Journal of Al-Qadisiyah for Computer Science and Mathematics VOL16(1) 2024, PP Comp. 124-136

Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Optimizing the Architecture of Convolutional Neural Networks Using Modified Salp Swarm Algorithm

Entesar H. Abdulsaed[1*], Maytham Alabbas[1], Raidah S. Khudeyer[2]

[1]Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq.
Email: hopidhadha@gmail.com, ma@uobasrah.edu.iq
[2]Department of Computer Information Systems, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq.
Email: raidah.khudayer@uobasrah.edu.iq

ARTICLE INFO

Article history:
Received: 3 /3/2024
Rrevised form: 21 /3/2024
Accepted : 28 /3/2024
Available online: 30 /3/2024

Keywords:

Deep learning

Convolutional neural networks

Hyperparameters optimization

Salp swarm algorithm

ABSTRACT

Deep learning is highly effective in dealing with complex tasks such as image classification and recognition. However, finding the optimal architecture's hyperparameters for Convolutional Neural Networks (CNNs) to achieve the best performance and parameter regularization can be challenging. Metaheuristic optimization algorithms can be utilized to find solutions in this context. In this research, a computerized CNN was adjusted using an improved Salp Swarm Algorithm (SSA) to enhance crucial CNN settings, like dropout rate, hidden units, learning rate, and batch size. The refined design was tested on two standard datasets. MNIST and Fashion MNIST. The outcomes displayed model performance achieving accuracy levels of 99.6% for MNIST and 94.08% for Fashion MNIST. This tuned system outperformed the existing practices by 0.2% and 0.04% for each dataset while also cutting down on computational expenses. The fusion of SSA with CNNs displayed adaptability and resilience opening up possibilities, in image classification and consistently delivering outstanding outcomes.

MSC.

https://doi.org/10.29304/jqcsm.2024.16.11450

1. introduction

Deep Learning (DL) falls under the umbrella of machine learning. Relies, on Artificial Neural Networks (ANNs) to learn from data and enhance performance on specific tasks [1]. These networks imitate the structure and functionality of the brain enabling them to process information and make decisions in a manner. DL can be viewed as a group of interconnected classifiers each tasked with recognizing features or patterns within the input data [2]. These classifiers use regression in conjunction with activation functions to extract more advanced features ultimately producing the desired output. It should be emphasized that achieving optimal performance from CNNs relies greatly on fine-tuning their hyperparameters [3, 4]. These settings, including the learning rate, batch size, and kernel size, dictate how the network functions and directly affects its overall performance [5, 6]. Consequently, finding the ideal

*Corresponding author *Entesar H. Abdulsaed Almuhsen*

Email addresses: hopidhadha@gmail.com

Communicated by 'sub etitor'