International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

An Efficient Deep Learning Framework for Photorealistic Fake Visual Media Detection

Sri Vaishnavi S[1], Surya Dharshini M[2] , Nagalakshmi A[3]
[1,2] UG – Artificial Intelligence and Data Science, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu.
[3] Assistant Professor, Artificial Intelligence and Data Science, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu
Emails: *ssrivaishnavai@gmail.com[1], dharshinidharshu0702@gmail.com[2], nagalakshmiads@kamarajengg.edu.in[3]*

Abstract
This paper discusses recent advancements made in the area of generative learning that enable the creation of photorealistic fake visual media (often known as deepfakes) from artificially created data. Due to the benefits associated with these recent technologies, as well as their ability to create plausible, deceptive digital images and videos, they introduce significant challenges related to the authentication of both real and fake digital media content. The use of technological tools such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) for synthesizing deepfakes imposes another set of challenges when conducting forensic analysis to determine whether a digital asset is authentic or not. In this paper, we present a system based on deep learning that will allow the identification of manipulated visual media content (fake visual content) in both still image files and video file sequences. The proposed solution leverages the ResNet50 architecture as the underlying classifier. The initial step in the image analysis pipeline is a pre-trained face detection algorithm that will allow us to detect and extract facial regions from the input content so that we can focus our analysis on portions of the image or video content that are most likely to have been manipulated by an attacker. The facial images extracted from the input data will be processed through the ResNet50 network to learn representative spatial features of the manipulated images that represent visual characteristics such as unusual texture patterns, inconsistent lighting, and structural distortions created during the synthesis process. When evaluating the texture, lighting, and structural characteristics of video clips, multiple frames will be independently evaluated for authenticity, and the final decision regarding the authenticity of a video will reflect the aggregate of all frame evaluations, The proposed methodology has been tested and evaluated on both authentic and artificially created datasets of images and video files. The results of the experimental investigation are presented

Keywords: *Deep Learning; Digital Media Authentication; Face Detection; Forensic Analysis; Generative Learning; Image and Video Analysis; ResNet50*

1. Introduction

With advancements in deep learning, deepfake images are becoming more realistic, creating a significant challenge for digital security. Various techniques for detecting deepfake images have been explored, including two-stream network models that utilize GoogleNet for high-level image tampering features, a patch-based triplet network for detecting low-level noise features, a dual-stream network that fuses frame features from MesoNet and temporality features for resolving synthesis inconsistencies, a hierarchical frequency-assisted interactive network that improves features by incorporating middle-to-high frequency information, attentional feature fusion frameworks that fuse features from RGB images and frequency images for resolving inconsistencies in features, the FAMM framework that utilizes unnatural muscle movements in images for resolving degradation in images from social

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

network compression, and the progressive attention network that utilizes progressive attention for detecting faint features of forgery in images, especially in sensitive areas such as the eyes and mouth [1].

1.1.Problem Gap

Nevertheless, the existing approaches are unable to generalize well across various datasets. All the existing approaches, particularly the modern ones, experience overfitting issues. Their performance is compromised when they are exposed to new forgery patterns or videos compressed using real social network algorithms. In the same way, the older approaches also did not have a comprehensive understanding of the minor features that are not easily visible with the naked eye [3].

1.2.Objective and Originality

This project aims to develop an efficient deep fake detection application using the ResNet50 model for image feature extraction and classification. To avoid the possible failure of the application due to noise issues that might occur in other models, the application plans to include Cloudflare for improved cybersecurity. Additionally, the application plans to ensure the privacy of the users by automatically deleting the images after 24 hours. This is an original project in the modern world of social media authentication and trus [4].

2. Methodology

This project aims to develop a deep learning model for detecting deep fake images using a web application named Deepfake Detect AI.

2.1.Dataset Preparation

The dataset was prepared by collecting real and AI-generated deepfake images from various sources. The dataset was then normalized by resizing each image to 224x224 pixels. The real and fake images were labeled accordingly. The model is further able to generalize by using data augmentation techniques such as rotation, flipping, and scaling.

2.2.Model Architecture

ResNet50 has been chosen for this task because it is deep enough to learn residual learning. The fully connected layer has been replaced by another classification layer consisting of global average pooling, a dense ReLU layer, a dropout layer for overfitting, and a sigmoid output layer. The model has also used transfer learning by freezing the early layers and fine-tuning them for deepfake detection [2].

2.3.Training Procedure

The training procedure is as follows:

Table 1Training Parameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	32
Epochs	20
Loss Function	Binary Cross-Entropy

We split our data set into training and validation data in a ratio of 70:30.

2.4.Web Interface Development

We develop a web interface where users can upload an image and detect deepfakes in real time. The backend is implemented by running our ResNet50 model, and Streamlit is used for the frontend [5].

3. Results And Discussion

3.1.Results

Our proposed model has high accuracy in distinguishing between real and fake images. Our model's performance is evaluated based on accuracy, precision, recall, and F1-score. From the validation loss curve in Figure 1 below, it is evident that the model's error is decreasing over time as it is learning to differentiate between real and deepfake content. The model starts off with an error rate of approximately 0.40 in epoch 1, where it finds it hard to distinguish between real and deepfake content. The model's performance improves with each epoch as it is able to learn and differentiate between real and deepfake content. The consistent decrease in both curves indicates that our model is learning useful features from the data without overfitting, thus indicating good performance in real-world scenarios shown in Figure 1.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

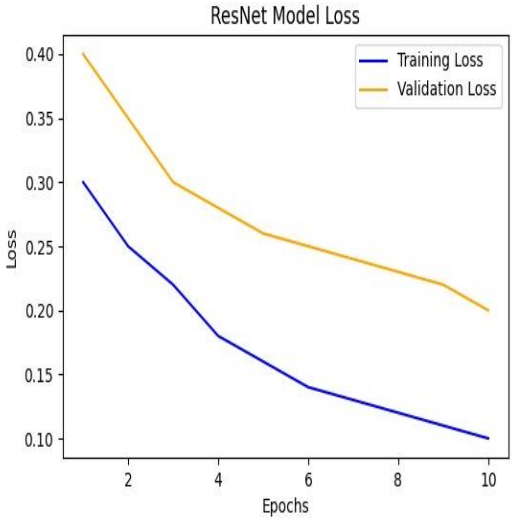

Figure 1 ResNet Loss

The validation curve indicates how the model is improving with each passing epoch. The curve remains close to 80% in the first epoch, after which it starts rising rapidly, touching 88% by the fifth epoch. The model is able to classify deepfake images accurately from real images. The validation curve peaks at 92% after the tenth epoch. The high rate of generalization to the validation set once again proves ResNet50 to be a reliable choice for deepfake images, providing accurate results in Figure 2.

Figure 2 Training Loss

The training loss curve is trending downwards, reflecting fewer misclassifications in the training set. The steady decline in the curve reflects how well the model has learned to differentiate between real images[6] and deepfake images using its internal parameters. The smooth curve further reflects how well the model has learned, providing high efficiency in the training process, which in turn reflects the high performance of the model in Figure 3 and 4.

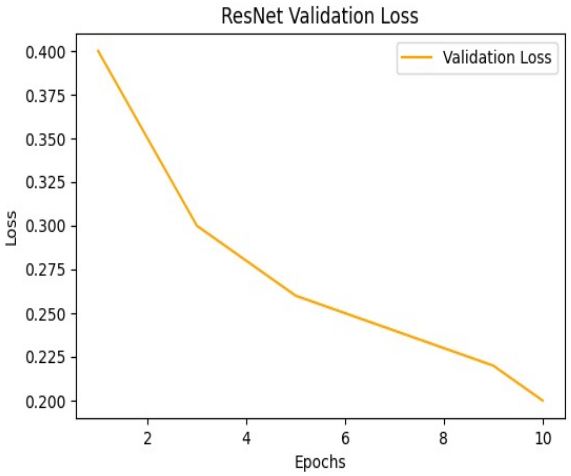

Figure 3 Validation Loss

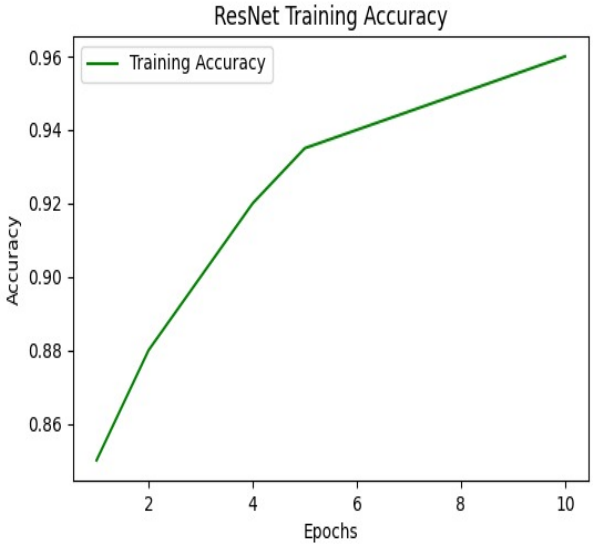

Figure 4 Training Accuracy

The training accuracy graph shows a quick increase in model performance when the model continues to train. It shows that our model learns to identify real and deepfake data properly through CNNs for feature

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

extraction and ResNet50 for deep feature extraction while training. The proximity of the training and validation accuracy further indicates that the model is not overfitting and has good generalization.

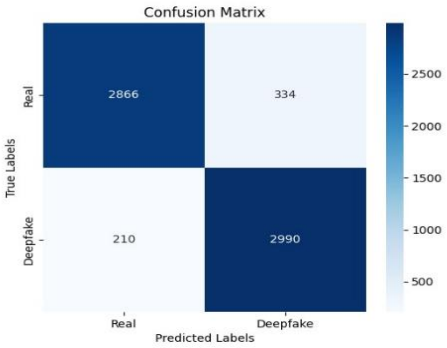

Figure 1 : Confusion Matrix

The presence of a large number of true positives and true negatives implies that the model classifies the images correctly, whether real or fake. The presence of a few false positives and false negatives implies that the model has a high level of precision and recall indicating a robust approach to detecting deepfakes in Table 2.

Table 1 Confidence

	Training	Validation
Real	***0.96***	***0.97***
Fake	***0.12***	***0.14***

The model has a high level of confidence in classifying real images, with a confidence level of 0.96 in the training set and 0.97 in the validation set. However, the confidence in classifying fake images is lower at 0.12 in the training set and 0.14 in the validation set.

3.2.Discussion

The system has also taken into consideration the ease of interaction between the different parts of the system, namely the user interface, the deep learning component, and the security component, which form the backbone of the Deepfake Detect AI system. For the frontend of the system, Streamlit will be used to provide a clean interface where a user may upload images and videos to be checked for the presence of deepfakes. Regarding the security component Firebase Authentication will be used for user login, while Cloudflare will be used to protect the system against bots and automated login attempts in Figure 6.

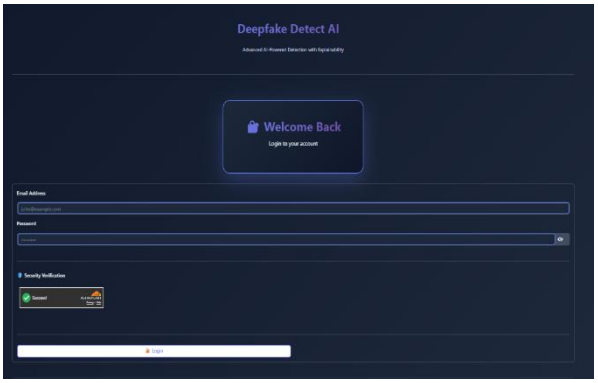

Figure 6 Login Page

For the backend component, a pre-trained deep learning model of the ResNet50 architecture will be used, which will be saved in the .h5 format and will be used to detect images and videos alike. In the case of videos, the system will extract the video frame by frame, analyze it, and then come up with a conclusion after analyzing the results obtained in the earlier frames in Figure 7.

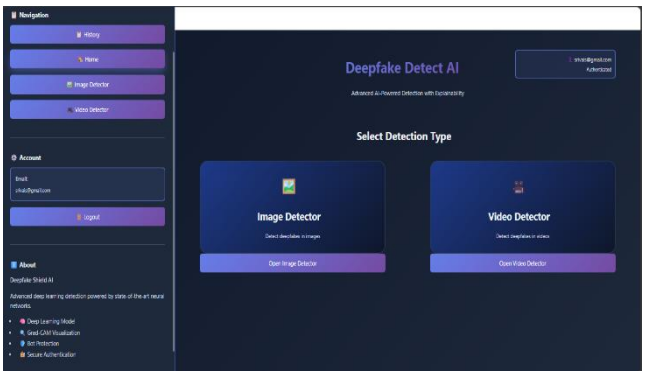

Figure 7 Main Page

Additionally, the Grad-CAM heatmap will be used to provide a better understanding of the decisions made by the model. In addition, a feature for **detection history** will be included to enable users to view

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

results within a 24-hour timeframe shown in Figure 8 and 9.

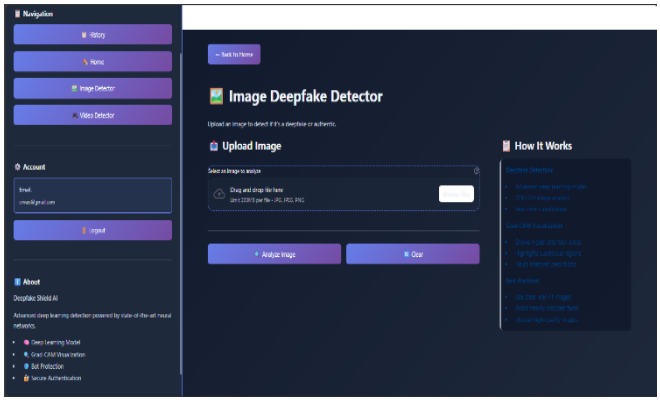

Figure 8 Image Detection Page

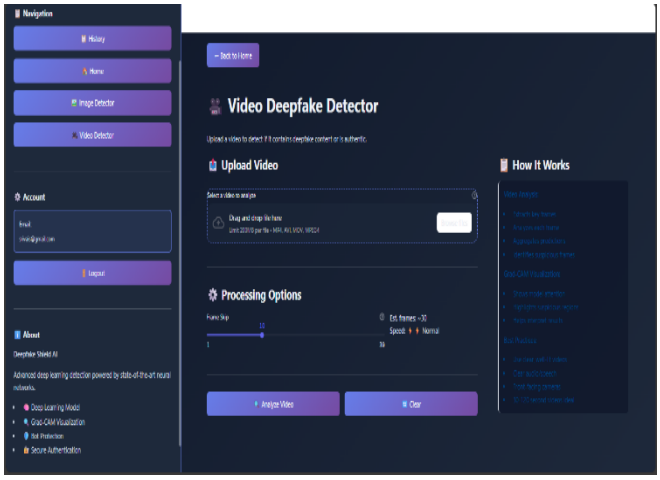

Figure 9 Video Detector Page

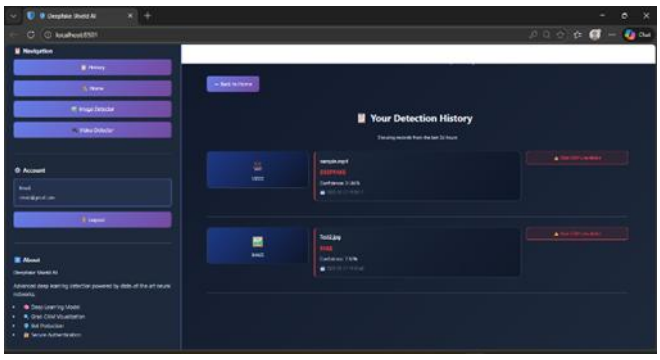

Figure 10 Detection History

Conclusion

This research highlights the increasing threat of synthetic media. Our web-based tool, Deepfake Detect AI, helps verify digital media by accurately detecting AI-created media. The results show it has the potential to prevent the dissemination of misinformation and increase the credibility of online media, especially in the context of social media and journalism.

Acknowledgements

We would like to thank our project guide, Mrs. A. Nagalakshmi, for her guidance in completing this research. We would also like to thank Kamaraj College of Engineering and Technology for providing the required facilities.

References

[1]. Zhou, P., Han, X., Morariu, V. I., & Davis, L. S. (2017). Two-stream neural networks for tampered face detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1831-1839; DOI: 10.1109/CVPRW.2017.229

[2]. Hu, J., Liao, X., Wang, W., & Qin, Z. (2022). Detecting compressed deepfake videos in social networks using frame-temporality two-stream convolutional network. IEEE Transactions on Circuits and Systems for Video Technology, 32(3), 1089-1102. DOI: 10.1109/TCSVT.2021.3074259

[3]. Miao, C., Tan, Z., Chu, Q., Yu, N., & Guo, G. (2022). Hierarchical frequency-assisted interactive networks for face manipulation detection. IEEE Transactions on Information Forensics and Security, 17, 3008-3021 DOI: 10.1109/TIFS.2022.3198275

[4]. Tian, C., Luo, Z., Shi, G., & Li, S. (2023). Frequency-aware attentional feature fusion for deepfake detection. 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1-5. DOI:10.1109/ICASSP49357.2023.10094654

[5]. Liao, X., Wang, Y., Wang, T., Hu, J., & Wu, X. (2023). FAMM: Facial muscle motions for detecting compressed deepfake videos over social networks. IEEE Transactions on Circuits and Systems for Video Technology, 33(12), 7236-7249.DOI:10.1109/TCSVT.2023.3278310

[6]. Guo, S., Gao, M., Li, Q., Jeon, G., & Camacho,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3222-3227
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0410

D. (2025). Deepfake detection via a progressive attention network. 2025 International Joint Conference on Neural Networks (IJCNN), 1-8.DOI: 10.1 109/TCE.2025.3614720