Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Deep Learning-Based Siamese Neural Network for Masked Face Recognition

Ali Haitham Abdul Amir [1], Zainab N. Nemer[2]
[1,2] *Department of Computer Science, Basra University, Basra, Iraq*
Corresponding Author:
[1] *Master in the Department of Computer Science at the University of Basrah*
[2] *Assist. Prof.Dr. in the Department of Computer Science at the University of Basrah*
Email: pgs.ali.hatham@uobasrah.edu.iq
Email: zainab.nemer@uobasrah.edu.iq

ARTICLE INFO	ABSTRACT

Received: 29 Dec 2024
Revised: 12 Feb 2025
Accepted: 27 Feb 2025

Face recognition is one of the essential elements of security and surveillance, and is widely used in mobile devices and public surveillance systems. However, face blockage poses a significant challenge for developing practical and effective applications in this area. With the spread of the COVID-19 pandemic, wearing face masks has become essential in public places to reduce the transmission of the virus. Despite their health importance, masks have complicated the task of facial recognition, as the face is only partially exposed, hindering traditional recognition systems and increasing the difficulty of the work of security personnel. In this study, the Siamese neural network was used as an innovative approach to human face recognition under partial face-occlusion. The model's performance was tested using RMFRD and MFR2 databases, and the results showed high accuracy. On the RMFRD dataset, the ResNet50, EfficientNet, and Xception models achieved 99% accuracy, while the MobileNet model achieved 99% accuracy. On the MFR2 dataset, the ResNet50, EfficientNet, and Xception models achieved 99% accuracy, while the MobileNet model achieved 98% accuracy. This approach shows great effectiveness in dealing with the challenges posed by face obscuration, as a result enhancing the capabilities of facial recognition systems in real-life scenarios and enabling their more efficient use in security and surveillance. The Siamese network has proven to be highly effective at recognizing masked faces, making it highly relevant for applications in security, human-computer interaction, and many other areas affected by the COVID-19 pandemic.

Keywords: Siamese neural network , Deep learning ,Couvolution neural network, Mask face recognition , COVID-19.

INTRODUCTION

Since face recognition is a basic computer vision problem that has advanced significantly over time, utilising sophisticated deep learning algorithms and feature extraction across numerous applications and programs, face recognition technology has emerged as one of the most popular artificial intelligence applications [1]. The World Health Organisation advised wearing a mask in public places while keeping a safe and effective distance to prevent the virus and infection from spreading, so this new challenge has been in use since the spread of the corona virus in a capacity and posed a serious risk to people. With the use of masked face recognition, facial recognition algorithms could be more realistic. Furthermore, the many types of masks employed in this technology prevent typical facial recognition systems from functioning effectively, forcing researchers to investigate alternate methods. Computer vision systems, artificial intelligence algorithms, and deep learning were created to recognise individuals [2][3]. Facial recognition is the most significant method for identifying people in public locations and preventing the infection when wearing a mask, such as ATMs, train stations, and airports. Blocked face recognition is a popular research topic among the computer vision field.

Previously, blocked face recognition systems were designed to identify and recognise an individual's face in the wild, where the contained portion of the face was randomly shaped and positioned. Meanwhile, the masked face frequently obscures the nose, lips, and cheeks, posing limits such as sensitivity to severe settings, lighting conditions, and partial

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

blockage [4][5]. To address the issue, a Siamese neural network-based face recognition model was created using half-faceted pictures. The Siamese neural network was combined with convolutional network models to detect the masked face. The suggested method's generalisability was examined using the MFR2, RMFRD data set, with an emphasis on assessing the approach's performance.

Several aspects in our study have helped to improve the accuracy of masked face recognition. Most notably, adopting the Siamese Neural Network, which depends on collecting characteristics from photos and analysing similarities rather than traditional classification, helped overcome the issue of partial face occlusion and enhance performance. Advanced models, such as ResNet50, EfficientNet, and Xception and MobileNet, have also been used for feature extraction since they are very accurate at extracting correct features even in the presence of face blockages. These models performed admirably, with the MobileNet model proving particularly useful in situations with restricted resources. Layer normalisation, data augmentation, and dropout layers were utilised to reduce over-adaptation. Furthermore, bespoke databases with partially masked faces, such as RMFRD and MFR2, have trained the model on real-world scenarios, considerably increasing its accuracy.

This paper is divided into six sections: Section 2 discusses previous studies in masked face recognition, Section 3 explains what the Siamese network is, Section 4 Suggested methods of masked face recognition, Section 5 problem statement, Section 6 discusses results, tool measures, and databases, and Section 7 concludes and suggests possible directions for future research in face recognition using masks.

RELATED WORK

Face recognition technologies have undergone significant development in recent years, supported by the use of deep learning methods that have contributed to improved performance and increased accuracy. However, the COVID-19 pandemic has posed new challenges, most notably the widespread use of face masks, which has negatively affected the efficiency of traditional facial recognition systems. In response to these problems, research has concentrated on building specialised deep learning approaches for recognising mask-covered faces. These approaches examine face characteristics altered by the use of masks, allowing for improved performance under different settings. Recent advancements also aim to better understand the aspects that influence face characteristics, such as environmental obstacles, and to improve the accuracy and adaptability of smart systems. This study attempts to overcome present limitations and produce more accurate and effective facial recognition systems in a post-pandemic future through the incorporation of deep learning technology.

Rajdeep Chatterjee et al. [6] proved that the Siamese Neural Network (SNN) can accurately detect the similarity or difference between two faces using little data. The suggested model focusses on the top half of the face and achieves noteworthy results in recognising persons with only the visible portion above the nose. Dataset 2 contains the faces of 20 celebrities, varying in age, brightness, emotions, and image quality. Each collection includes around 100 photos per individual, with masks worn at varying heights to replicate real-world circumstances. A test set was created with 30 images per individual from the same source. The model outperforms VGGFace2 in search time for face detection, achieving 93% accuracy and an average accuracy of 84.80±4.71% for partial face matching.

Mohammed Eman et al. [7] presented a novel deep learning-based solution to masked face recognition. The SSD-MobileNetV2 model was utilised for initial mask detection, while RPCA was employed to distinguish between concealed and clear components in pictures. To improve speed, the Gazelle Optimisation Algorithm (GOA) was used to optimise KNN features and the value of k. Experiments proved the superiority of the suggested technique, which achieved 98% accuracy on the LFW-SMFD dataset, 95% on ORL Face, and 97% on CASIA-WebFace, resulting in better precision and resilience in masked facial recognition.

Mehreen Fatima et al. [8] developed a facial recognition system using SVM and Random Forest to identify masked and unmasked faces. It achieves 98.2% accuracy on a dataset of 28 categories and compares favorably with state-of-the-art methods.

Saja Mohsen Abbas [9] proposed using CNN to recognize masked faces by extracting features from the eyes, forehead, and eyebrows. Using the "Yale B Extended Database" with medical masks, the method achieved 95% accuracy, showing its effectiveness.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Sammy V. Militante et al. [10] developed a system to detect face masks with 25,000 images at 224x224 resolution, achieving 96% accuracy. Using a Raspberry Pi, it triggers an alert and captures an image when no mask is detected, helping prevent virus transmission.

Mabirizi Vicent et al. [11] developed a deep learning model to detect masked faces, trained on 1,000 images from Kabale University students. Using pre-trained CNNs, VGG19 achieved 91.2% accuracy, Inception V3 90.3%, and VGG16 89.69%. The enhanced model reached 90.32%.

Ratnesh Kumar Shukla et al. [12] obtained 99.82% accuracy on the CIFAR, MIST, RMFRD, and SMFRD datasets. They tested VGG16, VGG19, ResNet50, and ResNet101, and the suggested model scored 99.30%. VGG19, ResNet50, and ResNet101 achieved accuracies of 99.54 percent, 78.70 percent, and 98.64 percent, respectively. The study emphasises the suggested model's superiority, which uses MobileNet V2 with transfer learning to monitor face mask usage for COVID-19 prevention.

Hoai Nam Vu et al. [13] proposed combining deep learning models with Local Binary Pattern (LBP) features for masked facial recognition. The method uses deep models for face detection and feature extraction, along with LBP features from the eyes and eyebrows. It achieved an 87% F1-score on the COMASK20 dataset and 98% on the Essex dataset, outperforming techniques like Dlib and Insight Face.

Omar Adel Mohi et al. [14] proposed a unique approach to masked face recognition utilising the BoCNN framework, which integrates Convolutional Neural Networks (CNNs) with the Bag of Words (BoW) idea to extract distinguishing characteristics from visible facial parts such as the eyes and brows. The work emphasises the importance of deep learning approaches, namely CNNs, in face recognition tasks because of their capacity to analyse and extract hierarchical and spatial information from pictures. The models achieved excellent accuracy across many datasets, scoring 97.35%-99.33% on LFW, 96.5%-98.48% on RMWMF, 97.86%-98.91% on MFR2, and 97.78%-99.37% on the Celebrity dataset, demonstrating their usefulness in accurate target categorisatio

Table 1. A summary result of related work

Study/ Ref	Year	Methodology	Database	Performance	Limitations
Sammy V. Militante et al	2020	VGG-16 CNN	25,000 photos	96% accuracy	Accuracy, data, performance, privacy, environments.
Hoai Nam Vu et al	2022	deep learning CNN and Local Binary Pattern (LBP).	COMASK20 Essex dataset.	87% f1-score 98% f1-score	Limited reliance on face features Performance varies between groups High computational complexity
Rajdeep Chatterjee et al	2022	Siamese neural network with Resnet50 model	The dataset2 contains 20 celebrity faces	results of 93%	Limited Circular, Fluctuations in accuracy, Limited Data
Mehreen Fatima et al	2022	Support vector machine (SVM) and Random Forest (RF)	The dataset is collected in the form of images for 28 classes	97% accuracy	Blocking features, generalization, environmental factors, privacy,

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

					susceptibility, data, time.
Saja Mohsen Abbas et al	2023	CNN	Extended Yale B database	accuracy of 95%	Lack of features, lighting, limited data, versatility.
Mohammed Eman et al	2023	KNN and robust principal component analysis (RPCA)	LFW-SMFD and real dataset	97%	Limitations include modeling, environmental challenges, and computational complexity.
Hayat Al-Dmour et al	2023	Convolutional Neural Network (CNN)	RMFD CFR Masked-FaceNet	97% accuracy	Data Bias, Risk of overtraining, Computational resource requirements, Big Data Expansion
Mabirizi Vicent et al	2023	VGG19 Inception V3 VGG16	1000 photos from student of Kabale University	VGG19: 91.2% Inception V3: 90.3% VGG16: 89.69%	generalization, limited data, performance, masks.
Ratnesh Kumar Shukla et al	2023	ResNet50, ResNet101, InceptionV3 VGG19, MobileNet V2	CIFAR10 MNIST RMFRD	Accuracy between 78.70% and 99.82 %	Data dependence, detection accuracy, generalization, complexity, over-personalization, realistic performance.
Omar Adel Muhi et al	2024	ResNet50, ResNet101, InceptionV3, VGG19	Celebrity MFR2 RMWMF LFW	Accuracy between 97.78% and 99.37% 97.86% and 98.91%, 96.5% and 98.48% 97.35% and 99.33%	Complexity of processing, diversity of masks, limited generalization, impact of the environment.

SIAMESE NEURAL NETWORK

Bromley et al. introduced the Siamese Neural Network (SNN) in 1994. For signature verification reasons. This network trains two similar neural networks with the same parameters and weights, resulting in two feature vectors when two separate items are added. These vectors are compared using distance measures, such as Euclidean distance, which serve as a loss function throughout the learning process to assess model performance [15]. The Siamese network's core function is to learn generic feature representations based on the closeness or distance between the feature vectors created by the input, since this process is critical to comprehending the connections between distinct elements. The Siamese network consists of two identical neural structures that share the same layers and parameters, which means that any parameter update in one network is reflected directly on the other [16]. This design proved effective in various applications, as it was suitable for applications that rely on learning similarity, such as signature verification. With the passage of time and recent developments in neural networks, the uses of the Siamese network have expanded to include multiple areas, such as facial verification, where it is used to determine whether two images belong to the same person, and systems A recommendation that relies on presenting elements similar to user preferences using similarity metrics [17].

The Siamese network is very flexible, as it can be divided into two main models according to the input method. The first model relies on the input of two images to calculate their similarity using a distance scale such as the Euclidean distance. The second model relies on the introduction of three images (reference image, similar image, and different image) to improve the measurement of similarity and enhance the accuracy of the model. This method is used to provide a more accurate representation of the similarity and difference between the inputs [18]. To measure similarity, multiple methods such as cosine similarity or Pearson's correlation coefficient are used. However, these methods may not be sufficient when comparing elements with different features. Therefore, Siamese networks were introduced in 1993 as a solution to measure similarity between two signatures , relying on identical feeder networks that share weights [19]. The outputs of these networks are compared using distance measures, such as Euclidean distance, to determine whether the inputs are similar (value close to 0) or different (close value). of 1). During training, the score is compared with the scores attached to the data to evaluate the model's performance through the loss function. Siamese networks can be divided into two main models based on the number of input parameters: Pairs: Two images are inserted to determine their similarity, Trigeminus: Three images (reference, similar, and different) are inserted to improve the measurement of similarity [20].

Figure 1 depicts the broad structure of the Siamese Neural Network, which is made up of numerous fundamental components that work together to compare two pictures. Initially, two pictures, Image A and Image B, are shown, with a convolutional neural network (CNN Encoder) extracting distinguishing information from each independently. The convolutional network's output is subsequently transformed to flat representations via the flatten layer, allowing for easier post-processing. The embedding layer is then utilised to turn the pictures into low-dimensional vector representations containing the images' retrieved characteristics. Then, these representations are compared using the matching layer, which determines how similar the two images are. In the end, the Similarity Score is calculated, which is the degree of similarity between the two images, where the score closest to 1 indicates that the two images are similar, while the score closest to 0 indicates that the two images are not the same.

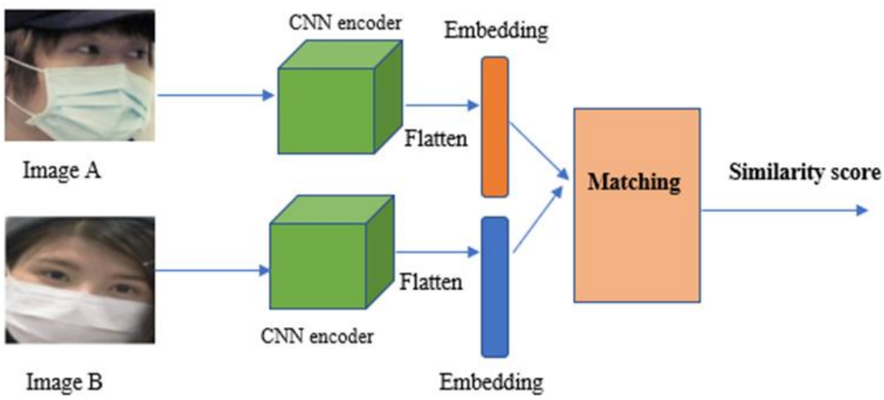


Figure 1. The general architecture of the Siamese Neural Network system

Siamese networks have proven to be highly successful in reducing dimensions and in poorly supervised metric learning, making them an effective tool in areas such as identity verification, recommendation systems and multimedia. The network relies on exploiting the similarity between the elements more efficiently than traditional

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

methods, making it an ideal choice for dealing with complex challenges in computer vision. **Figure 2** Architecture of a Siamese Neural Network where the input is a pair of elements [18]

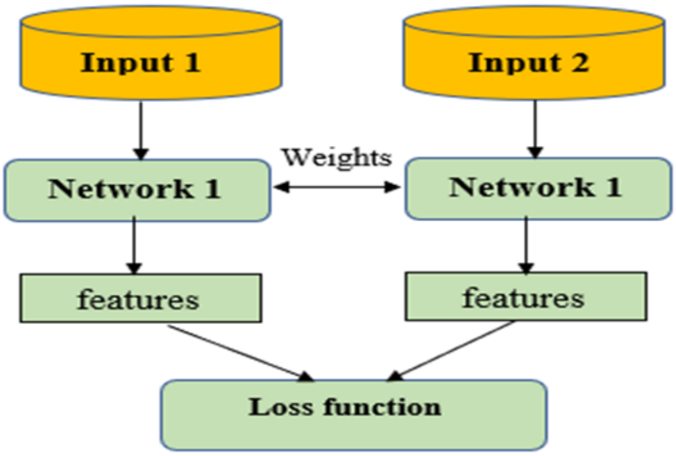


Figure 2. Architecture of a Siamese Neural Network where the input is a pair of elements

METHOD OF MASKED FACE RECOGNITION

As shown in Figure 3. the proposed method consists of three main steps: preprocessing, extracting features using the Siamese network based on Relu activation, and determining similarity between images using a distance meter

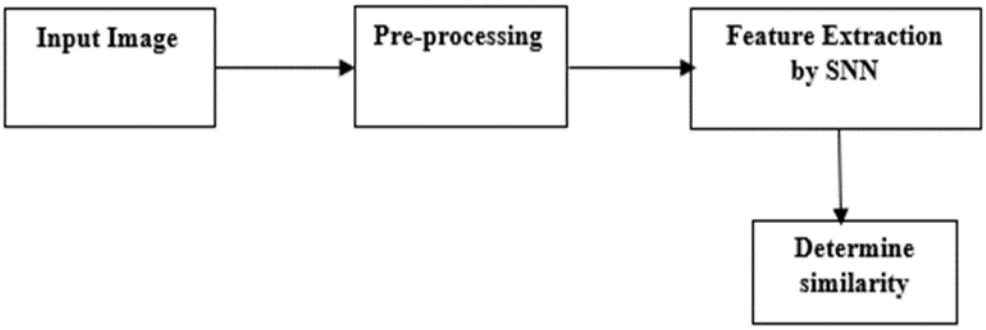


Figure 3. Proposed approach for masked face recognition

Pre-processing phase

Various steps were carried out on the face mask images, and they can be summarized as follows:

- Face Detection: Use the OpenCV library to detect the face in the photo.
- Add Face Masking: Add effects such as GaussianBlur and GaussNoise on photos
- Grayscale Conversion
- Normalization
- Data Augmentation

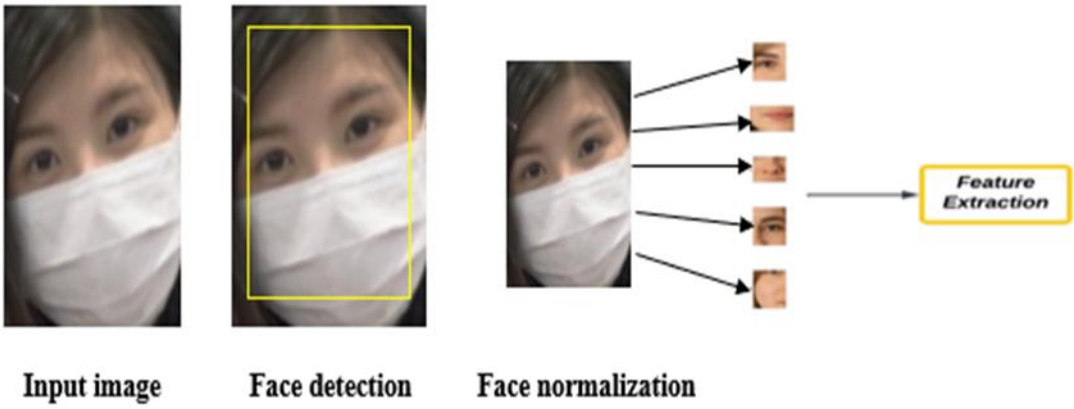
Feature Extraction

- Input image: The input image is a color image of 105x105 pixels after applying transformations.
- Convolutional Layer: Pre-trained models are used as a feature extraction layer, as this layer contains multiple convolution operations internally.
- Activation Function: The ReLU activation function is used after each fully connected layer to stimulate the network.
- MaxPooling has internal pooling processes that gradually reduce image dimensions.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Fully Connected Layers: After extracting features from convolution layers, they are passed through fully connected layers with dimensions of 512 and 256 at the end.
- Dropout: A dropout layer is used after each fully connected layer with the aim of reducing overfitting and improving model generalization. In Figure 4, we show the most significant features found in an extracted image.



Figurer 4. Steps to extraction features found in an image.

PROBLEM STATEMENT

Siamese networks have struggled to recognise masked faces because medical masks typically conceal the lips and nose, two essential elements used by the network to extract features and assess whether or not the images reflect the same person. Masks can cover substantial portions of the face, limiting the amount of information available to the model and making it difficult to discern between related photos. The Siamese grid's purpose is to compare attributes retrieved from two photos to see how similar they are, but with the mask hiding the core features, the model struggles to perform an accurate comparison. Masks can also blur the form of the face, complicating the comparison of photographs.

Also, changes in lighting and shadows may affect the accuracy of the model, especially when facial features are covered by a mask. These challenges make it necessary to train the Siamese network using data that includes masked faces in a variety of settings to strengthen the model's ability to adapt to these variables. On the other hand, you have difficulty collecting training data that contains images of masked faces under different circumstances, which further complicates the model's task in dealing with these diverse styles. The effect of different angles on the shape of the masked face also adds difficulty in the comparison and recognition process. Finally, current models may lack the flexibility to handle multiple types of masks and changing lighting conditions, adding to the challenges facing the Siamese network in improving its accuracy when recognizing masked faces.

PROBLEM STATEMENT

In this section, the Siamese neural network for masked face recognition is done on Python through Google Collab, and two databases and performance measurements were used, and the results were discussed.

Data Preparation

Data preparation is an essential step in machine learning projects, as this stage focuses on preparing data to be ready for use by the model. The main goal of data preparation is to improve the quality of data and reduce possible errors during training. In **figure 5** The data preparation process usually includes the following stages

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

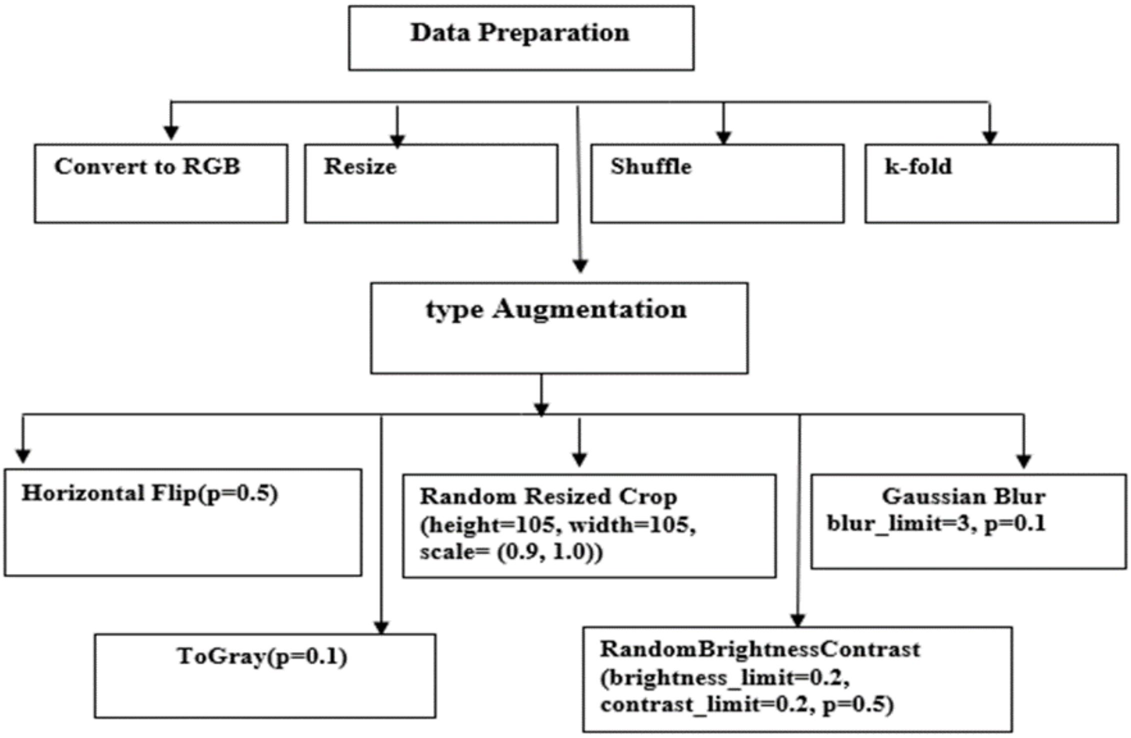


Figure 5. Shows the types of Augmentation

First dataset MFR2

The MFR2 dataset is one of the most prominent datasets used in real-world masked face recognition. The collection contains 269 images collected from the Internet, and includes 53 distinguished celebrities and politicians. Each character contains about five images, so that both the masked and exposed faces of the individuals are included. This collection has been enhanced by significant image modifications, where the dimensions of the images have been standardized to 105×105×3 pixels, with precise alignment of faces within images, which contributes to enhancing the accuracy and efficiency of models in recognizing faces in different conditions. **Figure 6** shows samples from the RMFRD dataset.

Figure 6. Samples of masked faces used in masked face recognition.

MFR2 dataset The accuracy values achieve by the ResNet-50 models 99%, EfficientNet 99%, Xception 99%, and MobileNet demonstrating their effectiveness in correctly categorizing or predicting the target. These models showed exceptional taxonomic performance in the task assigned to him. Table 1 indicates to a comparison of the performance of different deep learning models. Models compared include ResNet50, EfficientNet, Xception, and MobileNet With the Siamese network to achieve the best results.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

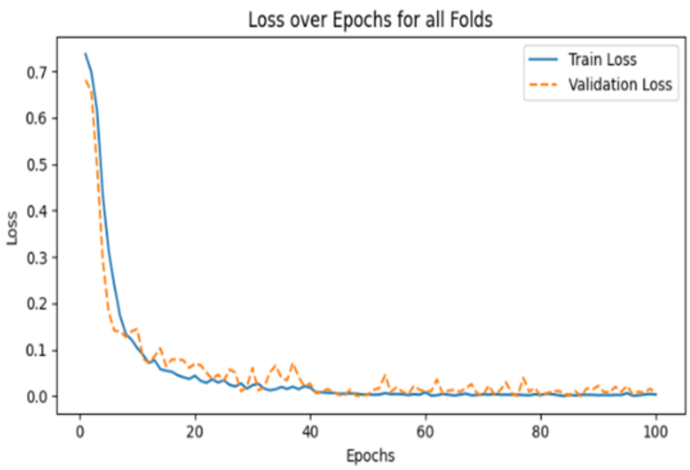
Research Article

Tabel 2. Result MFR2 dataset

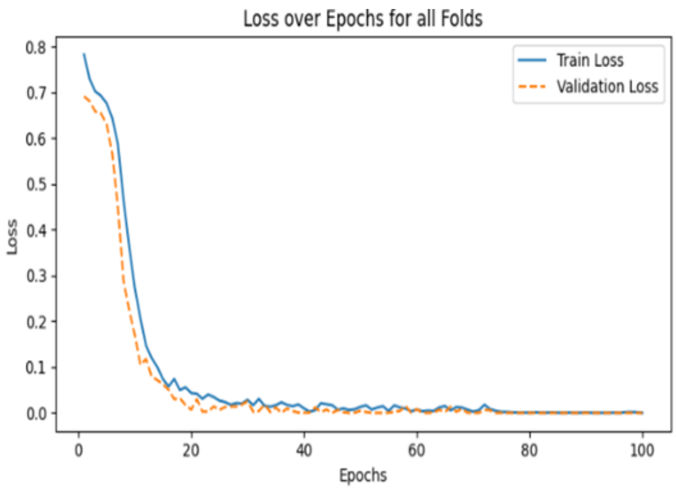
model	accuracy	precession	recall	F1-score	loss	MSE
Res-Net50	**99%**	**99%**	**99%**	**99%**	**0.03**	**0.01**
EfficientNet	**99%**	**99%**	**99%**	**99%**	**0.2**	**0.006**
Xception	**99%**	**99%**	**99%**	**99%**	**0.004**	**0.007**
MobileNet	**99%**	**99%**	**99%**	**99%**	**0.03**	**0.0096**

Figure 7. Dual network model performance in masked face recognition using MFR2 database. The graph explains the rate of training data loss, showing a gradual decrease during the training process, as well as verification data loss that is constantly decreasing as the number of epochs progresses. This reflects an improvement in the overall performance of the model and its increases the accuracy in data verification

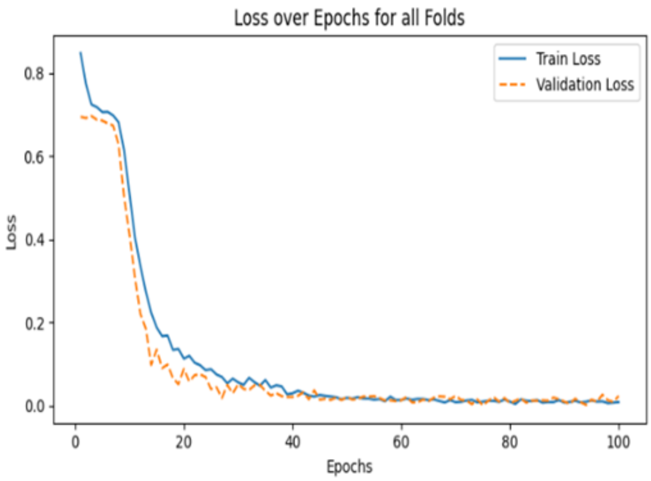
(A) Loss function in ResNet-50



(B) Loss function in EfficientNet



(C) Loss function in Xception

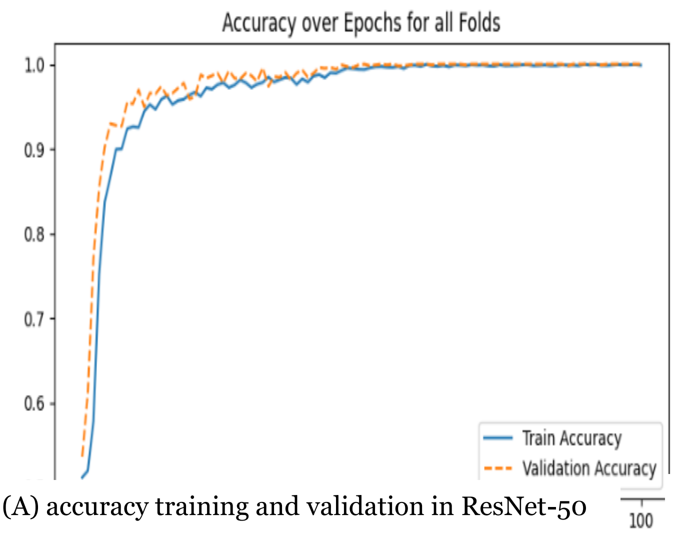


(D) Loss function in MobileNet

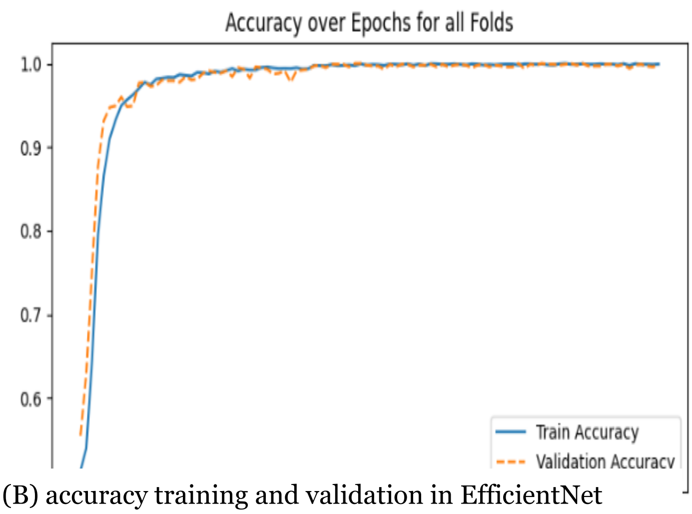
Figure 7. (A) explains the loss function in RwsNet-50 using 100 epochs. (B) is showing the loss function in EfficientNet using 100 epochs. (C) represents the loss function in Xception using 100 epochs. (D) indicates to the loss function in MobileNet using 100 epochs

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

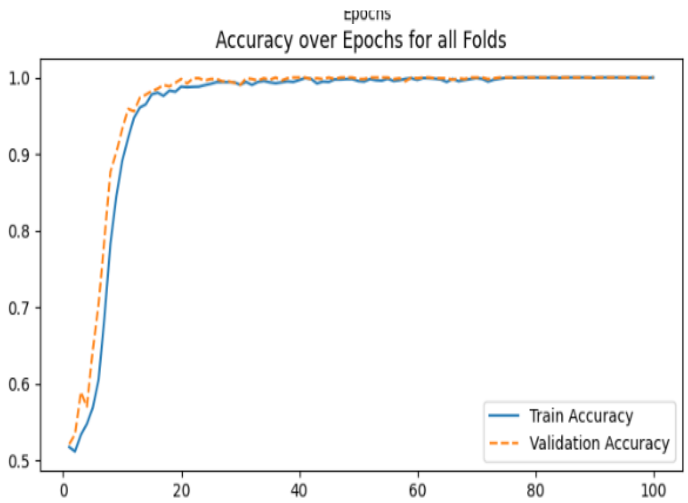
Research Article



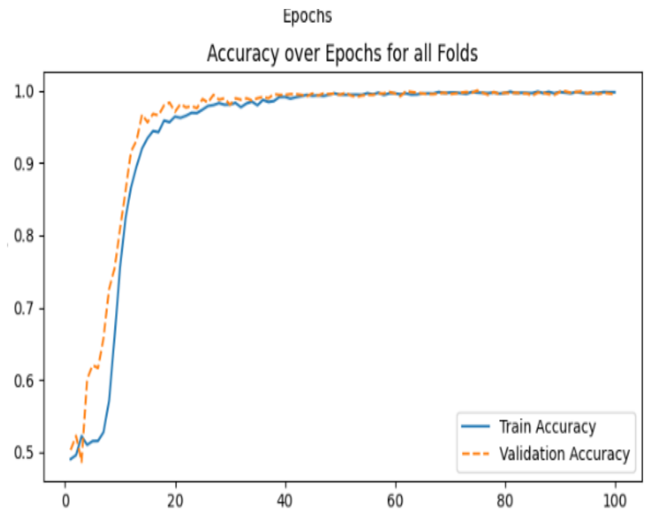
(A) accuracy training and validation in ResNet-50



(B) accuracy training and validation in EfficientNet



(C) accuracy training and validation in Xception



(D) accuracy training and validation in Xception

Figure 8. (A) is representing model accuracy value utilizing ResNet-50 in 100 epochs. (B) is indicating to the model accuracy value using EfficientNet in 100 epochs. (C) shows the model accuracy value using Xception in 150 epochs. (D) is referring to the model accuracy value using in MobileNet 100 epochs

In **Figure 8**, is explaining that the model learns effectively and nearly overcomes the problem of overfitting. Accuracy reaches very high values (close to 1.0) in the final stages of training, indicating the model's strong classification performance. However, it is essential to closely monitor model performance, especially by analyzing the difference between training accuracy and validation accuracy to ensure that the model generalizes well to unseen data. Several improvements were applied to enhance the efficiency and effectiveness of the Siamese network model. Layer normalization was added after extensive experiments involving different layers representing specific categories. The output was then passed to a fully connected layer that reduces dimensions to 512 units, providing better stability during training. Cross-validation using K-fold cross-validation with K=5 for each k content on 20 epochs the total is 100 was employed to partition the training data and ensure effective evaluation. The model processes images of size 105x105 pixels and uses the Adam optimizer with a learning rate of 0.0001, a ReLU activation function for hidden layers, a batch size of 32, and trains over 30 epochs. Dropout rates were adjusted to suit different architectures, with the following values: 0.6 for ResNet-50, 0.4 for EfficientNet, 0.6 for Xception, and 0.7 for MobileNetV3. These adjustments resulted in a more robust and efficient model.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Second dataset RMFRD

RMFRD is one of the richest collections of masked faces data in the real world. The collection contains 5,000 images of 525 people wearing masks, as well as 90,000 images of people without masks representing the same 525 people. The semi-automatic annotation strategy was used to crop parts of the face that carry important information, such as the eyes, nose and mouth, helping to improve the accuracy of facial recognition systems in real-world conditions. This set of data is extremely important in recognition research Faces in imperfect environments, such as masked face recognition, is a major challenge in security and authentication applications. **Figure 9** shows samples some pairs of face images from the RMFRD dataset, which show a variety of positions, lighting and ambient conditions, reflecting the challenges facing face recognition technologies.

Figure 9. Samples of masked faces used in masked face recognition.

RMFRD dataset. The accuracy values achieved by the ResNet5 models 99%, EfficientNet 99%, Xception 99%, and MobileNet99%. demonstrating their effectiveness in correctly categorizing or predicting the target. These models showed exceptional taxonomic performance in the task assigned to him. Table 3 represents a comparison of the performance of different deep learning models. Models compared include ResNet50, EfficientNet, Xception, and MobileNet With the Siamese network to achieve the best results.

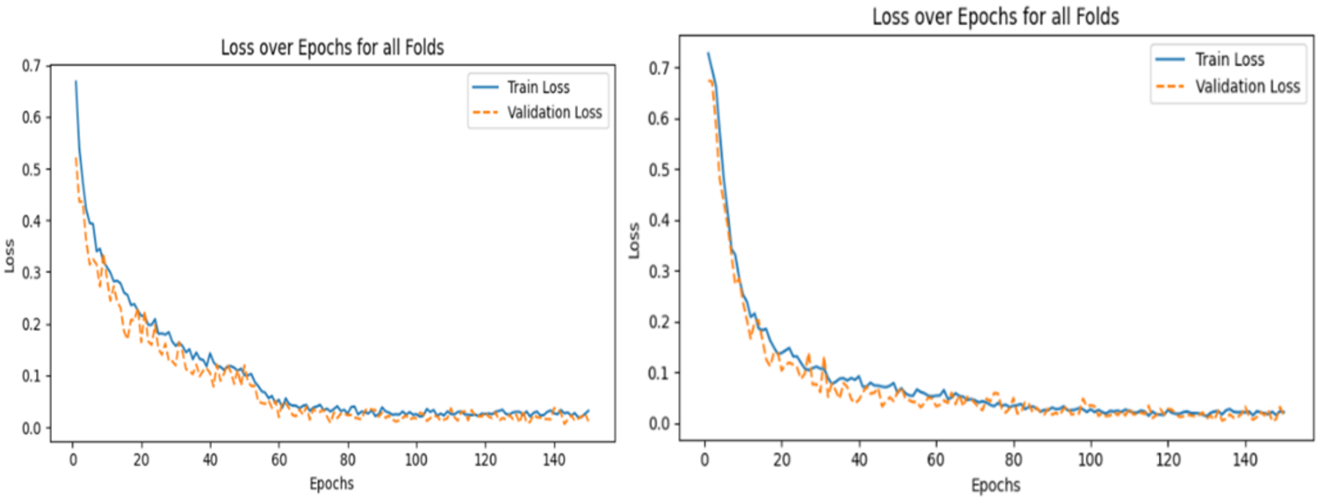
Tabel 3. Result RMFRD dataset

model	accuracy	precession	recall	F1-score	loss	MSE
Res-Net50	99%	99%	99%	99%	0.04	0.02
EfficientNet	99%	99%	99%	99%	0.03	0.01
Xception	99%	99%	99%	99%	0.02	0.007
MobileNet	98%	98%	98%	98%	0.07	0.024

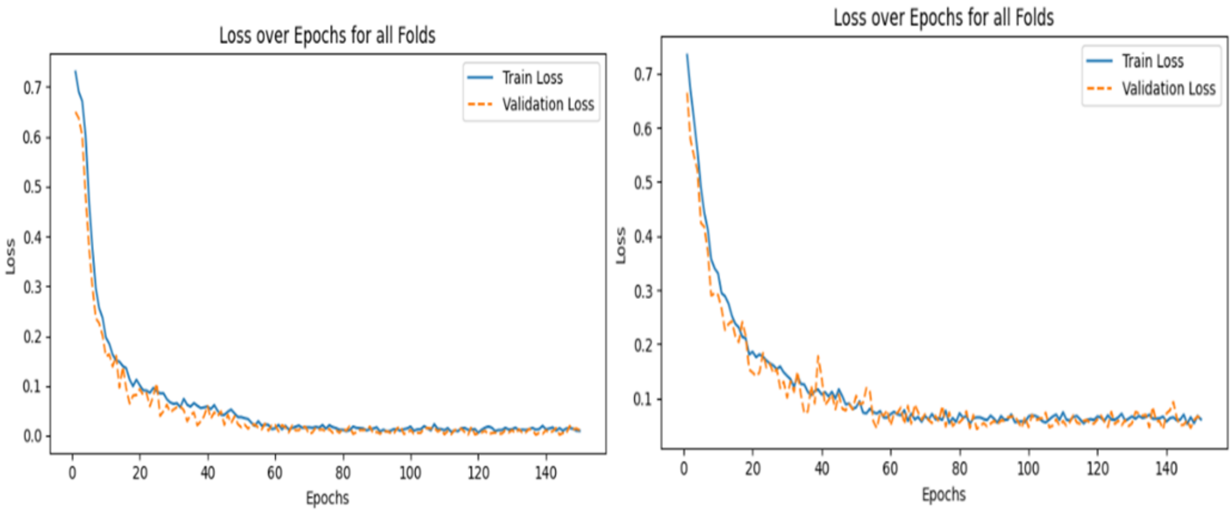
In **Figure 10**. Dual network model performance in masked face recognition using MFR2 database. The graph shows the rate of training data loss, showing a gradual decrease during the training process, as well as verification data loss that is constantly decreasing as the number of epochs progresses. This reflects an improvement in the overall performance of the model and its increased accuracy in data verification

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



(A) Loss function in ResNet-50 (B) Loss function in EfficientNet



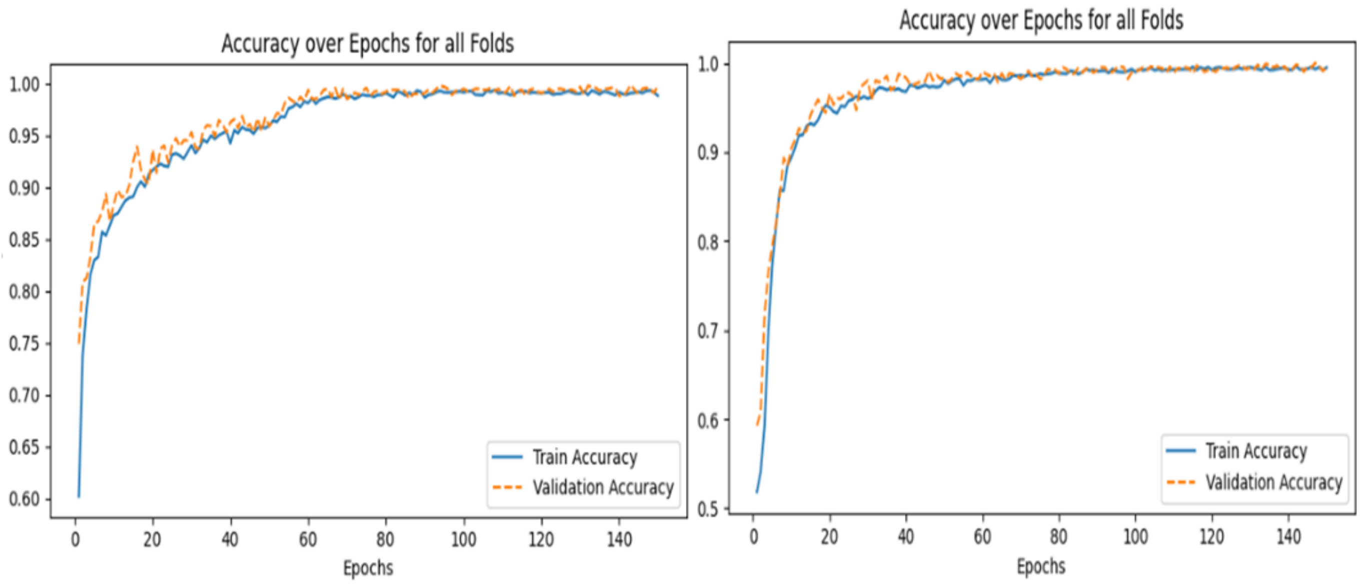
(C) Loss function in Xception (D) Loss function in MobileNet

Figure 10. (A) is showing loss function in RwsNet-50 using 150 epochs. (B) explains the loss function in EfficientNet using 150 epochs. (C) refers to the loss function in Xception using 150 epochs. (D) represents the loss function in MobileNet using 150 epochs**.**

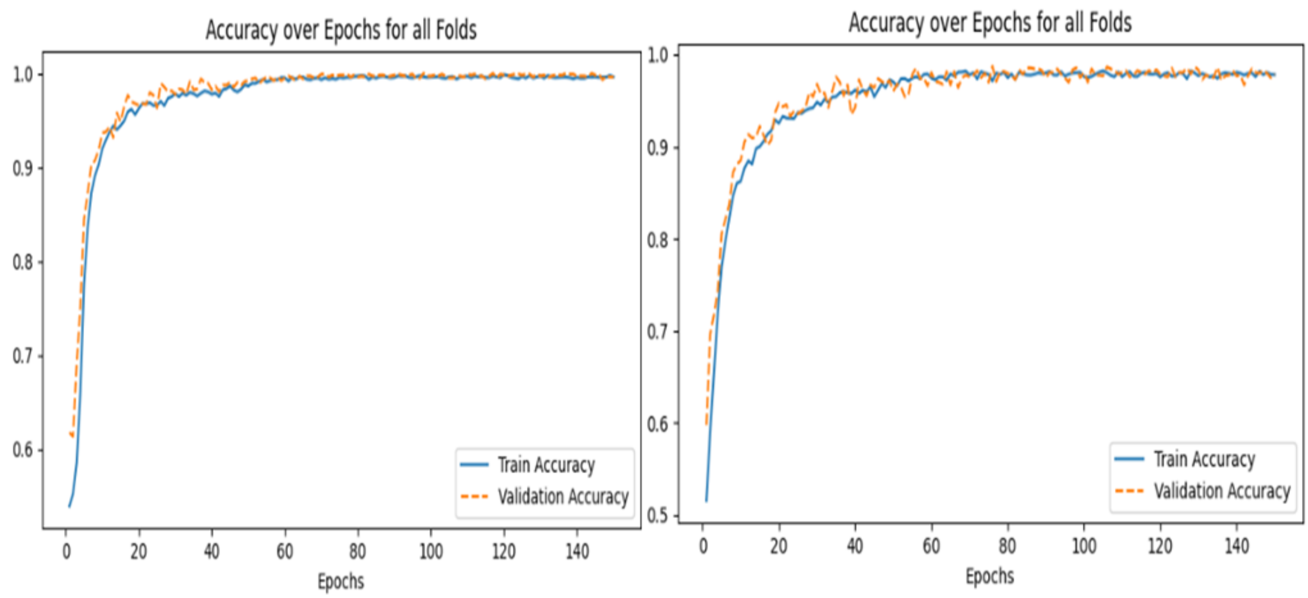
In **Figure 11** demonstrates that the model learns successfully and nearly overcomes the problem of overfitting. Accuracy reaches very high levels (almost 1.0) in the final phases of training, reflecting the model's excellent categorisation ability. However, it is critical to regularly evaluate model performance, particularly the difference between training and validation accuracy, to verify that the model generalises successfully to new data. Several enhancements were made to increase the efficiency and efficacy of the Siamese network model. Layer normalisation was implemented following extensive trials with several layers representing certain categories. The output was then passed to a fully connected layer that reduces dimensions to 512 units, providing better stability during training. Cross-validation using K-fold cross-validation with K=5 for each k content on 30 epochs the total is 150 was employed to partition the training data and ensure effective evaluation. The model processes images of size 105x105 pixels and uses the Adam optimizer with a learning rate of 0.0001, a ReLU activation function for hidden layers, a batch size of 32, and trains over 30 epochs. Dropout rates were adjusted to suit different architectures, with the following values: 0.2 for ResNet-50, 0.3 for EfficientNet, 0.4 for Xception, and 0.3 for MobileNetV3. These adjustments resulted in a more robust and efficient model.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article



(A) accuracy training and validation in ResNet-50 (B) accuracy training and validation in EfficientNet



(C) accuracy training and validation in Xception (D) accuracy training and validation in MobileNet

Figure 11. (A) is referring to the model accuracy value utilizing ResNet-50 in 150 epochs. (B) is indicating to the model accuracy value for using Efficient Net in 150 epochs. (C) indicates the model accuracy value using Xception in 150 epochs. (D) is representing model accuracy value using in MobileNet 150 epoch

COMPARISON WITH RELATED WORKS

The proposed framework shows its superiority when evaluated against different strategies and databases. The comparative procedure proved that the proposed technique outperforms many other methods, especially in identifying individuals wearing masks. **Table 4** shows the results of the comparative procedure in detail.

Journal of Information Systems Engineering and Management
2025, 10(50s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 4. Comparing result with related work

Reference	methodology	accuracy	dataset
[14]	CNN (ResNet-50, ResNet101, InceptionV3, VGG19)	98%	MFR2 Celebrity RMWMF
[21]	CNN (ResNet-50 VGG-16, AlexNet,)	91%	RMFRD SMFRD
[6]	SNN (ResNet-50)	93%	contains 20 celebrity faces
Our work	SNN (ResNet-50 EfficientNet, and MobileNetV3 Xception)	99%	RMFRD MFR2

PERFORMANCE

1. Accuracy: is a fundamental evaluative metric used in several fields, including facial recognition applications. It signifies the proportion of correct predictions to the total number of samples, formally defined as shown in Equation (1) [22].

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \quad (1)$$

2. Precision : is quantifying the ratio of correct positive identifications to the total positive matches found, formally represented as shown in Equation (2) [23]

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}} \quad (2)$$

3. Recall : include the measurement of the ratio of correctly identified positive matches to all positive matches found, as formally stated by Equation (3) [23].

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \quad (3)$$

4. F1-Score: the symphony of recall and accuracy, is an essential evaluation metric. All aspects of the face recognition model's performance, including false positives and false negatives, are evaluated by this metric. Because it provides a comprehensive assessment of model performance, the F1-Score is particularly useful for imbalanced datasets. The F1-Score is a more reliable indicator of a model's performance than accuracy, which could ignore certain types of mistakes. This is because it takes both types of mistakes into consideration. This calculation is shown in Equation (4) [22]

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision + Recall}} \quad (4)$$

5. Mean Square Error MSE

The use of Mean Squared Error is pertinent in situations requiring the detection of outliers. MSE is notably proficient at attributing increased significance to the data points [24]. Utilizing Equation (5).

$$MSE = \frac{1}{m} \sum (\mathrm{Xi} - \mathrm{Yi})^2 \quad (5)$$

CONCLUSION

Recognizing masked faces is one of the most prominent modern challenges in the field of artificial intelligence, as the importance of this field has increased significantly with the widespread use of masks in public environments as a result of the COVID-19 pandemic and other conditions that require face covering. The primary challenge is the ability of smart models to recognize individuals who use only limited features, such as eyes, eyebrows, and forehead, which may be insufficient to accurately distinguish between people. This requires the development of advanced technologies and innovative methods to cope with such complex situations. This research aims to provide effective solutions for masked face recognition by designing an advanced model that combines the Siamese neural network and the convolutional neural network (CNNs). The model is based on extracting intrinsic features from images, enabling it to recognize in high resolution even when there are obstacles such as masks. To achieve superior performance, data augmentation and model organization (regulation) techniques were used, along with the application of the cross-entropy loss function, enhancing the model's ability to distinguish between similar and different pairs. The proposed model was checked on a range of different databases to evaluate its performance in multiple conditions. In particular, the RMFRD database where the ResNet-50 model achieved 99% accuracy, reflecting its high effectiveness in recognizing masked faces. In addition, other models, EfficientNet, Xception and MobileNet, showed similar results with 99% accuracy. When testing the models on the MFR2 database, the results showed that the models achieved ResNet-50 accuracy of 99%, while other models recorded 99% EfficientNet, Xception 99% and MobileNet accuracy of 98%. The model was tested in diverse conditions such as poor lighting and different capture angles, demonstrating high performance flexibility.

In conclusion, the proposed approach addresses the challenge of accurately and reliably recognizing masked faces, especially under the circumstances imposed by the COVID-19 pandemic. By manipulating profound learning techniques as well as integrating Siamese neural networks with convolutional networks, these results confirmed the strength of the proposed model in balancing accuracy and generalization, making it suitable for use in security applications and human-computer interaction. This contributes to the challenges associated with the widespread use of face masks

REFERENCES

[1] Z. Wang and Y. Xu, "Studies advanced in face recognition technology based on deep learning," vol. 0, pp. 121–132, 2023, doi: 10.54254/2755-2721/31/20230133.

[2] S. G. H. Tang *et al.*, "Prerequisite for COVID-19 Prediction: A Review on Factors Affecting the Infection Rate," *Int. J. Environ. Res. Public Health*, vol. 19, no. 20, 2022, doi: 10.3390/ijerph192012997.

[3] B. Kocacinar, B. Tas, F. P. Akbulut, C. Catal, and D. Mishra, "A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System," *IEEE Access*, vol. 10, no. June, pp. 63496–63507, 2022, doi: 10.1109/ACCESS.2022.3182055.

[4] O. M. Assim and A. F. Mahmood, "Epileptic detection based on deep learning: A review," *Iraqi J. Electr. Electron. Eng.*, vol. 20, no. 2, pp. 115–126, 2024, doi: 10.37917/ijeee.20.2.10.

[5] A. Alzu, F. Albalas, T. Al-hadhrami, L. B. Younis, and A. Bashayreh, "Masked Face Recognition Using Deep Learning : A Review," 2021.

[6] R. Chatterjee, S. Roy, and S. Roy, "A Siamese Neural Network-Based Face Recognition from Masked Faces," *Commun. Comput. Inf. Sci.*, vol. 1534 CCIS, no. February, pp. 517–529, 2022, doi: 10.1007/978-3-030-96040-7_40.

[7] M. Eman, T. M. Mahmoud, and T. Abd-el-hafeez, "A Novel Hybrid Approach to Masked Face Recognition Using Robust PCA and GOA Optimizer," vol. 13, no. 3, pp. 25–35, 2023, doi: 10.21608/SJDFS.2023.222524.1117.

[8] M. Fatima, S. A. Ghauri, N. B. Mohammad, H. Adeel, and M. Sarfraz, "Machine Learning for Masked Face Recognition in COVID-19 Pandemic Situation," *Math. Model. Eng. Probl.*, vol. 9, no. 1, pp. 283–289, 2022, doi: 10.18280/mmep.090135.

[9] P. Poonpinij, S. Tarnpradab, N. N. Lumpoon, and N. Wattanapongsakorn, "Masked Face Recognition and Identification Using Convolutional Neural Networks," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 10, no. 1, pp. 571–576, 2023, doi: 10.1109/EECSI59885.2023.10295922.

[10] S. V. Militante and N. V. Dionisio, "Real-Time Facemask Recognition with Alarm System using Deep Learning," *2020 11th IEEE Control Syst. Grad. Res. Colloquium, ICSGRC 2020 - Proc.*, no. August, pp. 106–110, 2020, doi:

10.1109/ICSGRC49013.2020.9232610.
[11] M. Vicent, A. R. Brooks, and M. B. Gloria, "Masked and unmasked face recognition model using deep learning techniques . A case of black race Masked and unmasked face recognition model using deep learning techniques . A case of black race," no. May, 2023.
[12] R. K. Shukla and A. K. Tiwari, "Masked Face Recognition Using MobileNet V2 with Transfer Learning," no. December 2019, 2023, doi: 10.32604/csse.2023.027986.
[13] H. N. Vu, M. H. Nguyen, and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Appl. Intell.*, vol. 52, no. 5, pp. 5497–5512, 2022, doi: 10.1007/s10489-021-02728-1.
[14] M. B. Nadzeri, M. Musa, C. C. Meng, and I. M. Ismail, "Interactive Mobile Technologies," *Int. J. Interact. Mob. Technol.*, vol. 17, no. 15, pp. 135–154, 2023.
[15] L. Utkin, M. Kovalev, and E. Kasimov, "An Explanation Method for Siamese Neural Networks," *Smart Innov. Syst. Technol.*, vol. 220, pp. 219–230, 2021, doi: 10.1007/978-981-33-6632-9_19.
[16] H. Jansen, M. P. Gallee, and F. H. Schroder, "Analysis of sonographic pattern in prostatic cancer: Comparison of longitudinal and transversal transrectal ultrasound with subsequent radical prostatectomy specimens," *Eur. Urol.*, vol. 18, no. 3, pp. 174–178, 1990, doi: 10.1159/000463903.
[17] N. Serrano and A. Bellogín, "Siamese neural networks in recommendation," *Neural Comput. Appl.*, vol. 35, no. 19, pp. 13941–13953, 2023, doi: 10.1007/s00521-023-08610-0.
[18] A. Fedele, R. Guidotti, and D. Pedreschi, *Explaining Siamese networks in few-shot learning*, vol. 113, no. 10. Springer US, 2024. doi: 10.1007/s10994-024-06529-8.
[19] M. C. Tourad, A. Abdelmounaim, M. Dhleima, C. A. A. Telmoud, and M. Lachgar, "DeepSL: Deep Neural Network-based Similarity Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, pp. 1394–1401, 2024, doi: 10.14569/IJACSA.2024.01503136.
[20] M. Fisichella, "Siamese coding network and pair similarity prediction for near-duplicate image detection," *Int. J. Multimed. Inf. Retr.*, vol. 11, no. 2, pp. 159–170, 2022, doi: 10.1007/s13735-022-00233-w.
[21] W. Hariri, "Efficient masked face recognition method during the COVID-19 pandemic," *Signal, Image Video Process.*, vol. 16, no. 3, pp. 605–612, 2022, doi: 10.1007/s11760-021-02050-w.
[22] Y. Xin *et al.*, "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, vol. 6, no. May 2021, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.
[23] M. Kubat, *An Introduction to Machine Learning*. 2017. doi: 10.1007/978-3-319-63913-0.
[24] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.