Mesopotamian journal of Computer Science
Vol. 2025, **pp**. 83–91
DOI: https://doi.org/10.58496/MJCSC/2025/005 ; ISSN: 2958-6631
https://mesopotamian.press/journals/index.php/cs

Research Article

Approach for Detecting Face Morphing Attacks Using Convolution Neural Network

Essa M. Namis [1,*], Khalid Shaker [1], Sufyan Al-Janabi [1]

[1]*College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq.*

ARTICLE INFO

Article History

Received 18 Jan 2025
Revised 21 Feb 2025
Accepted 10 Mar 2025
Published 01 Apr 2025

Keywords

Face Recognition Systems (FRSs)

Convolution Neural Network (CNN)

Morphing attacks

Deep learning

Face detection

ABSTRACT

The facial morphing method combines at least two images of the face to get a singular altered facial image that exposes the vulnerabilities of face recognition systems (FRS). The extensive implementation of face recognition algorithms, particularly in Automatic Border Control (ABC) systems, has raised apprehensions over potential threats, as modified passports present significant risks to national security. In this paper, a new face morphing attack detection approach has been proposed using two different datasets (StyleGAN and AMSL) for testing and validation. A new model for face morphing attack detection based on a special Convolutional Neural Networks (CNNs) architecture has been used for classifying real and morphing images. Experimental results indicate that our suggested approach is extremely generalized and markedly robust in detecting face morphing attacks produced by different techniques .

1. INTRODUCTION

Banking, hospitality, transportation, and various other sectors have extensively employed face recognition, a significant biometric technology, for identification verification applications. The International Civil Aviation Organization (ICAO) selected the face as a biometric characteristic for electronic machine-readable travel documents (eMRTD) for easier detection and verification.This led to the Automatic Border Control (ABC) system using facial recognition technology more and more. Recently, numerous attacks have emerged on face recognition systems, with face morphing assaults posing a significant risk regarding the security of current face recognition systems (FRS) [1, 2]. Fig.1 demonstrates a situation involving the morphing of two facial images.

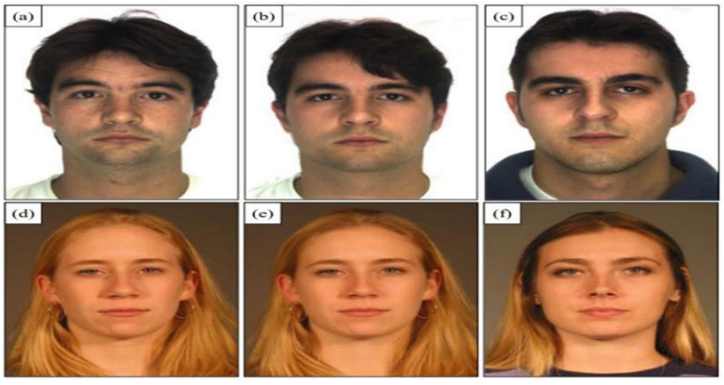

Fig. 1. shows two examples of morphing images: b gathered from subjects (a) and (c); e obtained from subjects (d) and (f) [3].

*Corresponding author. Email: ess22c1002@uoanbar.edu.iq

The morphing procedure will execute the integration process on the provided facial images from contributing contingent upon produce a singular Face Morphing Image (FMI). Because of this, the final FMI involves the facial features of all the people who provided information, showing that both automated FRS and human evaluators work well. The existing detection algorithms primarily rely on two methods: standard machine learning and deep learning [4]. In conventional machine learning methodologies, hand-crafted features denote the attributes and information extracted from the image by different techniques during the feature extraction phase prior to classification. Based on the classifier's judgment, the retrieved features distinguish between fraudulent morphing and bona fide images. In deep learning, convolutional neural network (CNN) and other approaches frequently do not require hand-crafted features, as they may autonomously learn features quickly from images, particularly with the advent of very deep CNN. As the difficulty of classification and recognition tasks escalated, the depth of deep learning architectures also rose, facilitating their management of these complexities by the integration of extra layers [5]. As the network expands, training becomes ever more challenging, and the computational expenses rise substantially.

In this work, the proposed model consists of a residual block with a stride of 1 and a non-residual block with a stride of 2. The global average pooling layer and rectified linear unit (ReLU) activation follow this structure. The network then connects to a fully connected (FC) layer containing two neurons. These two neurons correspond to predictions, one for "morphed" and the other for "real." The softmax function generates the predicted probabilities, with the final output provided by the classification layer. This research primarily contributes to the following topics:

1. A custom architectural approach using CNN has been presented for the detection of face morphing attacks.
2. Method Comparison: We have conducted comparisons with earlier related studies to determine whether the proposed CNN outperforms previous strategies in detecting face morphing attacks.

We have organized the following sections of this paper: Section 2 reviews relevant information on the detection of face morphing attacks. Section 3 outlines our suggested convolutional neural network methodologies for a face-morphing attack detection system. Subsequently, Section 4 explores the outcomes of the experiments and discussion, comparing them with recent methodologies. Finally, Section 5 offers a brief conclusion for this research.

2. RELATED WORKS

This section explores relevant studies related to this topic. Many studies have used texturing methodologies, such as binarized statistical image features, scale-invariant feature transformations (SIFT), and steerable pyramids, among others. Most research utilizes pre-existing CNN models, like VGGNet, AlexNet, and ResNet, for feature extraction. We examine some of them:

Tian Ma et al. [6] used neural networks and occlusion detection together and got excellent results. They explored pretrained network-trained approaches for improved precision, generality, resilience, and decision-making. Their methods may not be perfect, but they make the network more resilient to threats. The proposed methods made the system more resilient to attacks with accuracy up to 94%. Also, Singh et al. [7] proposed an approach to face-morphing attack production and detection, particularly in three-dimensional scenarios. Their point cloud-based 3D face-morphing model construction method. The vulnerability research uses 2D and 3D facial recognition systems to evaluate the new 3D face-morphing attacks. Human observers are studied to determine the value of 3D data in morphological detection. The results demonstrate the susceptibility of 3D face morphing models. They also automatically assessed the outstanding performance of 3D morphing models, which match actual 3D scans. To identify 3D morphing attacks utilizing pretrained point-based CNN models, they propose three 3D MAD approaches. Comprehensive research shows that 3D MAD algorithms can detect of 3D face morphing threats. Tapia et al. [8] presented and analyzed a single morphing attack detection (SMAD) technique for morphed facial images from various individuals. They created morphing face images for K = 2, 4, 8, and 16 people using facial photos from K subjects and their associates. Morphing images with more contributors conceal faces. They built AlphaNet, a multiclass CNN, using MobileNetV3. They then applied three alpha filters to RGB channels and evaluated AlphaNet within and between datasets. They also tried the approach on phony photos and got 4.41% and 4.56% BPCER10 and 20 in cross-dataset testing. Moreover, Jia et al. [9] developed a high-frequency detection framework with an advanced two-branch network design. They used the RGB stream and high-frequency information stream to detect morphing faces simultaneously, strengthening their interaction with SEM and IEM. They tested efficient and generalizable methods on the HNU and FEI datasets. The TSCNN had 0.67% ACER, 0.32% EER, and 98.93% ACC in the FEI dataset. The HNU (MDB1) dataset had 1.95% ACER, 0.88% EER, and 98.26% accuracy. Both datasets favored the two-branch network over the single-branch network. Iman et al. [10] proposed an enhanced facial feature extraction method. The proposed method includes four phases, which include the generation of morph pictures from real-life photographs using automatic selection landmarks, StyleGAN, and manual selection landmarks. We trust StyleGAN for optimum, artifact-free photos. A Faster Region The second phase employs a convolutional neural network to cut face landmarks (eyes, nose, mouth, and skin) while preserving ears, hair and image background for each database image. The third phase extracts features using PCA, eigenvalue, and eigenvector methods to create a two-dimensional matrix with one layer per method. Create a three-layer image from each image's extracted characteristics. The layers (1) represent principal component analysis, (2) eigenvalue characteristics, and (3) eigenvector features. Finally, optimize convolutional neural networks by inserting features. In the

fourth phase, the DNN and SVM second classifiers are used for classification. The DNN classifier averaged 99.02% accuracy versus SVM's 98.64%. The FRA and RFF evaluations demonstrate the proposed work's strength. This resulted in a reduction of the DNN (FAR 0.018, FRR 0.003) and SVM (FAR 0.023, FRR 0.06) error rates. The detection accuracy is improved when these ratios are below one. DNN had 95.8% accuracy, FAR 0.039, FRR 0%, and SVM 95.2% accuracy, FAR 0.047, FRR 0.98 on the AMSL dataset. Singh et al. [11] proposed an S-MAD architecture that detects face-morphing attacks utilizing multiple features, classifiers, and comparison scores at many layers. Human post-processing creates artifact-free face-morphing photographs in our new dataset. The collection includes digital, print-scan (PS-1 re-digitized by DNP and PS-2 by CANON), and compression images. They extensively evaluated the proposed approach using two evaluation methods and compared it with existing methods. In two evaluation protocols, the new strategy outperforms existing methods. Ibsen et al. [12] proposed a framework and loss function to strengthen deep learning-based facial recognition systems against morphing attacks. They made a neural network structure better that uses a certain TetraLoss function to tell the difference between the subject and morphing attack embeddings in morphed data. Results show that the suggested technique can increase state-of-the-art face recognition systems' morphing attack resistance while maintaining high performance. At FMR = 0.1%, two backbone topologies increase RIAPAR by at least 45%, making ArcFace, MagFace, and AdaFace operationally relevant. Indeed, Ramesh et al. [13] proposed a deep convolutional neural network-based morphing attack detection solution. Training and testing components of their image-morphing process were interchangeable. Due to compression and anti-forensic measures, the networks used semantic artifact-focused data. They scaled, rotated, and cropped the photos before feeding them into traditional manipulation traces. They also added noise and blur to the training and test data sets and trained three convolutional neural network architectures from pretrained networks. Our trained networks' FRR ranges from 3.5% to 16.2%, and FAR from 0.8% to 2.2%. The VGG19 The trained model has the best FRR and FAR, 3.5% and 0.8%, respectively.

Senthil et al. [14] developed a CNN-based method to protect personal data. It differentiates between a human and an AI character in the image. This restricts access to their data to approved users. Security, law enforcement, finance, education, government, and retail use facial recognition. Unauthorized access to these sections is dangerous. They assessed the proposed system based on its accuracy, precision, and sensitivity. Experiments showed that CNN-based facial recognition works well.

In our proposed approach, we employ a deep learning-based Convolutional Neural Network (CNN) to enhance efficiency and reduce computational costs. It combines residual and non-residual blocks to get more features out of the data and make face morphing attack detection more accurate.

3. THE PROPOSED APPROACH

The main objectives of our model are to enhance the accuracy and generalizability of the suggested detection model. We will utilize the CNN model to achieve these objectives. The proposed model architecture is illustrated in Fig. 2. The suggested model architecture integrates a sequence of specialized layers to effectively detect face morphing threats.

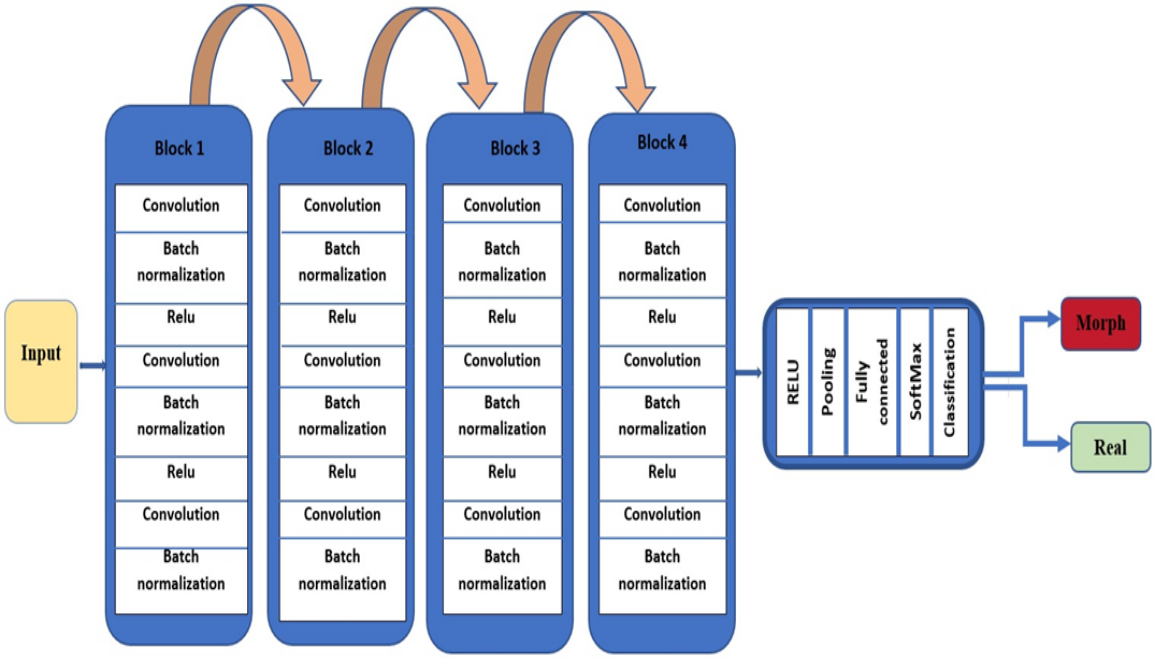

Fig. 2. The proposed model architecture

The model involves a blend of residual and non-residual blocks, enhancing the feature extraction process for face morphing detection. The residual block uses a stride of 1 to preserve spatial dimensions, whereas the non-residual block utilizes a stride of 2 for down sampling and extracting higher-level features. After the convolutional layers, the ReLU activation function adds non-linearity, which makes it easier for the model to find complex patterns in the input data.
A global average pooling layer follows the network, minimizing the spatial dimensions of the feature map to reduce computing complexity and mitigate overfitting. This pooling layer subsequently transmits the global features to a fully connected (FC) layer, which includes two neurons, one for categorizing the image as "morphed" and the other as "real." We link these neurons to a SoftMax activation function, which generates the anticipated probability for each class. The classification layer produces the final result, wherein the model categorizes the image as either "morphed" or "real". Table 1 explores the framework of the model we have suggested.

TABLE I. THE STRUCTURE OF PROPOSED MODEL

Block Name	Conv			Conv			Conv		
	size	filter	stride	size	filter	stride	size	filter	stride
Bottleneck1	1×1	96	1	3×3	96	2	1×1	24	1
Bottleneck2	1×1	144	1	3×3	144	1	1×1	24	1
Bottleneck3	1×1	144	1	3×3	144	2	1×1	32	1
Bottleneck4	1×1	192	1	3×3	192	1	1×1	32	1

This approach adeptly addresses the issues of detecting face morphing attacks by utilizing a lightweight and effective network architecture. The bottleneck layers employ residual and non-residual blocks, a feature not previously seen in similar convolutional neural network (CNN) models. They let the model pull out important features across multiple scales while keeping the efficiency of the computation.

4. DATASETS AND EXPERIMENTAL RESULTS

This section explores the selected datasets and the experimental results obtained.

4.1 Datasets

This study used two datasets to detect face morphing attacks. The AMSL dataset [15] of face morph images includes the real images of 201 people from the face research lab London set and 2000 morphed images.After we perform augmentation, the total number of real images becomes 2175, and the total number of morphed images is 2175. Fig. 3 demonstrates a sample from the AMSL dataset.

Fig. 3. Sample from AMSL dataset.

The second dataset employed was generated with the StyleGAN approach [16]. The approach produces an image that closely resembles the real one and exhibits a higher level of complexity. The collection comprises 1373 morph images derived from 1,445 real images. Fig. 4 demonstrates a sample from the Style GAN dataset.
The second dataset employed was generated with the StyleGAN approach [16]. The approach produces an image that closely resembles the real one and exhibits a higher level of complexity. The collection comprises 1373 morph images derived from 1,445 real images. Fig. 4 demonstrates a sample from the Style GAN dataset.

Fig. 4. Sample from StyleGAN dataset.

4.2 Experimental Results

We conducted experiments using MATLAB MathWorks 2024a. The CNN-based face morphing attack detector was optimized using the Adam optimizer. During training, the starting learning rate was 0.001, the maximum epochs were 30, and the mini-batch size was 128. Table 2 summarizes study parameters.

TABLE II. COMMON PARAMETERS

Parameter	Value
Optimizer	ADAM
Epoch	30
Batch size	128
Learning rate	0.001
Data split	70% training, 15% validation and 15% testing

The following metrics evaluated the efficacy of the proposed model in identifying morphing face attacks including Accuracy (ACC),F1- score, Recall, and Precision. Numerous metrics may be obtained from the confusion matrix.True positivee (TP), true negativee (TN), false positivee (FP), and false negativee (FN) are the four representations of the binary classification confusion matrix [17]. Equations (1)-(6) describe the evaluation metrics:

$$ACC = (TP+TN)/(TP+TN+FP+FN) *100 \quad (1).$$

$$Recall\backslash\backslash = TP/(TP+FN) *100 \quad (2).$$

$$Precision\backslash = TP/(TP+FP) *100 \quad (3).$$

$$F1\text{-}score = 2\ (precision * R e c a l l)/\ (precision + Recall) *100 \quad (4).$$

$$FPR = FP/(FP+TN) \quad (5).$$

$$FNR = FN/(FN+TP) \quad (6).$$

We trained the proposed model on two datasets: The StyleGAN dataset yielded results after 30 epochs of feature training. The overall accuracy rate reached 99.76%, recall 100%, precision 99.55%, F1-score 99.75%, FPR 0, and FNR 0.003. Fig. 5 illustrates the outcome of training the StyleGAN dataset and Fig. 6 illustrates the confusion matrix.

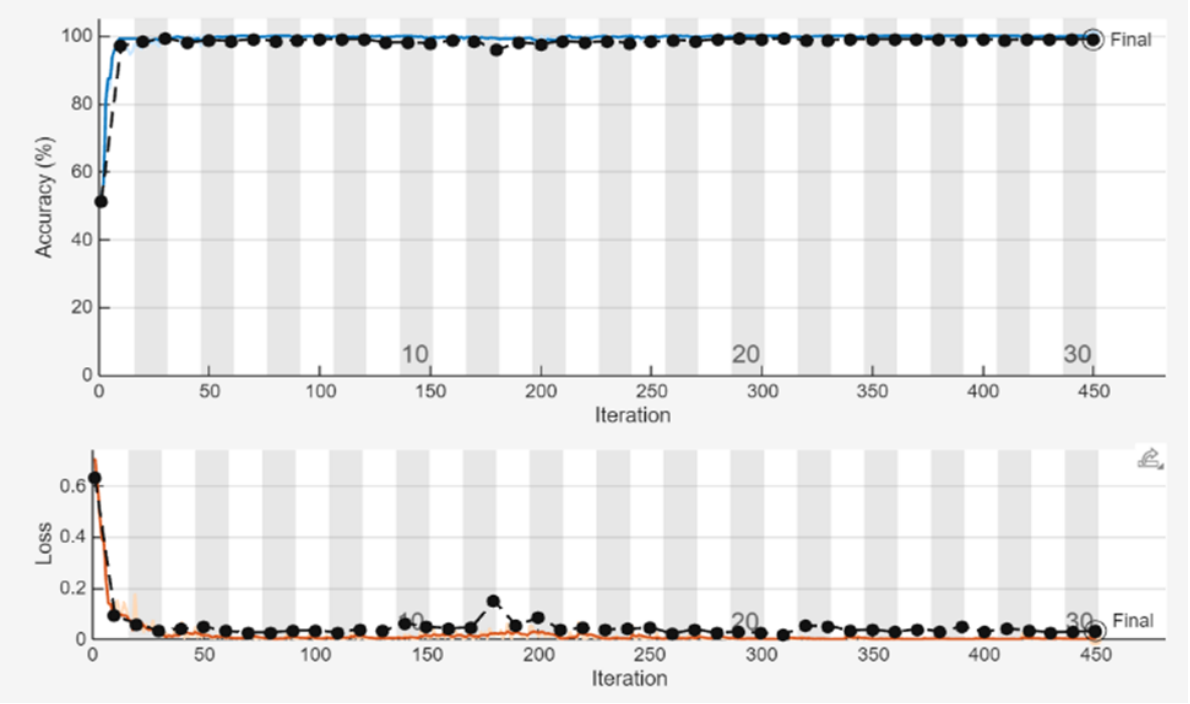

Fig. 5. Result of training styleGAN dataset.

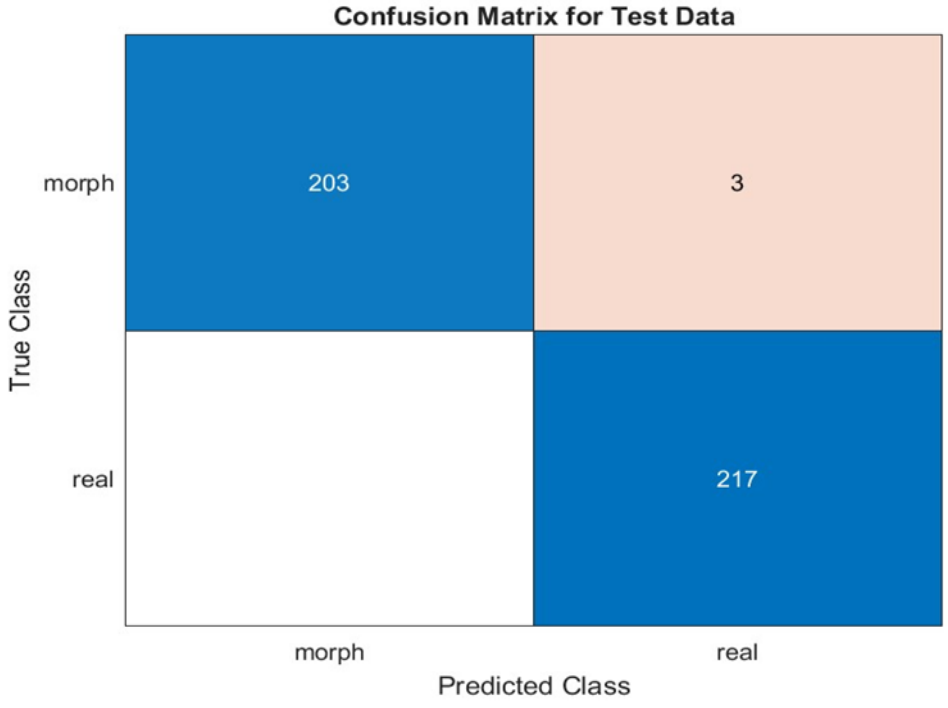

Fig. 6. Confusion matrix for styleGAN dataset.

For the AMSL dataset. We obtained the results after 30 epochs of feature training. The overall accuracy rate reached 100%, recall 100%, precision 100%, F1-score 100%, FPR 0, and FNR 0. Fig. 6 illustrates the outcome of training the AMSL dataset and Fig.7 shows the confusion matrix.Table 3 shows the results of our propoded model for both datasets.

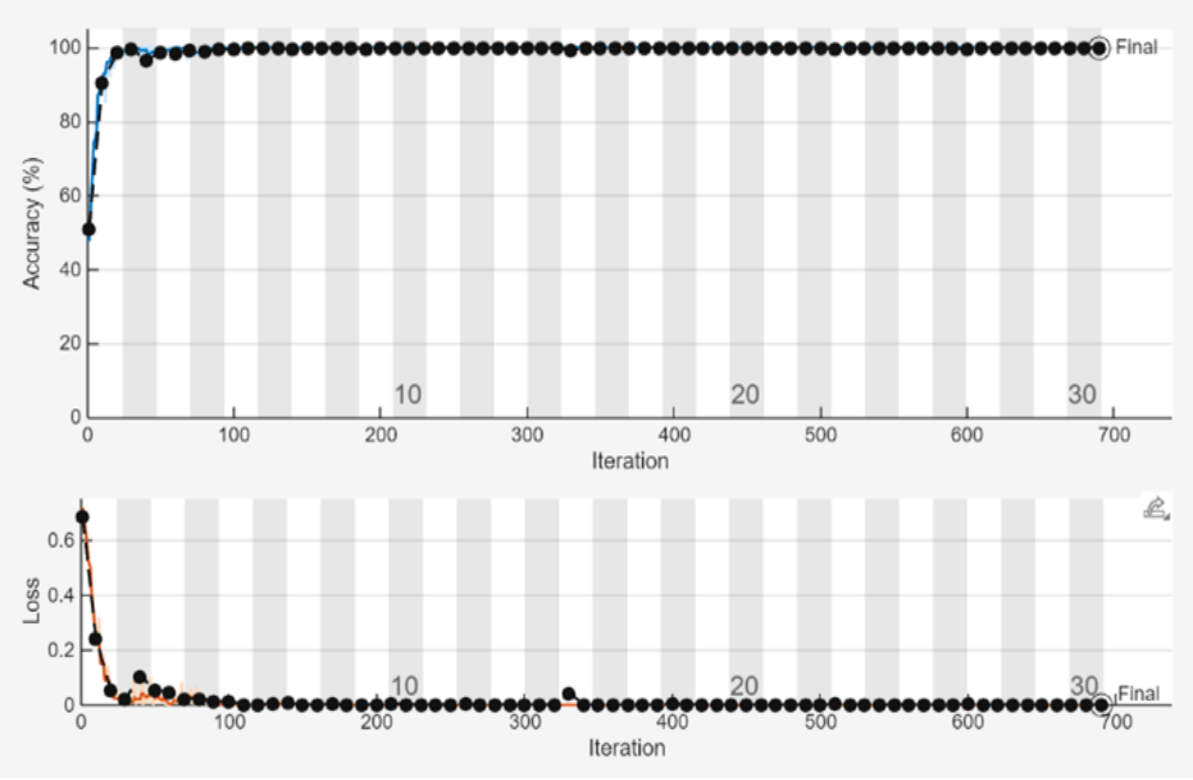

Fig. 7. Result of training AMSL dataset.

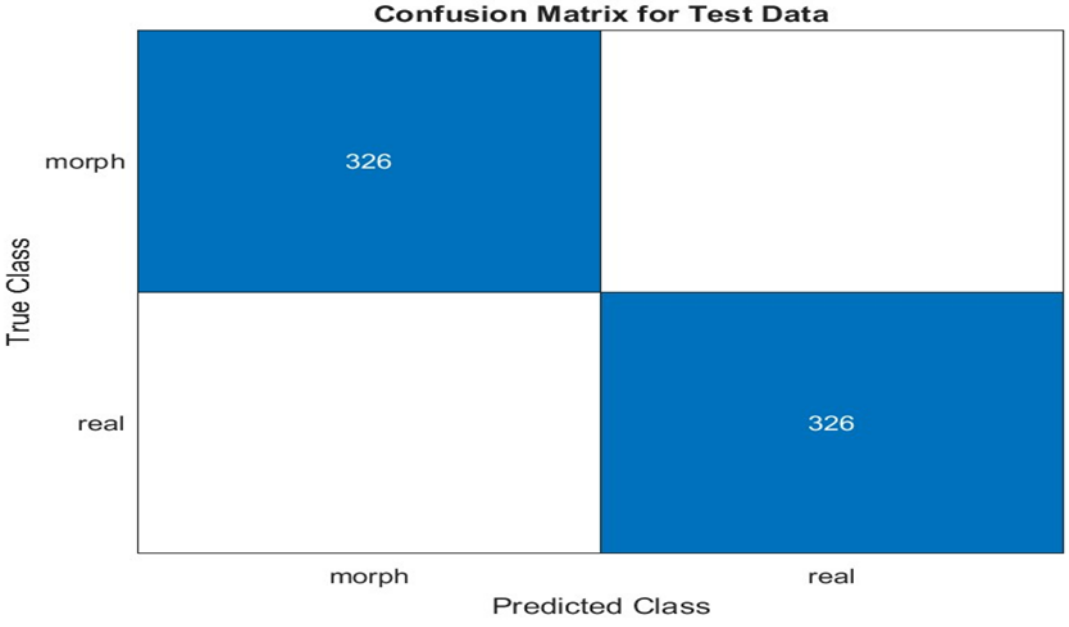

Fig.8. Confusion matrix for AMSL dataset.

TABLE III. PERFORMANCE METRICS OF OUR PROPOSED CNN MODEL.

Dataset	Accuracy	Recall	Precision	F1-score	FPR	FNR
StyleGAN	99.76	100	99.5	99.75	0	0.003
AMSL	100	100	100	100	0	0

Our proposed model effectively classified all real images in the styleGAN dataset, as measured by the resulting error amount. The model achieved an error rate of 0% but misclassified some morph images, leading to an error rate of 0.003. The AMSL dataset, on the other hand, correctly classified all morph and real images with a 0% error rate.

TABLE IV. COMPARISON OF THE PERFORMANCE OF OUR MODEL (ACCURACY) WITH OTHER MODELS OF CNN.

Model	Accuracy (%)	FPR (%)	FNR (%)
Hosny et al. [18]	94.65	0.095	0.033
Rangarajan et al. [19]	95.12	0.088	0.072
Rangarajan et al. [19]	94.20	0.230	0.041
Li et al. [20]	93.11	0.114	0.243
Li et al. [20]	95.5	0.075	0.099
Iman et al.[10] AMSL dataset	95.8 0	0.039	0
Iman et al.[9] StyleGAN dataset	98.62	0.030	0.001
Our Proposed Model AMSL dataset	100	0	0
Our Proposed Approach StyleGAN dataset	99.76	0	0.003

We compared the proposed work with a set of ready-made CNN models, as shown in Table 8. These results clearly demonstrate that our proposed approach outperformed the state-of-the-art methods in accuracy, FPR and FNR.

5. CONCLUSION

This paper introduced a convolutional neural network (CNN) model for the detection of face morphing attacks. The strength of our model lies in its ability to diagnose altered images with high accuracy and minimal error. We have selected two datasets to assess our model: AMSL and StyleGAN. The suggested model can spot face morphing attacks by using non-residual and residual blocks to extract features, ReLU activation to deal with non-linearity, and global average pooling to lower the risk of overfitting. Fully connected softmax layers categorize images as "morphed" or "real." For the AMSL dataset, the model attained an accuracy of 100%, with a false negative rate (FNR) of 0% and a false rejection rate (FRR) of 0%. For the StyleGan dataset, the model attained an accuracy of 99.70%, a false negative rate (FNR) of 0.003%, and a false rejection rate (FRR) of 0%. Our forthcoming efforts aim to enhance the system to function effectively with live photos and videos.The experimental results indicate that the CNN-based facial recognition system demonstrates superior performance.

Authors contribution

As a testament to the cooperative environment, every author made equally significant contributions. The researchers carefully designed and implemented the study framework, followed by a thorough analysis of the data and integration of their findings into a cohesive report. Their smooth cooperation and combined expertise propelled every phase of our undertaking, solidifying this effort as a genuine tribute to our shared dedication.

Funding

There was no outside funding for the research that led to the writing or publishing of this article.

Conflict Of Interest

None.

References

[1] E. M. Namis, K. S. Jasim, and S. Al-Janabi, "Face Morphing Attacks Detection Approaches: A Review," Mesopotamian Journal of Big Data, vol. 2024, pp. 82-101, 2024.

[2] E. J. Mohammed and I. T. Ahmed, "Comparative Analysis of Face Recognition Based on Multiple Feature Domains," in 2024 20th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), 2024, pp. 69-74.

[3] M. Ferrara and A. Franco, "Morph creation and vulnerability of face recognition systems to morphing," in Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks, ed: Springer International Publishing Cham, 2022, pp. 117-137.

[4] A. Agarwal and N. Ratha, "Face Morphing Detection in Social Media Content," in 2024 IEEE International Conference on Image Processing (ICIP), 2024, pp. 801-806.

[5] F. Hidayat, U. Elviani, G. B. Situmorang, M. Z. Ramadhan, F. A. Alunjati, and R. F. Sucipto, "Face Recognition for Automatic Border Control: A Systematic Literature Review," IEEE Access, 2024.
[6] T. Ma, A. Bamweyana, M. Guo, and K. Benon, "A Face Morph Detection Method Based on Convolutional Neural Networks and Occlusion Test," in 2022 7th International Conference on Image, Vision and Computing (ICIVC), 2022, pp. 158-165.
[7] J. M. Singh and R. Ramachandra, "3d face morphing attacks: Generation, vulnerability and detection," IEEE Transactions on Biometrics, Behavior, and Identity Science, 2023.
[8] J. Tapia and C. Busch, "AlphaNet: Single Morphing Attack Detection Using Multiple Contributors," in 2023 IEEE International Workshop on Information Forensics and Security (WIFS), 2023, pp. 1-6.
[9] C.-k. Jia, Y.-c. Liu, and Y.-l. Chen, "Face morphing attack detection based on high-frequency features and progressive enhancement learning," Frontiers in Neurorobotics, vol. 17, p. 1182375, 2023.
[10] I. S. Razaq, "Improved face morphing attack detection method using PCA and convolutional neural network," Karbala International Journal of Modern Science, vol. 9, p. 15, 2023.
[11] J. M. Singh, S. Venkatesh, and R. Ramachandra, "Robust Face Morphing Attack Detection Using Fusion of Multiple Features and Classification Techniques," in 2023 26th International Conference on Information Fusion (FUSION), 2023, pp. 1-8.
[12] M. Ibsen, L. J. González-Soler, C. Rathgeb, and C. Busch, "TetraLoss: Improving the Robustness of Face Recognition against Morphing Attacks," arXiv preprint arXiv:2401.11598, 2024.
[13] M. A. Ramesh, B. S. Lakshmi, D. Narendar, M. M. Najeeb, and V. Sai, "DETECTION OF FACE MORPHING USING DEEP LEARNING."
[14] S. S. Pandi, M. Monesh, and B. Lingesh, "A Novel Approach to Detect Face Fraud Detection Using Artificial Intelligence," in 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024, pp. 1-6.
[15] T. Neubert, A. Makrushin, M. Hildebrandt, C. Kraetzer, and J. Dittmann, "Extended StirTrace benchmarking of biometric and forensic qualities of morphed face images," Iet Biometrics, vol. 7, pp. 325-332, 2018.
[16] S. Price, S. Soleymani, and N. M. Nasrabadi, "Landmark enforcement and style manipulation for generative morphing," in 2022 IEEE International Joint Conference on Biometrics (IJCB), 2022, pp. 1-10.
[17] X. Liang, Z. Zhang, and R. Xu, "Multi-task deep cross-attention networks for far-field speaker verification and keyword spotting," EURASIP Journal on Audio, Speech, and Music Processing, vol. 2023, p. 28, 2023.
[18] K. M. Hosny, M. A. Kassem, and M. M. Fouad, "Classification of skin lesions into seven classes using transfer learning with AlexNet," Journal of digital imaging, vol. 33, pp. 1325-1334, 2020.
[19] A. Krishnaswamy Rangarajan and R. Purushothaman, "Disease classification in eggplant using pre-trained VGG16 and MSVM," Scientific reports, vol. 10, p. 2322, 2020.
[20] B. Li and D. Lima, "Facial expression recognition via ResNet-50," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 57-64, 2021.