IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 4 (2025)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 4, 2025

Research Paper

FACE SWAPPING & DEEP FAKE DETECTION

[1] Devarakonda Lakshmi Tirupathamma, [2]Chelli Jhansi, [3]Gudisa Yaswanth, [4]Mr. K. Ramesh Babu

[1,2,3]U. G Student, Dept COMPUTER SCIENCE AND ENGINEERING, St.Ann's College Of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh– 523187, India

[4]Assistant professor, COMPUTER SCIENCE AND ENGINEERING, St.Ann's College Of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh– 523187, India

ABSTRACT

The project entails the development of an AI-based Face Swapping & Deepfake Detection System that does two things, creates face-swapped outputs, and detects altered content very well. The face-swapping component employs deep learning-based models to capture facial characteristics of a source image and inlay them with desired level of realism onto the corresponding target facial region keeping the expressions and lighting conditions unaltered. The deepfake detection module employs transfers learning with pre trained models like Res Net or GANs trained on large datasets of authentic and fake images.

Keywords: *Artificial intelligence, Deepfake Detection, Face swapping, Machine learning, Deep learning, MongoDB.*

Received: 16-10-2025 Accepted: 26-11-2025 Published: 05-12-2025

I. INTRODUCTION

Deep learning, generative adversarial networks (GANs), and high-quality facial synthesis techniques have all contributed to the development of face swapping and deepfake technologies, which have become important advancements in the field of computer vision. Face swapping is the technique of substituting another person's face for another in a picture or video while maintaining realistic head positions, facial expressions, and lighting. Face swaps are among the most common and socially significant types of artificial intelligence- generated synthetic media, which are collectively referred to as "deepfakes." These technologies were first created for creative and entertaining objectives, but they have since progressed to provide incredibly convincing results, which has led to worries about their misuse in identity fraud, disinformation, and reputational damage. An "arms race" between techniques for detecting and creating fraudulent content has resulted from the quick advancement of generative models. Because of this, deepfake detection is an important field of study that aims to recognise and lessen the dangers associated with faked facial media. Numerous generation techniques have surfaced in recent years, ranging from sophisticated GAN and 3D model-driven frameworks to pipelines based on autoencoders. At the same time, deepfake detection techniques have advanced from basic hand-crafted forensic clues to complex hybrid models that include physiological, frequency- domain, temporal, and spatial signal analysis. In order to create and assess both generative and detection methods, benchmark datasets like Face Forensics++, Celeb-DF, and DFDC have been essential. The goal of this project is to develop a deepfake detection module and a reliable face swapping system. The objective is twofold: to comprehend and apply contemporary face synthesis methods, and to investigate efficient detection approaches that can recognize modified information in realistic settings.

II. LITERATURE SURVEY

Taking into account both scholarly research and commercial applications, the literature review investigates the evolution, potential, and constraints of current face swapping and deepfake detection systems. The background information required to support the architecture and methods used in the suggested Face Swapping & Deepfake Detection system is provided by this review. Here from the past books I collected some books and I explore the three books. They are: ResNet-18 – He et al.(2015) from the paper of Deep Residual Learning for Image Recognition by the Author Jian Sun and the second one MTCNN- Zhang et al. (2016) from the paper of joint face detection and alignment using multi-task cascaded convolutional network by the author - Zhan Peng Zhung and the third one is vision transformers(Vit) – Zhan et al. (2021) from the paper of multi – scale vision transforms from the deep fake detection by the author – Hong zoon, Lijun wang, et al.

III. RELATED WORK

Proposes correlating signals from the facial region with the surrounding context (hair, ears, neck) in order to detect identity modifications, including face swapping. makes use of two networks: one for context recognition and one for face identification. Inconsistencies brought about by swapping are seen in the differences between their results. And the technologies used in it are the – Artificial Intelligence, Both the production and detection of face-swapped or deepfake video heavily rely on artificial intelligence. The problem is both complex and fascinating because of these two involvements. Deepfakes and face- swapped media are both produced and detected using machine learning (ML). A NoSQL database called MongoDB is essential for organizing and storing the vast amounts of varied data created and utilized throughout the project. MongoDB is primarily utilized in the context of face swapping and deepfake detection.

IV. EXISTING SYSTEM

The previous book titled ResNet-18 – He et al.(2015) from the paper of Deep Residual Learning for Image Recognition by the Author Jian Sun, the disadvantages are As a shallower version, ResNet-18 usually performs less accurately than deeper versions such as ResNet-50 or ResNet-101.Its limited capacity may cause it to perform poorly on sophisticated tasks (such as high-resolution detection or fine- grained categorization).The model's 18 layers might make it difficult to grasp abstract elements or extremely intricate patterns found in deeper neural networks. Our approach uses deeper and more expressive convolutional neural networks, which can extract richer spatial data and recognize complex objects linked to synthetic content, to get around these restrictions. Our methodology guarantees higher generalization across different deepfake techniques and face swapping methods and greatly increases the accuracy of detecting modified facial regions by using models with more depth and sophisticated feature extraction capabilities. The architecture is also tailored to balance detection performance and computing economy, which makes it appropriate for real-time applications as well as offline analysis.

PROPOSED SYSTEM

The intended system has two main objectives: to create deepfake media by swapping faces and deepfake content detection using classification based on deep learning. These two modules are executed independently as part of a single unified architecture that demonstrates both the creation, and the identification of AI-altered visual media. The face swapping, the first part of our system, takes in two input files - a source image or video that contains the face to be swapped, and a target video or image to place the new face onto. The face swapping process begins with applying facial detection using models such as MTCNN or Dib to accurately identify and localize facial landmarks (eyes, nose, and mouth), which helps prepare the face swap step.

SYSTEM ARCHITECTURE

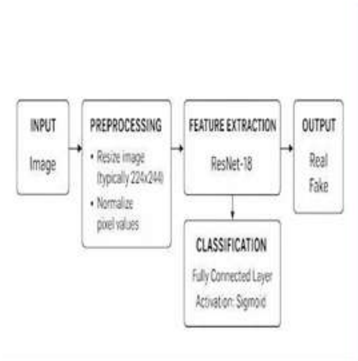

Fig:1 Architecture of AI-based Face Swapping & Deepfake Detection

METHODOLOGY DESCRIPTION

Input:

The input to the system consists of images (either derived from videos or otherwise). Images can either be from real or manipulated (deepfake) origins. The key question that the system seeks to answer is whether the input image either has been manipulated or has been created synthetically.

Preprocessing:

Prior to model input, the image undergoes preprocessing to ensure homogeneity and compatibility to neural network input specifications.

Preprocessing consists of Resizing:

Images need to be resized to fit a fixed resolution, typically 224×224 pixels to fit the input size of the ResNet-18 architecture. Normalization: Pixel values must be normalized, (typically to a range of [0, 1] or [-1, 1]) to stabilize learning and convergence during training. Normalization ensures that all input features contribute equally.

Feature Extension:

Following preprocessing, the picture is transferred to the feature extraction module, which uses ResNet-18, a deep residual network that contains 18 layers. ResNet-18 was selected due to its efficiency, as well as its ability to learn hierarchical visual representations and to use skip connections to alleviate the vanishing gradient problem. The model extracts high-level features from the input image that represent the types of patterns and anomalies that appear in deepfake media, such as lighting inconsistencies, blending artifacts, and unnatural facial textures.

Classification:

After that, the features are passed to a dense layer, which is supervised for binary classification using a sigmoid activation function that returns a probability score between 0 and 1: a score close to 0 indicates a real image, while a score close to 1 indicates a fake image. The output probability is threshold (generally at 0.5) to make the final classification.

Output:

The output of the system will always be a label: Real – if the image is determined to be authentic. Fake – if the image is determined to be a manipulated image (deepfake). The classification result can also be saved, visualized or built into a bigger system for real-time video forensics or content moderation.

V. RESULTS AND DISCUSSION

Fig :2 Home page

The login page is the main interface for user authentication, allowing registered users to log in securely with valid credentials. There are also validation mechanisms in place to help prevent unauthorized access and protect user information.

Fig :3 Registration page

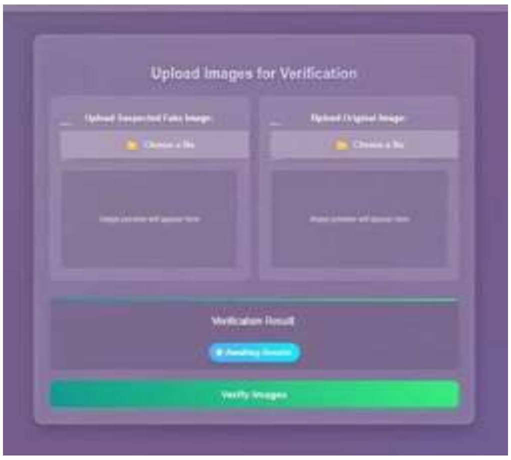

Fig 4: Upload image for verification

The registration page allows new user accounts to be created by asking for the necessary details of the user including name, contact details, and login credentials. It ensures that only authenticated users are allowed to use the features of the system by validating the information the user provided in account creation.

Fig 5: Forgery Detection

Here you can upload the image that you want to test real or fake. Here you can see the fear image and the fake image.

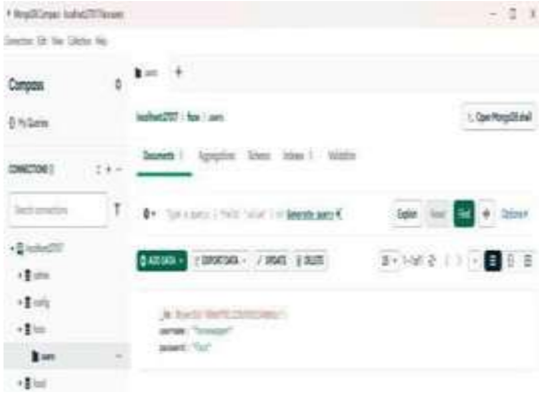

Fig 6: Database Storage

MongoDB Compass stores and manages project data, which includes uploaded images, user information, and detection results. All objects are saved to the database as documents with fields such as filename, upload time, and forgery status for simple, efficient storage and retrieval.

VI. CONCLUSION

The Face Swapping and Deepfake Detection system was effectively developed and implemented as a robust framework to identify and verify manipulated facial images. Leveraging advanced deep learning techniques, particularly the ResNet-18 model, the system successfully classifies deepfakes and authentic faces with high accuracy. Additionally, the use of SSIM (Structural Similarity Index) enhances pixel-level verification, ensuring improved reliability and confidence in detection results. This platform addresses the growing threat of AI-based face manipulations, offering a powerful defense mechanism against misinformation, identity theft, and digital fraud. Overall, the system contributes significantly to maintaining digital authenticity and trust in multimedia content.

FUTURE SCOPE

In the future, the Face Swapping and Deepfake Detection system can be further enhanced to achieve greater accuracy, robustness, and real-time performance. Optimizing the model for low-latency real-time processing will enable the system to detect and prevent deepfakes instantaneously, making it suitable for live applications such as video conferencing and social media content verification. The emotional and expression consistency between the source and target faces can be improved using advanced 3D deep face models or GAN-based techniques to ensure more realistic and coherent results. Additionally, incorporating multi-modal detection methods that combine visual, audio, and biometric features such as lip-sync accuracy and blink rate can provide a more comprehensive and reliable deepfake identification. To further strengthen system reliability, enhancing adversarial robustness through ensemble learning and adversarial training will help defend against attacks designed to deceive AI models. Finally, the system can be scaled for broader applications, including digital forensics, media authentication, and law enforcement, contributing to a safer and more trustworthy digital ecosystem.

REFERENCE

[1] Harini, J. D. P. Adaptable Face Picture Recovery utilizing Quality Upgraded Meager Code words.

[2] Dolhansky, B., Howes, R., Pflaum, B., Baram, N., & Ferrer, C. C., "The Deepfake Detection Challenge (DFDC) Dataset," arXiv preprint arXiv:2006.07397, 2020. [Online]. A

[3] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I., "MesoNet: A Compact Facial Video Forgery Detection Network," IEEE International Workshop on Information Forensics and Security (WIFS), pp. 1–7, 2018.

[4] Korshunov, P., & Marcel, S., "Deepfakes: A New Threat to Face Recognition? Assessment and Detection," arXiv preprint arXiv:1812.08685, 2018. [Online]

[5] Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., & Natarajan, P., "Recurrent Convolutional Strategies for Face Manipulation Detection in Videos," IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 80–87, 2019

[6] OpenAI, "Deepfake Detection Research Overview." [Online].

[7] MongoDB, "MongoDB Atlas Documentation." [Online].

[8] Flask, "Flask Web Development Framework." [Online].

[9] Python Software Foundation, "Python Documentation." [Online].

[10] W3Schools, "HTML, CSS,JavaScript, and Flask Tutorials." [Online].

[11] Geeks for Geeks, "Flask Tutorials and Python Web Development." [Online].

[12] W3Schools, "HTML, CSS, JavaScript, and Flask Tutorials." [Online].