Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 02 – Issue No: 01 – May 2025

Detection of Morphed Facial Images Using Convolutional Neural Networks

Tamburu Sreechandana[1], Yepuri Neha[2], Vankudothu Vaishnavi[3], Yasa Sreya[4], G. Manoj[5], K. Padmaja[6*]

[1,2,3,4,5]Department of CSE, Guru Nanak Institute of Technology, Hyderabad, Telangana, India.
[6*]Department of AIML, College Address: Mohan Babu University, Tirupati, India.
[1]srichandhanasrinivas15@gmail.com, [2]nehachoudary15805@gmail.com,
[3]vankudothuvaishnavivaishnavi@gmail.com,
[4]sreyayasa@gmail.com, [5]manoj.csegnit@gniindia.org, [6*]padmaja.k@mbu.asia
Corresponding author: [6*]padmaja.k@mbu.asia

Abstract:

Face morphing attacks are posing a threat to the contemporary biometric systems particularly in e-passports and digital identity verifications. These assaults also combine faces of several people to form synthetic images which can override face recognition models. To reduce this risk, this paper develops a lightweight Convolutional Neural Network (CNN)-based model to identify morphed facial images. This approach contains preprocessing, automated feature extraction and binary classification of bona fide and morphed faces. Evaluation of the experimental outcome of sample facial data reveals that the model proposed holds high recall rates and a reasonable overall accuracy. As shown by ROC and confusion matrices, morph attacks are strongly sensitive, but there are still false positives, and the precision can still be improved by using larger datasets and more sophisticated lightweight architectures like MobileNet. On balance, the results substantiate the efficacy of the deep learning in the reliable morph attack detection, as well as its possible usage in the real-life biometric security.

Keywords: Face morphing attack, CNN, biometric security, deep learning, fake face detection, MobileNet.

Submitted on: 31 March 2026 Accepted on: 14 May 2026 Published on: 16 May 2026

1. INTRODUCTION

Heart disease is one of the leading causes of death worldwide because it must be identified early and treated with the right prognostic tools. Given the rapid advancement of data-driven healthcare solutions, it is evident that machine learning (ML) has developed into a powerful tool that can assist physicians in the diagnosis of cardiovascular conditions. Supervised learning methods especially classification and regression models have shown great promise of predicting the occurrence and severity of heart disease. Classification algorithms are mainly designed to estimate a patient whether he has a heart disease (binary or multi-class prediction), but regression algorithms predict the severity of the disease, risk rating, or the probability. Previously, it has been demonstrated that ensemble models, SVMs, and decision-tree-based models can be more effective than more traditional statistical tools at predictive accuracy (Ali et al., 2021;

DoI: 10.63949/crinfo.v2i1.005

Katarya and Meena, 2021). Preprocessing techniques and feature selection are also crucial for improving predictive performance. Dissanayake and Md. Johar (2021) highlighted how feature selection techniques can increase the classification accuracy of datasets related to heart disease. Similarly, Hossain et al. (2023) verified that all models vary in terms of computing efficiency and model resilience while concentrating on the comparative performance study of artificial intelligence techniques.Though there are many studies aimed at the classification-based prediction of heart diseases, only a few articles combine the regression-based risk modeling in the same analysis. Thus, the paper seeks to develop an in-depth comparative evaluation of the two classification and regression algorithms to compare predictive power, computational power, and model robustness to the prediction of heart diseases.

2. REVIEW OF LITERATURE

At the beginning of the last decade, face morphing attack detection has become the focus of attention of many research works. Initial survey reviews by Hamza et al. (2022) and Venkatesh et al. (2021) have emphasized the development of morph generation methods and divided the detection methods as texture-based, feature-based, and deep learning methods. These works underlined the fact that deep learning models especially CNNs have better performances in identifying advanced morphs than the conventional handcrafted features techniques.

Seibold et al. (2020) suggested the correct and strong morph attack detection neural network designs. Their method showed that well trained CNNs are able to detect delicate blending artifacts and are able to obtain a high detection rate on datasets. On the same note, Razaq (2023) used Principal Component Analysis (PCA) with CNN to enhance feature representation and displayed better detection rates than when using either technique.

Recent studies have been more concerned with lightweight and application-specific CNN models. Namis et al. (2025) proposed a CNN-based system, which is effective in separating authentic and spoctic faces, and demonstrates encouraging results on the benchmark data. The effectiveness of CNN-based pipelines in real world use was further supported by Pokhytun et al. who implemented a neural network method of detecting altered facial images (2024).

Applying to the actual environment, Agarwal and Ratha (2024) explored the morph detection in social media content, which revealed difficulties in compression artifact and the uncontrollable quality of images. Their results show that detection models should be resistant to various real-life distortions.

Morais et al. (2025) introduced an extensive literature survey of the deep learning method in morph detection and reported that the hybrid and deep CNN models are currently the most effective, but the problem of their cross-dataset generalization still persists. Alkishri et al. (2023) investigated the idea of fake face detection based on color texture analysis with deep CNNs and showed that color-space features could also contribute to the increase in the detection accuracy.

DoI: 10.63949/crinfo.v2i1.005

The recent conference publication by Kadiri et al. (2024) used several machine learning algorithms to perform morphed image recognition and proved the effectiveness of deep learning models compared to standard ones. Hakradhar et al. (2025) made another contribution to the field by introducing a CNN-based morph detection system in the ICT4SD environment with the focus on enhanced robustness and practical deployment aspects.

In general, it is evident in the literature that there is a tendency to move away with the handcrafted feature methods to deep learning-based solutions. Even with the tremendous advancements, the issues of cross-dataset generalization, resistance to post-processing, and real-time deployment are problems in research. It is such gaps that encourage further studies on CNN-based morph detecting systems.

3. RESEARCH GAP

- **Extremely poor cross-dataset generalization:** It is not only because the most CNN models can be applied successfully to certain datasets and fail to be applied successfully to unseen data.Post processing sensitivity Compression, image resizing and social media filters decrease accuracy of detection.
- **High computation cost:**Existing deep models are too costly to run on real time or edge computation machines.
- **Poor working conditions with complex morphs:** patent comparison against the state-of-the-art GAN-based morphing attacks.
- **Absence of explainability:** The majority of the models are black boxes and it reduces the confidence of the security applications.
- **Expansions of datasets:** Datasets are constrained in most of the studies or not diverse, hence this is constraining the strength.
- **Not well real world sensitive:** No real world testing of images which are not controlled is done adequately.
- **sparse literature of hybrid methodology:** There is not a lot of literature of the application of CNNs and multimodal/frequency-based features.

4. OBJECTIVES OF THE STUDY

- To come up with an effective Convolutional neural network (CNN)-based system, which possesses an efficient morphed face image detector.
- It should be recollected that the work of the detecting in the post processed image of compression, resizing and noise, should be improved to the maximum.
- To extend the cross-dataset generalization of the proposed model. This has been determined to be used in the development of a light weight model that could be utilized in real time/ edge devices.

DoI: 10.63949/crinfo.v2i1.005

- To evaluate the suggested methodology on the traditional metrics of the performance: accuracy, precision, recall and F1-score.

5.PROPOSED METHODOLOGY

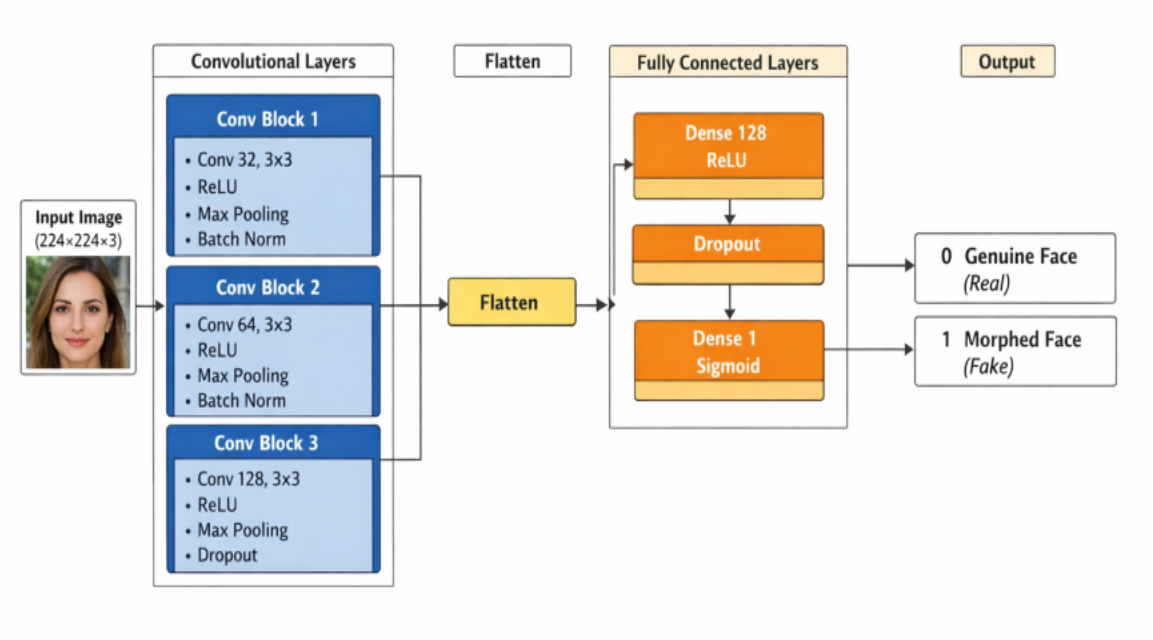

Figure 1. System Architecture

The proposed morphed facial images detecting system will use the lightweight Convolutional Neural Network (CNN) and will have a sequence of steps as follows:

Collection of Data: Original and distorted face pictures are acquired through the use of the public datasets that have variety in terms of demographics, light, and quality. The data is separated into training, validation and testing sets.

Data Preprocessing: Faces are oriented and recognized and resized (e.g. 224 x 224). To improve the generalization, the images are augmented and normalized (flipping, rotation, compression, noise).

CNN Feature Extraction: CNN: This is merely a lightweight CNN that is automatically trained to extract discriminative spatial features and it operates on convolutional layers that are initialized with ReLU activation, batch normalization and dropout to curb overfitting.

DoI: 10.63949/crinfo.v2i1.005

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 02 – Issue No: 01 – May 2025

Model Training: The network is trained with labels with binary cross-entropy loss and Adam optimizer. These are the validation monitoring and early stopping.

Morph Classification: The images are classified as genuine and morphed image usingmorph classifier consisting of full connected layers with a sigmoid/softmax output.

Model Assessment: The performance of a model is measured by Accuracy, Precision, Recall, F1-score, and ROC-AUC and assessments of robustness and cross-dataset testing are also done.

Explainability Analysis: Grad-CAM visualization can be used to compute the visualization of components of the face that influence the decision made by the model to increase transparency.

Deployment Consideration: It is pruning/quantizing which is optimization of the model is intended to be deployed in real-time and is testable on low-resource or edge devices.

7. RESULTS AND DISCUSSIONS

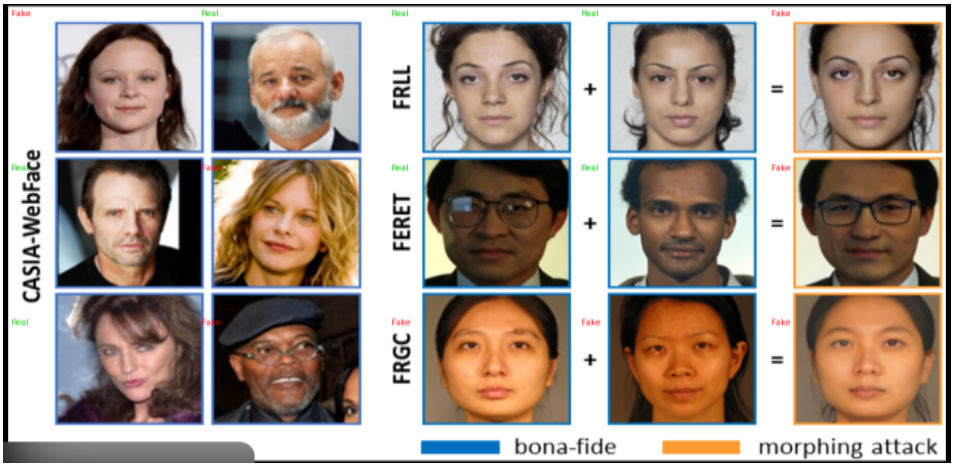

Figure 1. Classification of Real(bona-fide) and Fake (morphing attack)

The model divides the image into regions and classifies all of them into **Real(bona-fide)** and **Fake (morphing attack)** ones.

The labels in the image that can be expected are the following ones:

DoI: 10.63949/crinfo.v2i1.005

- Green → Real
- Red → Fake

As expected:

- Left-side original faces are greatly predicted Real.
- Fake images of right-side morphing attacks are the ones that are predicted.

ROC curve

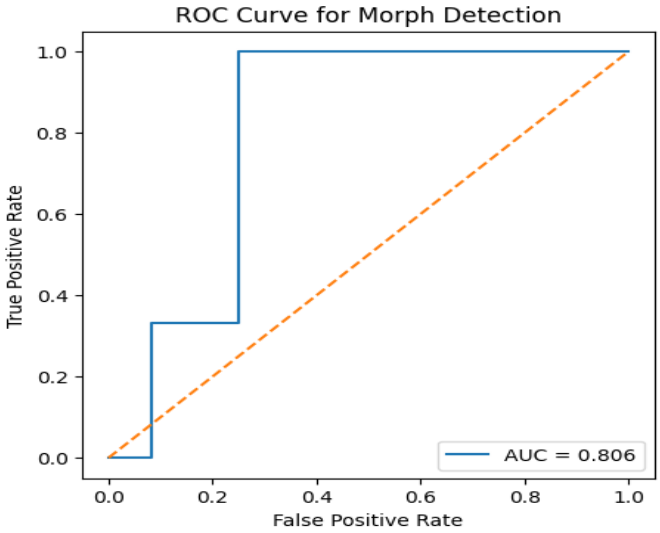

Figure 2. ROC Curve for Morph Detection

- ROC curve is the trade off entity between **the true positive rate(TPR)**and **the false positives rate(FPR).**
- The indication of the discrimination capability of the model is the figure of the calculated **AUC (Area Under Curve).**
- The higher the morph detecting curve is on the left hand side, the better.
- This demonstrates that the CNN-type classifier can distinguish the **bona-fide** and the **morphed faces.**

Table1. Performance metrics for the CNN-based morph detection

Metric	Value
Accuracy	0.667
Precision	0.375
Recall	1

DoI: 10.63949/crinfo.v2i1.005

- **Accuracy (66.7)** → In general, correct predictions are mediocre.
- **Precision (37.5)** → There were false images that were pronounced as real (there were false positives).
- **Recall (100)** →100% of the morph (fake) images were identified.

Confusion Matrix- Morph Detection:

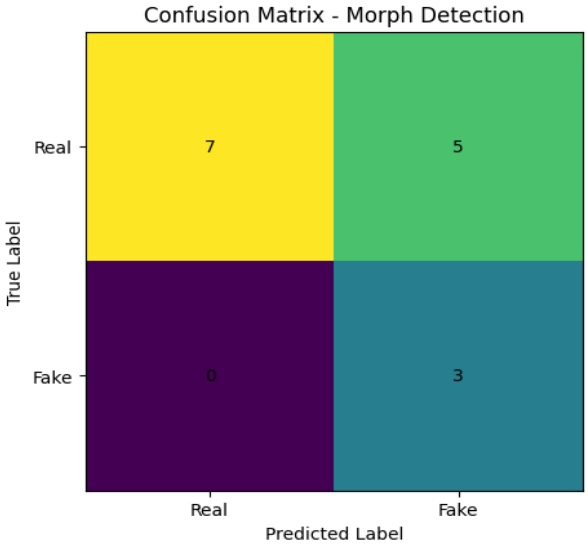

Figure 3. Confusion Matrix – Morph Detection

- **True Positives (TP):** Morph images which were recognized.
- **True Negatives (TN):** Real images which are rightfully recognized.
- **False Positives (FP) :**False images, which have been classified as real.
- **False Negatives (FN):** Morphs that were not taken into consideration by the model.

8. CONCLUSION

This analysis was done using the CNN-based technique to detect face morph attacks. The findings of the experiment confirm the fact that it is possible to distinguish between bona fide and morphed faces with the help of deep learning methods. High recall was achieved in the model that could identify all morph attack samples meaning that it is suitable in security critical system where false detection of an attack can be catastrophic.

Despite the overall medium level of accuracydue to false positive prediction, the results prove that the convolutional neural networks can be applied in the morph detection problems. Additional optimization,

DoI: 10.63949/crinfo.v2i1.005

larger datasets, and effective lightweight architectures can be used to develop the proposed framework into an effective and powerful solution to biometric security systems.

Altogether, the provided work is one of the results of the growing body of research on face morphing detection with the help of deep-learning-based approaches, and it can lead to further improvements to provide stable and trustworthy identity verification systems.

REFERENCES

[1] Alkishri, W., Widyarto, S., Yousif, J. H., & Al-Bahri, M. (2023). Fake face detection based on colour textual analysis using deep convolutional neural network. Journal of Internet Services and Information Security, 13(3), 143–155.

[2] Agarwal, A., &Ratha, N. (2024). Face morphing detection in social media content. In 2024 IEEE International Conference on Image Processing (ICIP) (pp. 801–806). https://doi.org/10.1109/ICIP51287.2024.10648209

[3] Hamza, M., Tehsin, S., Humayun, M., Almufareh, M. F., &Alfayad, M. (2022). A comprehensive review of face morph generation and detection of fraudulent identities. Applied Sciences, 12(24), 12545. https://doi.org/10.3390/app122412545

[4] Hakradhar, K., Tharun, K. T., Reddy, P. S. K., Anvitha, S. S., &Thangam, S. (2025). Morphed face detection. In S. Fong, N. Dey, & A. Joshi (Eds.), ICT analysis and applications (ICT4SD 2024) (Lecture Notes in Networks and Systems, Vol. 1161). Springer. https://doi.org/10.1007/978-981-97-8602-2_46

[5] Kadiri, P., Anusha, P., Prabhu, M., Asuncion, R., Pavan, V. S., &Suman, J. V. (2024). Morphed picture recognition using machine learning algorithms. In Second International Conference on Advances in Information Technology (ICAIT) (pp. 1–6). https://doi.org/10.1109/ICAIT61638.2024.10690845

[6] Morais, P., Domingues, I., & Bernardino, J. (2025). Deep learning techniques for detecting morphed face images: A literature review. IEEE Access, 13, 105952–105981. https://doi.org/10.1109/ACCESS.2025.3578199

[7] Namis, E. M., Shaker, K., & Al-Janabi, S. (2025). Approach for detecting face morphing attacks using convolution neural network. Mesopotamian Journal of Computer Science, 2025, 83–91. https://doi.org/10.58496/MJCSC/2025/005

[8] Pokhytun, A., Mazurets, O., Molchanova, M., &Tyschenko, O. (2024). Method for neural network detecting changed images of people's faces using CNN. https://elar.khmnu.edu.ua/handle/123456789/16938

[9] Razaq, I. S. (2023). Improved face morphing attack detection method using PCA and convolutional neural network. Karbala International Journal of Modern Science, 9(2), 15. https://doi.org/10.33640/2405-609X.3298

[10] Seibold, C., Samek, W., Hilsmann, A., &Eisert, P. (2020). Accurate and robust neural networks for face morphing attack detection. Journal of Information Security and Applications, 53, 102526. https://doi.org/10.1016/j.jisa.2020.102526

DoI: 10.63949/crinfo.v2i1.005

[11] Venkatesh, S., Ramachandra, R., Raja, K., & Busch, C. (2021). Face morphing attack generation and detection: A comprehensive survey. IEEE Transactions on Technology and Society, 2(3), 128–145. https://doi.org/10.1109/TTS.2021.3066254

DoI: 10.63949/crinfo.v2i1.005