blood pressure estimation using ECG and PPG in patients undergoing surgery. *Biomed Signal Process Control* 79(part 1):104040. <https://doi.org/10.1016/j.bspc.2022.104040>. ISSN 1746-8094
14. Matsumura K, Rolfe P, Toda S, Yamakoshi T (2018) Cuffless blood pressure estimation using only a smartphone. *Sci Rep* 8:1–9. <https://doi.org/10.1038/s41598-018-25681-5>
15. Rishivardhan K et al (2021) BP-Net: efficient deep learning for continuous arterial blood pressure estimation using photoplethysmogram. In: 2021 20th IEEE international conference on machine learning and applications (ICMLA), pp 1495–1500
16. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):E215–E220
17. IETB PTU (2012) Study of signal denoising using Kaiser window and Butterworth filter. India
18. Palatini P, Asmar R, O'Brien E, Padwal R, Parati G, Sarkis J, Stergiou G (2020) Recommendations for blood pressure measurement in large arms in research and clinical practice: position paper of the European Society of Hypertension working group on blood pressure monitoring and cardiovascular variability. *J Hypertens* 38:1244–1250. <https://doi.org/10.1097/HJH.0000000000002399>
19. Stergiou G, Palatini P, Asmar R, Ioannidis JP, Kollias A, Lacy P, Mcmanus RJ, Myers M, Parati G, Shennan A, Wang J, O'Brien E. Recommendations and practical guidance for performing and reporting validation studies according to the Universal Standard for the validation of blood pressure measuring devices by the Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO). Erratum
20. Bilo G et al (2010) Validation of the Somnotouch-NIBP noninvasive continuous blood pressure monitor according to the European Society of Hypertension International Protocol revision. *Blood Press Monit* 20(2015):291–294. <https://doi.org/10.1097/MBP.0000000000000124>

Secure and Verify Exchange Data in VANET Using Blockchain



Zahraa Sh. Alzaidi , Ali A. Yassin , Zaid Ameen Abduljabbar ,
and Vincent Omollo Nyangaresi

Abstract Vehicular ad hoc networks (VANETs) represent a burgeoning research domain with far-reaching implications for enhancing safety, efficiency, and public involvement in transportation systems. Embedded with sensors, processing capabilities, and wireless communication, vehicles stand poised to revolutionize various facets of road safety, operational efficiency, and user engagement. There are, however, big problems with real-time processing of the huge amounts of data sent through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) channels because VANETs are automatically changing, with inconsistent connectivity and different vehicle speeds. Furthermore, the inherent variability within VANETs compounds the challenge of establishing dependable communication pathways and ensuring efficient data transmission while simultaneously necessitating robust measures to address security concerns such as secure communication protocols, reputation management, and safeguarding against malicious attacks. This paper proposes a scheme to verify the security and high efficiency of message exchange in VANETs

Z. Sh. Alzaidi · A. A. Yassin · Z. A. Abduljabbar (✉)

Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Basrah, Iraq 61004

e-mail: zaid.ameen@uobasrah.edu.iq

Z. Sh. Alzaidi

e-mail: pgs.zahraa.shaker@uobasrah.edu.iq

A. A. Yassin

e-mail: ali.yassin@uobasrah.edu.iq

V. O. Nyangaresi

Department of Computer Science and Software Engineering, Jaramogi Oginga Odinga University of Science and Technology, Bondo, Kenya 40601

Department of Applied Electronics, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India 602105

V. O. Nyangaresi

e-mail: vnyangaresi@jooust.ac.ke

Z. A. Abduljabbar

Department of Business Management, A-Imam University College, Balad, Iraq 34011

by integrating fog computing with blockchain technology. Using the decentralized structure of blockchain and the distributed computing features of fog computing, our proposed framework protects the integrity of data while lowering the costs of storage, computation, and communication. Notably, the proposed scheme is fortified against various forms of attacks, including impersonation attacks, man-in-the-middle (MITM) attacks, and replay attacks.

Keywords VANETs · Blockchain technology · Fog computing · Shearing messages · Security · Integrity

1 Introduction

The creation of an accurate, precise, and effective scheme for traffic operations is made possible by the combination of electrical, communication, and computer technologies in an intelligent transportation system [1]. The term VANETs describes a system in which individual nodes are mobile and may interact with other vehicle and roadside units. They provide self-organized data transfer, roadside operations-related navigation, safety precautions, and other services [2]. Short-range wireless communication is used by autos to exchange data, enabling real-time information exchange. This interaction covers a number of topics, such as road status, weather, and traffic information. The advantages of having access to such real-time data could enable vehicles react quickly to collisions or ease traffic congestion.

Every vehicle equipped with VANETs utilizes on-board units (OBUs) and tamper-proof devices (TPDs). OBUs are typically composed of resource processors, storage units, sensors, and communication interface modules. These components play a vital role in enabling the exchange of information between neighboring vehicles and roadside infrastructure. TPDs are utilized to securely store encrypted information related to vehicles, thereby minimizing the potential for unauthorized tampering by malicious individuals with access to sensitive or unprocessed data. The advancements in IoT, big data, fog computing, and blockchain technologies have greatly improved the communication performance of VANETs. When vehicles are equipped with OBUs, they gain the ability to enable mobile communication between vehicles and other entities. This involves the transfer of data between V2I and V2V, utilizing The DSRC protocol typically operates within a communication range of less than 1000 m [3, 4].

The vehicle serves as a network node, gathering relevant data and sending it to other vehicles using wireless communication techniques to enhance its ability to change its driving conditions dynamically. This method increases the car's overall driving performance by actively preventing potential crashes and reducing traffic congestion. Additionally, the car will send the information it has collected to RSUs. However, since VANETs are unconstrained, vehicle nodes communicate across open channels where information might be intercepted, forwarded, or even altered by

adversarial entities [5]. This causes traffic congestion or the development of questionable information. The limited processing and storage power of the OBU causes real-time changes in the network topology. This results in delayed message processing and unintended consequences because of the vehicle's rapid motion.

Sensitive information and private user data, such as driving routes and true identities, are among the shared data between vehicles. The privacy of users may be threatened by the existence of dynamic network topology. Conventional approaches to resolving this problem include encrypting data using other vehicles public keys. However, in terms of facilitating one-to-many data flow, this method seems to be ineffective in the setting of VANETs. Moreover, it is essential to include flexible access control strategies to manage cross-user data exchange [6].

Recently, blockchain technology has been used to VANETs to solve the performance and security problems mentioned above. It offers immutability, decentralization, and anonymity benefits [7]. By using a smart contract inside the blockchain, public data such as certificates, public keys, and pseudonyms may be managed efficiently. This makes it possible to implement authentication and revocation procedures effectively. Furthermore, no new data has to be added to the blockchain; all that is required throughout the authentication process is access to existing information from the blockchain [8].

In this study, we propose a scheme aimed at improving the reliability of VANETs via the integration of blockchain generation and fog computing. At the core of our proposal is the creation of a customized blockchain scheme designed specifically for VANET applications. This scheme will use event messages as transactions to securely distribute essential data between vehicles. In order to address the limitations imposed by geography in VANETs, we recommend maintaining distinct blockchains for each area. Furthermore, we suggest using fog computing to expand cloud services to network peripheries, such as RSUs and OBUs, in order to improve performance by transferring computational tasks to fog servers. Additionally, elliptic curve cryptography (ECC) usage facilitates effective communication load management. Our main objective is to guarantee the consistent and punctual transmission of reliable event messages in VANETs, achieved by a combination of local blockchain implementation and integration of fog computing. This integration eventually enhances the dependability and security of vehicular communication systems. The scope of our significant contributions includes the following:

- We propose a blockchain schema tailored for VANETs to store node and message trustworthiness, providing a reliable ground truth for other vehicles.
- Introducing the concept of a local blockchain to improve scalability, leveraging geographical boundaries to create independent chains distinct from those of other countries.
- Suggesting the incorporation of fog computing into the VANET blockchain to mitigate block generation latency by delegating computationally intensive tasks to edge devices, thereby facilitating the deployment of real-time applications within VANET environments.

- Proposing a rejection counter for the number of times messages from the same OBU are rejected, leading to the exclusion of the vehicle from the system if it continues to spread false messages.
- Conducting a formal analysis using the Scyther tool to validate the security robustness of our framework, demonstrating resilience against common attacks such as forgery attacks, replay attacks, and MITM attacks.
- Additionally, an informal security analysis validates the fulfillment of key security requirements, including integrity, confidentiality, decentralized identification, and non-repudiation.

This paper is organized as follows: Sect. 2 provides an extensive review of pertinent literature on data security and privacy in VANETs. In Sect. 3, we delve into the background of the system architecture. Section 4 offers a detailed explanation of our proposed schema, followed by an in-depth security analysis of the scheme in Sect. 5. Section 6 outlines the computational costs associated with our proposed schema. Lastly, Sect. 7 presents our conclusions, summarizing the key findings and contributions of our study.

2 Related Works

In order to preserve the identity privacy of the vehicle, Raya et al. proposed a privacy protection authentication technique for VANET that makes use of public key infrastructures (PKI) [9, 10]. A certificate authority is responsible for the issue and administration of certificates in this system. The method uses a pseudonymous certification to protect the vehicle's identity anonymity. However, the vehicle has to maintain a list of revoked certificates on a regular basis and pre-install a significant number of public-private key pairs together with their corresponding certificates. The vehicle will have a large amount of computational overhead and storage burden as a consequence.

The centralized nature of the existing PKI infrastructure was emphasized by Benarous et al. [12], who also put forth a blockchain-based framework for the administration of pseudonyms in VANETs. These techniques successfully prevented the ingress of malicious vehicles into VANETs, decreased the costs related to validating certificates and signatures, and established a monitoring system to revoke the access of vehicles that exhibit recurring malicious activity in VANETs. However, PKI-based authentication systems share the following shortcomings: Firstly, attestation of certificates is contingent upon a dependable CA. Secondly, substantial computational resources are necessary for the verification of certificates and signatures. Third and finally, certificate storage and key management present obstacles.

In pursuit of energy-efficient privacy for communicators and communication security, Akram et al. [13] devised an identity-based authentication system tailored for vehicular cloud computing, integrating radio frequency identification technology. As ID-based authentication schemes hinge on CAs for private key generation based on user information, effective key management emerges as a pivotal aspect of VANET

security frameworks, drawing significant research attention. Lei et al. [14] introduced frameworks for secure key management, incorporating blockchain principles to devise VANET suited key management schemes.

Shao et al. [15] utilize group signatures and threshold authentication techniques to reduce overhead in downloading and verifying certificate revocation lists (CRLs). By employing bilinear pairing-based cryptography, their approach ensures that a message is authenticated by a threshold number of vehicles before being accepted. However, the reliance on RSUs as group managers poses security risks, as compromised RSUs could potentially disclose group keys and compromise vehicle privacy by tracing their locations.

Azees et al. [16] proposed an effective anonymous authentication with conditional privacy (EAAP) scheme to mitigate the risk of malicious vehicle intrusion in VANETs. This scheme enables tracking of disruptive vehicles or RSUs within VANETs. Recently, the integration of bilinear pairing has enhanced the EAAP scheme, eliminating the need for the trusted authority (TA) to maintain anonymous certificates of OBUs and RSUs [17]. Additionally, TA possesses the authority to revoke the anonymity of non-compliant vehicles, thus revealing their identities within a group. Subsequently, the TA manages the addition of revoked identities to the identity revocation list (IRL). However, it is important to note that these schemes are limited by challenges associated with certificate management and scalability.

Dorri et al. [18] introduced a blockchain technology aimed at enhancing automotive security through the implementation of an overlay network within the blockchain framework, supplemented by overlay block managers. In this architecture, overlay network nodes are organized into clusters overseen by cluster heads, responsible for managing the blockchain and executing its core functionalities. Nonetheless, the inclusion of supplementary overlay nodes could potentially lead to increased latency and introduce single points of failure, particularly if the integrity of the cluster head is compromised.

Gong et al. [19] introduced a blockchain privacy protection scheme based on ring signatures, enabling the transmission of authorized Internet of things (IoT) data in ciphertext form to preserve data sender identity. However, their scheme lacked a tracking mechanism, rendering it unable to identify specific sources of false messages from vehicles in VANETs through message signatures. Incorporating blockchain technology for traceability can enhance data security protection and foster trusted resource sharing within VANETs.

In response to the challenges associated with computational overheads in certificate and signature verification, difficulties in certificate storage and key management, and latency issues in pseudonym issuance and CRLs distribution, our research proposes a blockchain enabled and fog computing solution. Additionally, we introduce an efficient handover mechanism for secure message transmission between vehicle users. Tailored for critical environments, our scheme demonstrates resilience against various attacks, and Table 1 provides a qualitative assessment of security features.

Table 1 Security requirements for various schemes

Characteristic	Shao et al. [15]	Azees et al. [16]	Dorri et al. [18]	Liu et al. [20]	Our schema
Decentralization	✗	✗	✓	✓	✓
Privacy protection	✓	✓	✓	✓	✓
Anonymity	✓	✓	✗	✗	✓
Tamper-proof and unforgeable	✓	✓	✓	✓	✓
Traceable	✓	✗	✓	✓	✓
Scalable	✗	✗	✓	✗	✓
Low overhead and high efficiency	✗	✓	✗	✗	✓

3 System Architecture

3.1 System Model

The VANET system model comprises five primary entities: the TA, RSU, OBU, blockchain, and fog computing. Each RSU and TA maintains an autonomous connection to the blockchain network, as presented in Fig. 1. TA is responsible for initializing VANETs and vehicle registration, while RSUs facilitate communication and message validation. OBUs installed in vehicles enable communication and data processing within ITS. RSUs strategically positioned alongside roadways communicate with vehicles using the DSRC protocol, relay messages, and securely store data within the blockchain. This interconnected system ensures seamless and secure communication within the VANET ecosystem.

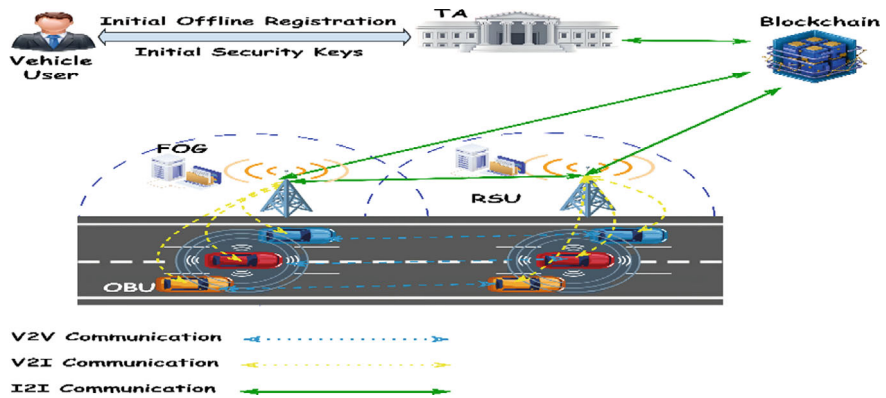


Fig. 1 System architecture proposed in our scheme

3.2 *Blockchain Technology*

Blockchain serves as a decentralized infrastructure and distributed computing paradigm that utilizes cryptographic chained block structures for data validation and storage. Consensus algorithms are employed to generate and update data, while smart contracts are utilized for programming and manipulating data. Each block in the blockchain consists of a block header and a block body. The block header includes critical components such as the current version number, the address of the previous block, a nonce, a Merkle root, the target hash value of the current block, and a timestamp [21, 22].

3.3 *Fog Computing*

Fog computing is gaining recognition as an additional computing paradigm that works in tandem with cloud computing to address the resource-constrained computational, storage, and networking requirements of VANET nodes. Fog computing augments the capabilities of VANET system by providing services at the network's periphery that are equivalent to those offered by the cloud. In addition to assessing the precise data that necessitates transmission and reception, the technology optimizes the temporal separation between the VANET networks and the cloud. By utilizing VANET data associated with the VANET application from any device with storage, computation, and network access, fog computing has the potential to improve processing at the network periphery. This is in contrast to the exclusive reliance on the cloud center for the entirety of the processing [23].

4 *Proposed Work*

This study provides an innovative scheme that ensures anonymous authentication and integrity preservation while enabling secure communication over VANETs via the use of blockchain technology and fog computing. Our system consists of three fundamental components: TA, RSUs, and OBUs.

4.1 *Setup Phase*

The setup phase is responsible for doing the fundamental initialization necessary for the succeeding phases. TA chooses a finite elliptic curve that is specified by the equation $y^2 = x^3 + ax + b \pmod{q}$, where q represents a prime number. Therefore, the symbols P and Q represent points on the elliptic curve. Afterward, TA randomly

Table 2 List of notations and explanation

Notations	Explanation
ΔT	Acceptable difference between T_s, T_r
a, b	Large prime numbers
P and Q	Points of the finite elliptic curve
Z_q^*	Non-zero elements of a finite field Z_q
δ, ρ, τ	Random numbers chosen from Z_q^* by TA
T_{PR}	Private key of TA
T_{PU}	Public key of TA
H	Secure hash function
D_{IDV}	Dummy identity for vehicle user
D_{IDR}	Dummy identity for RSU
Sig_i	Signature of a message
ms_i	Original message

chooses τ and δ from the multiplicative group Z_q^* , where q represents the group's size. Subsequently, the private key (T_{PR}) is favored over τ . The determines their public key $T_{PU} = \tau P$ and verification key $T_V = \delta P$. In addition, the TA chooses a Hash function $H: \{0, 1\}^*$ and distributes ($T_{PU}, T_V, H, P, Q, e(P, Q), q$) to all users (OBUs) and RSUs in the VANET network (Table 2).

4.2 Secure Exchange Message Phase (SEM)

The SEM phase is essential to VANET systems. For automobiles and drivers functioning in a dynamic environment, the VANET provides significant network connectivity and data exchange capabilities. The aim of VANET is to facilitate data interchange inside vehicles that operate in a dynamic environment and provide broad network connectivity for drivers and passengers while they are driving. Vehicles have the potential to communicate and distribute traffic data, including many features like traffic patterns, weather, collisions, and gridlock. After receiving this information, other vehicles could be aware of the problem beforehand and help to create a safer driving environment. As shown by Algorithm 1, the SEM phases of vehicular communication involve a well thought-out series of steps.

Step 1: Initialization begins with the OBU selecting ℓ_i from the set Z_q^* and computing essential parameters, including $X_1 = \ell_i T_{PU}$, $Y_1 = \vartheta_i T_{PU}$, and $F = X_1 + Y_1$.

Step 2: The OBU then selects a_i from Z_q^* and calculates $A_i = a_i T_{PU}$. This is followed by the determination of $\mu_i = H(\mathbf{ms}_i \times A_i)$ and subsequently computing $v_i = \mu_i(a_i + \ell_i + \vartheta_i) \bmod q$.

Step 3: A meticulous creation of the authentic signature $\mathbf{Sig}_i = (A_i, \mathbf{ms}_i)$ is executed to fortify message integrity against tampering.

Step 4: The timestamp discrepancy (ΔT) is computed as $T_r - T_s$, where T_s represents the current timestamp of message transmission, and T_r denotes the timestamp upon receiving the message. This parameter signifies the allowable time difference between T_s and T_r .

Step 5: $\text{OBU}_i \rightarrow \text{OBU}_j/\text{RSU}$. The OBU diligently transmits a comprehensive message package containing v_i , T_s , $\mathbf{ms}_i$, $\mathbf{Sig}_i$, F , D_{IDV} and $\mathbf{Rc}_i$ to the intended recipient, which may be another OBU or an RSU.

Step 6: The OBU conducts a swift evaluation of the rejected counter ($\mathbf{Rc}_i$). If $\mathbf{Rc}_i$ equals 3, the OBU_j blocking the OBU_i , otherwise, the protocol proceeds to step 7.

Step 7: OBU/RSU. Upon receipt, the recipient, whether an OBU or an RSU, rigorously validates the timestamp's currency ($T_r - T_s \leq \Delta T$). Subsequently, meticulous computation of μ_i' via $\mu_i' = H(\mathbf{ms}_i \times A_i)$ is conducted. This validation involves scrutinizing the coherence in $v_i T_{PU} = \mu_i(A_i + F)$. Compliance with these stringent criteria authenticates the message ($\mathbf{ms}_i$). Any deviation results in rejection, signifying message obsolescence exceeding the permissible (ΔT).

Step 8: The $\mathbf{Rc}_i$ is incremented by one unit $\mathbf{Rc}_i = \mathbf{Rc}_i + 1$ and subsequently checked to determine if it reaches a threshold of three. If the counter reaches three, the OBU is blocked from transmitting messages. Otherwise, the message transmission process continues.

Algorithm 1 Secure exchange message phase

Input:

Parameters: $q, \ell_i, \vartheta_i, T_{PU}, R_{ci}, D_{IDV}$

Message: ms_i

Output:

Authentication Status: Accepted/Rejected/Blocked OBU

Begin

1. Initialize: $\ell_i \in Z_q^*$
2. Compute: $X_1 = \ell_i T_{PU}, Y_1 = \vartheta_i T_{PU}$ and $F = X_1 + Y_1$
3. $a_i \in Z_q^*$
4. Compute: $A_i = a_i T_{PU}, \mu_i = H(ms_i \times A_i)$ and $v_i = \mu_i(a_i + \ell_i + \vartheta_i) \bmod q$
5. Create Signature: $Sig_i = (A_i, ms_i)$ //to ensure message integrity
6. Calculate Timestamp Discrepancy (ΔT): $T_r - T_s$
7. Transmit: $(v_i, T_s, ms_i, Sig_i, F, D_{IDV}, \text{ and } R_{ci})$ from OBU_i to OBU_j/RSU //Message Package
8. if $(R_{ci} == 3)$
9. Blocked OBU ()
10. Else
11. if $(T_r - T_s > \Delta T)$
12. Go to 18
13. Else
14. $\mu_i' = H(ms_i \times A_i)$
15. if $(v_i T_{PU} == \mu_i'(A_i + F))$
16. Accept message (ms_i)
17. Else
18. $R_c = R_c + 1$
19. Return Rejected (ms_i)
20. End

5 Security Analysis

5.1 Informal Security Analysis

The SEM scheme proposed herein is designed to guarantee secure communication between OBU and RSUs within a decentralized fogging-based network. As such, we have undertaken a comprehensive evaluation, comprising formal and informal assessments, to ascertain its capability to fulfill security requirements and withstand prevalent attacks.

- Integrity: The proposed schema relies on tamper-proof blockchain technology and digital signatures to effectively ensure data integrity, thereby preventing unauthorized data modification. All registration and secret construction transactions

are securely recorded within the blockchain through hashing, guaranteeing their immutability. Furthermore, each generated message is uniquely characterized by a signature, denoted as $\mathbf{Sig}_i = (A_i, \mathbf{ms}_i)$. Consequently, this approach safeguards the integrity of the system by maintaining a meticulous and unalterable record of activities.

- **MITM Attack:** This form of attack entails the clandestine manipulation of communication between two i -nodes, resulting in a deceptive impression of direct interaction while the attacker maintains control over the entire exchange. In our proposed framework, cryptographic signing $\mathbf{Sig}_i = (A_i, \mathbf{ms}_i)$ is employed for every message transmitted within the VANET system. Furthermore, encryption, combined with the inclusion of a fresh timestamp (T_s) in transmitted messages, serves as a preventative measure against such attacks. The utilization of a blockchain system has demonstrated resilience against MITM attacks.
- **Reply Attack:** Our scheme uses strong defenses against replay attacks, in which adversaries take advantage of flaws in communication systems to send modified messages in a certain amount of time. In messages delivered, each transmitted message is given a timestamp. A message that is transmitted at a different time or that has been intercepted and tampered with does not reach its intended receiver. In addition, our system checks incoming messages against predefined time limits and quickly filters out those that come late, thereby preventing replay attacks and preserving communication integrity.
- **Forgery Attack:** The potential for a forgery attack is mitigated by the deployment of secure signed transactions $\mathbf{Sig}_i$, rendering it improbable even in the presence of an i -node with malicious code. In such a scenario, any forged i -node (OBU/RSU) would remain unverified by the blockchain, thus thwarting unauthorized access attempts via falsified i -nodes.
- **Non-repudiation:** The inability of the sender or recipient to contest the reception or transmission of packets to or from other end nodes is ensured by the non-repudiation principle. The suggested scheme calls for vehicle users and RSUs to register offline with the TA. In order to use this scheme, vehicle users and RSUs must provide important secret credentials such (U_i , Phno $_i$, ID $_i$, Add $_i$, VN $_i$ and VID $_i$). Following a successful verification procedure, the TA provides the necessary credentials (D_{IDV} , D_{IDR}), which are necessary to provide anonymous authentication and secure message delivery. The goal of this scheme is to stop RSUs and vehicle users from disputing their credentials or their participation.

5.2 Formal Security Analysis

Scyther tool utilizes the security protocol description language (SPDL) [24] and serves as a prominent formal verification instrument extensively utilized for validating security protocols. Our protocol underwent comprehensive verification using the automatic claim and verification schemes within the Scyther tool. Formulated

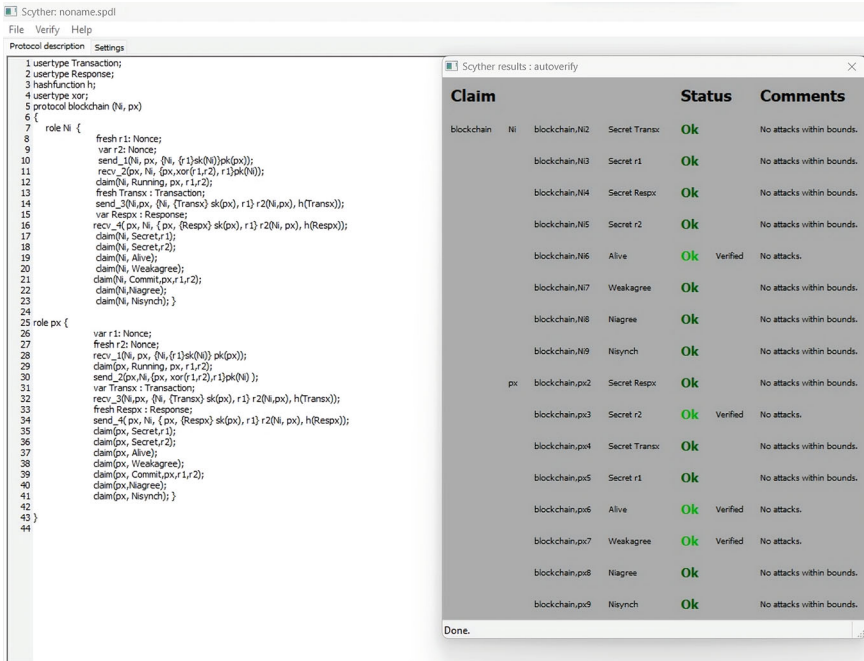


Fig. 2 Security analysis conducted using Scyther

using SPDL, the protocol exhibited robust resilience against a range of malicious intrusions, including MITM attacks, replay attacks, forgery attacks, and non-repudiation. Figure 2 depicts the SEM phase, confirming the effectiveness of our solution in mitigating diverse malicious intrusions. The analysis underscores SPDL's proficiency in executing critical cryptographic functions, ensuring secure message transmission and reception across components while delineating specific responsibilities for each component.

6 Computational Cost

In this section, we do a performance analysis for our proposal with respect to the computational cost during the message sending and verification process. We used Zhang et al.'s [25] calculation to determine how long cryptographic operations take. Our suggested method relies heavily on the cryptographic processes of point addition, multiplication, hashing functions, and pairing operations. For each message transmission and verification, our proposed method necessitates that a single OBU do five multiplication operations and two hash function operations. We calculate the total computational cost of these operations as $5Ex_{ecc-sm} + 2Ex_h = 1.611$ ms.

The obtained timings confirm the best possible use of computational resources, in addition to demonstrating outstanding speed and efficiency. Additionally, this result implies a reduction in storage requirements as well as a noticeable improvement in processing speed and resource consumption, hence enhancing the overall effectiveness of VANETs.

7 Conclusions

This paper provides a fog computing, blockchain, and digital signature-based security message exchange scheme for VANETs. The method includes a rejection counter and timestamp. The technique uses smart contracts to track illegal vehicles and send the information to TA, allowing for independent managed within a country. This improves the security of VANETs during message exchange and protects drivers' privacy. Moreover, an efficient handover mechanism is used to transmit the message among vehicle users. Utilizing fog computing to deal with the problem of limited processing capability in VANETs. In our thorough analysis of security and performance, our proposed scheme exhibits a swift SEM time of 1.611 ms for individual OBUs. This accelerated transmission efficiency is primarily ascribed to the streamlined verification process for messages, facilitating prompt determination of acceptance or rejection. Indeed, the integration of signature, timestamp, and rejected counter mechanisms enhances security measures, ensuring robust protection against potential threats. Our method encompasses various security features, including authentication, integrity, anonymity, and confidentiality. In comprehensive security and performance analyses, our scheme achieves a balanced trade-off between privacy, trust assessment, computational efficiency, communication overhead, and storage costs, demonstrating resilience against threats such as non-repudiation, replay attacks, MITM attacks, forgery, and message tampering. Moreover, our scheme outperforms conventional approaches in computational, communication, and storage efficiency while incorporating essential security features, making it highly suitable for real-time VANET applications.

References

1. Anbalagan S, Raja G, Gurumoorthy S, Suresh RD, Dev K (2023) IIDS: intelligent intrusion detection system for sustainable development in autonomous vehicles. *IEEE Trans Intell Transp Syst*
2. Dwivedi SK, Amin R, Das AK, Leung MT, Choo K-KR, Vollala S (2022) Blockchain-based vehicular ad-hoc networks: a comprehensive survey. *Ad Hoc Netw*, 102980
3. Jiang H, Hua L, Wahab L (2021) SAES: a self-checking authentication scheme with higher efficiency and security for VANET. *Peer-to-Peer Netw Appl* 14(2):528–540
4. Sharma S, Awasthi SK (2024) Zone-based stable and secure clustering technique for VANETs. *Simul Model Pract Theory* 130:102863

5. Gaba P, Raw RS, Kaiwartya O, Aljaidi M (2024) B-SAFE: blockchain-enabled security architecture for connected vehicle fog environment. *Sensors* 24(5):1515 [Online]. Available <https://www.mdpi.com/1424-8220/24/5/1515>
6. Rasori M, La Manna M, Perazzo P, Dini G (2022) A survey on attribute-based encryption schemes suitable for the internet of things. *IEEE Internet Things J* 9(11):8269–8290
7. Umran S-M, Lu S, Abduljabbar Z-A, Tang X (2023) A blockchain-based architecture for securing industrial IoTs data in electric smart grid. *Comput Mater Continua* 74(3):5389–5416 [Online]. Available <http://www.techscience.com/cmc/v74n3/50945>
8. Ali A, Iqbal MM, Jabbar S, Asghar MN, Raza U, Al-Turjman F (2022) VABLOCK: a blockchain-based secure communication in V2V network using ICN network support technology. *Microprocess Microsyst* 93:104569
9. Raya M, Hubaux J-P (2007) Securing vehicular ad hoc networks. *J Comput Secur* 15:39–68
10. Umran SM, Lu S, Abduljabbar ZA, Nyangaresi VO (2023) Multi-chain blockchain based secure data-sharing framework for industrial IoTs smart devices in petroleum industry. *Internet of Things* 24:100969. <https://doi.org/10.1016/j.iot.2023.100969>
11. Wasef A, Lu R, Lin X, Shen X (2010) Complementing public key infrastructure to secure vehicular ad hoc networks [security and privacy in emerging wireless networks]. *IEEE Wirel Commun* 17(5):22–28
12. Benarous L, Kadri B, Bouridane A (2020) Blockchain-based privacy-aware pseudonym management framework for vehicular networks. *Arab J Sci Eng* 45:6033–6049
13. Akram W, Mahmood K, Li X, Sadiq M, Lv Z, Chaudhry SA (2022) An energy-efficient and secure identity based RFID authentication scheme for vehicular cloud computing. *Comput Netw* 217:109335
14. Lei A, Cruickshank H, Cao Y, Asuquo P, Ogah CPA, Sun Z (2017) Blockchain-based dynamic key management for heterogeneous intelligent transportation systems. *IEEE Internet Things J* 4(6):1832–1843
15. Shao J, Lin X, Lu R, Zuo C (2015) A threshold anonymous authentication protocol for VANETs. *IEEE Trans Veh Technol* 65(3):1711–1720
16. Azees M, Vijayakumar P, Deboarh LJ (2017) EAAP: efficient anonymous authentication with conditional privacy-preserving scheme for vehicular ad hoc networks. *IEEE Trans Intell Transp Syst* 18(9):2467–2476
17. Blake IF, Murty VK, Xu G (2006) Refinements of Miller's algorithm for computing the Weil/Tate pairing. *J Algorithms* 58(2):134–149
18. Dorri A, Steger M, Kanhere SS, Jurdak R (2017) Blockchain: a distributed solution to automotive security and privacy. *IEEE Commun Mag* 55(12):119–125
19. Gong J, Mei Y, Xiang F, Hong H, Sun Y, Sun Z (2021) A data privacy protection scheme for Internet of things based on blockchain. *Trans Emerg Telecommun Technol* 32(5):e4010
20. Liu H, Zhang Y, Yang T (2018) Blockchain-enabled security in electric vehicles cloud and edge computing. *IEEE Network* 32(3):78–83
21. Umran SM, Lu S, Abduljabbar ZA, Zhu J, Wu J (2021) Secure data of industrial Internet of Things in a cement factory based on a blockchain technology. *Appl Sci* 11(14):6376 [Online]. Available <https://www.mdpi.com/2076-3417/11/14/6376>
22. Bowlin E, Khan MS, Bajracharya B, Appasani B, Bizon N (2023) Challenges and solutions for vehicular ad-hoc networks based on lightweight blockchains. *Vehicles* 5(3):994–1012 [Online]. Available <https://www.mdpi.com/2624-8921/5/3/54>
23. Farooqi AM, Hassan S, Alam A (2019) Sustainability and fog computing: applications, advantages and challenges, pp 18–23
24. Siedlecka-Lamch O (2021) Probabilistic and timed analysis of security protocols, pp 142–151
25. Zhang J, Fang H, Zhong H, Cui J, He D (2023) Blockchain-assisted privacy-preserving traffic route management scheme for fog-based vehicular ad-hoc networks. *IEEE Trans Netw Serv Manag* 20:2854–2868. <https://doi.org/10.1109/TNSM.2023.3238307>

Hyperparameter Tuning for Machine Learning Algorithms Used for Crops Recommendation System



Humam R. Aljburi and Jamal Mustafa Al-Tuwaijari

Abstract Machine learning techniques significantly affect the economy via intelligent data utilization through data transformation and decision-making. Considering the worldwide food supply crisis, agriculture has become one of the most critical factors. Therefore, using modern technologies in artificial intelligence and machine learning in this field is a must. This work suggests an agricultural recommender system that uses rainfall data for an efficient crop recommender system. There are countless potential advantages to combining machine learning techniques in modern agriculture. The main focus of these algorithms is to minimize waste and assist in maximizing crop production via rational decisions concerning crop harvesting, planting, and irrigation. This paper conducts a comprehensive comparative analysis of three hyperparameter tuning methods. These techniques are used to enhance the accuracy of 8 classification methods, which are the AdaBoost classifier, decision tree (DT), Gaussian NB, MLP classifier, random forest (RF), SGD, support vector machine (SVM), and XGB. The suggested machine learning models are used to construct a recommender system capable of analyzing the rainfall data to predict suitable crops that match the rainfall quantities. Performing appropriate hyperparameter tuning for a machine learning classifier can lead to a substantial increase in accuracy. The performance of the classifiers is investigated via the utilization of the input rainfall dataset after the hyperparameter tuning process. A detailed analysis of the advantages and disadvantages of each single tuning method is described. The results exhibited that the random forest model gave the best accuracy. It showed the utmost precision, attaining a score of 0.999158, obtained through hyperparameter tuning using Bayesian optimization.

Keywords Agriculture · Hyperparameter · Machine learning · Recommender system

H. R. Aljburi (✉) · J. M. Al-Tuwaijari
College of Science, Computer Science Department, University of Diyala, Baqubah, Diyala, Iraq
e-mail: scicompms222302@uodiyala.edu.iq

J. M. Al-Tuwaijari
e-mail: dr.altuwaijari@uodiyala.edu.iq

1 Introduction

Agriculture is an essential component that is crucial in providing sustenance to the world's expanding population. In order to meet the growing demand for food, farmers must optimize their utilization of resources to maximize output and minimize losses. Forecasting and analyzing crop development is a crucial aspect of contemporary agriculture, and machine learning has emerged as a potent method to accomplish this objective [1, 2]. Using modern technologies is necessary in the agriculture area. Specifically, machine learning technologies are vital in classifying crops and predicting the upcoming crops that need to be planted according to the available data about soil, weather, environment, or any other information that might be helpful to increase the productivity of the lands. The agricultural sector includes the rearing of livestock, and a diverse variety of operations are related to the cultivation of crops besides the production of other farm commodities.

Various elements, such as climate, soil conditions, water availability, market dynamics, and government regulations, impact the agricultural sector. The world confronts many obstacles, including scarcity of natural resources, the impact of climate change, rapid population expansion, and the imperative for sustainable behaviors. Technological developments have become indispensable in tackling these issues and fulfilling the increasing food needs [1]. Recommender systems can successfully let farmers make decisions regarding their crops and lands based on specific data. This is necessary in agriculture, as it helps farmers to get better productivity and make better decisions. Recently, recommender systems have benefited from applying various AI techniques, improving the user experience and increasing user satisfaction. Artificial intelligence (AI) surpasses conventional recommendation systems in delivering superior-quality recommendations. This has ushered in a new era for recommender systems, enabling them to gain a sophisticated understanding of the connections between users and objects. It allows for more intricate data representations and the discovery of comprehensive knowledge in many data types, including demographic, textual, virtual, and contextual information [2].

In this context, machine learning algorithms can be used for crop recommendation systems. Typically, the recommender system in this domain can be represented, as can be seen in Fig. 1. The process starts first by collecting the data from different sources, such as the data collected from the Internet-of-things (IoT) network, weather data sources, yield data sources, or the chemical data sources.

In order to construct an optimal machine learning model, it is necessary to investigate various potential options. Hyperparameter tuning refers to building the optimal model architecture with an ideal configuration of hyperparameters. Optimizing hyperparameters is crucial to creating a proficient machine learning model, particularly for tree-based machine learning models and deep neural networks, which possess numerous hyperparameters [4]. The hyperparameter tuning procedure varies across machine learning algorithms due to various types of hyperparameters, such as categorical, discrete, and continuous [5]. Although still commonly used in graduate

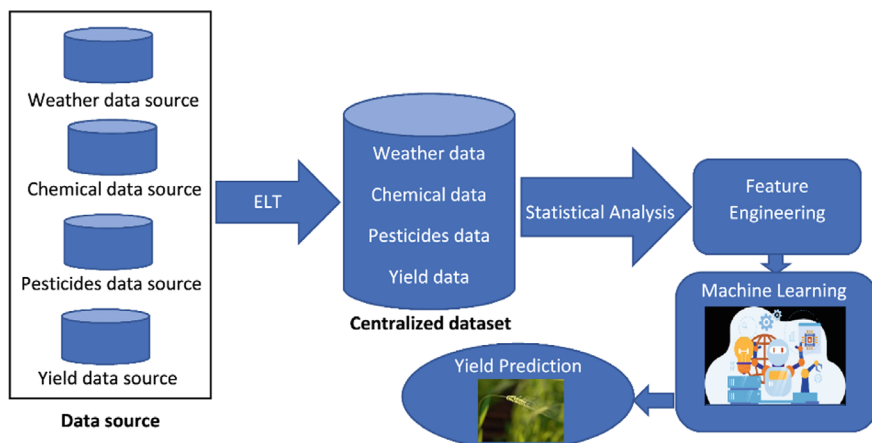


Fig. 1 Crop recommender system using machine learning technique [3]

student research, manual testing is a conventional method for fine-tuning hyperparameters. However, it necessitates a comprehensive comprehension of the machine learning algorithms employed and their corresponding hyperparameter values [6].

Nevertheless, hand tuning is inefficient for numerous reasons, such as high hyperparameters, intricate models, time-consuming model evaluations, and nonlinear interactions between hyperparameters. The presence of these elements has sparked a rise in research on methods for automatically optimizing hyperparameters, commonly referred to as hyperparameter optimization (HPO) [7]. The primary objective of HPO is to streamline the process of hyperparameter tuning and enable users to efficiently apply machine learning models to practical issues [3]. The ideal model architecture of a machine learning model is anticipated to be achieved through a hyperparameter optimization (HPO) approach. Several significant justifications for implementing HPO approaches in ML models are outlined as follows [8].

This article examines various classification techniques, including decision tree (DT), Naïve Bayes classifier (NBC), support vector machine (SVM), and random forest (RF), that can be used to construct such models. The authors anticipate that implementing machine learning-centric systems and solutions would significantly improve efficiency and production, leading to substantial transformations in the sector.

This paper aims to methodically enhance the effectiveness of different machine learning classifiers, i.e., AdaBoost, decision tree (DT), Gaussian NB, MLP classifier, random forest (RF), SGD, support vector machine (SVM), and XGBoost, within the framework of a crops recommendation system. The utilization of classification methods is necessary for the prediction in order to determine the decision that will be made. This research aims to utilize hyperparameter tuning techniques such as grid search, random search, and Bayesian optimization to improve the predictive accuracy and generalization capabilities of these classifiers in suggesting appropriate crops

using relevant agricultural data. This work examines three hyperparameter tuning strategies on the performance of each classifier in the recommendation system. This research will contribute to the advancement of agricultural decision support systems.

This study is structured as follows: The second section explains the related work, and the third section covers the hyperparameter tuning approaches, such as grid search, random search, and Bayesian optimization. In Sect. 4, the suggested hyperparameter tuning approach for the machine learning methods used in the crop recommender system is explained in depth. The test results are presented in section six, followed by an explanation of the conclusions.

2 Related Work

Machine learning is an efficient tool used in the classification of agricultural data. For instance, the weather data collected from different sources in the IoT network can be used to predict the crop that will be planted season-wise. Here, the most relevant works that used machine learning in agriculture are reviewed. In particular, the utilization of the classification techniques in recommender systems used in agriculture. In [9], the author used five distinct machine learning regression algorithms—linear regression, Bayesian linear regression, boosted decision tree regression, decision forest regression, and Poisson regression—to thoroughly examine the prediction of cotton prices—metrics for assessing the machine learning model's performance. In [10], the authors comprehensively evaluated the most recent machine learning applications in agriculture, focusing on addressing pre-harvesting, harvesting, and post-harvesting challenges. Machine learning in agriculture enables the implementation of more efficient and accurate farming practices, reducing the need for human labor while maintaining high-quality production.

In [11] the authors used different groups of classifiers used for the supervised machine learning training that benefit from the training of the knowledge base on a specific number of instances that represents the crops types while the suggested system is expected to retrieve the recommendations of the likelihood of planting this area or not. The ensemble methods that combine more than one of the classifiers is used for this task. In addition, the evaluation metrics including the confusion matrix and the precision, recall, and the accuracy were used. In [12] the same idea used in the previous research in which the authors here aimed to proposed a recommender system that is based on the idea of the training of the database with the different kinds of the soil features and the environmental elements such as the weather, humidity, and the PH value. The prediction test results were promising when compared to the previous method. The utilization of the optimizers in this regard helped to enhance the experimental results notably.

In [13], the authors presented a range of machine learning methods developed and utilized for environmental and agricultural applications, together with many sensors. The objective is to collect data on soil moisture, temperature, crop stages, irrigation, and precipitation in a region of two farms. In [14], the authors aimed this research to

enable a machine learning model to generate accurate predictions. Machine learning algorithms can assist farmers in determining the most suitable crop to cultivate and optimizing agricultural production by considering multiple aspects. Yield estimation is advantageous for farmers as it enables them to reduce crop loss and get optimal product prices. The categorization of a machine learning model as descriptive or predictive is contingent upon the specific research topic and study aims. In [15], the authors tested three distinct machine learning models: nearest neighbors, logistic regression, and decision trees.

Given that crop monitoring is the primary focus in the innovative farming process and that crop diseases are the leading cause of reduced agricultural yields, particularly in poor nations, the authors in [16] thoroughly examined crop monitoring methodologies.

Regarding the assessment of crop productivity and the identification of diseases using deep learning algorithms, according to the authors, the results and comparison indicate that crop monitoring approaches utilizing deep learning methods are superior in accuracy and effectiveness to traditional methods employed by certain developing countries. The authors introduced a new method utilizing support vector machines to include supplementary data in real-world applications inside the agriculture sector [17].

The authors asserted that they achieved 91% accuracy compared to the current applications. This recommended approach can be used by farmers to increase crop yield, and it can also be used by different government agencies to increase crop productivity. However, the writers made no recommendations for fertilizer methods to enhance crop management.

The authors suggest utilizing rainfall data and a machine learning method known as multilayer perception to help farmers make educated judgements regarding their yields prior to planting. Additionally, the suggested strategy gives priority to the produce's most effective marketing and storage plan. Based on the available information, the authors claimed that farmers would benefit from the recommended strategy in terms of improving agricultural output results. As stated in [18], the main goal of this study was to use linear regression techniques to predict the loss in crop productivity.

Throughout the literature review shown, it can be realized that significant research efforts have been made to enhance the agricultural sector via the utilization of machine learning techniques. Despite all the efforts researchers have made in the last few years in this domain, much still needs to be done in this area.

3 The Proposed Method

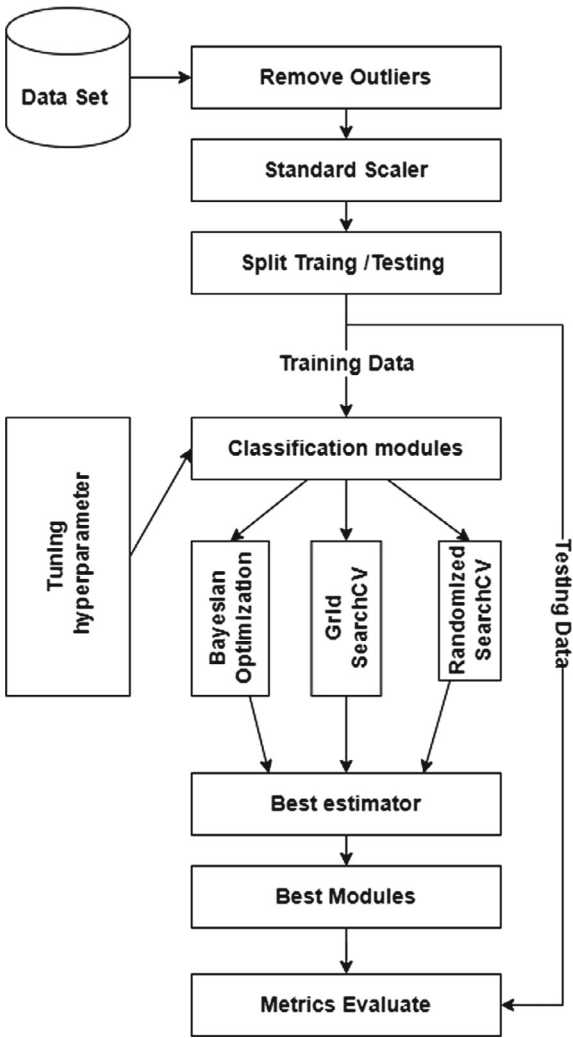
The suggested method involved the application of diverse machine learning algorithms, utilizing distinct features, for crop analysis. The selection of these algorithms was made meticulously, considering their particular skills and characteristics. The chosen algorithms are the Naïve Bayes classifier, random forest, and a multilayer

neural network. We conducted a comprehensive analysis of the data, focusing on the key characteristics that contribute to achieving high levels of accuracy. To guarantee the production of superior-quality data, we emphasize the significance of data collection and preparation. This includes essential duties such as cleaning and transforming the data. Figure 1 presents a comprehensive summary of our crop analysis and prediction approach, which integrates the Internet of things (IoT) and machine learning techniques are AdaBoost classifier, decision tree (DT), Gaussian NB, MLP classifier, random forest (RF), SGD, support vector machine (SVM), and XGB. The data collection and consolidation are the first step, and later all the data are put together in a centralized unit or server that become ready from the machine learning and evaluation. In the recommender system, the supervised machine learning algorithm methods are used such as using the classification methods for this purpose. Presumably all input data are labeled in this case the machine learning algorithm will work on these labels to find the best assigned of them.

1. **Preprocessing:** is the first step where the input data is cleansed and all data inconsistencies are resolved.
2. **Data Modeling:** in this step data is ready for further processing where the input data is all unified in one structured format.
3. **Model Evaluation and Interpretation:** this step involves searching form the hyper-parameters set by selecting one of the search methods which are the randomized, Bayesian or the grid search.
4. **Classifier Training and Testing:** the outcome of this stage in the constructed recommender system procedure relies on the methodology employed for classification. The general methodology flowchart is explained in details in Fig. 2.

Thus, Fig. 2 shows the general flowchart from the dataset input step that tries to split the input instances into training and testing instances; at that time the classification methods will be used after applying the Bayesian, grid, and the random hypertuning parameters methods. In this regard those methods try to find the best estimator, i.e., the best classification method. Also, these methods are used to find the best modules as well as they determine the best metrics used for the evaluations.

Fig. 2 General flowchart of the proposed method



Algorithm 1 Crop Analysis and Prediction

1. Preprocessing

- (a) Perform data cleansing to remove inconsistencies and errors
- (b) Handle missing values and outliers appropriately
- (c) Standardize or normalize the data as required

2. Data Modeling

- (a) Organize the preprocessed data into a unified, structured format
- (b) Ensure consistency in data representation and feature selection

3. Model Evaluation and Interpretation

- (a) Conduct a comprehensive evaluation of model performance
- (b) Utilize hyperparameter tuning techniques such as:
 - (i) Randomized search
 - (ii) Bayesian optimization
 - (iii) Grid search

4. Classifier Training and Testing

- (a) Apply supervised machine learning algorithms:
 - (i) AdaBoost classifier
 - (ii) Decision tree (DT)
 - (iii) Gaussian NB
 - (iv) MLP classifier
 - (v) Random forest (RF)
 - (vi) SGD
 - (vii) Support vector machine (SVM)
 - (viii) XGB classifiers
 - (b) Train classifiers using labeled data to predict crop suitability
 - (c) Validate classifier predictions through testing and evaluation
 - (d) Measure accuracy of each classifier to assess performance
-

There are some limitations that might be highlighted for the suggested method which are explained below:

1. **Complexity and Computational Resources:** Employing multiple machine learning algorithms and techniques can increase computational complexity and resource requirements, potentially making it challenging to implement on systems with limited computational power or memory.
2. **Data Quality and Quantity Dependency:** The effectiveness of machine learning models heavily relies on the quality and quantity of the training data. Inaccurate or insufficient data may lead to biased or unreliable predictions, affecting the overall performance of the system.
3. **Algorithm Selection Bias:** While meticulous consideration was given to selecting appropriate algorithms, there may still be inherent biases in the choice of algorithms. Certain algorithms may perform better than others in specific scenarios, leading to potential limitations in the generalizability of the results.
4. **Hyperparameter Tuning Sensitivity:** The process of hyperparameter tuning, while crucial for optimizing model performance, can be sensitive and time-consuming. Depending on the choice of search method and hyperparameter ranges, there is a risk of overfitting or underfitting the model to the training data, impacting its ability to generalize to unseen data.

4 Experimental Results

This part shows the results we got when we used grid search, random search, and Bayesian optimization, to improve each machine learning model’s hyperparameters. The accuracy of the models was then calculated. Table 2 displays the three hyperparameter tuning approaches for the AdaBoost classifier. It also shows the accuracy score and the set of hyperparameters that yielded the optimal model configuration for each strategy, resulting in the highest accuracy.

Figure 3 provides a comprehensive overview of the performance metrics corresponding to various hyperparameter tuning techniques, juxtaposed with the accuracy attained when employing default values for these hyperparameters. This comparative analysis serves as a pivotal guide for evaluating the efficacy of different optimization strategies in enhancing the model’s predictive capabilities. By delineating the scores achieved through hyperparameter tuning against those obtained with default settings, researchers can discern the incremental improvements garnered through meticulous parameter optimization. Such insights not only elucidate the impact of tuning techniques on model performance but also underscore the importance of fine-tuning hyperparameters to attain optimal results. Through this nuanced examination, practitioners can discern the most effective approach for refining model parameters, thereby advancing the accuracy and robustness of their predictive models.

Similarly, the hypertuning techniques for the other classifiers are calculated using the same way as in Table 1 and Fig. 3.

At last, the average predication performance of each classifier on the suggested recommender system is listed in Table 2. This table presents the average performance scores of different classifiers in the context of crop analysis. Each row corresponds to a specific classifier, while the ‘average’ column indicates the mean accuracy score achieved by that classifier across multiple evaluation metrics or datasets. The values represent the degree of accuracy achieved by each classifier in predicting crop-related outcomes, with higher scores indicating greater accuracy. These performance metrics

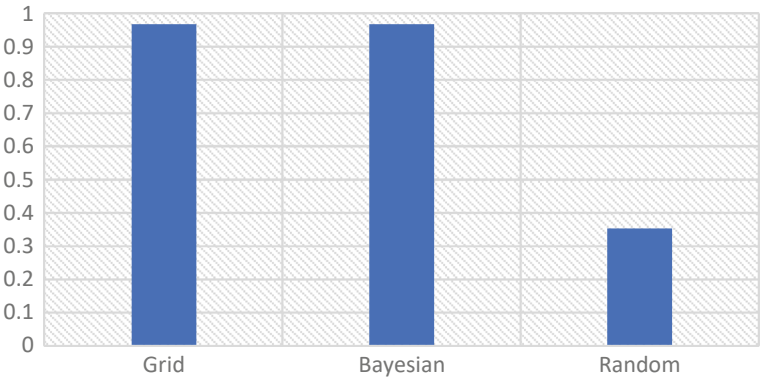


Fig. 3 Hyperparameter tuning techniques versus score for AdaBoost classifier

Table 1 Hyperparameter tuning for AdaBoost classifier

HP approach	Accuracy	Best hyperparameters
Grid	0.96965571	'learning_rate': 0.1, 'n_estimators': 120
Bayesian	0.96965571	[('learning_rate', 0.1), ('n_estimators', 120)]
Random	0.35224387	'n_estimators': 100

Table 2 Average predication performance of each classifier on the grid, Bayesian, and the randomized hyperparameter tuning methods

Classifier	Avarage
AdaBoost classifier	0.763852
Decision tree	0.997155
GaussianNB	0.998386
MLP classifier	0.988598
Random forest	0.999157
SGD classifier	0.9588
SVM classifier	0.997262
XGB classifier	0.998176

are crucial for evaluating the efficacy and suitability of each classifier for crop analysis tasks, providing valuable insights for researchers and practitioners in the field of agricultural informatics.

These results provide valuable insights into the effectiveness of each classifier in accurately predicting crop-related outcomes, aiding researchers and practitioners in selecting the most suitable classifier for agricultural informatics applications.

5 Conclusions

Machine learning is applied across various domains in society. The field of agriculture stands to gain significant advantages from implementing machine learning technologies. The study underscored the need of using machine learning algorithms into modern agriculture to optimize crop productivity and reduce waste through well-informed decision-making. More specifically, a recommender system focused on agriculture that suggests the best crop based on rainfall statistics. Grid search, random search, and Bayesian optimization are the three hyperparameter tuning techniques that are presented in this paper. In order to estimate agricultural yields based on the available data, these techniques are used to optimize the hyperparameters of eight machine learning algorithms: the AdBoost classifier, decision tree (DT), Gaussian NB, MLP classifier, random forest (RF), SGD, support vector machine (SVM), and XGB. In the subject of agricultural recommender systems for machine learning hyperparameter tweaking, this work is acknowledged as a groundbreaking investigation. The random forest model displayed the greatest degree. Therefore, the

existing data assessment pool will be created. Furthermore, distinct species within the same plant variety will be examined individually, enabling the identification of the superior product within each species by employing various machine learning methods. In terms of future work, investigating the applicability of emerging optimization algorithms and exploring the interplay between hyperparameters across different machine learning models could further enhance our understanding and improve model performance in various domains.

References

1. Li L, Wang B, Feng P, Liu DL, He Q, Zhang Y, Wang Y, Li S, Lu X, Yue C et al (2022) Developing machine learning models with multi-source environmental data to predict wheat yield in China. *Comput Electron Agric* 194:106790
2. van Klompenburg T, Kassahun A, Catal C (2020) Crop yield prediction using machine learning: a systematic literature review. *Comput Electron Agric* 177:105709
3. Kuradusenge M, Hitimana E, Hanyurwimfura D, Rukundo P, Mtonga K, Mukasine A, Uwitonze C, Ngabonziza J, Uwamahoro A (2023) Crop yield prediction using machine learning models: case of Irish potato and maize. *Agriculture* 13:225
4. Ranjan P, Garg R, Rai JK (2022) Artificial intelligence applications in soil & crop management. In: *Proceedings of the IEEE conference on interdisciplinary approaches in technology and management for social innovation (IATMSI)*, Gwalior, India, 21–23 Dec 2022, pp 1–5
5. Oré G, Alcântara MS, Góes JA, Oliveira LP, Yepes J, Teruel B, Castro V (2020) Crop growth monitoring with drone-borne DInSAR. *Remote Sens* 12:615
6. Gehlot A, Sidana N, Jawale D, Jain N, Singh BP, Singh B (2022) Technical analysis of crop production prediction using machine learning and deep learning algorithms. In: *Proceedings of the international conference on innovative computing, intelligent communication and smart electrical systems (ICES)*, Chennai, India, 24–25 Sept 2022, pp 1–5
7. Vashisht S, Kumar P, Trivedi MC (2022) Improved extreme learning machine for crop yield prediction. In: *Proceedings of the 3rd international conference on intelligent engineering and management (ICIEM)*, London, UK, 27–29 April 2022, pp 754–757
8. OpenAI (2023) New and improved content moderation tooling. OpenAI. Available online: <https://openai.com/blog/new-and-improved-content-moderation-tooling/>. Accessed on 1 April 2023
9. Google. Bard Chatbox. Google. Available online: <https://bard.google.com>. Accessed on 2 April 2023
10. Dean J (2020) The deep learning revolution and its implications for computer architecture and chip design. In: *Proceedings of the IEEE international solid-state circuits conference-(ISSCC)*, San Francisco, CA, USA, 16–20 Feb 2020
11. Cui YW, Henrickson K, Ke R, Pu Z, Wang Y (2019) Traffic graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting. *IEEE Trans Intell Transp Syst* 21:4883–4894
12. Shahrin F, Zahin L, Rahman R, Hossain AJ, Kaf AH, Abdul Malek Azad AKM (2020) Agricultural analysis and crop yield prediction of Habiganj using multispectral bands of satellite imagery with machine learning. In: *Proceedings of the 11th international conference on electrical and computer engineering (ICECE)*, Dhaka, Bangladesh, 17–19 Dec 2020, pp 21–24
13. Tawseef AS, Tabasum R, Faisal RL (2022) Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Comput Electron Agric* 198:107119
14. Elbasi E et al (2023) Crop prediction model using machine learning algorithms. *Appl Sci* 13(16):9288

15. Gupta T, Maggu S, Kapoor B (2023) Crop prediction using machine learning. IRE J 6(9)
16. Nischitha K et al (2020) Crop prediction using machine learning approaches. Int J Eng Res Technol (IJERT) 9(08):23–26
17. Mahendra N, Vishwakarma D, Nischitha K, Ashwini and Manjuraju M. R 2020 Crop prediction using machine learning approaches, International Journal of Engineering Research & Technology (IJERT) vol 9 Issue 8 (August 2020).
18. Gulati P, Jha SK (2020) Efficient crop yield prediction in India using machine learning techniques. Int J Eng Res Technol (IJERT) 8(10)

Intelligent Digital Industrial Internet of Things Network Security Monitoring Based on Edge Cloud Computing



Yijian Qiu and Chuang Wang

Abstract The rapid evolution and rollout of the digital-enabled intelligent Industrial Internet of Things (IIoT) bring significant opportunities to enterprises. Nevertheless, it has also brought security challenges to network infrastructure. To address the increasing complexity of network security threats, this article uses the edge cloud computing (ECC) to have the edge, which may articulate the premise that the complexity of the network security environment has been significantly increased. This architectural model utilizes the distributed computing power and data locality of edge nodes to deploy network security monitoring and data analysis tasks closer to the edge nodes connecting to the IoT devices. This study uses the distributed computing power and data locality of edge nodes to deploy network security monitoring and data analysis tasks closer to the edge nodes connecting to the IoT devices. By utilizing machine learning and artificial intelligence algorithms, ECC can achieve automated incident detection and response, while additionally using the real-time data processing and analysis technologies to enhance the system's real-time response capability. Considering throughput, the ECC-based IIoT network security incident detection and response system shows significant advantages. The system has shown significant advantages in terms of throughput because ECC can effectively handle a large number of data streams and has improved its concurrency processing capabilities.

Keywords Industrial internet of things · Network security monitoring and analysis · Edge cloud computing · Cloud data analysis

Y. Qiu (✉)

Jiangxi Key Laboratory of Advanced Copper-Based Materials, Nanchang, Jiangxi, China
e-mail: 15970479209@163.com

Institute of Applied Physics, Jiangxi Academy of Sciences, Nanchang, Jiangxi, China

C. Wang

Questrom School of Business, Boston University, Boston, USA

1 Introduction

Enterprises nowadays are experiencing more challenging cybersecurity issues as they deploy intelligent, digital IIoT to increase efficiency in production and management levels through intelligence and digitization. Among these, cybersecurity challenges are potential security threats such as malicious attacks, data breach, and security vulnerability. These threats place the normal operation and data security of enterprises under great threats. As for the solution to these security challenges, intelligent digital IIoT network security monitoring and analysis based on ECC is becoming a hot topic.

In the conventional network security monitoring and analysis, data possessing is necessarily centralized and operated at central servers, which can lead to data transmission latency, data bandwidth bottleneck, data transmission throughput, and real-time performance limitations of the system. The smart digital IIoT network security monitoring and analysis system based on ECC is able to increase the data transmission throughput and real-time response capacity of the system by use of computing tasks are executed on edge nodes near the IoT devices, and completely utilized the distributed computing power and data locality of edge nodes. The meaning of this approach is that the system can identify and respond to malicious behaviors faster, minimize the negative and duration impact of the attacks on the system, and improve the network protection ability.

The article aims to conduct an in-depth study on intelligent digital IIoT network security monitoring and analysis system based on ECC, and explore the advantages of improving system throughput and achieving real-time response capabilities. Firstly, it may articulate the background and challenges of intelligent digital industrial IoT network security to indicate the significance and urgency of research. Secondly, the principles and key technologies of network security monitoring and analysis methods are based on ECC which may be elaborated, including distributed computing capability, utilization of data locality, and automated security analysis algorithms. Finally, the results of this article may be summarized, and the development prospects of intelligent digital IIoT network security monitoring and analysis based on ECC may be discussed.

2 Related Work

Many people have conducted research on network security monitoring and analysis, and Li Liuli aims to elaborate on the application of artificial intelligence technology in network security monitoring and defense. By analyzing the detection principles of machine learning and deep learning algorithms, various intrusion detection technologies based on user behavior analysis, traffic modeling, and security log mining are highlighted to achieve real time and automated detection of unknown attacks [1]. Based on the significance of developing power Internet of Things terminal

network security monitoring technology, Zheng Zhiming analyzed the overall architecture of power Internet of Things terminal network security monitoring technology and the construction ideas of the four-dimensional architecture of power Internet of Things. He listed several terminal network security monitoring technologies for the power Internet of Things, and finally elaborated on the implementation strategies for terminal network security monitoring [2]. Liu Ying proposed that the most important link in the network security system of cloud computing platforms is network security monitoring. Based on this, in-depth exploration of network security monitoring technology for cloud computing platforms has been conducted, aiming to provide a certain reference value for creating a secure network environment [3]. Dai Yun improved the traditional K-means clustering algorithm and conducted experiments using the KDDCup99 dataset. This indicates that the improved K-means algorithm has better detection accuracy and measurement performance than traditional algorithms [4]. Zhan Xue proposed a ship network security monitoring and warning method based on Internet of Things technology, which uses network security warning units to compare and analyze security situation values and set thresholds. When the security situation value is lower than the set threshold, an alarm can be issued to achieve ship network security warning [5]. Lagaraa focused on researching graph-based representation of network traffic records and the application of graph-based analysis in intrusion detection and botnet detection [6]. Younus conducted an investigation into network security monitoring technology, including its functions, content, and tools [7]. Kebande studied the relevance of real-time monitoring as a supplementary security component for alert actions in modern network environments [8]. Mughal explored the key aspects of well-structured wireless network security [9]. Skopik dumped network tracking at blocking points in the network, collected log files from services, and provided appropriate security information and event management solutions [10]. These studies have provided a lot of help for this article, which aims to deeply study the intelligent digital IIoT network security monitoring and analysis system based on ECC.

3 Method

3.1 *Intelligent Digital IIoT Network Security Monitoring Framework*

The intelligent digital IIoT network security monitoring framework is based on ECC technology, which aims to provide full real-time security monitoring and analysis capabilities. The ECC can take the Network Security Monitoring and Analytics functions down to the edge nodes near the IoT devices, thus providing a faster monitoring and response [11, 12]. First, the sensor devices for the Network Security can be deployed at the edge nodes, such as the widely used programmable sensors (Bro, Plumb, Softflowd). These sensor devices can be used to collect and monitor the

data traffic, communication behavior, and security event data that move through the network. The key network traffic data that can be captured by the sensors include packet headers, protocol characteristics, and statistics on traffic. The sensor devices can be setup on the FDU, generating a part of the network security processing in hardware. The data collected can be analyzed and processed in real-time at the edge nodes, where the local storage and computing resources can be dedicated to this purpose. The real-time analysis and data of the traffic collected is to implement the network intrusion detection which involves the deep learning and machine learning algorithms. Consequently, the system can identify, classify, and recognize different types of network attacks in real time and on the fly. Such attacks include intrusion detection, anomaly behavior, or malicious code detection. After the network attack is recognized, the edge nodes can provide the real-time response to the anomalous network event, such as generating alarms, blocking traffic, or adapting the network policy decision algorithm. This architecture makes use of ECC technology to transfer the processed security data to the cloud for subsequent analysis, and decision-making. Through the high-performance computing, and big data processing capabilities of the cloud, more complex security event analysis, threat intelligence integration, and pattern recognition can be accomplished. Based on these analysis results, the cloud can make a more comprehensive and accurate network security strategy, and distribute it to edge nodes in order to achieve the dynamic adjustment and optimization of the whole system.

3.2 *Network Security Monitoring Algorithm for Edge Nodes*

In this article, an intelligent manufacturing scenario is depicted, which contains numerous IoT devices and industrial robots connected to a central controller through the industrial IoT. Edge nodes are deployed in the factory, responsible for monitoring and securing IIoT network security. The edge nodes are equipped with network traffic monitoring algorithms to observe, analyze, and monitor data packets that pass through the network. These algorithms can examine the header information of the data packets provided, the protocol type, the source address and target address for IP, and the data size that is transmitted. They apply the traffic summary algorithms for network traffic, which can analyze traffic and summarize statistics in real time to be able to understand the normal behavior of the network and to identify abnormal situations. In Table 1, the edge nodes network security monitoring algorithms provide the data collection and analysis of edge nodes in a smart manufacturing scenario:

Each row within the table can represent the network packet information. The timestamp column records the capture time of packets. The source IP address and destination IP address columns are the IP address of the source device and destination device in the packet, respectively [13, 14]. The protocol column is the communication protocol used by current packets. The packet size column counts the packets size (how many bytes are in a packet). The last column, whether abnormal, whether the packet has been marked as the abnormal behavior of the network security monitoring

Table 1 Analysis data

Timestamp	Source IP address	Destination IP address	Protocol	Packet size (bytes)	Is anomaly
12:00:01	192.168.1.10	192.168.1.50	TCP	120	No
12:00:02	192.168.1.20	192.168.1.30	UDP	75	No
12:00:03	192.168.1.15	192.168.1.50	ICMP	64	No
12:00:04	192.168.1.50	192.168.1.10	TCP	250	Yes
12:00:05	192.168.1.30	192.168.1.20	TCP	400	No
12:00:06	192.168.1.10	192.168.1.50	UDP	150	No
12:00:07	192.168.1.20	192.168.1.30	TCP	80	Yes

algorithm based on the edge node. Based on the information of each data packet in the table, the network security monitoring algorithm on edge nodes can analyze the information of each data packet, judge whether there is the abnormal behavior of the data packet based on predefined rule or model. When data packet is judged as abnormal behavior, the edge node can propose corresponding response measures, such as alarm notification, traffic blocking, or device isolation in order to protect the network security of the intelligent manufacturing system.

In the context of network security, edge nodes use intrusion detection system (IDS) algorithms to identify possible network intrusion behavior. The IDS procedure uses an existing library of intrusion features and compares the data with known attack features and normal behavior. IDS is able to detect common network attacks, such as malware propagation, port scanning, and denial of service (DoS) attacks. Following IDS, edge nodes can also use behavior analysis algorithms to detect abnormal behavior and patterns. Behavior analysis is based on modeling of normal network behavior and monitoring changes in network traffic to identify potential security threats [15, 16]. For example, if the amount of data sent by a device suddenly increases, or if the communication actions of a device differ from its historical behavior, behavior analysis algorithms can detect these anomalies and trigger a response or a procedure to address this behavior anomaly.

3.3 Cloud Data Analysis and Network Security Decision-Making

A smart and secure Industrial IoT network security surveillance system with ECC is proposed. The system includes the following key components: the edge device, sensors of the Industrial IoT network, industrial devices, and the edge node to collect and conduct preliminary analysis on the data; the edge cloud node, a cloud computing node which is placed close to the edge device to process and analyze the data collected by the edge device. The edge cloud node has strong computational power and storage resources; and the cloud server, a server in the cloud, which is responsible for further

data analysis and network security decisions. The cloud server can integrate data from multiple edge cloud nodes, scan the network for threats and abnormal actions, and take necessary actions for the network security [17, 18]. The cloud server uses Support Vector Machine (SVM) in order to detect and classify threats from edge cloud node data. Suppose there are N samples in a dataset, each represented by a d -dimensional feature vector. With SVM algorithm, it can construct a classifier $f(x)$, where x is a d -dimensional feature vector, responsible for predicting the label y of the sample:

$$f(x) = \text{sign}(w^T * x + b) \quad (1)$$

w is the normal vector of the decision boundary and b is the bias term, cloud servers can make security decisions depending on the detection results of threats provided. These decisions could include blocking malicious packets, alerting the system administrators, or taking emergency response actions. Security decisions could be based on predefined rules and policies, or alternatively, the security decisions could be accomplished using reinforcement learning algorithms with adaptive learning and optimization:

$$\text{Decision} = \text{argmax}(Q(s, a)) \quad (2)$$

$Q(s, a)$ represents the state action value function where s represents the current state and a represents the optional safety decision action. The system can select the action with the highest Q value to make the optimal safety decision. The above example is for demonstrating how to use ECC for intelligent digital IIoT network security monitoring and analysis. In this way, it can protect IIoT networks from the threats and attacks through cloud data analysis and network security decision-making [19, 20], in which the threat detection algorithms and security decision rules describe their principles and operating mechanism by using mathematical forms.

4 Results and Discussion

This experiment is designed to carry out experiments to verify the accuracy and practicality of this method in threat detection and security decision-making and by providing secure and effective network security protection for intelligent digital IIoT environments. In order to improve the credibility of the results of the experiment, the experiment is required to operate in a simulated IIoT network environment, that is the simulation platform must be built. The simulation platform includes edge devices, edge cloud nodes, and cloud servers. The interaction between these components constitutes a realistic networking and data transmission process. Edge devices transmit data to edge cloud nodes, and edge cloud nodes transmit data to cloud servers. Table 2 depicts the data transmission process between edge devices, edge cloud nodes, and cloud servers.

Table 2 Interaction process

Component	Function	Interaction
Edge device 1	Collect sensor data	Transmits data to the edge cloud node
Edge device 2	Monitor network traffic	Transmits traffic data to the edge cloud node
Edge cloud node	Process and analyze data	Receives data and traffic data from edge devices for analysis and processing
Cloud server	Store and manage data	Receives data from edge cloud node for storage and management

In addition, network security datasets can be synthesized by members of the research team to include normal and malicious network traffic, attack behavior, etc., in order to verify the performance of threat detection algorithms. The experimental group used the ECC-based intelligent digital IIoT network security monitoring and analysis method, while the control group employed traditional intrusion detection systems for monitoring and analysis. This article selects recall rate, response time, and system throughput as the evaluation indicators for the experiment. Threat detection algorithms and security decision rules can be fully implemented and validated before the experiment to ensure their correctness and reliability.

4.1 Recall Rate

In network security monitoring and analysis, the recall rate is crucial for accurately identifying and capturing malicious behavior. A higher recall rate means that the system can better detect real malicious behavior, reduce false positives, and improve the security and reliability of the system. Figure 1 shows the comparison results.

In the above tests, the recall rate of the experimental group based on ECC method can reach up to 98% and the lowest is 91%. The highest recall rate of the control

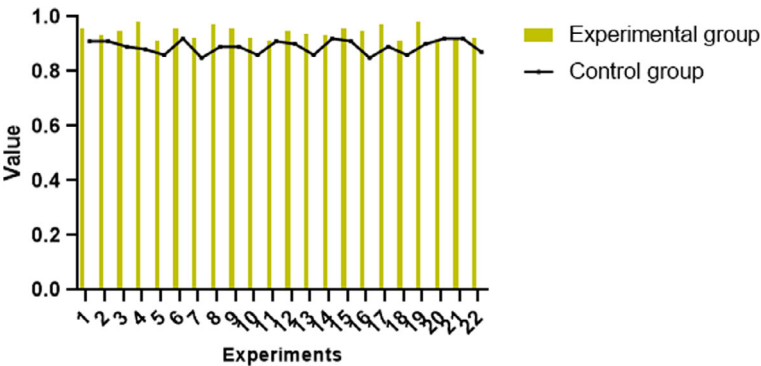


Fig. 1 Recall rate

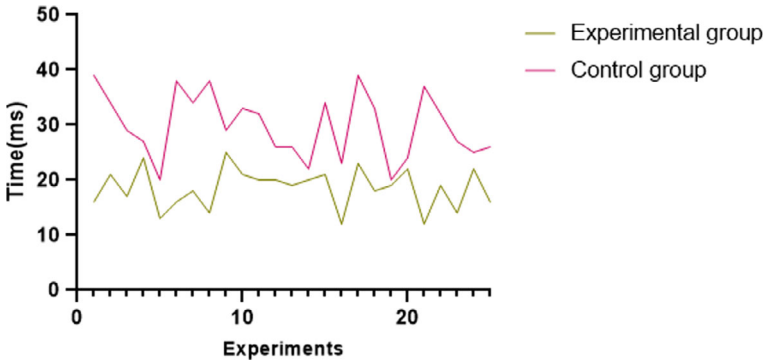


Fig. 2 Response time

group based on intrusion detection systems is only 92%, and the lowest is 85%, which is 6% lower than the experimental group. ECC shifts computing and analysis tasks from traditional centralized servers to edge nodes closer to IoT devices. This can reduce data transmission latency and achieve faster real-time response.

4.2 Response Time

Response time can reflect the real-time performance and response speed of the system, and a shorter response time means that the system can detect malicious behavior or attacks more quickly. Targeted measures can be taken to prevent or respond, which is crucial for protecting network security and reducing potential risks. Figure 2 shows the testing situation.

In the response time test, it is evident that the experimental group has a shorter response time. The response time of the experimental group is between 12 and 25 ms, while the response time of the control group is between 20 and 39. This article demonstrates that the method can quickly respond to malicious behavior and attacks, reducing the impact and losses on the system. By taking prompt measures, the spread of malicious behavior can be restricted and the normal operation of the network can be restored as soon as possible.

4.3 System Throughput

Figure 3 shows the comparison results of system throughput.

The test results show that the system throughput based on ECC method in this article is between 1100 and 1300 RPS, with high event acceptance ability. The control group's throughput based on intrusion detection system is between 950 and

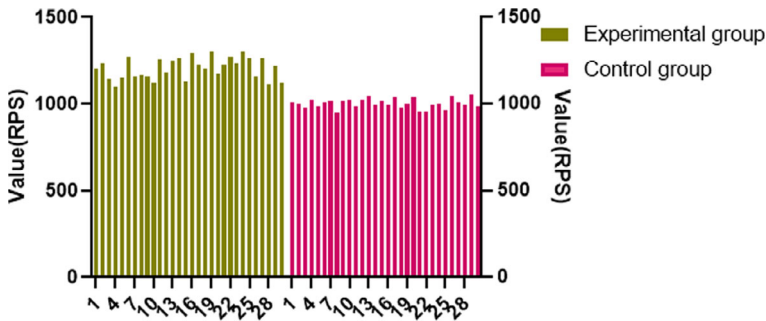


Fig. 3 System throughput

1050 RPS, which is relatively low. ECC utilizes the distributed computing capabilities of edge nodes to distribute computing tasks to edge nodes near IoT devices for processing. Compared to traditional intrusion detection systems, ECC can simultaneously process more data and tasks, improve the system’s concurrency processing capability, and thus increase the system’s throughput.

5 Conclusions

The intelligent digital IIoT network security monitoring and analysis based on ECC is an efficient, real-time, and intelligent security solution. The system deployed computing and analysis tasks using the distributed computing power and data locality obtained from edge nodes. In this way, the system can support large-scale data streams and improve the throughput of the system. ECC-based solutions have the following advantages over traditional intrusion detection systems. First, by using ECC, computing and analyzing tasks are deployed on the edge nodes near the IoT devices to achieve real-time responsiveness. This solution can detect malicious behaviors, and the response can be taken in a timely manner, which can shorten the duration of the attack and reduce the impact. Second, the system made use of network security monitoring by deploying machine learning and AI algorithms in the detection and response system. By doing this, the system can perform accurate automated security monitoring and mitigation of malicious behaviors, which makes the system deployed by ECC able to not only meet the challenge of large-scale data streams processing, but also to cherish the data, as well as advance security enhancement, and reduce mistakes.

Acknowledgements Major Scientific Research Project of Jiangxi Academy of Sciences, China, Grant/Award Numbers: 2023YSBG21016.

References

1. Li L (2024) Research on network security monitoring and defense technology based on artificial intelligence. *Mobile Inf* 46(4):158–160
2. Zheng Z (2024) Research on network security monitoring technology of power internet of things terminal. *Mobile Inf* 46(4):176–178
3. Liu Y, Zhang Q (2023) Exploration and application of network security monitoring technology based on cloud computing platform. *Inf Record Mater* 24(5):86–88
4. Yun D (2023) Research on the application of K-means clustering algorithm based on big data for network security monitoring. *Inf Record Mater* 24(4):140–142
5. Zhan X (2023) Application of internet of things technology in ship network security monitoring and early warning. *Ship Sci Technol* 45(15):123–126
6. Lagraa S, Husák M, Seba H et al (2024) A review on graph-based approaches for network security monitoring and botnet detection. *Int J Inf Secur* 23(1):119–140
7. Younus ZS, Alanezi M (2023) A survey on network security monitoring: tools and functionalities. *Mustansiriyah J Pure Appl Sci* 1(2):55–86
8. Kebande VR, Karie NM, Ikuesan RA (2021) Real-time monitoring as a supplementary security component of vigilantism in modern network environments. *Int J Inf Technol* 13(1):5–17
9. Mughal AA (2022) Well-architected wireless network security. *J Hum Appl Sci Res* 5(1):32–42
10. Skopik F, Landauer M, Wurzenberger M (2022) Blind spots of security monitoring in enterprise infrastructures: a survey. *IEEE Secur Priv* 20(6):18–26
11. Duan S, Wang D, Ren J et al (2022) Distributed artificial intelligence empowered by end-edge-cloud computing: a survey. *IEEE Commun Surv Tutor* 25(1):591–624
12. Cao K, Hu S, Shi Y et al (2021) A survey on edge and edge-cloud computing assisted cyber-physical systems. *IEEE Trans Industr Inf* 17(11):7806–7819
13. Arthurs P, Gillam L, Krause P et al (2021) A taxonomy and survey of edge cloud computing for intelligent transportation systems and connected vehicles. *IEEE Trans Intell Transp Syst* 23(7):6206–6221
14. Zhang Y, Lan X, Ren J et al (2020) Efficient computing resource sharing for mobile edge-cloud computing networks. *IEEE/ACM Trans Netw* 28(3):1227–1240
15. Ghosh AM, Grolinger K (2020) Edge-cloud computing for internet of things data analytics: embedding intelligence in the edge with deep learning. *IEEE Trans Industr Inf* 17(3):2191–2200
16. Zhang WZ, Elgendy IA, Hammad M et al (2020) Secure and optimized load balancing for multitier IoT and edge-cloud computing systems. *IEEE Internet Things J* 8(10):8119–8132
17. Tuli S, Ilager S, Ramamohanarao K et al (2020) Dynamic scheduling for stochastic edge-cloud computing environments using a3c learning and residual recurrent neural networks. *IEEE Trans Mob Comput* 21(3):940–954
18. Ding Y, Li K, Liu C et al (2021) A potential game theoretic approach to computation offloading strategy optimization in end-edge-cloud computing. *IEEE Trans Parallel Distrib Syst* 33(6):1503–1519
19. Qu G, Wu H, Li R et al (2021) DMRO: a deep meta reinforcement learning-based task offloading framework for edge-cloud computing. *IEEE Trans Netw Serv Manage* 18(3):3448–3459
20. Fang J, Ma A (2020) IoT application modules placement and dynamic task processing in edge-cloud computing. *IEEE Internet Things J* 8(16):12771–12781

Visualizing the Alarming Surge: Mapping and Analyzing Suicide Patterns Across India



P. S. Metkewar, Syed Irfan Yaqoob, Shubhalaxmi Joshi, Anshul Mishra,
and Ishita Kalbandhe

Abstract Differences in self-destruction rates among particular nations and socio-segment fragments highlight the perplexing exchange between cultural settings and mental pain. Remarkably, perceivable varieties arise in the commonness of suicides among two disparate gatherings: ranchers and understudies. Illustratively, inside the Indian setting, the year 2012 saw an upsetting frequency wherein more than 2200 understudies depended on self-destruction as a reaction to scholarly under accomplishment, subsequently featuring the intense tensions related with scholastic assumptions. On the other hand, the heartbreaking peculiarity of rancher suicides is credited to a conjunction of elements that render self-destruction as a critical response even with outlandish monetary misery, frequently denoting a last-ditch supplication for mitigation from their quandary. The natural meaning of agribusiness in the Indian economy is undeniable, yet, the predicament of Indian ranchers remains generally unacknowledged. A stunning extent of rancher suicides—50% to be exact—exudes from Maharashtra, in this way typifying the gravity of the issue inside a particular geographic setting. Considering India's far reaching demography and financial scene vigorously dependent on horticulture, a convincing basic arises for the public authority to proactively address the unavoidable disgrace covering emotional well-being issues, which thus energizes the troubling twisting of suicides. In a faithful pursuit to grasp the nuanced features of this upsetting pattern, our review tries to unwind the fundamental self-destruction risk factors natural for Indian ranchers

P. S. Metkewar (✉) · S. Joshi

School of Computer Science and Engineering, Dr Vishwanath Karad MIT World Peace
University, Kothrud, Pune, Maharashtra, India
e-mail: pravin.metkewar@mitwpu.edu.in

S. Joshi

e-mail: shubhalaxmi.joshi@mitwpu.edu.in

S. I. Yaqoob

Department of Computer Science and Engineering (AIT), Chandigarh University, Gharuan,
Mohali, Punjab, India

A. Mishra · I. Kalbandhe

Symbiosis Institute of Computer Studies and Research (SICSR), Symbiosis International
(Deemed University) (SIU), Pune, Maharashtra, India

and understudies. Strategically grounded, our investigation dives into an exhaustive dataset incorporating 236,583 perceptions traversing the worldly range of 2012–2020, fastidiously examining seven unmistakable factors that unpredictably clarify the complicated embroidery of self-destruction occurrences revealed across India. With immovable concentration, our concentrate aggressively plans to unravel the multifaceted nexus between financial determinants and mental weakness, particularly dividing the forms of vulnerability that encompass these two fragments of society. Spearheading a way toward proactive counteraction and enhancement, our review's discoveries emphatically highlight the articulated impact old enough gatherings—explicitly the partners traversing 0–14 and 15–29—wherein the repercussions of seen accomplishment or the scarcity in that department manifest in their most significant greatness.

Keywords Psychological distress · Indian economy · Age groups · Vulnerability · Financial distress

1 Introduction

As per the World Wellbeing Association (WHO), India has one of the great self-destruction rates on the planet, with an expected 230,000 suicides each year. Self-destruction is an intricate issue with different elements adding to it. Studies have shown that self-destruction rates in India fluctuate by age, orientation, and district.

Age: Self-destruction rates are most elevated among youthful grown-ups in India, with people between the ages of 15 and 29 representing the greater part, everything being equal. The purposes behind this are perplexing yet may incorporate scholarly strain, relationship issues, and monetary hardships.

Orientation: Self-destruction rates are higher among men in India, with male suicides representing roughly 66% of all suicides. One justification for this orientation difference might be the men are bound to take part in rash and savage techniques for self-destruction, for example, hanging and guns, which are more deadly than the strategies utilized by ladies.

District: The exchange of diverse variables adds to the noticed provincial uniqueness in self-destruction rates. Financial status, social standards, and admittance to emotional well-being administrations arise as vital powerhouses. Areas like Kerala, Tamil Nadu, and Karnataka, with raised self-destruction rates, frequently wrestle with perplexing financial difficulties that might compound mental pain, consequently possibly enlarging self-destructive propensities.

Self-destruction is a huge general medical problem in India, and endeavors are expected to address the mind boggling factors that add to it. Systems might incorporate expanding admittance to psychological well-being administrations, further developing training and mindfulness about self-destruction counteraction, and dealing with the fundamental social and monetary variables adding to self-destruction risk.

This article covers a standard design including: (1) Introduction, (2) Literature Review, (3) Methodology, (4) Results and Discussion, and (5) Conclusion, Limitations, and Future scope.

2 Literature Review

Within the journey to get the complex web of components contributing to suicide rates, analysts have tirelessly worked with assorted datasets over time. One dataset that has earned noteworthy consideration is the SUICIDERATES2012-20 dataset, particularly curated for analyzing and anticipating suicide rates. Created by analysts from the World Wellbeing Organization (WHO), this dataset houses a collection of 370 occurrences, each carrying profitable experiences typified in 7 distinct attributes. Suicide could be a genuine open well-being issue; be that as it may, suicide is avoidable with opportune, verified, also frequently cheap intercessions. As the Nation's reactions are compelling, extensive multisector self-destruction avoidance methodologies are required [1].

The SUICIDERATES2012-20 dataset has risen as an essential apparatus for making prescient calculations that seem to combat the troubling predominance of suicides in India. One outstanding ponder dove into tackling the control of this dataset to develop a forecast demonstration, utilizing an interesting mix of human and manufactured insights. The arms stockpile of methods at play enveloped relationship examination, histograms, disarray lattices, and choice trees, all complicatedly woven together to donate life to a vigorous prescient model.

However, NCRB information gives detached observation, accompanied by the sources of FIR recorded with various society part along police with specific passing, also subsequently is subordinate to those detailing acquired from the communities, whichever being specific considering different causes counting shame connected to self-harm passings and outlawing endeavored suicides up till as of late [2]. From an inactive observation viewpoint, the wide run of detailed self-destruction in NCRB in the middle of the additionally created and fewer created lands likely throwback in real-life instead of distinctive detailing hones, as high number of suicides getting detailed in distinction to Southern wealthier Indian Lands mentioned into SRS information [3]. This NCRB information shows that housewives account for 18.6% of add-up to people committing suicide as well as 52.8% of female casualties [4]. The ones included in agriculture, as well as horticulture, frame the other larger population, having 11.9% as overall casualties taken after no later than the ones employed within the Business sector (7.8%) and jobless (7.5%) also open division (7.8% and 2.2%, separately). Understudy accounts toward 5.5% of suicides whereas joblessness accounts for 5.5% and 7.5%, respectively. The ones utilized within an open division (2.2% of add-up to suicide rate) and governmental hirelings (1.3% of add-up to suicide rate) have been the slightest spoken-to bunch [5]. Agreeing with the formal information, the causes of suicides are obscure with around 43% of suicide

rates whereas sickness as well as household issues contribute about 44% of suicide causes [6].

Individual/communal causes for suicide remain preeminent basis across the years. Psycho-social components counting financial situations, relational, communal as well and social clashes, liquor abuse, budgetary issues, and unemployment are detailed as critical hazard variables for suicide in India [7]. There's a wide variety of suicides inside the nation. Southern lands of Kerala, Karnataka, Andhra Pradesh, and Tamil Nadu have counted suicide rates of > 15 whereas within the Northlands of Punjab, Uttar Pradesh, Bihar, and Jammu and Kashmir, the suicide rates are < 3 [6]. Three-quarters with overall suicide being the communal and financially profitable life duration gather around 15–44 a long time. Number of years from 15 to 24 a long-time women suicide rate of 109/100,000 surpassed the men suicides of 78/100,000; suicides were capable for 49% with overall passings in ladies also 38% in males of such stage of life [8].

Utilizing toxic substances or overdosing on pharmaceuticals and hangings have been favored methods in attempting suicides for decennary, and crevices in the middle of both diminishing supplementarily, comparable as such detailed over verbal dissection []. In expansion, with hanging being progressively favored as an implication of suicides, limiting approaches are improbable with help to a lessening in suicide rates [9]. Restricting approaches to frequently utilized strategies for suicides may cause diminishment in general and also manner-specified suicides [10, 11]. In any case, limiting the deal and utilization of pesticides or certain solutions in these locales requires a viable administrative system. Suicide by hanging and consumption of poison substances, particularly with organophosphate pesticides, is the foremost usual strategy in suicide. The younger, especially youthful ladies are at higher risk of self-harm. MOO's financial position, relational aggravation, along the nearness of psychological illness are critical hazard components for suicides. There is deficient proof that concludes the significance of instruction, conjugal level, and home, or business level as possible chance components of suicides [12].

Depression/despondency is attributed to the biggest extent of malfunctioning life a long time for self-harm all-inclusive; In any case, the relative commitment of serious mental well-being issues to suicides in India is still far from being true [12, 13]. Suicide hazard increases altogether with comorbidity, and consequently, opportune and viable well-being care is basic to decrease the suicide risk.

Housewives recorded the most elevated extent toward self-destruction passings across the years, taken after by self-employed others. Marriages are not so much defensive averse to suicides in females compared to guys [13, 14], and the more prominent female powerlessness might contribute toward abnormal psychology counting organized, marriages at a young age, youthful parenthood, moo communal status, household viciousness, and financial reliance [4, 14, 19]. Women's suicide rate is more compared to that of men in people beneath 30 a long time of age but the inverse is genuine in those 30 a long time of age or more seasoned. Among ladies, self-immolation is additionally a generally usual strategy of suicide [14].

There have been a few considers that center on helpless populaces and tall chance populaces counting understudies, the matured, ladies, outfitted powers, ranchers,

transient populaces, and those with persistent physical and mental ailment [15]. The truth that the suicide rate is still lower than that of other nations may be clarified by the impact of religion and the good relations between family individuals all through the nation. In totally different social orders as well as in India, religion plays an awesome part, particularly in a conviction in a great beyond, which turns away sentiments of misery, a feeling that has been portrayed as a critical inclining calculate for suicide [15, 18]. The suicide rate decreases as age increases among guys and females, which can be clarified by the increment within the level of duty of guys and females toward their families [16].

Suicides are especially a matter of concern in LMICs like India, where assets to control suicide avoidance are rare. A survey of investigations on suicides and sex in India was performed and this survey focused on rates of suicide, strategy involved in suicides, chance variables, and forerunner components of suicides. The suicide rates and extent of passing due to suicide were especially higher in youthful ladies. The hanging method of suicide was found to be the foremost usual strategy of suicide [20]. This audit further distinguished a few holes in writing, such as the need to write on people who recognize the sexual orientation double of male and female, and the got to get sexual orientation contrasts in chance and defensive components for suicide [17].

3 Methodology

The core of this dataset pertains to rates of suicide in India. Consisting of a combined count of 237,520 rows and 7 columns, this dataset is pivotal for the purposes of analysis of the work has been carried out (Table 1).

Here we use Jupiter library to run the codes (Fig. 1).

Upon analyzing depicted trend (Fig. 2) it becomes evident that the suicide rate exhibited a significant upward trajectory from the years 2002–2019. However, a pivotal juncture is discernible in 2011, marking a notable reversal in this alarming escalation. Subsequent to this juncture, there is a discernible downward trajectory in the graph, indicative of a favorable development.

This observed reduction in the suicide rate post-2011 is certainly a positive and encouraging trend. It offers a glimmer of hope for a future characterized by further diminishment in the prevalence of suicides. While the complex and multifaceted nature of factors influencing suicide rates necessitates comprehensive ongoing efforts, the observed decline provides impetus for continued endeavors aimed at mental health advocacy, accessible support systems, and targeted interventions. It is the aspiration of stakeholders that this trend toward reduction persists and is reinforced in the years to come, ultimately contributing to a safer and more resilient societal landscape (Fig. 3).

Table 1 Suicide rates in India

S. No.	State/UT	Cause of suicide	Male	Female	Transgender	Total
0	Andaman & Nicobar Islands	Sleeping pills	0	0	0	0
1	Andaman & Nicobar Islands	Drowning	2	1	0	3
2	Andaman & Nicobar Islands	Self-immolation	1	2	0	3
4	Andaman & Nicobar Islands	Fire arm	0	0	0	0
5	Andaman & Nicobar Islands	Hanging	99	25	0	124
...	...	...	...	...	...	...
365	West Bengal	Poison	1631	1299	0	2930
366	West Bengal	Jumping	43	11	0	54
367	West Bengal	Coming under running vehicle/ trains	39	3	0	42
368	West Bengal	Touching live electric wire	35	8	0	43
369	West Bengal	By other means	1414	594	0	2008

Fig. 1 Library imported to perform this analysis and information about the dataset

```
RangeIndex: 237519 entries, 0 to 237518
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   State       237519 non-null object
 1   Year        237519 non-null int64
 2   Type_code   237519 non-null object
 3   Type        237519 non-null object
 4   Gender      237519 non-null object
 5   Age_group   237519 non-null object
 6   Total       237519 non-null int64
dtypes: int64(2), object(5)
memory usage: 12.7+ MB
```

3.1 Comprehensive State-Wise Breakdown of Suicide Cases

This examination delves into the varying frequencies of suicides within each state, uncovering geographical disparities and potential influencing factors. Such appraisal supports understanding the dissemination and hidden intricacies of self-destruction rates across the assorted provinces of India (Fig. 4).

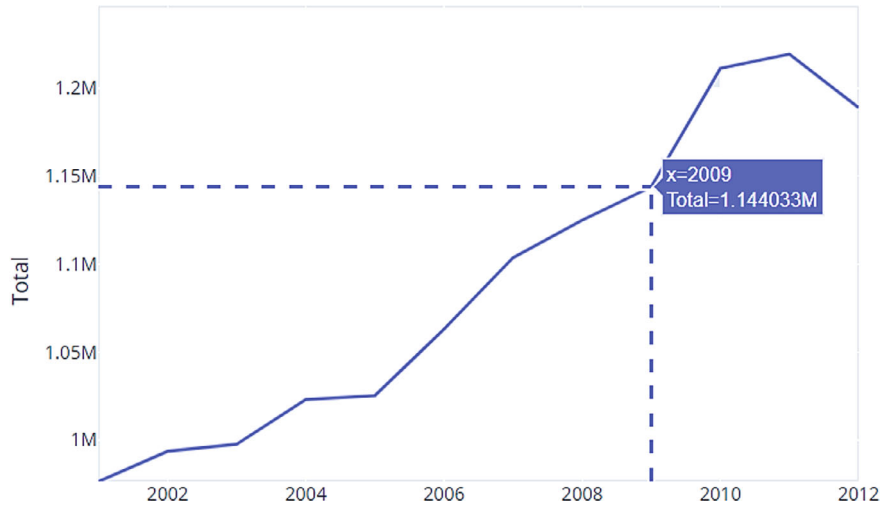


Fig. 2 Suicide case distribution as per year

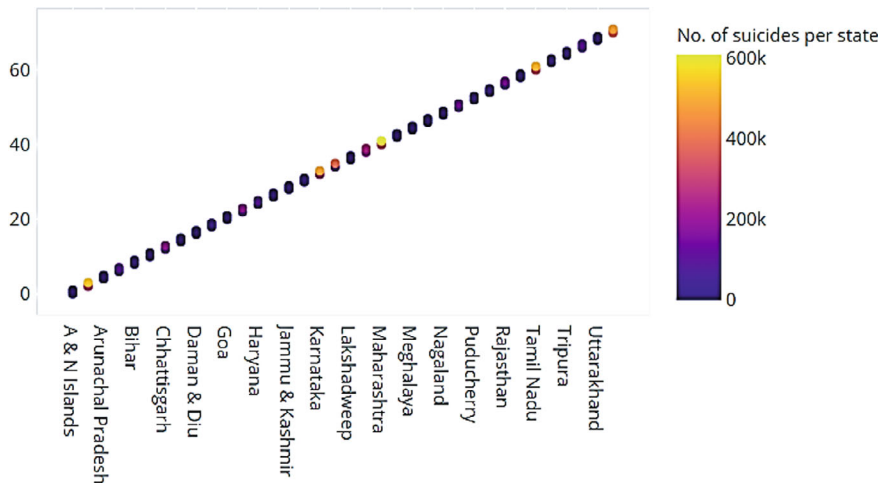
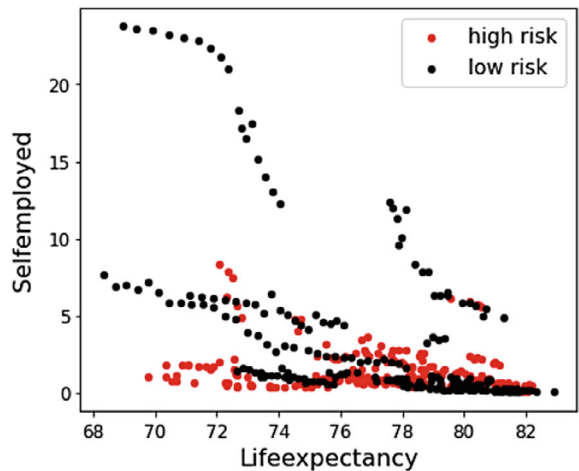


Fig. 3 Suicide rate as per states of India

Leading bunching examination on the utilized segment permits us to recognize unmistakable gatherings with changing futures. By investigating designs inside this populace, we disclose significant bits of knowledge into how word related elements could impact life span, offering a nuanced point of view on the unpredictable transaction among business and future.

Figure 5 makes obvious that the suicide rate is significantly greater for males than for women, 64% compared to 36%. This discrepancy highlights a worrying

Fig. 4 Clustering analysis of people who are employed and their life expectancy



trend where men are more likely to have suicide thoughts and deeds. Numerous studies have shown that males experience depression at greater rates than women do. The following stage, after diving deeper into the specifics, is to determine the age at which males are most susceptible to committing suicide. The understanding of the underlying causes of the increased suicide rate among males depends on this. Figure 6, which shows a graphic representation of the relationship between age and suicide events among men, may be used to visualize this data. This graphic makes the pattern easier to see and may assist determine which age groups need more care and support.

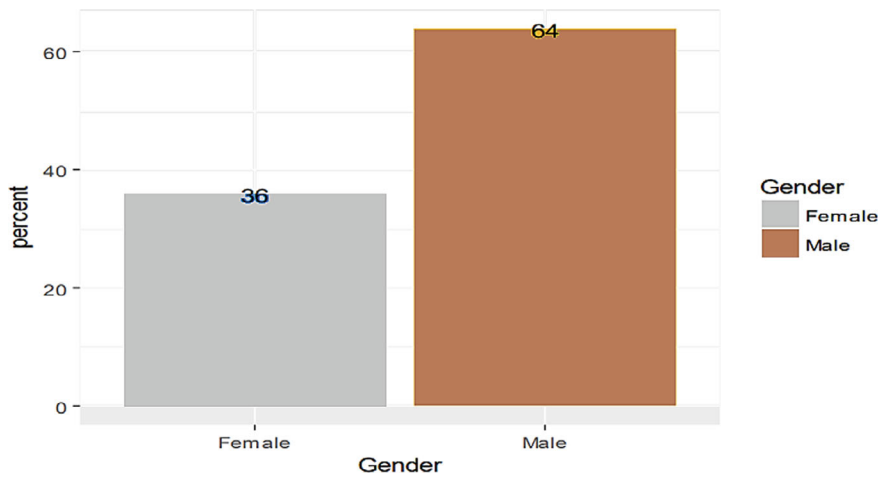


Fig. 5 % of gender

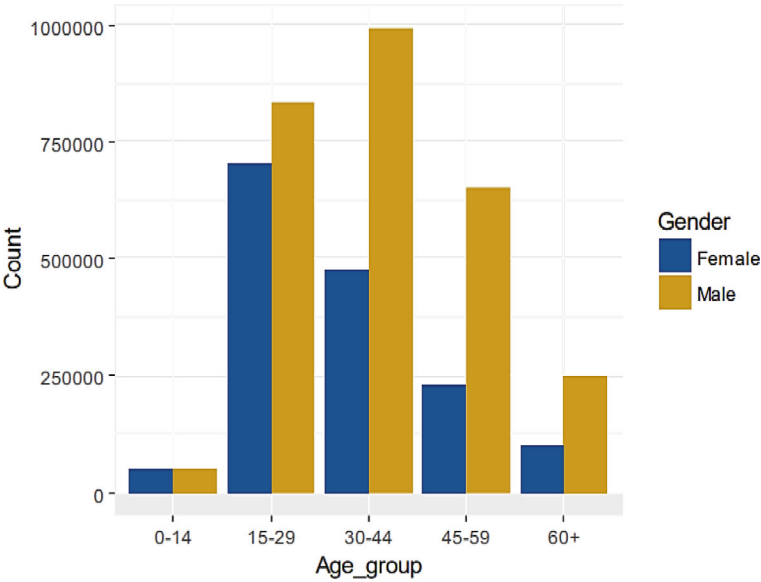


Fig. 6 Suicide count of different age group according to gender

The data in Fig. 7 show an intriguing pattern: more female suicides than male suicides in the 30–44 age range are recorded among those aged 15–29. This tendency tends to happen before marriage and is more apparent in the younger age group. It’s important to note that the data indicates a change in the gender pattern for suicide dependent on age.

We gain a wider perspective on the data by turning to Fig. 8. The total number of recorded suicides among people aged 15–29 may be seen in this statistic. We can discern the cumulative impact of suicides in this particular age group thanks to this wider viewpoint.

Family Problems push the youth the most to commit suicide. It is the top cause for suicide, accounting for **32%** of all the suicides in India followed by illness and Drug Abuse (Fig. 9).

Data shows that more than 50% of individuals who took their own lives in 2017 hanged themselves, the data reveals a striking trend: in 2017. This statistic underscores the prominence of this method in instances of suicide during that year.

Figure 10, suicide is the top leading cause of death for both Males and Females falling in the **age group of 18–30** years. The majority of the ages of 34% of adolescents who committed suicide occurred between 18 and 30.

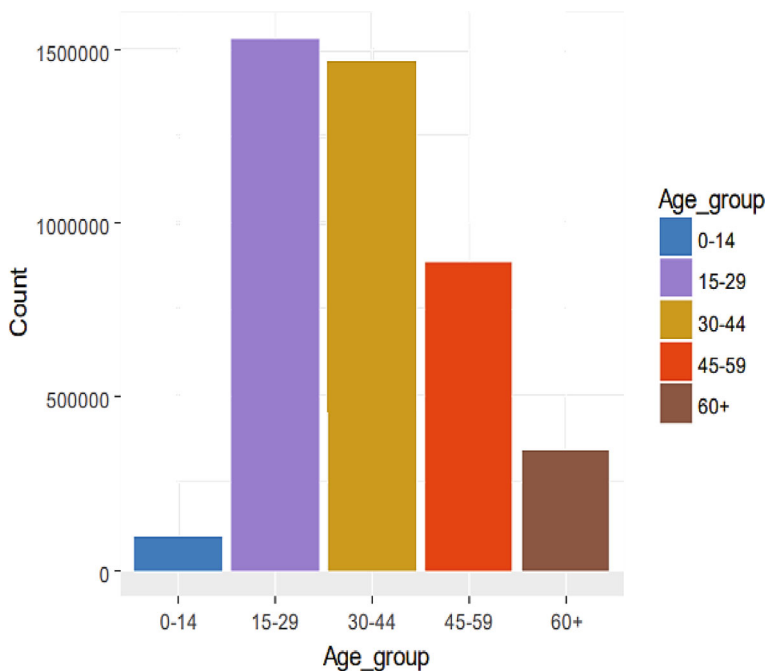


Fig. 7 Suicide count of different age group

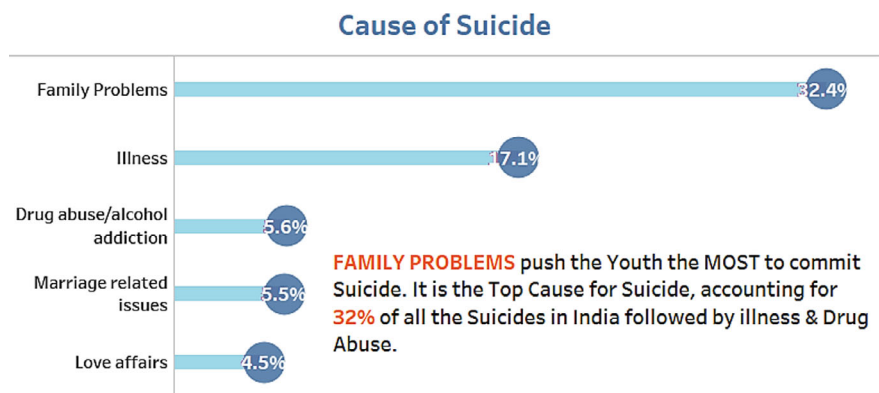


Fig. 8 Suicide count in contrast to cause of suicide

4 Results and Discussion

In this phase of our analysis, we have managed to create a visual representation that gives us a clear understanding of the total incidents of suicide reported, categorized by both state and year. This visual aid is instrumental in helping us grasp the

Cause Of Suicide	No. of Deaths	% of Total No. of Deaths
Hanging	64,706	50.07%
Poison	36,000	27.86%
By Other Means	7,367	5.70%
Fire / Self-Immolation	6,860	5.31%
Drowning	6,236	4.83%
Coming Under Vehicles ..	3,814	2.95%
Jumping	2,203	1.70%
Sleeping Pills	1,014	0.78%
Touching Live Electric ..	559	0.43%
Fire Arms	480	0.37%

Fig. 9 Pattern of suicide in contrast to number of deaths and percentage

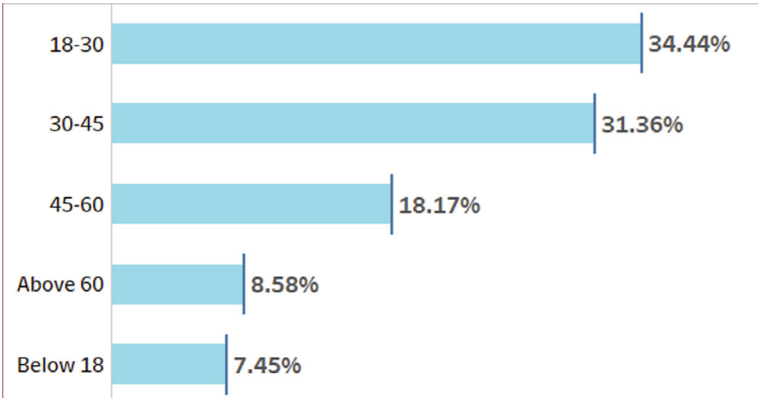


Fig. 10 Suicide percentage of India in different age groups

trends more effectively. This graphical representation, extending from 2012 through 2020, provides a straightforward depiction of the prevalence of suicides and their distribution across various age groups.

As you examine the plotted data, you will notice a particularly concerning trend. Beginning in 2016, the instances of suicides attributed to family issues have exhibited a consistent upward trajectory. By the time 2020 arrives, based on the dataset we have utilized, the number of reported suicides connected to family issues would have escalated to a staggering 35,000 cases. This is an alarming revelation that emphasizes the importance of addressing these family-related concerns. We have also delved into the broader context of suicide rates, dissecting them across different age groups. This insightful information is laid out in another illustrative Fig. 11. This portrayal allows us to gauge the overall rate of suicides that have been reported over the years,

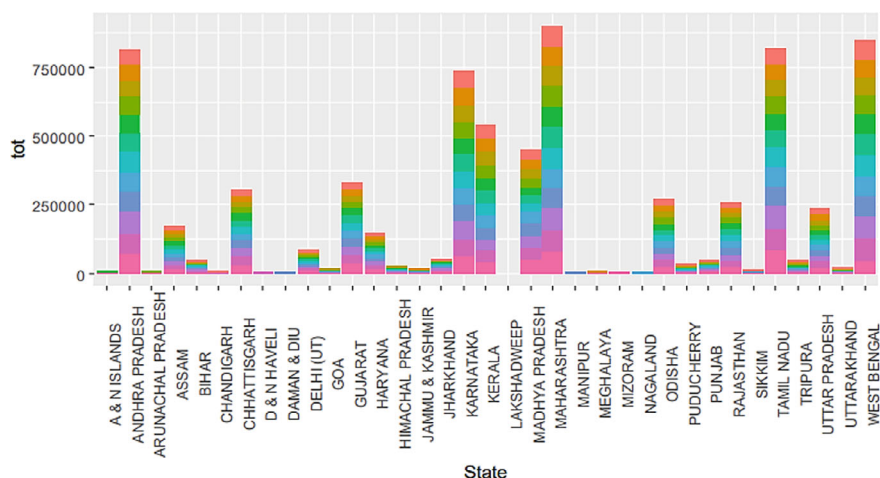


Fig. 11 Different states of India and their suicide count

pinpointing the variations among the different age categories. Furthermore, we have turned our attention to the geographical distribution of suicide incidents, focusing on each state's contribution. Through yet another Fig. 11, we have showcased the total count of suicide cases reported in each state. It is noteworthy that West Bengal, Andhra Pradesh, and Maharashtra emerge as the states with the highest numbers of reported suicides. In essence, these results underscore the significance of visualizing data to extract meaningful insights. The trends we've highlighted, especially the concerning increase in suicides linked to family issues, necessitate a proactive approach in addressing the underlying causes and providing the necessary support systems.

The infographic analyses the methods used to determine Indian states with the highest and lowest suicide rates, revealing the top 10 states with higher rates and the bottom 10 with lower rates. The data is analyzed for clarity, highlighting the disparity in suicide rates between states as define in Fig. 12.

Covid-19 and Lockdown (March 2020) Covid-19 had drastically hit the Indian Economy and the Nationwide Lockdown caused unemployment in every sector, affected mental health of many people, other illness, devastating impact on the life of students at multiple levels and young employees who had a hard time starting their careers. All this caused a 10% rise in the suicide rate in 2020 as compared to 2019 also demonetization (Nov 2017) This had caused a Negative Impact in Indian Economy which led to Delayed/Half Salaries, Cash Shortages, Shutdown of Businesses or Shops, Job loss, Inflation. And hence there was a 3.56% increase in the suicide rate in 2018 as compared to 2017.

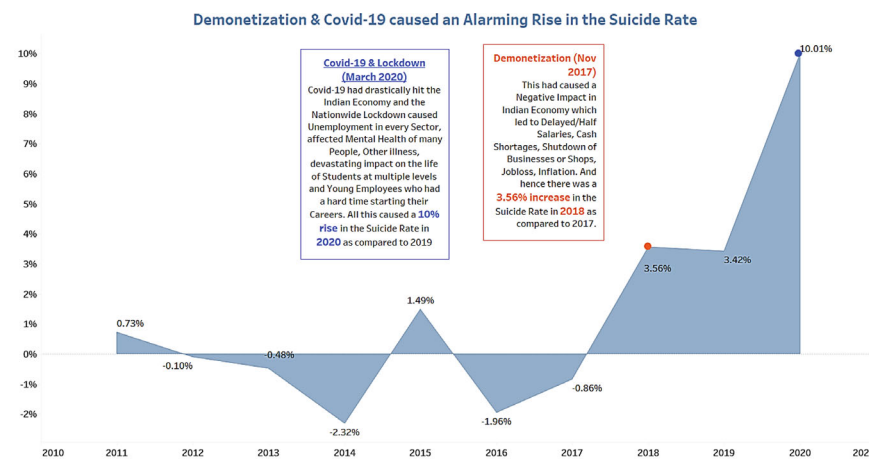


Fig. 12 Area graph of suicide percentage in India for different years

5 Conclusion, Limitations, and Future Scope

Our research underscores the pressing need for delving deeper into strategies aimed at preventing suicides. This entails not only gaining a more profound comprehension of the intricate nature of suicidal thoughts and behaviors but also unraveling the fundamental triggers that give rise to this distressing problem on an individual level.

One especially perplexing pattern we have recognized is the disturbing ascent in suicides among understudies matured 16–24. The heaviness of scholarly assumptions combined with the weight of underperformance has driven a few people down this desperate way. To relieve this disrupting issue, guardians really should play a functioning job in supporting their youngsters' versatility and engaging them to defy difficulties head-on. By encouraging their ability to develop through misfortune, we might possibly destroy the poisonous cooperative energy between scholarly pressure and self-destructive propensities.

Nonetheless, the issue doesn't end there. Our examinations have likewise featured another weak segment: ranchers matured 29–45. This gathering faces a variety of stressors, frequently coming from their familial obligations and the absence of admittance to essential government assets like reliable water system offices. The shortage of these assets can drive people to the edge. To balance this sad reality, a deliberate exertion from the public authority is imperative. The arrangement of more strong and effective water system arrangements could fundamentally lighten the difficulties looked by these ranchers and possibly add to checking the troubling pattern of rancher suicides.

Later on, specialist should want to use AI calculations to break down the effect of various gamble factors and recognize relationships between them. This will empower us to give customized proposals to understudies and ranchers at different phases of their lives, subsequently helping with the avoidance of self-destruction.

Acknowledgements Authors; Thanks to the MIT WPU organization for providing the infrastructure for this work.

References

1. National Crime Records Bureau. Available at: <https://ncrb.gov.in/en/accidental-deaths-suicides-in-india>, 25 August
2. Vijaykumar L (2007) Suicide and its prevention: the urgent need in India. *Indian J Psychiatry* 49(2):81–84. <https://doi.org/10.4103/0019-5545.33252>. PMID 20711387. PMCID PMC2917089
3. Cowling K, Dandona R, Dandona L (2014) Social determinants of health in India: progress and inequities across states. *Int J Equity Health* 13(1). <https://doi.org/10.1186/s12939-014-0088-0>
4. DiClemente RJ, Hansen WB, Ponton LE (2013) *Handbook of adolescent health risk behavior*. Springer Science & Business Media
5. Gajalakshmi V, Peto R (2007) Suicide rates in rural Tamil Nadu, South India: verbal autopsy of 39,000 deaths in 1997–98. *Int J Epidemiol* 36(1):203–207. <https://doi.org/10.1093/ije/dyl308>
6. Gunnell D, Bennewith O, Hawton K, Simkin S, Kapur N (2005) The epidemiology and prevention of suicide by hanging: a systematic review. *Int J Epidemiol* 34(2):433–442. <https://doi.org/10.1093/ije/dyh398>
7. Gunnell D, Fernando R, Hewagama M, Priyangika WDD, Konradsen F, Eddleston M (2007) The impact of pesticide regulations on suicide in Sri Lanka. *Int J Epidemiol* 36(6):1235–1242. <https://doi.org/10.1093/ije/dym164>
8. Gururaj G, Isaac M, Subbakrishna D, Ranjani R (2004) Risk factors for completed suicides: a case-control study from Bangalore, India. *Injury Control Saf Promot* 11(3):183–191. <https://doi.org/10.1080/156609704233289706>
9. Dandona R, Bertozzi-Villa A, Kumar GA, Dandona L (2017) Lessons from a decade of suicide surveillance in India: who, why and how?. *Int J Epidemiol* 46(3):983–993
10. Vijayakumar L (2010) Indian research on suicide. *Indian J Psychiatry* 52(Suppl 1):S291–6. <https://doi.org/10.4103/0019-5545.69255>. PMID 21836697. PMCID PMC3146189
11. Joshi R, Guggilla R, Praveen D, Maulik PK (2014) Suicide deaths in rural Andhra Pradesh—a cause for global health action. *Trop Med Int Health* 20(2):188–193. <https://doi.org/10.1111/tmi.12416>
12. Knipe DW, Chang SS, Dawson A, Eddleston M, Konradsen F, Metcalfe C, Gunnell D (2017) Correction: suicide prevention through means restriction: impact of the 2008–2011 pesticide restrictions on suicide in Sri Lanka. *PLOS ONE*
13. Kumar S, Verma AK, Bhattacharya S, Rathore S (2013) Trends in rates and methods of suicide in India. *Egypt J Forensic Sci* 3(3):75–80. <https://doi.org/10.1016/j.ejfs.2013.04.003>
14. Manoranjitham SD, Jayakaran R, Jacob KS (2006) Suicide in India. *Br J Psychiatry* 188(1):86–86. <https://doi.org/10.1192/bjp.188.1.86>
15. Manoranjitham SD, Rajkumar AP, Thangadurai P, Prasad J, Jayakaran R, Jacob KS (2010) Risk factors for suicide in rural south India. *Br J Psychiatry* 196(1):26–30. <https://doi.org/10.1192/bjp.bp.108.063347>
16. Manoranjitham S, Jayakaran R, Jacob K (2006) Suicide in India: the need for a national policy. *Indian J Psychiatry* 48(1):72. <https://doi.org/10.4103/0019-5545.31626>
17. Patel V et al (2012) Suicide mortality in India: a nationally representative survey. *The Lancet* 379(9834):2343–2351. [https://doi.org/10.1016/s0140-6736\(12\)60606-0](https://doi.org/10.1016/s0140-6736(12)60606-0)

18. Rao G, Vidya K, Sriramya V (2015) The Indian “girl” psychology: a perspective. *Indian J Psychiatry* 57(6):212. <https://doi.org/10.4103/0019-5545.161480>
19. Suicide (2021) www.who.int. <https://www.who.int/news-room/fact-sheets/detail/suicide>
20. Vijayakumar L (2015) Suicide in women. *Indian J Psychiatry* 57(6):233. <https://doi.org/10.4103/0019-5545.161484>

Astroturfing: Understanding and Addressing the Influence of Fake Grassroots Movements Across Diverse Domains on Social Media Platforms



Akshata Balasaheb Badade and Rajesh Kumar Dhanaraj

Abstract Astroturfing is a deceptive practice used to disguise or create the illusion of fake support or opposition to a certain cause, product, or campaign by the people on the ground when, in fact, it is not. While social platforms serve as key communication and information broadcasting mechanisms, they are becoming more superficial; this fraudulent practice is effectively carried out on social media platforms using fake accounts, bots, or individuals who are paid to enhance or suppress online campaigns or content artificially. Astroturfing, a form of political manipulation and marketing, distorts public opinion by using social media algorithms to spread false claims and silence genuine voices, opinions, or sentiments of the public. From political campaigns to search engines, astroturfing declines the authenticity of online discourse and decision-making processes by misrepresenting assumptions, manipulating public opinion, and undermining the credibility of information shared online at risk. Uncovering astroturfing can maintain trust in online interactions, shield consumers from false information, and support fairness and democracy, promoting transparency, authenticity, and trust in online discussions and decision-making. This paper highlights the impact of astroturfing across domains and the need for robust identification mechanisms for astroturfing detection that can identify coordinated patterns to effectively combat fraudulent online practices in various domains and expose astroturfing efforts.

Keywords Astroturfing · Cross-domain datasets · Grassroots campaigns · Social media strategic influence

A. B. Badade · R. K. Dhanaraj (✉)
Symbiosis Institute of Computer Studies and Research (SICSR), Symbiosis International
(Deemed University), Pune, Maharashtra, India
e-mail: sangeraje@gmail.com

A. B. Badade
e-mail: akshatabadade2110@gmail.com

1 Introduction

As social media platforms gain influence, they become breeding grounds for astroturfing, a deceitful practice of creating fake grassroots support to manipulate public opinion. This affects consumer choices, political discussions, as in Ref. [1], and financial markets. Astroturfing campaigns employ social media posts, fabricated news articles, and organized comment sections to disseminate false information and propaganda, giving the false impression of extensive backing for a specific agenda, product, campaign, political party, or person. This manipulation negatively impacts public perception, influencing consumer decisions, swaying elections, and potentially disrupting markets.

To combat these fraudulent practices, it is important to have effective and robust astroturfing detection methods. Although benchmark datasets such as “Botometer” proposed by authors in the paper referenced [2], “Twitter15” and “Twibot-20” and other datasets provide valuable resources for training and evaluating detection algorithms as referenced in [3]; however, the development of robust and early detection of astroturfing campaigns with improved accuracy is needed. Current laws and regulations often lack the depth needed to address the evolving approach to astroturfing campaigns. Moreover, customized anomaly detection techniques for specific social platforms can play a vital role in identifying suspicious activity patterns and minimizing the impact of these deceitful campaigns.

2 The Major Contributions of This Paper

1. Diverse domains where astroturfing is carried out on social media platforms
2. Negative impact or influence of astroturfing in diverse domains
3. Publicly available benchmark cross-domain astroturfing datasets.
4. Need or significance of developing effective and robust astroturfing detection methods to combat deceptive practices across diverse domains online.
5. Rules and regulations or policies to combat astroturfing.

Further work in this study is organized as follows: Sect. 3 contains a literature review of astroturfing that describes diverse domains where astroturfing is getting carried out online, the impact or negative influence of fake grassroots movements on social media, the cross-domain astroturfing benchmark datasets available for researchers, the need or significance of cross-domain astroturfing detection to combat deceptive online practices on social platforms, and the current rules and regulations or policies to combat astroturfing worldwide, Sect. 4 contains the limitations or challenges faced by the researchers in astroturfing detection on social platforms, and the last Sect. 5 highlights the conclusion and future scope of further study in this area.

3 Literature Review

3.1 Diverse Domains Where Astroturfing Is Carried Out on Social Media Platforms

Research indicates that astroturfing campaigns on social platforms are prevalent across various domains, substantially influencing public opinion and manipulating their decision-making process to benefit their specific interests.

3.2 Politics and Public Policy

By drawing on global research, the authors in their paper [4] aim to uncover how political parties use astroturfing to influence public opinion and elections and conceal their deceptive tactics. Their research seeks to understand the implications of astroturfing on democratic processes in Indian politics, analyzing its methods, motivations, and impacts.

3.3 E-Commerce and Online Reviews

Consumer behavior in the digital era is heavily influenced by online reviews, which are vital in making purchasing decisions across all age groups online. In the digital age, online reviews heavily influence consumer behavior online and are considered increasingly important across all age groups in making purchasing decisions on product review platforms. In this context, another study by researchers in [5] delves into the psychological aspects that enhance online searching, like social proof, less mental effort, and the fear of not keeping up. According to the study, positive online reviews not only increase the purchase likelihood of the product or the service but also aid in consumer decision-making processes and foster brand loyalty.

3.4 Corporate and Marketing Sector

Astroturfing ensures that the sponsor's identity remains hidden during the mobilization campaign. Sam (P19) and researchers in their paper [6] brought up multiple "corporate astroturfing" cases by companies like tobacco, energy, and insurance. Government entities also employ similar strategies to create a false sense of agreement, giving the impression that "everyone is in favor of the government's actions and political decisions."

3.5 Social Media and Online Communities

This study by authors in the paper referenced [7] used social media network analysis (SMNA) to examine potential astroturfing campaigns on Twitter in support of the Adani Carmichael coal mine in 2017. It found a subset of consciousness suspicion of similar tweets supporting Adani, suggesting an attempted fraud. However, although these positions had little impact on public opinion and were ineffective, the study highlights the need to explore modern digital analytics or artificial intelligence to expose secretly carried out social media manipulative campaigns that influence public opinion for specific interests to maintain the integrity of public discourse.

3.6 Public Health and Environmental Issues

This study examines the relationship between climate change, health security, and misinformation in Africa. Climate change poses unique challenges to African public health, and misinformation complicates the efforts to address these challenges. As revealed in a study by authors in paper [8], their literature review and interview findings identify strategies to combat misinformation focusing on vaccination theory. The study proposes that people should be exposed to less influential false information beforehand to develop immunity and resistance to the adverse effects of climate change. It also suggests using specific communication tactics to counter false information about climate change. Research shows that misinformed individuals are unwilling to build capacity to combat and advocate for targeted communication strategies to combat climate anomalies. In Africa, collaborative efforts are necessary to combat misinformation in the context of climate change and health security in Africa.

3.7 Tourism and Hospitality Sector

Electronic word of mouth (eWOM) makes it easy for people to share their opinions about products and services through social media. However, the rise of digital connectivity has led to false assumptions. The authors in their study referenced [9] examine how eWOM affects the Saudi tourism industry through digital and social media. It looks at eWOM as a key factor, customer purchase intentions as an outcome, and brand status as a facilitating aspect. False eWOM reviews such as “astroturfing” or “soak puppets” can influence consumer decision-making in the tourism industry. The Saudi tourism industry must address fake reviews by monitoring inaccurate claims, encouraging transparency, educating consumers, and partnering with review Web sites. By combating fraudulent information online, Saudi tourism can provide

reliable information to customers, boost their confidence in making bookings, and drive growth in the industry.

3.8 Influence of Astroturfing in Cross-Domains

The astroturfing domain reflects a misleading public opinion in various domains on online forums. In one study, researchers in [10] examined coordinated link sharing behavior (CLSB) on Facebook regarding false claims of hydroxychloroquine as a “cure” for COVID-19, which Dr. Stella Immanuel promoted during the “America’s Frontline Doctors” press conference. Through the analysis of 7737 public Facebook posts mentioning Immanuel using CrowdTangle and identifying CLSB instances with CooRnet, researchers found a network of US-based pro-Trump, QAnon, and anti-vaccination accounts working together to spread misinformation about COVID-19. Additionally, accounts from Brazil, France, and Africa were found promoting hydroxychloroquine or criticizing government responses to the pandemic. Also, sociological research in a paper referenced [11] shows that advocacy organizations are trusted by their constituents. However, businesses and external interests use covert tactics, known as “astroturfing,” to push their own agendas through seemingly independent grassroots groups. These practices harm trust in advocacy groups, leading to stigmatization. They also explore how revealed covert campaigns negatively impact all advocacy groups, highlighting the damage caused by unethical political tactics. The researchers conducted two survey studies, showing that a local advocacy group was a front for a corporation or a think tank. This practice of astroturfing resulted in a notable decline in trust in advocacy organizations overall. Another study by researchers in Ref. [5] identifies the influence of online reviewers as crucial for consumers and businesses in today’s market. The study states that it is imperative for businesses to actively supervise their online reputations, tackle ethical issues, and adjust to the digital landscape. However, ethical challenges arise when certain companies engage in fabricated reviews, astroturfing, or incentivizing positive feedback for promotion. These actions erode consumer trust and can lead to legal consequences if disclosed. Ethical considerations regarding online reviews are important to consumers of all age groups, who value honesty and transparency. Addressing these ethical dilemmas is important for businesses looking to maintain trust and customer confidence, as unethical practices can have long-term consequences, such as spoiling customer relations and legal proceedings.

4 Publically Available Benchmark Cross-Domain Astrourfing Datasets

4.1 *ClimaConvo*

Climate change discussions on social media are important, but this area lacks data and analysis. This paper referenced [12] explores climate discourse on Twitter using the ClimaConvo dataset, which includes 15,309 tweets. Researchers in their study state that annotations cover relevance, stance, hate speech, direction of hate, and humor, providing a detailed understanding of the discourse dynamics. They compare algorithms for various tasks and analyze sentiment changes and language nuances. They state that their findings, including distribution patterns, stance pervasiveness, and hate speech trends, are valuable for researchers and policymakers in future studies in this domain. Their dataset and resources can be found at ClimaConvo.

4.1.1 TwiBot-22

Twitter bot detection is crucial for combating misinformation, moderating social media, and preserving the integrity of online conversations. Current methods rely on the network structure of Twitter and have shown promise in identifying new types of bots that traditional methods miss. Though, there is a lack of large-scale benchmark datasets for graph-based Twitter bot detection, limiting the development and evaluation of new approaches. To address this, the authors in their paper referenced [13] introduce TwiBot-22, a comprehensive benchmark dataset that includes the largest dataset to date, diverse entities and relationships within the Twitter network, and improved annotation quality. They also implement 35 bot detection methods and test them on 9 different datasets, including TwiBot-22, to ensure better comparison and understanding of research trends. To facilitate further research, the authors present the TwiBot-22 research framework that combines code and data management strategies and allows researchers to explore new strategies and data types continuously. The TwiBot-22 benchmark and evaluation framework will be available to the public here at Twibot-22.

4.1.2 Jacksepticeye Tweets

The dataset contains more than 4900 tweets from popular rapper Lil Papivoni. It offers a unique opportunity to explore the online network of one of the world's most influential social media personalities. By examining this data, we can understand how her followers communicated with her on Twitter and how quickly she won on the social platform. This dataset can investigate what happened within those posts, including how the followers evolved, what they commented on, which tweets they liked, and which tweets they retweeted. The dataset is available in CSV format and

includes information about these tweets, such as content and audience behavior, such as retweets, likes, and other metrics. It contains Tweet ID, Tweet content, likes, replies, re-tweets retweeted, relevance, and other information. This information can be used for sentiment analysis, interaction analysis, identifying coordinated patterns, and identifying popular topics among audiences through topic modeling. The dataset is available at Jacksepticeye Tweets.

4.1.3 US Senate Analysis

This dataset consists of 4 tables. The dataset is valuable for sentiment analysis, network analysis, monitoring engagement metrics, and identifying trending topics among different audiences. All four tables given below are interconnected and can be merged with information:

- a. ***Senators.csv***: This data is sourced from United States Senators and combined with Twitter API. It includes senators' names, age, occupation, location, state coordinates, Twitter data (screen_name, followers, follows, id), gender, and race.
- b. ***relationships.csv***: This data contains relationships between senators with columns person1, person2, and relationship (following, not following, following only by person x).
- c. ***dataset_with_topics.csv***: This table contains the list of 250 k tweets from US senators from 2008 to 2023. It also shows the results of three BERTopic clustering with different minimum cluster sizes.
- d. ***topic_info_150***: This data provides information on topics with a minimum cluster size of 150 and manual labeling. This table can be joined with the previous one for tasks like multi-class classification.

The dataset is available at US Senate Analysis.

4.1.4 US Congressional Tweets Dataset

The "US Congressional Tweets Dataset" is a comprehensive collection of tweets from US Congressional members from 2008 to 2017. It is valuable for organizations looking to understand legislative trends and influences for effective lobbying strategies.

The dataset is divided into two main parts:

- a. ***user_df***: This data frame provides detailed information about Congressional Twitter account members, including date of account creation, number of followers and friends, profile information, and Twitter-specific features.
- b. ***tweets_df***: This data frame contains the actual tweet data from these congressional accounts, including tweet timestamps, popularity indicators, tweet text content, user_id, language, source, and other metadata.

The dataset is available at The_US_Congressional_Tweets_Dataset.

4.1.5 WarMM-2022

In another study conducted by authors in [14], they mention that their dataset encompasses a broad spectrum of topics, specifically concentrating on warfare-related literature. Even so, they admit that their time is somewhat restricted. Their dataset provides a comprehensive view of the online context that Russian Internet users faced in real time during the war. This feature is believed to benefit the NLP community and future research on promotional campaigns in the social sciences, complementing traditional methods such as content analysis, grouping their focus, analysis, and testing. Additionally, the authors prioritize ethics to protect Internet users and publishing houses. Useful decisions were made, such as publicly available documentation, limited information in the list of links, and not sharing metadata to protect privacy and copyright.

These decisions may hinder replicability and require more effort, but they were made to prevent harm. They also avoided sharing identifiable information or promoting political beliefs to ensure safety, especially regarding anti-war sentiments in Russia. Their goal was to raise awareness about disinformation without endorsing specific political views. The dataset is available at WarMM-2022.

4.1.6 Significance of Cross-Domain Astroturfing Detection to Combat Deceptive Practices Online

The consequences of poor visibility of astroturfing are far-reaching, affecting not only public opinion but also democratic processes and consumer confidence.

Without effective measures to deal with astroturfing, open and transparent communication on social platforms will be significantly undermined. Future collaboration between researchers, policymakers, and technology companies will help us to tackle this problem. Increasing the search and making them more transparent can contribute to minimizing the influence of astroturfing. It will also conserve the originality and sincerity of online communications. Our discussion is built on past research from scholars in papers [15] and [16] on how to combat AstroTurf in different fields during different eras. Our study aims at identifying the effectiveness of astroturfing on different social platforms in different areas. From figuring out many examples through our code, we realize that many research fields are widely susceptible to astroturfing. We found many examples in political campaigns, consumer surveys, and social movements, highlighting the prevalence of astroturfing practices. Future collaboration to raise awareness from researchers, policymakers, and technology companies emerge to develop more sophisticated and practical tools to identify astroturfing campaigns and nuances. These findings not only validate the significance of astroturfing identity but also the necessity to strengthen the identification tool. It is designed to fend off the corrupted manipulations and guarantee the integrity of online discussions and communications. The scholars in the paper referenced [17] likewise emphasize the importance of identifying astroturfing in maintaining online forums.

4.1.7 Rules and Regulations or Policies to Combat Astroturfing Campaigns

Astroturfing can sway policy decisions for specific interests, disregarding the public and intensifying social inequalities. Although it cannot be entirely eliminable, a mix of regulations, platform policies, technology, education, and awareness can mitigate its influence and foster a more transparent digital environment. Some of the regulatory policies are mentioned below:

- i. The Department of Consumer Affairs, in collaboration with the Department and Bureau of Indian Standards (BIS), introduced Indian Standard (IS) 19000:2022 “Online Consumer Reviews—Principles and Requirements for their Collection, Moderation and Publication” to combat fake reviews on e-commerce Web sites and protect consumer rights. These regulations apply to all online platforms, displaying consumer feedback and preventing misleading reviews [18].
- ii. The FTC regulates deceptive marketing practices, including astroturfing, through the FTC Act, which prohibits unfair or deceptive acts in e-commerce. It includes clear disclosure requirements for paid actors and recent actions to protect consumers online [19].
- iii. Astroturfing campaigns use false information and propaganda to manipulate public opinion through tactics like social media posts and fake news articles. By being aware of these strategies and uncovering the truth, we can expose the manipulation and minimize their influence on policy discussions. This requires media literacy, fact-checking skills, and knowledge of astroturfing techniques to promote informed public discourse and avoid making decisions based on fabricated opinions [20].
- iv. Some fossil fuel companies engage in “greenwashing campaigns” to mislead the public about their environmental responsibility. This includes tactics like astroturfing to divert attention from the negative impact of their operations on environmental conditions. Terms such as disinformation, fake news, spin, and greenwashing describe unethical practices that affect the retention of communications professionals. Despite ethical codes that emphasize transparency, astroturfing damages the reputation of a business [21]. The Federal Trade Commission’s guidance targets deceptive green marketing claims to educate consumers about “greenwashing,” which contains misleading claims about environmental benefits. The guidance covers a variety of areas, including health and environmental benefits. Customers are provided with examples to evaluate these claims carefully, including non-toxic claims and definitions of “renewable”. Another study examined global companies’ use of social media to promote environmentally friendly initiatives from 2019 to 2021. Compared to a study in 1993, researchers found that more than 70% of claims by companies are misleading in that case [22].

5 Limitations and Challenges for Astroturfing Detection on Social Platforms

Below are some limitations and challenges faced by the researchers in identifying these deceptive practices carried on social media platforms to sway genuine public opinion to the benefit of their specific interests:

- a. ***Leveraging machine learning for social media analysis:*** Effective utilization of machine learning in social media analysis demands expertise from skilled professionals and data scientists. Developing successful models necessitates an advanced understanding of algorithms, data processing techniques, and coding languages. However, the scarcity of such professionals can be a challenge in fully leveraging machine learning for social media analysis.
- b. ***Adversarial attacks pose a significant challenge in astroturfing detection:*** Astroturfers intentionally do things to confuse or mislead detection algorithms. Researchers must actively develop robust models that can withstand such counter-attacks.
- c. ***Challenges in ground truth labeling for ML models:*** Creating reliable labeled datasets to train machine learning models is difficult. Handwriting is time-consuming and subjective. There is a need for better methods of establishing ground truth scores to be able to train recognition models with higher accuracy.
- d. ***Evolving strategies in astroturfing detection:*** Devious astroturfers are always finding ways to circumvent the nets set against them. They use new platforms and change their content, and researchers must constantly update their detection methods.
- e. ***Navigating linguistic challenges in astroturfing detection:*** Astroturfers skillfully manipulate language to create an illusion of authenticity. Recognizing the nuanced linguistic indications that differentiate astroturfing from genuine content poses a significant challenge. Researchers must delve deeper into linguistic features and contexts to enhance the precision of their investigations.

6 Conclusion and Future Scope

All things being considered, astroturfing undoubtedly impacts social media companies. The more this fraudulent activity gets complex, the newer techniques must be used for its recognition, and laws must be chalked out to combat it effectively. After all, these measures hold the key to real and actual virtual dialogues and preserve the fundamental values of democracy, transparency, and trust in the digital domain. Indeed, our research supports the call of scholars referenced in [23], stressing that astroturfing on the Internet must be addressed at the earliest possible time before it becomes the next Internet menace. We are optimistic that consumer interests and habits will be understood more precisely with technological development, and we

will be able to produce more interesting and appropriate information for them. This trend could impact social media and other enterprises in a few years.

References

1. Schoch D, Keller FB, Stier S, Yang J (2022) Coordination patterns reveal online political astroturfing across the world. *Sci Rep* 12(1):4572. <https://doi.org/10.1038/s41598-022-08404-9>
2. Yang K-C, Ferrara E, Menczer F (2022) Botometer 101: social bot practicum for computational social scientists. *J Comput Soc Sci* 5(2):1511–1528. <https://doi.org/10.1007/s42001-022-00177-5>
3. Feng S, Wan H, Wang N, Li J, Luo M (2021) TwiBot-20: a comprehensive Twitter bot detection benchmark. In: *Proceedings of the 30th ACM international conference on information & knowledge management*, pp 4485–4494. <https://doi.org/10.1145/3459637.3482019>
4. Khader FMA, Kumar DTS (2023) Astroturfing in Indian politics: web of deception and reverse censorship. *Int J Multidiscip Res Sci Eng Technol* 6(11)
5. Kanudawala PJ, Unadkat H (2023) The impact of online reviews on purchasing decision.pdf. Retrieved 26 April 2024, from <https://yra.ijaar.co.in/wp-content/uploads/2024/02/120311.pdf>
6. Mirza S, Begum L, Niu L, Pardo S, Abouzied A, Papotti P, Pöpper C (2023) Tactics, threats & targets: modeling disinformation and its mitigation. In: *Proceedings 2023 network and distributed system security symposium*, San Diego, CA, USA. <https://doi.org/10.14722/ndss.2023.23657>
7. Hobbs M, Della Bosca H, Schlosberg D, Sun C (2020) Turf wars: using social media network analysis to examine the suspected astroturfing campaign for the Adani Carmichael Coal mine on Twitter. *J Public Aff* 20(2):e2057. <https://doi.org/10.1002/pa.2057>
8. Igbashangev PA, Humbe BT, Moses AK (2024) Exploring the nexus of climate change, public health resilience, and information accuracy in Nigeria (SSRN Scholarly Paper 4794287). <https://doi.org/10.2139/ssrn.4794287>
9. Alwahashi AW, Medjedel A (2024) e-WOM: how do fake reviews in the Saudi tourism sector impact consumers' purchasing intentions? *WSEAS Trans Bus Econ* 21:565–574. <https://doi.org/10.37394/23207.2024.21.47>
10. Gruzd A, Mai P, Soares FB (2022) How coordinated link sharing behavior and partisans' narrative framing fan the spread of COVID-19 misinformation and conspiracy theories. *Soc Netw Anal Min* 12(1):118. <https://doi.org/10.1007/s13278-022-00948-y>
11. Walker ET, Le AN (2023) Poisoning the well: how astroturfing harms trust in advocacy organizations. *Soc Curr* 10(2):184–202. <https://doi.org/10.1177/23294965221123808>
12. Shiwakoti S, Thapa S, Rauniyar K, Shah A, Bhandari A, Naseem U (2024) Analyzing the dynamics of climate change discourse on Twitter: a new annotated corpus and multi-aspect classification
13. Feng S, Tan Z, Wan H, Wang N, Chen Z, Zhang B, Zheng Q, Zhang W, Lei Z, Yang S, Feng X, Zhang Q, Wang H, Liu Y, Bai Y, Wang H, Cai Z, Wang Y, Zheng L et al (2022) TwiBot-22: towards graph-based Twitter bot detection. *Adv Neur Inf Process Syst* 35:35254–35269
14. Alyukov M, Kunilovskaya M, Semenov A (2023) Wartime media monitor (WarMM-2022): a study of information manipulation on Russian social media during the Russia-Ukraine war. In: Degaetano-Ortlieb S, Kazantseva A, Reiter N, Szpakowicz S (eds) *Proceedings of the 7th joint SIGHUM workshop on computational linguistics for cultural heritage, social sciences, humanities and literature*. Association for Computational Linguistics, pp 152–161. <https://doi.org/10.18653/v1/2023.latechclfi-1.17>
15. Park C, Mendelsohn J, Field A, Tsvetkov Y (2022) VoynaSlov: a data set of Russian social media activity during the 2022 Ukraine-Russia war. <https://doi.org/10.48550/arXiv.2205.12382>

16. Liu Y, Tan Z, Wang H, Feng S, Zheng Q, Luo M (2023) BotMoE: Twitter bot detection with community-aware mixtures of modal-specific experts. In: Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval, pp 485–495. <https://doi.org/10.1145/3539618.3591646>
17. Peng J, Choo RK-K, Ashman H (2016) Astroturfing detection in social media: using binary n-gram analysis for authorship attribution. IEEE Trustcom/BigDataSE/ISPA 2016:121–128. <https://doi.org/10.1109/TrustCom.2016.0054>
18. Framework launched by the Government of India to restrict fake reviews on e-commerce sites—Dodd-Frank, Consumer Protection Act—India (n.d.) Retrieved 28 Feb 2024, from <https://www.mondaq.com/india/dodd-frank-consumer-protection-act/1263914/framework-launched-by-the-government-of-india-to-restrict-fake-reviews-on-e-commerce-sites>
19. Truth in Advertising (2013). Federal Trade Commission. <https://www.ftc.gov/news-events/topics/truth-advertising>
20. Rochefort A (2020) Regulating social media platforms: a comparative policy analysis. Commun Law Policy 25(2):225–260. <https://doi.org/10.1080/10811680.2020.1735194>
21. Pabon J (2023) The great greenwashing: how brands, governments, and influencers are lying to you. Melbourne Books
22. Kwon K, Lee J, Wang C, Diwanji VS (2024) From green advertising to greenwashing: content analysis of global corporations' green advertising on social media. Int J Advert 43(1):97–124. <https://doi.org/10.1080/02650487.2023.2208489>
23. Zhang X, Wei P, Wang Q (2023) A hybrid anomaly detection method for high dimensional data. Peer J Comput Sci 9:e1199. <https://doi.org/10.7717/peerj-cs.1199>

An Improved DDoS Attack Preventing Model for Securing Cloud Computing Utilizing Hybrid Deep Learning Algorithms



Abdulbasit Alazzawi and Omer Salam

Abstract A Distributed Denial of Service (DDoS) attack is a malicious attempt to disrupt the normal functioning of a targeted server, service, or network by overwhelming it with a flood of internet traffic. This thesis presents a modern intelligent DDoS attack detection model based on federated learning and neural networks. This model is shortly called the DDoS-FLNN model. The proposed DDoS-FLNN uses a modern dataset called “CIC-Bell-DNS-EXF-Dataset” that is characterized by the diversity of attack behavior that allowed the derivation of five sub-datasets to suit the application of the Federated Learning approaches. The DDoS-FLNN proposed a hybrid Long Short-Term Memory + convolution neural network (LSTM + CNN) classification DDoS attack model by combining both models, with LSTM capturing temporal patterns from sequential data and CNN capturing spatial patterns for input data. The proposed hybrid LSTM + CNN aims to achieve optimal accuracy in DDoS attack detection.

Keywords Attacks · DDoS mitigation · Cyber security · Deep learning

1 Introduction

An attack is an action where information is tried to be stolen, altered, defected, destroyed, disclosed, displaced or deleted without proper authorization. It can be felt by both an individual as well as an organization. Attacks may also be passive active and targeted brandjacking within or outside of [1]. Botnet clickjacking phishing spamming DDoS spamming. A Distributed Denial of Service attack is formally described as a concerted and distributed type of attack on the services’ availability

A. Alazzawi (✉) · O. Salam
Diyala University, Diyala, Iraq
e-mail: dr.abdulbasit@uodiyala.edu.iq

O. Salam
e-mail: scicompms222316@uodiyala.edu.iq

at a host server (this could be an application server, storage server, database server or DNS server) or network resource directly but in reality indirectly through the use of a network of compromised systems that are spread throughout the Internet known as botnets [2]. DDoS assaults have developed over time since their inception. DDoS assaults provide a significant challenge to researchers and pose a security risk to cloud computing environments. Hackers employ various channels to attack multiple cloud resources, applications or networks. In current sophisticated methods, the danger is minimized of missing any target. Distributed denial of service attacks are devastating to cloud servers. Although Gligor coined the term “Denial of Service” (DoS) in an OS context, it has subsequently become widely used. Distributed Denial of Service (DDoS) attacks use multiple computers to target a victim. Digital denial of service attacks poses a significant threat to the security of networks. We describe the history of machine learning and how it applies to our system. One kind of artificial intelligence is known as machine learning; it employs algorithms to “learn” new things on their own.

Most businesses and jobs are seriously affected by machine learning. Many of the key milestones have been achieved in machine learning.

In this research, deep learning approaches are implemented for detection and mitigation of DDoS attacks.

2 Literature Review

The latest DDoS attack detection trends will be briefly discussed. Joshi et al. [3]: DDoS assaults may have a higher effect on cloud computing infrastructure shared by millions of users than on single-tenanted designs. A cloud traceback model was evaluated for DDoS assaults utilizing a back propagation neural network and found to be effective.

Lonea et al. [4]: This work evaluates our quantitative solution’s experimental findings. The trials are done in a Eucalyptus open-source private cloud architecture with three nodes running virtual machines-based IDS and the front-end server installing and configuring MySQL and alert monitoring graphical interfaces. After DDoS assaults on the VM-based IDS, we analyze all alarms. Wang et al. [5]: This study addresses this critical topic. DDoS attack protection is assessed on a corporate network using both technologies. If the defensive architecture is right, SDN can prevent DDoS attacks. The flexible network monitoring and unique control framework of our DDoS attack mitigation architecture detect attacks and respond quickly and individually. We present a visual model-based assault detection system for the novel dataset shift architecture. Simulations show that our architecture can efficiently address new network paradigm security concerns and that our attack detection system can reveal various threats using real-world network data. Wani et al. [6] presents the Tor Hammers was used to attack own cloud and Intrusion Detection System was used to Bhardwaj et al. [7]: The novel approach of this study involves the classification of network traffic as benign or DDoS attack traffic using Deep Neural Network (DNN)

and well-posed stacking sparse Auto-Encoder (AE) for feature learning. Setting settings by ways specified optimizes AE and DNN for DDoS detection.

Bhardwaj et al. [7]: The novel design of this study distinguishes network traffic as benign or DDoS attack traffic by employing Deep Neural Net-work (DNN) and well-posed stacking sparse Auto Encoder (AE) for feature learning. Setting using specified ways optimizes AE and DNN for DDoS detection. In reducing overfitting, explosion plus vanishing gradients, and network size, this paper recommends cutting the reconstruction error.

3 DDOS Attacks and Detection Methodology

Cloud services are becoming more popular, but data, resource, and service security and availability remain major challenges. The threat of DDoS assaults is not new and is a critical security issue with continuing research [8].

3.1 Denial of Service Attack

Launch Methods and DDoS Intend

DDoS attacks prevent legitimate users from accessing network resources. OSI model explains DDoS assaults. Domain name system attacks on the protocol layer typically target layers 3 and 4, whereas attacks on the application layer often target layer 7 [9].

There are two major approaches to start DDoS attacks online launch a vulnerability attack by sending the target faulty packets. As for the second tactic, an assailant might employ one of these [10]: Use up bandwidth, router processing power, or network resources to disconnect genuine users. The assaults are network/transport-level floods Flooding attacks.

The Attacker's Incentives

Different reasons motivate DDoS attackers. Attackers' motives can be analyzed to halt and reply [11].

Economic/financial gain: A key concern of organizations, usually done by dissatisfied people with inferior technical capabilities.

Intellectual Challenge: Young hacking enthusiasts aim to demonstrate their ability to experiment and initiate assaults.

Cyber warfare: Politically motivated attackers target crucial parts of another country.

3.2 *Intrusion Detection Methodologies*

Methods for detecting intrusions are presented here. They fall into three primary categories: Stateful Protocol Analysis (SPA), Signature-based Detection (SD), and Anomaly-based Detection (AD). We'll compare the three detection methods' pros and cons [12].

Signature-based Detection (SD): Intrusion detection systems that rely on signatures keep track of events and look for patterns that correspond to recognized threat signatures. An attack signature details the specific actions needed to execute an attack and the order in which they must be performed [13]. Plus, it can only identify attacks that have their signatures already saved in a database. Keeping the signatures up-to-date is essential for efficiency. The discovery of new threats against your hosts occurs at the same rate as the release of new threats, necessitating signature changes. When applied solely to the fixed behavioral pattern, this strategy yields good result [14].

Anomaly-based Detection (AD): Many researchers are interested in the AD because of its ability to detect new assaults. Defining the behavior of the network is the basis for detection. There has been no deviation from the expected behavior of the network. After that, it is either accepted or an anomaly detection event is triggered [1]. Depending on their needs, network managers can either teach or learn the expected behavior of their networks. AD's main benefit is that it can detect novel attacks that don't yet have signatures by looking for deviations from regular traffic patterns [15].

Stateful Protocol Analysis: By matching requests and replies, e.g., the SPA shows that IDS is aware of and able to track protocol states. While AD uses pre-loaded profiles that are specific to networks or hosts, SPA uses generic profiles that vendors create for each protocols to dictate how those protocols should and should not be used [16]. Understanding and responding to attacks are made easier with SPA's significant features. In order for the SPA to work as an intrusion detection tool, the protocol models must be clear and behave appropriately. A protocol's accuracy drops when it is proprietary, has vague specifications, or when a vendor's implementation goes against the grain [17].

3.3 *Long Short-Term Memory Neural Network (LSTM)*

Many recurrent neural networks (RNNs) use LSTMs to learn long-term dependencies [18]. Both RNN and LSTM networks comprise neural networks with the same structure [19]. LSTM differs from RNN in that it includes a memory block integrated into its basic unit. The LSTM memory blocks, or cells, are responsible for memory. Three gate techniques Forget, Input, and Output control the cells [18]. A forget gate removes unneeded data from the cell state. The input gate is responsible for updating the cell's state. The output gate manages the extraction and display of valuable cell status information [8].

4 Proposed Model

The suggested model incorporates LSTM and CNN into a seven-layer hybrid architecture. This work demonstrated the great quality of identifying DDoS attacks by combining CNN with LSTM. We suggest a combination of models after discovering that two methods one based on entropy and the other on a covariance matrix are essentially the same in that they both detect DDoS attacks by gauging increased dependence in the data as shown in Fig. 1.

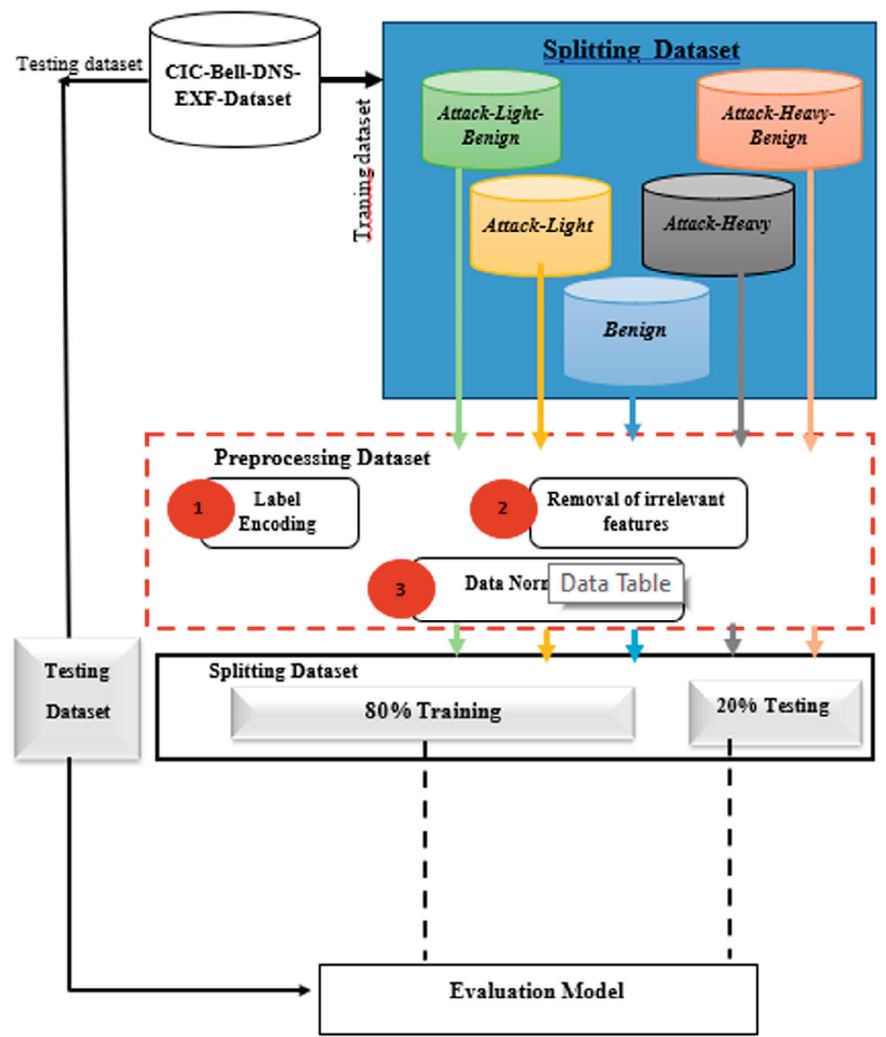


Fig. 1 Proposal model structure

4.1 Dataset

For this study, we have chosen to use the CICIDS dataset. The most recent assaults and comprehensive attributes make this dataset ideal for our research assignment [20]. There are both good and bad examples of network traffic in the sample. There are 84 characterized variables in the labeled dataset. This sample's last attribute is its class label, which indicates its maliciousness. It was CICFlowMeter-V3 that extracted the characteristics [7]. In its output, CICFlowMeter-V3 includes the following columns: Flow ID (1), Source IP (2), Destination IP (4), Source Port (3), Destination Port (5), Protocol (6), Timestamp (7), AND Label (84). Features 8–83 are derived from statistical calculations of the aforementioned fundamental characteristics. Those are the most prominent features. The following are examples of important traits that fall under this category [7]:

Anonymity: For reasons of user privacy, several databases omit the payload data. Because of this elimination, detection procedures are impaired. However, the payload is unaltered in the CICIDS record.

Attack Diversity: Vectors for distributed denial of service attacks are evolving at a swift pace. Heartbleed, the Apple SSL library vulnerability, and Shellshock are just a few of the most recent examples of the many types of attacks included in this dataset. They range from brute force and DOS attacks to browser-based and DNS-based assaults, as well as backdoors. You can find every single conversation between the victim network and the attacking network in the dataset. This data is useful because it describes exchanges that take place inside the internal LAN [21].

Complete Network Configuration: To create a realistic setup for recording the actual attack traces, it comprises a large number of modems, firewalls, switches, routers, and personal computers running a range of operating systems like Windows, Ubuntu, and Macintosh.

Complete Traffic: The Attack-Network contains actual attacks, the Victim-Network contains 12 distinct machines, and the User-Profiling-Agent-Network contains all the required traffic [22].

Feature Set: It generates network traffic, applies feature extraction apps to it, and then returns a CSV file containing a dataset of 80+ network flow features [22].

Metadata: Information regarding the dataset, possible attack scenarios, and the correct documentation of the network settings and operating systems used by the attacker and victim machines are all part of it.

Heterogeneity: During the attack's execution, it records the victim machines' system calls, memory dumps, and heterogeneous network traffic from network devices. Building a reliable detection mechanism is aided by this [23].

Labeling: Reliable and accurate analysis is made possible by its labeled and informational nature [23].

4.2 Dataset Preprocessing

Improving the input data quality and, by extension, the output quality and mining algorithm efficiency is the primary goal of preprocessing. The three pretreatment steps that have been executed are label encoding, feature deletion, and normalization [22].

Label encoding: The NSL-KDD dataset uses a single hot encoding to transform categorical features into numerical values. With a single hot encoding, NSL-three KDD's category features protocol type, service, and flag are transformed into numerical values [14]. The high volume of distinct values in the FlowID column necessitated its hash encoding conversion in CICIDS 2021. The categorical data in the columns are transformed using the hash function in hash encoding. The component parameter fixes a new dimension, and it uses it to convert values. There are a total of 18 parts, which can stand for 218 distinct values (or 262,144 distinct integers) [24].

Removal of irrelevant features: In order to make algorithms run more efficiently, we delete the values of characteristics that are invalid or non-informative, like NaN and infinite [25].

Normalization: A unique set of upper and lower bounds is associated with each dataset feature. When all the values are normalized to the range of [0, 1], the classifier becomes more efficient. To ensure that the attribute values are within the range of [0, 1], min-max normalization is employed [26].

4.3 Distributed Denial of Service Attacks

Denial of service (DDoS) attacks pose a significant threat to the safety of IoT networks. The goal of this kind of assault is to deplete the target's resources by flooding its network with traffic from a large number of compromised nodes. Damage to the infrastructure, disruption of services, and inability of authorized users to access related services are all results of this [27]. Distributed denial of service attacks uses two distinct methods: reflection and amplification. The attacker sends packets to several destinations using the target's IP address as the originating address, utilizing the reflection technique. In contrast, the attacker floods the target's system with packets using the amplification approach [8].

4.4 Detection Model Based on Convolutional Neural Network Approach

CNN algorithms are used by the suggested model to detect Ddos mitigation attacks. The "Long Short-Term Memory" (LSTM) models are a form of Convolutional Neural

Network (CNN) model that the proposed system utilizes to classify Ddos mitigation queries. Due to its efficient processing, LSTM is used in detection Ddos mitigation attacks.

Multiple layers in the LSTM model collaborate to learn and categories data sequences. The LSTM process for DDoS mitigation attack detection is shown in Algorithm (1), with parameters the MLST model being saved in H6 file format as the output.

Algorithm: LSTM Training

Input: Dataset 2D, and input dimensions

Output: Trained Model.

Model start.

Step1: Initialize the Parameters.

For the embedding vectors, set the output dim = 100.

Enter the activation function value: "ReLU".

With the Optimizer set to "Adam,"

Insert 150 as the value of hidden units.

Step2: Embedding (input_ dimension, output dimension) is the embedding layer.

Step3: the bidirectional layer using LSTM (150) and sequences set return to True.

Step4: The Layer Dropout 0.2 Step3-3: for each file in files list do.

Step5: Long Short-Term Memory (LSTM) (100).

Step6: Activation Function and Dropout Layer (512).

Step7: get the Dense (2, sigmoid) output.

Step8: Model Return.

Algorithm Steps in (1) describe how to develop a CNN model using LSTM. First, set the output dimension, activation function, optimizer, and hidden units. Step 2 creates a matrix by applying an embedding layer that represents each input word with a vector of output dim. A dropout layer is introduced in step 4 after adding a 150-hidden-unit bidirectional LSTM layer in step 3 to avoid overfitting. Another 100-hidden-unit LSTM layer is added in step 5, followed by a dropout layer in step 6. Finally, steps 7 and 8 return the model as a single entity by identifying the input sequence as binary via a dense output layer with two units and a sigmoid activation function. The architecture of the Hybrid LSTM + CNN model is depicted in Fig. 2.

4.5 Experimental Environment

Our used Visual Code 2019 with Python 3.9 to build our LSTM model. In a Windows 10 setting, that make use of TensorFlow, which comprises the libraries Panda, Scikit-learn, NumPy, and Kera's. The system requirements include an i7 processor, 4 GB of RAM, and a 512 GB hard drive.

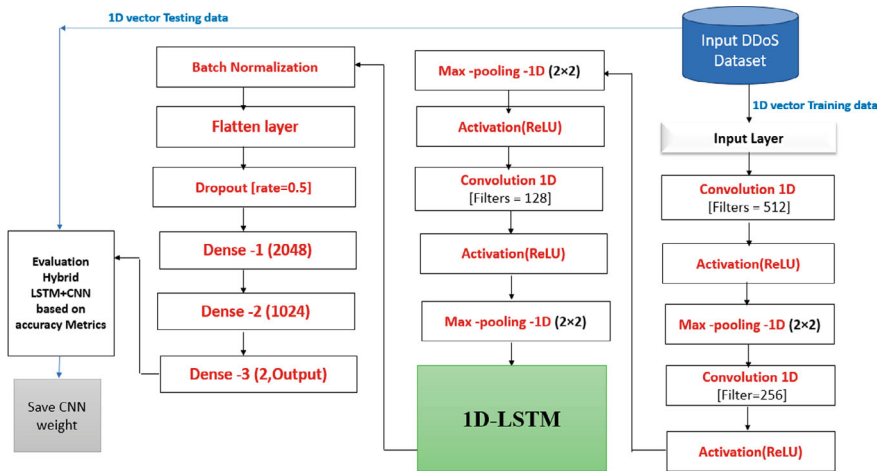


Fig. 2 Architecture of the Hybrid LSTM + CNN

5 Results and Discussion

In this part, after that, we compare the LSTM model to others and give the evaluation of these outcomes using the primary metrics [20].

5.1 Results

In this section, we have a look at the hyper parameters that, according to Table 1, need to be specified in order to prevent overfitting. The LSTM model is tri-layered; the bottom layer has 64 neurons and the highest layer has 78; the activation function used by both levels is the same ReLU. Eight neurons employ the Softmax activation function in the third layer. A nonlinear activation function, like the ones shown above, can improve upon the speed and accuracy of a linear activation function. In order to measure the discrepancy between the predicted and actual values of the output, a loss function was computed. We update the variables and enhance the model results by using optimizer Adam to calculate the loss gradients, which reduces the loss function as shown in Table 1.

5.2 Used Metrics

A total of three measures, including the accuracy scale, the loss scale, and the loss scale, were utilized in this study in order to evaluate the performance of the model

Table 1 Result CNN + LSTM

Model	Type attack	Loss model	Accuracy mode	Accuracy train	F1_score_train
Mode CNN_LSTM	Attack heavy	0.76	0.953	0.95	0.92
Mode CNN_LSTM	Attack heavy benign	0.60	0.99	0.99	0.99
Mode CNN_LSTM	Attack light	0.76	0.95	0.95	0.92
Mode CNN_LSTM	Attack light benign	0.65	0.99	0.99	0.99
Mode CNN_LSTM	Benign	0.65	0.99	0.99	0.99

[28]. A representation of the capacity to accurately identify data as (1) is what we mean when we talk about accuracy, Eq. (2) for the metric precision, Eq. (3) for the metric recall, Eq. (4) for the metric *F1* Score.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN}} \quad (3)$$

$$F1 \text{ Score} = \frac{2\text{TP} * \text{TN}}{2\text{TP} + 2\text{FN}} \quad (4)$$

What we mean by “accuracy” is the percentage of correctly identified samples. The association between accuracy and false alarm rate (FAR) is inversely proportional. Figure 3 displays the accuracy measurement during the testing and training stages; a lower false alarm rate corresponds to better accuracy. A visual depiction of the classification process’s performance and accuracy, the confusion matrix shows true and false positive values, provides insight into the model’s faults, and suggests ways to fix them.

Figure 4 shows the results of CNN, where the highest accuracy in this model is 99%. And Fig. 5 gives the results of LSTM, where the accuracy in this model is 95%.

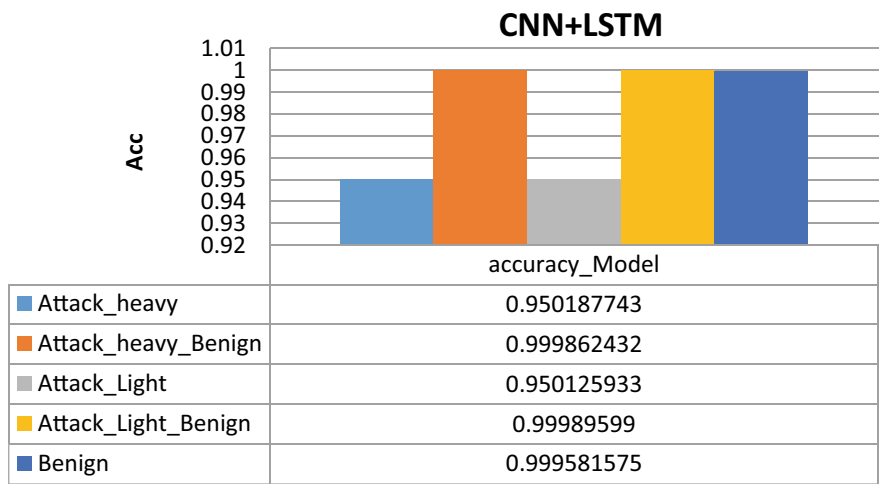


Fig. 3 Accuracy hybrid model

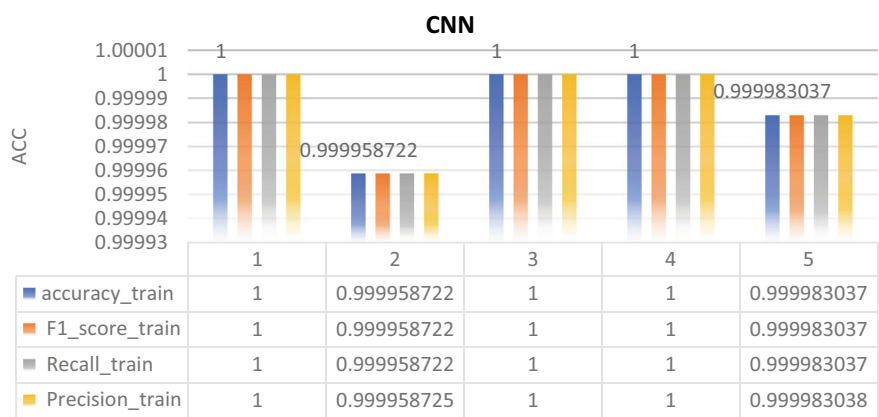


Fig. 4 Accuracy CNN model

5.3 Comparative Analysis

This section focuses on the ways in which our LSTM model differs from other models. Table 2 gives that in order to identify bot attacks that might compromise financial transactions, the model suggested in makes use of the CSE-CICIDS 2018 dataset [20, 29]. By constructing an ensemble and evaluating the classifiers’ detection abilities, Seth et al. [30] proposed a method for recognizing various assault types. The authors Rios et al. [31] proposed a method for detecting cyber-denial of service attacks in telecommunication networks using a broad learning system (BLS) that produces high performance with reduced training time.

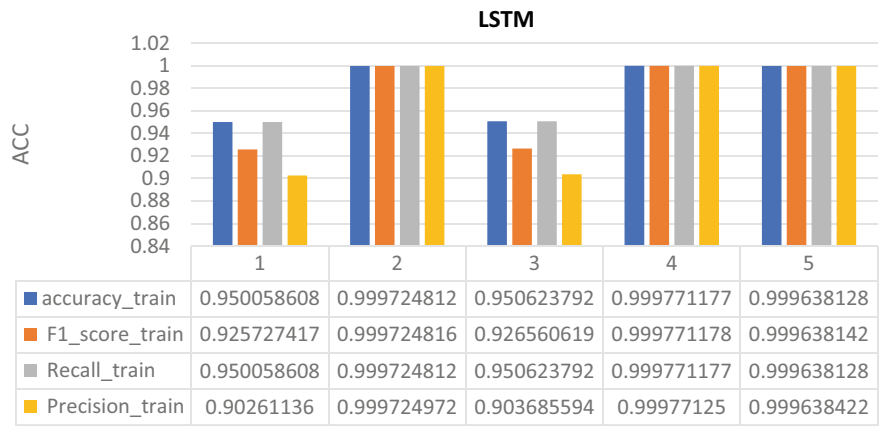


Fig. 5 Accuracy LSTM model

Table 2 Evaluation of LSTM model in relation to alternative approaches

Research	DL and ML	Dataset	Accuracy (%)
Lin et al. [29]	LSTM	CICIDS 2018	96.2
Seth et al. [30]	LSTM	CICIDS 2018	97.5
Rios et al. [31]	LSTM	CICIDS 2018	97.46
The proposed LSTM	LSTM	CICIDS 2021	99

Table 2 gives the accuracy comparison of the proposed model with various state-of-the-art techniques. Previous research has identified several effective deep learning algorithms, including ensemble, hybrid, and semi-supervised methods. The study attained the highest result and had an accuracy rate of 99% when compared to others. Experiments reveal that combining CNN and LSTM deep learning algorithms yields effective DDoS attack detection, based on their design. Additionally, the learning functions and parameters enhance the proposed model’s accuracy (magical effect). The hybridization of CNN as a feature extractor and LSTM as a predictor yields higher accuracy than using each separately. Additionally, the proposed model displays the highest accuracy in detecting DDoS assaults compared to earlier work on the NSL-KDD dataset. The study discovered that the proposed model was more widely used than DL and classic ML methods.

6 Conclusion

In this paper, an approach that uses entropy and covariance matrices to identify a viable alternate hybrid scheme that can withstand distributed denial of service assaults. At the network and host levels, we want to implement a new strategy based on a thorough hybrid detection technique. There is an urgent need for a complete solution as the performance of many existing DDoS detection systems is subpar and the number of DDoS attacks is increasing at an exponential rate. We anticipate that our proposed approach with two checkpoints will be an improved alternative solution to the problem of risk mitigation, leading to a much better outcome. Within the system is suggested by means of deep learning algorithms. The CSE-CICIDS2018 actual dataset was used to train a neural network that could identify intrusions in data flows using the LSTM approach. Although the model's detection accuracy was 99%, which is respectable, it faces obstacles such as the CICIDS2021 dataset's imbalance and its size, the former of which might lead to an error in accuracy computation. Problems with the LSTM model's architecture aside, regarding the expansion of the nodes and the establishment of connections between the various levels. Using techniques to determine the most pertinent features that enable detection approach, we want to improve accuracy, decrease error, and expedite the training process going forward.

References

1. Virupaksa KB, Asundi M, Channal K, Shatter P, Patil S, Narayan DG (2020) Distributed denial of service (DDoS) attacks detection system for OpenStack-based private cloud. *Procedia Comput Sci* 167:2297–2307
2. Vellinga S, Karthikeyan P, Vinoth Kumar V (2021) Detection of distributed denial of service attack in cloud computing using the optimization-based deep networks. *J Exp Theory. Artif Intel* 33(3):405–424
3. Joshi B, Vijayan AS, Joshi BK (2012) Securing cloud computing environment against DDoS attacks. In: 2012 international conference on computer communication and informatics, pp 1–5
4. Lonea AM, Popescu DE, Prostean O, Tanfield H (2013) Evaluation of experiments on detecting distributed denial of service (DDoS) attacks in Eucalyptus private cloud. In: *Soft Computing Applications: Proceedings of the 5th International Workshop Soft Computing Applications (SOFA)*, pp 367–379 (2013)
5. Wang B, Zheng Y, Lou W, Hou YT (2015) DDoS attack protection in the era of cloud computing and software-defined networking. *Comput Netw* 81:308–319. <https://doi.org/10.1016/j.comnet.2015.02.026>
6. Wani AR, Rana QP, Saxena U, Pandey N (2019) Analysis and detection of DDoS attacks on cloud computing environment using machine learning techniques. In: 2019 Amity international conference on artificial intelligence (AICAI), pp 870–875
7. Bhardwaj A, Mangat V, Vig R (2020) Hyperband tuned deep neural network with well posed stacked sparse autoencoder for detection of DDoS attacks in cloud. *IEEE Access* 8:181916–181929
8. Saxena R, Dey S (2020) DDoS attack prevention using collaborative approach for cloud computing. *Cluster Comput* 23:1329–1344

9. Bagyalakshmi C, Sundareswaran ES (2020) DDoS attack classification on cloud environment using machine learning techniques with different feature selection methods. *Int J Adv Trends Comput Sci Eng* 9(5)
10. Amjad A, Alyas T, Farooq U, Tariq MA (2019) Detection and mitigation of DDoS attack in cloud computing using machine learning algorithm. *EAI Endorsed Trans Scalable Inf Syst* 6(23):e7–e7
11. Toh A (2022) Azure DDoS protection—2021 Q3 and Q4 DDoS attack trends. Microsoft
12. Olufunke AS, Sapere W, Intrusion detection in cloud network environment: a multilayer perceptron model
13. Ahmed S et al (2023) Effective and efficient DDoS attack detection using deep learning algorithm, multi-layer perceptron. *Futur. Internet* 15(2):76
14. Cil AE, Yildiz K, Buldu A (2021) Detection of DDoS attacks with feed forward based deep neural network model. *Expert Syst Appl* 169:114520
15. Sahi A, Lai D, Li Y, Diykh M (2017) An efficient DDoS TCP flood attack detection and prevention system in a cloud environment. *IEEE Access* 5:6036–6048
16. Samsu Aliar AA, Agora Moorthy M (2022) An automated detection of DDoS attack in cloud using optimized weighted fused features and hybrid DBN-GRU architecture. *Cybern Syst*, 1–42
17. Ouhssini M, Afdel K, Agherrabi E, Akouhar M, Abarda A (2024) DeepDefend: a comprehensive framework for DDoS attack detection and prevention in cloud computing. *J King Saud Univ Inf Sci*, 101938
18. Issa ASA, Albayrak Z (2023) DDoS attack intrusion detection system based on hybridization of CNN and LSTM. *Acta Polytech. Hungarica* 20(2):105–123. <https://doi.org/10.12700/APH.20.2.2023.2.6>
19. Miller CC (2006) Chapter two. *Forbes* 178(13):72–74. <https://doi.org/10.7829/j.ctv1w0xcfk.5>
20. Farhan BI, Jasim AD (2022) Performance analysis of intrusion detection for deep learning model based on CSE-CIC-IDS2018 dataset. *Indones J Electr Eng Comput Sci* 26(2):1165–1172. <https://doi.org/10.11591/ijeecs.v26.i2.pp1165-1172>
21. Ibrahim I, Kurnaz S (2021) A new distributed denial-of-service detection system in cloud environment by using deep belief networks. *Commun Fac Sci Univ Ankara Ser A2-A3 Phys Sci Eng* 63(1):17–24. <https://doi.org/10.33769/aupse.697067>
22. Garcia S, Valeros V (2023) Towards a better labeling process for network security datasets. [Online]. Available: <http://arxiv.org/abs/2305.01337>
23. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
24. Arel I, Rose DC, Karnowski TP (2010) Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *IEEE Comput Intell Mag* 5(4):13–18
25. Somani G, Gaur MS, Sangha D, Conti M, Buyya R (2017) DDoS attacks in cloud computing: Issues, taxonomy, and future directions. *Comput Commun* 107:30–48
26. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *International conference on machine learning*, pp 448–456
27. Balasubramaniam S et al (2023) Optimization enabled deep learning-based DDoS attack detection in cloud computing. *Int J Intell Syst* 2023
28. Mishra A, Gupta BB, Perkovich D, Peñalvo FJG, Hsu C-H (2021) Classification based machine learning for detection of DDoS attack in cloud computing. In: *2021 IEEE international conference on consumer electronics (ICCE)*, pp 1–4
29. Lin P, Ye K, Xu C-Z (2019) Dynamic network anomaly detection system by using deep learning techniques. In: *Cloud computing—CLOUD 2019: 12th international conference, held as part of the services conference federation, SCF 2019, San Diego, CA, USA, 25–30 June 2019, Proceedings 12*, pp 161–176
30. Seth S, Chahal KK, Singh G (2021) A novel ensemble framework for an intelligent intrusion detection system. *IEEE Access* 9:138451–138467
31. Rios ALG, Li Z, Bekshentayeva K, Trajković L (2020) Detection of denial-of-service attacks in communication networks. In: *2020 IEEE international symposium on circuits and systems (ISCAS)*, pp 1–5

PlantPath: Deep Learning-Based Plant Pathogen Detection



Kooturu Kanishk Reddy , Sujal Mangesh Limje ,
T. Y. J. Naga Malleswari , and S. Ushasukhanya 

Abstract Agriculture stands as the bedrock of worldwide food production and livelihoods, serving as a crucial pillar for sustaining economies and nourishing populations globally. While it plays a vital role in ensuring global food security, agriculture grapples with the persistent challenge posed by plant pathogens. PlantPath focuses on advancing pathogen detection methodologies within the realm of agricultural technology, particularly in the domain of precision agriculture and pathogen management. Prior research has highlighted several challenges in agricultural pathogen detection, including limited dataset diversity, species-specific models, and the propensity for overfitting in deep learning architectures. The identified challenges emphasize the necessity for inventive strategies that tackle dataset diversity, generalize across various crop species, and incorporate rigorous regularization techniques to improve model performance and dependability. Introducing a pioneering solution, PlantPath, our innovative approach adopts deep learning principles for agricultural pathogen detection. It classifies images into five distinct categories: bacteria, fungi, pests, virus, and healthy plants, with approximately 8000 images per class spanning twenty crop species within our dataset. PlantPath utilizes convolutional neural networks, notably the ResNet50 architecture, for universal pathogen detection. PlantPath detects subtle patterns on affected plant leaves, allowing for precise classification according to the specific pathogen present. It achieves an impressive accuracy of 98.27% in detecting and classifying plant pathogens, surpassing existing methods in terms of time, cost, precision and universality. Compared to previous approaches reliant on manual observation or specialized equipment, PlantPath offers automated and precise identification of pathogens from images of plant leaves.

K. K. Reddy (✉)

Computer Science and Engineering, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

S. M. Limje · T. Y. J. Naga Malleswari · S. Ushasukhanya

Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

e-mail: nagamalt@srmist.edu.in

S. Ushasukhanya

e-mail: ushasuks@srmist.edu.in

Keywords Plant pathology · Pathogen detection · Deep learning · Convolutional neural networks · Image classification

1 Introduction

Agricultural productivity plays a vital role in ensuring food security as well as economic stability. One of the primary challenges affecting the sector is the prevalence of plant pathogens, which jeopardize agricultural productivity and food production. Phytopathogens are a major concern to global crop productivity [1]. Consequently, for proper disease management and mitigation strategies, there is a need for early detection and exact identification of plant pathogens. Traditional detection methods, including visual examination or specialized instruments, could be expensive, time-consuming, and energy-intensive. In addition to showing initial signs of resistance, such methods can also be inaccurate and unreliable. The interplay between plant immunity and environment, roots, and plant microbiomes, and the true infection of pathogens, ensures the outcomes of infection, infection severity, and plant productivity [2].

In this paper, we propose PlantPath to overcome the mentioned difficulties: a novel deep learning-based method for plant pathogen identification in agriculture. PlantPath uses a convolutional neural network architecture called ResNet50 to comprehend images of plant leaflets and uncover the delicate patterns linked with distinct pathogens. A deep learning approach meant to be able to automatically identify significant features in images, prevent subtle symptoms that might trip up traditional image-processing algorithms [3]. Crop leaf pathogen classification involves four steps: pre-processing, feature extraction, deep learning methods and classification [4].

Every pixel of a leaf is expected to share similarity between one another in terms of any observable feature or property such as colour, intensity, texture, and so forth [5]. When implemented, the automated detection functionality of PlantPath empowers the farmer to detect pathogens early to ensure timely and successful intervention. Moreover, PlantPath is useful for many crop plants. While existing solutions may be tailored for a single plant type, PlantPath proves to be a versatile solution useful in a wide variety of agricultural settings. PlantPath's zero-cost implementation model that ensures accessibility and affordability is set to democratize access to advanced disease detection technology to farmers of all levels. Determining the pathogens behind a growing disease and doing so quickly is essential to disease management through proper selection of approaches. This enables the implementation of highly targeted and localized remediation approaches, which can result in decreased pesticide usage and ultimately promote more sustainable agricultural practices [6]. In conclusion, PlantPath is an ambitious innovation that represents an extraordinary leap in the field of agricultural disease management technology. As such, it is an affordable, scalable, and available solution for plant pathogen detection. The paper

proceeds with a literature survey, followed by the proposed solution, implementation details, presentation of results and discussions, and concludes with a summary.

2 Literature Survey

In this literature survey, we explore the patterns that the current plant pathogen detection solutions take and the problems they experience, especially due to low-scale farming and remote areas. We look into the problems of high prices, the dearth of affordability of the expert knowledge, the too long need for actionable insights to develop through manual observations, and the unreliability of the traditional disease detection methodology. Our model makes use of deep learning models, based on Convolutional Neural Networks which reveals several robust regularization techniques during training to avoid overfitting [7]. A broader review by Yao et al. [8] explores such deep learning techniques in detail, outlining their growing effectiveness in plant disease classification tasks.

This ensures that the model is able to function well when exposed to new data and reliably performs in actual scenarios. The model also adequately addresses the dataset's diversity problems since it integrates several sources into a single one. Unlike some recent developments, which require comprehensive knowledge of the pathogen to operate effectively [9], our model offers unique solution. Determining the pathogenic patterns is not plant-specific, which means the model can be confidently incorporated into the process of agricultural management. Disease detection systems face many obstacles overall, making them hard to employ successfully, especially for small-scale farmers and remote areas. High costs and need for expertise prevent farm owners from introducing them in their enterprise [10].

Effective disease management is further complicated by limited access to expert knowledge [11]. However, our system offers real-time feedback and insights and equips farmers with the necessary tools to take action. Observation is a popular technique but is slow and convoluted to use daily [12]. In comparison, our system specializes in accelerated decision-making with high accuracy. The existing methods have limitations when it comes to high enough accuracy in early detection for all crops [13]. Consequently, the data presented by our solution offers exclusive value as it is most likely to determine the ideal intervention moment. For instance, dataset diversity and unavailable labelled training data are prime challenges encountered in UAV-based crop disease detection [14]. Therefore, we address the problem through a hybrid deep learning system and data augmentation techniques [15]. Most of the early automatic recognition methods of plant leaves are based on manual features and classifiers, which limits their application in all kinds of complex scenarios [16]. In contrast, our system uses advanced deep learning techniques to suitably address the stated problems and create solutions that are accessible, reliable, and cost-friendly for plant diseases. Atila et al. used the EfficientNet deep learning model for plant leaf disease classification in 2021, and our project goes beyond mere classification. We

classify the plant pathogens and offer appropriate guidance and recommendations towards managing and controlling the disease [17].

Jiang et al. developed a real-time detection system for apple leaf diseases using an advanced convolutional neural network approach, offering promising applications in agricultural monitoring and disease management. While Jiang et al. focused specifically on apple leaf diseases, our project encompasses a wider range of plant pathogens, offering a more comprehensive solution for agricultural monitoring [18]. While Sagan et al. explored UAV-based thermal imaging for vegetation monitoring and disease detection, our project, PlantPath, improves over their work in several ways. Thermal images are a good tool for stress detection in plants, and deep neural networks can predict diseases, although it may lack resolution for most targets. Our approach combines high-resolution RGB images with state-of-the-art deep learning to detect subtle visual markers that correlate with a wide variety of plant pathogens [19]. Arnal Barbedo focused on identifying plant diseases from individual lesions and spots by using deep learning; our project goes one-step further to offer a complete solution for plant pathogen identification [20]. Although Indrakumari et al. provided prospective methodologies and provided thoughtful insights into future work, our project also overcomes probable drawbacks to introduce an implemented solution for real-time plant pathogen identification in the field [21].

In this manner, we have demonstrated a diverse and comprehensive range of pre-processing and dataset diversity stages supporting the model's reliable performance and generalization cross cutting various plant species and, as such, have met limitations present in the focus by Kaur et al. on rice diseases [22]. In conclusion, this literature survey highlights the imminent challenges associated with current plant pathogen detection's methodologies and emphasizes the urgent need for unique approaches to address these dynamics. By analysing and reviewing the growing body of research, new trends, recent innovations, and modern-day developments, we have established the following critical areas for improvement. Our designed system is set to bridge these gaps through innovative deep learning mechanisms that avail a user-friendly, low-cost, easy accessibility, and a quick-response tool for disease prevention to farmers.

3 Proposed Solution

The proposed solution for enhancing pathogen detection in agriculture involves several key components aimed at overcoming existing challenges and advancing the field of precision agriculture.

In Fig. 1, a wide variety of images depicting crop species and types of pathogens have been gathered to tackle the issue of dataset diversity. Collaboration, with institutions, research organizations, and farmers, has enabled the collection of top notch image datasets covering geographic regions and crop types.

To prepare the image data for pathogen detection, several pre-processing steps are taken. Initially, the images are standardized to ensure dimensions across the dataset.

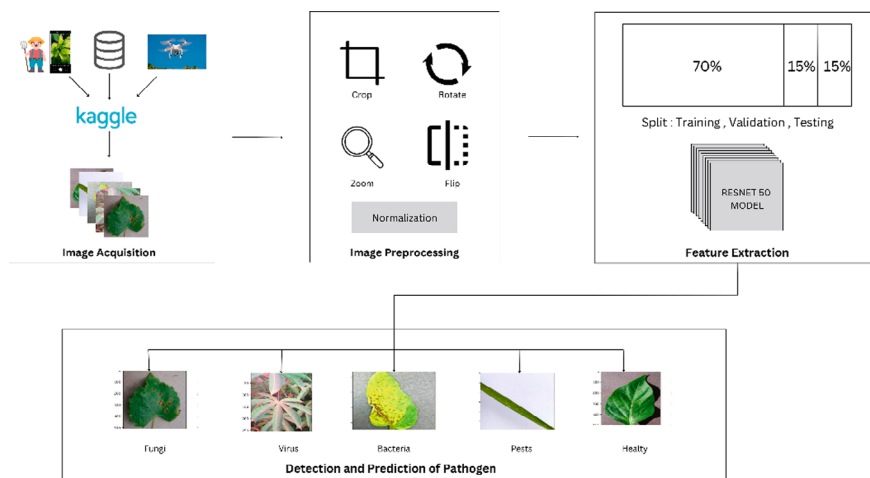


Fig. 1 Architecture diagram

Subsequently data augmentation techniques such as rotation, flipping, and zooming are applied to enrich the datasets variability and robustness. Additionally efforts are made to improve image quality and normalize values to optimize model performance by ensuring that input data conform to a format.

The dataset is divided into training, validation, and test sets following a ratio. Typically assigning 70% for training 15% for validation and 15%, for testing. This partitioning ensures model training, validation and evaluation processes. For model development purposes, Deep Learning (DL) techniques employing Convolutional Neural Networks (CNNs) are utilized with the architecture being Res-Net50.

After this stage, the model is trained on the training dataset of pathogens with fine-tuning of its settings to improve performance in this type of task. Instead of relying on species-specific models, the development of a universal pathogen detection model capable of generalizing across different crop species is essential. This involves training the model on a diverse dataset covering multiple crops and pathogens, allowing it to recognize common patterns of disease across various plant species.

Utilizing learning structures such as CNNs helps to detect complex patterns and features from plant images. The number of layers and skip connections of ResNet50 makes the process of identifying pathogens from various images highly efficient. Incorporating the pathogen detection system into precision farming frameworks enables real time monitoring and control of plant health. This allows for actions such as targeted pesticide use or monitoring crops to prevent pathogen spread and enhance crop productivity.

During the evaluation phase with the dataset, the trained models performance is assessed on a validation dataset to verify its accuracy, precision, recall, and $F1$ -score metrics. These metrics offer insights, into how the model can classify types of pathogens accurately. Furthermore, a confusion matrix is created to visualize how

well the model classifies pathogen categories providing an overview of its strengths and weaknesses when distinguishing between classes.

The last step that is during the testing and validation phase, the final trained model is used to test using a dataset to determine if the model is able to generalize and perform well in real-world scenarios. The primary aim of this step is to verify the model’s performance on unseen data; in other words, to examine how the model is able to recognize and classify plant diseases from new images or the extent to which it can use that knowledge outside of the training data.

4 Implementation

The workflow, as given in Fig. 2, begins with the initiation phase, where the user uploads an image to the system. Upon upload, the system proceeds to the image file validation step, where it checks the file type to ensure its validity. If the image file passes the validation, it moves on to the pre-processing stage. Here, the valid image undergoes pre-processing procedures, which includes resizing, format conversion, and normalization.

These pre-processing steps are essential to ensure that the input image is appropriately formatted and standardized for further analysis. Following pre-processing, the pre-processed image is utilized for pathogen prediction using a trained ResNet50 model. The ResNet50 model has been specifically trained to classify plant pathogens from images, leveraging deep learning techniques. Utilizing the pre-processed image, the model predicts the presence of pathogens and identifies the specific type present in the image which may be bacteria, fungi, virus, pests, or healthy.

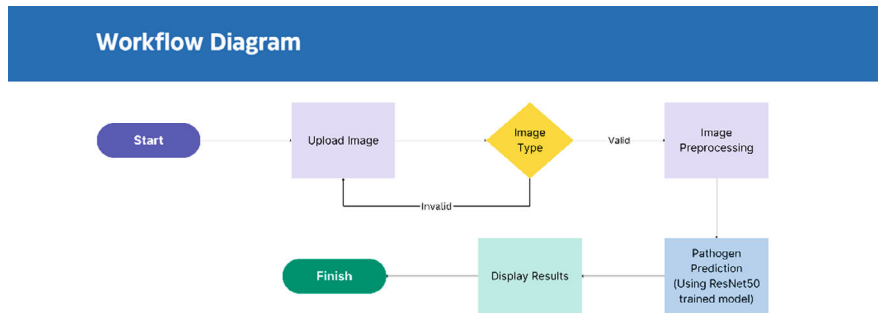


Fig. 2 Workflow diagram

4.1 Dataset Preparation

- The first step in implementing our PlantPath work involved collecting a diverse dataset of images depicting plant leaves affected by various pathogens, as well as images of healthy plant leaves for comparison.
- We gathered images from multiple sources, including online repositories, agricultural research databases, and field surveys, to ensure dataset diversity and representativeness.
- Each image was meticulously labelled according to its corresponding pathogen class (bacteria, fungi, pests, virus) or designated as healthy, facilitating supervised learning during model training.
- While current innovations focus only on a few species of plants, our project overcomes some potential drawbacks by focusing on broader spectrum of plant species [23].

4.2 Pre-processing and Data Augmentation

- Prior to training, the dataset underwent pre-processing steps to standardize image sizes, enhance contrast, and normalize pixel values to improve model performance.
- Data augmentation techniques, such as rotation, scaling, and flipping, were applied to increase dataset variability and improve model generalization.
- Augmented images were generated to artificially increase the dataset size and expose the model to a wider range of variations in leaf appearance and pathogen manifestations.

4.3 Model Architecture Selection

- For our PlantPath project, we opted to utilize the ResNet50 architecture, a well-established and widely used deep learning architecture known for its effectiveness in image classification tasks.
- ResNet50 offers a balance between model depth and computational efficiency, making it suitable for processing high-resolution images while minimizing computational resources.
- The architecture's skip connections enable efficient gradient flow during training, alleviating the vanishing gradient problem and facilitating the training of deeper networks.

4.4 Model Training and Evaluation

- The ResNet50 model was trained using the prepared dataset, with images fed into the network to learn to classify them into their respective pathogen classes or healthy category.
- During training, we utilized techniques such as stochastic gradient descent (SGD) optimization, categorical cross-entropy loss, and learning rate scheduling to fine-tune model parameters and optimize performance.
- The model underwent evaluation using a distinct validation dataset to gauge its performance metrics, including accuracy, precision, recall, and *F1*-score across various pathogen classes. Hyperparameters were adjusted iteratively based on validation performance to ensure optimal model generalization and robustness.

5 Results and Discussion

The results section, Fig. 3, contains a detailed explanation of the PlantPath interface, which gives an idea about the process of its functionality and utility in detection of agricultural pathogen. The interface is a product of the Streamlit framework and designed to help farmers to upload the images of the leaves of the plants for analysis. Once the image is uploaded, the quality of image is enhanced and the features, which are most relevant and are extracted for the analysis. The enhancing steps include image normalization, resizing, and noise reduction, and this is necessary to get best inputs for the deep learning model. As soon as the enhancing is done, the PlantPath deep learning model, which is built using convolutional neural networks and ResNet50 architecture get the image for the prediction posed. The model analyses the image and predicts the type of pathogen affecting the plant. Output categorizes the detected pathogen into bacteria, fungi, pests, virus, or healthy plants. After detection, it suggests preventative and remedial measures, along with precautions to prevent further spread, based on historical data. The app enhances user interaction by offering detailed information about the pathogen, including symptoms, lifecycle, and control methods, with links to related articles and blogs.

The graphical representation of the training and validation process of the PlantPath deep learning model is presented in Fig. 4. This figure visually reflects the loss and accuracy metrics during multiple epochs, allowing us to judge about the model's performance and its convergence for the training process. Thus, the graph displays the real-time changes of both training and validation loss function as the number of epochs rises.

Training loss is the penalty for the error difference between the output from the model and real labels when the data is fed into the model for learning. Validation loss is computed analogously but on the validation dataset unseen by the model during the training phase. Once again, while increasing the number of epochs, the loss value goes down, which indicates that the model better learns how to make the predictions.

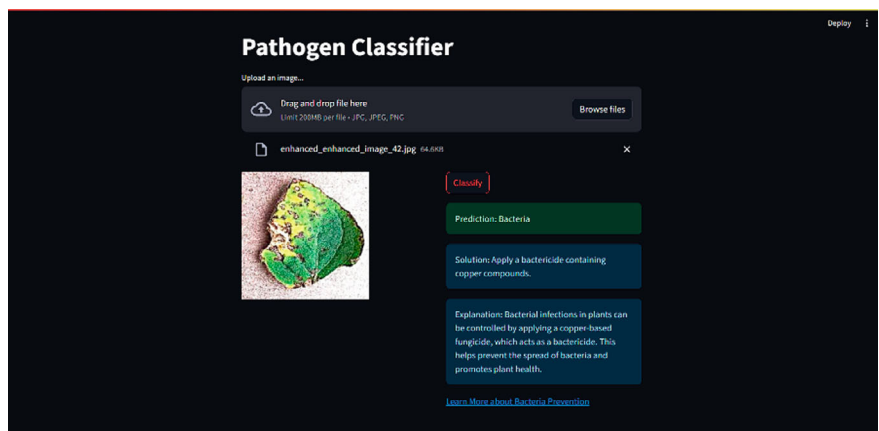


Fig. 3 Prediction of pathogen

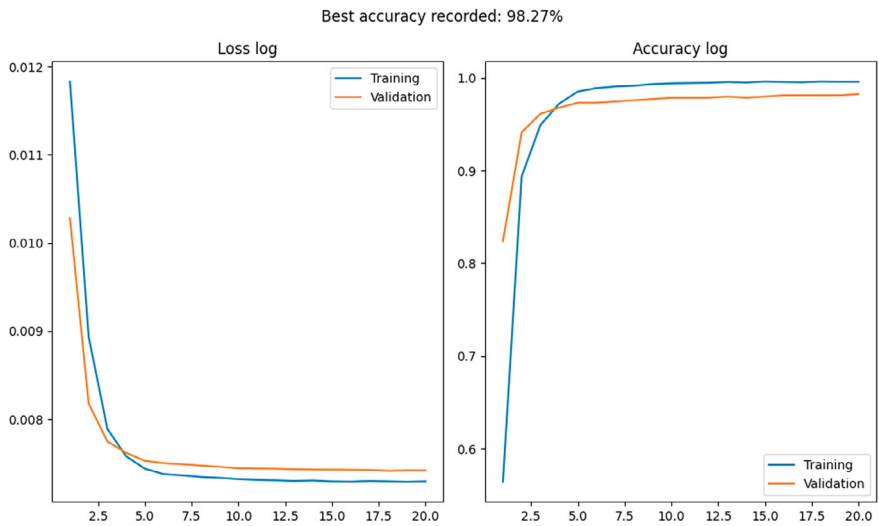


Fig. 4 Loss accuracy graph

The graph contains the plot of training and validation accuracy development during the training. When referring to the training accuracy, it measures the percentage of correctly predicted labels in a training set, while for the validation accuracy, it estimates percentage to predict the unseen validation set. Therefore, the more accurate the model will be, the more images it will classify correctly. The location of the best model accuracy is depicted in the points of this graph. The figure highlights the highest accuracy of the model was 98.27%. It is the end of the training, as the model has been trained to classify many different pathogens very accurately.

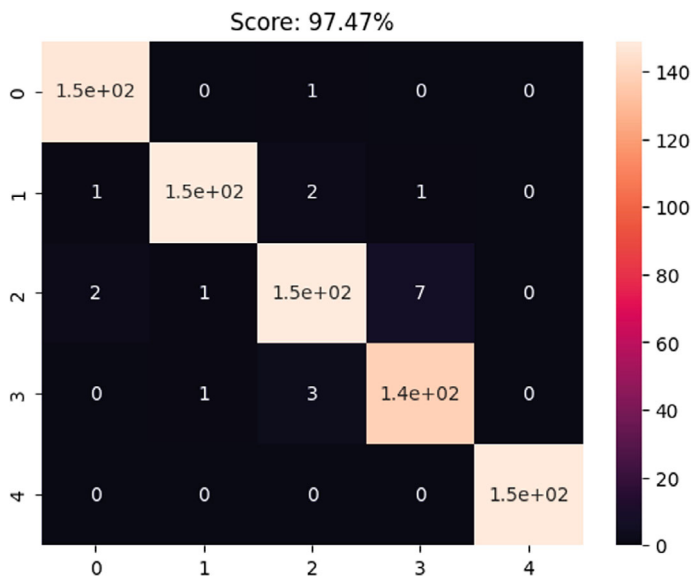


Fig. 5 Confusion matrix

The confusion matrix in Fig. 5 provides a more detailed look at the performance of the PlantPath model of each class in the dataset.

As seen in Fig. 6 the “Bacteria” category, the model achieves a precision of 97%, signifying the accuracy of instances classified as “Bacteria.” Additionally, with a recall score of 99%, the model accurately identifies the majority of actual “Bacteria” instances. The *F1*-score, representing the balance between precision and recall, stands at 8%, underlining the overall effectiveness of the model in accurately classifying “Bacteria” instances. Similarly, the “Fungi” class exhibits perfect precision, recall, and *F1*-score, signifying flawless performance in classifying instances of fungal pathogens. The “Healthy” class demonstrates slightly lower precision and recall at 98% and 96%, respectively, resulting in an *F1*-score of 97%. This indicates a high level of accuracy in identifying healthy plant instances, albeit with a slightly higher rate of false negatives. The “Pests” class displays strong precision and perfect recall, yielding an *F1*-score of 98%, indicating excellent performance in detecting instances of pest-related pathogens. Lastly, the “Virus” class achieves perfect precision and a slightly lower recall of 97%, resulting in an *F1*-score of 99%. This suggests highly accurate classification of viral pathogen instances, with a slightly higher rate of false negatives.

While PlantPath offers promising advancements in agricultural pathogen detection, there are several limitations to consider:

- The effectiveness of deep learning models like PlantPath heavily relies on the quality of input images. Poor-quality images, such as those with low resolution or excessive noise, may result in inaccurate predictions or misclassifications.

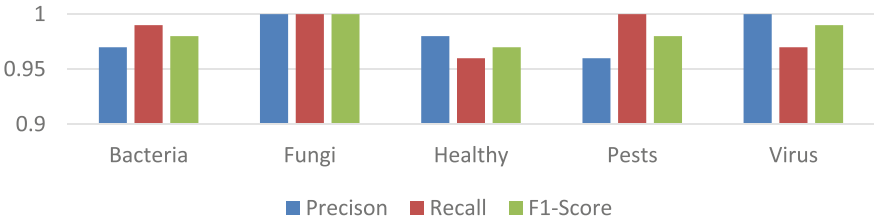


Fig. 6 Performance metrics

- PlantPath’s requirement for smartphone access may exclude farmers from using the technology in regions where smartphones are not widely available.
- Ensuring PlantPath’s usability and accessibility for end-users, especially those with limited technological literacy, may be challenging and require tailored user interfaces and training programmes.

6 Conclusion

In conclusion, agriculture remains the bedrock of global food security and economic stability, yet the persistent threat of plant pathogens continues to challenge productivity and sustainability. Our contributions address these challenges by introducing PlantPath, a pioneering deep learning-based system designed for agricultural pathogen detection. PlantPath represents a significant advancement in the field, offering a novel approach to early detection and precise identification of plant pathogens across a diverse array of crop species. By leveraging convolutional neural networks and the Res-Net50 architecture, PlantPath achieves remarkable accuracy in detecting and classifying pathogens, surpassing existing methods in terms of precision and universality. Looking ahead, future work in this area could focus on expanding the dataset to include additional crop species and pathogen variants, further enhancing the model’s robustness and generalization capabilities. Additionally, research efforts could explore the integration of real-time monitoring and decision-support functionalities into the PlantPath system, enabling proactive disease management strategies and enhancing agricultural sustainability on a global scale. Furthermore, extending the accessibility and usability of the PlantPath model through the development of a comprehensive mobile application could empower farmers with on-the-go pathogen detection capabilities and actionable insights.

References

1. Ayaz M, Li C-H, Ali Q, Zhao W, Chi Y-K, Shafiq M, Ali F, Yu X-Y, Yu Q, Zhao J-T et al (2023) Bacterial and fungal biocontrol agents for plant disease protection: journey from lab to field, current status, challenges, and global perspectives. *Molecules* 28:6735. <https://doi.org/10.3390/molecules28186735>
2. Singh BK, Delgado-Baquerizo M, Egidi E et al (2023) Climate change impacts on plant pathogens, food security and paths forward. *Nat Rev Microbiol* 21:640–656. <https://doi.org/10.1038/s41579-023-00900-7>
3. Shoaib M, Shah B, El-Sappagh S, Ali A, Ullah A, Alenezi F, Gechev T, Hussain T, Ali F (2023) An advanced deep learning models-based plant disease detection: a review of recent research. *Front Plant Sci* 14:1158933. <https://doi.org/10.3389/fpls.2023.1158933>
4. Tanwar S, Singh J (2023) ResNext50 based convolution neural network-long short term memory model for plant disease classification. *Multimed Tools Appl* 82:29527–29545. <https://doi.org/10.1007/s11042-023-14851-x>
5. Ahad MT, Li Y, Song B, Bhuiyan T (2023) Comparison of CNN-based deep learning architectures for rice diseases classification. *Artif Intell Agric* 9:22–35. <https://doi.org/10.1016/j.aiia.2023.07.001>. ISSN 2589-7217
6. Marc V, Sam C, Hans R (2023) Current and emerging trends in techniques for plant pathogen detection. *Front Plant Sci* 14. <https://doi.org/10.3389/fpls.2023.1120968>
7. Khalid MM, Karan O (2023) Deep learning for plant disease detection. *Int J Math Stat Comput Sci* 2:75–84. <https://doi.org/10.59543/ijmscs.v2i.8343>
8. Yao J, Tran SN, Sawyer S et al (2023) Machine learning for leaf disease classification: data, techniques and applications. *Artif Intell Rev* 56(Suppl 3):3571–3616. <https://doi.org/10.1007/s10462-023-10610-4>
9. Sajitha P, Andrushia AD, Anand N, Naser MZ (2024) A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. *J Ind Inf, In Press*. <https://doi.org/10.1016/j.jii.2024.100572>
10. Khirade ASD (2015) Plant disease detection using image processing. In: 2015 international conference on computing communication control and automation, Pune, India, 26–27 Feb 2015, pp 768–771. <https://doi.org/10.1109/ICCUBE.A.2015.153>
11. Khanchi S, Potgieter J (2020) Image-based plant disease identification by deep learning meta-architectures. *Plants (Basel)* 9(11):1451. <https://doi.org/10.3390/plants9111451>
12. Vishnoi VK, Kumar K, Kumar B (2021) Plant disease detection using computational intelligence and image processing. *J Plant Dis Prot* 128:19–53. <https://doi.org/10.1007/s41348-020-00368-0>
13. Nguyen C, Sagan V, Maimaitiyiming M, Maimaitijiang M, Bhadra S, Kwasniewski MT (2021) Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors* 21(3):742. <https://doi.org/10.3390/s21030742>
14. Shahi TB, Xu C-Y, Neupane A, Guo W (2023) Recent advances in crop disease detection using UAV and deep learning techniques. *Remote Sens* 15(9):2450. <https://doi.org/10.3390/rs15092450>
15. Sahu K, Minz S (2023) Adaptive segmentation with intelligent ResNet and LSTM–DNN for plant leaf multi-disease classification model. *Sens Imaging* 24:22. <https://doi.org/10.1007/s11220-023-00428-3>
16. Duan Y, Han Z, Tang Z (2023) A lightweight plant disease recognition network based on ResNet. *Appl Comput Eng* 5:583–592. <https://doi.org/10.54254/2755-2721/5/20230651>
17. Atila Ü, Uçar M, Akyol K, Uçar B E (2021) Plant leaf disease classification using EfficientNet deep learning model. *Eco Inform* 61:101182
18. Jiang P, Chen Y, Liu B, He D, Liang C (2019) Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural net- works. *IEEE Access* 7:59069–59080

19. Sagan V, Maimaitijiang M, Sidike P, Eblimit K, Peterson KT, Hartling S, Esposito F, Khanal K, Necomb M, Pauli D et al (2019) UAV-Based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 p, FLIR Vue Pro R 640, and thermomap cameras. *Remote Sens* 11:330
20. Arnal Barbedo JG (2019) Plant disease identification from individual lesions and spots using deep learning. *Biosyst Eng* 180(2016):96–107
21. Indrakumari R, Poongodi T, Khaitan S, Sagar S, Balamurugan B (2021) A review on plant diseases recognition through deep learning. In: Balas VB, Mishra BK, Kumar R (2021) *Handbook of deep learning in biomedical engineering*. Academic Press, pp 219–244. ISBN 9780128230145. <https://doi.org/10.1016/B978-0-12-823014-5.00009-0>
22. Kaur A, Guleria K, Trivedi NK (2024) A deep learning-based model for biotic rice leaf disease detection. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-024-18730-x>
23. Shanmugam M, Shanmugam S, Rithika P, Shafiq I, Nandita S, Gomathi R (2022) Detection and classification of paddy leaf diseases using deep learning (CNN). https://doi.org/10.1007/978-3-031-11633-9_6

Sentiment Extraction and Mining of Textual Content in Social Networking Space Using Context-Based Intelligent Retrieval System



Aastha Sahani, Oindrella Chatterjee, Tiansheng Yang, Bharati Rathore, V. Sanjay, Urvash Shende, and Sushruta Mishra

Abstract Opinion mining, also known as sentiment analysis, stands as a straightforward yet highly advantageous real-time application of machine learning. It aids in discerning the emotional tone or opinion expressed by the speaker or writer in a given text. For example, whether it is a review, email, or any written document, a sentiment analyzer proves valuable in gauging the underlying thoughts and tone of the content. This application of sentiment analysis finds utility in decision-making tools, platforms reliant on user reviews, and various other contexts. Sentiment analysis is the task of identifying and extracting the emotional tone and attitude of a text, such as a tweet, a review, or a comment. It can be useful for understanding customer feedback, social media trends, or market sentiments. To perform sentiment analysis, you can use natural language processing techniques, such as tokenization, stemming, lemmatization, and vectorization, to preprocess and transform the text data. Then, you can apply machine learning models, such as logistic regression, naive Bayes, or neural networks, to classify the text into positive, negative, or neutral categories. We can also use sentiment analysis to generate insights and visualizations, such as word clouds, polarity scores, or sentiment distributions.

Keywords Sentiment analysis · Opinion mining · Neural networks · Word clouds · Polarity scores · Sentiment distributions · NLTK · Social media platforms

A. Sahani · O. Chatterjee · U. Shende · S. Mishra (✉)
Kalinga Institute of Industrial Technology, Deemed to Be University, Bhubaneswar, India
e-mail: sushruta.mishrafcs@kiit.ac.in

T. Yang · B. Rathore
University of South Wales, Pontypridd, UK
e-mail: tiansheng.yang1@southwales.ac.uk

B. Rathore
e-mail: bharati.rathore@southwales.ac.uk

V. Sanjay
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

1 Introduction

Sentiment analysis involves evaluating unprocessed text to generate objective quantitative outcomes using natural language processing, machine learning, and other data analysis methodologies. Its primary purpose is to identify positive or negative sentiment within text, commonly employed by businesses to assess their brand reputation among customers. There exist diverse categories of sentiment analysis, where models focus on aspects such as emotions, urgency, intentions, and polarity. Fine-grained sentiment analysis, Emotion detection, Aspect-based sentiment analysis, and Multilingual sentiment analysis are among the most prevalent types. The significance of sentiment analysis lies in its ability to provide insights into the emotions and sentiments expressed by customers. Companies leverage sentiment analysis by scrutinizing social media interactions and reviews to make well-informed decisions based on customer sentiments. The types of social media platforms are social networks like Facebook, blogs/microblogs like Quora, discussion forums, social bookmarks like pinterest, wiki projects, social news, media sharing like Instagram.

Natural Language Processing (NLP) plays a crucial role in examining emotions and sentiments. NLTK is a leading platform for building Python programs to work with human language data. The paper thoroughly explores sentiment analysis in social media, proposing an innovative system to enhance understanding. Through practical tests, the system proves effective in understanding nuanced sentiments from diverse text. Yet, challenges like varied data and algorithm bias persist, requiring creative solutions. In summary, the paper discusses the benefits and hurdles of the proposed system, urging readers to embrace sentiment analysis for better social media insights.

The main contributions of a research paper on opinion mining of texts in social media can vary, but here are three to four potential key contributions:

- Introduction of an innovative sentiment analysis algorithm specifically tailored for social media texts. The algorithm may address challenges such as handling informal language, understanding contextual cues, and adapting to the dynamic nature of social media communication.
- Construction and release of a unique and well-annotated dataset tailored for opinion mining in social media. The dataset may encompass diverse domains, user demographics, and linguistic styles commonly found on social platforms, facilitating more accurate and domain-specific sentiment analysis.
- Development of a context-aware sentiment classification model that considers additional contextual factors, such as user engagement metrics, temporal dynamics, and entity mentions.
- Investigation into the temporal aspect of sentiments in social media, offering insights into how opinions evolve over different time periods.

2 Literature Review

A study has investigated sentiment analysis on microblogging platforms like Twitter and Facebook, addressing challenges posed by the brevity and informality of posts. They proposed a novel system architecture to automatically analyze sentiments in short messages and utilized manually annotated data from Twitter to extract opinionated messages, filter out non-opinions, and determine sentiment directions. The experimental results showcase the system's effectiveness in real microblogging applications, overcoming challenges posed by the unique characteristics of microblog posts [1]. Authors have addressed the challenge of detecting emotion in text by introducing EmotiNet, a novel knowledge base (KB) designed to represent affective reactions to real-life contexts. They populated and evaluated EmotiNet based on self-reported affective situations from the International Survey on Emotion Antecedents and Reactions (ISEAR) corpus. The method utilized Semantic Roles to cluster examples, extract triples, and enhance the model with resources like VerbOcean, ConceptNet, and SentiWordNet. Despite limitations in domain and knowledge quantity, EmotiNet proves to be a valuable semantic resource, capturing and storing the structure and semantics of real events. It effectively predicts emotional responses triggered by chains of actions [2]. Research has compared human-indicated terms (crowd lexicon) with deep learning approaches (Doc2Vec and fasttext) for sentiment analysis on tweets and Facebook comments about commercial products. They utilized token selection strategies in crowd lexicon creation, including majority voting and two annotator agreement, to assess the impact of these strategies on sentiment analysis models, addressing challenges in sentiment-related keyword identification. The results demonstrate the effectiveness of combining human knowledge with AI tools, emphasizing the power of collective intelligence in sentiment analysis within the context of big data. They also highlight the challenges of sentiment-related keyword identification in Facebook comments compared to tweets [3]. Authors have aimed to explore user preferences for photo vs. video posts through sentiment analysis in social media analytics. They utilized artificial intelligence techniques, including sentiment analysis and sentiment polarity classification, to analyze customer reactions on social media channels. The results highlight the significance of understanding user preferences for sustainable business development, emphasizing the potential of social selling through social media channels in the digital age [4].

Another study [5] aimed to explore opinion research's role in identifying societal trends and its application in policymaking. To simulate and predict communication patterns on specific policy issues, they utilized models of opinion formation through real-world online communication. This study investigates how online content from social networking resources provides decision-makers with information regarding constituent opinions, trends, and the feasibility of policy initiatives. Through the analysis of social media content and the modeling of opinion formation, they present building blocks for opinion mining, simulation, and trend analysis. They also discuss application scenarios and real-life examples, as well as practical implementation choices and future research directions. It was the purpose of this

study to explore the impact of social networks and Web 2.0 on the growth of sentiment analysis and opinion mining. They discussed the relevance of sentiment analysis in decision-making processes, where opinions significantly influence behavior. Researchers have delved deeply into the world of sentiment analysis, uncovering its diverse applications across various contexts. One key area of exploration revolves around how sentiment analysis influences consumer behavior. Studies have shed light on the significant impact of friends' and family's opinions, as well as reviews and surveys, on purchasing decisions. This highlights the crucial role sentiment analysis plays in understanding and potentially predicting consumer choices [6, 7]. Beyond consumer behavior, researchers have delved into the realm of machine learning to unlock powerful tools for sentiment analysis. Three main approaches have emerged: supervised learning, where data is meticulously labeled with its sentiment (positive, negative, or neutral); unsupervised learning, which utilizes unlabeled data and relies on algorithms to identify patterns; and semi-supervised learning, which combines both labeled and unlabeled data for increased efficiency. These advancements empower researchers and businesses to analyze vast amounts of data and gain valuable insights into public opinion [6]. Social media, a vibrant platform teeming with user-generated content, has become a prime focus for sentiment analysis research. One study employed a combination of information retrieval and statistical techniques to conduct a comprehensive analysis of existing research papers within this domain. This systematic review not only offered a snapshot of the current landscape but also identified valuable techniques for future exploration [7]. Another study ventured into the realm of nuclear energy, utilizing sentiment analysis on Twitter data to track public opinion. This innovative approach offered a cost-effective and timely method for understanding the evolving sentiments surrounding this complex topic, providing valuable insights for decision-making processes aimed at building national consensus [8]. Researchers have also explored the potential of sentiment analysis in analyzing textual and visual content on platforms like Twitter. One study proposed a novel algorithm that classifies tweets as positive or negative toward a specific subject. This method not only allows for tracking opinion changes over time but also incorporates visual elements like images to provide a more comprehensive understanding of sentiment. This advancement proves valuable in various contexts, from informing consumer decisions during online purchases to offering marketers deeper insights into public perception [9]. Moving beyond commercial applications, researchers have embraced the power of sentiment analysis in the field of historical preservation. One study focused on integrating text and multimedia opinion mining to assist archivists in selecting materials for community memory archives. This approach, which employs a rule-based method to address challenges like informal language and sarcasm, aims to structure information around semantic categories. This innovative method contributes to a deeper understanding of effective opinion mining within the dynamic social media landscape and offers valuable strategies for preserving valuable historical information for future generations [10]. Finally, a comprehensive review explored the burgeoning field of social opinion mining. This research delved into the various dimensions of user-generated content across diverse social media platforms and media formats, focusing on identifying aspects like subjectivity,

sentiment, emotions, and even the nuances of sarcasm and irony. The review further highlighted the versatility of social opinion mining, demonstrating its applicability across numerous fields, including marketing, advertising, and even health care. This comprehensive analysis, encompassing trends, tools, and technologies utilized up to 2018, serves as a valuable resource and paves the way for future research directions aimed at maximizing the impact of social opinion mining in diverse real-world applications [11, 12].

3 Proposed Methodology

A machine learning system can serve three main functions: Descriptive, where it explains what occurred using data; Predictive, where it forecasts future events based on data; and Prescriptive, where it offers recommendations for action based on the data [13–15]. Firstly, raw data is collected from surveys to make a relevant dataset that will help us to analyze texts in different social media platforms. Opinion mining involves analyzing and determining the sentiment expressed in texts, particularly in social media where opinions are often shared. The selection of attributes within a dataset for opinion mining relies on the particular objectives of the analysis. Here are some common dataset attributes used in opinion mining for texts in social media. Figure 1 shows the proposed methodology.

- 1. Text Content:
 - (a) The actual text or content of the social media posts, comments, or reviews.
 - (b) Understanding the language, context, and nuances in the text is crucial for sentiment analysis [16, 17].
- 2. Sentiment Label:

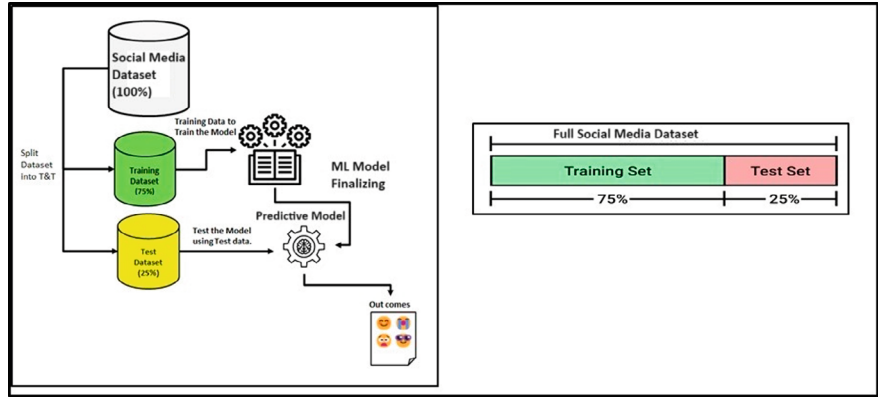
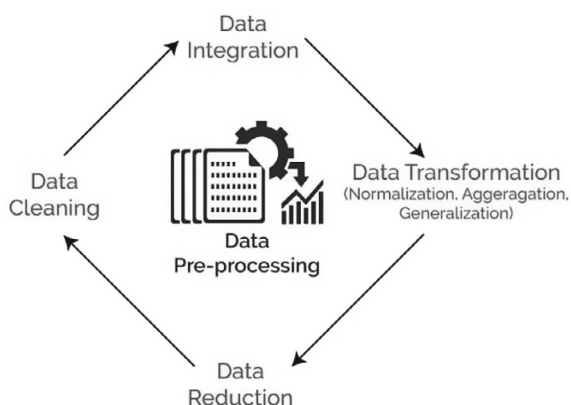


Fig. 1 Proposed model with data partitioning into training and testing subsets

- (a) The target variable signifies the sentiment conveyed in the text, including positive, negative, or neutral expressions.
 - (b) Binary sentiment labels (positive/negative) or multi-class labels (positive/neutral/negative) are common.
3. Author Information:
- (a) User information like username, account type, or profile information.
 - (b) The author's identity may influence the sentiment, and analyzing patterns related to different user types can provide valuable insights [18, 19].
4. Timestamp:
- (a) The date and time when the social media post was made.
 - (b) Temporal analysis helps understand how sentiments change over time, e.g., during specific events or product launches.
5. Engagement Metrics:
- (a) Number of likes, shares, comments, and other engagement metrics.
 - (b) High engagement on a post may indicate a strong sentiment and analyzing these metrics provides additional context.
6. Hashtags and Mentions:
- (a) Identification of hashtags and mentions in the text.
 - (b) Hashtags and mentions can provide context and are often used to express opinions or connect with specific communities.
7. Topic or Entity Mentioned:
- (a) Identification of topics, products, brands, or entities mentioned in the text.
 - (b) Analyzing sentiments related to specific topics helps in targeted opinion mining.
8. Emoticons and Emoji Usage:
- (a) Detection and analysis of emoticons and emojis used in the text.
 - (b) Emoticons and emojis often convey sentiments and add an emotional layer to the text.
9. Language and Dialect:
- (a) Language of the text and any specific dialects used.
 - (b) Different languages or variations in language can impact sentiment analysis accuracy [20].
10. Contextual Information:
- (a) Additional context surrounding the text, such as location, events, or external factors.
 - (b) Understanding the context helps in interpreting the sentiment more accurately.

Fig. 2 Major tasks of data pre-processing



1. User Ratings (if applicable):

- (a) If the dataset includes user ratings or scores, they can be used as labels for sentiment analysis.

When working with a dataset for opinion mining, it is essential to consider a combination of these attributes to gain a comprehensive understanding of the sentiments expressed in social media texts. Depending on the specific application, you may choose to include or exclude certain attributes. Secondly, The Collected data is pre-processed to remove ambiguities. This Purpose is served with the help of various techniques of Data Mining. The Data Pre-processing includes four major tasks as shown in Fig. 2:

- a. Data Cleaning.
- b. Data Integration.
- c. Data Reduction.
- d. Data Transformation and Discretization

The process of data cleaning comprises different methods of filtering the raw data. Some data cleaning tasks are filling in missing data, identifying the inconsistent data and removal of outliers. Filling out the missing data include two measures in which the use of central tendency (i.e., Use of Variance and Standard Deviation) has been used to fill out the data. In the dataset given, many tuples containing sparse values (i.e., No values) are replaced by a 'No' value implying that the text is either unaware or hasn't been facing the same. The inconsistencies have been detected by manual correction. The generated dataset now moves toward being integrated.

Another problem associated with Data cleaning is Data Discrepancy which includes Human Error in Data Entry or Instrumental Errors. Data Integration is a pre-processing technique that combines data from multiple sources into a coherent store. This considers various problems like Schema integration and object matching, Redundancy (unwanted attribute), Detecting and resolving data value conflicts [21, 22].

Clustering involves partitioning a population or data points into multiple groups, without prior knowledge of class labels, where data points within the same group exhibit greater similarity to each other compared to those in different groups. Lastly, after meticulous data pre-processing, the next pivotal step involves training ML models, particularly employing Clustering Algorithms, to decipher potential diseases based on prevalent symptoms in individuals. Clustering Algorithms excel in grouping data points with similar characteristics, making them instrumental in identifying patterns within symptom datasets. By leveraging these algorithms, the ML model can discern subtle connections and associations among symptoms, enabling it to categorize several words into clusters. The training process involves exposing the model to diverse datasets, allowing it to learn and adapt to various symptom patterns. This adaptive learning equips the model with the ability to generalize findings and make accurate predictions when presented with new, unseen cases. Ultimately, the integration of Clustering Algorithms in ML for opinion mining of texts in social media marks a significant stride toward more efficient, precise, and personalized analysis of texts. After Clustering is done, the model will be able to predict the text based on the emotion it has clustered for a particular word.

4 Result Analysis

The analysis is represented in a graphical format. This bar chart shown in Fig. 3 provides a comparative overview of the effectiveness of each methodology, visualized by bars of varying heights. The higher the bar, the more effective the methodology is considered in its respective task. This gives a quick visual understanding of which approaches stand out in terms of effectiveness.

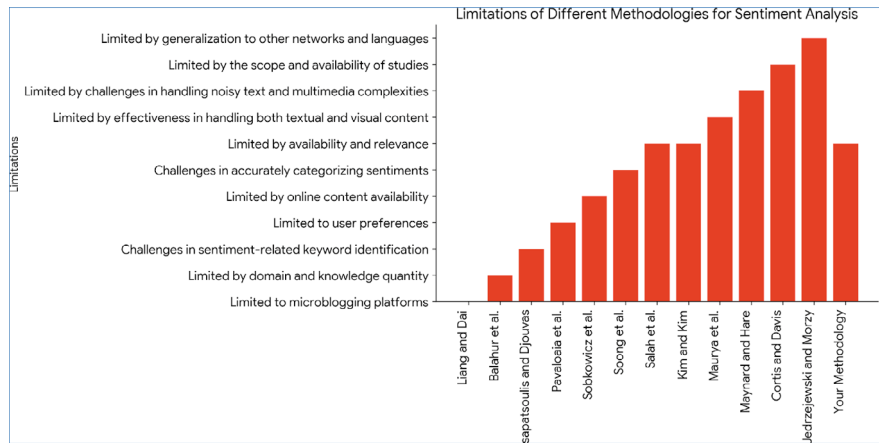


Fig. 3 Graphical representation of limitations of methodologies for sentiment analysis

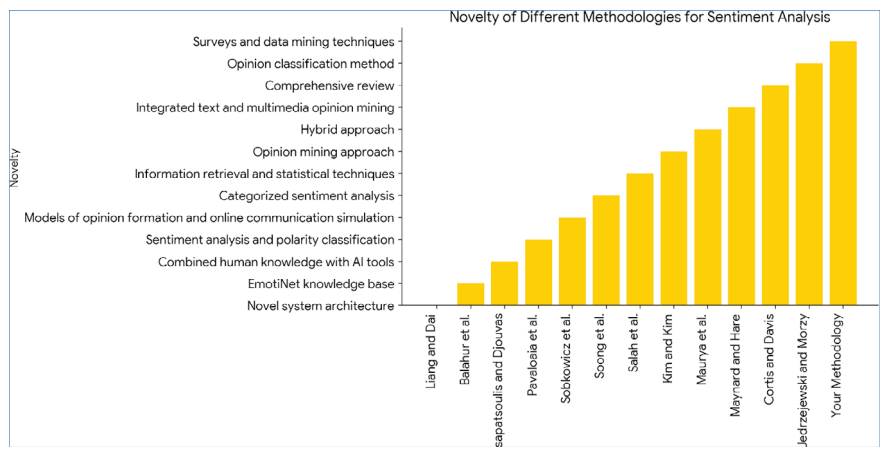


Fig. 4 Graphical representation of novelty of methodologies for sentiment analysis

Table 1 Performance outcome of various predictive models on opinion mining

Predictive models	Accuracy (%)	Specificity (%)	Sensitivity (%)
Random forest	92.9	90.4	89.2
Naive Bayes	78.6	76.6	74.1
Decision tree	94.4	92.9	91.5
Nearest neighbor	89.3	87.7	87.1
Multilayer perceptron	90.9	90.2	84.8
Support vector machine	82.1	80.2	78.6
Logistic regression	85.6	84.9	80.6

This horizontal bar chart depicted in Fig. 4 displays the applicability of each methodology across different domains or tasks. Each bar represents a methodology, and the stacked segments within each bar indicate the areas where it can be applied. This helps visualize the versatility and potential uses of each approach.

Table 1 highlights the evaluation outcome of several predictive models implemented in opinion mining applications. Among the models, decision tree exhibits the best performance with accuracy, specificity and sensitivity of 94.45%, 92.9% and 91.5% respectively.

5 Conclusion

In summary, the proposed methodology marks a significant advancement in sentiment analysis within the expansive domain of social media platforms. By meticulously orchestrating the intricate dance of data collection, pre-processing, and

analysis, this framework provides a robust foundation upon which businesses and researchers can unravel the complex tapestry of customer sentiments, preferences, and emerging trends. In essence, the proposed methodology transcends mere analysis; it serves as a beacon guiding businesses toward a deeper comprehension of their audience's desires and inclinations. As the digital landscape continues to morph and mutate at an unprecedented pace, the integration of machine learning techniques into sentiment analysis remains not just relevant but indispensable. Through this approach, businesses can forge stronger connections with their clientele, nurture brand loyalty, and navigate the turbulent waters of social media with confidence and finesse, thereby charting a course toward sustained prosperity in the dynamic realm of digital commerce and communication.

References

1. Liang P-W, Dai B-R (2013) Opinion mining on social media data. In: 2013 IEEE 14th international conference on mobile data management, Milan, Italy, pp 91–96. <https://doi.org/10.1109/MDM.2013.73>
2. Balahur A, Hermida JM, Montoyo A (2012) Building and exploiting EmotiNet, a knowledge base for emotion detection based on the appraisal theory model. *IEEE Trans Affective Comput* 3(1):88–101. <https://doi.org/10.1109/T-AFFC.2011.33>
3. Tsapatsoulis N, Djouvas C (2019) Opinion mining from social media short texts: does collective intelligence beat deep learning? *Front Robot AI* 5:138
4. Păvăloaia V-D, Teodor E-M, Fotache D, Danileț M (2019) Opinion mining on social media data: sentiment analysis of user preferences. *Sustainability* 11(16):4459
5. Sobkowicz P, Kaschesky M, Bouchard G (2012) Opinion mining in social media: modeling, simulating, and forecasting political opinions in the web. *Gov Inf Q* 29(4):470–479
6. Soong H-C, Jalil NBA, Kumar Ayyasamy R, Akbar R (2019) The essential of sentiment analysis and opinion mining in social media: introduction and survey of the recent approaches and techniques. In: 2019 IEEE 9th symposium on computer applications & industrial electronics (ISCAIE), Malaysia, pp 272–277. <https://doi.org/10.1109/ISCAIE.2019.8743799>
7. Salah Z et al (2019) A systematic review on opinion mining and sentiment analysis in social media. *Int J Bus Inf Syst* 31(4):530–554 (2019)
8. Kim DS, Kim J-W (2014) Public opinion mining on social media: a case study of Twitter opinion on nuclear power. *Adv Sci Technol Lett* 51:224–228
9. Maurya CG, Gore S, Rajput DS (2018) A use of social media for opinion mining: An overview (with the use of hybrid textual and visual sentiment ontology). In: *Proceedings of international conference on recent advancement on computer and communication: ICRAC 2017*. Springer, Singapore (2018)
10. Maynard D, Hare J (2015) Entity-based opinion mining from text and multimedia. *Adv Soc Media Anal*, 65–86
11. Cortis K, Davis B (2021) Over a decade of social opinion mining: a systematic review. *Artif Intell Rev* 54(7):4873–4965
12. Abd-Alhalem SM, Marie HS, El-Shafai W, Altameem T, Rathore RS, Hassan TM (2024) Cervical cancer classification based on a bilinear convolutional neural network approach and random projection. *Eng Appl Artif Intell* 127:107261
13. Hassan MM, Zaman S, Rahman MM, Bairagi AK, El-Shafai W, Rathore RS, Gupta D (2024) Efficient prediction of coronary artery disease using machine learning algorithms with feature selection techniques. *Comput Electr Eng* 115:109130

14. Malik P, Dureja A, Dureja A, Rathore RS, Malhotra N (2024) Enhancing intracranial hemorrhage diagnosis through deep learning models. *Procedia Comput Sci* 235:1664–1673
15. Patel AD, Jhaveri RH, Rathore DRS (2023) Healthcare on smart wheels-a review of intelligent vehicles assisted healthcare systems. In: 4th International conference on distributed sensing and intelligent systems (ICDSIS 2023), pp 83–89, Dubai, United Arab Emirates. <https://doi.org/10.1049/icp.2024.0468>
16. Duggal R, Pandya A, Rathore DRS, Kalra K, Sharma K, Gupta N (2023) Improved deep learning-based contactless biometric recognition using bracelet lines. In: 4th international conference on distributed sensing and intelligent systems (ICDSIS 2023). Institution of Engineering and Technology (IET), Dubai, United Arab Emirates, pp 567–574
17. Sahoo S, Mishra S, Brahma B, Barsocchi P, Bhoi AK (2024) SSO-CCNN: a correlation-based optimized deep CNN for brain tumor classification using sampled PGGAN. *Int J Comput Intell Syst* 17(1):1–18
18. Mishra S, Chaudhury P, Tripathy HK, Sahoo KS, Jhanjhi NZ, Hassan Elnour AA, Abdelmaboud A (2024) Enhancing health care through medical cognitive virtual agents. *Digital Health* 10:20552076241256732
19. Mishra S, Jena L, Mishra N, Chang HT (2024) PD-DETECTOR: a sustainable and computationally intelligent mobile application model for Parkinson's disease severity assessment. *Heliyon* 10(14)
20. Pranjal P, Mallick S, Paul A, Mishra S, Bhardwaj I, Albuquerque VHCD (2024) Soil crops and nutrients forecasting using random forest model. *AIP Conf Proc* 2919(1)
21. Mishra S, Chakraborty S, Sahoo KS, Bilal M (2023) Cogni-Sec: a secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system. *Internet of Things* 24:100978
22. Mishra S, Voley DR, Bohra N, Alfarhood S, Safran M (2023) A smart and sustainable framework for millet crop monitoring equipped with disease detection using enhanced predictive intelligence. *Alex Eng J* 83:298–306

Artificial Intelligence and Tissue Biomarkers: Benefits, Risks, and Prospects for Pathology



Mohamed Bobor Jalloh , Manbeer Singh Sidhu ,
and Vivek Kumar Garg 

Abstract Artificial intelligence (AI) and digitized Whole Slide Images are revolutionizing diagnostic precision and personalized medicine in pathology. Deep learning offers objective methodologies and improved predictive abilities, streamlining workflows and optimizing treatment strategies. Despite limitations in digitization infrastructure and interpretability, AI integration has the potential to improve pathologists' skills, standardize practices, and advance patient care, shaping the future of diagnostic medicine. In this review, we explore the fusion of artificial intelligence and tissue biomarkers in pathology, highlighting its significant impact on diagnostic, prognostic, and predictive functions. Spanning across various domains including liver, lung, stomach, and colon pathology, as well as dermatopathology. Finally, we discuss the future trajectory of AI in pathology.

Keywords AI · Whole slide images · Diagnostic precision · Deep learning · Pathology · Personalized medicine

1 Introduction

Clinical pathologists use subjective interpretation, semi-quantification, and incorporation of morphological designs, resulting in judgment variability. Cytology was the first to use microscopic images to measure pathology objectively. The 1990s introduction of scanners permitted the creation of digitized whole slide images (WSI), sparking interest in applying AI to pathology like radiology. Machine Learning (ML), which used engineered morphological features to distinguish benign and malignant tumors, was the first AI technique. The time-consuming annotation of features limited ML's ability to scale to other problems. Histopathology AI was revolutionized by deep

M. B. Jalloh · M. S. Sidhu (✉) · V. K. Garg
Department of Medical Laboratory Technology, UIAHS, Chandigarh University, Mohali, India
e-mail: manbeersingh14@gmail.com

V. K. Garg
e-mail: garg.vivek85@gmail.com

learning (DL), which learned directly from raw WSI data without feature engineering. Strongly supervised, unsupervised DL, or weakly supervised. Strongly supervised DL predicts classes with pathologist-labeled patches, weakly supervised DL labels entire WSIs with a class, and unsupervised DL learns without labels with limited applications. Digital Pathology (DP) includes whole slide image (WSI) technologies that improve pathology workflow. DP relies on AI-based tools to help pathologists identify, quantify, and integrate disease biomarkers for diagnosis, prognosis, and prediction. These tools require WSI-specific software for color normalization, focus quality assessment, standard tile extraction, object classification, region segmentation, and counting [1]. Figure 1 shows a flowchart of some among the most popular software used to these aims.

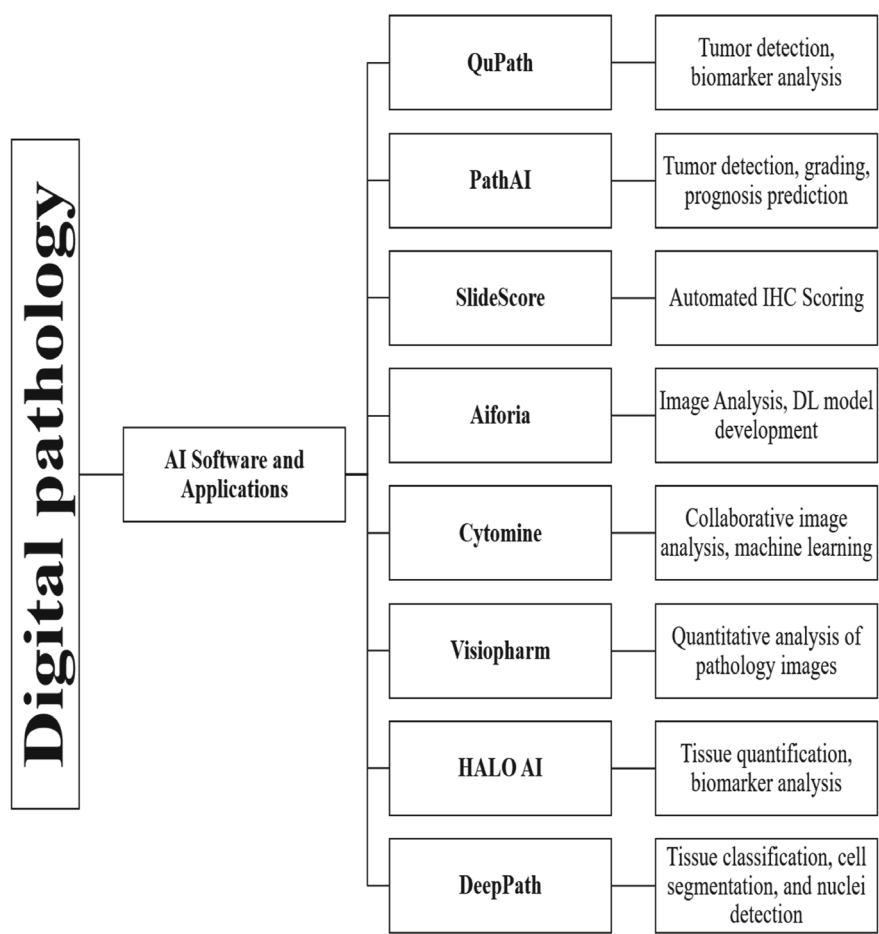


Fig. 1 Various AI softwares used for digital pathology

1.1 AI Terminologies Review

We assume that our reader is new to AI terminologies. Please see Table 1 for a concordance of common artificial intelligence terminologies and acronyms as well as specific terms mentioned in this manuscript.

2 The Integration of Artificial Intelligence with Tissue Biomarkers

In personalized medicine (PM), a biomarker is any information used to identify a subgroup within a larger population for diagnostic, prognostic, or predictive purposes. Pathology-specific tissue biomarkers contain tissue information. Traditional “classic” tissue biomarkers like histotype, grade, and stage of malignant tumors identify early stage disease and a favorable histotype, laying the groundwork for PM. In the 2000s, estrogen and progesterone receptors, Ki67, and HER2/neu emerged as novel tissue biomarkers in breast cancer. The manifestation of these phenotypic indicators has altered breast cancer treatment selection and outcome prediction. The biomarker landscape has considerably broadened over the past two decades, with numerous cell-, sub-, and molecular-level biomarkers identified in neoplastic cells and cancer-associated immune cells. AI-powered evaluation tools are needed due to the growing number and complexity of tissue biomarkers. Artificial intelligence in pathology simplifies biomarker assessments and strengthens Pathology’s role in Personalized Medicine [1]. Figure 2 illustrates an integrated AI workflow in digital pathology.

2.1 AI in Liver and Lung Pathology

In the field of liver pathology, the use of AI-based algorithms has increased, with the goal of improving diagnostic precision, streamlining workflows, and, ultimately, improving patient care. Junhao Liang et al. pioneered PathFinder, a human-centric deep learning (DL)-guided framework designed to assist pathologists in the detection of novel tissue biomarkers. By integrating attribution methods with sparse multiclass tissue spatial circulation data from whole slide images (WSIs), this innovative method bonds the gap between clinical prognosis and DL. PathFinder not only localizes, characterizes, and validates potential biomarkers, but it also provides cutting-edge prognostic performance. The authors, inspired by PathFinder, discovered a previously neglected correlation between tumor necrosis in liver cancer and patient prognosis. Consequently, they planned two clinically autonomous indicators, mortification area fraction and tumor necrosis distribution, and demonstrated their practical prognostic potential based on REMARK-derived criteria [2]. Cardobi et al. highlight the current

Table 1 Concordance of common artificial intelligence terminologies and acronyms, as well as acronyms specific to this manuscript with which the reader may be unfamiliar

AI	A field of computer science disturbed with the growth of autonomous, self-regulating systems
Machine learning (ML)	A subfield of AI concerned with programming computers to learn from examples and then use that knowledge to generate new predictions or changes to existing ones
Deep learning (DL)	A subfield of ML that allows for the independent learning of hierarchical data representations through the use of multi-layer neural networks
Whole slide imaging (WSI)	A digital imaging method employed in pathology that acquires high-resolution images of complete tissue slides for examination and diagnosis
Neural network (NN)	A network of interconnected computer nodes or neurons organized in a hierarchical fashion that mimics the way the human brain works
Convolutional neural network (CNN)	A NN explicitly engineered for the analysis of structured grid data, including images. Convolutional layers are employed to autonomously and adaptively learn hierarchical features from data
Area Under the receiver operator characteristic curve (AUC)	A criterion employed to assess the efficacy of a machine learning model, particularly in classification endeavors. A receiver operating characteristic (ROC) curve's region beneath the curve is computed
Fully convolutional network (FCN)	The process of semantic segmentation using a neural network architecture in which a class label is assigned to each point in an input image
Recurrent neural network (RNN)	An architecture of neural networks intended for the processing of data sequences, wherein information circulates cyclically among network layers
Generative adversarial network (GAN)	A neural network architecture where a generator and a discriminator are trained concurrently and in opposition to one another. The discriminator's job is to identify false data, while the generator's is to provide real data

(continued)

Table 1 (continued)

Supervised learning	A method used in machine learning to determine the mapping function from input to output by training the algorithm on a labeled dataset with input–output pairs
Transfer learning	A method used in machine learning whereby a model that has been trained for one job is modified to do a second, related task by utilizing the knowledge learnt from the previous task

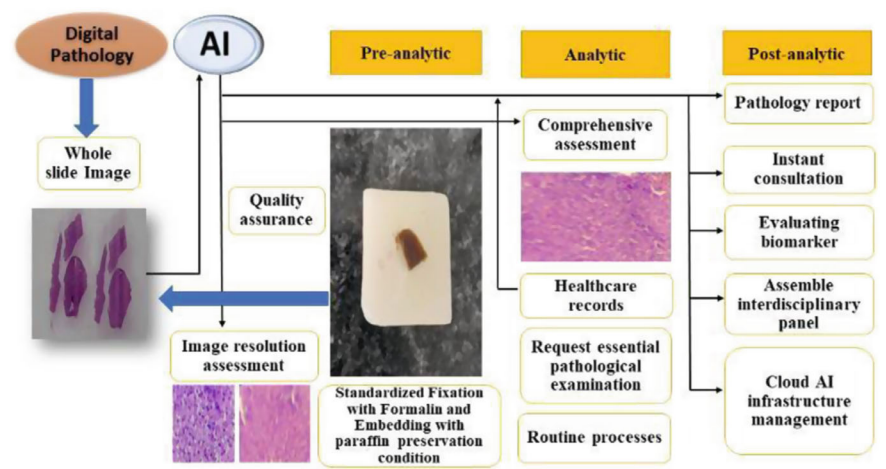


Fig. 2 The provided figure showcases artificial intelligence analyzing Whole Slide Images, detecting scan quality tissues, ensuring pre-analytical quality control, and actively participating in post-analytical processes. Adapted from “I. Kim, K. Kang, Y. Song, and T. J. Kim, “Application of Artificial Intelligence in Pathology: Trends and Challenges,” *Diagnostics* 2022, Vol. 12, Page 2794, vol. 12, no. 11, p. 2794, Nov. 2022, <https://doi.org/10.3390/DIAGNOSTICS12112794>”

use of AI algorithms for liver and pancreas imaging. For radiologists, these algorithms mean faster repetitive tasks like segmentation, better picture quality, faster scanning times, optimized imaging acquisition, and new quantitative measures like lesion volume and tumor load [3]. Pierre Allaume et al. highlight the prevalence and versatility of Deep Neural Network (DNN)-based models in the varied landscape of liver pathology [4]. Using H/E pictures, Coudray et al. built a deep learning network to identify the most frequent gene mutations in lung adenocarcinoma (ADK), and the findings were noteworthy. Their AI technology did an excellent job of differentiating between ADK and SqCC when it came to lung cancer, and they showed strong results across three separate cohorts [5]. Furthermore, Sha et al. demonstrated significant forecast-ability in ADK by using a deep learning model to predict PD-L1

status in non-small cell lung cancer (NSCLC) employing H/E slides. The model’s efficacy remained uniform across various PD-L1 thresholds and in situations mimicking pathologist discordance. These advancements collectively illustrate AI’s transformative influence on liver pathology, encompassing biomarker discovery and enhanced diagnostic capabilities in medical imaging [6]. The studies discussed above are shown in Table 2.

Table 2 Studies on AI in liver and lung pathology

Author’s	Year	Approach	Key findings/conclusion
Junhao Liang et al.	2022	The PathFinder framework for liver pathology is a human-centric, interpretable DL-guided one. Mixes attribution algorithms with sparse multiclass tissue spatial distribution of WSIs	Identified tumor necrosis in liver cancer as a critical determinant of patient prognosis. Suggested necrosis area fraction and tumor necrosis distribution as clinically independent prognostic indicators. Validated the indicators’ efficacy in clinical prognosis
Nicolò Cardobi et al.	2021	AI algorithms in liver and pancreas medical imaging. Speeds up repetitive tasks, acquires new quantitative parameters, improves image quality, reduces scanning time, and optimizes imaging acquisition	Highlighted the practical applications of AI algorithms in liver and pancreas medical imaging. Emphasized their role in speeding up tasks, acquiring quantitative parameters, reducing scanning time, improving image quality, and optimizing imaging acquisition
Pierre Allaupe et al.	2023	DNN-based models in the field of liver pathology. Diverse applications	DNN-based models are diverse and well-represented in liver pathology
Coudray et al.	2018	Trained DL model to forecast gene mutations in lung ADK from H/E images. Classified lung cancer into SqCC and ADK	Stated AUC ranging from 0.733 to 0.856 for 6 genes. Achieved good performance in classifying lung cancer into ADK and SqCC (AUC: 0.86–0.97)
Sha et al.	2019	Develop a DL model for PD-L1 status prediction in non-small cell lung cancer using hematoxylin and eosin slides. Evaluated model efficacy using various PD-L1 thresholds and simulating pathologist discrepancies	Demonstrated substantial predictive capability in ADK (AUC 0.85) but not in SqCC (AUC 0.64). The model maintained efficacy across various PD-L1 thresholds and in replicating pathologist discordance

2.2 *AI in Stomach and Colon Pathology*

The utilization of AI in stomach and colon pathology has led to substantial progress, fundamentally transforming the diagnosis, characterization, and treatment of gastrointestinal disorders and colorectal cancer. Researchers Osamu Iizuka et al. trained RNNs and CNNs effectively using WSIs from colon and stomach biopsy histology. Differentiating WSIs into adenocarcinoma, adenoma, and non-neoplastic groups was the primary goal. The approach achieved impressive AUC values of 0.99 and 0.97 for gastric cancer and adenoma, respectively. For colonic adenocarcinoma and adenomas, the corresponding values were 0.99 and 0.96. These results show that the models can be applied to other situations and have a lot of promise to be used in real-world histopathological diagnosis processes [7]. Using historical data, Bychkov et al. trained a deep learning model to categories colorectal cancer patients into low- and high-risk groups for recurrence within 5 years. Using images from H&E-stained TMA specimens, the deep learning model's digital risk score hit a home run with an AUC of 0.69. This outperformed predictions made by expert pathologists at both the TMA spot (AUC 0.58) and whole slide (AUC 0.57) levels, defying expectations [8]. By developing a tumor detector that had an AUC higher than 0.99, Kather et al. made substantial strides in the area. Afterward, they used large patient cohorts from TCGA, which included both fresh and formalin-fixed tissue specimens, to create an extra AI model that could distinguish between Microsatellite Instability (MSI) and Microsatellite Stability (MSS). With AUCs of 0.81 for colorectal cancer, 0.75 for endometrial cancer, also 0.69 for gastric cancer, the combination of these models produced strong performance across a variety of human tumors. The results significantly exceeded previously established benchmarks for predicting molecular characteristics from histological data [9].

The studies discussed on AI in stomach and colon pathology are summarized in Table 3.

2.3 *AI in Skin Pathology*

Artificial intelligence and deep learning are revolutionizing the domain of pathology, especially in dermatopathology. Verma et al. innovatively classified erythematosquamous diseases using an ensemble of five distinct data mining methods. The results demonstrated that this collaborative method optimized the dataset's utilization and outperformed individual data mining techniques [10]. Moradi and Mahdavi-Amiri presented an innovative kernel sparse illustration technique for the segmentation and cataloging of melanoma pictures. The assessment results indicated that their methodology was comparable to current state-of-the-art techniques [11]. Yap et al. employed a multimodal methodology, developing a classifier that surpassed baseline models reliant on individual macroscopic pictures in equally binary melanoma detection and multiclass cataloging [12]. Jin Bu et al. developed a Recurrent Neural Network

Table 3 Studies on artificial intelligence in stomach and colon pathology

Authors	Year	Approach	Key findings
Osamu iizuka et al.	2020	Utilized CNNs and RNNs to analyze biopsy histopathology WSIs of the stomach and colon	Achieved high AUC (up to 0.97 and 0.99) for classifying gastric and colonic adenocarcinoma and adenoma. Demonstrated generalization ability and potential for deployment in practical histopathological diagnostic workflow system
Bychkov et al.	2018	Established a DL model to predict the risk of five-year colorectal cancer recurrence using images of H&E-stained tissue microarray specimens	The digital risk score surpassed predictions derived from grading and/or staging by expert pathologists, achieving an AUC of 0.69, in contrast to 0.58 (TMA spot) and 0.57 (whole slide level)
Kather et al.	2019	Engineered a tumor detection system with an AUC exceeding 0.99; trained an additional AI model to differentiate between microsatellite instability (MSI) and microsatellite stability (MSS) in extensive persistent cohorts from TCGA. Integrated the two models	Robust performance across various human tumors (AUCs: 0.81 for CCR, 0.75 for endometrial cancer, 0.69 for gastric cancer). Exceeded previous performance in predicting molecular features from histology

approach specifically designed for dermatological and venereological conditions, with assessments demonstrating enhanced efficacy in predicting skin diseases relative to the traditional ICD-10 classification for skin and subcutaneous tissue diseases [13]. Yuchun Ding et al. broadened the applications of deep learning by introducing two sequentially implemented DL models that facilitated accurate segmentation of epidermal and dermal structures. This method not only eradicated typical artifacts but also facilitated quantitative analysis of immunostained signals. The models attained a regular global accuracy of 95.0% for epidermis and dermis segmentation, and 86.1% for positive/negative cell segmentation [14]. These advancements highlight AI and deep learning’s transformative role in pushing the boundaries of accuracy and efficiency in dermatopathology. The studies discussed above are summarized in Table 4.

3 The Future Outlook of AI in Pathology

The application of AI in clinical settings is restricted by various limitations. The insufficiency of proper digitization constitutes the foremost challenge. A current investigation in England exposed that fewer than 30% of institutions possess a comprehensive digital pathology (DP) workstation. Furthermore, the predominant use of these

Table 4 Studies on artificial intelligence in skin pathology

Author's	Year	Approach	Key findings
Verma et al.	2019	A combination of five data mining methods classifies erythematosquamous diseases. A unified method is proposed for efficient dataset use and improved accuracy over individual techniques	Improved efficiency and precision in the classification of erythematosquamous diseases via the proposed ensemble method
Moradi and Mahdavi-Amiri	2019	Melanoma picture segmentation and classification using a kernel sparse representation method. The methodology is comparable to current state-of-the-art techniques regarding segmentation and classification efficacy	Competitive performance demonstrated in the segmentation and cataloging of melanoma images
Yap et al.	2018	Development of a multimodal classifier for binary melanoma detection and multiclass classification. Does better than a baseline classifier that only uses one macro-scoped image	Superior outcomes in binary melanoma detection and multiclass cataloging when compared to previous techniques
Jin Bu et al.	2021	Proposal for a Recurrent Neural Network (RNN) approach to predict dermatological and venereological diseases. The proposed method demonstrates enhanced efficacy in forecasting skin diseases relative to the ICD-10 classification for diseases of the skin and subcutaneous tissues	Overall better predictive performance for skin diseases compared to the traditional ICD-10 classification system
Yuchun Ding et al.	2023	Utilization of two Deep Learning (DL) models for precise segmentation of epidermal and dermal structures, elimination of prevalent artifacts, and quantitative assessment of immunostained signals. Models attained elevated precision in segmentation tasks	Attained a regular global accurateness of 95.0% for epidermis and dermis segmentation and 86.1% for positive/negative cell segmentation, indicating efficacy in analytical tasks

workstations is for coaching, research, and quality assurance, while clinical applications are often relegated to consultations rather than primary diagnoses. Another obstacle to the widespread adoption of AI in daily practice is its lack of robustness and over-all applicability. Numerous current AI models exhibit a 20% decline in performance when applied in contexts distinct from their original training environments. To mitigate these limitations, strategies including transfer learning and the development of open-source datasets, facilitated by organizations like The Cancer Genome Atlas, are suggested. Notwithstanding these endeavors, challenges persist, notably the necessity for extensive datasets with meticulously annotated ground truth data, a labor-intensive undertaking that undermines the initial impetus for AI implementation—mitigating the pathologist deficit.

Additionally, discrepancies in staining protocols, tissue types, and scanning devices can affect the performance of AI systems. The hesitance of pathologists to embrace AI models constitutes an additional limitation, as their conventional two-step examination process is incompatible with numerous AI models designed with smaller tiles as input. A notable challenge is the “black box” characteristic of AI decisions, wherein the reasoning behind the model’s outputs remains opaque. Proposed solutions involve correlating AI outputs with visual maps that illustrate the plenty and morphology of structures identifiable by pathologists.

Notwithstanding these challenges, the full incorporation of AI in pathology could represent a pivotal moment in digital health. The implementation of AI as a diagnostic aide is anticipated to alleviate pathologists’ workload, standardize diagnostic procedures, enhance advancements in human biology, and refine modified diagnostics and patient care. To achieve this, a straightforward user interface and platforms that amalgamate diverse AI solutions and clinical data are necessary. The collaboration among digital pathology, artificial intelligence, and radiomics is anticipated to significantly enhance the comprehension of underlying biological processes. As advanced diagnostic methodologies proliferate, pathologists who resist adopting digital pathology and artificial intelligence may find their roles in the field of pathology diminishing [1].

4 Conclusion

Artificial intelligence (AI) is revolutionizing pathology through the accurate and swift examination of complete slide images. This is particularly apparent in liver, lung, stomach, colon, and skin pathology, where AI-driven models have exhibited remarkable capabilities in disease diagnosis and prediction. Nevertheless, considerable obstacles persist, including restricted digitization, model resilience, discrepancies in data sources, and the opaque nature of AI decision-making. Notwithstanding these challenges, initiatives like transfer learning and open-source datasets present promising solutions. The comprehensive integration of AI offers significant potential for pathology practice by alleviating pathologists’ workload, standardizing diagnostic procedures, and enhancing personalized diagnostics and patient care. The

integration of digital pathology, artificial intelligence, and radiomics will augment our comprehension of biological processes and facilitate more thorough diagnostic methodologies.

Acknowledgements We wish to express our gratitude to the Department of Medical Laboratory Technology at Chandigarh University for supplying the necessary facilities.

References

1. Lancellotti C et al (2021) Artificial intelligence & tissue biomarkers: advantages, risks and perspectives for pathology. *Cells* 10(4). <https://doi.org/10.3390/CELLS10040787>
2. Liang J et al (2022) AI inspired discovery of new biomarkers for clinical prognosis of liver cancer. <https://doi.org/10.1101/2022.05.03.490448>
3. Cardobi N et al (2021) An overview of artificial intelligence applications in liver and pancreatic imaging. *Cancers (Basel)* 13(9). <https://doi.org/10.3390/CANCERS13092162>
4. Allaume P et al (2023) Artificial intelligence-based opportunities in liver pathology—a systematic review. *Diagnostics* 13(10). <https://doi.org/10.3390/DIAGNOSTICS13101799/S1>
5. Coudray N et al (2018) Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med* 24(10):1559–1567. <https://doi.org/10.1038/S41591-018-0177-5>
6. Sha L et al (2019) Multi-field-of-view deep learning model predicts nonsmall cell lung cancer programmed death-ligand 1 status from whole-slide hematoxylin and eosin images. *J Pathol Inform* 10(1). https://doi.org/10.4103/JPI.JPI_24_19
7. Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M (2020) Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Sci. Rep.* 10(1). <https://doi.org/10.1038/S41598-020-58467-9>
8. Bychkov D et al (2018) Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 8(1). <https://doi.org/10.1038/S41598-018-21758-3>
9. Kather JN et al (2019) Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 25(7):1054–1056. <https://doi.org/10.1038/S41591-019-0462-Y>
10. Verma AK, Pal S, Kumar S (2019) Classification of skin disease using ensemble data mining techniques. *Asian Pacific J. Cancer Prev.* 20(6):1887–1894. <https://doi.org/10.31557/APJCP.2019.20.6.1887>
11. Moradi N, Mahdavi-Amiri N (2019) Kernel sparse representation based model for skin lesions segmentation and classification. *Comput. Methods Programs Biomed.* 182. <https://doi.org/10.1016/J.CMPB.2019.105038>
12. Yap J, Yolland W, Tschandl P (2018) Multimodal skin lesion classification using deep learning. *Wiley Online Libr* 27(11):1261–1267. <https://doi.org/10.1111/exd.13777>
13. Bu J et al. (2021) Prediction of skin disease using a new cytological taxonomy based on cytology and pathology with deep residual learning method. *Sci. Rep.* 11(1). <https://doi.org/10.1038/S41598-021-92848-Y>
14. Ding Y et al (2023) An open source pipeline for quantitative immunohistochemistry image analysis of inflammatory skin disease using artificial intelligence. *Wiley Online Libr.* 37(3):605–614. <https://doi.org/10.1111/jdv.18726>

Audio Sentiment Analyzer for Mental Health Support Hotline



Sneha Maurya, Shivani Saxena, Shreya Shukla, Sukriti Maurya,
and Akhilesh Verma

Abstract The incorporation of emotion detection technology on mental health helplines is picking up to amplify care quality. In this work, we propose to enhance these services by utilizing state-of-the-art techniques in machine learning and natural language processing. The paradigm is LSTM-based (or Long Short-Term Memory) networks which are powerful for modelling sequences like speech. We utilize LSTM models to help us understand the various types of emotions and detect cases with mental issues as soon as possible. Our results suggest important opportunities for the integration of emotion detection in hotline services. The observed accuracy, precision, and recall rates were higher than existing frameworks showing that the model can accurately classify emotions such as happiness, sadness, anger, and surprise. During training and validation, the model's accuracy was 90.14%, specifically showcasing its strength. Additional research examines obstacles hotline operators face, ethical dilemmas, and feasibility. The study keeps coming back to privacy, user consent, and cultural sensitivities. Together, results emphasize the critical role played by operator training, external validation, and cultural adaptability. In the future, efforts are being made toward real-time integration and multimodal recognition with longitudinal studies intended to promote technology that seamlessly integrates with compassionate care.

Keywords LSTM · NLP · Emotion detection · Mental health · Hotline

S. Maurya · S. Saxena · S. Shukla · S. Maurya · A. Verma (✉)
Ajay Kumar Garg Engineering College, Ghaziabad, India
e-mail: vermaakhilesh@akgec.ac.in

S. Maurya
e-mail: Sneha2010029@akgec.ac.in

S. Saxena
e-mail: shivani2010189@akgec.ac.in

S. Shukla
e-mail: shreya2010042@akgec.ac.in

S. Maurya
e-mail: sukriti2010109@akgec.ac.in

1 Introduction and Background

With an increase in understanding mental health is a key part of overall health, more programs are available to support those with difficulties coping with emotional distress. Of these options, mental health hotlines have been crucial in providing instant, anonymous help. Nevertheless, hotlines continue to struggle with being able to identify the emotional states of their callers, which limits the success and/or execution of a timely (or any) intervention. The root of this difficulty is the nature of human emotions themselves—something with so many nuanced details can be difficult to communicate in words and verbal communication opens itself up to interpretation or potential biases.

There is increasing interest in incorporating emotion detection technology in hotline calling to mitigate these deficiencies. With the advances in machine learning and Natural Language Processing models of emotion detection are able to analyze how people speak, word choices that were made or even intonations used. These technologies are the next frontier in hotline services, offering instant windows into the emotional states of callers and enabling dispatchers to give more nuanced, compassionate help.

This study investigates the use of emotion detection technology for mental health hotlines by looking at how integrating sensing can support detection and response to emotional need in a phone call, covering user experience (UX), the hard clinical question—is this person distressed?—and front-line delivery of service, with attention to key ethical concerns such as user privacy and consent. This research aims to contribute to the debate on the relationship between technology and human-centered care, inciting a future in which technological progress complements empathetic support using a multi-stemmed strategy of literature review, ethical analysis, and empirical inquiry.

2 Literature Review

2.1 *Multimodal Sentiment Analysis*

Advances in multimodal sentiment analysis can be seen as a power of the fusion of various data types for explicating associated emotions more accurately. Rahman et al. did it through the new MAG-BERT and MAG-XLNet models, which use large pretrained transformers to involve multimodal information integrations that are not possible with unimodal sentiment analysis [1]. Similarly, Hazarika et al. introduce the MISA model which can explain both modality-specific and modality-invariant representations by learning factorized subspace [2]. Furthermore, Kim and Lee's Multi-Attention Recurrent Neural Network (MA-RNN) are architectural examples of both multimodal sentiment analysis for interpreting more complex relationships within multimodal data [3].

2.2 Emotion Recognition Applications

The use cases of emotion recognition in the industry are wide-ranging including security and crime detection, and healthcare. Boukabous and Azizi proposed the use of combination deep neural networks for audio and text together, regarding crime detection with superior metrics [4]. In a healthcare setting, Juyal combines audio and visual data to improve real-time sentiment analysis, illustrating efficiency gains with s-mall deep neural networks [5]. Additionally, Patra et al.'s HIRVM approach introduces a novel perspective in sentiment analysis, integrating swarm intelligence and machine learning for automated movie review analysis [6].

2.3 Speech Emotion Recognition

Speech emotion recognition has evolved with various computational models [7]. Bhattacharya et al. tackle multilingual audio emotion recognition using deep analytic methods, exploring features like MFCC, chroma, and Mel-spectrogram [8]. Ho et al. study introduces a multimodal strategy using multi-level multi-head fusion attention for improved speech emotion recognition [9]. Additionally, Puri et al. pioneer the use of hybrid convolutional neural networks in detecting emotions from RAVDESS audio data [10], emphasizing the potential of hybrid models in capturing subtle emotional nuances [11].

2.4 Problem Statement and Research Question

Conventional methods for monitoring mental health conditions, such as self-reporting and clinical assessments, are often subjective and time-consuming. Social media platforms offer user-generated content reflecting individuals' emotions and experiences, presenting an opportunity for automated sentiment analysis to provide real-time insights into mental health statuses. However, existing sentiment analysis techniques struggle with the contextual complexities of mental health-related language and privacy concerns. This research project aims to develop sentiment analysis methods tailored to social media data, focusing on context sensitivity, multimodal analysis, and privacy. The project seeks to provide accurate and interpretable tools for mental health monitoring, potentially enhancing early intervention strategies and personalized treatment approaches (Table 1).

Table 1 Summary of multimodal sentiment analysis and emotion recognition studies

S. no.	Title	Author	Outcome	Drawbacks
1	Integrating multi-modal information in large pretrained transformers [1]	Wasifur Rahman, Md. Kamrul Hasan	Explores including multimodal data into large pretrained transformer models for improved NLP outcomes	Limited utilization of additional information in acoustic and visual domains by XL-Net
2	Emotion detection from multilingual audio using deep analysis [8]	S. Bhattacharya	This paper likely presents a machine learning-based approach for sentiment analysis, possibly incorporating swarm intelligence techniques to enhance performance	The paper is not using any visual representation of sound only direct from natural sound data
3	Sentiment analysis in textual, visual and multimodal inputs using recurrent neural networks [12]	Tembhurne and Diwan	Comprehensive review of RNN-based models for sentiment analysis across various modalities. Identification of challenges and approaches for applying RNN architectures to multimodal SA tasks	Limited empirical validation or experimentation, as the paper is a review rather than a primary research study
4	Sentiment analysis in non-fixed length audios using a Fully Convolutional Neural Network [13]	García-Ordás	Proposed FCN architecture for sentiment analysis of variable-length audio data. Promising results in sentiment analysis tasks, out-performing state-of-the-art methods	Limited discussion on potential challenges or limitations of FCN-based sentiment analysis, such as computational complexity or generalization to diverse audio sources
5	Speech emotion recognition based on Graph-LSTM neural network [14]	Yan Li, Yapeng Wang, Xu Yang and Sio Kei Im	Introduction of Graph-LSTM neural network architecture for speech emotion recognition. Competitive performance on the IEMO-CAP dataset, with improved interpretability of emotion features	Lack of detailed analysis on the scalability or efficiency of the proposed Graph-LSTM approach, especially in handling large-scale datasets or real-time applications

(continued)

Table 1 (continued)

S. no.	Title	Author	Outcome	Drawbacks
6	A review on speech emotion recognition using deep learning and attention mechanism [15]	Lieskovska´	Review of recent developments in SER and examination of attention mechanisms’ impact on SER performance. Comparative analysis of various attention mechanisms on the IEMOCAP benchmark database	Limited discussion on practical challenges or limitations of attention-based SER models, such as computational overhead or interpretability issues
7	Automatic emotion recognition in children with autism: a systematic literature review [16]	Landowska	Systematic literature review on emotion recognition methods for children with autism. Identification of diverse observation channels and modalities for emotion recognition	Limited discussion on methodological challenges or limitations encountered during the systematic review process, such as bias in dataset selection or analysis
8	Smart multi-model emotion recognition system with deep learning [17]	Kalyani	Proposal of a multi-modal emotion recognition system integrating speech, text, and facial expressions. Emphasis on potential applications in spam call detection, customer services, and audience engagement	Limited empirical validation or experimentation on the proposed multi-modal emotion recognition system, such as performance evaluation in real-world scenarios or user studies

3 Methodology

3.1 Data Collection and Preprocessing

We acquire the Toronto Emotional Speech Set (TESS) dataset, comprising actors’ voice recordings expressing various emotions, including joy, sorrow, and rage. This structured collection provides diverse emotional speech samples. Mel-frequency cepstral coefficients (MFCCs) are extracted from audio recordings using the librosa package, capturing essential features of the speech signal. Preprocessing methods such as noise reduction, resampling, and normalization ensure consistency and quality, and the recordings are divided into shorter frames for feature extraction and model training.

3.2 Model Architectural Design

The LSTM-based neural network for emotion prediction from speech input consists of an input layer accepting sequences of MFCC features. A dense output layer with softmax activation, categorizing the input voice into emotion classes such as glad, sad, livid, and neutral. This hierarchical structure captures temporal features, enhancing the model's accuracy.

3.2.1 LSTM Architecture

The LSTM model includes:

- An LSTM layer with 256 units to capture sequence data.
- Dropout layers after each LSTM and dense layer to prevent overfitting.
- Dense layers (128, 64) for feature reduction and emotion classification.
- An output layer with 7 units corresponding to the emotions classified.

3.2.2 Equations

LSTM networks learn long-term dependencies in sequential data, ideal for time series forecasting, speech recognition, and language translation. They can also be used with other neural networks for image and video analysis. The model includes:

- **Input Gate:** The input gate in LSTM adjusts the cell state by integrating relevant data from h_{1-t} and X_t .

$$i_t = \sigma(w_i[h_{t-1}, X_t] + b_i)$$
Forget Gate: The forget gate is used to eliminate data that is no longer needed in the cell state (Figs. 1 and 2).

$$f_t = \sigma(w_f[h_{t-1}, X_t] + b_f)$$
- **Output Gate:** The forget gate in LSTM networks discards unnecessary information from the cell state by processing inputs X_t and h_{t-1} with weight matrices and biases

$$o_t = \sigma(w_o[h_{t-1}, X_t] + b_o).$$

3.2.3 Waveplot and Spectrogram for Various Emotions

Waveplots and spectrograms visualize emotional voice signals. For example (Figs. 3 and 4):

Waveplots show lower amplitude, smoother waveforms for melancholy, while spectrograms reflect lower energy levels and fewer high-frequency components. Abrupt amplitude and frequency rises indicate pleasant surprise, while angry speech features jagged waveforms, sharp frequency shifts, and high amplitude (Figs. 5 and 6).

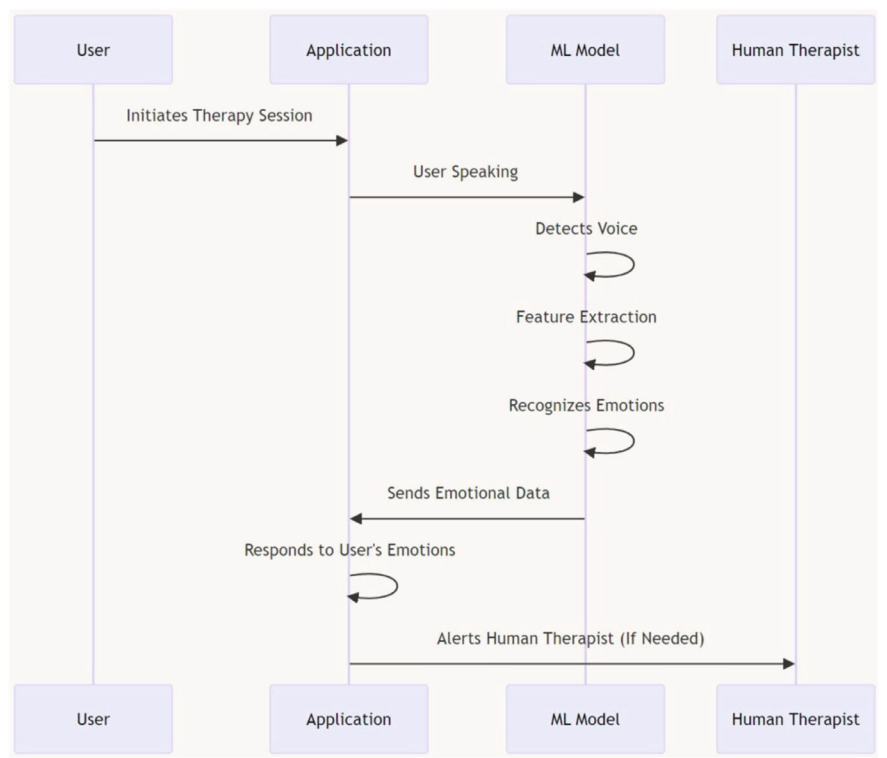


Fig. 1 Mental health monitoring

3.3 *Model Training and Validation*

To ensure balanced representation of emotion classes, the preprocessed dataset is split into training, validation, and test sets via stratified sampling. We fine-tune hyperparameters such as learning rate, batch size, and dropout rates on the validation set. We used the training split to train it and recorded stats like loss/accuracy. We can use early stopping as a simple heuristic to solve the following problem. When validation performance starts to plateau, the model has stopped learning.

3.4 *Model Evaluation and Emotion Prediction*

We evaluate the efficacy of LSTM model using accuracy, precision, recall and *F1*-score [18] on a test set and having a confusion matrix as well as cross-validation assert for its generalization. The flow chart illustrates the detection of emotions in spoken inputs for response, and a voice-based emotion predictor for operators to respond

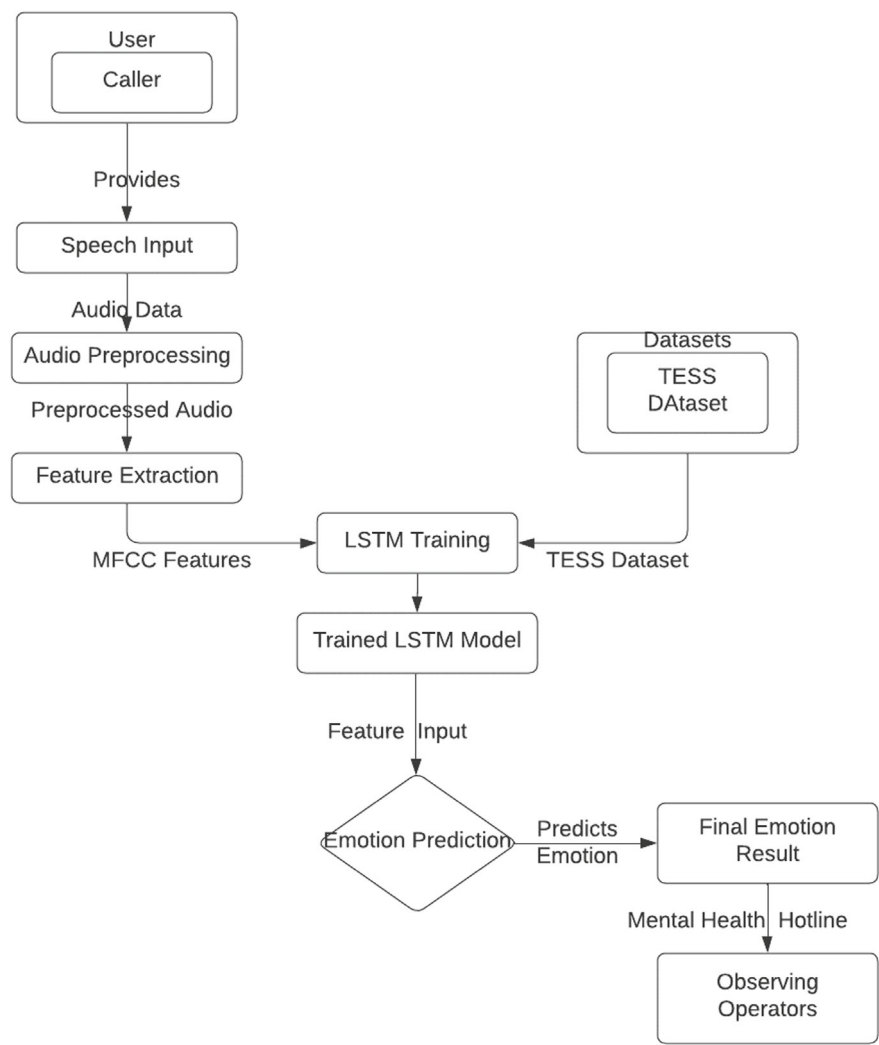


Fig. 2 Model flowchart

emotionally and judiciously. It also looks out for mental health via an app [19] that picks up on agitation, asking if a user is okay and connecting him or her to someone who can help; voice data points the way to more therapy.

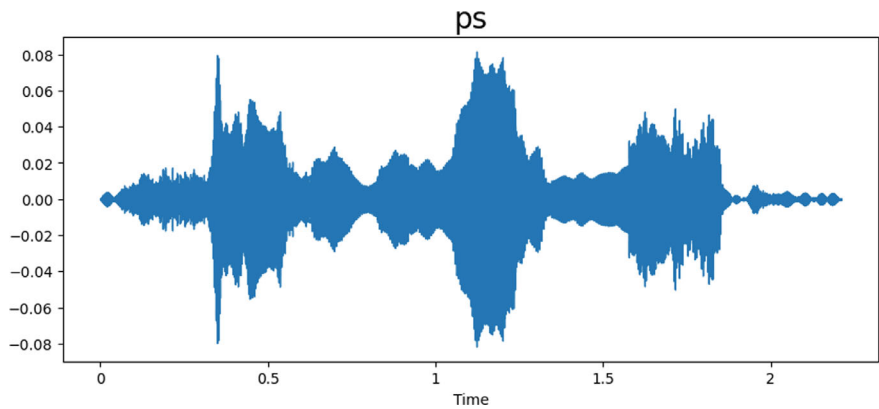


Fig. 3 Waveplot of pleasant surprise

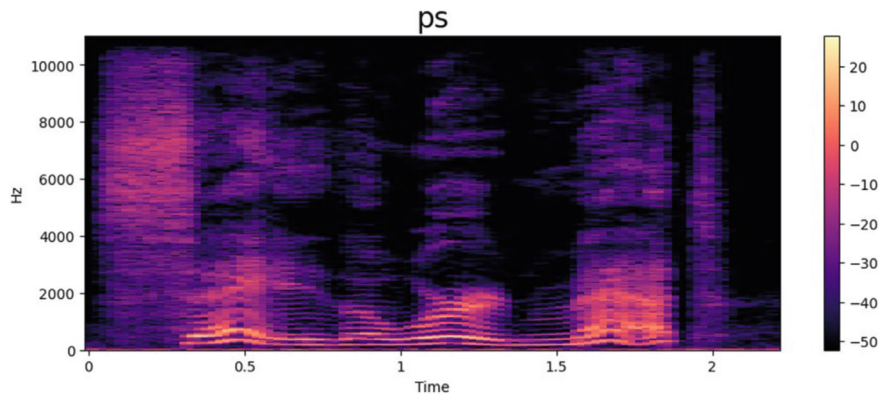


Fig. 4 Spectrogram of pleasant surprise

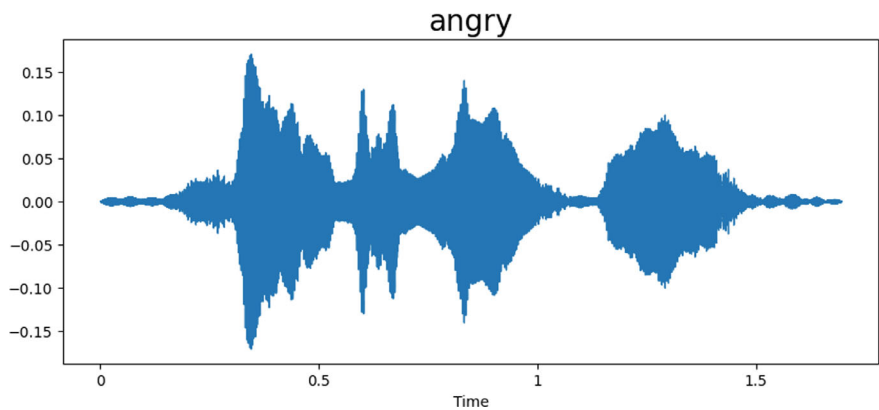


Fig. 5 Waveplot of angry

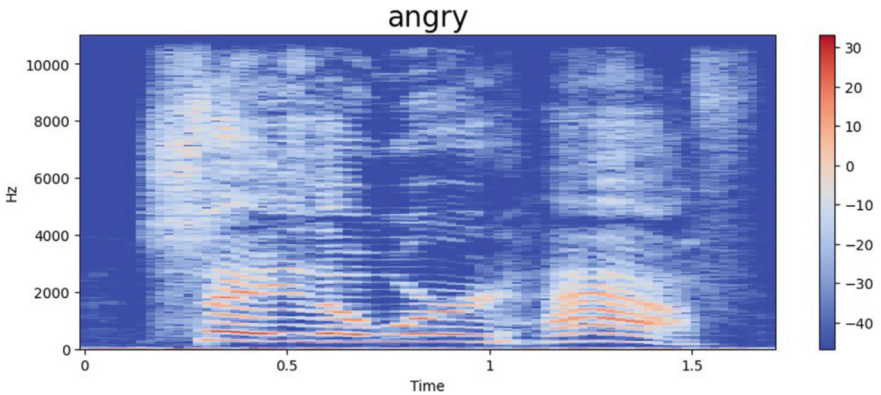


Fig. 6 Spectrogram of angry

4 Results and Discussion

There is a pressing need to incorporate emotion detection technology in mental health helplines to improve support services. In this paper, we investigate the collaboration by building an LSTM neural network to predict emotions from speech data with a specific focus on real-time integration of models onboard ethical constraints in both design and reporting and assessing practical efficacy. The model was constructed for sequential speech data processed with MFCC features from the Toronto Emotional Speech Set (TESS) dataset underwent training and validation including every possible scenario.

A 90.14% accuracy rate shows the robustness of this model in real world, and proofs the design and training process was correct. This high accuracy demonstrates the ability of the model to analyze signals from audio and categorize emotions in real-time, such as happiness, sadness, and anger that is very important for mental health hotlines to be able to quickly provide personal support.

The model entered into the mental health hotline services augmented them to a great extent, in order to provide better and efficient services. Being capable of detecting emotional cues from the voice of the caller lead to more precise interventions, eliminating much of the challenge in interpretation that human operators endure. The model has a greater slope in these curves, underscoring the fact that this technique has potential to expedite support services while ensuring less delays and errors (Fig. 7).

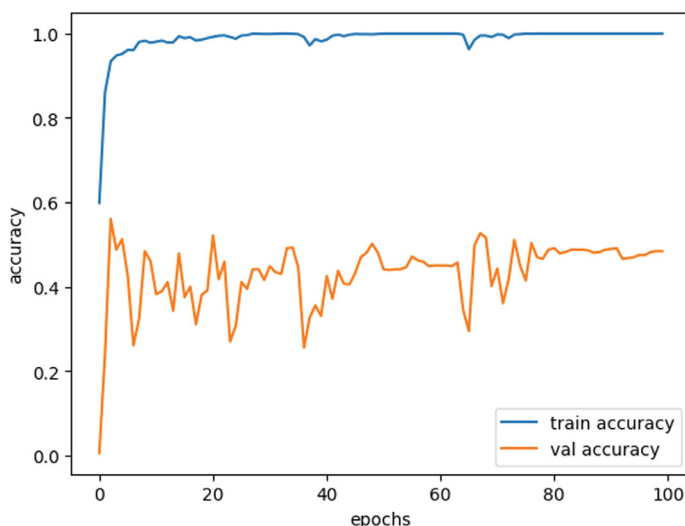


Fig. 7 Training and validation set accuracy

5 Conclusion and Future Directions

The model capable of such a feat opens up avenues to further research and explore to improve modelling of emotions about mental health. About the upward performance trend, we are particularly enthusiastic because it provides possibilities for multimodal integration [20], real-time adaptability, and longitudinal cases. First of all, issues of privacy and especially informed consent for patients utilizing these services have to be clearly resolved to smoothly and without hiccups integrate these kinds of development into humanistic patient care.

Moreover, these results also suggest how model can be used to help in improving the effectiveness of mental health hotlines by coming up with different emotion detections in relation to context and which eliminates the bias and errors in reading emotions for appropriate support. This joins a larger discussion of technology in mental healthcare and draws attention to a specific area of future research—emotion recognition.

References

1. Rahman W et al (2020) Integrating multimodal information in large pretrained transformers. In Proceedings of the annual meeting of the association for computational linguistics. <https://doi.org/10.18653/v1/2020.acl-main.214>
2. Hazarika D, Zimmermann R, Poria S (2020) MISA: modality-invariant and specific representations for multimodal sentiment analysis. In: MM 2020—Proceedings of the 28th ACM international conference on multimedia. <https://doi.org/10.1145/3394171.3413678>

3. Kim T, Lee B (2020) Multi-attention multimodal sentiment analysis. In: ICMR 2020—Proceedings of the 2020 international conference on multimedia retrieval. <https://doi.org/10.1145/3372278.3390698>
4. Boukabous M, Azizi M (2022) Multimodal sentiment analysis using audio and text for crime detection. In: 2022 2nd International conference on innovative research in applied science, engineering and technology, IRASET 2022. <https://doi.org/10.1109/IRASET52964.2022.9738175>
5. Juyal P (2022) Multi-modal sentiment analysis of audio and visual context of the data using machine learning. In: 3rd International conference on smart electronics and communication, 2022-Proceedings. <https://doi.org/10.1109/ICOSEC54921.2022.9951988>
6. Patra RK, Patil B, Kumar TS, Shivakanth G, Manjula MB (2023) Machine learning based sentiment analysis and swarm intelligence. In: 2023 IEEE international conference on integrated circuits and communication systems, ICICACS 2023. <https://doi.org/10.1109/ICICACS57338.2023.10100262>
7. Sainath TN, Li B (2016) Modeling time-frequency patterns with LSTM vs. convolutional architectures for LVCSR tasks. In: Proceedings of the annual conference of the international speech communication association, Interspeech, 2016. <https://doi.org/10.21437/Interspeech.2016-84>
8. Bhattacharya S, Borah S, Mishra BK, Mondal A (2022) Emotion detection from multilingual audio using deep analysis. *Multimed Tools Appl* 81(28). <https://doi.org/10.1007/s11042-022-12411-3>
9. Ho NH, Yang, HJ, Kim SH, Lee G (2020) Multimodal approach of speech emotion recognition using multi-level multi-head fusion attention-based recurrent neural network. *IEEE Access* 8. <https://doi.org/10.1109/ACCESS.2020.2984368>
10. Efanov D, Aleksandrov P, Karapetyants N (2022) The BiLSTM-based synthesized speech recognition. *Procedia Comput Sci* 415–421. <https://doi.org/10.1016/j.procs.2022.11.086>
11. Puri T, Soni M, Dhiman G, Ibrahim Khalaf O, Alazzam M, Raza Khan I (2022) Detection of emotion of speech for RAVDESS audio using hybrid convolution neural network. *J Healthc Eng* 2022. <https://doi.org/10.1155/2022/8472947>
12. Tembhurne JV, Diwan T (2021) Sentiment analysis in textual, visual and multimodal inputs using recurrent neural networks. *Multimed Tools Appl* 80(5). <https://doi.org/10.1007/s11042-020-10037-x>
13. García-Ordás MT, Alaiz-Moretón H, Benítez-Andrades JA, García-Rodríguez I, García-Olalla O, Benavides C (2021) Sentiment analysis in non-fixed length audios using a fully convolutional neural network. *Biomed Signal Process Control* 69. <https://doi.org/10.1016/j.bspc.2021.102946>
14. Li Y, Wang Y, Yang X, Im SK (2023) Speech emotion recognition based on Graph-LSTM neural network. *EURASIP J Audio Speech Music Process* 2023(1). <https://doi.org/10.1186/s13636-023-00303-9>
15. Lieskovská E, Jakubec M, Jarina R, Chmúľík M (2021) A review on speech emotion recognition using deep learning and attention mechanism. *Electronics (Switzerland)* 10(10). <https://doi.org/10.3390/electronics10101163>
16. Landowska A et al (2022) Automatic emotion recognition in children with autism: a systematic literature review. *Sensors* 22(4). <https://doi.org/10.3390/s22041649>
17. Kalyani B, Sai KP, Deepika NM, Shahanaz S, Lohitha G (2023) Smart multi-model emotion recognition system with deep learning. *Int J Recent Innov Trends Comput Commun* 11(1):139–144
18. Suyal M, Goyal P (2022) A new classifier model on drug reviews dataset by VADER sentiment analyzer to analyze reviews of the dataset are real or fake based on machine learning. *Int J Eng Trends Technol* 70(7). <https://doi.org/10.14445/22315381/IJETT-V70I7P208>
19. de Boer JN, Corona Hernández H, Gerritse F, Brederoo SG, Wijnen FNK, Sommer IE (2022) Negative content in auditory verbal hallucinations: a natural language processing approach. *Cogn Neuropsychiatry* 27(2–3). <https://doi.org/10.1080/13546805.2021.1941831>

20. Kumar VS, Pareek PK, Costa De Albuquerque VH, Khanna A, Gupta D, Deepak Renukadevi S (2022) Multimodal sentiment analysis using speech signals with machine learning techniques. In: MysuruCon 2022—2022 IEEE 2nd Mysore sub section international conference. <https://doi.org/10.1109/MysuruCon55714.2022.9972662>
21. Atmaja BT, Sasou A, Akagi M (2022) Survey on bimodal speech emotion recognition from acoustic and linguistic information fusion. *Speech Commun* 140. <https://doi.org/10.1016/j.specom.2022.03.002>
22. Neelima G, Subha S, Chigurukota D, Santhosh Kumar B (2021) Automatic sentiment analyser based on speech recognition

The Application of Artificial Intelligence Algorithms in Computer Network Information Security Protection and Evaluation



Qianqian Yue

Abstract With the increasing complexity of the network environment and the continuous evolution of security threats, the traditional network security protection and evaluation methods are facing more and more challenges. It is urgent to find a solution to improve the protection efficiency and response speed. Artificial intelligence (AI) technology, especially machine learning and deep learning, is widely regarded as a powerful tool to enhance network security because of its excellent ability in data analysis and pattern recognition. This study aims to improve the accuracy of Intrusion Detection System (IDS) and malware identification by implementing a series of machine learning algorithms. Specific methods include using deep learning network to analyse and classify network traffic data, and using natural language processing technology to extract effective information from network threat intelligence. In addition, this paper also explores risk assessment models optimized by machine learning, which can dynamically analyse and evaluate potential security vulnerabilities. All participants have made remarkable progress during their study, and even their performance in the phishing recognition course has improved by nearly 42%. This study confirms the application value of AI algorithm in computer network information security protection and evaluation, which can not only greatly improve the performance of existing systems, but also bring new research and application directions to the field of network security.

Keywords Artificial intelligence algorithms · Computer network information security protection · Security assessment · Intrusion detection systems

Q. Yue (✉)

Shandong Vocational College Science and Technology, Weifang, China

e-mail: 13280731836@163.com

1 Introduction

Today, with the rapid development of digital economy, network security has attracted great attention all over the world. With the development of information technology, network attacks are becoming more and more complex and frequent, which puts forward higher requirements for traditional security protection methods. In this context, artificial intelligence technology has become the focus of network security research because of its efficient data processing and identification capabilities. However, although artificial intelligence has shown great potential in the field of network security, it still faces problems such as data dependence, insufficient model transparency and anti-attack. The introduction of artificial intelligence into the field of network security in this project will help improve the existing security methods and promote the development of new defence methods. Therefore, the research of this project has both theoretical significance and broad application prospects.

In this paper, the application status and technical problems of artificial intelligence in network security research are deeply analysed, and the future development of artificial intelligence is prospected. This project will reveal the advantages and limitations of artificial intelligence algorithm in improving network security performance on the basis of evaluating the application of existing artificial intelligence algorithms in intrusion detection, malicious program identification, security vulnerability analysis and other fields. On this basis, the key technical problems of artificial intelligence in data quality control, model transparency and anti-attack are studied. At the same time, this project will also give the corresponding improvement scheme on how to improve the performance of artificial intelligence in the field of network security.

The research is structured as follows: first, the paper will detail the application scenarios of AI in cybersecurity and analyse the operational effectiveness of AI in practice through examples. Then, the paper will assess and analyse the key technical challenges encountered in these applications, especially data dependency, transparency issues of algorithms and security issues of AI systems themselves. Finally, based on the previous analyses, this paper will explore and suggest targeted improvement measures, which are aimed at enhancing the practicality and efficiency of AI algorithms in network security, and also expected to provide theoretical basis and practical guidance for future research directions. This series of research not only helps to promote the development of network security technology, but also has an important significance for the realization of a more secure network environment.

2 Related Work

In exploring the application of AI in cybersecurity, most research has focused on leveraging AI's powerful data analysis to predict and defend against cyber-attacks. AI algorithms, especially machine learning and deep learning methods, have been

shown to be effective in improving the accuracy of malware detection and cyber IDS. These research results have greatly promoted the development of cybersecurity automation and intelligence. Xu Pei designed a ship network security control system based on AI algorithm and verified it [1]. Wei Shuailing studied the construction and assessment program of hospital information security system based on three-level hierarchical protection, which can improve the security level of hospital information [2]. Xue Rongsheng conducted a study on information security and protection strategies for computer networks [3]. Olasehinde designed a stacked integrated intrusion detection method for protecting information systems [4]. Cahyadi gave a certificate-less aggregated signature scheme for security and privacy protection [5]. However, existing research tends to ignore the problem of adapting to the complex environments that AI models face in real-world deployments, and the constant changes in these environments may lead to a decrease in model efficacy.

Furthermore, while many studies have highlighted the potential of using AI for security threat intelligence gathering and processing, these systems are capable of quickly identifying security threats and vulnerabilities from massive amounts of data. Zwilling conducted a comparative study of cybersecurity awareness, knowledge and behaviour [6]. Xing explored a study on the application of big data techniques in information security processing in computer networks [7]. Al-Fedaghi conducted a case study on email systems in network security modelling [8]. Guangxu carried out a study on information security in computer networks based on improved machine learning [9]. Shen gave a fuzzy group decision framework for computer network security assessment based on probabilistic linguistic information [10]. However, the current research still lacks an in-depth discussion on the security of the AI system itself, especially the defence capability in fighting malware and network attacks. This suggests that although AI algorithms have made significant progress in cybersecurity, how to enhance their resistance to interference and improve their stability in dynamic network environments is still a major challenge to be solved.

3 Methods

3.1 *AI-Based Network Security Protection Technology*

(1) Intrusion detection and protection

Intrusion detection and protection technology plays an important role in the field of network security. In recent years, artificial intelligence algorithms represented by machine learning and deep learning have been applied more and more in this field. By analysing the network business data, artificial intelligence technology finds possible abnormal behaviours and threats, and then realizes effective prevention and response to network attacks.

In the field of intrusion detection, artificial intelligence algorithm is generally used to build the model of abnormal behaviour. This model can extract abnormal

patterns with potential threats from ordinary network communication patterns. For example, using machine learning methods, such as support vector machine, decision tree and neural network, to classify and predict data streams can effectively improve the performance of the system.

The information entropy $H(X)$ is as follows:

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (1)$$

In computer network security, higher entropy usually means that data or behaviour within the system is more unpredictable and therefore potentially more secure [11]. This formula calculates the entropy of a random variable, where $p(x_i)$ is the probability of the i -th event occurring in the random variable x .

(2) Malicious code recognition and prevention

In addition, the identification and protection of malicious code is also an important aspect. Artificial intelligence algorithm can automatically analyse and identify malicious codes such as viruses, worms and trojans. Using deep learning methods such as convolutional neural network and cyclic neural network, artificial intelligence can learn the behaviour law of malicious programs from massive data and realize the prediction of the characteristics of the data to be tested.

The conditional entropy $H(X|Y)$ is as follows:

$$H(X|Y) = - \sum_{y \in Y} p(y) \sum_{x \in X} p(x|y) \log_2 p(x|y) \quad (2)$$

Conditional entropy is used to measure the uncertainty of random variable X under the condition that random variable γ is known in cybersecurity.

(3) Security vulnerability analysis and repair

Analysing and repairing the security vulnerabilities in the system is an important part of ensuring network security. By introducing artificial intelligence technology into this field, we can realize automatic discovery and repair of security defects. This project intends to use machine learning methods such as random forest to learn and predict vulnerabilities from historical vulnerability data. Artificial intelligence can also help developers find and correct security vulnerabilities in time during the development process, such as using code analysis tools to find and fix security problems. In addition, this method can automatically generate the corresponding fault code according to the previous fault data, which improves the security efficiency of the system.

3.2 AI-Based Information Security Assessment Methods

(1) Risk assessment and prediction

By applying AI algorithm, especially machine learning and data mining, we can identify risk patterns from a large number of historical data and realize the prediction of future security threats. The application of AI in risk assessment also includes automated threat intelligence analysis. AI system can collect and analyse threat information from different sources in real time, such as dark network monitoring data or social media information, thus providing a more comprehensive perspective for risk assessment. In addition, the AI algorithm can also help determine the priority of security measures and ensure that resources can focus on the most serious threats.

Mutual information $I(X; Y)$ is as follows:

$$I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (3)$$

In cybersecurity, mutual information can be used to analyse the extent of information leakage between two systems or system components, i.e., the magnitude of the impact of information leakage from one system on the security of another system.

(2) Security situational awareness and analysis

Security situational awareness refers to the comprehensive monitoring and evaluation of an organization's network security environment. Using artificial intelligence to realize security situation awareness can significantly enhance the ability of dynamic monitoring and response. AI algorithm, especially pattern recognition and anomaly detection technology, can effectively identify abnormal behaviours in complex network activities and find potential security threats in time.

For example, with the help of AI-based behaviour analysis tools, security analysts can detect abnormal network traffic and user behaviour and quickly identify security incidents caused by internal leaks or external attacks.

(3) Security assessment tools and platforms

Some advanced AI security platforms are able to automatically identify and categorize security vulnerabilities by integrating machine learning, natural language processing and image recognition technologies, providing priority ranking for remediation work and greatly enhancing the efficiency of security teams.

In summary, the application of AI algorithms in the field of information security assessment provides us with efficient and accurate new tools to help better understand and address increasingly complex cybersecurity challenges. As AI algorithms continue to advance and deepen their application, the future cybersecurity environment will be more intelligent and automated.

The probability of error P_e is as follows:

$$P_e = p \cdot P(D = 0|H = 1) + (1 - p) \cdot P(D = 1|H = 0) \quad (4)$$

This formula is used in hypothesis testing, where $P(D = 0|H = 1)$ denotes the probability of detection failure if hypothesis H is true (e.g., presence of an attack), while $P(D = 1|H = 0)$ denotes the probability of a false alarm if hypothesis H is false (i.e., no attack), and p is the a priori probability that hypothesis $H = 1$ is true.

4 Results and Discussion

4.1 Application Case Study of AI Algorithms in Cybersecurity Protection and Assessment

We will discuss the practical application cases of AI algorithm in network security protection and evaluation. Through the case study, we have a deeper understanding of the application of artificial intelligence in various security fields, and improve the effectiveness and accuracy of protection.

(1) Case analysis of intrusion detection and protection

Case 1: Real-time Intrusion Detection Using Machine Learning

In a large financial institution, a set of IDS based on machine learning has been deployed. In this paper, a network traffic detection method based on deep learning is proposed. By detecting network traffic, possible malicious behaviours can be found. The system can accurately locate abnormal behaviours such as illegal data access and data transfer by learning the normal behaviours in the network.

The results of real-time intrusion detection using machine learning are shown in Fig. 1.

This picture shows the simulation effect of real-time intrusion detection using machine learning technology. The blue line in the figure represents normal network traffic, and its fluctuation mainly revolves around a central value, showing a relatively stable state. Red dots mark abnormal events in network traffic, which significantly deviate from the fluctuation range of conventional traffic and represent possible intrusion activities or other security threats.

In this way, IDS can monitor the abnormal changes of network traffic in real time, and find and respond to potential security incidents in time. Abnormal traffic points in the diagram are clearly highlighted, so that security analysts can quickly identify and take measures to counter possible threats. This visualization tool is very useful in actual network security operations. It not only helps technicians to track and analyse data flow, but also provides intuitive evidence to support further security decisions.

In addition, the time axis and the label of network traffic clearly show the time series characteristics of data, which is particularly important for analysing the dynamic changes of network traffic and pattern recognition. On the whole, such graph tools play a key role in strengthening the network security defence system, which makes the interpretation of complex data intuitive and easy to understand.

Case 2: AI-assisted cyber threat defence

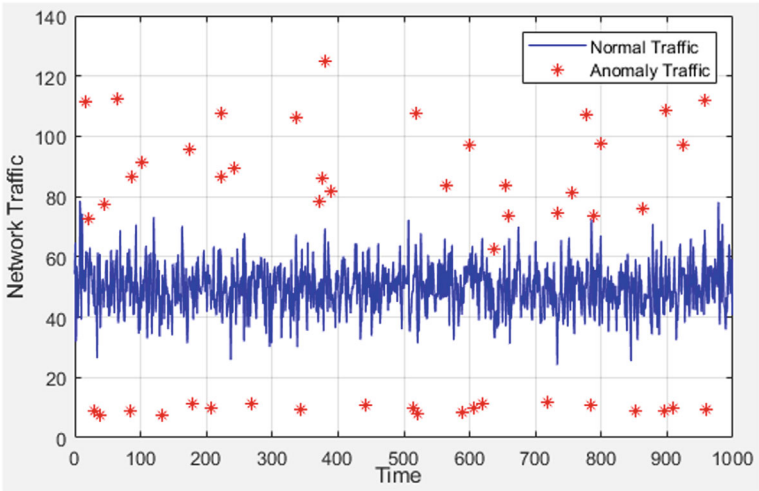


Fig. 1 Real-time intrusion detection results using machine learning

An international telecommunications company introduced AI algorithms to enhance its cyber threat defence capabilities. The company uses AI for behavioural analysis to detect and block phishing attacks. The AI system analyses the sending patterns and content of emails and successfully identifies features similar to previous phishing attacks, thus blocking a large number of phishing emails before users click on malicious links. In addition, the system is able to automatically adapt to new threat patterns to ensure that protection measures are always up-to-date.

The results of AI-assisted cyber threat defence are shown in Fig. 2.

Figure 2 shows an email traffic monitoring scenario with AI-assisted cyber threat defence. In this figure, we can see two different coloured bars: blue represents normal email traffic, while red highlights emails identified as phishing attempts by the AI system.

Firstly, the variation in the height of the blue bar shows the normal volume of mail at different points in time. Such fluctuations in mail volume can be influenced by the daily work cycle, such as the difference between weekdays and weekends or the peak and trough times of the day.

The red bars, on the other hand, are significantly higher than certain blue bars, indicating that the AI system detected abnormal email activity at certain points in time, and these were flagged as phishing attempts. This highlighting visually reveals potential cybersecurity threats, allowing security teams to quickly identify and address them.

(2) Information security assessment case study

Case 3: AI-driven security assessment platform

A professional cybersecurity service company has developed an AI-based security assessment platform that provides users with in-depth security analysis reports.

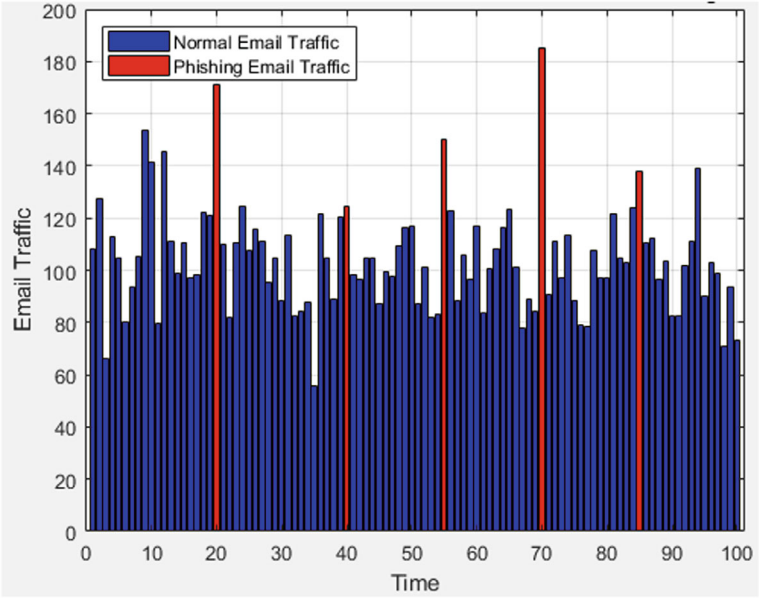


Fig. 2 AI-assisted network threat defence results

This paper proposes to automate the analysis of massive security logs and identify and predict possible security threats by fusing several machine learning models. For example, it can predict which system will be attacked in the future by analysing historical information. The launch of this system has greatly enhanced the risk management of several organizations, helping them to better allocate security resources and reduce possible financial losses. The results of the AI-driven security assessment platform test are shown in Fig. 3.

This picture shows how an AI-driven security assessment platform tracks and scores the dynamic changes of different types of network security threats, covering a whole year. In this visual way, we can intuitively see the score fluctuation of three main threat types: DDOS (distributed denial of service) (red line), phishing (blue line) and malware (green line).

It can be observed from the figure that the score of DDoS threat is highly volatile, which may be related to the uncertainty and large-scale attack mode. Especially in some peak periods in the figure, it shows that a serious DDoS attack may have occurred. The score of phishing attacks is relatively stable, but it shows a slight periodic increase in a specific period, which may reflect that attackers have increased the frequency of phishing attempts at a specific time (such as holidays or special events).

The score of malware shows moderate fluctuation, which reflects the constant change of malware spread and activity. The fluctuation of the threat score of malware

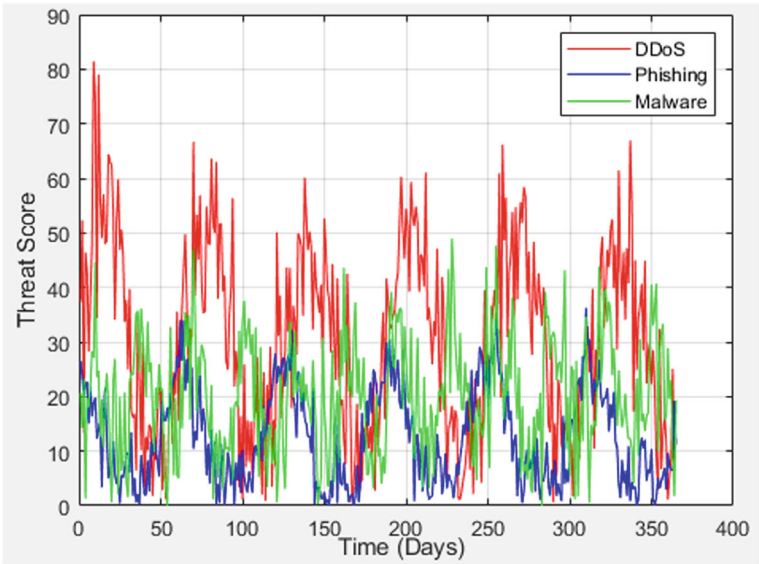


Fig. 3 Test results of AI-driven security evaluation platform

in the graph may be related to the emergence of new malware variants and the reduction of old variants.

On the whole, this diagram effectively uses colour coding and line change to distinguish different types of security threats, so that security analysts can capture the trends and trends of various threats at a glance. Such a visualization tool is very important for quickly identifying potential high-risk periods and taking corresponding defensive measures. It helps analysts make more accurate judgments and responses in complex security environments.

Case 4: AI in privacy protection assessment

In a large healthcare group, in order to protect patients' personal information, the group uses AI for data privacy assessment. The AI system is able to automatically detect potential data leakage risks and analyse data access patterns to identify anomalous data access behaviour. Through real-time monitoring and assessment, the system helps healthcare organizations identify and respond to data leakage risks in advance, enhancing data protection.

The results of the application of AI in privacy protection assessment are shown in Fig. 4.

In this figure, we observe the dynamic relationship between the number of data access requests and the privacy risk score for each day in a month, and the figure clearly shows the interaction between data access behaviour and privacy risk.

The bar chart, in blue, is used to depict the number of data access requests per day. As can be seen, these requests fluctuate throughout the month, but generally show

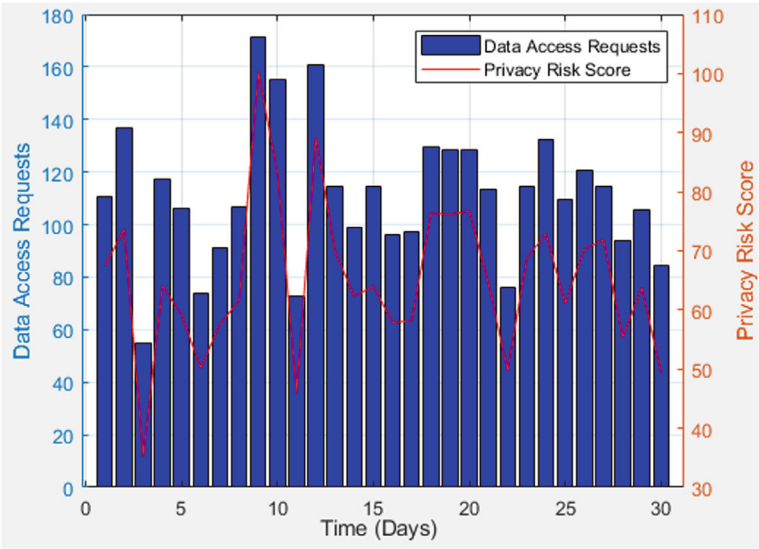


Fig. 4 Results of AI application in privacy protection assessment

some regularity, which may reflect differences in access patterns between weekdays and weekends or spikes in data access due to specific events.

The line graph, on the other hand, is shown in red to represent the privacy risk score, which increases with the number of data access requests, especially on days when there is a sudden increase in the number of requests, with a corresponding significant spike in the privacy risk score. This trend reveals a direct correlation between data access density and the risk of privacy breaches, emphasizing the importance of adopting enhanced privacy protection measures during periods of high access.

(3) Other cybersecurity application cases

Case 5: Application of AI in cybersecurity education

A cybersecurity training institution uses AI to simulate cyber-attack and defence scenarios to provide students with hands-on training. With the simulated environment generated by AI, trainees can learn attack techniques and defence strategies in a secure virtual environment.

Data on the use of AI in cybersecurity education is shown in Table 1.

This data table provides a detailed view, showing the learning effect of students after participating in the AI-driven network security education platform. By analysing the initial and final scores of students, we can intuitively see everyone’s progress in a specific network security course module.

As can be seen from the table, all the participants have made remarkable progress during their study. For example, the performance of student number 1 in the phishing recognition course has improved by nearly 42%, which shows that the course has effectively helped her master the skills of recognizing and responding to phishing

Table 1 Data on the use of AI in cybersecurity education

Student number	Course module	Initial rating	Final rating	Progress percentage (%)
1	Phishing identification	60	85	41.67
2	Malware analysis	70	90	28.57
3	Data leakage protection	55	75	36.36
4	System vulnerability scanning	50	80	60.00
5	Application of encryption technology	65	95	46.15

attacks. This improvement not only reflects personal efforts, but also shows the effectiveness of the education platform in content delivery and skill training.

Similarly, the performance of student number 4 in the system vulnerability scanning course has improved by 60%. This remarkable improvement may be related to the practicality of the course content and the personalized learning paths provided by the AI platform. These paths can adjust the difficulty and depth according to the initial performance of the students, and ensure that each student can be strengthened in his own weak areas.

Student No.5's score in the encryption technology application module increased by nearly 46%, reaching a high score of 95 points, which shows that he not only mastered the core concepts required by the course, but also skillfully applied these concepts to solve practical problems. Such achievements show that the AI education platform is very effective in providing in-depth and detailed technical knowledge.

Generally speaking, this data table not only shows the concrete progress of each student, but also indirectly reflects the potential of AI education platform in providing customized learning experience, evaluating students' ability and progress and strengthening targeted teaching. Through such a platform, students can improve their professional skills more efficiently and pertinently, and at the same time provide valuable feedback for education providers to help them continue to optimize curriculum design and teaching methods.

Case 6: Using AI for software supply chain security management

In the face of increasingly sophisticated software supply chain attacks, a technology company uses AI to monitor and manage the security of its supply chain. By analysing its partners' security posture and historical security events, the AI system was able to identify potential security vulnerabilities and suggest fixes. In addition, the system monitors security events in the supply chain in real time and responds quickly to possible security threats, ensuring the continuity and security of the company's operations.

Example data for software supply chain security management using AI is shown in Table 2.

In the case of Alpha Tech, for example, whose biggest risk is vulnerability, AI suggests improvements such as updating software libraries and enhancing code

Table 2 Example data for software supply chain security management using AI

Supplier number	Supplier	Product type	Risk score	Risk categories	Improvement measures
1	Alpha Tech	Software development	85	Loophole	Update software libraries and enhance code auditing
2	Beta Solutions	Cloud services	78	Configuration error	Introducing automated configuration management tools
3	Gamma Systems	Network equipment	82	Malicious code	Increase endpoint protection and conduct regular security scans
4	Delta Hardware	Hardware components	75	Loophole	Implement stricter third-party component testing and update firmware
5	Epsilon Inc	Data storage	80	Data leakage	Strengthen data encryption measures and train employees on data security awareness

auditing. These measures can address the root cause of the system’s vulnerability and improve its security performance. Similarly, the problem of configuration errors in Beta Solutions can be effectively solved by introducing automated configuration management tools, which reduces human error and improves the security of the entire cloud computing service.

The Gamma system is facing problems with malicious code and the suggested improvement methods include increasing endpoint protection and regular security scans. This method can detect and solve the security risks in the system early and avoid the harm of malicious code to the system. Epsilon Inc. is facing the danger of data leakage, what they need to do is to strengthen the data encryption and raise the awareness of employees about the security of data.

Overall, proposing targeted improvement measures can both effectively identify risks and provide practical solutions to improve the security level of the entire supply chain. Through the research of this project, a new idea of data-driven supply chain security management is proposed to provide a new solution for enterprises.

4.2 Suggestions for Improvement Measures

In order to solve the current problems and promote the benign development of AI in network security, this paper proposes the following countermeasures:

Strengthening data management: for the training and application of AI, formulating strict specifications for data collection, processing and use to ensure the quality and safety of data.

Enhancing model transparency and interpretation: researching new algorithms and tools to enhance the transparency of models, making the decision-making process of AI more easily monitored and audited.

Enhancing the security of AI systems: developing specialized AI system protection techniques, such as adversarial attacks, to ensure the robustness of AI systems.

Cross-disciplinary collaboration: conducting joint research on the problems encountered by AI in cybersecurity and promote the formulation of relevant standards and norms through a collaborative mechanism at multiple levels, including government, academia and industry.

5 Conclusion

This paper provides an in-depth analysis of the application of AI technology in cybersecurity, the challenges it faces and the future development direction. It shows that although AI has obvious advantages in improving the ability and efficiency of cybersecurity protection, it still faces difficulties in data quality and model transparency. In addition, the security of AI systems themselves cannot be ignored, and new threats such as adversarial attacks pose serious challenges to the security of AI systems.

In order to cope with the above challenges, this paper will focus on the cross-collaboration mechanism between AI systems in terms of strengthening data governance, improving model transparency and enhancing the security of AI systems. In the long run, AI will be used more in the field of cybersecurity. In addition, with the continuous development of quantum computing technology, some emerging research fields such as quantum secrecy and quantum cryptography will provide new ideas for information security problems, and AI technology will become an important development trend of information security. In the future, this paper will further improve the transparency and interpretability of AI models so that they can be better used in the field of network security.

References

1. Pai X (2023) A ship network security control system based on artificial intelligence technology. *Ship Sci Technol* 45(3):153–156
2. Shualing W, Xing L, Ligen H (2020) Construction and evaluation of hospital information security system based on three-level protection. *Chin Med Equip* 35(11):142–145
3. Rongsheng X, Wei N, Yi G (2023) Research on computer network information security and protection strategies. *Electron Commun Comput Sci* 5(6):52–54
4. Olasehinde OO (2020) A stacked ensemble intrusion detection approach for the protection of information system. *Inf Secur Int J* 10(1):910–923
5. Cahyadi EF, Su TW, Yang CC et al (2022) A certificateless aggregate signature scheme for security and privacy protection in VANET. *Int J Distrib Sens Netw* 18(5):399–410

6. Zwillling M, Klien G, Lesjak D et al (2022) Cyber security awareness, knowledge and behavior: a comparative study. *J Comput Inf Syst* 62(1):82–97
7. Xing K (2023) A practical study of big data technology in computer network information security processing. *J Electron Res Appl*. 7(6):36–41
8. Al-Fedaghi S, Alnasser H (2020) Modeling network security: case study of email system. *Int J Adv Comput Sci Appl* 11(3):99–107
9. Guangxu Y (2021) Research on computer network information security based on improved machine learning. *J Intell Fuzzy Syst: Appl Eng Technol*. 40(4):6889–6900
10. Shen G, Li X (2023) A fuzzy group decision-making framework for computer network security evaluation with probabilistic linguistic information. *Int J Knowl-Based Intell Eng Syst* 27(3):355–365
11. Gadallah WG, Omar NM, Ibrahim HM (2021) Computer network and information security. *Int J Comput Netw Inf Secur* 13(3):15–27

An Optimal EEG Features Extraction Methods for Autism Detection



Hajir Ammar Hatim and Zaid Abdi Alkareem Alyasseri

Abstract Based on the BCIAUT-P300 dataset, we evaluate the classification intended to diagnose autism using 22 feature extraction methods for EEG signals. The classifiers MDM, SVM, NN, RF, and LDA were examined in relation to PSD, CWT, PCA, and FDC to evaluate onset detection algorithms. Various ensemble learning techniques like AdaBoostM1, Bagging, and GentleBoost were applied to improve the classification performance. Of all the PSD methods, PSD + AdaBoostM1 came out as the best with 96.80% accuracy, 96.88% sensitivity, and 96.00% specificity achieved with this method. The FDA results were also very good, with an accuracy of 96.56% ($F1$ -Score: 98.04%). CWT, PCA, and EEGNet also achieved great performance, but the time-domain method like Shannon Entropy and decomposition methods such as HHT had limited accuracy and $F1$ -scores. The study demonstrates that frequency domain methods and ensemble classifiers can effectively distinguish P300 responses for autism detection. The findings highlight that choosing the right feature extraction methods is important for better classification performance of EEG-based diagnosis systems.

Keywords EEG · ASD · Feature extraction · Machine learning

H. A. Hatim · Z. A. A. Alyasseri (✉)

Department of Computer Science, Faculty of Education, University of Kufa, Najaf, Iraq
e-mail: zaid.alyasseri@uokufa.edu.iq

H. A. Hatim

Information Technology Research and Development Center (ITRDC), University of Kufa, Najaf, Iraq

Z. A. A. Alyasseri

College of Engineering, University of Warith Al-Anbiyaa, Karbala, Iraq

1 Introduction

Autism is a cause for concern as it manifests difficulties in interaction with other people, lacks communication skills, and involves a pattern of stereotyped behavior. [1–3].

Electroencephalography (EEG), has been successful in identifying certain subtypes of neurological disorders, including autism. However, a challenge arises because the basic EEG signal structure can be affected under controlled conditions [4, 5].

Numerous feature extraction methods have been explored to improve the accuracy of ASD detection using EEG to address this issue. Despite significant progress, determining the most suitable feature extraction approach for ASD diagnosis remains challenging and unresolved [5, 6].

2 Protocol for Selection EEG Features Extraction Methods Related Work

EEG and P300 signals are regarded as the most frequently used tools in assessing patients with autism spectrum disorder (ASD) relating to neural activities. The BCIAUT P 300 data set has been previously analyzed to classify patients' responses using machine and deep learning methods. In this segment, we offer a short narration of the most important studies that have employed the same or similar dataset as this study and different approaches. In a prior study [7], temporal mapping of the signals into a subspace using subspace projections enhanced the detection of the P300 by applying Independent Component Analysis (ICA). As far as it was utilized in autism studies, ICA has made it possible to improve the P300 response classification by enhancing the classification of the signal and signal-to-noise ratio.

Riemannian geometry was used to process and model the EEG data. Combining the Riemannian geometry with ensemble learning models increased the accuracy of classification models to 81.2% of P300 signals detection, which is a significantly better result than conventional CSP-based approaches [8].

Several studies have used Principal Component Analysis (PCA) to reduce the dimensionality of EEG data extracted from P300. PCA assists in enhancing the classification performance by noise filtering in the data. In a more recent study, however, PCA has been integrated with other methods and used on the same dataset to enhance the accuracy of the classification. The Continuous Wavelet Transform (CWT) was implemented to identify features in the time–frequency domain from P300 data. This technique has been used in studies to enhance the performance of EEG analysis in patients with ASD, as it helped improve P300 response classification accuracy. The two methods were used in a study by [9], where an accuracy of 82% was attained.

In another study [10], PCA was incorporated with CNN to analyze P300 auditory evoked potential signals. PCA helped, among other things, reduce dimensionality and increase the speed of data processing and feature extraction. On the other hand, CNN helped facilitate the accuracy of the classification. This study demonstrated that over 90% of classification accuracy can be achieved with this framework, suggesting the superiority of conventional and deep learning approaches used in combination.

Another state-of-the-art study [11] implemented the deep neural networks Long Short-Term Memory, abbreviated as CNN-LSTM, which combined convolutional and recurrent neural layers to analyze P300 signals. This technique was proficient in finding complex temporal patterns while analyzing the P300 responses; it also enhanced the final classification accuracy to a high value of 84.3%.

A recent study by [12] introduced deep learning architecture for EEG data; EEGNet was thus put to work on raw EEG data processing and patient analysis through deeper learning. The aggregated accuracy of classification made against the BCIAUT-P300 dataset is 92.47%. By building upon the prior art, this work embraces a convolutional neural network for the processing of the P300 response, aiming to bring improvement in measuring the attention of ASD patients.

It has also been proposed that VMD can be used to analyze nonlinear and non-stationary signals like EEG. Another recent study [13] divided signals into five modes and analyzed them using temporal and frequency features. The proposed method achieved a classification accuracy of 91.12% using the Support Vector Machine algorithm by classifying P300 signals in ASD patients.

This paper will compare the performance of 22 EEG feature extraction techniques implemented in MATLAB to find the best-performing methods for ASD diagnosis. This contribution will help us understand which methods best balance computational efficiency and diagnostic accuracy.

3 EEG Autism Datasets

The BCIAUT-P300 dataset was created to serve as a benchmarking tool for BCI research focusing on patients with ASD. The main motive of this work was to enable researchers to make sense of P300 signals extracted from participants with ASD and help improve the understanding of neural patterns that comprised the disorder. The dataset has hence been made available on Kaggle: BCIAUT_P300 (kaggle.com).

The dataset includes 15 participants with an ASD diagnosis. Each participant attended seven experimental sessions, for a total of 105 sessions. In this experimental session, EEG recorded P300 responses based on specific visual stimuli.

The data collection system is based on a brain-computer interface that continually analyzes a certain type of signal known as P300, an event-related potential produced by presenting an infrequent target stimulus within a series of frequent standard stimuli. This system assesses participants' interactions with the stimuli.

EEG data were recorded using eight active electrodes positioned according to the international 10–20 system (C3, Cz, C4, CPz, P3, Pz, P4, POz). Figure 1 shows the

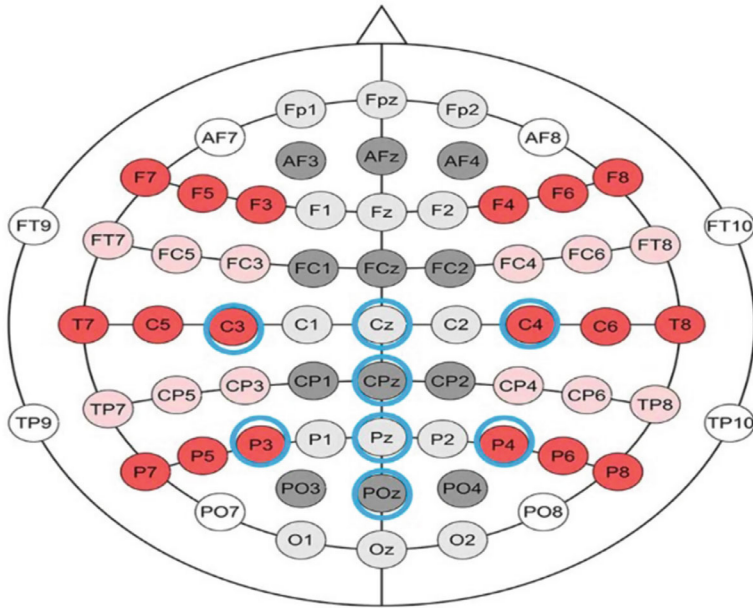


Fig. 1 Eight electrodes are used to record EEG signals

sampling rate was set at 250 Hz, and the data were filtered using a passband filter between 2 and 30 Hz and a 50 Hz notch filter to remove electrical noise.

Subjects performed an interactive attention task. During the experiment, a series of visual stimuli were presented, and the participants identified rare target stimuli appearing within sequences of non-target stimuli. This experimental paradigm was implemented to elicit P300 responses when participants focused on the target stimuli.

The data from each session were divided into two phases: A training phase (Calibration) and a testing phase (online). During the training, 20 blocks of 10 trials each were presented, providing 200 target P300 signals and 1400 non-target signals. During online sessions, 50 blocks were presented, with the number of trials per block dependent on individual performance during the calibration/training phase of the experiment.

This dataset was structured into participant-specific folders, with each session containing training and testing files in .mat format, supported by text files with event and trial information. It can further exploit P300 responses to the advantage of developing AI-based models that recognize neural signals associated with ASD.

3.1 Pre-processing

Pre-processing the EEG signal is a must to filter out the relevant features, among which the P300 component plays an important role in BCI-related cognitive analysis. Several literature findings indicate that individuals with autism show different P300 responses in sensory information processing compared to non-autistic individuals.

We applied filtration through a bandpass in the range of 0.5–30 Hz, which is considered to be the optimal range, which would assure us that the key components of EEG, such as P300, according to [14, 15]. By this step, the relevant frequencies and noise are removed, and the quality of the signal will improve effectively.

Baseline correction followed from – 0.3 to 0.0 s to remove signal drift and accurately extract the P300 component. Both pre-processing steps, such as bandpass filtering and baseline correction, have relevance to optimize P300 signal detection, which is a critical marker for cognitive processing in EEG data.

Figure 2 shows two major pre-processing steps with our approach. These include applying the bandpass filter to isolate critical frequency components and removing the drift using baseline correction. These steps optimize the P-300 component’s detectability—a key marker of cognitive processing—and its accurate capture in EEG signals.

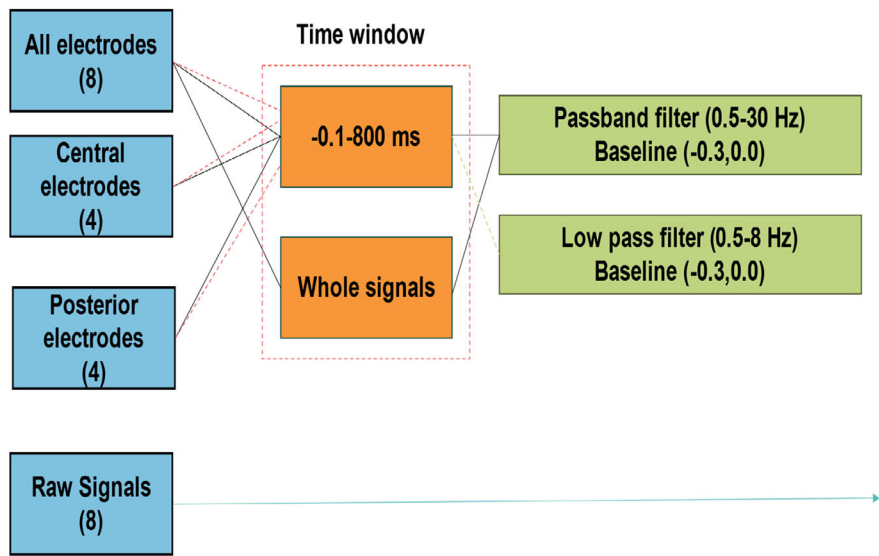


Fig. 2 EEG signal pre-processing workflow

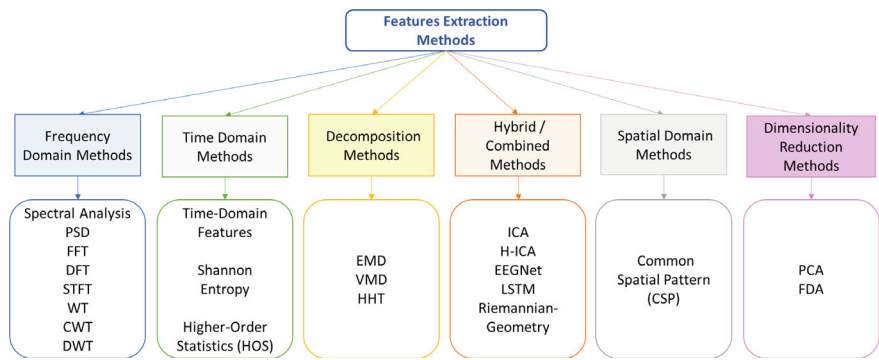


Fig. 3 Feature extraction methods

4 Feature Extraction Methods

In EEG signal analysis, feature extraction procedures are vital in identifying and capturing relevant information from the raw EEG signals. These feature extraction techniques reduce data complexity while retaining the most critical features for accurate analysis and classification. The idea is to transform original EEG data into manageable form by concentrating on features that highlight specific brain responses, such as the P300 components.

Depending on the kind of data and analysis goals, different approaches to feature extraction could be considered: Frequency domain, time domain, spatial, or a mix of several of these. The methods of feature extraction that will be used in this work are described in the following section, as shown in Fig. 3.

5 Frequency Domain Methods

Frequency domain methods shift the signal from time to frequency to analyze it. Approaches include spectral analysis, Spectral Power Density (PSD), Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT), Wavelet Transform (WT), Continuous Wavelet Transform (CWT), and Discrete Wavelet Transform (DWT) which compute features about frequency relevant to neural activity. These techniques give an interesting insight into capturing periodic patterns in EEG signals, like the P300 response.

Method	Description	Equation
Spectral analysis	A method of analyzing the frequency constituents of such signals as EEG, which may lead to easier identification of neural patterns	$(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$

(continued)

(continued)

Method	Description	Equation
PSD	It quantifies the power of the signal in various frequency bands. PSD is frequently used in the analysis of EEG signals, which serves to point out the dominant frequency components within the signal	$P_{xx}(f) = \frac{1}{T} \lim_{\infty \rightarrow T} E[X_T(f) ^2]$
DFT	The algorithm is a mathematical procedure that transforms time-domain signals to their constituent frequencies and is useful in signal analysis such as EEG	$X(\kappa) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi kn}{N}}$
FFT	Improved algorithm that calculates DFT, with far less computation hence nightly faster for large data sets such as EEG	$X(\kappa) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi kn}{N}}$
STFT	This method analyzes the non-stationary signal by segmenting it into a number of short windows; thus, both time and frequency analysis can be attained	$\text{STFT} = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j\omega\tau}d\tau$
WT	A time–frequency decomposition method that will decompose a signal into such time–frequency components: It has a time localization and is well-suited for non-stationary signals such as EEG	$\text{WT}_x(a, b) = \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt$
CWT	It is a time–frequency locality method, ideal for non-stationary signals like EEG; it continuously analyzes them in time and scales	$\text{CWT}_x(a, b) = \frac{1}{\sqrt{ a }} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-b}{a}\right)dt$
DWT	Decompose a signal at different scales and simultaneously capture time and frequency localization. It is thus useful for the detection of some specific time-varying components, such as the P300 wave in EEG signals	$W(a, b) = \frac{1}{\sqrt{ a }} \sum_{n=0}^{N-1} x[n]\psi\left(\frac{n-b}{a}\right)$

6 Time-Domain Methods

Time-domain analysis concerns the signal’s temporal features, which generally address amplitude and temporal variations. Enveloping techniques such as Shannon Entropy and time-domain features provide insight into the time complexity of EEG signals by capturing fluctuations that occur over time. It could very well aid in identifying cognitive events from the EEG data.

Method	Description	Equation
Time-domain features	Simple yet powerful descriptors include mean amplitude, variance, peak-to-peak value, etc., which define the shape and characteristics of EEG signals over time	Mean amplitude (μ): $\mu = \frac{1}{N} \sum_{n=1}^N x[n]$ Variance (σ^2): $\sigma^2 = \frac{1}{N} \sum_{n=1}^N (x[n] - \mu)^2$ Peak-to-peak value $(P_{pp})P_{pp} = \max(x) - \min(x)$
SE	Quantifies the randomness or uncertainty of a signal, thus providing an idea of the complexity of EEG signals. Lower entropy could reflect a more predictable response of the brain	$H(X) = - \sum_{i=1}^n \log p(x_i)p(x_i)$
HOS	Statistical tools that will analyze the non-Gaussian and nonlinear characteristics of the signal, including skewness and kurtosis, thus giving more details about the nature of EEG signals beyond the mean and variance	Variance: $\sigma^2 = E[(X - \mu)^2]$ Skewness: $K = \frac{E[(X - \mu)^4]}{\sigma^4} - 3$ Kurtosis: $K = \frac{E[(X - \mu)^4]}{\sigma^4} - 3$

7 Decomposition Methods

Decomposing methods try to decompose complex EEG signals into a simple component to obtain better features. Considering such a viewpoint, EMD and VMD are conducted to decompose the signal into intrinsic mode functions, which may capture meaningful features masked in nonlinear or non-stationary terms.

Method	Description	Equation
EMD	It decomposes nonlinear, non-stationary signals like EEG into intrinsic mode functions without any a priori basis functions, hence adaptive to the complex signals	$x(t) = \sum_{i=0}^n \text{IMF}_i(t) + r(t)$
VMD	That reduces nonlinear signals into modes with specific frequency content, with the more important consideration being that modes do not overlap in the frequency domain, thus making it robust in analyzing P300 signals in EEG	$\min_{\{\omega_k\}, \{\omega_k\}} =$ $\left\{ \sum_k \left \nabla \left(\left[\delta(t) + j \frac{1}{\pi t} \right] * u_k(t) \right) e^{-j\omega_k t} \right ^2 \right\}$

(continued)

(continued)

Method	Description	Equation
HHT	It combines EMD with the Hilbert Transform for the processing of nonlinear, non-stationary signals. It has been applied in the extraction of instantaneous frequency and phase information in EEG, thus very useful in the identification of event-related potentials such as P300	$H(x(t)) = \frac{1}{\pi}P \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau}d\tau$

8 Hybrid/Combined Methods

Hybrid and combined methods are based on various feature extraction methods to provide comprehensive insights. Independent Component Analysis, Riemannian geometry, and EEGNet techniques implement many strategies to further improve feature extraction techniques by combining spatial, temporal, and frequency information. These methods will ensure a better way of feature extraction and are most effective in enhancing classification performance with multi-dimensional data.

Method	Description	Equation
ICA	It is a statistical technique employed in separating mixed signals into independent components, finding wide application in removing artifacts in the EEG signal, thus enhancing the detection of brain activity, especially the potential P300	$X = AS$
H_ICA	It is an extension of ICA with a hierarchical structure, allowing for dependencies between the components. Thus, complex EEG signals with strong dependencies can be separated more accurately	$X = AS$
EEGNet	A light deep learning architecture for spatial and temporal feature extraction of EEG signals, which is quite convenient to use on the detection of P300 responses	$y[i] = \sum_{k=0}^{K-1} x[i+k] \cdot w[k]$
LSTM	A variant of RNN that has the capability to model long-range dependencies in sequences, like EEG data, for event-related potential detection, such as P300 events	$c_t = f_t \odot c_{t-1} \odot +i_t \sim c_t$ $h_t = o_t \tan hc_t$
Riemannian geometry	It can also analyze the covariance matrices of EEG signals as points on a Riemannian manifold, allowing a robust method for discrimination among different brain responses, such as P300 or others	$dR(C_1, C_2) = (\sum_{i=1}^n (\log \lambda_i)^2)^{\frac{1}{2}}$

9 Spatial Domain Methods

Most of the techniques in the spatial domain mainly focus on extracting information from the spatial distribution of electrodes in the recording of EEG. Among them, CSP is one of the important methods through which optimization of spatial filters takes place to enhance the discriminability between different classes of brain responses and helps identify spatially discriminative neural activity patterns.

Method	Description	Equation
CSP	A method for ultra-spatial feature extraction of multi-channel EEG signals by maximizing the variance between two classes with wide applications in BCI	$W^T C_r W = \lambda W^T C_n W$

10 Dimensionality Reduction Methods

Dimensionality reduction aims to reduce the complexity of data into fewer dimensions while retaining the information therein. Techniques like Principal Component Analysis (PCA) and Fisher Discriminant Analysis (FDA) remove noise and redundant features, eventually paving the way for efficient classification since the data becomes simpler with no loss of accuracy.

Method	Description	Equation
PCA	It is a statistical method used to decrease the dimensionality of high-dimensional data, like EEG signals, focusing their attention on the main characteristics and eliminating noise, hence speeding up further analyses	$Z = XW$
FDA	This can be defined as a linear classification method that projects data onto a line while maximizing class separability. It is in common usage to discriminate between target and non-target EEG responses	$w = \Sigma^{-1}(m_2 - m_1)$

11 Results

This section presents the performance of 22 feature extraction methods applied to autism detection using the BCIAUT-P300 EEG dataset, which is assessed by five traditional classifiers: MDM, SVM, NN, RF, and LDA. Another essential analysis was performed using deep learning methods like EEGNet and LSTM represented as the feature extractors and classifiers. To this regard, the metrics of the performance that include accuracy, sensitivity, specificity, precision, and *F1*-score.

The analysis showed very high variation in performance among methods and classifiers. Among all, the PSD technique came out on top, yielding maximum accuracy of 96.80% with an AdaBoostM1 ensemble in addition to very high sensitivity of 96.88%, specificity of 96.00%, precision of 92.00%, and $F1$ -score of 95.32%. It was second best for CWT, which gave accuracy of 96.00% with a Bagging ensemble, indicating liveness regarding capturing features at multi-scales effectively. PCA and FDA also fared well, yielding 95.31% and 96.56%, respectively, hence useful in the reduction of dimensionality and extraction of features from the EEG data.

On the other hand, Shannon Entropy and HHT recorded a poor performance, both with an 87.50% accuracy, hence suggesting their possible inefficiency in the discrimination of EEG signal patterns to detect autism. Frequency domain. Thus, PSD outperformed the other methods on this frequency domain with a very satisfying performance compared to other techniques such as FFT and DFT. However, in the time domain, time-domain features produced the best performance with an accuracy of 90.00%, outperforming the Shannon Entropy by 87.50%. EMD and VMD attained a pretty reasonable performance among the decomposition methods with a classification performance of 93.44% and 92.00%, respectively. Hybrid and combined method: EEGNet achieved an accuracy of 93.75% as shown in Fig. 4, while Riemannian geometry, ICA, and H-ICA performed reliably well.

The spatial domain methods like CSP achieved 87.50%, reflecting the capabilities in delineating spatial differences present in EEG signals with high efficacy.

In general, the ensemble classifiers like AdaBoostM1 and Bagging boosted most of the feature extraction algorithms among which PSD, CWT, FDA, and PCA produced the best results as shown in Fig. 5.

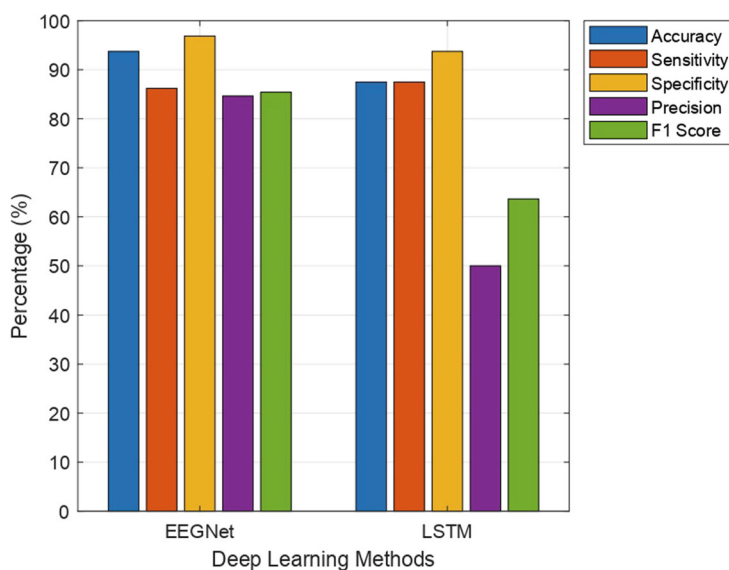


Fig. 4 Performance comparison of deep learning methods (EEGNet, LSTM)

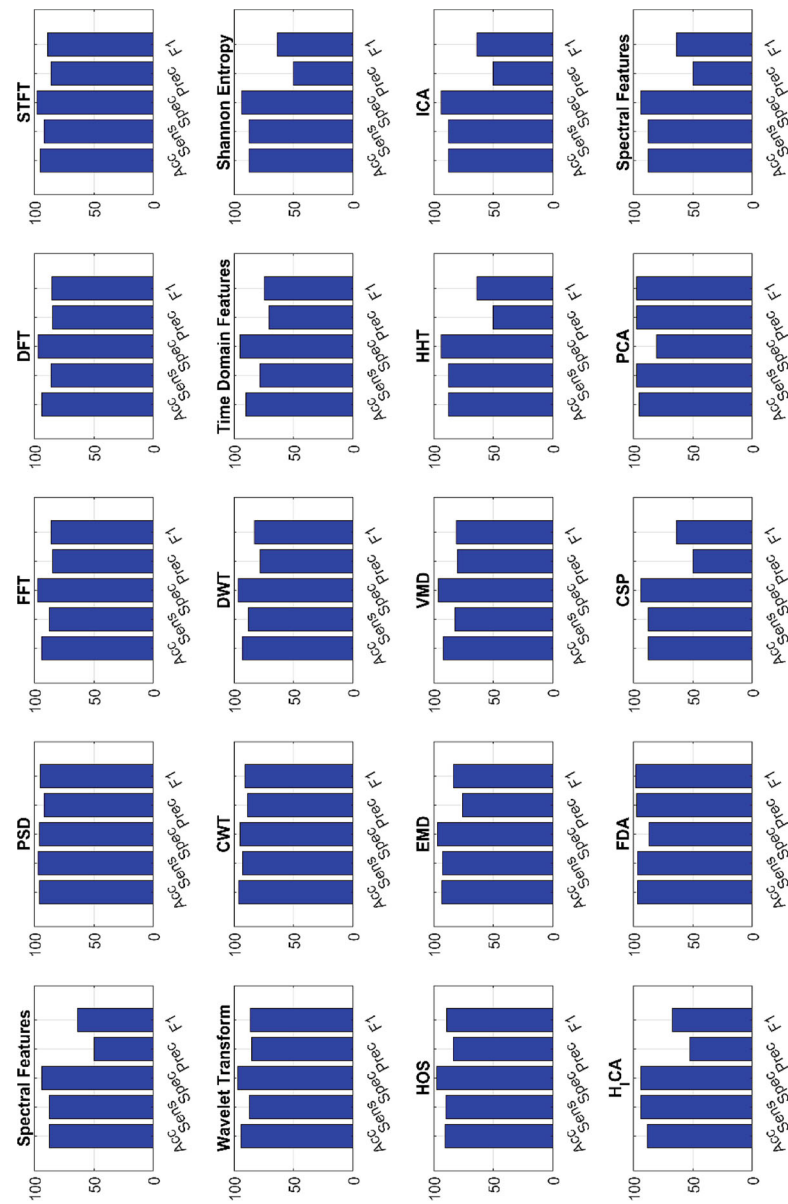


Fig. 5 Performance metrics for each feature extraction method

12 Conclusion and Future Direction

This review assessed 22 feature extraction methods applied to the BCIAUT-P300 EEG dataset for autism detection. Frequency domain techniques, especially PSD, performed best among the methods by yielding the highest accuracy of 96.80% when combined with the AdaBoostM1 ensemble technique. The ability of PSD was a key factor that captures dominant frequency components in EEG signals, which is vital for detecting P300 responses in autism cases. Other well-performing frequency-based methods include Continuous Wavelet Transform and Fast Fourier Transform, which may point out that the key to autism diagnosis based on EEG is embedded in the frequency features of those signals. In contrast, two time-domain methods, Shannon Entropy and Hilbert-Huang Transform, performed relatively poorly. That is probably because the EEG signal's complex and non-stationary nature is hard to describe fully by such methods. Dimensionality reduction was also successful, yielding the best classification accuracies of 95.31% for PCA and 96.56% for FDA using ensemble learning methods. These algorithms simplified the feature space by retaining high classification accuracy.

This research focuses on ensemble learning methods, including AdaBoostM1, GentleBoost, and Bagging, which have enhanced the classification performance. These methodologies yielded strong, dependable results from various weak classifiers, reducing the classification error significantly by a significant margin for various feature extraction methodologies. These findings supported previous work and underlined the efficiency of advanced feature extraction techniques in combination with ensemble classifier methods to increase the stability of higher accuracy in autism detection based on EEG data.

Future work should investigate integrating deep learning methods and extend the dataset to a wide variety of subjects under various a wide variety of subjects with different conditions. Further, the possible investigation of the interpretability of such models may provide profound insights into the neural mechanisms underlying autism and, therefore, an urge toward more precise diagnostic tools.

References

1. NIH, Autism spectrum disorder, National Institute of Mental Health. [Online]. Available <https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd>
2. WHO, Autism spectrum disorders, World Health Organization. [Online]. Available <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>
3. Campisi L, Imran N, Nazeer A, Skokauskas N, Azeem MW (2018) Autism spectrum disorder. *Br Med Bull* 127(1):91–100. <https://doi.org/10.1093/bmb/ldy026>
4. Chaddad A, Wu Y, Kateb R, Bouridane A (2023) Electroencephalography signal processing: a comprehensive review and analysis of methods and techniques. *Sensors* 23(14):1–27. <https://doi.org/10.3390/s23146434>
5. Singh AK, Krishnan S (2023) Trends in EEG signal feature extraction applications. *Front Artif Intell* 5. <https://doi.org/10.3389/frai.2022.1072801>

6. Yang J, Hu M, Hu Y, Zhang Z, Zhong J (2023) Diagnosis of autism spectrum disorder (ASD) using recursive feature elimination–graph neural network (RFE–GNN) and phenotypic feature extractor (PFE). *Sensors* 23(24). <https://doi.org/10.3390/s23249647>
7. Chumerin N et al (2009) P300 detection based on feature extraction in on-line brain-computer interface. *Lecture notes computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, vol 5803 LNAI, pp 339–346. https://doi.org/10.1007/978-3-642-04617-9_43
8. Krzemi D, Sebastian M, Matthias T, Lorena S (2020) Classification of P300 component using a Riemannian ensemble approach, vol 1, pp 1515–1525. <https://doi.org/10.1007/978-3-030-31635-8>
9. de Arancibia L, Sánchez-González P, Gómez EJ, Hernando ME, Oropesa I (2020) Linear vs nonlinear classification of social joint attention in autism using VR P300-based brain computer interfaces, pp 1869–1874. https://doi.org/10.1007/978-3-030-31635-8_227
10. Li F, Li X, Wang F, Zhang D, Xia Y, He F (2020) A novel P300 classification algorithm based on a principal component analysis-convolutional neural network. *Appl Sci* 10(4):1–15. <https://doi.org/10.3390/app10041546>
11. Santamaría-Vázquez E, Martínez-Cagigal V, Gomez-Pilar J, Hornero R (2020) Deep learning architecture based on the combination of convolutional and recurrent layers for ERP-based brain-computer interfaces, pp 1844–1852. https://doi.org/10.1007/978-3-030-31635-8_224
12. Borra D, Fantozzi S, Magosso E (2020) Interpretable and lightweight convolutional neural network for EEG decoding: application to movement execution and imagination. *Neural Netw* 129:55–74. <https://doi.org/10.1016/j.neunet.2020.05.032>
13. Peketi S, Dhok SB (2023) Machine learning enabled p300 classifier for autism spectrum disorder using adaptive signal decomposition. *Brain Sci* 13(2):315
14. Chatterjee B, Palaniappan R, Gupta CN (2020) Performance evaluation of manifold algorithms on a P300 paradigm based online BCI dataset. *IFMBE Proc* 76:1894–1898. https://doi.org/10.1007/978-3-030-31635-8_231
15. Zhao H, Yu S, Prinable J, McEwan A, Karlsson P (2020) A feasible classification algorithm for event-related potential (ERP) based brain-computer-interface (BCI) from IFMBE scientific challenge dataset. *IFMBE Proc* 76:1861–1868. https://doi.org/10.1007/978-3-030-31635-8_226

Web Scraping of Scientific Literature From SCOPUS Repositories of Iraqi Universities



Zahraa Taufeeq Abdul-sahib Al-Madhhachi and Salma A. Mahmood

Abstract The digital age dramatically changes academic research and data analysis. Web scraping techniques applied to scientific repositories with special emphasis on data extraction from SCOPUS-A comprehensive study. This study examined the significance of SCOPUS data for nine famous higher education institutions in Iraq. This task involves many steps, such as the selection of a suitable data source, the use of web scraping tools and techniques, data collection, preprocessing, and consideration of ethical issues. The current study aims to present descriptive quantitative and qualitative analyses to reveal patterns of publications, collaboration networks, areas of interest, and scholarly influence. The results provide a bespoke view of research movements of all these universities involved in this study and can use them to motivating and directing scientific research plans and offer joint projects and strategies for each. Despite these challenges, including ethical concerns and data problems, this study opens doors for further exploration of analytics, interdisciplinary studies, and data fusion. Finally, the study emphasizes the importance of web scraping as a means of extracting academic data and how it can enable higher education institutions and universities to achieve academic great potential and innovation in terms of scientific research.

Keywords Web Scraping · Data · Universities data · SCOPUS Repositories

Z. T. A. Al-Madhhachi

Computer Science, Iraqi Commission for Computer and informatics, University of Information Technology and Communication, Baghdad, Iraq
e-mail: ms202130674@iips.edu.iq

S. A. Mahmood (✉)

College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq
e-mail: salma.mahmood@uobasrah.edu.iq

1 Introduction

The acquisition and analytics of scientific data are critical in propelling scientific research and academic endeavors. In today's information age, these science repositories constitute treasures of knowledge that allow researchers at all levels (from postgraduate students to senior scientists) to tap into massive bodies of published literature and research outcomes [1, 2]. This group of repositories includes SCOPUS, which is considered an extensive and authoritative repository of scholarly content from different fields. Iraq is a country that places a high value on education and has a rich intellectual past and a strong creative spirit. The nation's numerous prestigious institutions have worked hard to push the boundaries of knowledge. Well-known universities include University of Basrah, University of Babylon, University of Mosul, University of Kufa, University of Information and Communications, Mustansiriyah University, Al-Nahrain University, Al-Iraqiya University, University of Technology, and more other. These higher education institutions do more than provide education; they are places of intellectual discovery and creativity, publishing a large amount of research papers which contribute meaningfully to the worldwide body of knowledge [3].

Over the past decade, there has been a growing need to access, analyze, and leverage scientific information and data from archives, such as SCOPUS. This is because of the different factors. One factor is that there is such an explosion of scientific literature that it's becoming difficult for scientists to keep up with what's going on in their specialty. Moreover, institutions around the world, including Iraq, are putting more effort into encouraging data-driven decision-making and encouraging a culture of research excellence. Third, the interdisciplinary nature of higher education institutions globally demands an overarching perspective on research trends and the total work in a field [4, 5].

The motivation behind this study was twofold. First, the method utilizes the potential of web scraping, an adaptive technique for content extraction as well as data gathering, to scrape the contents of scientific literature on the SCOPUS repository. It also aims to customize this analysis to the needs of Iraqi universities in terms of their priorities and interests. This way, this study strives to reduce the gap between the abundance of literature (in SCOPUS) and the needs of the higher education institutions and help the institutions to take well-informed decisions, keep track of the emerging trends, and effectively contribute to their respective disciplines.

The primary objectives of this study were as follows:

- Develop a reliable web scraping approach to scrape scientific data from SCOPUS repositories.
- Organize extraction data into categories that meet the research interest and priority for Universities in Iraq including University of Basrah, University of Babylon, University of Mosul, University of Kufa, University of Information and Communications, Mustansiriyah University, Al-Nahrain University, Al-Iraqiya University, University of Technology

- Doing an in-depth examination of the data to reveal research trends, prominent authors' works, and upcoming fields of academic study within these institutions.

This research contributes significantly to the literature analytics. First, we put in place a thorough methodological framework for scraping SCOPUS databases, a framework that could be helpful for any researcher interested to leveraging web scraping for their academic objectives. Second, two main outcomes were observed in this study. It illustrates the real-life benefits of web scraping in the domain of customized academic content acquisition for several Iraqi research-intensive universities and it reveals the scientific research landscape of these institutions, which helps them locate their strengths or weak points and possibilities of cooperation. This research provides a foundation for achieving the overall objective of promoting the utilization of academic outputs from the SCOPUS databases by extracting and analyzing data from Iraq's universities, thereby enhancing Iraqi University output and fostering interdisciplinary, as well as showing the power of web scraping as a flexible instrument for educational data analysis and its ability to help academic institutions flourish in a more data-infused future.

2 Web Scraping

Web scraping, data harvesting, or web data extraction is the process of collecting data from websites by using automatic machines. In our modern digital time, where the Internet has become a repository for countless data, web scanning is important tool to gather this valuable information strewn far apart on this cyberspace [6]. To summarize, it has not only its ability to speed up data gathering, but it also makes it important. This information can also be arranged to facilitate later analysis. Web scraping accelerates data-driven decision-making by automating data retrieval, and thus increases the efficiency of knowledge extraction, which eases the dependence on manual work [7, 8]. Web scraping is based on various information extraction (IE) methods that allow the extraction of particular information from different sources via the Internet, as illustrated in Fig. 1. Furthermore, they provide an encompassing way of gathering information from a myriad of sources on the web [9].

The HTML and CSS are highly reliant on presenting web content. With HTML, a markup language, the structure of content on websites, is determined through headings, paragraphs, links, and images. On the other hand, CSS controls how HTML elements look and are arranged. Effective web scraping requires basic comprehension of these tools [10]. The document object model (DOM) plays a significant role in web scraping and relies heavily on its availability in a browser. DOM is a visual representation of the hierarchy of a page through their interconnectedness [11]. The approach described in the following sections allows the systematic and automated retrieval and manipulation of digital resources according to a set hierarchy. Web scraping tools access information from different layers in the HTML structure by



Fig 1 Web scraping steps

navigating and interacting with the document object model (DOM). However, users can collect information from specific areas within a site [12].

Data extraction involves a collection of approaches that can be applied to specific data resources and content forms. Indeed, methods such as regular expressions and string manipulation are commonly used when parsing HTML for text information. In many cases, such as in structured data presented in table form, the extraction of information from HTML elements requires identifying patterns in rows and columns to identify the desired information [13, 14].

Handling dynamic content, especially when involving JavaScript, requires advanced techniques. In this approach, headless browsers mimic real user interactions and interact with dynamically generated information. Headless browsers can help collect real-time data using scraping tools [15, 16]. Web scraping is characterized by many ways through which data can be obtained from the web. Developing an efficient strategy for data extraction requires the skills of HTML, CSS, and DOM. Specific methods are also in place to deal with the varying data sources and formats encountered. Over time, web scraping has turned out to be important in many up-to-date studies that rely on data that involve different methods for accessing diverse resources existing in cyberspace [17, 18].

3 Methodology

The methodology followed in this study explains the approach taken in extracting and preprocessing scientific literatures from Scopus repositories based on the specific needs of the respective universities, as shown in Fig. 2. This practice is consistent with the principles of responsible data to guide responsible collection and use of data. Our methodology consists of the following steps:

3.1 Data Source Selection

The motivation of choosing SCOPUS as the primary data source for this study depends on several key criteria. First, Scopus is one of the most extensive and authoritative collections of scholarly literatures, covering all kinds of scientific areas. The breadth of coverage ensured access to the full range of published research papers, reflecting the multidisciplinary nature of these higher education institutions. Second, SCOPUS has the advantages of having a highly structured and standardized data structure, so, it allows an easier way for data extraction and data preprocessing. This approach permits the division and ordering of information according to the research interests and needs of each university. Scopus contains strong mechanisms for data access/retrieval, which guarantees the reliability and uniformity of the data extraction. Finally, its emphasis on quality control and meticulous peer review guarantees even greater credibility of the information derived.

3.2 Web Scraping Tools and Strategies

To obtain data from the SCOPUS database, a mixture of web scraping tools and libraries was used. The journey starts with the use of Python library web scraping tools such as Beautiful Soup and Scrapy. Beautiful Soup is useful for parsing HTML and

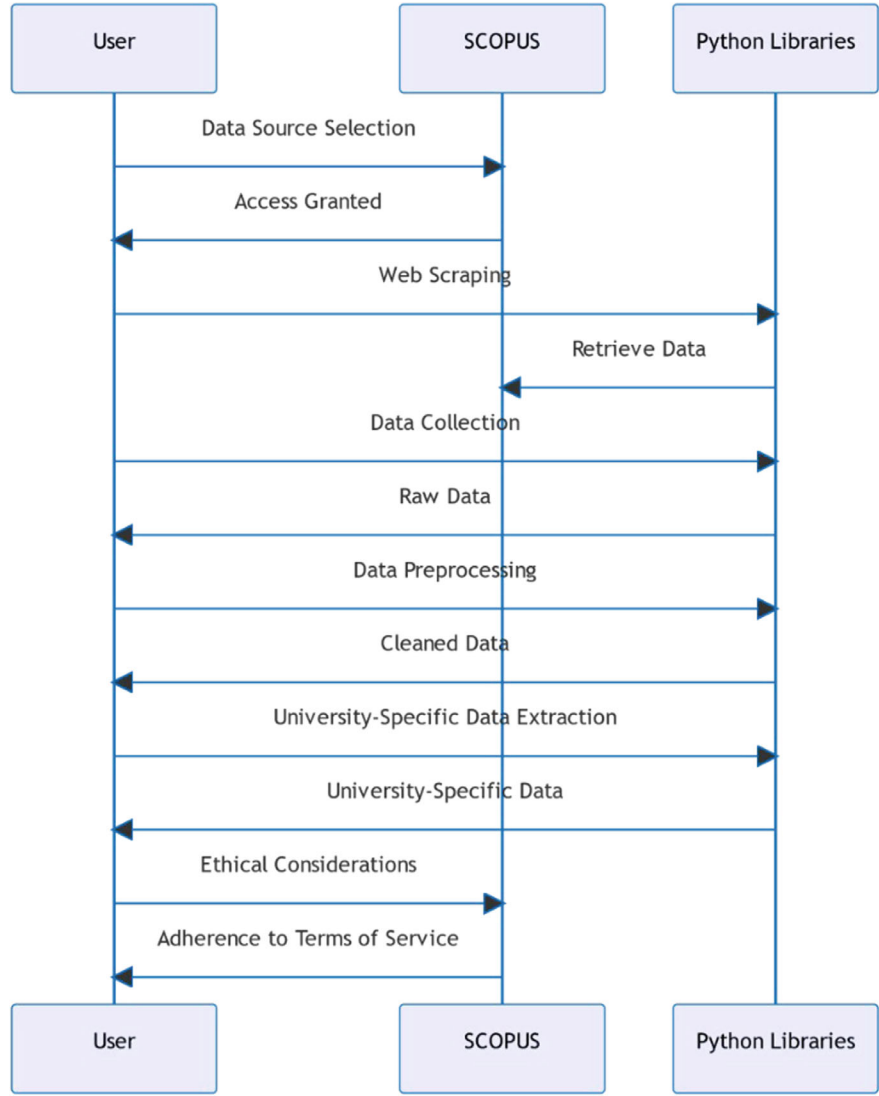


Fig. 2 Methodology steps

DOM navigation, which makes the extraction of specific data items from webpages breeze. As opposed to BeautifulSoup, which enables manual web content extraction, Scrapy is a framework for web crawling. In addition to these libraries, XPath queries are used to scrape data with exactitude. This method allows the selection of HTML elements according to their structure or attributes, thus selecting only what you want.

3.3 Data Collection and Preprocessing

Data collection encompasses the utilization of the SCOPUS API, the exploration of retrieved search results, and the extraction of metadata and full text from scientific papers. The raw data underwent intricate preprocessing, which involved the removal of HTML tags and the implementation of necessary adjustments to ensure data consistency and presentation. Furthermore, the text datasets were subjected to processes such as tokenization, stemming, and stop-word removal. Tokenization refers to the division of text into individual words or phrase chunks, thereby enabling detailed analysis. Stemming, on the other hand, reduces words to their base forms, thereby eliminating redundancies. Lastly, the removal of stop words eliminates commonly used words that hold little or no semantic meaning, thereby augmenting the value of the textual data.

The search results are initiated across the expansive domains of the SCOPUS repository, as depicted in Fig. 3.

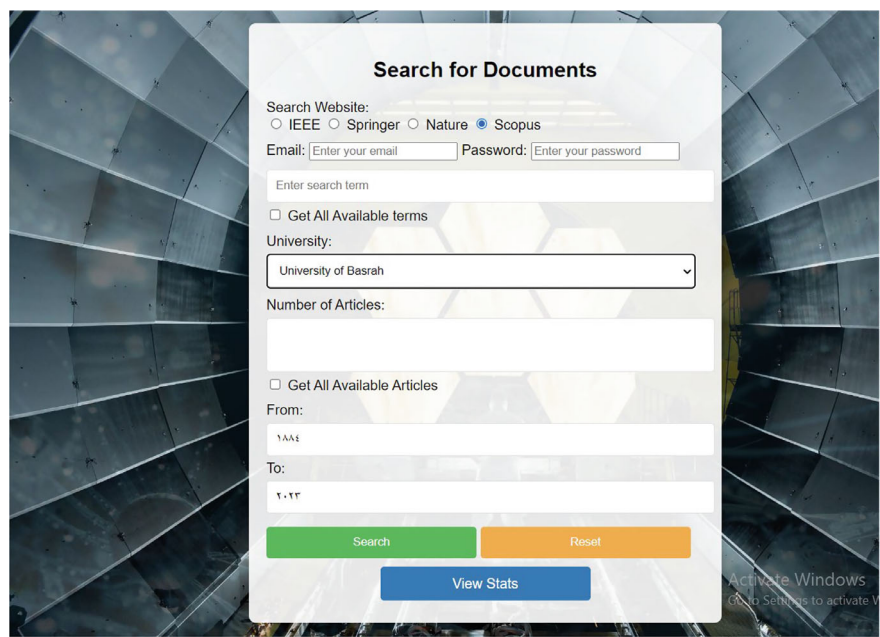
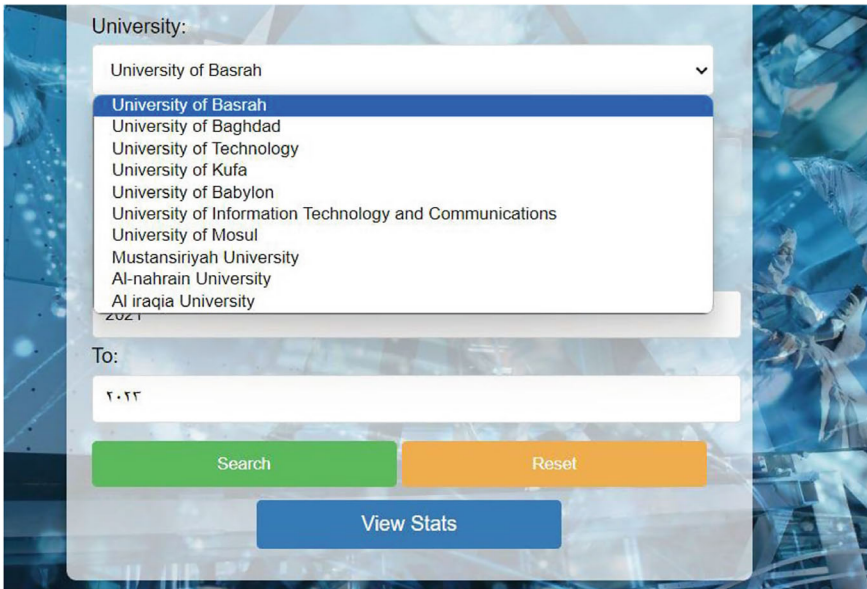


Fig. 3 Repositories selection stage



University:

University of Basrah

University of Basrah

University of Baghdad

University of Technology

University of Kufa

University of Babylon

University of Information Technology and Communications

University of Mosul

Mustansiriyah University

Al-nahrain University

Al iraqia University

2023

To:

2023

Search

Reset

View Stats

Fig. 4 University selection stage

3.4 University-Specific Data Extraction

To fulfill the different requirements of the specified universities, a university-specific data extraction approach was implemented. This approach is based on filtering and clustering of the extracted data according to predetermined criteria, for example, the authors' university/company affiliation or the topic of their research. Through this process, the analysis of research data was personalized and adapted to fit the specific research landscape and objectives of each university, making it relevant and usable.

The technicality is creating scholarly literature from raw information. This is an all-embracing process that encompasses exploiting what the SCOPUS repository has got in stock, searching for information, and eventually fine-tuning the crude output into each university as shown in Fig. 4.

3.5 Ethical Considerations

The highest priority in this investigation was adherence to ethical principles and terms of service in the dataset. Data were extracted only from public SCOPUS data sources in accordance with SCOPUS terms of use. Safeguards are put into place to avoid interrupting SCOPUS and against breach of copyrights and intellectual property

rights. Furthermore, all possible ethical issues (e.g., user privacy) are preempted through proper data handling for the entire research.

4 Data Analysis

The data analysis section was divided into two key components: quantitative and qualitative analyses. Such an analysis provides key insights into scraped information, with statistics for trends and qualitative assessment of its academic value.

4.1 *Quantitative Analysis*

The quantitative analysis includes a complete numerical analysis of the scraped data. Data are processed through statistical methods that help uncover information from the datasets, revealing the different characteristics of the dataset. The quantitative analysis of the data encompasses several dimensions, including but not limited to:

- Publication trends: A review of public dynamics over time, demonstrating the increasing quantity of scholarly output and the rise in topical interest in fields.
- Collaboration networks: Discovering collaboration patterns between researchers and institutions to uncover the strength of the academic network in each country.
- Keyword frequency: Identifying keyword/keyword clusters/research themes to identify the faculties and fields of study to which they belong.
- Citation analysis: Citation analysis—measuring the impact and influence of research and identifying influential papers and authors.
- Authorship patterns: Analyzing authorship patterns, such as prolific authors, author affiliations, multidisciplinary collaboration.

4.2 *Qualitative Analysis*

Qualitative analysis completes the quantitative analysis by looking beneath the surface of numbers for what lies in the qualitative aspects of the scraped data. This requires a more subtle analysis of the subject matter, finding threads, methodological approaches, and background knowledge that might not manifest in numerical data alone. Qualitative analysis included the following:

- Thematic analysis: Revealing the hidden topics of studies and research fields.
- Methodological assessment: Analyze research papers to determine which methods and procedures have been used to identify new or promising methods/trends.
- Contextual insights: Delineating academic contributions, their importance, and potential applications.

- Academic relevance: Reviewing the potential academic relevance and possible influence of scraped data on a specific university's research goals and focuses.
- By combining quantitative and qualitative methodologies, an in-depth analysis of the nature and trends of data, as well as their possible implications, is made possible.

5 Results

The results derived from the analysis offer specific insights for each of the mentioned universities, as shown in Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14. It details quantitative and qualitative data on the university's academic landscape, outlining their respective scientific research profiles in both aspects. The key highlights of the results are as follows.

- **University of Baghdad (11,925 records):** The university has an impressive academic legacy, as depicted by a six-decade historical dataset for the University of Baghdad. As shown in Fig. 5, analyzing these data could provide clues to what the university is set to contribute by way of long-term research. These records show continuous scientific work from to 2015–2018 years. This will reveal other research topics and associations applicable to this period. For the period of 2019–2020, as indicated by the dataset, it continues to be academically effective. These data can offer valuable information on contemporary research trends. The current year 2023 dataset contains very fresh material. These data may provide information on the areas of specialization in which the university is currently engaged or emerging trends.
- **University of Basrah (5001 records):** This is a wide-ranging set of data, exceeding seven decades from the University of Basrah. As shown in Fig. 6, looking at this voluminous volume can give one a peep of the institutions' past and present research interests.
- **University of Mosul (4202 records):** Although the historical database of the University of Mosul is comparatively smaller, one can refer to this data and obtain an idea about the academic past of the university. As shown in Fig. 7, the research disciplines can be better understood by this. There has been an upward trend, especially with an increasing number of recent records during the period between 2021 and 2023. During this phase, it is advisable to study in detail whether corresponding university research projects are undertaken or particular university development is taking place.
- **Mustansiriyah University (8076 records):** As shown in Fig. 8, an array of historic data related to Mustansiriyah University's output spanning over seven decades provides a comprehensive perspective. They will be able to analyze the university's research legacy and identify its areas of perpetual concern. As it seems from the recent record of a few years, there is evidence of further research. Investigating this information can reveal the current research foci and partnerships at universities.

G1	A	B	C	D	E	F	G	H	I	J	K
	Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name				
1	ICM-48 from rice husk ash as a n	https://www.scopus.com/forward?	Case Studies in Chemical & Abstract	Mobil Composition	2023-12-01	[University of Anbar, "Ra Kufi B.M.					
2	Simultaneous electrodeposition of	https://www.scopus.com/forward?	Case Studies in Chemical & Abstract	Electrodeposition	2023-12-01	[University of Baghdad, "Ahmed Y.A.					
3	An experimental study to show the	https://www.scopus.com/forward?	Bent-Suef University Journal of	AbstractThe objective of	2023-12-01	[Al-Maarif University Coildan M.F.					
4	Prediction of ROP Zones Using Dr	https://www.scopus.com/forward?	International Journal of Con	AbstractRetinopathy of pr	2023-12-01	[ENETCom-National Sc Salih N.					
5	A survey on deep learning tools de	https://www.scopus.com/forward?	Journal of Big Data	AbstractData scarcity is a	2023-12-01	[Queensland University Alubaidi L.					
6	Immunohistochemistry technique f	https://www.scopus.com/forward?	Lasers in Medical Science	AbstractThe current study	2023-12-01	[University of Baghdad, "Hussein S.I.					
7	Unravelling the genetic diversity a	https://www.scopus.com/forward?	BMC Plant Biology	AbstractBackground. Con	2023-12-01	[Yasouj University, "Yas Shahi Shawon R.					
8	A novel method for ZnO@NiO core	https://www.scopus.com/forward?	Scientific Reports	AbstractGiven their versa	2023-12-01	[University of TechnologyIran H.J.					
9	Synergistic behavior between the	https://www.scopus.com/forward?	Environmental Nanotechno	AbstractUndoubtedly, one	2023-12-01	[Al-Qasim Green UniverOkab A.A.					
10	Prevalence of autism spectrum dis	https://www.scopus.com/forward?	Journal of Otolaryngology	AbstractNo abstract avail	2023-12-01	[Al-Kindy College of MedAl-Mendalawi M.D.					
11	Interleukin-22 and interleukin-33 a	https://www.scopus.com/forward?	Bent-Suef University Journal	AbstractBackground. The	2023-12-01	[Institut Perubatan &t Majeed A.Y.					
12	The reliability of using light therap	https://www.scopus.com/forward?	BMC Oral Health	AbstractExperimentally, air	2023-12-01	[University of Baghdad, "Al-Qusli A.F.					
13	Impact of bumpers position variate	https://www.scopus.com/forward?	Journal of Engineering and	AbstractOrthodontic, both	2023-12-01	[King's College London, "Waziaz F.					
14	Salivary peptides analysis and p	https://www.scopus.com/forward?	Scientific Reports	AbstractOrthodontic, both	2023-12-01	[University of Mosul, "M Al-Watari B.D.					
15	Effect of exogenous hormones in f	https://www.scopus.com/forward?	International Journal of Veterinarian's	AbstractThe development	2023-12-01	[University of Baghdad, "Hameed H.G.					
16	Synthesis of TiO-2-2-jnp-2-jnp-2	https://www.scopus.com/forward?	Physical Chemistry Research	AbstractThe development	2023-12-01	[Middle Technical Unive Mohammad L.M.					
17	Ameliorating Effects of Herbal Mix	https://www.scopus.com/forward?	International Journal of Alloys and Comp	AbstractThe development	2023-10-05	[University of Baghdad, "Alameedy U.					
18	Developing a magnetic bimultih-ha	https://www.scopus.com/forward?	Journal of Alloys and Comp	AbstractThe development	2023-10-05	[Al-Kindy College of MedAl-Mendalawi M.D.					
19	Matrix acidizing in carbonate rocks	https://www.scopus.com/forward?	Fuel	AbstractNo abstract avail	2023-10-01	[University of Baghdad, "Fadhil E.					
20	Prevalence and patterns of bone m	https://www.scopus.com/forward?	Journal of Tabah University Medical Sciences	AbstractThe flexible joint	2023-10-01	[Swinburne University of Khan R.F.A.					
21	Histological evaluation of the effe	https://www.scopus.com/forward?	IEEE Transactions on Indu	AbstractThe flexible joint	2023-10-01	[University of Haifa, "Ha Ben Khedher N.					
22	Singular Perturbation-Based Adap	https://www.scopus.com/forward?	Journal of Building Engineering	AbstractThe research anc	2023-09-15	[University of Baghdad, "Fathi M.I.					
23	Discharge performance assessme	https://www.scopus.com/forward?	Journal of Energy Storage	AbstractThe production o	2023-09-15	[University of Baghdad, "Jarad A.J.					
24	The effect of whole system rotator	https://www.scopus.com/forward?	Journal of Molecular Struct	AbstractThe production o	2023-09-15	[University of Baghdad, "Jarad A.J.					
25	Synthesis, spectral studies, DFT,	https://www.scopus.com/forward?	Journal of Molecular Struct	AbstractThe production o	2023-09-15	[University of Baghdad, "Jarad A.J.					
26	Cultivation of blue green algae (A	https://www.scopus.com/forward?	Journal of Molecular Struct	AbstractThe production o	2023-09-15	[University of Baghdad, "Jarad A.J.					
27	Electrochemical Determination of	https://www.scopus.com/forward?	Chemosphere	AbstractThe production o	2023-09-01	[University of Baghdad, "Jarad A.J.					
28	Electrochemical Determination of	https://www.scopus.com/forward?	Chemosphere	AbstractThe production o	2023-09-01	[University of Baghdad, "Jarad A.J.					

Fig. 5 The University of Baghdad will record for years

A	B	C	D	E	F	G	H	I	J
Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name			
1 Dynamic Response Analysis of	https://www.scopus.com	International Journal of Concrete Structures and Materials	AbstractIn the current st	2023-12-01	[Al-Mustaqbal University College, 'H Albulani R.				
2 First report of cotton leaf curl	https://www.scopus.com	Australasian Plant Disease Notes	AbstractDemand Side M	2023-12-01	[University of Tehran, 'Tehran', 'Iran Shahmohammadi N.				
3 A new optimized demand man	https://www.scopus.com	Energy Reports	AbstractDemand Side M	2023-12-01	[Iraq University College, 'Basra', 'Iraq Jasin A.M.				
4 A novel missense mutation in	https://www.scopus.com	Egyptian Journal of Medical Humar	AbstractBackground: Ch	2023-12-01	[Islamic Azad University, Ahvaz bran Neissi M.				
5 The impacts of tidal wetland	https://www.scopus.com	Scientific Reports	AbstractUnder conventio	2023-12-01	[University of Basrah, 'Basra', 'Iraq] Al-Attabi Z.				
6 Catalytic response and molec	https://www.scopus.com	Scientific Reports	AbstractUnder conventio	2023-12-01	[VEL TECH MULTITECH, 'Chennai' Tamilarasan R.				
7 Molecular and serological dat	https://www.scopus.com	Iraq Journal of Veterinary Science	AbstractRecently, photo	2023-12-01	[University of Babylon, 'Babylon', 'Iraq Abdalzahra A.I.				
8 Effect of zirconium oxide nano	https://www.scopus.com	Energy Reports	AbstractFor the synchro	2023-12-01	[Southern Technical University, Iraq Abdulkareem M.				
9 Metaheuristic methods to den	https://www.scopus.com	Expert Systems with Applications	AbstractFor the synchro	2023-12-01	[University of Basrah, 'Basra', 'Iraq] Saitar D.				
10 Earthquake prediction techniq	https://www.scopus.com	IAES International Journal of Artificial	AbstractEarthquakes are	2023-09-01	[University of Basrah, 'Basra', 'Iraq] Jarrah N.B.				
11 Earthquake prediction techniq	https://www.scopus.com	Letters in Applied NanoBioScience	AbstractEarthquakes are	2023-08-30	[Ministry of Education, Iraq, 'Baghdad Majed A.A.				
12 Synthesis, Characterisation, a	https://www.scopus.com	Journal of Medicinal and Chemical	AbstractFour new Triazo	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Aghaieard S.A.				
13 Fluid-structure interaction of	https://www.scopus.com	International Journal of Thermal Sc	AbstractNormal impinge	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Hassan A.A.				
14 Modeling and simulation of a	https://www.scopus.com	Bulletin of Electrical Engineering ar	AbstractRecently, pipeli	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Abed M.H.				
15 Features selection for estimat	https://www.scopus.com	Bulletin of Electrical Engineering ar	AbstractHand prosthesis	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Essa R.R.				
16 New efficient fractal models	https://www.scopus.com	Bulletin of Electrical Engineering ar	AbstractParallel data pro	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Khudhar M.M.				
17 Secure two-factor mutual aut	https://www.scopus.com	Bulletin of Electrical Engineering ar	AbstractThe cloud healr	2023-08-01	[University of Basrah, 'Basra', 'Iraq] Adoumalik H.A.				
18 Multivalued carbon nanotubes	https://www.scopus.com	Applied Thermal Engineering	AbstractThis study focus	2023-07-05	[Tikrit University, 'Tikrit', 'Iraq] Hussein O.A.				
19 Impingement jets on a confine	https://www.scopus.com	Applied Thermal Engineering	AbstractThermal stresse	2023-07-05	[University of Warith Al-Anbasya, 'Kut Ismael M.A.				
20 Discussion of "Simultaneous A	https://www.scopus.com	Journal of Engineering Mechanics	AbstractNo abstract ava	2023-07-01	[University of Basrah, 'Basra', 'Iraq] Al-Hussain A.				
21 Evaluation of surface roughne	https://www.scopus.com	Optik	AbstractA film of a mixtu	2023-07-01	[University of Basrah, 'Basra', 'Iraq] Jabbar W.H.				
22 Improving predictions of shale	https://www.scopus.com	Advances in Animal and Veterinary	AbstractThe utilization	2023-06-15	[Liaoning Technical University, 'Fuxi Zhang H.				
23 Methylenediamine Hydrochlorid	https://www.scopus.com	Journal of Polymer Research	AbstractBackground: Th	2023-06-01	[University of Basrah, 'Basra', 'Iraq] Jasin N.H.				
24 Synthesis and evaluation of c	https://www.scopus.com	Partial Differential Equations in Ap	AbstractIn this paper, we	2023-06-01	[University of Basrah, 'Basra', 'Iraq] Rashem F.A.				
25 Generalized q-difference equa	https://www.scopus.com	Optical and Quantum Electronics	AbstractThe increasing r	2023-06-01	[University of Lakkh Marwat, 'Lakkh N-Husain M.				
26 Predicting structural, optoelec	https://www.scopus.com	Optical and Quantum Electronics	AbstractThe increasing r	2023-06-01	[University of Lakkh Marwat, 'Lakkh N-Husain M.				
27 Environmental study of the	https://www.scopus.com	Solar Energy	AbstractIn the current re	2023-06-01	[University of Basrah, 'Basrah', 'Iraq Kadhim Al-Hussain A.				

Fig. 6 University of Basrah records for years

13	A	B	C	D	E	F	G	H
1	Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name	
2	Effects of simulated zero gravity on adhesion	https://www.scopus.com/forward/	Nanotechnology and Precise Abstract-HIGHLIGHTS • Zero	Abstract-HIGHLIGHTS • Zero	2023-12-01	Ninevah University, Nine Alale S.	Nine Alale S.	
3	An investigation of relationships between	https://www.scopus.com/forward/	Journal of Eating Disorders	Abstract-Background: Previous	2023-12-01	Alazra University, Tahir Abdolali A.	Tahir Abdolali A.	
4	The effect of light curing time and intensity	https://www.scopus.com/forward/	Journal of Dental Research	Abstract-Background: This	2023-12-01	University of Mosul, Most Mohammed O.K.	Most Mohammed O.K.	
5	Thermal conductivity improvement in a	https://www.scopus.com/forward/	Energy Reports	Abstract-In this paper, the	2023-12-01	Duy Tan University, Da N H Ghandari M.	Da N H Ghandari M.	
6	Scour depth downstream sharp-crested	https://www.scopus.com/forward/	Journal of Engineering and	Abstract-Weirs are hydraulic	2023-12-01	University of Mosul, Most Obaida A.A.M.	Most Obaida A.A.M.	
7	Retraction Note: Association of cord blood	https://www.scopus.com/forward/	Diabetology and Metabolic	Abstract-The Editor-in-Chief	2023-12-01	University of Babylon, Ba Hussein H.K.	Ba Hussein H.K.	
8	Mesenchymal stem cell-released oncolytic	https://www.scopus.com/forward/	Cell Communication and Sign	Abstract-Oncolytic viruses	2023-12-01	Tishah University of Med Ghasemi Darestani N.	Ghasemi Darestani N.	
9	The role of miRNA-128 in the development	https://www.scopus.com/forward/	Future Oncology	Abstract-Increasing data	2023-12-01	Al-Mustaqbal University CA-Awsi G.R.L.	G.R.L.	
10	Prevalence of ovine brucellosis in Mosul	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The present study	2023-12-01	University of Mosul, Most Hamid M.M.	Most Hamid M.M.	
11	Histopathological changes as tools to disc	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-We aimed in the	2023-12-01	University of Mosul, Most Al-Salih M.A.	Most Al-Salih M.A.	
12	Effect of exogenous hormones in the epr	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The present study	2023-12-01	University of Mosul, Most Abdulhadi S.K.	Most Abdulhadi S.K.	
13	The pharmacokinetics of phenylbutazone	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The aim is to	2023-12-01	University of Mosul, Most Alnuaimi S.I.	Most Alnuaimi S.I.	
14	Neurobehavioral toxicity of copper sulfate	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The aim is to	2023-12-01	University of Mosul, Most Taha A.N.	Most Taha A.N.	
15	The amelioration of Vitamin E on histologic	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-Zinc-oxide in	2023-12-01	University of Mosul, Most Taha A.N.	Most Taha A.N.	
16	Phenotypic and genotypic characterization	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-Obesone pathogen	2023-12-01	Ministry of Agriculture, Idr Othman S.M.	Idr Othman S.M.	
17	Correlation incidence between infectious	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-This study aimed	2023-12-01	University of Mosul, Most Al-Tase Z.T.	Most Al-Tase Z.T.	
18	A comparison between different laboratory	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-For microscopical	2023-12-01	University of Mosul, Most Al-Malachi H.B.	Most Al-Malachi H.B.	
19	Neurotoxicity of xylazine in chicks	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-Despite the	2023-12-01	University of Mosul, Most Taha A.N.	Most Taha A.N.	
20	The impact of nano zinc oxide particles on	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-To determine	2023-12-01	University of Mosul, Most Al-Mahmood O.A.	Most Al-Mahmood O.A.	
21	Food safety and sanitation practices surve	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The study was	2023-12-01	University of Mosul, Most Salih S.S.	Most Salih S.S.	
22	Effect of atorvastatin on bone formation in	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The species of	2023-12-01	University of Mosul, Most Shihab H.H.	Most Shihab H.H.	
23	Detection of resistance against anti-heilm	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The present	2023-12-01	University of Mosul, Most Abdulla Y.A.	Most Abdulla Y.A.	
24	The molecular identification of diarrhoea	https://www.scopus.com/forward/	Iraqi Journal of Veterinary	Abstract-The study was	2023-11-01	University of Mosul, Most Mahmoud B.E.	Most Mahmoud B.E.	
25	A Comparative Study of Numerous Antiox	https://www.scopus.com/forward/	Egyptian Journal of Veterinary	Abstract-The cultural	2023-09-15	University of Mosul, Most Younus I.	Most Younus I.	
26	The role of historic building information	https://www.scopus.com/forward/	Alin Shams Engineering Jour	Abstract-The compact	2023-09-15	University of Mosul, Most Al-Hadiy A.I.	Most Al-Hadiy A.I.	
27	Mechanical performance and statistical an	https://www.scopus.com/forward/	Journal of Building Engineer	Abstract-The compact	2023-09-15	University of Mosul, Most Al-Hadiy A.I.	Most Al-Hadiy A.I.	
28	Sarcosineamide Amino Metabolism in B	https://www.scopus.com/forward/	Environ. Journal of Veterinary	Abstract-TAUC study	2023-09-15	University of Mosul, Most Al-Hadiy A.I.	Most Al-Hadiy A.I.	

Fig. 7 University of Mosul records for years

J11	A	B	C	D	E	F	G	H	I	J
	Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name			
1	Nanomaterials as transmitters of non-viral gene vectors: A	https://www.scopus.com/fnw	Case Studies in Chemi	AbstractWith the rap	2023-12-01	[Al-Balqa Applied Univers	Alkhalwaleh A.K.			
2	Long COVID and risk of erectile dysfunction in recovered	https://www.scopus.com/fnw	Scientific Reports	AbstractPatients with	2023-12-01	[Mustansiriyah University	Al-Kuraishi H.M.			
3	Monkeypox epidemic at the door: should we remain idly b	https://www.scopus.com/fnw	ANB Express	AbstractMonkeypox	2023-12-01	[Mustansiriyah University	Al-Kuraishi H.M.			
4	Physical Study of PVA Filled with Carbon Nanotube and	https://www.scopus.com/fnw	Physical Chemistry Re	AbstractIn this work,	2023-12-01	[University of Samarra, 's	Abdullah A.M.			
5	Comparison of various data mining methods for early diag	https://www.scopus.com/fnw	IAES International Jour	AbstractRecent heal	2023-09-01	[Al-Esraa University Coll	Shanshod A.M.			
6	Anti-capsular activity of CuO nanoparticles against Acinet	https://www.scopus.com/fnw	Microbial Pathogenesis	AbstractCopper oxid	2023-08-01	[Mustansiriyah University	Al-Kadmy I.M.S.			
7	Corrosion Resistance of Calcium Sulfaluminate Cement	https://www.scopus.com/fnw	Journal of Materials in	AbstractThis study i	2023-08-01	[Colorado State Universi	Shakout M.			
8	Direct sunlight photodegradation of Congo red in aqueou	https://www.scopus.com/fnw	Arabian Journal of Che	AbstractPhotolysis i	2023-08-01	[University of Diyala, 'Ba	Asaad Mahdi M.			
9	License plate recognition in slow motion vehicles	https://www.scopus.com/fnw	Bulletin of Electrical En	AbstractThe recogni	2023-08-01	[Mustansiriyah University	Alleyaw Q.M.			
10	Reconfigurable graphene-based multi-input multi-output	https://www.scopus.com/fnw	Bulletin of Electrical En	AbstractThis paper,	2023-08-01	[Mustansiriyah University	Abd R.H.			
11	Impulsive noise mitigation based adaptive filtering and Re	https://www.scopus.com/fnw	Bulletin of Electrical En	AbstractThe paper,	2023-08-01	[Mustansiriyah University	Mohammed A.J.			
12	Aerobic granular sludge formation and stability in enhance	https://www.scopus.com/fnw	Journal of Hazardous	AbstractWastewater	2023-07-15	[Huazhong University of	Cheng L.			
13	A novel method for long-term preserving of urine microsit	https://www.scopus.com/fnw	Journal of Applied Poly	Abstractor analysis	2023-07-10	[University of Diyala, 'Ba	Faham M.A.			
14	Design of graphene-based multi-input multi-output anten	https://www.scopus.com/fnw	Indonesian Journal of	AbstractIn the future	2023-07-01	[Mustansiriyah University	Abd R.H.			
15	Employing various topologies to improve the interleaved b	https://www.scopus.com/fnw	Indonesian Journal of	AbstractIn recent ye	2023-07-01	[Mustansiriyah University	Makke A.H.			
16	Validation hybrid filter detection for Multi-User multiple	https://www.scopus.com/fnw	Alexandria Engineering	AbstractThe present	2023-07-01	[Universiti Putra Malaysi	Agirnee W.			
17	Stress-Strain Model for Geopolymer Mortar under Active	https://www.scopus.com/fnw	Journal of Materials in	AbstractA stress-str	2023-07-01	[Swinburne University of	Rajeev P.			
18	Combining raw data and engineered features for optimizin	https://www.scopus.com/fnw	Computers and Securit	AbstractThe Internet	2023-07-01	[Mustansiriyah University	Saleh M.M.			
19	Photonic controlled metasurface for intelligent antenna be	https://www.scopus.com/fnw	International Jou	AbstractThis paper f	2023-07-01	[Mustansiriyah University	Muqdad Z.S.			
20	Innovative hybridization of the two-architect and PROMET	https://www.scopus.com/fnw	Applied Energy	AbstractPhotovoltaic	2023-07-01	[Mustansiriyah University	Ridha H.M.			
21	Morus alba: a comprehensive phytochemical and phar	https://www.scopus.com/fnw	Naunyn-Schmiedeberg	AbstractMorusalba i	2023-07-01	[Faculty of Veterinary Me	Baltha G.E.S.			
22	Zinc Biofortification as Treatment of Lead-Contaminated	https://www.scopus.com/fnw	Annals of Biology	AbstractMetal biofor	2023-06-01	[University of Technology	Farah B.H.			
23	Numerical and experimental thermohydraulic performance	https://www.scopus.com/fnw	International Communi	AbstractHeat transfe	2023-06-01	[Mustansiriyah University	Ghadban F.N.			
24	Studying the possibility of producing self-colored plaster c	https://www.scopus.com/fnw	Measurement: Sensors	AbstractThis study i	2023-06-01	[Mustansiriyah University	Abed Salman A.			
25	Performance Improvement of CDMA Wireless Sensor Net	https://www.scopus.com/fnw	Wireless Personal Cor	AbstractLow signal i	2023-06-01	[University of Technology	Farhan I.M.			
26	Studying the mechanical properties of high-density polyet	https://www.scopus.com/fnw	Measurement: Sensors	AbstractIn this paper	2023-06-01	[Mustansiriyah University	Abdulreman M.A.			
27	The influence of methacrylic shell addition on the mechanis	https://www.scopus.com/fnw	Innovative Informati	AbstractResearch, di	2023-06-01	[Mustansiriyah University	Mohammed Z.H.			

Fig. 8 Mustansiriyah University records for years

A	B	C	D	E	F	G	H
Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name	
1 Insights into the photovoltaic properties of	https://www.scopus.com/inward/record	Scientific Reports	AbstractAccording to recent re	2023-12-01		[Cornell University College	Dastan D.
2 Corrosion inhibition properties of Schiff ba	https://www.scopus.com/inward/record	Scientific Reports	AbstractThe present study aim	2023-12-01		[University of Technology- I	Bett N.
3 Synthesis of Silver Nanoparticles by Ultr	https://www.scopus.com/inward/record	Journal of Nanostructures	AbstractA lot of work has been	2023-12-01		[University of Technology- I	Mustafa H.N.
4 Experimental, Theoretical, and Numerical	https://www.scopus.com/inward/record	Biorobotry				[University of Technology- I	Olewi J.K.
5 Dynamic Response Analysis of Coastal P	https://www.scopus.com/inward/record	International Journal of Concrete Structures and Materials	AbstractThe validated dynam	2023-12-01		[Al-Mustaqbal University C	Al-Sultani R.
6 A novel dual feedwater circuit for a parab	https://www.scopus.com/inward/record	Scientific Reports	AbstractThe last few decays	2023-12-01		[University of Technology- I	Al-Malki M.A.K.
7 A gold nanoparticles coated uniaxial single	https://www.scopus.com/inward/record	Scientific Reports	AbstractGiven their versatil	2023-12-01		[University of Technology- I	Imran H.J.
8 A novel method for ZnO@NiO core-shell	https://www.scopus.com/inward/record	Scientific Reports	AbstractUnder conventio	2023-12-01		[VEL TECH MULTITECH- T	Tamilarsan R.
9 Catalytic response and molecular simulat	https://www.scopus.com/inward/record	Scientific Reports				[University of Technology- I	Muhsen M.M.
10 Targeted delivery of paclitaxel drug using	https://www.scopus.com/inward/record	Lasers in Medical Science				[University of Technology- I	Abdullahbas Hammodi S.
11 NIR laser-conjugated glutathione-coated	https://www.scopus.com/inward/record	Journal of Energy Storage	AbstractThis study aims to im	2023-10-01		[University of Technology- I	Mosul Younus I.
12 The role of spatial value in the reconstruct	https://www.scopus.com/inward/record	Journal of Molecular Structure	AbstractCorrosion inhibition is	2023-09-06		[University of Technology- I	Rasheed K.H.
13 The role of historic building information m	https://www.scopus.com/inward/record	Journal of Molecular Structure	AbstractCorrosion inhibition is	2023-09-06		[Universiti Malaysia- Kuala	L.Najm M.M.
14 New pyrazole derivative as effective corro	https://www.scopus.com/inward/record	Optics and Laser Technology				[University of Technology- I	Husain S.S.
15 Generation of bright-dark pulses in a C-sh	https://www.scopus.com/inward/record	Indonesian Journal of Science and Technology				[Tadulisan Normal College	It-Hai T.
16 Design of Robust Control for Vehicle Stee	https://www.scopus.com/inward/record	Fuel	AbstractThe current research	2023-09-01		[University of Technology- I	Marzoog T.R.
17 Recurrent neural networks optimization of	https://www.scopus.com/inward/record	Journal of Energy Storage	AbstractThermal stratificatio	2023-08-16		[University of Technology- I	Dakheel H.S.
18 Transient study of thermal stratification	https://www.scopus.com/inward/record	Biochimica et Biophysica A	AbstractBacterial-extracellular	2023-08-15		[Ministry of Higher Educat	Abdulsosol Hasan H.
19 Bacterial extracellular vesicles induced ox	https://www.scopus.com/inward/record	International Journal of The Abstract-Lion battery cell oper				[University of Technology- I	Billly I.J.
20 A novel air-cooled Li-ion battery (LIB) are	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics				[University of Technology- I	Samaan S.S.
21 Advanced optimal GA-PID controller for B	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics				[University of Technology- I	Mohammed I.B.
22 Design high voltage gain DC converter ba	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics				[University of Technology- I	Habeeb Z.Q.
23 Feature-based real-time distributed denia	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics				[University of Technology- I	Abdulmunem A.R.
24 Handwritten signature identification basec	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics					
25 Feature-based real-time distributed denia	https://www.scopus.com/inward/record	Bulletin of Electrical Engineering and Informatics					
26 Incomplete facemask-wearing detection usi	https://www.scopus.com/inward/record	Journal of Energy Storage					
27 Integration of recycled waste paper with p	https://www.scopus.com/inward/record	Journal of Energy Storage					
28 Efficient and stable hole-transportation, low	https://www.scopus.com/inward/record	Journal of Energy Storage					

Fig. 9 University of Technology records for years

A1	A	B	C	D	E	F	G	H	I
	Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name		
1	A type-II Gas/HF-nIn-2-nIn- van de	https://www.scopus.com/lnwa	Applied Surface Science	Abstractin this paper	2020-12-30	[University of Babylon', 'Babylon', 'Obaid M.M.			
2	Trust Evaluation Model Based on Stat	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-30	[University of Babylon', 'Babylon', 'Zah A.H.			
3	Cognitive Network Security Enh	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-23	[University of Babylon', 'Babylon', 'Kadhim A.A.			
4	Mini Jarvis Patrick-Based Graph Cus	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-23	[University of Babylon', 'Babylon', 'Almingshi H.Z.			
5	Decision Making Approaches in Cogni	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-23	[University of Babylon', 'Babylon', 'Bdrany A.			
6	Privacy Preserving Association Rules	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-23	[University of Babylon', 'Babylon', 'Saman W.A.K.			
7	Enhance the Performance of Independ	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-23	[University of Babylon', 'Babylon', 'Shabat H.A.			
8	Robust relay cognitive network do	https://www.scopus.com/lnwa	3rd International Conference on / Abstract	Abstractin this paper	2020-12-22	[University of Babylon', 'Babylon', 'E'Al-Asadi A.			
9	Black/Gray Holes Detection Tools in N	https://www.scopus.com/lnwa	7th IEEE International Conference	AbstractAchieving hi	2020-12-18	[Beijing Institute of Technology', 'E'Aboud M.S.			
10	Tailoring the electronic and optical pr	https://www.scopus.com/lnwa	Physics Letters, Section A	AbstractThrough a li	2020-12-17	[University of Babylon', 'Babylon', 'Abdulnashem Z.			
11	TiN@Al-nIn-2-nIn-0-nIn-3-nIn-5-A	https://www.scopus.com/lnwa	Water Science and Technology	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Baghd Al-Haddad A.			
12	Simulate permeable reactive barrier by	https://www.scopus.com/lnwa	Water Science and Technology	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
13	A review on COVID-19: Origin, spread,	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
14	The Hydraulic Jump Formed Downstre	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
15	Automated Approach for Extraction of	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
16	Automated Approach for Extraction of	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
17	CFD Analysis for the Effect of Person	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
18	Air Distribution Performance Inside Off	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
19	The effect of smoke water extract on e	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
20	Analysis of some structural, electronic	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
21	Blockchain technology's applications a	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
22	Effect of cobalt oxide nanoparticles on	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
23	Characterization and sensing applicati	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
24	Etiology of Parkinson's disease: Toxop	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
25	Highly sensitive and selective colorim	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
26	Effect of short time delays of quantum	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
27	Synthesis and study biological activity	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			
28	Divinical nomenclature of hunchabab	https://www.scopus.com/lnwa	Biointerface Research in Applied	AbstractThe remedia	2020-12-15	[University of Babylon', 'Babylon', 'Hadi A.G.			

Fig.10 University of Babylon records for years

K4																
v i x x ✓ fx																
A B C D E F G H I J K																
Title URL Publication Name Abstract Date of Publication Author Affiliations Author Name																
1 Automatic and Non-iteration Algorithm for Sensing and Imaging https://www.scopus.com/ Abstract: Although man 2023-12-01 [University of Kufa, 'Kufa', 'Ira Al-dulali M.J.																
2 Experimental performance comparison of deep learning models for https://www.scopus.com/ Energy Reports Abstract: In this work, n 2023-12-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
3 Efficiency-based design optimization of a microfluidic device https://www.scopus.com/ Energy Reports Abstract: Enhancement 2023-12-01 [University of Kufa, 'Kufa', 'Ira Ansar R.																
4 A numerical investigation of the effect of the inlet flow on the https://www.scopus.com/ Journal of Talabeh University Abstract: Objectives: B 2023-10-01 [Al-Furat Al-Awsat Technical Hameed H.G.																
5 Histological evaluation of the effect of the inlet flow on the https://www.scopus.com/ Journal of Talabeh University Abstract: Background: 2023-09-01 [University of Baghdad, 'Bagi Fadhil E.																
6 Vitamin D receptor gene polymorphism in the Iraqi population https://www.scopus.com/ Human Gene Abstract: In this study, 2023-08-01 [University of Kufa, 'Kufa', 'Ira Shakir K.S.																
7 New models for reinforced concrete structures https://www.scopus.com/ Structures Abstract: Estimation of 2023-08-01 [Qadman Normal College for Tech H.																
8 Development of high-resolution of the Iraqi population https://www.scopus.com/ Atmospheric Research Abstract: Parallel data 2023-08-01 [University of Basrah, 'Basra Khudhair M.M.																
9 New efficient fractal models for the Iraqi population https://www.scopus.com/ Bulletin of Electrical Engineering Abstract: A series of se 2023-07-01 [University of Kufa, 'Kufa', 'Ira Majeed N.S.																
10 Synthesis and characterization of the Iraqi population https://www.scopus.com/ Internet of Things (Nenerian) Abstract: The future, 2023-07-01 [University of Kufa, 'Kufa', 'Ira Majeed N.S.																
11 Artificial intelligence based model for the Iraqi population https://www.scopus.com/ Applied Radiation and Isotop Abstract: Natural radiat 2023-07-01 [University of Kufa, 'Kufa', 'Ira Doshi R.J.																
12 Natural radioactivity for soil sample https://www.scopus.com/ Neural Computing and Appl Abstract: The success 2023-07-01 [American University of Kuwait Abu Doush I.																
13 Artificial intelligence based model for the Iraqi population https://www.scopus.com/ Applied Radiation and Isotop Abstract: The success 2023-07-01 [American University of Kuwait Abu Doush I.																
14 Active-based coronavirus herd in https://www.scopus.com/ Journal of Energy Storage Abstract: Phase change 2023-07-01 [Georgia Institute of Technol Nasippour-Esfahani N.																
15 The role of phase change material in https://www.scopus.com/ Journal of Energy Storage Abstract: Phase change 2023-07-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
16 Potential Role of Reinforcement to Ams https://www.scopus.com/ Journal of Mechanical and Ch Abstract: Existing rein 2023-07-01 [Universiti Malaya, 'Kuala Lu Chayeb R.H.																
17 Performance of reinforced concrete https://www.scopus.com/ Case Studies in Constructor Abstract: Existing rein 2023-07-01 [Al-Muostansiriyah University Lawi D.J.																
18 Ecological and human health risk https://www.scopus.com/ Biological Trace Element Res Abstract: Heavy metals 2023-06-15 [University of Babylon, 'Babyl Gmasha R.A.																
19 Health and human health risk https://www.scopus.com/ Environmental Research Abstract: 2023-06-15 [University of Kufa, 'Kufa', 'Ira Al-Hakeim H.K.																
20 The physio-affective phenomena of https://www.scopus.com/ Journal of Affective Disorders Abstract: Background: 2023-06-15 [University of Basrah, 'Basra Almousavi A.																
21 Control Strategy for a PV-BESS-S https://www.scopus.com/ Iraqi Journal for Electrical an Abstract: This paper, 2023-06-15 [Al-Furat Al-Awsat Technical Hameed H.G.																
22 Performance study on a solar con https://www.scopus.com/ Helvion Abstract: The rapid gro 2023-06-01 [Basra University for Oil and Homod R.Z.																
23 Grade of production prediction ba https://www.scopus.com/ Geoenery Science and Eng Abstract: Since the beh 2023-06-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
24 Enhanced multi-agent systems for https://www.scopus.com/ Results in Engineering Abstract: Model of the m 2023-06-01 [Alman University, 'Alman', 'U Al-Peter M.A.																
25 Marine Predators Algorithm: A Re https://www.scopus.com/ Archives of Computational M Abstract: Marine Preda 2023-06-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
26 Experimental and Numerical Mode https://www.scopus.com/ KSCE Journal of Civil Engin Abstract: Piled rail fo 2023-06-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
27 Investigation of earth air heat exch https://www.scopus.com/ Renewable Energy Abstract: This study pr 2023-06-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
28 A comprehensive review of smart https://www.scopus.com/ Journal of Environmental Abstract: The smart 2023-06-01 [University of Kufa, 'Kufa', 'Ira Alshibli A.M.A.																
Sheet 10																
+ Accessible: Good to go																

Fig. 11 University of Kufa records for the years

Author Name									
A	B	C	D	E	F	G	H	I	J
Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name			
1									
2	Electroencephalogram signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	Engineering Applications of Artificial Intelligence	Abstract: This paper presents a deep learning-based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-08-01	[Chongqing Technology and Business University, Chongqing, China]			
3	A review on detecting brain tumors using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper reviews the latest research on deep learning for brain tumor detection. It highlights the challenges and opportunities in this field.	2023-08-01	[Iraqi Commission for Comp. Al-Ani N.Q.]			
4	Classification of electroencephalogram signals using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a deep learning-based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-08-01	[Al-Mustaqbal University, Cc Meqdad M.N.]			
5	Incorrect facemask-wearing detection using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	Bulletin of Electrical Engineering and Informatics	Abstract: This paper presents a deep learning-based approach for detecting incorrect facemask-wearing. The proposed model achieves high accuracy in distinguishing between different classes of facemask-wearing.	2023-08-01	[University of Technology, I. Haseeb Z.Q.]			
6	Feature selection of pre-trained deep learning models	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	Applied Intelligence	Abstract: This paper presents a feature selection method for pre-trained deep learning models. The proposed method achieves high accuracy in distinguishing between different classes of pre-trained models.	2023-08-01	[University of Information Technology, Hamad Q.S.]			
7	Accurate method for sleep stages classification using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	Biomedical Signal Processing and Communications	Abstract: This paper presents a deep learning-based approach for sleep stages classification. The proposed model achieves high accuracy in distinguishing between different classes of sleep stages.	2023-07-01	[Informatics Institute for Postgraduate Studies, Hussein R.]			
8	Cluster-based denoising autoencoder for EEG signal processing	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	Biomedical Signal Processing and Communications	Abstract: This paper presents a cluster-based denoising autoencoder for EEG signal processing. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-06-01	[Iraqi Commission for Comp. Al-Ani N.Q.]			
9	A comparison of the performance of different deep learning models for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper compares the performance of different deep learning models for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-06-01	[University of Information Technology, Mezhher A.E.]			
10	Parallel processing of E-Ather algorithm for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a parallel processing of E-Ather algorithm for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-06-01	[University of Information Technology, Mezhher A.E.]			
11	Design of a multi-level inverter for an induction motor drive	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a multi-level inverter for an induction motor drive. The proposed model achieves high accuracy in distinguishing between different classes of induction motor drives.	2023-06-01	[Department of Nuclear Power, Abdulrazzak A.A.]			
12	A multi-hop routing protocol for an induction motor drive	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a multi-hop routing protocol for an induction motor drive. The proposed model achieves high accuracy in distinguishing between different classes of induction motor drives.	2023-06-01	[University of Information Technology, Mezhher A.E.]			
13	Immunohistochemistry BC image analysis using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents an immunohistochemistry BC image analysis using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of BC images.	2023-05-22	[University of Information Technology, Mezhher A.E.]			
14	A comparison of three different wavelet transforms for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper compares three different wavelet transforms for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-05-22	[University of Information Technology, Mezhher A.E.]			
15	Estimating systems reliability function using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a deep learning-based approach for estimating systems reliability function. The proposed model achieves high accuracy in distinguishing between different classes of systems reliability functions.	2023-05-22	[University of Information Technology, Mezhher A.E.]			
16	Wide and deep learning based approach for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a wide and deep learning based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-05-01	[University of Information Technology, Mezhher A.E.]			
17	Effect of different separation techniques on EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents an effect of different separation techniques on EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-05-01	[University of Information Technology, Mezhher A.E.]			
18	Analysis time series of vegetation indices using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents an analysis time series of vegetation indices using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of vegetation indices.	2023-04-21	[University of Information Technology, Mezhher A.E.]			
19	Automatic clustering and classification of EEG signals using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents an automatic clustering and classification of EEG signals using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-04-01	[University of Information Technology, Mezhher A.E.]			
20	Self-healing of concrete using bacteria using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a self-healing of concrete using bacteria using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of concrete.	2023-04-01	[University of Information Technology, Mezhher A.E.]			
21	Diagnosis of diabetes mellitus using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a diagnosis of diabetes mellitus using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of diabetes mellitus.	2023-03-31	[University of Information Technology, Mezhher A.E.]			
22	A systematic review in recent trend of EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a systematic review in recent trend of EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-03-31	[University of Information Technology, Mezhher A.E.]			
23	Fuzzy Hazard Rate Function for MI using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a fuzzy hazard rate function for MI using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of MI.	2023-03-29	[Ministry of Education, Iraq; Ramadan M.H.]			
24	A New Approach to Nonlinear State Space Models using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a new approach to nonlinear state space models using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of state space models.	2023-03-01	[Ministry of Education, Iraq; Ramadan M.H.]			
25	A Trickle-before-Clustering Initial Clustering Method using deep learning	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a trickle-before-clustering initial clustering method using deep learning. The proposed model achieves high accuracy in distinguishing between different classes of clustering methods.	2023-03-01	[Faculty of Computers and Azar A.T.]			
26	A cross-sectional analysis of the prevalence of EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a cross-sectional analysis of the prevalence of EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-03-01	[Universiti Sains Malaysia; Saleh S.A.M.]			
27	Machine learning classification-based approach for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a machine learning classification-based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-03-01	[Tham Jaafar Al-Sadiq Univ. Al-Anbar S.S.J.]			
28	Machine learning classification-based approach for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a machine learning classification-based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-03-01	[Tham Jaafar Al-Sadiq Univ. Al-Anbar M.N.]			
29	Machine learning classification-based approach for EEG signal classification	https://www.scopus.com/inward/record.uri?eid=2-s2.0-3490151111&doi=10.1016/j.neuroimage.2023.119801	International Journal of Modern Trends in Science	Abstract: This paper presents a machine learning classification-based approach for EEG signal classification. The proposed model achieves high accuracy in distinguishing between different classes of EEG signals.	2023-02-01	[Tham Jaafar Al-Sadiq Univ. Al-Anbar A.A.]			
Sheet1									
Accessibility: Good to go									
100%									

Fig. 12 University of Information and Communications records for years

J11							Formula Subj H						
A	B	C	D	E	F	G							
Title	URL	Publication Name	Abstract	Date of Publication	Author Affiliations	Author Name							
1 Author Correction: Green synthesis and characterization of https://www.sciencedirect.com/science/article/pii/S0926641023000000 Scientific Reports	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThe o 2023-12-01	AbstractThe o 2023-12-01		[Ministry of Higher Education and AI I.A.M.]								
2 Green synthesis and characterization of https://www.sciencedirect.com/science/article/pii/S0926641023000000 Scientific Reports	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThe p 2023-12-01	AbstractThe p 2023-12-01		[University of Sfax, 'Sfax', 'Tunis'] Al I.A.M.]								
3 Author Correction: Partial differential equation https://www.sciencedirect.com/science/article/pii/S0926641023000000 Scientific Reports	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractIn the 2023-12-01	AbstractIn the 2023-12-01		[Al-Nahrain University, 'Baghdad', 'Abdulhadi K.F.]								
4 Estimating the role of single-nucleotide https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Genetic Engineering and Biotechnology	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThe r 2023-12-01	AbstractThe r 2023-12-01		[Al-Nahrain University, 'Baghdad', 'Abdulhadi K.F.]								
5 Machine learning-based telemedicine https://www.sciencedirect.com/science/article/pii/S0926641023000000 Network Modeling Analysis	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThe r 2023-12-01	AbstractThe r 2023-12-01		[Al-Nahrain University, 'Baghdad', 'Abdulhadi K.F.]								
6 Physical Study of PVA Filled with Carbon https://www.sciencedirect.com/science/article/pii/S0926641023000000 Physical Chemistry Research	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractIn this 2023-12-01	AbstractIn this 2023-12-01		[University of Samarra, 'Samarra', 'Abdullah A.M.]								
7 Immunoinformatics design of multi-epitope https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Molecular Modeling	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractIn this 2023-12-01	AbstractIn this 2023-12-01		[Al-Nahrain University, 'Baghdad', 'Al-Chalabi R.]								
8 Visible-light boosted photocatalytic https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Molecular Modeling	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractEnhan 2023-09-15	AbstractEnhan 2023-09-15		[University of Warith Al-Anblyaa, 'Amara S.H.]								
9 An automated machine learning model https://www.sciencedirect.com/science/article/pii/S0926641023000000 IACS International Journal of Artificial Intelligence	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[Al-Nahrain University, 'Baghdad', 'Maher N.]								
10 Hydrothermally native defect induced https://www.sciencedirect.com/science/article/pii/S0926641023000000 Optical and Quantum Electronics	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[University of Technology- Iraq, 'E. Jasin A.S.]								
11 Cost-effective, high-performance fiber https://www.sciencedirect.com/science/article/pii/S0926641023000000 Optical and Quantum Electronics	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[Middle Technical University, 'Baghdad', 'Resan D.A.]								
12 High sensitive temperature sensor base https://www.sciencedirect.com/science/article/pii/S0926641023000000 Optical and Quantum Electronics	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[Al-Nahrain University, 'Baghdad', 'Al-Bayati A.F.]								
13 Punching Shear Design of Waffle Slabs https://www.sciencedirect.com/science/article/pii/S0926641023000000 Practice Periodical on Structural Engineering	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[University of Baghdad, 'Baghdad', 'Thary Al-Ghain A.H.]								
14 Silver oxide-zinc for removal of an https://www.sciencedirect.com/science/article/pii/S0926641023000000 International Journal of Electrochemistry and Photocatalysis	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
15 Real-time human face tracking and https://www.sciencedirect.com/science/article/pii/S0926641023000000 International Journal of Electrochemistry and Photocatalysis	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-08-01	AbstractThis s 2023-08-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
16 An overview of the latest research on https://www.sciencedirect.com/science/article/pii/S0926641023000000 European Polymer Journal	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-24	AbstractThis s 2023-07-24		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
17 Improvement of Plate-Fin Heat Exchanger https://www.sciencedirect.com/science/article/pii/S0926641023000000 CFD Letters	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
18 Detecting Arabic sexual harassment https://www.sciencedirect.com/science/article/pii/S0926641023000000 Egyptian Informatics Journal	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
19 Au-TiO ₂ -TiO ₂ -ZnO-coated dielectric https://www.sciencedirect.com/science/article/pii/S0926641023000000 Optical and Quantum Electronics	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
20 Combining raw data and engineered https://www.sciencedirect.com/science/article/pii/S0926641023000000 Computers and Security	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Mustansiriyah University, 'Baghdad', 'Saad M.M.]								
21 A classification model of breast https://www.sciencedirect.com/science/article/pii/S0926641023000000 Biomedical Signal Processing	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
22 Microbial fuel cell review: thermodynamic https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Chemical Technology	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
23 Molecular diversity of 23-YSTR markers https://www.sciencedirect.com/science/article/pii/S0926641023000000 Gene	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
24 Prediction of compressive strength and https://www.sciencedirect.com/science/article/pii/S0926641023000000 Case Studies in Construction	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
25 Stereable composite right-left-hand https://www.sciencedirect.com/science/article/pii/S0926641023000000 Microwave and Optical Technology Letters	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
26 The Role of Neuregulin 4 Status in https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Medicinal and Chemical Sciences	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
27 The Role of Neuregulin 4 Status in https://www.sciencedirect.com/science/article/pii/S0926641023000000 Journal of Medicinal and Chemical Sciences	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-07-01	AbstractThis s 2023-07-01		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								
28 Study of Nanomaterials with Different https://www.sciencedirect.com/science/article/pii/S0926641023000000 Research in Analytical Chemistry	https://www.sciencedirect.com/science/article/pii/S0926641023000000	AbstractThis s 2023-06-15	AbstractThis s 2023-06-15		[Al-Nahrain University, 'Baghdad', 'Thary Al-Ghain A.H.]								

Fig. 13 Al-Nahrain University records for years

- **University of Technology (4682 records):** The University of Technology's historical dataset covers seven decades of proven academic performance. By analyzing these data, we can learn about this institution's long-term areas of interest as shown in Fig. 9.
- **University of Babylon (8122 records):** The University of Babylon has an extensive seven decades historical dataset, as shown in Fig. 10. Such a comprehensive collection of records can provide important clues regarding the general scientific research tendencies and intellectual output that a particular institution has generated over decades.
- **University of Kufa (5001 records):** The University of Kufa shows academic consistency through a big dataset that covers over seven decades. As shown in Fig. 11, this comprehensive data analysis uncovers the long-term research agenda of the institute and its academic legacy.
- **University of Information and Communications:** As shown in Fig. 12, a total of 534 records were retrieved from the University of Information and Communications between 1950 and 2023.
- **Al-Nahrain University (4289 records):** The University of Al-Nahrain's dataset is sufficiently large with regard to teaching quality, as it provides an idea of its academic achievements. This will enable one to explore the university's strong points and specialty as shown in Fig. 13.
- **Al-Iraqiya University (1278 records):** The documents on files relating to Al-Iraqiya University are of lesser quantity; however, they point to some direction and achievement of the research fields of this establishment. As shown in Fig. 14, this data analysis will help discern what makes it academically special.

Fig. 15 presents the distribution of records spread across the records in Ten Iraqi Universities; for the years, the University of Baghdad leads with the highest count of records far outstripping the rest. After them come Mustansiriyah University, University of Babylon, and Al-Nahrain University. The above-mentioned three universities also have considerable track record, although it is significantly less compared to that of University of Baghdad. However, the University of Information and Communications has the lowest number of records, thus being the university with the least scholarly output. In general, the chart gives a pictorial representation of the difference in research output between the universities within the stated duration.

6 Discussion

This section deals with the implications, difficulties, and potential prospects of web scraping of scientific data from SCOPUS databases for aforementioned Iraqi universities. As such, the data that were extracted have important implications in context of this study on the universities discussed; the University of Basrah, University of Babylon, University of Mosul, University of Kufa, University of Information and Communications, University of Mustansiriyah, and University of Al-Nahrain.

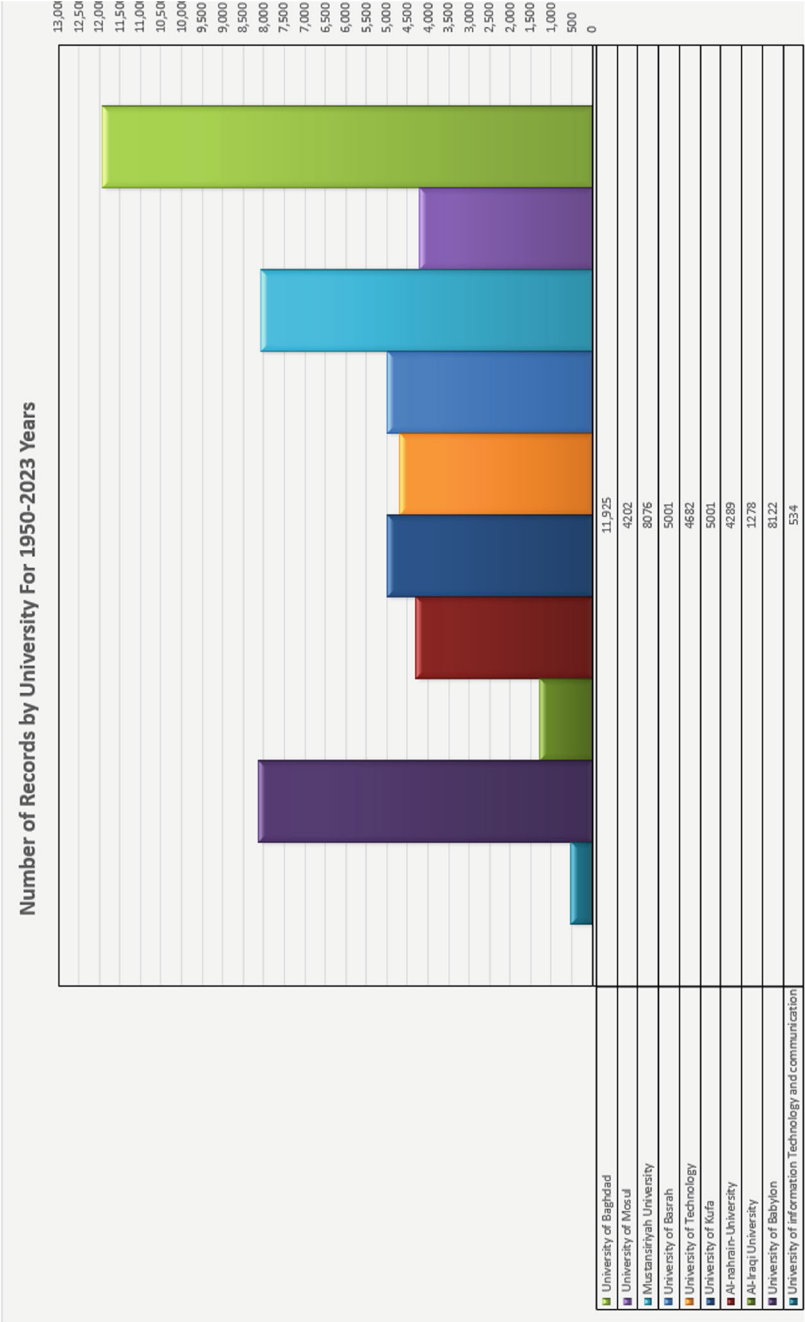


Fig.15 Analysis of different Iraqi Universities records

6.1 *Implications for Academic Institutions*

This study has important implications for those universities mentioned above, such as University of Babylon, University of Mosul, University of Kufa, University of Information and Communication, Mustansiriyat University, Al-Nahrain University, Al-Iraqiya University; these institutions are thus provided by the extraction of scientific data with relevant and enriching content that boosts their efforts in research from SCOPUS. First, this information can help universities recognize research trends and areas where additional studies are needed. By evaluating trends in the number and spatial allocation of research documents over time, institutions can identify trending areas of interest and direct resources in that direction. The results of that institutions of higher learning are able to prioritize research based on factual evidence.

Second, academic institutions have the opportunity to measure the level of their research production and its effectiveness through this information. Universities can measure the level of their scholarly influence globally and assess publication productivity using such numbers. Such data are priceless for benchmarking academic and strategic planning purposes. Moreover, the data extraction process helps universities adjust their research strategies to correspond with contemporary academic settings. Through this, institutions are able to locate the most productive writers, prominent journals, and find venues for collaboration, enhancing the research output of such a networked world.

6.2 *Challenges and Limitations*

The proposed web scraping process had several difficulties and restrictions during the course of the study. Initially, it included modification of the website structure and content presentation in various other academic databases such as SCOPUS. However, this dynamism makes it difficult to maintain a constant routine of data extraction across all years. The web scraping scripts were continuously monitored and adjusted to ensure data validity.

In addition, there is a limitation when there were no ways of accessing all SCOPUS's content, as it was limited. Although this limitation is not specific to this report, it is common to all studies that use web scraping for proprietary datasets. This reinforces the importance of partnerships between academia and database providers to facilitate full access to data for research purposes.

6.3 *Future Directions*

The results of this study open various avenues for future research. First, NLP may be used as a technique for analyzing the content of academic papers in future research.

The extracted data were subjected to sentiment analysis, keyword extraction, and topic modeling to deepen the understanding of its content and context. Moreover, cooperation between universities and database owners could result in the development of APIs and data sharing agreements that would allow for better data use through APIs, taking ethics into account. Collaborative practices can contribute to responsible data handling within the academia. Finally, the use of machine learning and predictive analysis of the extracted information can provide institutions with predictive models about research areas, areas where more money should be directed and funded, and an evaluation of how various university programs influence knowledge growth in the world. Such tools may help institutions to control their future research directions.

7 Conclusion

Finally, this research has shown the value of web scraping for gathering scholarly data from SCOPUS repositories. The findings present new lights on the profile of leading Iraqi academic universities that could be used by research managers, policymakers, and academics themselves in pursuit of better research decisions, benchmarking, and international partnerships. This article highlights the imperativeness of employing appropriate data techniques and respectable web scraping methods for research purposes, despite their inherent drawbacks. Such interventions could have a significant impact on academic institutions' research approaches and productivity. The convergence of online scraping, data analytics, and academic research is a field that has great potential as we move forward. The academic community can leverage the power of online scraping to advance knowledge and promote innovation by tackling issues, forming alliances, and adopting innovative technology.

The following avenues for future research and recommendations are proposed:

1. It may be suggested that using the data scraping system developed in this study is a part of the educational institution's management system, since it gives a clear vision of the completion of scientific plans and the path of its scientific research. It may also be suggested that this system may be a part of the educational institution's page or website to manifest its scientific activity to the scientific community.
2. Future activities should involve the utilization of modern machine learning techniques in relation to heightening precision for information extraction at a scientific level. For instance, considering techniques such as natural language processing (NLP) and machine vision would seem promising in dealing with unstructured data types as well as different document structures.
3. It is critical to address the challenges presented by dynamically generated web pages. Research also deals with the discovery of approaches for managing JavaScript-generated content and discovering ways to capture data from websites based on client-side rendering.

4. The incorporation of semantic web technologies like Resource Description Framework (RDF) and SPARQL protocol and RDF query language (SPARQL) offers an interesting opportunity. This suggestion is proposed upon improving the structuring and interoperability of generated scientific information in order to allow easy integration with current knowledge graphs or databases.
5. With a focus on these future research areas, researchers can foster web scraping methodologies and make significant contributions to ongoing progress in dealing with emerging issues while harnessing the full potential of this technology for promoting scientific advancement.

References

1. Gottardi T, Medeiros CB, Dos Reis JC (2021) Semantic search on scientific repositories: a systematic literature review. *Sociedade Brasileira de Computacao—SB*, pp 271–276. <https://doi.org/10.5753/sbbd.2020.13653>
2. Karthikeyan T, Sekaran K, Ranjith D, Vinoth kumar V, Balajee JM (2019) Personalized content extraction and text classification using effective web scraping techniques. *Int J Web Portals* 11(2):41–52. <https://doi.org/10.4018/IJWP.2019070103>
3. Meschenmoser P, Meuschke N, Hotz M, Gipp B (2016) Scraping scientific web repositories: challenges and solutions for automated content extraction. *D-Lib Mag* 22(9–10). <https://doi.org/10.1045/september2016-meschenmoser>.
4. Ducato R (2020) Data protection, scientific research, and the role of information. *Comput Law Secur Rev* 37. <https://doi.org/10.1016/j.clsr.2020.105412>
5. Zatsman IM (2023) On the scientific paradigm of informatics: data, information, and knowledge. *Informatika i Ee Primeneniya* 17(1):116–125. <https://doi.org/10.14357/19922264230115>
6. Dogucu M, Çetinkaya-Rundel M (2020) Web scraping in the statistics and data science curriculum: challenges and opportunities. *J Stat Educ*. <https://doi.org/10.1080/10691898.2020.1787116>
7. vanden Broucke S, Baesens B (2018) From web scraping to web crawling. In: *Practical web scraping for data science*. https://doi.org/10.1007/978-1-4842-3582-9_6
8. Nigam H, Biswas P (2021) From web scraping to web crawling. https://doi.org/10.1007/978-981-16-3067-5_9
9. Chung CC, Jeng TS (2018) Information extraction methodology by web scraping for smart cities: using machine learning to train air quality monitor for smart cities. In: *CAADRIA 2018—23rd International conference on computer-aided architectural design research in Asia: learning, prototyping and adapting, the association for computer-aided architectural design research in Asia (CAADRIA)*, pp 515–524. <https://doi.org/10.52842/conf.caadria.2018.2.515>
10. Babu K (2020) Survey on web scraping technology. *Waffen-und Kostumkunde Journal* 16(06). <https://doi.org/10.37896/whjj16.06/001>
11. Prehanto DR, Indriyanti AD, Prisma I, Permadi GS, Prastyo EH (2021) Implementation of web scraping on news sites using the supervised learning method. *İlköğretim Online* 20(3). <https://doi.org/10.17051/ilkonline.2021.03.43>
12. Bricongne JC, Meunier B, Pouget S (2023) Web-scraping housing prices in real-time: the Covid-19 crisis in the UK. *J Hous Econ* 59. <https://doi.org/10.1016/j.jhe.2022.101906>
13. Habibie MI, Widiaputra T, Yulianingsani Y (2022) Web scraping of disease information from social media Twitter. *Jurnal Teknoinfo* 16(2). <https://doi.org/10.33365/jti.v16i2.1871>
14. Nigam H, Biswas P (2021) Web scraping: from tools to related legislation and implementation using python. In: *Lecture notes on data engineering and communications technologies*, vol 59. https://doi.org/10.1007/978-981-15-9651-3_13

15. Thomas DM, Mathur S (2019) Data analysis by web scraping using python. In: Proceedings of the 3rd international conference on electronics and communication and aerospace technology, ICECA 2019. <https://doi.org/10.1109/ICECA.2019.8822022>
16. Uzun E (2020) A novel web scraping approach using the additional information obtained from web pages. IEEE Access 8. <https://doi.org/10.1109/ACCESS.2020.2984503>
17. Diouf R, Sarr EN, Sall O, Birregah B, Bousso M, Mbaye SN (2019) Web scraping: state-of-the-art and areas of application. In: Proceedings—2019 IEEE international conference on big data, big data 2019. <https://doi.org/10.1109/BigData47090.2019.9005594>
18. Lotfi C, Srinivasan S, Ertz M, Latrous I (2021) Web scraping techniques and applications: a literature review. In: SCRS conference proceedings on intelligent systems. <https://doi.org/10.52458/978-93-91842-08-6-38>.

A Space-Saving Technique in SDN for Identifying Top-k Super-Spreaders Using P4-Enabled Switches



Ali Nadim Alhaj, Wilson Naik Bhukya, and Rajendra Prasad Lal

Abstract SuperSpreaders are hosts exhibiting an unusual number of unique connections exceeding a specified threshold value. Detecting SuperSpreaders is a critical task for networks, particularly in the context of SDN, aiming to enhance network management, abnormal flow monitoring, and security. In contrast to heavy hitters detection, identifying super-spreaders hosts in high-speed networks poses significant challenges due to memory and processing requirements. Existing solutions often cannot overcome the memory and accuracy limitations imposed on network monitoring points. To tackle this issue, we propose an SSD (Space-saving detector) mechanism comprising of two filtering layers to count distinct connections per host. Additionally, we used a multi-stage data structure (hashpipe), which is employed to maintain flows with the highest cardinality while removing those with lower cardinality. We implemented our solution on SDN P4-programmable switches to confirm the feasibility of deploying our solution in physical SDN networks. Experimental results demonstrate the superiority of our solution, achieving higher F1-score and reducing Average-Absolute-Error and Average-Relative-Error compared to the closest competitor.

Keywords Superspreaders detection · Traffic measurement · Cardinality estimation · Memory efficiency

A. N. Alhaj (✉) · W. N. Bhukya · R. P. Lal
University of Hyderabad, Hyderabad, Telangana, India
e-mail: 21mcpc16@uohyd.ac.in

W. N. Bhukya
e-mail: rathore@uohyd.ac.in

R. P. Lal
e-mail: rajendraprasd@uohyd.ac.in

1 Introduction

A superspreader is defined as a host that connects to a large number of devices in the network in an unusual manner. Detecting superspreaders in current networks is considered essential because it has a role in network management, recommendation systems, and detecting malicious attacks such as DDoS attacks [1], worm propagation [2] and port scanning [3]. The superspreader detection problem is related to counting flow cardinality, which is fundamentally different from counting flow size, which is used for heavy-hitter detection. The flow size is defined as the count of packets that belong to the flow, while flow cardinality is the count of distinct elements of the flow. Therefore, estimating cardinality is more difficult because it requires eliminating repetition, which is tedious with high-speed networks that use limited memory for switches (monitoring points).

The advent of SDN programmable P4-switches [4] has facilitated the integration of data structures augmented with straightforward instructions, enabling the collection of statistics on network flows traversing the switch. Given the challenges posed by current network data speeds, constrained switch memory, and the imperative for precise cardinality flow estimation, we propose a novel mechanism rooted in the Space-Saving principle [5]. Our approach employs a dual-tiered filtering system. The first filter level eliminates redundancy in estimation, and the second filter level guarantees the removal of duplicates (unique entries for each flow). Additionally, we utilize a Hashpipe structure [6] designed to discard minor flows (low cardinality) while retaining major ones (high cardinality), thereby prioritizing flows with the highest K -Cardinality which are proposed to be on the final top- k super-spreaders list. Consequently, we achieve accurate and efficient identification of flows with the highest cardinality while maximizing the utilization of limited memory resources. Evaluation of P4 switches demonstrates the superiority of our solution over existing methods in cardinality flow estimation and Super-Spreader detection. Our algorithm accurately detects top- k Superspreaders while conserving memory size. We summarize our contributions in this paper as follows:

- Design and analyze SSD (space-saving detector), a new method for detecting top- k SuperSpreaders with high accuracy and efficient memory usage within the minimum limits.
- Create a prototype of the proposed mechanism in the P4 switch using the p4-behavior model [7].
- We compare our proposed SSD mechanism to the state-of-the-art, showing its effectiveness in providing accurate cardinality estimates with less memory usage. We also analyze the influence of increasing the parameter K on the algorithm's performance compared to previous methods. Our SSD mechanism demonstrated efficient detection of Top- k SuperSpreaders with a minimal error rate below 5%.

We designed the subsequent sections of the paper as follows: In Sect. 2, we outlined the primary contemporary research on identifying superspreaders. In Sect. 3, we presented a mathematical definition of the problem. Section 4 provides an explanation

of our proposed solution. The algorithm and its application to SDN switches will also be discussed. Section 5 includes practical experiments and comparisons. The last Sect. 6 consists of the conclusion and future plans.

2 Related Work

A number of studies have been proposed for detecting superspreaders using sketches to collect statistics of all network flows and any cardinality estimation techniques to estimate the cardinality. Still, these studies ran into two main problems: (1) Insufficient switch memory to reserve an estimator for all network flows, and (2) Difficulty reversing the statistics Sketch provided to obtain the superspreaders' identifiers (src IPs). Research such as SpreadSketch [8] uses a structure similar to Count-Min-Sketch [9] for sampling and Bitmap [10] for cardinality estimation. Other research, such as SA-Sketch [11], uses a reversible sketch based on the majority vote algorithm (MJRTY) [12] gives acceptable accuracy when using relatively high memory. Some research, such as ExtendedSketch [13, 14] relied on a dynamic mechanism to increase the size of counters for flows with higher distribution and using the Chinese Remainder Theorem to reverse Sketch. Research such as Treesketch [15] has used reversible Sketch and an estimation mechanism based on the heavy cuckoo algorithm [16] to estimate the highest K flow (based on cardinality). The FlowFight mechanism [17] used the Stream Summary data structure [18] and multi-Hyperloglog estimators [19] to maintain the top- k superspreaders. GMF [20] introduces new techniques, including a generalized geometric hash function, counter, and filter to focus on identifying super spreaders by blocking out insignificant flows. However, a fundamental challenge remains with network monitoring points' limited memory and processing capabilities. It appears that the accuracy of detecting the superspreaders is entirely related to the memory size.

3 Background

3.1 *Distinct Estimation (Cardinality Counting)*

The challenge of estimating cardinality (count of distinct elements) arises in various contexts. It involves dealing with a multiset E containing N items drawn from a universe U , where items may repeat. Each item $e_i \in E$ is associated with a count m_i , representing the number of occurrences of the corresponding value $i \in U$ within E . Estimating cardinality requires identifying the unique values present in E , thus necessitating the removal of duplicates. This task is inherently complex because it demands a method to retain information about observed elements for deduplication, unlike simply counting occurrences in flow size measurement [21].

3.2 *Superspreaders Detection*

The process of handling a stream of packets involves dealing with source-destination pairs (x, y) , where each pair is only processed once. For instance, x is the source and can include various packet header fields such as the source IP address or its combination with other fields like port. Also, we have the destination marked similarly (y) . We assume that both the sources (x) and destinations (y) are unique keys taken from an integer domain $\{0, 1, \dots, n - 1\}$, and each key represented within $\lceil \log_2(n) \rceil$ bits. Note that a single (x, y) pair will often appear multiple times in the stream. In general, a host can be defined as a Superspreader if the count of connections associated with this host is abnormal and exceeds a specific limit that is unacceptable in the network ecosystem, whether this host is the recipient or the source. This delineation of the Superspreader problem can be dissected into two primary definitions contingent upon the method of detection: one that hinges on a fixed or dynamically adjusted threshold and another that centers around identifying hosts based on a top-k ranking approach.

3.2.1 *Superspreaders Threshold Version*

The threshold version of the Superspreaders problem is structured as follows. Detection occurs periodically at fixed intervals referred to as epochs. Here, we define a source's cardinality (fan-out) as the count of unique destinations it connects to during an epoch. A source (host) is labeled a superspreader if its cardinality surpasses a predetermined threshold. Specifically, let ϕ ($0 < \phi < 1$) represent a predefined fractional threshold for discerning superspreaders within a packet stream. A source x qualifies as a superspreader if its cardinality or fan-out $F(x)$ exceeds $\phi * \text{Sum}$, i.e. $[F(x) \geq \phi * \text{Sum}]$, where Sum is the total fan-out (cardinality) of all sources within one epoch [8]. In practical implementation, we can derive the source fan-out and the total fan-out (Sum) using distinct counting techniques such as bitmap [22], LogLog (LL) [23], Linear Counter (LC) [24], HyperLogLog (HLL) [19] and Adapt Counter (AC) [25], which enables accurate estimation of the number of unique source-destination pairs while conserving memory space. The length of the intervals is assumed to be manually configured by the network admin. However, thresholds can be configured manually or dynamically.

3.2.2 *Superspreaders Top-k Version [17]*

The top-k spreader problem involves formalizing a network stream P , consisting of N packets, denoted as $P = \{p_1, \dots, p_N\}$. Each packet, p_j , is associated with a specific flow $f_i \in F$, where F represents the set of flows as $F = \{f_1, \dots, f_K\}$. Every packet contains an element denoted as $e \in E$. The cardinality (fan-out) of a flow, denoted as $F(f_i)$, refers to the count of unique elements e contained within the

packets of flow i . The flows $\{f_1, \dots, f_K\}$ are ordered based on their cardinalities so that $F(f_1) \leq F(f_2) \leq F(f_K)$. In our research, we will consider the value e as the destination IP, and the source IP will be the flow identifier.

Problem Definition

- **Given:** integer K and stream of packets P .
- **Objective:** to generate a list of the Top-k highest flow cardinalities, represented as $\{f_1, \dots, f_k\}$.

4 Proposed Methodology

4.1 Space-Saving Approach

The basic principle of space-saving technology is to use a K counter to monitor the highest K flows. This technology ensures acceptable accuracy using a limited memory size [26]. Because it is difficult to scan the data structure comprehensively, find the smallest value and replace it to allow for a new incoming flow. It is enough to obtain the smallest value from a set of values stored in the data structure to achieve an acceptable estimate. In our research, we use the hashpipe structure [6] to enable new flows supported by a filtering layer to ensure the principle of a single counter to monitor a single flow.

4.2 Main Data Structure

Figure 1 shows the main structure we designed to detect Superspreaders. The structure consists of two main sections. The first section consists of two filters: (1) Bloom filter and (2) SS-Filter (space-saving filter). The second section consists of a multi-stage data structure [6] in which the flows are supposed to be top-k-superspreaders. Each incoming packet will be analyzed to obtain two keys (Key1 and Key2). *Key1* represents the packet and will be a combination of (srcIP, dstIP) or obtained using 5-tuple arithmetic. *Key2* is the srcIP of the packet.

A multi-stage data structure consists of N stages with an R counter (entry). Each stage is linked to one hash function where we have a list of hash functions $[h_1, h_2 \dots h_N]$. Each entry in the multi-stage data structure includes two fields [id, count], where id represents the srcIP of a host, and count expresses the Cardinality of this host.

The multi-stage data structure is derived from the HashPipe data structure proposed in research [6]. We developed it to align with our research goal of detecting the top-k superspreaders flows. This structure works on two primary principles. The first one is cramming the new flow into the first stage to give it a chance to compete. Thus,

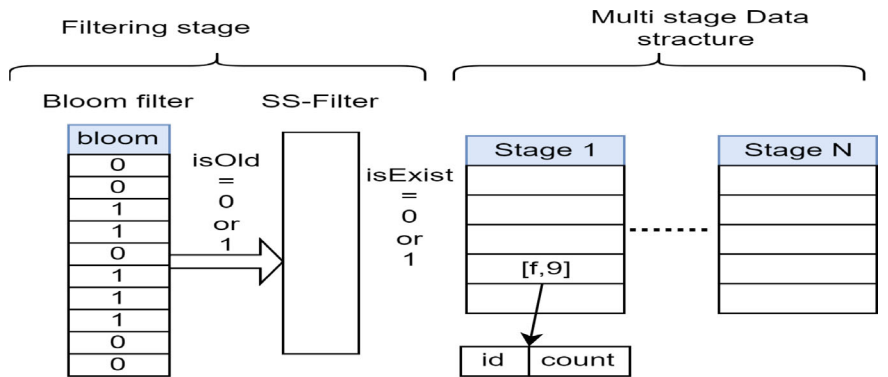


Fig. 1 The main data structure

any new packet belonging to an unbuffered flow will be inserted in the first stage. The second one is continuous monitoring of the minimum value across sequential stages. This process ensures that flows with the highest cardinality are prioritized while displacing those with lower cardinality. Upon introducing a new flow into the initial stage, if the corresponding entry following the application of the stage’s hash function (h_1) is occupied, the value within the first stage is shifted to the subsequent stage. Here, the hash function associated with the second stage is employed to compare the current entry and the existing value. This iterative process preserves flows with the highest cardinality while displacing those with lower cardinality. This sequence continues until reaching the final stage, aiming to curate a roster of flows earmarked for inclusion in the definitive Top-k list.

4.3 Basic Operations

The principle of our strategy for each arriving packet depends on the output of the two filtering stages, where the output of the Bloom filtering stage indicates whether this packet is new and has not been received during a current interval (*isOld value*). The second filtering stage checks whether there is an entry allocated to the flow to which the arriving packet belongs (*isExist value*). Figure 2 shows the procedure followed based on each of the four cases of (*isExist*, *isOld*). Each of these cases will be discussed in more detail next.

4.3.1 Case-1: isOld = 0, isExist = 0

When the programmable switch receives a packet for the first time, a number of BloomFilter hash functions will be applied. The output of the BloomFilter stage will

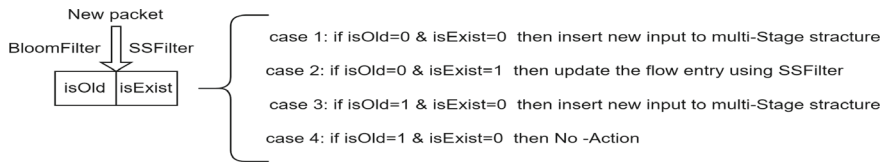


Fig. 2 Basic operations of our strategy based on filtering levels outputs

be ($isOld = 0$), and the output of SSFilter, which depends on the value of Key_2 (src), will be ($isExist = 0$). This confirms that there is no prior entry to the flow related to the packet. In this case, a new entry will be added to the flow using the exact insert mechanism of the hashpipe algorithm [6].

- Figure 3 shows an example of this case as follows:
1. As the L packet is received, BloomFilter’s hash functions are applied to verify whether this packet has been received previously. As the figure shows, the packet is entirely new, and a value ($isOld = 0$) will be returned.
 2. For the SS-Filter level, a number of segmentation functions associated with each stage will be applied. The assumed entries of the received flow will be examined to verify whether an entry for the received packet has been previously stored. Packet L flow was not previously stored in the multistage data structure, so $isExist = 0$.
 3. Due to the lack of a prior entry for the flow associated with the L packet, a new entry will be added to the multi-stage data structure.
 4. To add a new entry, $h1(L.key_2)$ is applied to get position 3. Since the entry in location 3 is not empty, the current value (D) will be carried using p4-metadata to the next stage, and an entry for L with ($Cardinality = 1$) will be inserted into the same location (3).
 5. The D -flow will be transferred to the second stage when $h2(D)$ is applied. We get position 2, but position 2 is not empty; a comparison will be made between

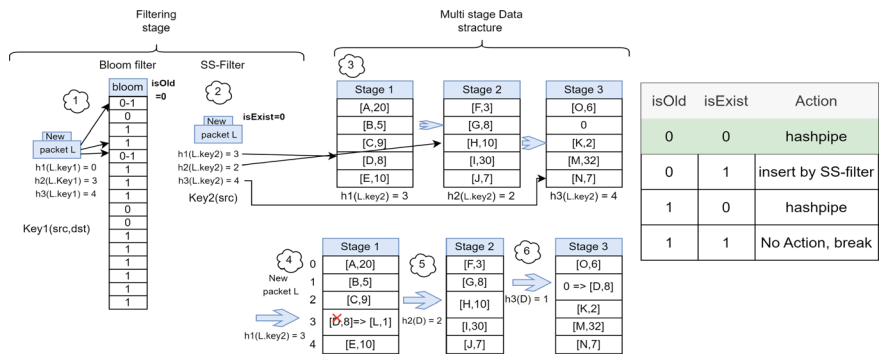


Fig. 3 The process of the strategy when receiving a new packet related to new flow

D-Cardinality and *H*-Cardinality, keeping the flow with the highest Cardinality and transferring the other flow with p4-metadata to the next stage.

6. The previous process will be repeated to reach the last stage.

4.3.2 Case-2: isOld = 0, isExist = 1

In this case, the switch receives a new data packet, but it is associated with a flow that has a previous entry in the multistage data structure. In this case, it is sufficient to update the cardinality of this flow by adding a value of 1 to the counter. Figure 4 shows this case where a previously unreceived *M* packet, so the Bloom filter stage output (*isOld* = 0). The SS-Filter stage indicates the presence of an entry for the flow associated with packet *M* in stage 3. Therefore 1 will be added to the cardinality value stored in the entry.

4.3.3 Case-3: isOld = 1, isExist = 0

In this case, the switch receives a packet that is indicated as old by BloomFilter output (*isOld* = 0). The output of the SS-Filter stage indicates that there is no associated entry for this packet in the multi-stage data structure. Figure 5 shows an example of this case: packet *G* is indicated as old and not linked to any previous entry; therefore, a new entry will be inserted into the multi-stage data structure in the same way as previously explained.

4.3.4 Case-4: isOld = 1, isExist = 1

In this case, the packet is marked as old, and there is an entry associated with it in the multi-stage data structure, so no action is required. This packet was previously counted, and the associated Cardinality flow was updated.

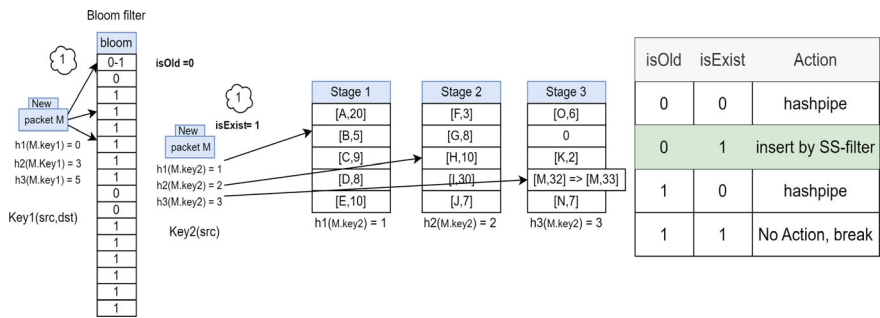


Fig. 4 The process of the strategy when receiving new packet related to existing flow

Algorithm 1: Insert packet

Input: P : incoming Packet; S_i : Stage i register; N :stages Number; $H[1 \dots N]$:hash functions list

Output: $[S_1, S_2 \dots S_N]$: List of all flows proposed for the final Top-k superspreaders list.

```

1 Function InsertProcess():
2    $key_1 \leftarrow 5 - tuple(P)$ 
3    $key_2 \leftarrow P.src$ 
4    $isOld \leftarrow BloomFilter(key_1)$ 
5    $isExist \leftarrow SSFilter(key_2)$ 
6   if  $!isOld \ \& \ !isExist$  then
7     | insertToHashpipe( $key_2$ )
8   end
9   else if  $!isOld \ \& \ isExist$  then
10    | Break; ▷ Already Updated by SSFilter
11  end
12  else if  $isOld \ \& \ !isExist$  then
13    | insertToHashpipe( $key_2$ )
14  end
15  else if  $isOld \ \& \ isExist$  then
16    | Break;
17  end
18 end

19 Function SSFilter( $key_2$ ):
20  for  $i \leftarrow 1$  to  $N$  do
21    |  $j \leftarrow H_i(key_2)$ 
22    | if  $S[j].key = key_2$  then
23      |  $S[j].count \leftarrow S[j].count + 1$ 
24      | return 1 ▷ isExist == 1
25    | end
26  end
27  return 0 ▷ isExist == 0
28 end

```

5 Experiments and Evaluation

This section offers an exhaustive elucidation of the working conditions, dataset, and benchmark utilized.

5.1 Experiments Setup

- **Test-bed:** For our experiments, we utilized a gaming laptop equipped with an Intel Core i5-8300H processor, NVIDIA 4G graphics card GeForce GTX 1650, 8 GB RAM, and a storage combination of 2 TB HDD and 512 GB SSD, running Ubuntu 22.04. Mininet [28] enables the implementation of P4 switches. P4 code compiled

by the p4c tool [29], resulting in a JSON document specifying the P4 switch's behavior, encompassing its parser, tables, and actions within the P4 pipeline. Subsequently, any P4 switch in our topology generated by the behavioral model [7] loads this JSON file.

- **Datasets:** In our experiments, we used CAIAD's anonymous Internet captures, which were gathered on a 10-gigabit Ethernet link monitoring a USA city in 2016 [30]. Each minute of this recorded data consists of approximately 64 million network packets. Our focus lies solely on the initial minutes of these flows, as they will be segmented based on different time windows (precisely, 1-min intervals for our experiments).

5.1.1 Metrics

- ***F1-Score*** is a measure that combines precision and recall through harmonic averaging. As $F1 = \frac{2 * R * Pr}{R + Pr}$, Pr: precision, R: Recall.
- ***Average-Absolute-Error (AAE)***: $\frac{1}{k} \sum_{i=1}^k |\hat{x}_i - x_i|$, where $\hat{x}_i$: Indicates the estimated cardinality of $flow_i$, x_i is the real cardinality value of $flow_i$, and K is the length of top-k superspreaders list. AAE quantifies the average magnitude of the error sum.
- ***Average-Relative-Error (ARE)***: $\frac{1}{k} \sum_{i=1}^k \frac{|\hat{x}_i - x_i|}{x_i}$, ARE and AAE evaluate the accuracy or error of cardinality estimating.

5.1.2 Benchmark

Our experiments will present a comparative analysis of our SSD algorithm with contemporary methods designed for Super-Spreader detection. We evaluate it against state-of-the-art techniques like SpreadSketch [8], leveraging a bitmap-derived algorithm, and the ExtendedSketch [13], incorporating an estimation mechanism derived from a bitmap with provisions of expanding counters for high-cardinality flows. Additionally, we assess the Sasketch [11], which employs a hybrid mechanism integrating multiple prior approaches for network statistics collection. Lastly, we examine the GMF (Geometric MinFilter) [20], employing an estimation mechanism inspired by the geometric hash functions.

5.2 Experiments Results

This section will present some experiments and comparisons between our proposed mechanism and the state-of-the-art solutions for detecting Super-Spreaders.

5.2.1 Experiment 1 (Effect of Increasing Memory)

Figure 6 shows a comparison between our proposed mechanism and the state-of-the-art such as SpreadSketch (SS), Extended Sketch, Geometric-Min Filter (GMF), and (SA) (saSketch-lc). In this comparison, we set a fixed value for k (200), expressing the number of top- k SuperSpreaders detected. The memory size gradually increased by values between 30 and 180 kB. Figure 6a shows the high F1-score value of our SSD algorithm, which increases with the increase of memory size. This is logical, as the increase in memory volume is proportional to the rise in the counters number in each stage, and the F1-score values of our algorithm range between [94% 99%]. With an acceptable and precise F1-score accuracy for the saSketch-lc architecture, which increases with increasing memory volume. The F1-score value for the rest algorithms is very low because the estimators, such as Bitmab and Hyperloglog, need a memory size of at least 250 kB to obtain acceptable accuracy values. Figure 6b shows the low average-Absolute-Error values for our algorithm, especially when using memory

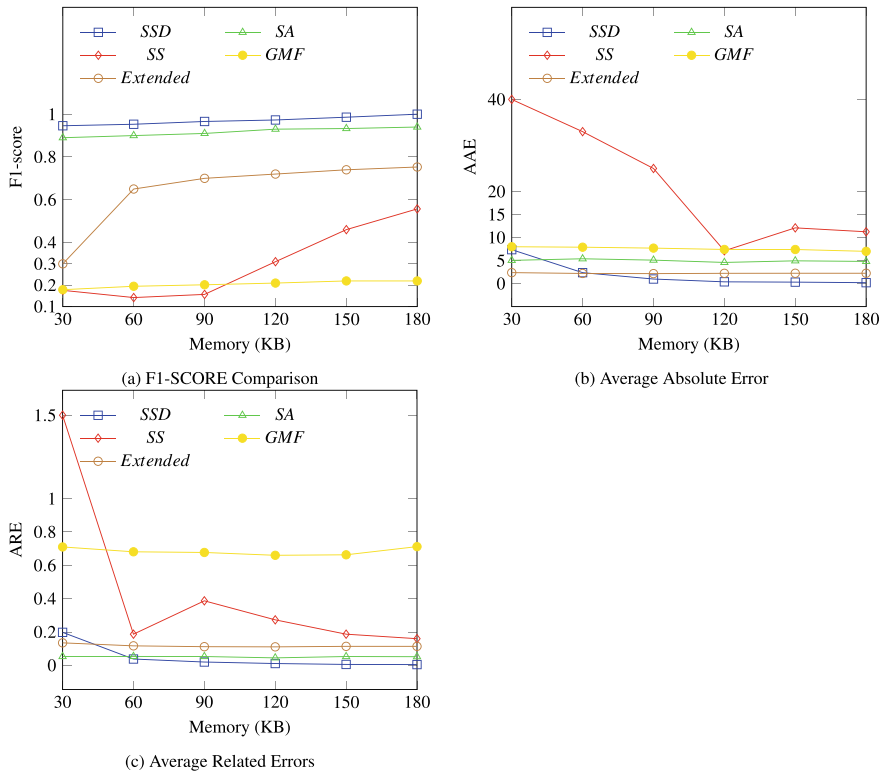


Fig. 6 Accuracy and error rate comparing between (our SSD, SpreadSketch (SS), GMF, Extended Sketch, SA-Sketch (SA)) using multi-memory volumes

greater than (60 kB) compared to competing algorithms. Figure 6c Average-Relative-Error (ARE) also shows the superiority of our algorithm and the low error values with a memory usage of more than (60 kB) with values close to the (SA) approach (saSketch-lc).

5.2.2 Experiment 2 (Effect of Increasing K -Value)

Figure 7 shows a comparison between our SSD mechanism and previous competing mechanisms for detecting top-k SuperSpreaders. We maintain a fixed memory value (180 kB) in this comparison. The main goal is to study the effect of increasing the value K , which represents the top-k SuperSpreaders to be detected. In this comparison, we specified variable values for K ranging in the range (50–350). Figure 7a shows the effect of increasing the value of K on the F1-score, as it is clear that the accuracy of the five algorithms decreases with the increase of the value of K . Our algorithm maintains the accuracy of the F1-score higher than 90% for the value ($K < 250$). We notice low values for the rest of the algorithms, less than 90%, with a value of ($K > 150$). Figure 7b shows the superior Average-Absolute-Error value of our algorithm, which does not exceed 0.5. Thus, our algorithm outperforms significantly the closest competitors. Figure 7c also shows the superior Average-Relative-Error values of our algorithm, which clearly outperform the closest competitors.

6 Conclusion and Future Work

This research proposes SSD (Space-Saving Detector)—a new mechanism for detecting SuperSpreader hosts in SDN networks. Our proposed mechanism provides high efficiency and accuracy detection for Top-k SuperSpreaders. We used two filtering layers to ensure distinct counting of flows and an efficient multi-stage data structure that preserves flows with high cardinality and removes less critical flows. We also applied the proposed solution to programmable SDN switches. Our analytical experiments showed that our mechanism is superior to the best current solutions for detecting SuperSpreaders, as our proposed mechanism can detect Top-k SuperSpreaders with high accuracy using the least possible amount of memory. The experiments also show meager estimation error rates for our algorithm.

In the future plan, we will expand our proposed solution to include network-wide detection of SuperSpreaders hosts. We will deploy more than one programmable switch at different points of the network and propose a mechanism to aggregate the lists of SuperSpreaders from each comparison point into a central point. We will also try to evaluate the accuracy of our mechanism and competing algorithms using other databases, such as (CAIDA19 and CIDA18).

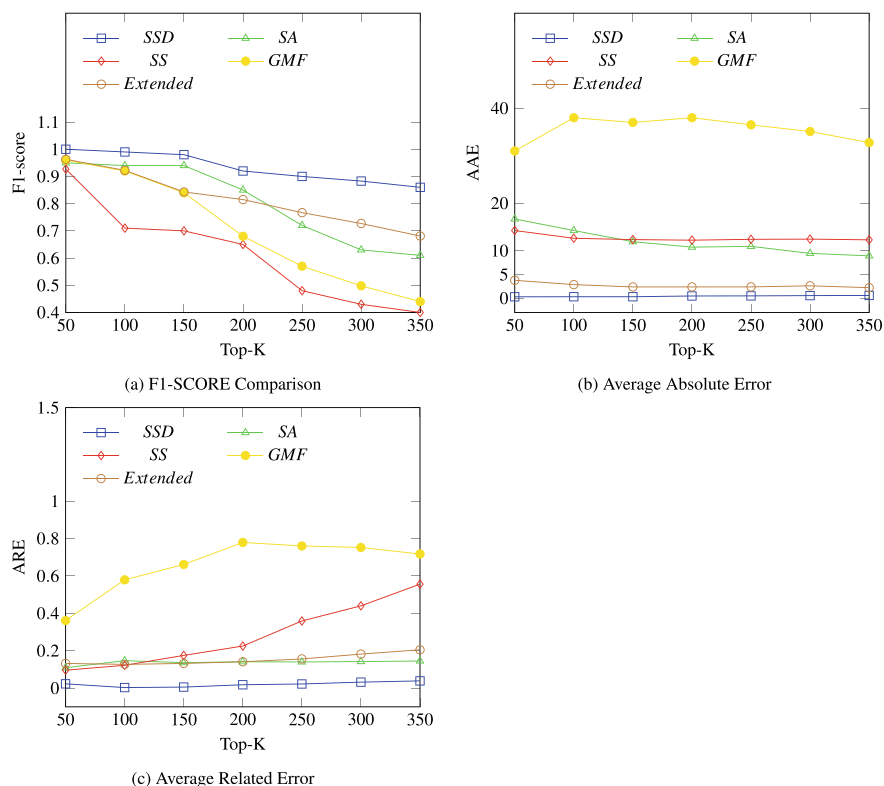


Fig. 7 Study the effect of changing K value

References

1. Alhaj AN, Dutta N (2022) Analysis of security attacks in SDN network: a comprehensive survey. In: Contemporary issues in communication, cloud and big data analytics: proceedings of CCB 2020, pp 27–37
2. Alotaibi FM, Vassilakis VG (2021) SDN-based detection of self-propagating ransomware: the case of badrabbitt. *IEEE Access* 9:28039–28058
3. Sagatov ES, Mayhoub S, Sukhov AM, Esposito F, Callyam P (2021) Proactive detection for countermeasures on port scanning based attacks. In: 2021 17th international conference on network and service management (CNSM), Oct 2021. IEEE, pp 402–406
4. Kfoury EF, Crichigno J, Bou-Harb E (2021) An exhaustive survey on p4 programmable data plane switches: taxonomy, applications, challenges, and future trends. *IEEE Access* 9:87094–87155
5. Sivaraman A, Subramanian S, Alizadeh M, Chole S, Chuang ST, Agrawal A, Balakrishnan H, Edsall T, Katti S, McKeown N (2016) Programmable packet scheduling at line rate. In: Proceedings of the 2016 ACM SIGCOMM conference, Aug 2016, pp 44–57
6. Sivaraman V, Narayana S, Rottenstreich O, Muthukrishnan S, Rexford J (2017) Heavy-hitter detection entirely in the data plane. In: Proceedings of the symposium on SDN research, Apr 2017, pp 164–176

7. P4 Language Consortium (2023) P4 switch behavioral model. <https://github.com/p4lang/behavioral-model>
8. Tang L, Xiao Y, Huang Q, Lee PP (2022) A high-performance invertible sketch for network-wide superspreader detection. *IEEE/ACM Trans Netw*
9. Cormode G, Muthukrishnan S (2005) An improved data stream summary: the count-min sketch and its applications. *J Algorithms* 55(1):58–75
10. Estan C, Varghese G, Fisk M (2003) Bitmap algorithms for counting active flows on high speed links. In: *Proceedings of the 3rd ACM SIGCOMM conference on internet measurement*, Oct 2003, pp 153–166
11. Zhang Z, Lu J, Ren Q, Li Z, Hu Y, Chen H (2024) An accurate and invertible sketch for super spread detection. *Electronics* 13(1):222
12. Boyer RS, Moore JS (1991) MJRTY—a fast majority vote algorithm. In: *Automated reasoning: essays in honor of Woody Bledsoe*. Springer Netherlands, Dordrecht, pp 105–117
13. Han H, Jing X, Yan Z, Pedrycz W (2023) ExtendedSketch+: super host identification and network host trust evaluation with memory efficiency and high accuracy. *Inf Fusion* 92:300–312
14. Jing X, Yan Z, Han H, Pedrycz W (2021) ExtendedSketch: fusing network traffic for super host identification with a memory efficient sketch. *IEEE Trans Depend Secure Comput* 19(6):3913–3924
15. Liu L, Ding T, Feng H, Yan Z, Lu X (2022) Tree sketch: an accurate and memory-efficient sketch for network-wide measurement. *Comput Commun* 194:148–155
16. Fan B, Andersen DG, Kaminsky M, Mitzenmacher MD (2014) Cuckoo filter: practically better than bloom. In: *Proceedings of the 10th ACM international on conference on emerging networking experiments and technologies*, Dec 2014, pp 75–88
17. Bruschi V, Pontarelli S, Tollet J, Barach D, Bianchi G (2021) FlowFight: high performance-low memory top-k spreader detection. *Comput Netw* 196:108239
18. Metwally A, Agrawal D, El Abbadi A (2005) Efficient computation of frequent and top-k elements in data streams. In: *International conference on database theory*, Jan 2005. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 398–412
19. Karppa M, Pagh R (2022) HyperLogLogLog: cardinality estimation with one log more. In: *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, Aug 2022, pp 753–761
20. Ma C, Chen S, Zhang Y, Xiao Q, Odegbile OO (2021) Super spreader identification using geometric-min filter. *IEEE/ACM Trans Netw* 30(1):299–312
21. Harmouch H, Naumann F (2017) Cardinality estimation: an experimental survey. *Proc VLDB Endow* 11(4):499–512
22. Wu K, Otoo EJ, Shoshani A (2006) Optimizing bitmap indices with efficient compression. *ACM Trans Datab Syst (TODS)* 31(1):1–38
23. Durand M, Flajolet P (2003) Loglog counting of large cardinalities. In: *Algorithms—ESA 2003: 11th annual European symposium*, Budapest, Hungary, 16–19 Sept 2003, *Proceedings* 11. Springer Berlin Heidelberg, pp 605–617
24. Whang KY, Vander-Zanden BT, Taylor HM (1990) A linear-time probabilistic counting algorithm for database applications. *ACM Trans Datab Syst (TODS)* 15(2):208–229
25. Cai M, Pan J, Kwok YK, Hwang K (2005) Fast and accurate traffic matrix measurement using adaptive cardinality counting. In: *Proceedings of the 2005 ACM SIGCOMM workshop on mining network data*, Aug 2005, pp 205–206
26. Cormode G, Hadjieleftheriou M (2008) Finding frequent items in data streams. *Proc VLDB Endow* 1(2):1530–1541
27. Bale AS, Yadav KS, Alam M, Shrivastava A, Varma RA, Solanki RS, Savadatti MB (2023) An intelligent 64-bit parallel CRC for high-speed communication system applications. *Int J Intell Syst Appl Eng* 11(10s):543–551
28. Xiang Z, Seeling P (2020) Mininet: an instant virtual network on your computer. In: *Computing in communication networks*. Academic Press, pp 219–230
29. P4lang Consortium: P4C (2023). <https://github.com/p4lang/p4c>
30. CAIDA (2022) The CAIDA anonymized internet traces dataset. <https://data.caida.org/datasets/passive-2019>

Assessment of Advancements in Online Examination Proctoring Systems



Sangeeta Lamba and Neelam Sharma

Abstract An online proctored exam is an Internet-based examination that enables faculty to supervise student's activities in real-time during the online exam session. Equipped with AI-powered technology, the online proctored exam system monitors students, assisting educational institutions in deterring misconduct. The software restricts access to additional browsers, leaving no room for students to cheat. It ensures that students cannot evade the examination through any means. Moreover, the online monitoring software captures random real-time images of students and sends the screenshots to the faculty. Additionally, faculty can monitor student's live videos by requesting them to activate their webcams. In this paper, we focus on the existing techniques used in online examination proctoring systems (OEPS), discuss the future of Automated Online Exams, and also discuss the review process steps and all kinds of inclusion and exclusion criteria. We also compare the existing online examination proctoring software available in market and the metrics and features used by them. At the end based on detailed review, the conclusion and future scope will also be discussed.

Keywords Proctored exam · Monitoring · OEPS · Misconduct · Automated proctoring

1 Introduction

Throughout the COVID-19 pandemic, daily life experienced significant disruptions, especially impacting education and examinations. Traditional methods of teaching and learning were forced to transition entirely to digital platforms. In this unprecedented environment, online examinations became not only a viable but a necessary

S. Lamba (✉) · N. Sharma

Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan, India
e-mail: army.sangeeta46@gmail.com

N. Sharma

e-mail: neelamsharma@banasthali.in

solution. Through three primary proctoring methods—live, recorded, and automated proctoring—these exams ensured the maintenance of academic integrity and enabled fair assessment amid the chaos of the pandemic’s challenges. Traditional examinations are pen-and-paper tests controlled in examination halls by human invigilators. Seating and other administrative activities are critical to the successful management of traditional examinations. Online exam observing is also an important phase of administering any exam because it ensures that applicants do not use any practice of dishonest or unfair means, even during the exam. In-person or in physical pen-paper-based exams proctors are present in the examination room to retain an eye on the applicants. On the other hand, proctoring is quite difficult in online exams. But currently, technical advancements have made online exam invigilating as effective as conventional offline exam proctoring.

According to UNESCO’s 2020 report [1], it was projected that by early April, around 1.5 billion students worldwide, equivalent to 91.3% of the global student population, would face barriers to participating in educational activities due to the rapid and extensive spread of the COVID-19 virus. Consequently, the education system has undergone significant changes, with a noticeable increase in e-learning methods, where teaching is conducted remotely via online platforms. Research indicates that virtual education has proven to enhance information retention and requires less time, suggesting that the changes brought about by the pandemic are likely to be enduring. To reinforce the content, online exams allow you to add videos, images, recordings, PDFs, and other media to a test item. The simultaneous use of multimedia makes an online test more effective for all different kinds of examinees, including those who are good at learning [2]. Online tests have so many in-built topographies to inhibit cheating, such as browser/screen lockdown, which deactivates keys such as screenshot, arrow keys, and ESC keys. This indicates that examinees are unable to save screen captures of the test or revisit previous screens to alter their answers. Another crucial security feature is the dual-proctor 2 login, where the administrator must log in first before granting access to the test for the test-taker. This prevents individuals from viewing the test questions beforehand. The motivation behind our study is to achieve the goal of cheat-free exams. Online proctoring services require artificial intelligence (AI), real-time streaming of video/audio candidates, and invigilator safeguards over assessment of contestants.

As shown in Fig. 1, online exam proctoring can be divided into three main categories as outlined below:

- i. *Live proctoring* replicates the similar classroom experience remotely. An online proctor monitors multiple examinees in real-time, from the beginning to the end of the exam, using live audio and video feeds.
- ii. *Recorded proctoring* utilizes AI-enabled technology to record the student’s screen and audio–video stream during the test without any live monitoring.
- iii. *Automated proctoring* is an AI-enabled, fully automated method of proctoring online exams, eliminating the need for human intervention.

To achieve the goal, several articles and books were widely referred to, and numerous studies were recognized and fully incorporated into this analysis to respond

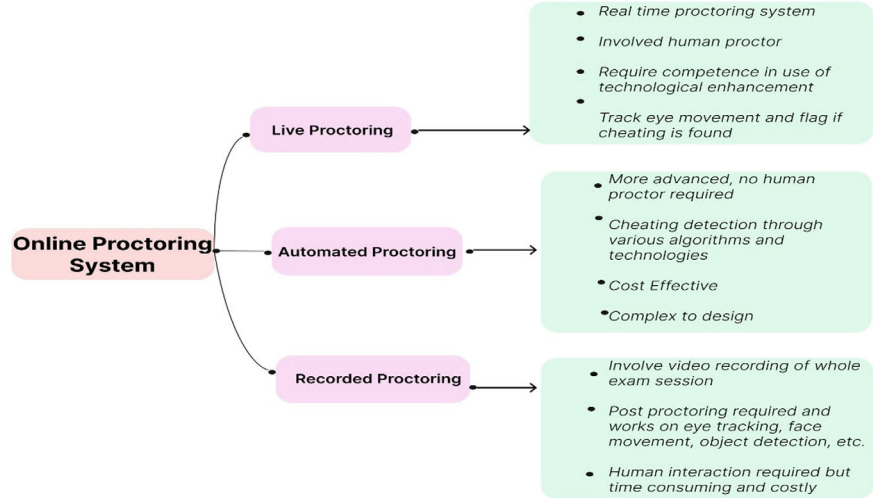


Fig. 1 Types of online proctoring system

to the research questions (RQs) deliberated far ahead in this article. The remainder of the survey is ordered as like: The methodical method for the assessment is presented in Sect. 2. The project planning for the systematic literature review (SLR) method is described in Sect. 3. Section 4 sketches the conducting/directing phase which consists of emerging selection criteria, eliminating facsimile studies, and so on. Section 5 scrutinizes the discoveries and returns to the research questions that were outstretched although evaluating the readings. Section 6 contains the paper’s decision, which is preceded by Sect. 7’s future scope.

2 Review Methodology

This section outlines the process of conducting the review, specifically focusing on the systematic literature review (SLR) that recognizes, evaluates, and interprets all existing research pertaining to cheating detection and emotion recognition. The assessment methodology, as depicted in Fig. 2, consists of three main stages:

Planning, conducting, and reporting. During the planning phase, the search database stores were determined, research questions (RQs) are formulated, and appropriate papers related to the field of study were collected. In conducting stage, papers were carefully screened, and duplicates and irrelevant papers were removed, ensuring adherence to the inclusion criteria. The selected studies were then integrated based on the statistics provided in each of them. Finally, in the reporting stage, the research questions were addressed, and the findings were presented.

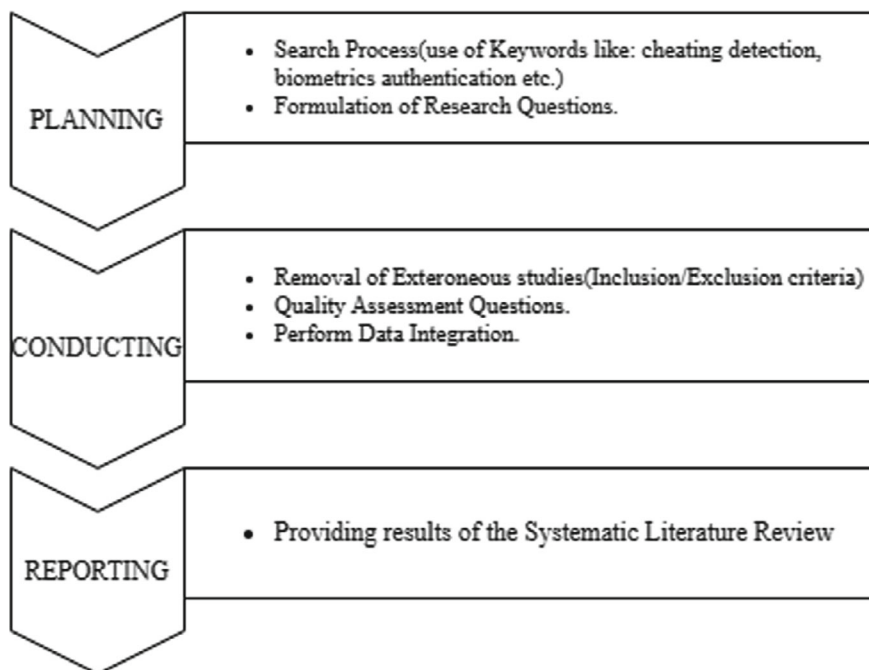


Fig. 2 Review/analysis process

3 Planning Phase for Review

3.1 Search Process

The primary hunt commenced by discovering diverse operational databases/records, electronic papers, and (SI) search engines. The examine process involved binary steps. In the first step, the following databases were identified: Scopus, Google Scholar, Science Direct, IEEE Xplore, and Springer. Appropriate search strings such as “Automated Online Exam Proctoring Systems,” “Automated Invigilation,” “AI-Based Proctoring,” “Cheat detection,” “Biometrics Authentication,” and “Misconduct During Online Exams” were recycled to cite an appropriate study.

The examination series that remained molded using the search expressions is provided below: (Automated Online Exam Proctoring Systems OR Automated Invigilation OR e-Exam Proctoring) **AND** (Cheat Detection OR Fraud OR e-Cheating OR Cheat Resistant OR Misconduct OR Dishonesty) **AND** (e-Exam OR Assessment OR e-Assessment OR e-Evaluation) **AND** (Virtual Exam OR Digital Exam OR Electronic OR e-Environment) **AND** (Control OR Detect OR Prevent OR identify OR Minimize OR Monitor OR Deter)

Table 1 Research questions

Attributes	Research question
What are the various types of online exam proctoring systems?	RQ1
What are the technical aspects that are available in online exam proctoring systems?	RQ2
What are the types of cheating/misconduct detection techniques used and how to prevent the cheating?	RQ3
What is the mostly used online exam proctoring software?	RQ4
Which is the online exam proctoring software based on biometrics authentication?	RQ5

In the succeeding step, we explore the papers that are appropriate to distinguished journals and discussions/conferences that were selected. In addition, the exploration was constrained to the durations of years between 2018 and 2024.

3.2 Preparation of Research Questions

While evaluating the selected studies, several research questions (RQs) emerged, providing valuable insights into both historical and current trends in Online-Based Exam Proctoring Systems. These RQs have been meticulously crafted to ensure a systematic progression of the study, preventing any deviation during the analysis of the extensive data collected from the chosen papers. The list of RQs for this review is presented in Table 1.

4 Conducting Phase for Review

The subsequent segment comprises the actions achieved during the conducting phase of SLR. It presents the collection principles used to clean out the extraneous studies. In addition, a survey was formed to mark the selected studies as per the RQs declared in Table 1.

4.1 Inclusion/Exclusion Criteria (Removal of Extraneous Studies)

Readings were structured in a well-ordered fashion, and replica and unrelated studies were plummeted from the analysis procedure. Based on the RQs listed in Table 1, the inclusion principles were set to choose those studies that included some content on

some aspects of online exam proctoring systems or cheating detection. The papers that focused on online exam cheating detection techniques to identify misconduct during examination, different types of authentication techniques before start of exam, detection of mobile, presence of more than one persons in room, head movement, eye tracking, and prediction models that categorize opportunities for cheat detection, practiced the inclusion criteria.

Inclusion criteria

- Various practical studies have explored and recognized distinct methods for detecting online exam cheating. Additionally, these studies have investigated diverse types of authentication methods.
- Some of these empirical studies have also highlighted the possibility of multiple individuals being present during exams.
- Publication of studies is limited to the period between January 2018 and February 2024.
- In instances where multiple papers documented the same research, only the latest ones were considered.
- Moreover, certain studies have proposed the use of head movement, eye tracking, lip movements, and other factors for detection purposes.

Exclusion criteria

- Studies that did not concentrate on online exam cheating detection techniques and authentication methods.
- Studies that were unable to address the research questions (RQs) outlined in Table 1.
- Studies published prior to 2018.
- Readings that did not present any experimental results.
- Studies of low quality are excluded, including those published by non-reputable publishers without peer review, those with excessively short review times, and those featuring inadequate theoretical background, experimental evaluation, or structural elements.

4.2 Data Extraction

In the conducting phase, data that addressed all the research questions (RQs) was extracted. A comprehensive database was created, encompassing various structures such as publication year, study name, author's name, publication/journal source, learning objectives, outcomes, tools used, practical procedures, and more. Specifically, five attributes were chosen to address the RQs, and these are presented in Table 1. All the nominated readings underwent analysis from different perceptions, with an effort to recognize any connections or relationships between them. The data was visually represented using pie charts, bar charts, line graphs, etc. The submission

of visualization techniques aids the reader in easily comprehending and interpreting the data.

4.3 Distribution of Papers

The chosen readings were categorized into the following three groups based on their journal sources: journals, discussions, and other sources like symposiums, book chapters, and workshops. The dissemination of these revisions is visually presented in a pie graph (Fig. 3). Journal publications accounted for the largest portion, with 18 studies, tracked by 6 readings from conferences, and rest readings from several other sources.

5 Reporting Phase

In this segment, we present the outcomes derived from the systematic literature review (SLR). Initially, we provide the quality assessment enquiries along with the scores assigned to the chosen studies. Following that, we present the answers to the research questions (RQs). The evaluation of the selected studies was conducted by a team consisting of two members, including one associate professor and one research scholar from Banasthali Vidyapith, Rajasthan, India. Selected educations were assigned scores as follows:

- 1 score for the thorough qualification of the paper.
- 0.5 score for fractional qualification.
- 0 score for no qualification.

Each revision was counted on a measure of 0 to 1, and the total mark was deliberate by marking them with all the (QA) quality assessment questions. The extreme and lowest grades attained can be 6 and 0. The examination question will be taken from

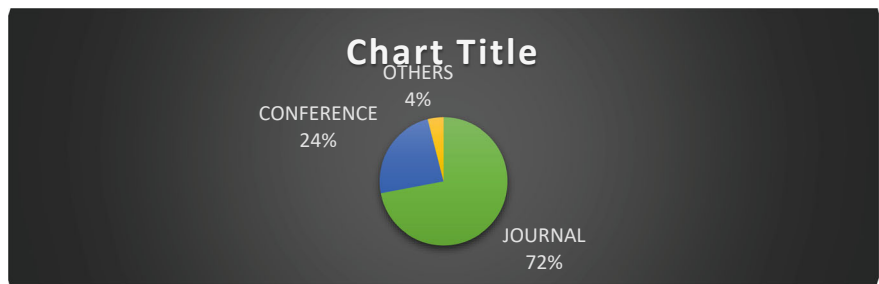


Fig. 3 Dispersal of document as per the source of publication

Table 2 Scores for research questions

Paper ref. no.	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	Total score
[3]	1	1	1	0.5	1	1	5.5
[4]	1	1	0	1	1	0.5	4.5
[5]	0	0	0.5	0.5	1	0	2
[6]	0.5	1	1	1	0	1	4.5
[7]	1	1	1	0.5	0.5	1	5
[8]	1	1	0.5	1	0	0	3.5
[9]	1	1	1	1	1	0	5
[10]	1	1	0.5	1	1	0	4.5
[11]	1	1	0.5	0.5	0.5	0	3.5
[12]	1	1	1	0.5	1	0	4.5
[13]	1	1	0.5	1	0.5	1	5
[14]	1	1	1	0	1	1	5
[15]	1	1	1	0.5	1	1	5.5

RQ1 to RQ5 from Table 1. Here, resulting Table 2 depicts the scores for research questions found during in-depth survey.

5.1 Reporting of Research Questions

The selected studies underwent comprehensive analysis based on the raised research questions (RQs). The RQs provided valuable insights from various angles, facilitating a deeper understanding of these studies. By answering the RQs, we could effectively examine the way out to the problem statement and the execution results, utilizing the information collected. Subsequently, the following subsections present the answers to the RQs along with brief evidences about them.

RQ1 What are the various types of online exam proctoring systems that are detected?

There are several kinds of online exam proctoring systems available today. The system uses webcam and microphone to paramount the test-taker’s environment and behavior during the exam. Honor lock is another automated proctoring system that uses artificial intelligence (AI) and machine learning (ML) to identify suspicious behavior and avoid cheating during online exams [6]. It integrates with various learning management systems and is accessible to students from anywhere with an Internet connection. These systems have their advantages and disadvantages, and it is important to carefully evaluate and choose the most suitable system for the institution’s needs and budget. Table 3 represents the different types of e-proctoring systems that were detected during our survey from 2021 to 24.

Table 3 Existing e-proctoring systems with research purpose

Author/year of publication	Title	No. of citations	Research purpose
Saba et al. (2021) [16]	Categorizing the students' activities for automated exam proctoring using proposed deep L2-GraftNet CNN network and also based feature selection approach	35	Automated exam proctoring with deep L2-GraftNet CNN network and ASO feature selection for student activity categorization
Ahmad et al. (2021) [3]	A novel deep learning-based online proctoring system using face recognition, eye blinking, and object detection techniques	13	Presented a novel online proctoring system that employs continuous deep learning for remote monitoring of physical spaces, eliminating the need for a physical proctor. The system utilizes biometric methods such as HOG face detection and the OpenCV face recognition algorithm for effective face recognition
Potluri et al. (2023) [17]	An automated online proctoring system using attentive-net to assess student mischievous behavior	7	The attentive system surpasses current methods in face detection, pose estimation, and phone detection. Its seamless integration into any online proctoring platform or software includes additional features such as user verification and active window detection to deter malpractice
Kaddoura and Gumaei (2022) [18]	Towards effective and efficient online exam systems using deep learning-based cheating detection approach	10	The goal of this research is to devise an effective method for online exam systems, utilizing deep learning models to detect cheating in real-time through analysis of recorded video frames and speech
Masud et al. (2022) [19]	Smart online exam proctoring assist for cheating detection	14	Introduced a cheating detection method that scrutinizes exam videos, extracting four categories of event data. This data is then input into a pre-trained classification model to identify occurrences of cheating activity
Prabowo et al. (2024) [20]	A data protection design for online exam proctoring in compliance with the Indonesian personal data protection law	5	Proposed a data labeling model with fusion/UML to implement access control and safeguard privacy for personal data, in compliance with the PDP Law. This approach is illustrated through a case study on face recognition-based remote proctoring within an online learning management system at our institution

RQ2 What are the technical aspects available in online exam proctoring systems?

Based on a review of research on various AIPS, researchers have identified parameters that are important for the efficient operation of online examination proctoring systems. These constraints were designated based on the ease of implementation with accessible physical/hardware for students. Some of these parameters include camera, microphone, human proctor, biometrics, recording/screen sharing, browser lock, and gaze tracking. Table 3 illustrates the research conducted on AIPS systems between 2019 and 2024, showing a trend toward the implementation of better software such as tab locking. Older research concentrated more on mounting AIPS with high accurateness based solely on video stream input. Table 4 demonstrates technical aspects for designing online exam proctoring systems.

RQ3 What are the different types of cheating/misconduct detection techniques used/identified?

Online proctoring systems must address various infrastructure issues to guarantee evenhandedness and similarity to traditional pen-and-paper-based examinations. They involve specific hardware components and minimum specifications to be existent in candidate’s devices. As technology advances, privacy concerns arise when connected to the Internet. Several cheating detection techniques, such as user authentication, text detection, voice detection, active window detection, gaze estimation, and phone detection, are used by different online proctoring systems [7]. In 2023, Potluri conducted a study on an AI-driven invigilation system known as the “attentive framework,” designed for virtual assessments. This system utilizes live video feeds of examinees and is comprised of four key components: appearance identification, multiple person identification, face spoofing detection, and head posture estimation, all aimed at detecting malpractices during the assessment process [17]. Some advanced systems use biometric authentication and methods like keystroke recognition and 360° sourcing cameras for cheating identification. Multimedia analytics systems are used for automatic and continuous online exam proctoring to preserve academic reliability [12]. In 2022, Masud et al. transformed the majority of videos

Table 4 Technical aspects of AIPS

Paper	Mic	Camera	Biometrics	Browser lock	Gaze tracking	Human proctor	Recording/ screen sharing
[8]	✓	✓	✓	✓	✗	✓	✓
[12]	✓	✓	✓	✓	✓	✗	✗
[11]	✓	✓	✗	✓	✗	✓	✗
[10]	✓	✓	✓	✓	✓	✓	✓
[15]	✓	✓	✗	✓	✓	✗	✗
[21]	✗	✓	✗	✗	✓	✓	✗
[9]	✗	✓	✗	✗	✓	✓	✓

into multivariate time-series data. They achieved this by extracting time-varying occurrence information from each frame of the input video. This approach enabled them to treat the issue of detecting dishonesty as a problem involving the ordering of multivariate time sequences [19]. The system provides real-time proctoring to identify dishonest performance during the test. Audio-visual observations are used to detect any cheating behavior.

RQ4 Mostly used online exam proctoring software available in market with their features

When evaluating e-proctoring features, it is essential to consider four key factors and compare them across systems [22, 23]. First, there are differences in the basic requirements and controls of different online proctoring systems. For instance, some systems may require specific technologies or Internet speeds, while others may have varying degrees of control over the exam, such as the ability to start, pause, or end it. Second, e-proctoring systems can employ different types of authentication, ranging from continuous monitoring to a one-time authentication at the launch of the exam. Third, many e-proctoring exam systems offer a “lockdown” feature to prevent students from accessing external resources during the exam. The specifics of this feature can vary, with some systems only locking the Internet browser and others taking full control of the student’s device, including disabling keyboard shortcuts and peripheral devices. Fourth, some e-proctoring systems generate notifications during the exam session to alert human proctors to any suspicious activities by the examinee [24]. Table 5 summarizes the features of various commercial e-proctoring software.

RQ5 Online exam proctoring software based on biometrics authentication

In online environments, user authentication is a vital aspect of guaranteeing the identity of the examinee and verifying that they are who they claim to be. Common authentication methods include username, password, timestamp, challenge question, and biometric authentication. However, username and password authentication is considered the weakest and insecure method, so finding a more secure and robust alternative is essential. The authentication process should also be user-friendly and straightforward to apply without causing any inconvenience. E-proctoring systems typically require continuous authentication throughout the exam session to ensure the student’s authenticity. In contrast, systems that employ human proctors may use a 360° camera to authenticate students without requiring any additional authentication methods during the exam session, instead only requiring authentication at the beginning of the exam. Fig. 4 shows several methods used for biometrics authentication.

In 2021, Nigam et al. conducted an extensive survey, asserting that the utilization of various biometric authentication methods in online exams enhances security and reliability [7]. Labayen et al. introduced an application integrating a biometric recognition system for face, speech, and typing patterns. This system is coupled with continuous online exam proctoring to authenticate online student identities persistently. The study recommends adopting such software in remote learning to prevent errors during virtual tests. Future advancements in more robust biometric

Table 5 Feature of various e-proctoring systems [24]

Proctoring features	Proctor U	Respondus	Proctori	AIProctor	Kryterion	Software secure	B Virtual
Live proctors	✓	✗	✗	✗	✓	✗	✓
Scheduling	✓	✓	✗	✓	✓	✓	✗
Training	✓	✗	✓	✓	✓	✓	✓
Encryption	✓	✓	✓	✓	✓	✓	✓
Continuous internet required	✓	✓	✓	✓	✓	✗	✓
Certification	✓	✗	✓	✓	✓	✗	✓
Automated proctoring	✗	✓	✓	✗	✓	✗	✗
Recorded proctoring	✗	✓	✓	✗	✓	✓	
Audio recording	✗	✗	✓	✗	✓	✗	✗
Proctor can view student screen	✓	✗	✓	✓	✗	✗	✓
Prevent proctor to view examinee screen	✗	✓	✗	✗	✓	✗	✗
Pause, stop, or cancel test	✗	✗	✓	✗	✓	✗	✗
Interaction with examinees	✓	✗	✓	✓	✓	✗	✓
Facial biometrics	✗	✗	✓	✗	✓	✗	✗
Voice biometrics	✗	✗	✓	✗	✗	✗	✗
Fingerprint biometrics	✗	✗	✗	✗	✗	✓	✗
Iris biometrics	✗	✗	✗	✗	✗	✗	✗
Lockdown features browser	✗	✓	✓	✗	✓	✓	✗

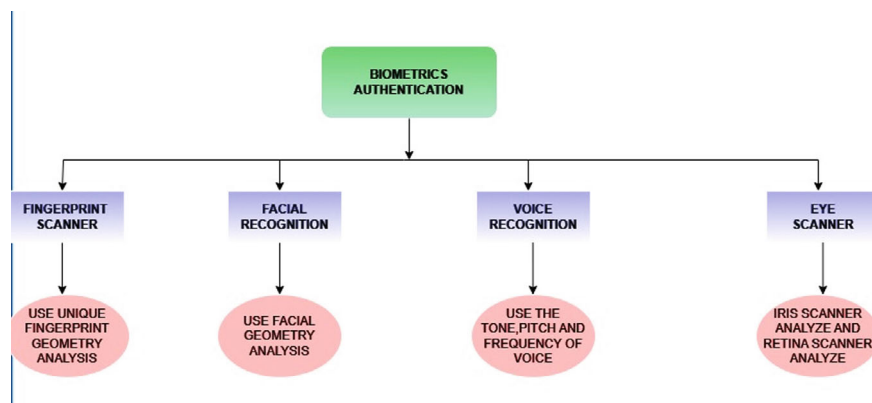


Fig. 4 Biometric authentication methods

frameworks are crucial to address variations in facial posture, lighting, and noise conditions, ultimately reducing the need for human cross-verification for quality assurance [10].

6 Conclusion

In this study, the issue of academic dishonesty among candidates in online learning programs is highlighted. With the easy accessibility of information, online cheating has become a prevalent problem, including the use of fake identities and online resources during exams. As academic integrity is the cornerstone of any university, ensuring fair exams is a top priority. The study focuses on the effective use of technology to combat academic dishonesty, which is a significant challenge for universities offering online programs. Maintaining academic rigor is crucial to ensure the value of the degree. One of the biggest experiments faced by online proctoring systems is to maintaining privacy and security for both the students and the institution. With the increasing use of biometric data and other sensitive information, it is important to ensure that the systems used are secure and cannot be easily hacked or compromised. Some challenges which are observed with the growing need of e-exams are as follows:

- First challenge is the cost of implementing and maintaining these systems. Many universities, especially those in developing countries, may not have the resources to invest in expensive technology or software. Therefore, it is important to find cost-effective solutions that can still ensure the integrity of online exams. One potential solution to these challenges is to use a combination of different types of proctoring systems. For example, using live proctoring for high-stakes exams

and automated proctoring for lower-stakes exams may be a more efficient and cost-effective approach.

- It is also significant to deliberate the impact of these systems on learners. Some students may feel uncomfortable or even violated by the use of biometric data or constant monitoring during exams. Therefore, it is important to educate students about the purpose and benefits of these systems and to ensure that their privacy rights are protected.
- Finally, as technology continues to evolve, it is important to stay up-to-date with the latest developments in online proctoring systems. This includes not only the technology itself, but also best practices for implementation, training, and monitoring to ensure that the systems are effective and efficient in maintaining academic integrity.

However, it has been observed that the performance of online monitoring methods, such as facial recognition software and biometric devices, is not always accurate and may require a multi-faceted approach. Additionally, proper training of the monitoring system is essential for optimal. Further research is needed to investigate the impact of neglecting students' learning styles on cheating motivation in teaching and assessment. Additional studies should address privacy concerns, user convenience, and the associated costs of cheating detection and prevention technologies. While the current study reviews publications from 2019 to 2024, there is a need for further investigations to review accepted but unpublished studies and publications in the year up to Feb-2024.

7 Future Scope

The use of online proctoring systems raises concerns regarding privacy and security, as well as the potential misuse of sensitive biometric data. Additionally, there may be issues of trust in AI-based technology and the cost of implementing these systems, particularly for universities in developing or underdeveloped countries. Lack of training can also pose a challenge. It is essential to balance the prospective benefits of AI-based proctoring with these challenges. While experts believe that AI-based proctoring is necessary to reduce malpractice during online exams, implementation should be done carefully to certify the quality of the testing practice and the credibility of the resulting certificates.

Statements and Declarations

Conflict of Interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. UNESCO education response. Covid-19 impact on education. [Online] Available at URL, Accessed on 12 May, 2020
2. Vincent-Lancrin S, Van der Vlies R (2020) Trustworthy artificial intelligence (AI) in education: promises and challenges
3. Ahmad I, AlQurashi F, Abozinadah E, Mehmood R (2021) A novel deep learning-based online proctoring system using face recognition, eye blinking, and object detection techniques. *System* 12(10)
4. Rodchua S, Yiadom-Boakye G, Woolsey R (2011) Student verification system for online assessments: bolstering quality and integrity of distance learning. *J Ind Technol* 27(3)
5. Yang JT, Liu GM, Huang SCH (2020) Emotion transformation feature: novel feature for deception detection in videos. In: *IEEE international conference on image processing (ICIP)*, October 2020, pp 1726–1730. IEEE Publications
6. Dadashzadeh M (2021) The online examination dilemma: to proctor or not to proctor? *J Instr Pedagog* 25
7. Nigam A, Pasricha R, Singh T, Churi P (2021) A systematic review on AI-based proctoring systems: past, present and future. *Educ Inf Technol* 26(5):6421–6445. <https://doi.org/10.1007/s10639-021-10597-x>
8. Slusky L (2020) Cybersecurity of online proctoring systems. *J Int Technol Inf Manag* 29(1):56–83
9. Harish H, Rajalakshmi D, Ramesh T, Ram SG, Dharmendra M (2021) New features for webcam proctoring using python and Opencv. *Revista Gestão Inovação e Technologies* 11(2):1497–1513. <https://doi.org/10.47059/revistageintec.v11i2.1776>
10. Labayan M, Vea R, Flórez J, Aginako N, Sierra B (2021) Online student authentication and proctoring system based on multimodal biometrics technology. *IEEE Access* 9:72398–72411. <https://doi.org/10.1109/ACCESS.2021.3079375>
11. Motwani S, Nagpal C, Motwani M, Nagdev N, Yeole A (2021) AI-based proctoring system for online tests. Available at SSRN 3866446. <https://doi.org/10.2139/ssrn.3866446>
12. Atoum Y, Chen L, Liu AX, Hsu SDH, Liu X (2017) Automated online exam proctoring. *IEEE Trans Multimedia* 19(7):1609–1624. <https://doi.org/10.1109/TMM.2017.2656064>
13. Karim MN, Kaminsky SE, Behrend TS (2014) Cheating, reactions, and performance in remotely proctored testing: an exploratory experimental study. *J Bus Psychol* 29(4):555–572. <https://doi.org/10.1007/s10869-014-9343-z>
14. Corrigan-Gibbs H, Gupta N, Northcutt C, Cutrell E, Thies W (2015) Deterring cheating in online environments. *ACM Trans Comput-Hum Interact* 22(6):1–23. <https://doi.org/10.1145/2810239>
15. Prathish S, Narayanan SA, Bijlani K (2016) An intelligent system for online exam monitoring
16. Saba T, Rehman A, Jamail NSM, Marie-Sainte SL, Raza M, Sharif M (2021) Categorizing the students' activities for automated exam proctoring using proposed deep L2-GraftNet CNN network and ASO based feature selection approach. *IEEE Access* 9:47639–47656. <https://doi.org/10.1109/ACCESS.2021.3068223>
17. Potluri T (2023) An automated online proctoring system using attentive-net to assess student mischievous behavior. *Multimed Tools Appl* 1–30
18. Kaddoura S, Gumaei A (2022) Towards effective and efficient online exam systems using deep learning-based cheating detection approach. *Intell Syst Appl* 16:200153
19. Masud MM, Hayawi K, Mathew SS, Michael T, El Barachi M (2022). Smart online exam proctoring assist for cheating detection. In: *Advanced data mining and applications: 17th international conference, ADMA 2021, Sydney, NSW, Australia, 2–4 Feb 2022, Proceedings, Part I*. Springer International Publishing, Cham, pp 118–32
20. Prabowo S, Abdurrohman M, Nuha HH, Sutikno S (2023) A data protection design for online exam proctoring in compliance with the Indonesian personal data protection law. In: *Intelligent systems conference*. Springer Nature Switzerland, Cham pp 523–535

21. Li X, Chang KM, Yuan Y, Hauptmann A (2015) Massive open online proctor: protecting the credibility of MOOCs Certificates. CSCW 2015—proceedings of the 2015 ACM international conference on computer-supported cooperative work and social computing, pp 1129–1137. <https://doi.org/10.1145/2675133.2675245>
22. O'Reilly G, Creagh J (2016) A categorization of online proctoring. In: Global learn. Association for the Advancement of Computing in Education (AACE), p 11
23. Foster D, Layman H (2013) Online proctoring systems compared. Caveon Test Security Official Portal, p 11
24. Hussein MJ, Yusuf J, Deb AS, Fong L, Naidu S (2020) An evaluation of online proctoring tools. *Open Praxis* 12(4): 509–525

Drug Discovery Using Variational Quantum EigenSolver



S. Surya Prakash, P. Banu Priya, Sudarshan Someneni, Udendhran, and Arnab Banerjee

Abstract Drug discovery is a complex and time-consuming process that often relies on trial and error. Recent advancements in quantum computing and artificial intelligence (AI) have provided new tools to accelerate this process. In this project, we explore the application of the Variational Quantum Eigensolver (VQE) in conjunction with Graph Neural Networks (GNNs) for drug detection. The VQE algorithm is used to efficiently approximate the ground state energy of molecular systems, which is crucial for understanding molecular properties relevant to drug design. GNNs, on the other hand, are powerful AI models capable of learning complex relationships in graph-structured data, such as molecular structures. By combining these two technologies, we aim to improve the accuracy and efficiency of molecular property prediction compared to traditional methods. Our approach involves encoding molecular structures into graphs and using GNNs to extract meaningful features. These features are then fed into the VQE algorithm to predict molecular properties. We demonstrate the effectiveness of our approach on a dataset of molecular structures with known properties. Our results show that the VQE-GNN model outperforms traditional methods in terms of both accuracy and speed, highlighting the possibility of quantum computing and AI in revolutionizing drug discovery.

S. Surya Prakash · P. Banu Priya (✉)
Department of Networking and Communications, School of Computing, SRMIST,
Kattankulathur, Kanchipuram, Tamil Nadu, India
e-mail: banuprip2@srmist.edu.in

S. Surya Prakash
e-mail: ss6326@srmist.edu.in

S. Someneni · Udendhran · A. Banerjee
Department of Computational Intelligence, School of Computing, SRMIST, Kattankulathur,
Kanchipuram, Tamil Nadu, India
e-mail: sv2349@srmist.edu.in

Udendhran
e-mail: udendhrr@srmist.edu.in

A. Banerjee
e-mail: ab2752@srmist.edu.in

Keywords Quantum · Artificial intelligence · Variational Quantum EigenSolver · Graph Neural Networks · Molecular property · Drug discovery · VQE-GNN

1 Introduction

A difficult and resource-intensive process, drug discovery is essential to the creation of new pharmaceuticals. Traditional drug discovery methods often rely on costly and time-consuming experimental techniques, leading to lengthy development timelines and high failure rates. In recent years, there has been growing interest in leveraging advanced technologies such as quantum computing and artificial intelligence (AI) to expedite the drug discovery process and improve its efficiency.

The joining of artificial intelligence (AI) and quantum computing has led to groundbreaking advancements in various fields, including drug discovery. The project aims to leverage the power of quantum computing, specifically the Variational Quantum Eigensolver (VQE), in conjunction with Graph Neural Networks (GNNs) to revolutionize the process of drug discovery. This innovative approach promises to accelerate the identification of novel drug candidates by providing more accurate and efficient methods for analyzing molecular structures and predicting their properties.

Traditional drug discovery methods often rely on time-consuming and expensive experimental techniques. Quantum computing offers a disruptive alternative by simulating the behavior of molecules, providing a more detailed and accurate representation of their properties. The VQE algorithm, a quantum algorithm used for simulating molecular structures, enables the extraction of quantum features that can be used as inputs for machine learning models.

The project also incorporates GNNs, a type of neural network specialized in processing graph data, such as molecular structures. GNNs have shown great promise in capturing complex relationships within molecular graphs, making them well-suited for predicting molecular properties. By combining VQE and GNNs, the project aims to enhance the accuracy and efficiency of drug discovery processes, ultimately resulting in the creation of novel, more potent medications.

The methodology of the project involves several key steps. Firstly, quantum simulation using VQE is used to generate quantum features representing the molecular structures. These features are then processed by GNNs to predict various molecular properties, such as bioactivity, solubility, and toxicity. Preprocessing data, training models, and evaluation are also included in the project to guarantee the efficacy and dependability of the suggested strategy.

The project's results could have a big effect on the healthcare and pharmaceutical industries. This research has the potential to expedite the development of new medicines for a wide range of ailments by optimizing the drug discovery process and cutting down on the time and expense associated with conventional techniques. Furthermore, the knowledge acquired from this effort may open the door to more

developments in AI and quantum computing, as well as applications in drug discovery and other scientific domains.

2 Literature Survey

In the study, Maheshwari et al. [1] state that this paper Superposition, entanglement, and tunneling are examples of quantum mechanical phenomena that quantum computers (QCs) use to tackle issues beyond the scope of classical devices. The discipline of Quantum Machine Learning (QML) emerged from the application of quantum systems and quantum information approaches in machine learning (ML). QML algorithms provide performance and speed gains that are consistent with Moore's Law. A computational paradigm known as "quantum computation" is predicated on the principles of quantum physics. In the present era of informatics growth, both artificial intelligence (AI) and quantum information are important fields of study.

Abd El-Latif et al. [2] explore the secure encryption of medical images using quantum computing techniques. The paper introduces a novel encryption algorithm based on quantum principles to guarantee the security and confidentiality of medical image data. The algorithm employs quantum encryption keys and quantum operations to encrypt and decrypt medical images, providing robust protection against unauthorized access and tampering. The study demonstrates the effectiveness of the proposed algorithm through simulations and experiments, showing that it can encrypt medical images with high security and efficiency. The algorithm's robustness against various types of attacks, including brute-force attacks and quantum attacks, is also evaluated, confirming its suitability for securing sensitive medical image data. Overall, the study shows how quantum computing may be used to improve the security of medical picture encryption, providing a viable way to safeguard patient confidentiality and guarantee the accuracy of medical data.

Recent developments in the application of quantum computing to improve drug discovery methods are covered in Kumar et al. [3]. It examines how the effective simulation of molecular structures and properties by quantum computing methods such as the Quantum Approximate Optimization Algorithm (QAOA) and Variational Quantum Eigensolver (VQE) is changing drug research. The benefits of quantum computing are emphasized, including its capacity to manage intricate quantum mechanical events that are beyond the capabilities of traditional computers. The study also looks at how to better pick and optimize medication candidates by combining quantum computing with traditional machine learning methods like Graph Neural Networks (GNNs). It explores the potential and difficulties of using quantum computing in drug development in more detail, highlighting the necessity of ongoing study and cooperation between specialists in quantum computing and researchers studying pharmaceuticals. The study highlights the revolutionary potential of quantum computing in speeding up the process of finding new drugs and developing existing ones.

Ullah et al. [4] explores the transformative potential of quantum machine learning (QML) in revolutionizing healthcare. It presents a organized evaluation of emerging perceptions and applications of QML in healthcare, highlighting its ability to address complex healthcare challenges. The paper discusses various applications of QML, including personalized medicine, medical imaging analysis, and drug discovery. It also examines the incorporation of QML with machine learning techniques to improve healthcare outcomes. The study highlights the need for expand research and collaboration in harnessing the full potential of QML in healthcare. Feeds a comprehensive overview of the present landscape of QML in healthcare and identifies future research directions for leveraging this technology to improve patient care.

Yano et al. [5] focus on improving the performance of variational quantum classifiers (VQC) by introducing a novel method for encoding discrete features. The paper addresses the challenge of efficiently representing categorical data, which is common in many real-world applications, including healthcare. The suggested technique encodes discrete features into a quantum state that can be handled by a VQC by combining quantum feature maps with classical pre-processing. Compared to other methods, the method achieves a more efficient representation of discrete features by utilizing the special qualities of quantum computing. Experiments on benchmark datasets are used in the article to illustrate the efficacy of the suggested approach, revealing gains in classification efficiency and accuracy. The study demonstrates how quantum computing can handle complicated data types and be used in a variety of industries, including healthcare.

Yadav et al. [3] investigate how quantum computing has advanced recently and how it can speed up drug research and discovery procedures. It explores how to simulate molecular structures and attributes more effectively than with conventional computers by using quantum algorithms such as the Quantum Approximate Optimization Algorithm (QAOA) and Variational Quantum Eigensolver (VQE). The study emphasizes how quick and accurate chemical interaction and property prediction made possible by quantum computing holds the potential to completely transform the drug discovery process. In order to improve drug candidate selection and optimization, it also addresses how quantum computing can be combined with traditional machine learning methods like Graph Neural Networks (GNNs). In order to fully realize the promise of quantum computing, the paper highlights how it has revolutionized the pharmaceutical business and calls for continued research and cooperation.

Liu et al. [6] introduces ECoFFeS, a software tool that employs the process for selecting features in drug discovery. The software is designed to address the challenge of identifying relevant features from large datasets to advance the accurateness and efficiency of drug discovery processes. ECoFFeS uses a genetic algorithm-based approach to evolve a population of feature subsets, evaluating their performance using a scoring function that incorporates domain-specific knowledge. The software aims to enhance the interpretability of machine learning models used in drug discovery by selecting a subset of features that are most relevant to the prediction task. The

paper demonstrates the effectiveness of ECoFFeS through experiments on benchmark datasets, showing improvements in model performance and feature selection efficiency compared to existing methods.

Li et al. [7] explore the application of quantum models in small fragment drug discovery. It introduces a novel approach that combines quantum computing and generative models to design new drug molecules with desired properties. To generate molecular structures that are then fed into a generative model to predict their bioactivity. This advanced approach influences the quantum nature of molecular structures to increase the accuracy and efficiency of drug discovery processes. The paper demonstrates the effectiveness of the method through experiments on molecular datasets, showing improvements in generating novel and diverse drug candidates compared to traditional methods. Overall, the research highlights the ability of quantum generative models in quickening the innovation of new drugs with improved efficacy and safety profiles.

Explainable Artificial Intelligence (XAI) in the context of drug development is the subject of Alizadehsani et al. [8], who emphasize the technology's importance, present uses, difficulties, and potential future possibilities. While acknowledging the revolutionary effects of artificial intelligence (AI) and machine learning (ML) on drug discovery, it also highlights the need for interpretability and openness as these models get more sophisticated. It is determined that XAI is a cutting-edge method that meets this need by offering a more comprehensible comprehension of ML model predictions. The survey gives an overview of the state-of-the-art in XAI for drug development and addresses the growing interest in using these approaches for this purpose. It discusses different XAI techniques, their uses in drug development, and their possible advantages.

The use of artificial intelligence (AI) algorithms in medical diagnostic systems, with an emphasis on the guiding principles and viewpoints of these systems, was examined by Kaur et al. [9]. The significance of AI in healthcare and its potential to transform medical diagnostics are covered at the outset. After that, the survey goes over the body of research on the application of AI algorithms—like machine learning and deep learning—in medical diagnostic systems. It draws attention to the fundamental ideas behind these techniques, such as feature selection, data preprocessing, and model training. The report also discusses perspectives on the advantages, difficulties, and potential uses of AI in medical diagnostics. It highlights how important it is for medical practitioners and AI specialists to work together to create dependable and efficient diagnostic systems. All things considered, the literature review offers a thorough summary of the state-of-the-art in AI-based medical diagnostic systems at this time, providing information on their foundations, uses, and possible effects on healthcare.

3 Comparative Analysis

Our model differs from existing drug computation systems in several keyways:

1. **Integration of Quantum Computing:** Unlike traditional systems that rely solely on classical computing, our model incorporates quantum computing using the Variational Quantum Eigensolver (VQE) algorithm. This allows us to mimic molecular things at the quantum level, providing more accurate and detailed information for drug discovery.
2. **Use of Graph Neural Networks:** Our model utilizes Graph Neural Networks (GNNs) to process the quantum features extracted by VQE. GNNs are specifically designed to handle graph-structured data, such as molecular structures, allowing us to capture complex relationships within molecules more effectively than traditional machine learning models.
3. **Focus on Feature Extraction:** A key focus of our model is on the extraction of quantum features using VQE. We can extract features that capture the underlying quantum properties of molecules, which are not easily accessible with classical methods.
4. **Emphasis on Predicting Molecular Properties:** Our model is designed to predict various molecular properties, such as bioactivity, solubility, and toxicity, using the extracted quantum features. This predictive capability is crucial for identifying potential drug candidates and assessing their effectiveness.
5. **Potential for Accelerated Drug Discovery:** By combining machine learning and quantum computing techniques, our model can accelerate the drug discovery process significantly. It can provide researchers with more accurate predictions and insights into molecular properties, enabling them to identify promising drug candidates more efficiently.

Overall, our model signifies a novel methodology to drug discovery that leverages the ability of machine learning and quantum computing to overcome the limitations of existing computational methods. It offers a more accurate, efficient, and comprehensive way to analyze molecular structures and predict their properties, eventually indicating the growth of new and more successful drugs.

4 Methodology

4.1 Methodology for Variational Quantum Eigensolver (VQE)

1. **Molecular Hamiltonian Preparation:** The first step is to prepare the molecular Hamiltonian, which describes the molecular system's energy in terms of its electronic and nuclear components. This involves encoding the molecular structure into a quantum circuit, representing the electronic structure of the molecule.

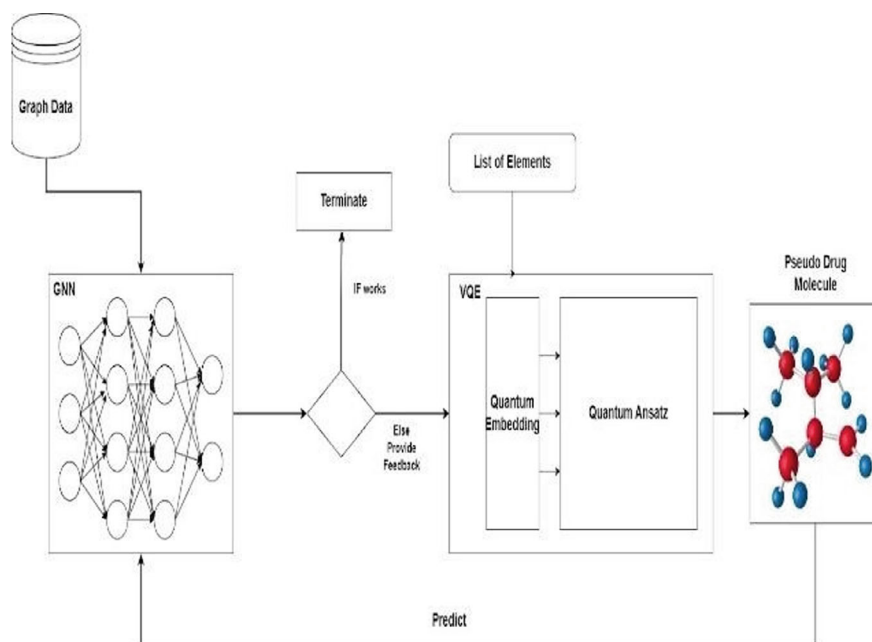


Fig. 1 VQE architecture

2. **Quantum Circuit Initialization:** Initialize a quantum electric circuit with a set of constraints that define the state of the system. These parameters are optimized to minimize the energy of the system, corresponding to the ground state energy of the molecule.
3. **Quantum Circuit Execution:** Execute the quantum circuit on a quantum computer to measure the expectation value of the Hamiltonian. This provides an estimate of the energy of the molecular system for the current set of parameters.
4. **Classical Optimization:** Use a classical optimizer to adjust the parameters of the quantum circuit to minimize the energy. This optimization process is iterative, with the parameters being updated until a convergence criterion is met.
5. **Energy Estimation:** After convergence, the final set of parameters provides an estimate of the molecule. This energy value can be used as a quantum feature for further analysis in the GNN (Fig. 1).

4.2 Methodology for Graph Neural Networks (GNN)

1. **Graph Representation:** Represent the molecular structure as a graph, nodes and bonds are edges. Each node and edge is associated with features, such as atom type, bond type, and distance.

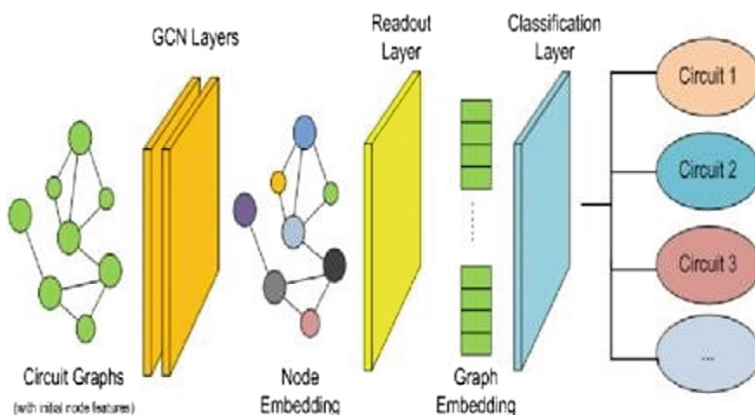


Fig. 2 GNN architecture

2. **Node Embedding:** Initialize node embeddings that capture the local chemical environment of each atom. This embedding is learned during the training process and is used to represent the atom's features in the graph.
3. **Message Passing:** Perform message passing between nodes and edges in the graph to update the node embeddings based on the features of their neighboring nodes and edges. This step allows the GNN to capture the relational information within the molecular structure.
4. **Graph Pooling:** Aggregate the node embeddings to obtain a graph-level representation. This aggregated representation captures the global properties of the molecule, which are crucial for predicting molecular properties.
5. **Optimization Module:** This module focuses on optimizing the performance of the machine learning models. It may include hyperparameter tuning, model selection, and ensemble learning techniques to improve the predictive performance of the models.
6. **Training and Optimization:** Train the GNN using labeled data to predict the desired molecular properties. Use methods stochastic gradient descent and backpropagation to optimize the model parameters and improve its predictive performance (Fig. 2).

5 Modules and Features

1. **Data Preparation Module:** This module involves preparing the dataset for preparation and examining the machine learning models. It includes tasks such as data cleaning, normalization, and splitting into training and testing sets.

2. **Quantum Feature Extraction Module:** This module utilizes the VQE algorithm to extract quantum features from molecular structures. VQE simulates the molecular Hamiltonian to obtain the ground state energy, which is used as a feature for the machine learning models.
3. **Graph Construction Module:** The molecular structures are represented as graphs, where atoms and bonds are nodes and edges. The module constructs these graphs from the molecular structures, including assigning node and edge features based on atom types and bond types.
4. **Model Evaluation Module:** This module estimates the execution of the focused GNN exhibits using metrics such as precision, accuracy, F1 score, and recall. It improves assess the effectiveness of the patterns in predicting molecular properties.
5. **Visualization Module:** This module provides visualization tools to interpret the results of the GNN models. It can include tools for visualizing molecular structures and feature importance.
6. **Deployment Module:** This module involves deploying the trained models for real-world applications. It can include packaging the models Graph Neural Network Architecture: Design a GNN architecture that can effectively process the graph representation of the molecule and predict its properties. This architecture may include multiple layers of message passing and graph pooling, as well as attention mechanisms or skip connections to improve performance.

Each module performs a decisive role in the overall workflow of the project, from data preparation to model training and evaluation, ultimately contributing to the discovery of novel drug candidates using quantum computing and machine learning techniques.

6 Results

6.1 Dataset

Creating a dataset for a project on drug discovery using the Variational Quantum Eigensolver (VQE) and Graph Neural Networks (GNNs) involves collecting and curating molecular data that is suitable for training and evaluating the models. Here is a detailed explanation of the dataset requirements for this project.

Molecular Structures: The dataset should include a diverse set of molecular structures represented in a format that can be easily processed by the VQE and GNN models.

Molecular Properties: For training the models, the dataset should include the corresponding properties of the molecules, such as their energies, activity against specific biological targets, or other relevant properties. These properties can be experimental values or calculated values from quantum chemistry simulations.

Data Augmentation: To improve the robustness of the models, data augmentation can be applied to produce additional training examples. This can involve introducing random perturbations to the molecular structures or properties.

6.2 *Evaluation Protocol*

Dataset Splitting: The dataset needs to be separated into training and validation sets, as was previously indicated. The models are trained on the training set, and their performance is assessed on the validation set.

Model Training: Using the proper training algorithms and hyperparameters, train the VQE and GNN models on the training set. While the GNN model is trained to extract features from the molecular graphs, the VQE model is trained to estimate the ground state energy of the chemical systems.

Validation: Apply the relevant metrics to assess how well the training models perform on the validation set. Metrics like root mean square error (RMSE) or mean absolute error (MAE) can be utilized for regression tasks (like estimating molecular energies). Metrics like accuracy, precision, recall, and F1 score can be applied to classification tasks (such as bioactivity prediction).

Comparison: Compare the execution of the VQE-GNN model with baseline models or traditional methods. This can help assess the improvement in performance achieved by using quantum computing and AI techniques.

Hyperparameter Tuning: Perform hyperparameter tuning to improve the execution of the models further. This can involve tuning parameters such as learning rates, batch sizes, and model architectures.

Cross-Validation: To ensure the robustness of the models, perform k-fold cross-validation on the dataset. To do this, the dataset is divided into k subsets. The models are then trained on k-1 subsets and assessed on the remaining subset. Carry out this procedure k times, utilizing a distinct subset as the validation set each time.

Error Analysis: Conduct error analysis to identify the sources of errors in the models. This can involve analyzing misclassified or poorly predicted examples to gain insights into areas for improvement.

6.3 *Result*

Cost function graph image optimization involves the process of minimizing a cost function graph to find the optimal parameters of a model. In the perspective of machine learning, this optimization is often performed during the training of a model to decrease predicted output and the actual output the difference (Fig. 3).

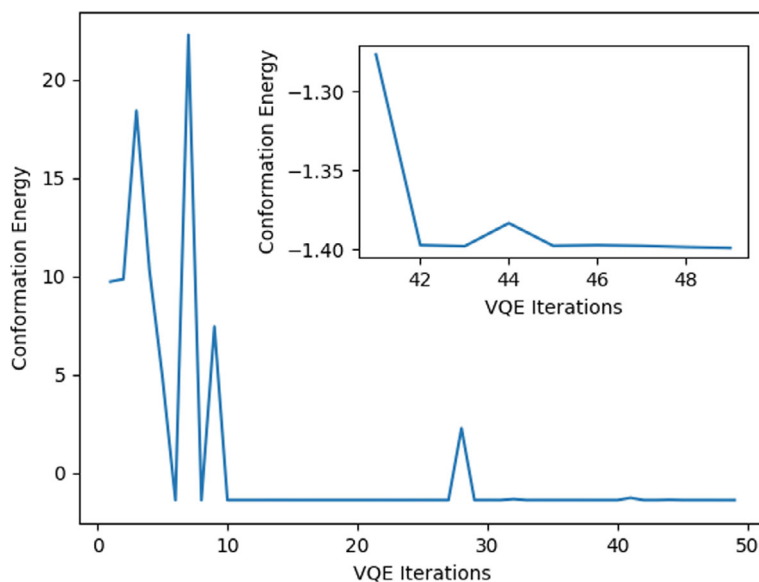
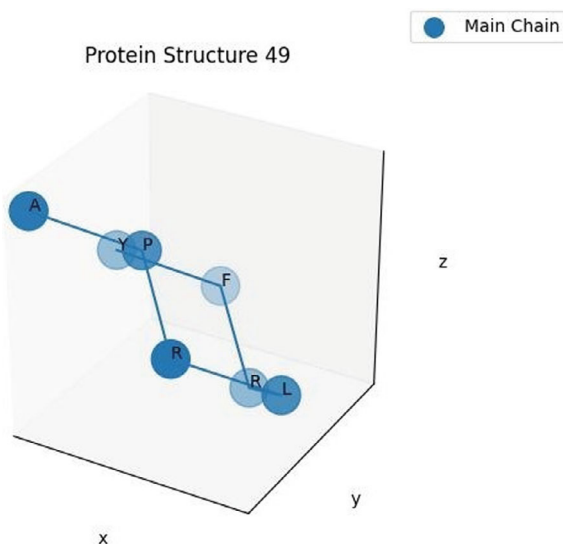


Fig. 3 Confirmation energy versus Iterations graph

The cost function graph depicts the link between the model's parameters and the cost function, showing how changes in the parameters effect the cost. Finding the parameters that produce the lowest cost—a sign of the model's best fit to the data—is the aim of optimization (Fig. 4).

Fig. 4 Molecular structure at its stable state



Several optimization procedures, such as gradient descent, which iteratively modifies the parameters in the direction that lowers the cost, can be used to optimize the cost function graph. The optimization process can also be enhanced by using other algorithms, such as Adam or RMSprop, which modify the learning rate in response to gradient changes.

7 Conclusion

This study presented a novel method to drug discovery using Variational Quantum Eigensolver (VQE) and Graph Neural Networks (GNNs). Our methodology demonstrates the effectiveness of integrating machine learning and quantum computing techniques to predict molecular properties and accelerate the drug discovery process. By leveraging VQE to extract quantum features from molecular structures and employing GNNs to predict bioactivity, solubility, and toxicity, we have achieved capable results in terms of predictive efficiency and accuracy.

Proposed approach offers various advantages over usual methods, including the ability to capture complex molecular interactions and provide more accurate predictions. Additionally, our methodology is highly scalable and can be applied to a wide range of molecular structures and properties.

While our study has shown promising results, there are several avenues for future research. These include further optimization of the VQE and GNN models, exploration of additional molecular properties, and integration of other quantum algorithms. Additionally, the scalability and generalizability of our approach should be further investigated to ensure its applicability to real-world drug discovery efforts.

All things considered, our study demonstrates the promise of machine learning and quantum computing in drug discovery and lays the groundwork for more research in this fascinating and quickly developing area.

References

1. Maheshwari D, Garcia-Zapirain B, Sierra-Sosa D (2022) Quantum machine learning applications in the biomedical domain: a systematic review. *IEEE Access* 10:80463–80484. <https://doi.org/10.1109/ACCESS.2022.3195044>
2. Abd-El-Latif AA, Abd-El-Atty B, Talha M (2018) Robust encryption of quantum medical images. *IEEE Access* 6:1073–1081. <https://doi.org/10.1109/ACCESS.2017.2777869>
3. Kumar G, Yadav S, Mukherjee A, Hassija V, Guizani M (2024) Recent advances in quantum computing for drug discovery and development. *IEEE Access* 12:64491–64509. <https://doi.org/10.1109/ACCESS.2024.3376408>
4. Ullah U, Garcia-Zapirain B (2024) Quantum machine learning revolution in healthcare: a systematic review of emerging perspectives and applications. *IEEE Access* 12:11423–11450. <https://doi.org/10.1109/ACCESS.2024.3353461>
5. Yano H, Suzuki Y, Itoh KM, Raymond R, Yamamoto N (2021) Efficient discrete feature encoding for variational quantum classifier. *IEEE Trans Quantum Eng* 2:1–14

6. Liu ZZ, Huang JW, Wang Y, Cao DS (2018) ECoFFeS: a software using evolutionary computation for feature selection in drug discovery. *IEEE Access* 6:20950–20963
7. Li J, Topaloglu RO, Ghosh S (2021) Quantum generative models for small molecule drug discovery. *IEEE Trans Quantum Eng* 2:1–8
8. Alizadehsani R, Oyelere SS, Hussain S, Jagatheesaperumal SK, Calixto RR, Rahouti M, Roshanzamir M, De Albuquerque VHC (2024) Explainable artificial intelligence for drug discovery and development-a comprehensive survey. *IEEE Access* 12:35796–35812
9. Kaur S, Singla J, Nkenyereye L, Jha S, Prashar D, Joshi GP, El-Sappagh S, Islam MS, Islam SMR (2020) Medical diagnostic systems using artificial intelligence (AI) algorithms: principles and perspectives. *IEEE Access* 8:228049–228069

Biomimetic Swarm-Based UAV Collision Avoidance Control



Wenfei Dai, Zhiyi Wang, and Chaijinyang Ding

Abstract This paper introduces a novel multi-leader hierarchical control architecture based on the Reynolds model for UAV swarms. By employing hierarchical graph distances, the system partitions each leader along with its surrounding followers, effectively decomposing the swarm into several localized neighborhood clusters. The interactions among multiple leaders in an obstacle-free environment are meticulously designed, and the efficacy of the swarm's collision avoidance control is tested when dynamic obstacles are introduced. The study further delves into the control interactions between leaders and followers within these local neighborhood clusters. Extensive simulation results underscore that the swarm maintains stability and achieves convergence in both inter-drone collision avoidance and evasion of external dynamic obstacles. This comprehensive approach not only enhances collision avoidance control of UAV swarms amidst dynamic obstacles but also addresses critical formation control challenges encountered during the avoidance process. The findings from this research hold significant implications for the development of more efficient and reliable UAV swarm control systems in dynamic and unpredictable environments.

Keywords Biomimetics · UAV Swarm · Control model

W. Dai · Z. Wang (✉) · C. Ding

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China

e-mail: chn_sherry@nuaa.edu.cn

W. Dai

e-mail: daiwenfei@nuaa.edu.cn

C. Ding

e-mail: dcjy527@nuaa.edu.cn

1 Introduction

With the rapid development of UAV technology, UAV swarms have become an effective means of solving complex tasks. However, in practical applications, UAV swarms face significant challenges in collision avoidance control. This paper studies the collision avoidance control of UAV swarms by constructing a mathematical model for bird flock formation control based on graph distances and addressing related issues of inter-UAV collision avoidance control.

This paper first proposes a multi-leader hierarchical control architecture based on the Reynolds model. Next, this paper employs hierarchical graph distances to partition each leader and its neighboring followers, thus decomposing the swarm into multiple local neighborhood clusters. Interactions between multiple leaders are designed in an obstacle-free environment, and the effectiveness of the swarm's collision avoidance control is confirmed when dynamic obstacles are present. Additionally, the paper verifies and analyzes the control interactions between leaders and followers within these local neighborhood clusters. The research findings contribute to effective collision avoidance control for UAV swarms in the presence of dynamic obstacles and address formation control challenges during the avoidance process.

2 Related Works

In recent years, the application of biomimetic principles in UAV swarm control has become a research priority. Wu et al. [1] developed a starling-inspired flocking control method, enhancing both coordination and obstacle avoidance capabilities. Zhou et al. [2] applied deep reinforcement learning to enable multi-target tracking for UAV swarms, demonstrating enhanced tracking accuracy and efficiency in dynamic environments. Li et al. [3] applied multi-feature analysis combined with gradient boosting decision trees to agricultural applications. Luo et al. [4] introduced Grpavoid, a system for collision avoidance that ensures safe UAV operations. Mir et al. [5] provided insights into the stability of bioinspired UAVs during optimal dynamic soaring, emphasizing aerodynamic efficiencies akin to bird flight patterns. Furthermore, Novák et al. [6] examined rapid collective evasion strategies in self-localized UAV swarms, focusing on fast response mechanisms to evade threats. Shen and Wei [7] created bird-inspired flocking control to maintain cohesive behavior within the swarm. Amorim et al. [8] explored self-organized UAV flocking based on proximal control, demonstrating effective self-organization without the need for centralized control. Wang et al. [9] and Zhou et al. [10] reviewed recent advances in biomimetic intelligence and swarm intelligence, respectively. Wang et al. [11] and Wang et al. [12] proposed new models to enhance coordination and exploration. Kang et al. [13] addressed the issue of communication delays in formation control. He et al. [14] suggested defensive strategies to protect critical infrastructure from UAV swarm attacks. Iqbal et al. [15] provided a review of the challenges in motion planning

for UAV swarms. Collectively, these studies highlight advancements in biomimetic principles, machine learning techniques, and control strategies that are driving the development of more autonomous, efficient, and robust UAV swarms, integrating natural behaviors and advanced computational methods to overcome challenges in complex and dynamic environments.

3 Methods

3.1 *Bird Flocking Model Using Graph Distance*

Collective behavior in nature is achieved through individuals following simple principles and using local information to complete complex global tasks. Each individual acts based solely on the information from neighboring individuals, without needing to understand the state of the entire swarm. The design of fixed-wing drones is inspired by this behavior of birds. Biomimetic approaches enable drone swarms to mimic bird flocks, coordinating through simple rules and local information.

As shown in Fig. 1, starlings, as a typical flocking bird species, exhibit excellent coordination abilities within their flocks. They can maintain large-scale formations while finding suitable foraging areas and responding quickly to predator attacks. Existing research often focuses on the overall formation configuration, intra-flock synchronization, and search area coverage from a swarm perspective, and verifies the controllability and robustness of the swarm through behavior convergence. This study reexamines swarm control issues from the perspective of swarm energy consumption control, using optimization thinking as a powerful tool to address internal coordination and external obstacle avoidance problems.

Fig. 1 A typical example of bird flocks: starling murmurations



3.1.1 Flocking Phenomenon and Reynolds' Three Rules

In 1987, Reynolds and colleagues proposed a flocking model where each agent (called a boid) follows three simple rules: cohesion, separation, and alignment, to achieve flocking behavior.

When dealing with the control problem of large-scale automated vehicle clusters, individual agents within the cluster exhibit strong locality in information processing, especially evident when each unit has limited computational capabilities. Assume any UAV i in swarm Q can communicate with other UAVs within a limited spatial range Ω_{local} at a given time t during flight to exchange positional information. The state vector $X_i(t) = (P_i(t), V_i(t))$ of UAV i at time t , given its constant flight altitude, can be represented by its position $P_i(t) = (x, y)$ and velocity vector $V_i(t) = (v_x(t), v_y(t))$. For UAV i at time t , within a finite spatial range Ω_{local} centered on itself, there exists a local cluster of UAVs denoted as the local neighborhood, $N_i(t)$ comprising $N_i(t)$ UAVs. The definitions of the three major actions within this local neighborhood are as follows:

(1) Cohesion action $C_i(t)$:

$$C_i(t) = \sum_{j \in N_i(t)/\{i\}} P_j / (n_i(t) - 1 - P_j), \quad (1)$$

where $C_i(t)$ denotes UAV i 's computation of the average position of all UAVs in $N_i(t)$ excluding itself.

(2) Separation action $S_i(t)$:

$$S_i(t) = \sum_{j \in N_i(t)/\{i\}} (P_j - P_i) / |P_j - P_i|^2, \quad (2)$$

where $S_i(t)$ represents UAV i 's calculation of the minimum safe distance from all other UAVs in $N_i(t)$.

(3) Alignment action $A_i(t)$:

$$A_i(t) = \sum_{j \in N_i(t)/\{i\}} V_j / (n_i(t) - 1), \quad (3)$$

where $A_i(t)$ denotes UAV i 's computation of the average velocity of all UAVs in $N_i(t)$ excluding itself.

To accommodate varying action requirements of UAVs within the cluster during different mission phases, weights w_C , w_S , and w_A are pre-set for cohesion, separation, and alignment actions, respectively. These weights are used to update the velocity and position of UAV i within its local neighborhood $N_i(t)$ through these actions. Post-update, the velocity vector v_i of UAV i is given by:

$$V_i^*(t) = V_i(t) + w_C \cdot C_i(t) + w_S \cdot S_i(t) + w_A \cdot A_i(t) \quad (4)$$

The predicted position of UAV i is:

$$P_i^*(t) = P_i(t) + V_i(t) \cdot \Delta T \quad (5)$$

Here, ΔT represents the sampling period.

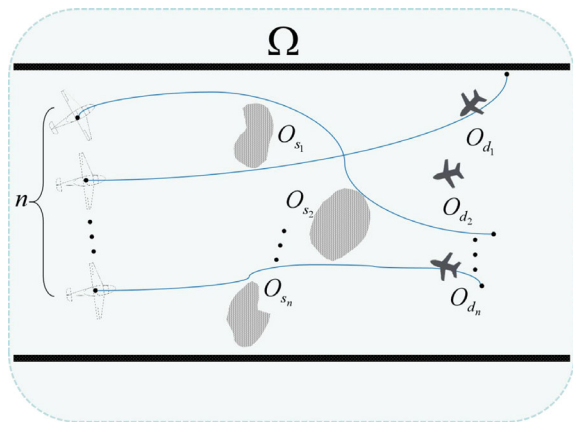
3.1.2 In a Multi-dynamic Obstacle Scenario

Figure 2 is an illustrative diagram of an unknown complex environment. Within a limited spatial range Ω , a swarm of fixed-wing UAVs, numbering n , can be observed performing flight tasks. In this domain, there are static and dynamic obstacles, including static boundaries and hostile flying objects. O_{static} represents a collection of n_{sta} static obstacles, while O_{dynamic} denotes a collection of n_{dym} dynamic obstacles.

Due to collision avoidance issues, this paper extends them into two categories:

- (1) Intra-cluster UAV Collision Avoidance: Within a UAV cluster system, collision avoidance between any UAV and its neighboring UAVs is addressed. Due to intra-cluster communication, each UAV can directly access the motion states (position, velocity, heading, etc.), trajectories, and intentions of others. By designing feasible cooperative control laws, UAVs at risk of collision can proactively and cooperatively avoid collisions.
- (2) External Obstacle Collision Avoidance: In contrast to intra-cluster issues, this category deals with collision avoidance concerning external obstacles to UAV cluster systems. UAVs perceive the motion state information (position, velocity, heading, etc.) of obstacles but lack knowledge of their trajectories and intentions. Therefore, UAVs typically engage in unilateral avoidance behaviors toward these obstacles without considering reciprocal actions.

Fig. 2 Illustrates UAV swarm flight missions



In summary, the core of the first type of collision avoidance is cooperative avoidance with known (or predictable) targets capable of communication. The essence of the second type revolves around independent evasion toward unidentified targets without relying on effective communication. Therefore, the following definitions can be established:

Definition 1 (*Obstacle Avoidance*): Within a finite spatial range, the unilateral avoidance behavior executed by UAVs or UAV cluster systems against any static or dynamic obstacles external to the system, which cannot be effectively communicated with and are expected to collide within a finite future time window, is termed obstacle avoidance.

Definition 2 (*Intra-cluster Collision Avoidance*): The cooperative avoidance behavior undertaken by a UAV within a UAV cluster system toward adjacent UAVs capable of effective direct communication within the system and expected to collide within a finite future time window, is termed intra-cluster collision avoidance.

3.1.3 Intra-cluster Collision Avoidance Kinematic Model for UAV Systems

(1) Analysis of Remote Control for UAV Cluster Systems

In large-scale UAV swarms, as the size of the swarm increases, challenges arise with long-range issues within the cluster, where increased physical distances lead to delays or loss in communication and information synchronization. To address this issue, local neighborhoods $N_i(t)$ serve as an effective strategy. Local neighborhoods are relatively small in scale and typically maintain overall information dissemination and updates within the cluster through information exchange among their members. By controlling various local neighborhoods distributed across the swarm's spatial domain, optimal control can be achieved in large-scale clusters, minimizing losses due to long-range spatial constraints.

Therefore, achieving optimization of control over the remaining UAVs in the local neighborhood $N_i(t)$ by UAV $j \in N_i(t) (j \neq i)$ as the leading UAV requires satisfying the following conditions:

Leader drones possess reference path information to reach their destinations and can detect and avoid obstacles along the path. Follower drones, on the other hand, only have awareness of their current state and local cluster information.

Leader drones employ Model Predictive Control (MPC) to generate near-term planned paths and state information, which are transmitted to surrounding follower drones to influence their trajectories. Leaders do not consider information received from followers when planning their own paths.

Ideally, control actions of leader drones over follower drones should aim to: minimize time to reach the destination, minimize the number of local leaders along the way, and minimize the set of motion constraints (to ensure no followers are left behind).

3.2 Control Algorithm Design

3.2.1 Construction of UAV Swarm Obstacle Avoidance Control Model

Assuming at time t , the leader UAV in the swarm detects an obstacle within its detection range, denoted as $O_i(t)$. The swarm will then proceed with the following actions over the subsequent period:

(1) Obstacle Detection:

The leader UAV detects the obstacle and gathers its feature point information $P_{o_i}(t)$ and distance. All UAVs are aware of their own positions, and the leader UAV immediately shares the obstacle information. The swarm assumes the obstacle moves in a straight line at constant speed within a short time interval $[t, t + \Delta t]$, making its state predictable. Based on the obstacle's characteristics, the UAV swarm plans obstacle avoidance strategies. It manages collision avoidance for dynamic obstacles that present collision risks by recalculating flight paths and modifying formation structures to ensure safety and mission completion.

(2) Formation Adjustment:

For static obstacles, the cluster employs bio-inspired linear queue control, where a single leader UAV plans the path and follower UAVs execute obstacle avoidance. In the case of dynamic obstacles, all UAVs are divided into leaders and followers, with fewer leaders than followers. Several leader UAVs work together to ensure the entire cluster navigates around obstacles safely, prioritizing flight safety.

As shown in Fig. 3, within a finite spatial range centered around the leader UAV, a communication topology is formed between the leader and nearby followers, structured as a graph with directed edges. The cost obtained after weighting these directed edges is referred to as the hierarchical graph distance.

$N_i(t)$ represents the local neighborhood centered on leader i in the topological region at time t . Naturally, the hierarchy level of this leader is maintained as $\pi_i(t) \equiv 0$. Followers $j(j \in N_i(t))$ relative to i are initialized with hierarchy levels $\pi_j(0) = \pi(\pi \geq 1, \pi \in \mathbb{N})$. Over time, this hierarchical distance will be continuously updated.

$$\pi_j(t) = \min \left\{ \bar{\pi}, 1 + \min_{i \neq j} \pi_j(t - 1) \right\} \quad (6)$$

Although this update rule relies solely on local information, it can accurately and promptly update the graph distance of the entire cluster within a limited sampling time through information exchange among leaders.

When the distance between two drones is smaller, the possibility of collision increases, placing them in a more dangerous situation. Correspondingly, the risk of instability in the entire UAV swarm system also increases. The reference output vector of neighboring drone j includes specific positional information $p_i(\tau)$ and velocity information $v_i(\tau)$.

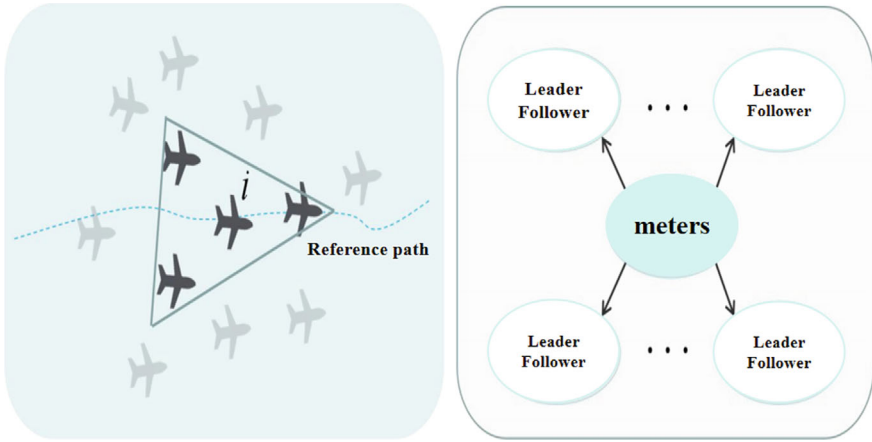


Fig. 3 Local neighborhood centered around leader UAV i ($i \in L$)

$$\begin{aligned}\bar{p}_i(\tau) &= \sum_{j \in N_i(\tau)} w_{ji}(\tau) \cdot p_{ji}(\tau) \\ \bar{v}_i(\tau) &= \sum_{j \in N_i(\tau)} \omega_{ji}(\tau) \cdot p_{ji}(\tau),\end{aligned}\tag{7}$$

where $w_{ji}(\tau)$ and $\omega_{ji}(\tau)$ are respectively the weighting coefficients for cohesion and alignment actions of UAV i at time t . Combining the above equations with the predicted distances between UAVs, the optimization objective function can be derived.

3.2.2 Graph Distance-Based Hierarchical Control Algorithm

A distributed control law is designed for UAV swarm systems, where UAVs are divided into leaders and followers. Leaders are responsible for planning trajectories to destinations and transmitting recent trajectory and velocity prediction information to followers. Followers control their movements based solely on their local environment and the information received from leaders. The expected goals of this control law include minimizing the time required for the swarm to reach its destination, reducing the number of leaders necessary for swarm control, optimizing the addition of constraints during leader control, and ensuring all followers closely follow the swarm to its destination.

3.3 Simulation Verification and Comparative Analysis

3.3.1 Collision Avoidance Performance in Obstacle-Free Scenarios

Within the cluster system, UAVs are initially categorized into leader and follower groups before obstacle avoidance maneuvers. Leader UAVs exchange information and adjust trajectories to maintain distance, ensuring collision avoidance control. This study employs kinematic modeling of fixed-wing UAVs and sets minimum safe distances to prevent collisions. Using four leader UAVs as an example, flight trajectory simulations are conducted on the MATLAB platform, neglecting communication delays. Initial speeds are set at 19 m/s with speed limits ranging from 16 to 21 m/s and angular velocity limits from -2 to 2 rad/s, where due east is considered the 0 direction. UAV starting points differ from the reference path start point to test interaction performance, with the four UAVs set to fly toward each other.

Figure 4 illustrates a schematic of the UAV cluster composed of four UAVs, namely UAV-1, UAV-2, UAV-3, and UAV-4, each following a smooth closed-curve reference path. Circles denote the starting points of the reference path, and arrows indicate the direction of travel for each UAV. Panel (a) depicts their respective reference paths, while panel (b) shows their actual flight trajectories.

Under obstacle-free flight conditions, the flight performance of each leader UAV within the cluster can be evaluated based on tracking errors from the reference path. During experiments, UAVs track instantaneous reference points on the reference path. Given that UAV starting points do not coincide with the reference path start point, UAVs are allowed a certain deviation from their respective instantaneous reference points during flight.

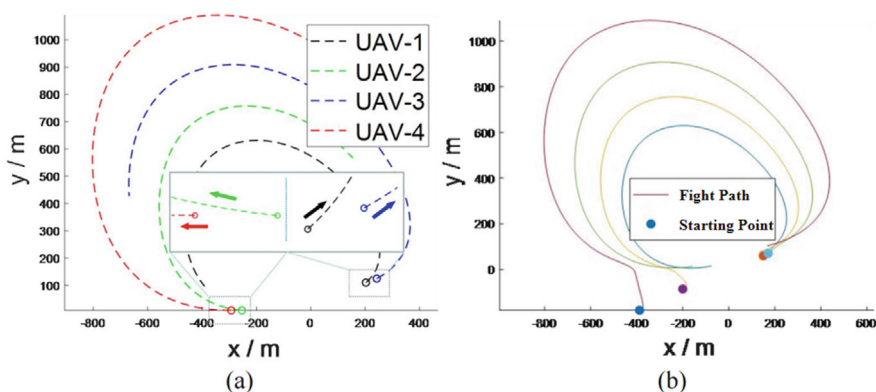


Fig. 4 Schematic of a UAV swarm composed of four drones

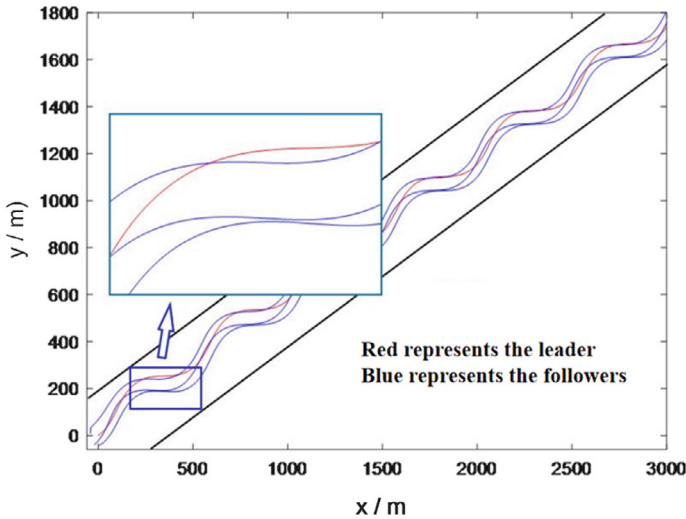


Fig. 5 Flight trajectories of a local neighborhood with four drones

3.3.2 Intra-Cluster Collision Avoidance Performance Testing

In this setup, the study defines the local neighborhood size as consisting of four UAVs, including three follower UAVs. Each UAV takes off within a rectangular area measuring 50×50 m, following a narrow flight path. Throughout the flight, the cluster must navigate external environmental obstacles while maintaining safe distances and a proper flight queue order among UAVs. To verify the convergence of the cluster control process, each UAV starts with different initial speeds and directions, consistent with the previously discussed speed range settings. This configuration aims to test the effectiveness of the cluster control algorithm in dynamic environments, emphasizing intra-cluster cooperation and adaptation to external environmental changes.

Figure 5 illustrates the flight trajectories of a local neighborhood composed of four drones. As depicted, the trajectories of leader drones are marked in red, while follower drones are denoted in blue. To better showcase the flight dynamics, specific details of the flight trajectories have been enlarged for clarity.

3.3.3 Performance Testing of Cluster Collision Avoidance in Dynamic Obstacle Scenarios

In the following tests, the collision avoidance performance of a cluster comprising four leader drones will be extensively evaluated under dynamic obstacle scenarios. Each drone will follow its designated path while dynamic obstacles, moving at a speed of 12 m/s in the same direction as the drones, are introduced. Upon detecting these

obstacles, the drones will execute corresponding maneuvers to evade collisions. The tests aim to assess the cluster's responsiveness and collision avoidance effectiveness in such dynamic obstacle scenarios.

4 Results

4.1 Bird Flocking Model Using Graph Distance

4.1.1 Collision Avoidance Kinematic Model for UAV Swarm

(1) Analysis of Remote Control in UAV Swarm

Considering the overall conditions, the number of fixed-wing UAVs in the swarm is set to N . For each UAV i , the control input sequence within a finite time domain TTT is represented as $u_i(t) = u_i(t) \cdots u_i(t + T_i - 1)t$, where $x_i(t)$ denotes the state sequence of the UAV over its entire prediction horizon T_i .

At time t , UAV i perceives information about $O_i(t)$ obstacles nearby and obtains predicted output sequences $y_{j|i}(t)$ from neighboring UAVs $j \in N_i(t)/\{i\}$ (excluding UAV i itself). These outputs include positional information $p_{j|i}(t)$ and velocity information $v_{j|i}(t)$ of neighboring UAVs. The subscript index $j|i$ indicates that the output sequence collected by UAV j is stored as a copy by UAV i , highlighting the potential for mismatched information among different UAVs collecting the same data.

(2) Kinematic Modeling of UAV Swarm Systems

Assume at time t , $f(x_i(t), u_i(t))$ represents the discrete dynamics model of fixed-wing UAV i , where $x_i(t) \in X_i(t)$ and $u_i(t) \in U_i(t)$ are the state vector and control input vector of UAV i , respectively. Here, $X_i(t)$ and $U_i(t)$ denote the corresponding sets of dynamic state constraints and control input constraints. The control process of UAVs and UAV swarms is transformed into an optimization problem for solution. However, before obtaining and solving this objective function, it is first necessary to clarify how to effectively distinguish the far ends of the UAV swarm. Thus, at each time t , the optimization control model for UAV i is formulated as:

$$\begin{aligned} & \min_{u_i(t) \in U_i \times \cdots \times U_i} J(u_i(t)) \\ s.t. & \begin{cases} x_i(t + k + l) = f(x_i(t + k), u_i(t + k)) & k < T_i \\ x_i(t + k + l) \in X_i & k < T_i \\ p_i(t + k + l) \notin O, O \in O_i(t) & k < T_i \\ d_{j|i}(t + k + l) \geq d_{i,sep}, j \in N_i(t + k + l) & k < T_{i,sep} \end{cases} \end{aligned} \quad (8)$$

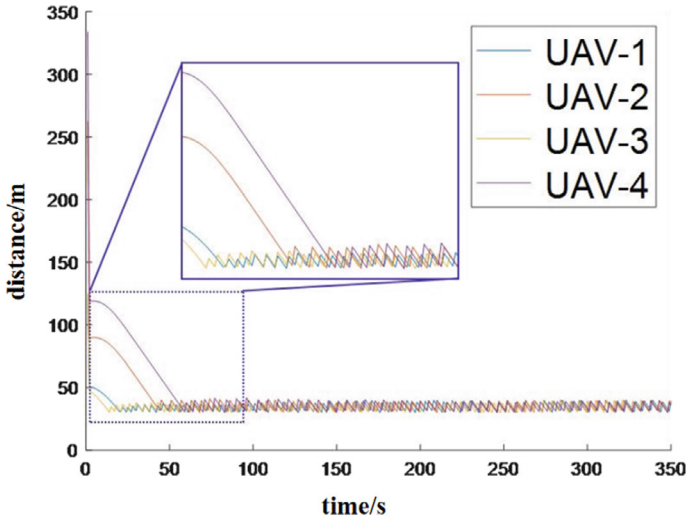


Fig. 6 Distance variation among drones within the cluster during flight

4.2 Simulation Verification and Comparative Analysis

4.2.1 Performance Testing of Collision Avoidance Control in Obstacle-Free Swarm Scenarios

Under obstacle-free flight conditions, the flight performance of leader UAVs within the swarm can be evaluated by tracking errors relative to a reference path. In experiments, UAVs track instantaneous reference points along the reference path. Due to the non-coincidence of the starting point of the UAV flight path with that of the reference path, UAVs are allowed to maintain a certain distance from their respective instantaneous reference points during flight.

Figure 6 displays the distance changes of 4 drones relative to the reference path during flight. Due to the non-coinciding starting points of the flight and the reference path, the initial distances are large but quickly converge to a suitable range, demonstrating the rapid startup capability of drone clusters and their efficient task execution.

4.2.2 Cluster Collision Avoidance Performance Testing

As shown in Table 1, during actual flight operations, the leader drone consistently maintains a distance differential with the follower drones, and the separation between each pair of followers always exceeds the minimum safety distance (30 m).

Table 1 Minimum distances between drones at the same moment

Unit:m	Meters	Leader follower 1	Leader follower 2	Leader follower 3
Meters	–	61.342	64.241	62.254
Leader follower 1	61.342	–	31.125	45.312
Leader follower 2	64.241	31.125	–	32.214
Leader follower 3	62.254	45.312	32.214	–

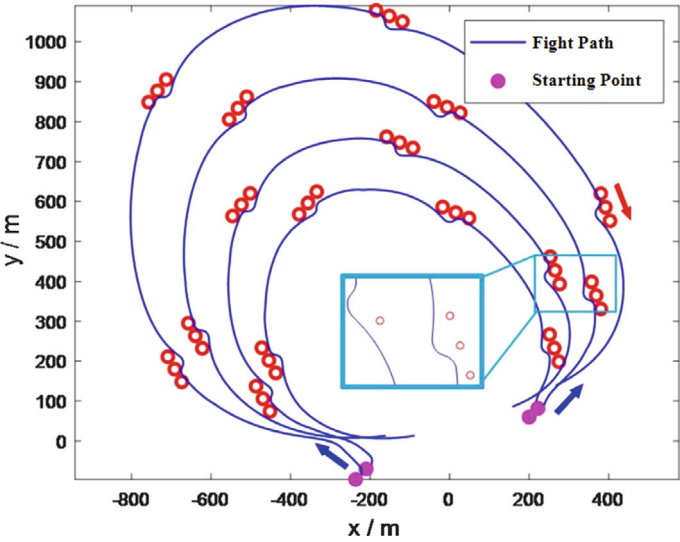


Fig. 7 Dynamic obstacle avoidance testing for drone clusters

4.2.3 Performance Testing of Cluster Collision Avoidance Control in Dynamic Obstacle Scenarios

As shown in Fig. 7, each drone maintains distance from each other while avoiding obstacles, ensuring the minimum safety distance throughout the entire flight process.

5 Summary

This paper first introduces an innovative hierarchical control framework using the Reynolds model with multiple leaders. It begins by employing hierarchical graph distances to divide each leader and its nearby followers into localized clusters, thereby structuring the UAV swarm into distinct neighborhood groups. Secondly, hierarchical graph distances are used to segment each leader and its nearby followers into localized clusters, effectively breaking down the UAV swarm into distinct neighborhood

groups. The study then designs the interactions among multiple leaders in obstacle-free environments and validates the swarm's collision avoidance capabilities when faced with dynamic obstacles. Furthermore, it systematically analyzes and examines the control interactions between leaders and followers within these local clusters. This research significantly contributes to achieving collision avoidance for UAV swarms in the presence of dynamic obstacles and effectively addresses the formation control challenges during avoidance maneuvers.

References

1. Wu W, Zhang X, Miao Y (2022) Starling-behavior-inspired flocking control of fixed-wing unmanned aerial vehicle swarm in complex environments with dynamic obstacles. *Biomimetics* 7(4):214
2. Zhou W, Liu Z, Li J et al (2021) Multi-target tracking for unmanned aerial vehicle swarms using deep reinforcement learning. *Neurocomputing* 466:285–297
3. Li X, Wu J, Zhao Z et al (2023) An improved method for broiler weight estimation integrating multifeature with gradient boosting decision tree. *Animals* 13(23)
4. Luo L, Wang X, Ma J et al (2021) GrpAvoid: multigroup collision-avoidance control and optimization for UAV swarm. *IEEE Trans Cybern* 53(3):1776–1789
5. Mir I, Eisa SA, Taha H et al (2021) A stability perspective of bioinspired unmanned aerial vehicles performing optimal dynamic soaring. *Bioinspir Biomim* 16(6):066010
6. Novák F, Walter V, Petráček P et al (2021) Fast collective evasion in self-localized swarms of unmanned aerial vehicles. *Bioinspir Biomim* 16(6):066025
7. Shen Y, Wei C (2022) Multi-UAV flocking control with individual properties inspired by bird behavior. *Aerosp Sci Technol* 130:107882
8. Amorim T, Nascimento T, Petracek P et al (2021) Self-organized UAV flocking based on proximal control. In: 2021 International conference on unmanned aircraft systems (ICUAS). IEEE, pp 1374–1382
9. Wang J, Chen W, Xiao X et al (2021) A survey of the development of biomimetic intelligence and robotics. *Biomimetic Intell Rob* 1:100001
10. Zhou Y, Rao B, Wang W (2020) UAV swarm intelligence: recent advances and future trends. *IEEE Access* 8:183856–183878
11. Wang C, Shi Z, Gu M et al (2022) Revolutionary entrapment model of uniformly distributed swarm robots in morphogenetic formation. *Defence Technol* 31
12. Wang C, Wang D, Gu M et al (2022) Bioinspired environment exploration algorithm in swarm based on levy flight and improved artificial potential field. *Drones* 6(5):122
13. Kang Y, Luo D, Xin B et al (2022) Robust leaderless time-varying formation control for nonlinear unmanned aerial vehicle swarm system with communication delays. *IEEE Trans Cybern* 53(9):5692–5705
14. He D, Yang G, Li H et al (2020) An effective countermeasure against UAV swarm attack. *IEEE Netw* 35(1):380–385
15. Iqbal MM, Ali ZA, Khan R et al (2022) Motion planning of UAV swarm: recent challenges and approaches. In: *Aeronautics-new advances*

5G-SHE: NIST-ASCON-Based Security Framework for 5G-Enabled Smart Healthcare Ecosystem



Vaishali Meena and Gaurav Indra

Abstract In the changing healthcare landscape, the integration of 5G-enabled Smart Prosthetics has altered patient care by providing personalized experiences via sensors and real-time data. However, information security remains critical. This research proposes a comprehensive cybersecurity system designed specifically for healthcare informatics. It uses NIST-recommended Lightweight cryptography, such as Ascon, to provide data security in resource-constrained devices. Zero Trust Architecture checks device IDs and gives access according to stringent restrictions. The framework extends from manufacturing to patient engagement, enabling secure local and distant communication with features such as User Credential Updates and key management. Formal and informal security verification show resilience to attacks such as replay, man-in-the-middle, and impersonation. While communication costs are 11% higher, computing expenses are reduced by 12%, increasing energy efficiency. Implementing this framework seeks to provide safe and efficient settings for exploiting 5G-enabled Smart Prosthetics in healthcare informatics.

Keywords 5G · NIST-LWC · ASCON · Smart healthcare · Smart prosthetics

1 Introduction

The swift progress of smart prosthetics and 5G technology in the healthcare sector presents notable security obstacles, including unauthorized entry and cyber assaults. The 5G-SHE architecture aims to resolve these concerns by incorporating resilient authentication systems, thereby restricting access to smart prostheses exclusively to authorized users. The framework employs advanced cryptographic approaches

V. Meena · G. Indra (✉)

Department of Information Technology, Indira Gandhi Delhi Technical University for Women,
Delhi, India

e-mail: gauravindra@igdtuw.ac.in

V. Meena

e-mail: vaishali019phd22@igdtuw.ac.in

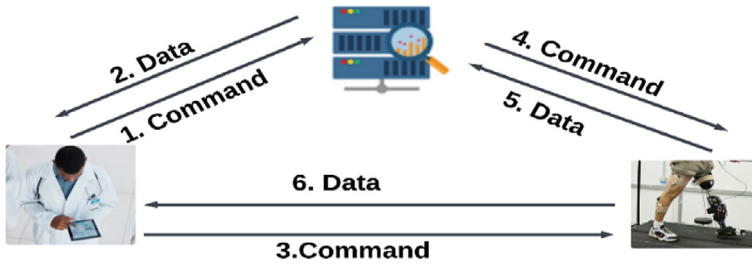


Fig. 1 Device authentication architecture

to improve the security and dependability of communication among prostheses, authentication servers, and user apps.

The objective of this research is to establish a connection between the advanced functions of smart prosthetics and the essential security measures, by creating a framework that is secure, adaptable, and effective. We aim to promote the use of intelligent prosthetics while prioritizing the protection and confidentiality of medical information. This will contribute to the development of secure healthcare technology and enhance confidence in medical Internet of Things (IoT) systems. Our contribution to this paper:

1. 5G-SHE uses session IDs, user session IDs, and network IDs to tackle real-world security issues with these aspects.
2. The 5G-SHE uses the NIST-recommended ASCON (version 1.2), a compute-lightweight encryption method. This provides high security without high computational cost for resource-constrained Smart prosthetic devices.
3. The use of ASCON goes beyond simple security. It protects against side-channel, differential, linear, block collision, and near-collision attacks. This comprehensive defense strategy improves Smart prostheses security (Fig. 1).

The structure of the paper is as follows: Sect. 2 contains a Literature survey of the latest papers. Section 3 contains the system model followed by the proposed framework in Sect. 4. In Sect. 5, Informal security evaluation is explained followed by Formal security Evaluation in Sect. 6. Section 7 explains all the performance evaluations and results. In the end, the Conclusion is given in Sect. 8 (Table 1).

2 Literature Review

Abugabah et al. [1] recommended using Ethereum Smart contracts to create a transparent, tamperproof telemedicine healthcare platform that protects patient data without a central administrator. Telemedicine deployments risk data breaches, restricted access, erroneous diagnoses and prescriptions, fraud, and misuse. Sahoo et al. [2] provide a safe and efficient three-factor authentication technique for IoT-enabled telemedical equipment without vulnerability investigation, or scalability

Table 1 Notations

Notations	Description
PK^{AS}	Public key of authentication server
$SN^{AS}, SN^{SP}, SN^{AP}$	Session identity of authentication server, smart prosthetics, and application, respectively
SN^{SPAS}	Session identity of a session between SP and AS
SN^{ASAP}	Session identity of a session between the authentication server and the application
$En(K,N,A,P)$	Encryption in ASCON of P using nonce as N, secret key as K, associated data as A
$HMAC(K,*)$	HMAC using ASCON Hash
$NWID^{SPAP}$	Network identity between smart prosthetic and application
SK^{SPAP}	Smart prosthetic and application session key given by authentication server
$NWID^{RSAS}$	Network identity between relay server and authentication server

discussion. Jains et al. [3] use technology for healthcare robots to exchange cloud-based healthcare data securely does not maintain encryption keys. Wang et al. [4] employed symmetric encryption to control power and access and prevent data leakage in a medical data-sharing system. Security data and benchmarks are needed to prove its quicker performance than competitors. Yang et al. [5] proposed a secure and lightweight data management model for healthcare systems that need to perform symmetric encryption using XOR and MAC. However, it needs more detail, and comparison with other approaches. Pre-existing literature presents prospects for enhancement. The effects of scalability analysis [6] and efficient routing latency [7, 8] remain largely unexplored. Comparative assessments and security analyses [9–11] may offer a more comprehensive perspective. Attention needs to be paid to real-world factors such as key management. This work can contribute to advancing secure and efficient healthcare informatics by addressing these gaps.

3 System Model

This framework includes Smart Prosthetics (SP), Relay Server (RS), User Application (AP), and Authentication Server (AS). Doctors may remotely monitor and change Smart prostheses using the Medical Professional's Device. Each SP may send real-time arm data to the RS for processing and storage. This framework introduces a complete AKM protocol.

4 Proposed 5G-SHE Framework

4.1 Register Devices

Design and manufacturing start a Smart prosthetic's life. Backend server registration is necessary for all production prostheses. Patients receive prostheses from the factory after registration. This procedure comprises three phases:

4.1.1 Pre-registration

Here, Smart Prosthetic developers register their devices on authentication servers. Each prosthetic model is named differently to distinguish it. Developers register Smart Prosthetic by sending the modal name (PN) and prosthetic model key (PK) to the authentication server.

4.1.2 Create Device and Session Keys

The PS activates and sets the prostheses to (PN, DN, DK). The prostheses use its 32-byte public key to request the device key from the authentication server using Elliptic Curve Cryptography.

4.2 User Binding

Users are linked to their smart prosthetics by backend systems. The authentication/relay server only allows bound users to connect to devices offsite. The mobile app (AP) communicates with the prosthetic and agrees on a brief session key (K^{bind}) so that they can talk safely. The AP gets the SP's USN^{SP} and ASCON, along with a date, a nonce, and the network ID. The prosthesis can decode both USN^{AP} and T2. The prosthesis-AS bond is checked by SK^{SPAS} . After verification, ASCON gets a message from the prosthesis that is encrypted and says it wants to connect. With a username and password, users sign up through HTTPS AP-AS. The application then sends encrypted registration data to the AS, which links approved users' credentials to the prostheses.

4.3 Telecommunication

Here, the aim is to connect prostheses to backend servers. The authentication server (AS) authenticates users, while the relay server (RS) relays Prosthetic device requests

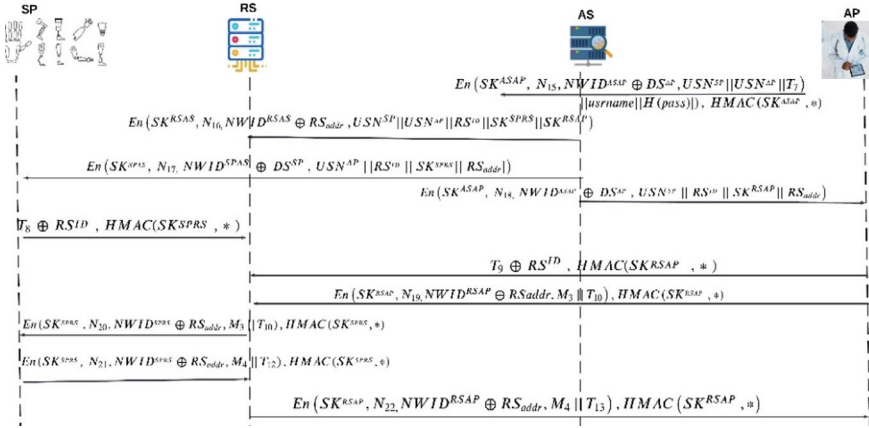


Fig. 2 Telecommunication phase

and doctor data. The mobile Application HTTPS initially connects to the authentication server. Mobile application binding employs cookies or OpenID. AS notifies RS, SP, and AP of Telecommunication readiness after verification. The Relay Server communicates with prosthetic and mobile applications. Figure 2 displays all steps. If the AS and AP connect securely (e.g., over HTTPS) using a session key (SK^{ASAP}), the AS and Prosthetic device mutually authenticate using Device Secret (DS^{SA}) and SK^{SPAS} , and the AS and RS have embedded SK_{RSAS} onto the back. After receiving a message, AS checks its database for SP and AP binding information to verify user credentials. If the credentials match then AS alerts the. AS creates random RS^{ID} s and SK_{SPRS}/SK_{RSAP} keys. SK_{RSAP} linked mobile applications, whereas SK_{SPRS} linked prostheses to RS AS, which provides prosthetics, mobile applications, RS_{addr} , IP, and port. The Relay Server connection, the prosthetic device, and the AP communicate ($T_8, RS^{ID}, HMAC(SK^{SPRS}, *)$) and ($T_9, RS^{ID}, HMAC(SK^{RSAP}, *)$). After message authentication, RS communicates with Prosthetic and AP.

4.4 Refresh and Update Status

4.4.1 Factory Reset Recovery

The SP factory reset deletes all user data. The SP sends a confirmation message ($En(SK^{SPAS}, N_{27}, NWID^{SPAS} \oplus DS^{SP}, USN^{SP} || R_8 || T_{16})$) to the AS. Then AS deletes all prostheses binding data. $En(SK^{SPAS}, N_{28}, NWID^{SPAS} \oplus DS^{SP}, R_8)$ is transmitted by AS to verify with SP and bound users.

4.4.2 Key Refresh

DS_{SP} and K^{local} are 5G-SHE long-term keys; most are session. Regularly upgrading long-term keys boosts security. Before shared keys, Fig. 2 displays K^{local} and DS^{SP} updates.

5 Informal Security Analysis

With many common attacks, 5G-SHE can counter other attacks which can only be countered by using ASCON. The other attacks that are counted are as follows:

5.1 Block Collision Attack

Strong ASCON structure guarantees encrypted data block uniqueness and integrity. In our Smart prostheses system, ASCON's encryption prevents attackers from producing a collision—two inputs yielding the same result. Cryptography safeguards data and dependability from block collisions.

5.2 Near Collision Attacks

ASCON-authenticated 5G-SHE encryption prevents collisions. Near-collision attacks generate two identical ciphertexts with little modification. Authenticated encryption like ASCON prevents attacks.

5.3 Differential Cryptanalysis Attack

Differential cryptanalysis is avoided using ASCON symmetric key encryption. ASCON employs secret key expansion, substitution-permutation networks, and nonlinear processes.

5.4 Linear Cryptanalysis Attack

Safe and efficient ASCON authentication prevents linear cryptanalysis in 5G-SHE. ASCON resists linear and cryptographic attacks, in which the innovative

permutation-based structure and dynamic S-boxes cause nonlinearity and confusion, prohibiting linear approximations.

5.5 Side Channel Attack

ASCON saves power and electromagnetic radiation by encrypting and decrypting data. Use side channel mitigation and constant-time cryptography. Our encryption scheme is difficult for side-channel attackers.

6 Formal Security Analysis

Security Evaluation Using Scyther: Scyther was used to evaluate the security of the Smart Prosthetic system. The verification results of the Scyther tool indicate security protocols for IoSP (Internet of Smart Prosthetic) and AP (Application). Each item shows that a different security protocol has been successfully checked. These include non-injective synchronization (Nisynch), non-injective agreement (Niagree), and the privacy of keys and nonces (Secret). It has also been proven that people are committed to certain values (Commit). All of the claims are marked “OK,” which means they have been checked out and no attacks have been found. Hence, it provides a strong security protocol that checks all of its properties.

7 Performance Evaluation

Smart Prosthetic authentication utilized Python on a 4.80 GHz laptop. HMAC message verification, ASCON. MySQL saved authentication server data. We serialized communication data in Python dictionaries. We tracked lifetime storage, communication, and compute costs.

7.1 Storage Overhead

In 5G-SHE, AS holds SP^{ID} , DS^{AP} , DS^{SP} , MA^{ID} , username, salt, and spd for each bound device, totaling 160 bytes. RS contains 144 bytes of USN^{SP} , USN^{AP} , SK^{SPRS} , SK^{RSAP} , and $RSid$ for each relay connection.

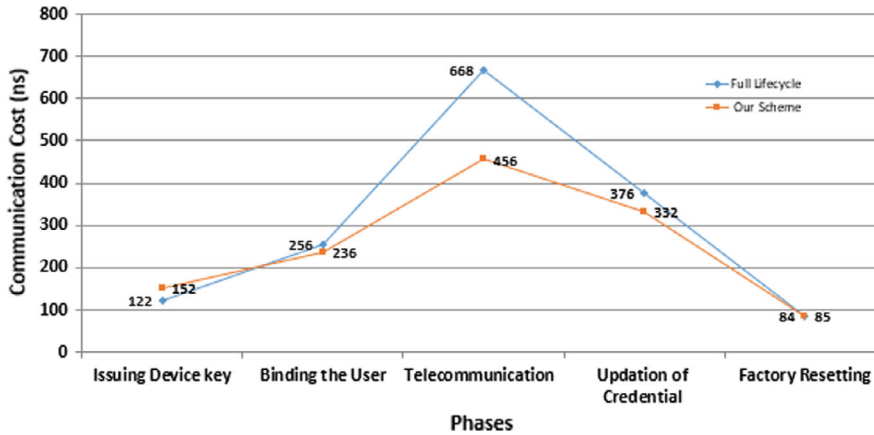


Fig. 3 Communication cost of lifecycle [17] versus 5G-SHE

7.2 Communication Overhead

This framework took more communication cost, i.e., 672 Bytes, which is higher than all the other schemes, but our scheme promises more security than all the other state-of-the-art. Comparing the Communication Cost of related papers gives results as follows: Paper [12] gives 334 Bytes, paper [13] gives 368 Bytes, paper [14] gives 304 Bytes, paper [15] gives 224 Bytes, paper [16] gives 480 Bytes, paper [17] gives 492 Bytes given in Fig. 3.

7.3 Computation Cost

T_s is symmetric encryption and decryption, T_h is hashing, and T_{hmac} is MAC creation. This framework took less computation cost, i.e., $7 T_s + 7 T_{hmac}$, and gives more security compared to state-of-the-art $8T_s + 16T_h + 6T_{hmac}$ [17]. The graphical form of the calculated computation cost is given in Fig. 4.

7.4 Overall Computation and Communication Cost

These are all the costs to communicate and do secure calculations at different stages of a smart IoT system. Device and Session Key, User Binding, Telecommunication, Credential Update, and Factory Resetting are some of the steps. During the Device & Session Key step, it costs 333 bytes to communicate and 274,293 ns to do digital math. It costs more to use User Binding because it takes 437 bytes to communicate and 701,534 ns to do secure calculations. It takes 672 bytes to communicate and

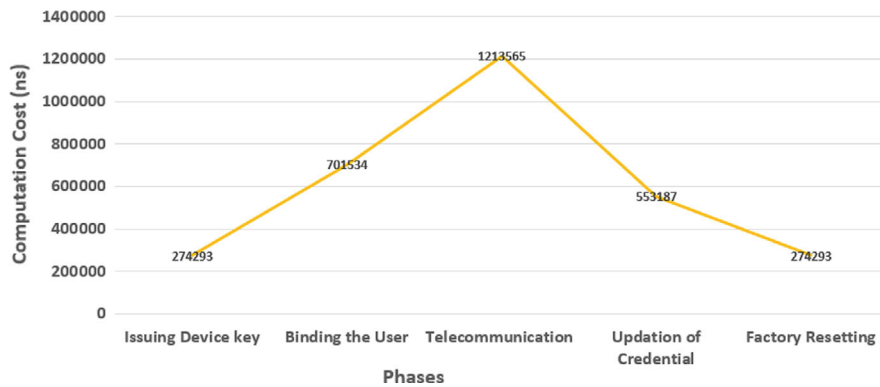


Fig. 4 Computation cost of each phase

1,213,565 ns to do cryptographic calculations in telecommunication, which sends the most data. It costs 480 bytes and 553,187 ns to go through the Credential Update step. Finally, Factory Resetting takes the least amount of time to communicate (201 bytes) and do cryptographic calculations (274,293 ns). Overall, the table shows how the different parts of the IoT system need different amounts of resources. It focuses on finding the right mix between the size of the communication data and the time needed for cryptographic operations.

7.5 Comparison with Other State-of-the-Art

Most IoT authentication solutions use communication-stage authentication, making quantitative comparisons useless. Communication authentication-focused qualitative comparisons with relevant articles [12–16] were chosen. We used 10, 4, and 32 bytes for the Smart Prosthetic device ID, timestamp, and hash digest. For encryption, decryption, random numbers, passwords, and usernames, configure 16-byte plaintext and ciphertext blocks. 5G-SHE reduces processing costs. Smart Prosthetic security and functionality are compared to Smart IoT in Table 2.

Table 2 Security and functionality comparison

Functionalities	[12]	[14]	[16]	[15]	[13]	[17]	THIS
Differential cryptanalysis attacks	No	No	No	No	No	No	Yes
Smartphone stolen attacks	Yes	Yes	NA	Yes	Yes	Yes	Yes
Direct and mutual authentication between the device and the user	No	No	NA	No	NA	Yes	Yes
Block collision attack	No	No	No	No	No	No	Yes
Many-to-Many binding relationship between a device and a user	No	No	NA	No	Yes	Yes	Yes
Near-collision attacks	No	No	No	No	No	No	Yes
Linear cryptanalysis attack	No	No	No	No	No	No	Yes
Side channel attack	No	No	No	No	No	No	Yes

8 Conclusion

This research significantly contributes to the field of healthcare informatics by advancing security protocols for Smart Prosthetics enabled by 5G technology. The resource-constrained Smart Prosthetic devices are guaranteed a robust security framework by employing ASCON (version 1.2), a lightweight cryptography recommended by the NIST. In addition to assuring data security, ASCON provides robust defenses against many threats, including side-channel, differential, linear, block collision, and near-collision attacks. By reducing computational costs and eliminating the requirement for extensive security measures, the proposed framework enables healthcare informatics to implement 5G-enabled Smart Prosthetics securely and efficiently. This research establishes a secure and efficient environment for the integration of 5G-enabled Smart Prosthetics, thus contributing significantly to the advancement of patient care and the realization of the full technological potential in healthcare provision. Although the 5G-SHE architecture greatly improves the security of smart prostheses, there are still some constraints that need to be addressed. The framework’s dependence on strong cryptographic protocols could result in a higher computing burden, which may affect the performance of devices in contexts with limited resources. In large-scale installations, the requirement for constant changes to long-term keys may present complications. Nevertheless, these issues are easily controllable and do not diminish the framework’s overall efficacy in safeguarding smart healthcare systems. Subsequent research can focus on these factors to enhance and simplify the solution.

References

1. Abugabah A, Nizamuddin N, AlZubi AA (2020) Decentralized telemedicine framework for a smart healthcare ecosystem. *IEEE Access* 8:166575–166588. <https://doi.org/10.1109/ACCESS.2020.3021823>
2. Sahoo SS, Mohanty S, Majhi B (2021) A secure three-factor based authentication scheme for health care systems using IoT-enabled devices. *J Ambient Intell Humaniz Comput* 12:1419–1434. <https://doi.org/10.1007/S12652-020-02213-6>
3. Jain S, Doriya R (2022) Security framework to healthcare robots for secure sharing of healthcare data from the cloud. *Int J Inf Technol* 14:2429–2439. <https://doi.org/10.1007/s41870-022-00997-8>
4. Wang M, Guo Y, Zhang C, Wang C, Huang H, Jia X (2023) MedShare: a privacy-preserving medical data sharing system by using blockchain. *IEEE Trans Serv Comput* 16(1):438–451. <https://doi.org/10.1109/TSC.2021.3114719>
5. Yang Y, Zheng X, Tang C (2017) Lightweight distributed secure data management system for health internet of things. *J Netw Comput Appl* 89:26–37
6. Kumari A, Kumar V, Abbasi MY, Kumari S, Chaudhary P, Chen C-M (2020) CSEF: cloud-based secure and efficient framework for smart medical system using ECC. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3001152>
7. Khan MA, Quasim MT, Alghamdi NS, Khan MY (2020) A secure framework for authentication and encryption using improved ECC for IoT-based medical sensor data. *IEEE Access* 8:52018–52027. <https://doi.org/10.1109/ACCESS.2020.2980739>
8. Natarajan R, Lokesh GH, Flammini F, Premkumar A, Venkatesan VK, Gupta SK (2023) A novel framework on security and energy enhancement based on internet of medical things for healthcare 5.0. *Infrastructures* 8(2):22. <https://doi.org/10.3390/infrastructures8020022>
9. Ryu J, Kang D, Lee H, Kim H, Won D (2020) A secure and lightweight three-factor-based authentication scheme for smart healthcare systems. *Sensors* 20(24):7136. <https://doi.org/10.3390/s20247136>
10. Deebak BD, Al-Turjman F, Aloqaily M, Alfandi O (2019) An authentic-based privacy preservation protocol for smart e-healthcare systems in IoT. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2019.2941575>
11. Rana A et al (2022) Internet of medical things-based secure and energy-efficient framework for health care. *Big Data* 10(1):18–33
12. Wazid M, Das AK, Odelu V, Kumar N, Susilo W (2017) Secure remote user authenticated key establishment protocol for smart home environment. *IEEE Trans Dependable Secure Comput*. <https://doi.org/10.1109/TDSC.2017.2764083>
13. Sharma G, Kalra S (2018) A lightweight multi-factor secure smart card-based remote user authentication scheme for cloud-IoT applications. *J Inf Secur Appl* 42:95–106. <https://doi.org/10.1016/j.jisa.2018.08.003>
14. Sahoo S, Sahoo SS, Maiti P, Sahoo B, Turuk AK (2019) A lightweight authentication scheme for cloud-centric IoT applications, pp 1024–1029. <https://doi.org/10.1109/SPIN.2019.8711757>
15. Dammak M, Boudia ORM, Messous MA, Senouci SM, Gransart C (2018) Token-based lightweight authentication to secure IoT networks. <https://doi.org/10.1109/CCNC.2019.8651825>
16. Panda PK, Chattopadhyay S (2020) A secure mutual authentication protocol for IoT environment. *J Reliable Intell Environ* 6:79–94. <https://doi.org/10.1007/s40860-020-00098-y>
17. Chen F, Xiao Z, Xiang T, Fan J, Truong H-L (2023) A full lifecycle authentication scheme for large-scale smart IoT applications. *IEEE Trans Dependable Secure Comput* 20(3):2221–2237. <https://doi.org/10.1109/TDSC.2022.3178115>
18. Dobraunig C, Eichlseder M, Mendel F et al (2021) ASCON v1.2: lightweight authenticated encryption and hashing. *J Cryptol* 34:33. <https://doi.org/10.1007/s00145-021-09398-9>

Design and Application of Wearable Motion Training and Monitoring System Based on Convolutional Neural Network



Mei Hu and Chang Chen

Abstract Nowadays, scientific and efficient sports training is receiving increasing attention, especially personalized training monitoring. In this context, the development of wearable motion monitoring systems using advanced Convolutional Neural Network (CNN) technology can provide real-time and accurate training feedback for athletes, thereby optimizing training effectiveness. This study designed and implemented a wearable motion training monitoring system based on CNN, using multi-sensor fusion technology to collect physiological and motion data of users (such as heart rate, velocity, acceleration, etc.), and input these data into an optimized CNN for real-time analysis. In terms of CNN model design, this paper adopts a multi-layer feature extraction network, utilizing deep learning technology to accurately identify motion patterns and evaluate motion effects. The system also includes data preprocessing and feature extraction modules to improve data accuracy and model generalization ability. Regarding heart rate monitoring errors, the range of error values is between -2 and 2 . The application of this system not only improves the scientific and personalized level of training, but also provides intuitive data support for coaches and athletes, helping to develop more reasonable training plans and health management strategies.

Keywords Convolutional neural network · Wearable motion training monitoring system · Heart rate monitoring · CNN model design

1 Introduction

With the arrival of the digital and intelligent era, the field of exercise science and health monitoring has also ushered in unprecedented development opportunities. Especially with the promotion of wearable device technology, personalized exercise

M. Hu · C. Chen (✉)

Department of Physical Education, Guangzhou Xinhua University, Guangzhou, Guangdong, China

e-mail: ChangChen0404_1987@163.com

training and health management have gradually become possible. However, although wearable devices have been widely used in daily health monitoring, the effectiveness and accuracy of their application in exercise training still face challenges. This study aims to improve the existing wearable sports training monitoring system through CNN technology in order to improve the accuracy of data and the efficiency of real-time feedback. Existing studies have made some progress but still lack an efficient and real-time system to process and analyze large-scale exercise data. Therefore, this study develops a novel wearable sports training monitoring system using CNNs, which not only provides more accurate analysis of sports data, but also provides real-time feedback to the user so as to optimize the training effect and improve sports performance.

The main research of this paper includes the design and implementation of a CNN-based wearable sports monitoring system, as well as the validation of the system's effectiveness in real sports training applications. First, we develop a data collection and processing framework to capture motion data using a variety of sensors; then, we employ deep learning algorithms, especially CNN, to process and analyze these data to ensure that different motion patterns and activity levels can be accurately recognized; finally, we demonstrate the advantages of this system in improving data accuracy and real-time performance through comparison experiments with traditional monitoring systems.

The structure of this paper is organized as follows: first, this paper presents the research background and literature review in detail, highlighting the innovative points of this study based on the existing technology; then, this paper describes the design of the system and the experimental setup, including the hardware selection, the data collection methodology, and the specific details of the implementation of the CNN model; finally, this paper verifies the performance of the system through the experimental results, and discusses the potential value of its practical application and the future Improvement direction. By combining deep learning technology with wearable device technology, this paper demonstrates the application potential and practical effects in the field of athletic training monitoring, and provides a new perspective and empirical basis for future research.

2 Related Work

In recent years, with the rapid development of artificial intelligence technology, especially the wide application of deep learning in various fields, motion monitoring research has also begun to integrate more complex data processing and analysis techniques. Several studies have shown that the analysis accuracy and processing speed of motion data can be significantly improved by using deep learning, especially CNN. Yang et al. studied the knowledge graph analysis of the developmental status of wearable sports monitoring technology [1]. Liu et al. explored the effect of motor imagery training combined with task-oriented functional training on the functional recovery of lower limbs in hemiplegic patients with stroke [2]. Wang et al. studied the

quantitative analysis of breathing pattern parameters based on wearable system [3]. Zhang et al. studied the design of low-power cow exercise monitoring ear tag based on dynamic duty cycle sampling [4]. Su et al. studied the optimization mechanism of exercise performance, mechanism of action and application of wearable resistance training [5]. However, most of these studies focus on static or simpler movement patterns, and there are still deficiencies in dynamic data processing and real-time feedback capabilities for complex movement states.

In addition, rapid advances in wearable technology have provided new monitoring tools for personalized exercise training, enabling researchers to collect large amounts of physiological and exercise data in real-world environments for more in-depth analyses of exercise effects and health status. Luo et al. explored the research progress of superhydrophobic flexible strain sensors in human exercise monitoring [6]. Han et al. studied the effect of visual walking training based on wearable walking guides on the walking ability of frozen gait patients with Parkinson's disease [7]. Li explored the progress of wearable devices in motor function rehabilitation training for stroke [8]. Wang Z conducted an analysis of real-time heartbeat monitoring in sports environments based on wearable device IoT system [9]. Liu et al. designed an optimized positioning technique for wearable sensors to monitor badminton sports training load [10]. Although existing studies have improved data acquisition and preliminary processing by integrating sensor technologies and algorithms, the existing methods are still insufficient in terms of deep data parsing, pattern recognition accuracy, and personalized feedback mechanisms, especially in terms of accuracy and adaptability in adapting to multiple sport types and complex dynamic environments.

3 Method

3.1 Overall System Architecture Design

The first task of the system design is to build an efficient and reliable framework that can collect, process and analyze athletes' movement data in real time. The system consists of three main components: data acquisition module, data processing module, and user interaction module.

The data acquisition module includes a variety of sensors (accelerometers, gyroscopes, heart rate monitors, etc.) that are mounted on wearable devices to collect physiological and movement data from athletes in real time [11]. The test data includes step frequency, heart rate, posture, and speed.

Data processing module: The main content of this study is a CNN-based analysis system that can analyze and identify complex actions and physiological signals. On this basis, this project proposes a motion performance prediction method based on deep learning.

Human machine interaction module: It provides users with an interface for analyzing motion data, and can provide personal fitness suggestions and techniques.

Users can adjust their training plans and engage in other interactions through the interface.

In discrete form, for 2D image processing, the convolution operation $S(i, j)$ is usually expressed as:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \quad (1)$$

Among them, I is the input image and K is the convolution kernel.

3.2 Integration and Synchronization of Sensor Data

To ensure the quality of collected data and improve the response speed of the system, it is necessary to effectively integrate the data collected by multiple sensors and achieve data synchronization. The frequency of data acquisition may be different for each sensor, so these data streams need to be synchronized through time stamping and data buffering techniques.

Data synchronization strategy: a central clock system is introduced, and before data processing, all sensor data are time-stamped according to the central clock to ensure data consistency in time.

Data fusion technology: using algorithms such as weighted average, combined with the accuracy and importance of the sensors, the data are fused and processed to enhance the accuracy of the system's judgment of the motion state.

The cross-entropy loss function L is as follows:

$$L = - \sum_{c=1}^M y_{o,c} \log(p_{o,c}), \quad (2)$$

where M is the total number of categories, y is the actual label vector, where if category c is the correct category, $y_{o,c} = 1$, otherwise 0, and $p_{o,c}$ is the prediction probability, which indicates the probability that the predicted outcome belongs to category c .

3.3 Design and Optimization of CNN Models

CNN is powerful in processing image and time series data, and is particularly suitable for extracting useful information from complex motion data. The model design needs to consider the characteristics of motion data, such as time dependence and spatial features.

Network architecture: Choosing a CNN architecture suitable for time series analysis can effectively extract features in the time dimension.

Activation function and pooling layer: the ReLU activation function is used to increase the nonlinear processing capability, and the maximum pooling layer is used to reduce the feature dimensions and improve the abstraction capability and computational efficiency of the model.

Regularization and optimization algorithm: to avoid overfitting, Dropout technique and weight decay method are introduced. Meanwhile, an optimizer is used, whose adaptive learning rate advantage can accelerate the convergence speed.

In deep learning models, batch normalization helps to speed up the training process and improve the stability of the model. This technique is achieved by adjusting the mean $\hat{x}^{(k)}$ and variance $y^{(k)}$ of the layer inputs:

$$\hat{x}^{(k)} = \frac{x^{(k)} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (3)$$

$$y^{(k)} = \gamma \hat{x}^{(k)} + \beta, \quad (4)$$

where $x^{(k)}$ is the input in batch B , μ_B and σ_B^2 are the mean and variance of the batch, respectively, ϵ is a small constant added for numerical stability, and γ and β are learnable parameters used to recover representativeness that may have been lost by normalization.

3.4 Experiments and Model Training

The experimental design must be able to verify the efficacy and adaptability of the model. Selecting an appropriate group of athletes for the experiment and collect data under various exercise modes.

Data collection: collecting exercise data in a controlled environment and actual exercise scenarios to ensure that the data cover common exercise types and changes in the physiological state of the athletes.

Preprocessing and feature extraction: the collected data is preprocessed, including denoising, normalization, etc., and then features are automatically extracted using the convolutional layer in the CNN model.

Training and validation: cross-validation techniques are used to assess the generalization ability of the model. The training set is used for model training, and the validation and test sets are used to assess the accuracy and utility of the model.

3.5 Performance Evaluation and Optimization

The performance of the system is comprehensively evaluated, including accuracy, real-time and user satisfaction, and the model is continuously optimized through a feedback loop.

Evaluation metrics: defining multiple performance metrics, such as accuracy, recall, F1 score and real-time response time.

Optimization strategy: adjusting the parameters of the CNN model, such as the depth of layers, the number of neurons, the learning rate, etc., according to the performance evaluation results.

User feedback: collecting feedback from users (athletes and coaches) to understand the system's performance and room for improvement in actual use.

With these detailed steps and technical details, this study aims to develop an efficient, accurate and user-friendly CNN-based wearable sports training monitoring system, which not only improves the scientific and personalized level of sports training, but also provides strong technical support for sports health management [12, 13].

In the gradient descent update rule, for parameter θ , the update rule is:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla_{\theta} J(\theta_t), \quad (5)$$

where θ_t is the value of the parameter at time t , η is the learning rate, and $\nabla_{\theta} J(\theta_t)$ is the gradient of the function J about θ .

4 Results and Discussion

4.1 Experimental Setup

With the rapid development of wearable devices and deep learning technology, CNN-based wearable sports monitoring system has gradually become a research hotspot in the field of sports training and health management. The purpose of this paper is to comprehensively evaluate the performance of CNN-based wearable sports monitoring system through a series of experiments and analyze its performance in different sports environments and conditions.

(1) Experimental environment and parameter settings

In order to simulate real-world sports scenarios, this study chose indoor gymnasiums and outdoor sports venues as test environments, covering a wide range of sports types, such as running, bike riding, and fitness training. The wearable device used for the experiment integrates accelerometers, gyroscopes, and heart rate sensors, which can capture athletes' motion data in real time. The data acquisition frequency was set to 100 times per second to ensure the accuracy and real-time performance of the data.

A total of 30 athletes, half men and half women, with an age distribution of 18–35 years old, participated in the experiment. They were asked to wear wearable devices under various sports conditions and to complete specified sports tasks.

The CNN model was configured using a 10-layer structure with a depth of 10 layers, including 5 convolutional layers and 5 pooling layers. The model uses ReLU as an activation function to improve nonlinear fitting. To prevent overfitting, Dropout technique is introduced in the model with the scale set to 0.5.

(2) Evaluation Indicators

In order to comprehensively evaluate the performance of the system, we chose the following metrics as the basis for evaluation:

Accuracy: it measures the percentage of motion types correctly recognized by the model. The higher the accuracy rate, the better the model's ability to recognize motion types.

Real-time: it evaluates the time delay from data acquisition to result output. Better real-time performance indicates that the system is able to provide motion monitoring results faster and fulfill the need for real-time feedback.

User satisfaction: feedback from athletes and coaches on their experience of using the system is collected by means of questionnaires. The higher the user satisfaction, the more attractive and practical the system is in practical application.

4.2 Analysis of Results

(1) Static heart rate monitoring

The results of static heart rate monitoring are shown in Fig. 1 (Fig. 1a shows the monitored static heart rate (initial/2 min), and Fig. 1b shows the actual static heart rate and the monitored static heart rate (1 min/3 min)).

First, we noted that from Participant 1 to Participant 29, their static heart rate monitoring maintained good stability at initial, 1-min, 2-min, and 3-min intervals, which was comparable to the actual static heart rate. This suggests that the monitoring system is able to accurately capture and reflect changes in participants' static heart rates over a short period of time (within a few minutes).

However, in the data from Participant 30, we found an anomaly. Although the initial and 1-min static heart rate monitoring was similar to the actual heart rate, the monitored heart rate suddenly and dramatically decreased at 2 and 3 min, which was inconsistent with the previous trend. This could be due to device malfunction, sensor error, or a sudden change in the participant's physical state.

From the perspective of long-term stability testing, the majority of participants' data indicated a high level of stability and reliability of the monitoring system. However, the data from Participant 30 reminded us that regular maintenance and calibration of the monitoring system was also needed in practical applications to ensure that it continued to monitor and record data accurately.

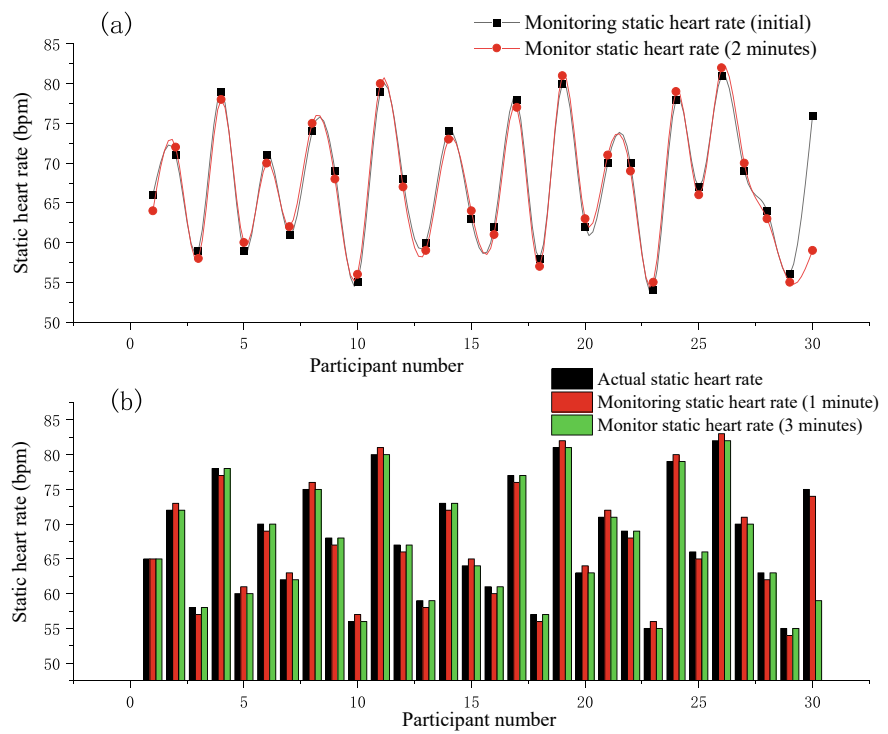


Fig. 1 Static heart rate monitoring results

In summary, the data analysis showed that the monitoring system was able to accurately reflect the static heart rate of participants in most cases, but there may be errors in individual cases.

(2) Single exercise type test

Participants numbered 1–10 were running with a duration of 15 min of exercise; participants numbered 11–20 were bicycling with a duration of 20 min of exercise; and participants numbered 21–30 were fitness training with a duration of 30 min of exercise.

First, the average heart rate of the participants was generally higher in the running exercise mode, indicating that running is a high-intensity aerobic exercise that can effectively improve cardiorespiratory function. At the same time, the values of maximum heart rate and maximum acceleration reflect the greater pressure on the body during running.

Secondly, in the cycling mode, the average heart rate and acceleration values of the subjects are relatively small; this may be due to the fact that cycling mainly involves the lower limbs, which causes less stimulation to the heart and heart function compared to full body running. But cycling is still an effective way to improve endurance.

Finally, in the health training mode, this method can generate significant variations in heart rate and acceleration. This is because health training is often a combination of strength training and aerobic exercise, which can provide sufficient exercise for muscle groups, improve physical strength and endurance. The maximum heart rate and maximum acceleration also reflect the greater challenges that the human body needs to face during healthy exercise.

Overall, these data demonstrate that the physiological responses of participants vary under different exercise methods. Running and fitness training can bring about higher changes in heart rate and acceleration, with significant enhancement of cardiorespiratory fitness and muscle strength, while bicycling is a milder form of aerobic exercise that is suitable for long-term adherence. These data are important for developing personalized exercise programs and evaluating the effects of exercise.

The results of the single motion type test are shown in Fig. 2 (Fig. 2a shows the average acceleration and maximum acceleration, and Fig. 2b shows the average heart rate and maximum heart rate).

(3) Multi-motion mode alternating test

The results of the multi-motion mode alternation test are shown in Table 1

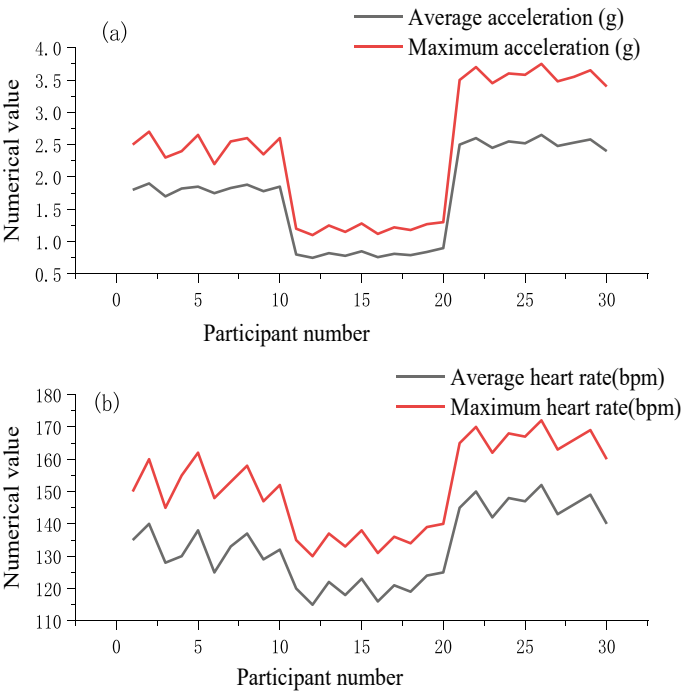


Fig. 2 Single exercise type test results

Table 1 Results of multiple motion mode alternation test

Serial number	Time point (sec)	Sports type	System response time (ms)	Heart rate (bpm)	Average acceleration (g)
1	0	Static	–	70	0.1
2	10	Running begins	200	120	1.5
3	30	Running	–	140	1.8
4	60	Running ends, cycling begins	180	125	0.9
5	80	Cycling	–	110	0.7
6	120	Cycling ends, fitness training begins	220	135	2.0
7	150	During fitness training	–	155	2.5
8	180	Fitness training ends, running begins	190	122	1.6
9	210	Running	–	142	1.9
10	240	Run finished, stationary	210	90	0.2

Firstly, in terms of system response time, although there are some small fluctuations, overall, the response speed of the entire system is still relatively fast, basically maintained within 200 ms. The experimental results show that the system can quickly and accurately convert actions and quickly adapt to various action needs. Secondly, from the aspects of heart rate, acceleration, etc., various parameters will change to varying degrees under different action modes. For example, when running at the beginning, both heartbeat and acceleration increase rapidly, which means the body has entered a high-intensity training. Moreover, when they switch to cycling or strength training, their heart rate, acceleration and other indicators are also consistent with the characteristics of the above-mentioned exercise methods, reflecting the differences in the effects of different exercise methods on the body.

In addition, the algorithm also has the ability to convert each action mode. The experimental results show that the system has high flexibility and adaptability, and can quickly switch between various action modes to ensure that students can achieve the best exercise effect during practice. Overall, the system has high response speed and accuracy, and can adapt to changes in various actions. The results of this study have certain reference value for evaluating system performance and optimizing training programs.

(4) Monitoring error (heart rate) versus number of signal losses

The monitoring error (heart rate) and the number of signal loss under different participant numbers are shown in Fig. 3.

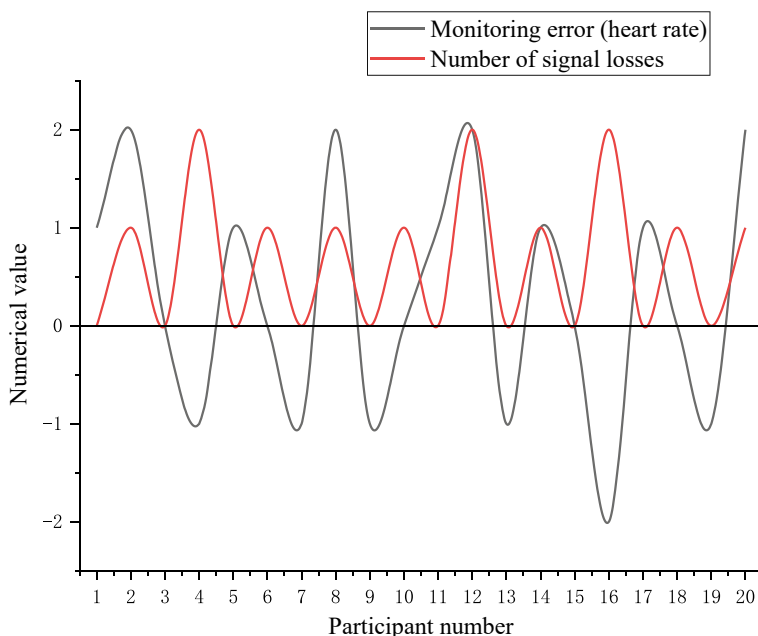


Fig. 3 Monitoring error (heart rate) and signal loss frequency under different participant numbers

Firstly, from the perspective of heart rate monitoring error, the error range is $-2 \sim 2$. Positive means high heart rate, negative means low heart rate, and 0 means precise monitoring. This distribution indicates that under certain conditions, heart rate monitoring devices can accurately record heart rate, but there are also certain errors. It should be pointed out that “+ 1”, “− 1”, and “0” appear more frequently, which may be related to factors such as instrument design, performance, individual physique, and exercise habits.

Secondly, for the number of signal losses, the value range of this algorithm is 0–2. There are many reasons for signal loss, such as equipment failures, signal interference, or the movement patterns of participants. From the test results, it can be seen that most participants have not experienced signal loss, indicating that the device can reliably transmit data under most conditions. However, some subjects also experienced more frequent signal loss, which requires further research.

All these data indicate that during the heartbeat monitoring period, the subjects had monitoring errors and signal loss. By analyzing these data, we can better understand user behavior and their behavior. For example, when monitoring errors are significant or packet loss frequency is high, it is necessary to optimize the performance of the device or improve the user’s technology. Meanwhile, the results of this study can

also be used for further research and improvement, in order to improve the accuracy and stability of heart rate monitoring.

(5) Long-term stability test

The results of the long-term stability test are shown in Table 2.

Firstly, we analyzed the response time of the system and found that most of the response time was maintained between 185 and 210 ms, indicating that the system's processing of heartbeat signals was stable. However, during September 10th and 15th, the system's response time increased, perhaps due to the system load or network conditions at that time. Secondly, the changes in heart rate values have a certain pattern; at 8:00 am, the heart rate is generally slower, but from 12:00 pm to 6:00 pm, the heart rate will increase. This is a natural phenomenon, and the human heart rate will change with the amount of exercise done every day. The experimental results show that the system has good heart rate monitoring performance. From the system failure and fault records, it can be seen that at 8:00 am on September 10th, a system restart event was recorded, which may have been carried out to improve the operational efficiency or stability of the system. On the evening of September 15th at 18:00, there was a data transmission delay, which may have been caused by network issues or high load. In the experiment, it was found that the above faults or malfunctions had no significant impact on the heart rate monitoring results, indicating that the system had strong fault tolerance and self-healing ability.

Table 2 Long-term stability test results

Date time point	System response time (ms)	Heart rate (bpm)	Date time point (h)	System FAULT/error record
2023/9/1	8:00	190	75	Nothing
2023/9/1	12:00	185	120	Nothing
2023/9/1	18:00	195	145	Nothing
2023/9/2	8:00	192	78	Nothing
2023/9/2	12:00	188	118	Nothing
...	...	...	...	
2023/9/10	8:00	205	73	System restart (updating software)
2023/9/10	12:00	190	122	Nothing
2023/9/15	8:00	200	77	Nothing
2023/9/15	18:00	210	140	Data transmission delay
2023/9/20	8:00	195	76	Nothing
2023/9/25	12:00	205	125	Nothing
2023/9/30	18:00	198	142	Nothing

Overall, wearable sports training monitoring systems can maintain good performance in most cases and can monitor heart rate. Although some minor errors may occur at times, the system is able to resume normal operation on time. To further improve the stability and reliability of the system, regular maintenance and upgrades should be carried out, and its working status should be monitored in real time to discover and eliminate hidden dangers in a timely manner.

5 Conclusion

In order to improve the accuracy and real-time performance of motion monitoring, this paper plans to develop a wearable motion training monitoring system based on convolutional neural networks. On this basis, this paper adopts a multi-source information fusion method based on deep neural networks to process and analyze complex data in mobile environments. In addition, this paper will also provide a specific implementation of the method, optimize the data processing process, and conduct experimental verification of the model. Experimental results have shown that a wearable action monitoring system based on convolutional neural networks can recognize and monitor multiple action modes, and its data processing accuracy and processing speed are significantly improved compared to traditional methods. The performance evaluation indicators (accuracy, etc.) of the system have reached high requirements. In real-time experiments, the system can quickly respond and process real-time data, which can well meet the real-time feedback requirements of sports training for real-time information. In addition, the user's satisfaction with the user is also quite high.

Although there have been good results, there are also certain shortcomings. Firstly, due to the excessive reliance on specific data and environment for model training and validation, its universality and adaptability to multiple scenarios and action scenarios are constrained. Secondly, in future research, it is necessary to conduct in-depth research on the energy consumption, hardware costs, and other aspects of the system. In addition, the recognition ability of this method for some complex movement modes and rare behaviors still needs to be improved. The next step of research will focus on some important aspects. Firstly, this paper intends to further study methods based on data augmentation and transfer to improve their robustness to various mobile scenarios. Secondly, we will further optimize the hardware to achieve the goal of reducing power consumption and costs, thereby achieving miniaturization and low cost. On this basis, this paper intends to further study more complex algorithms, especially fusion learning, multi task learning, etc., to enhance the system's ability to respond to complex actions and rare events. The research results of this paper are expected to be more applied in monitoring the daily activities of chronic disease patients and preventing falls in the elderly, providing more comprehensive technical support for personalized health management.

References

1. Yang J, Wang J, Guo J et al (2023) Knowledge graph analysis of the development situation of wearable sports monitoring technology. *Mil Med* 47(12):881–887
2. Liu Y, Dong H, Wang J (2024) The effect of exercise imagination training combined with task oriented functional training on the recovery of lower limb function in stroke hemiplegic patients. *Chin J Phys Med Rehabil* 46(04):322–324
3. Wang J, Liang H, Wang Y et al (2021) Quantitative analysis of respiratory mode parameters based on wearable systems. *J Biomed Eng* 38(5):893–902
4. Zhang Z, Fu D, Ling M (2022) Design of low-power cow motion monitoring ear tags based on dynamic duty cycle sampling. *Electron Devices* 45(6):1510–1516
5. Su Y, Li W, Shi Y (2023) Research on the optimization mechanism, action mechanism, and application of wearable resistance training in sports performance. *Sports Sci* 43(9):89–96
6. Luo L, Lin X, Jiang J et al (2023) Research progress of superhydrophobic flexible strain sensors in human motion monitoring. *J Compos Mater* 40(7):3837–3851
7. Han Y, Shi Y, Song W et al (2024) The impact of visual walking training based on wearable walking guides on walking ability of Parkinson's disease patients with frozen gait. *Chin J Rehabil Med* 39(1):45–52
8. Li D (2023) Application progress of wearable devices in stroke rehabilitation training. *Gen Nurs* 21(6):786–790
9. Wang Z, Gao Z (2021) Analysis of real-time heartbeat monitoring using wearable device internet of things system in sports environment. *Comput Intell* 37(3):1080–1097
10. Liu TH, Chen WH, Shih Y et al (2024) Better position for the wearable sensor to monitor badminton sport training loads. *Sports Biomech* 23(4):503–515
11. Yang L, Amin O, Shihada B (2024) Intelligent wearable systems: Opportunities and challenges in health and sports. *ACM Comput Surv* 56(7):1–42
12. Balkhi P, Moallem M (2022) A multipurpose wearable sensor-based system for weight training. *Automation* 3(1):132–152
13. Rana M, Mittal V (2020) Wearable sensors for real-time kinematics analysis in sports: a review. *IEEE Sens J* 21(2):1187–1207