ISSN 2277-2685
IJESR/April. 2026/ Vol-16/Issue-2s1/245-249
Mohd Faizan Ali *et. al.*, / International Journal of Engineering & Science Research

Full Length Article

Fake Face Detection Based On Videos Using OpenCV And Neural Network Architecture

Mohd Faizan Ali[1], Mohd Shoaib[2], Mohd Imamuddin Naveed[3], Mr. Syed Zaffar Mahmood[4]
[1,2,3]B.E.Students; Dept. of Information Technology ISL Engineering College, Hyderabad India.
[4]Associate Professor; Dept. of Information Technology ISL Engineering College, Hyderabad India.
Mail Id; mohdfaizaanali604@gmail.com ,mohdshoaibmz12345@gmail.com
Imamuddinnaveed595@gmail.com
Accepted 27-04-2026

Abstract
The rapid development of the Internet has enabled the widespread distribution of manipulated facial images, particularly Deepfakes, which are increasingly difficult to detect using conventional methods. While current approaches focus on spatial domain features or complex network architectures, they often lack robustness against sophisticated forgery techniques. To address this, we propose a MobileNetV2-based Deepfake detection framework that leverages efficient convolutional feature extraction for accurate classification of real and fake facial images. The framework begins with OpenCV-based preprocessing, including face detection, alignment, and normalization, to ensure consistent input quality and enhance the discriminative features for detection. MobileNetV2, a lightweight yet powerful convolutional neural network, is employed to automatically learn hierarchical spatial features from the preprocessed facial images, eliminating the need for handcrafted features. By combining OpenCV preprocessing with MobileNetV2, the proposed system effectively captures subtle visual artifacts and texture inconsistencies introduced by Deepfake manipulation. This approach enables robust and scalable detection, generalizing well across diverse datasets and real-world scenarios, providing a practical solution for automated Deepfake detection in security, media verification, and social media monitoring applications.

Keywords: *Deepfake Detection, MobileNetV2, OpenCV, Computer Vision, Deep Learning, Neural Network Architecture.*

Introduction

In the modern digital era, the advent of advanced artificial intelligence and deep learning technologies has transformed digital media generation, giving rise to hyper-realistic synthetic content known as Deepfakes. Deepfakes are manipulated facial images or videos created using deep generative models such as GANs and autoencoders, which convincingly alter or synthesize human appearances. Malicious use of Deepfakes in spreading misinformation, identity theft, and political manipulation has raised global concerns, making automated detection techniques a critical research area. Traditional detection approaches often rely on handcrafted features or simple spatial domain cues, but these methods struggle against modern forgery techniques that introduce highly subtle and nearly imperceptible artifacts.

In this study, we propose a Deepfake detection framework based on MobileNetV2, a lightweight yet efficient convolutional neural network architecture. The system is supported by OpenCV-based preprocessing techniques including face detection, alignment, and normalization, which ensure consistency and improve feature quality. MobileNetV2 is then employed to learn hierarchical spatial features that capture texture inconsistencies and visual artifacts introduced during manipulation. The lightweight nature of MobileNetV2 further ensures faster training and deployment, making it suitable for real-time applications. Ultimately, the study aims to restore trust in digital media and mitigate the harmful consequences of malicious Deepfake use.

Literature Survey

MesoNet: A Compact Facial Video Forgery Detection Network (2018): This work proposes MesoNet, a lightweight CNN targeting mid-level texture artifacts rather than heavy global features. Results show strong performance on early FaceSwap/DeepFake content with low compute cost, establishing a practical baseline for real-time deepfake screening.

Face Forensics++: Learning to Detect Manipulated Facial Images (2019): Introduces a large-scale benchmark for facial manipulation detection, evaluating popular detectors with Xception as a strong baseline. The dataset catalyzed research on robust preprocessing and augmentation.

Thinking in Frequency: Face Forgery Detection by Mining Frequency-Aware Clues (2020): Exploits frequency-domain artifacts that persist after spatial smoothing to extract multi-band spectral cues to flag forgeries. It improves robustness against compression and color post-processing.

Generalizing Face Forgery Detection with High-Frequency Features (2021): Proposes extracting stable high-frequency representations to combat overfitting in spatial CNNs. High-frequency filters and tailored losses encourage manipulation-invariant cues for robust generalization.

Protecting Celebrities from Deepfake with Identity Consistency Transformer (2022): Introduces an identity-consistency transformer for forgery detection, modeling cross-frame identity coherence to expose subtle temporal inconsistencies. It complements lightweight CNN backbones via late-fusion or distillation.

Methodology

Module Names

Assembling the Dataset
Data Interpretation
Data Conditioning
Model Execution
Model Calibration
Model Performance
Outcome Prediction

Module Explanations

Assembling the Dataset: Gathers a diverse dataset of real and deepfake videos or images, including variations in facial expressions, lighting conditions, and background settings to enhance model robustness.

Data Interpretation: Analyzes the dataset using statistical analysis, visualization, and metadata checks to understand patterns and unique characteristics between real and fake media.

Data Conditioning: Ensures data is cleaned, normalized, and prepared through noise removal, resizing frames, feature extraction (facial landmarks/texture analysis), and data augmentation (rotation, flipping, brightness adjustments).

Model Execution: Trains the deep learning model using frameworks such as TensorFlow to capture deepfake-specific artifacts, including inconsistencies in eye blinking, lip synchronization, and facial blending.

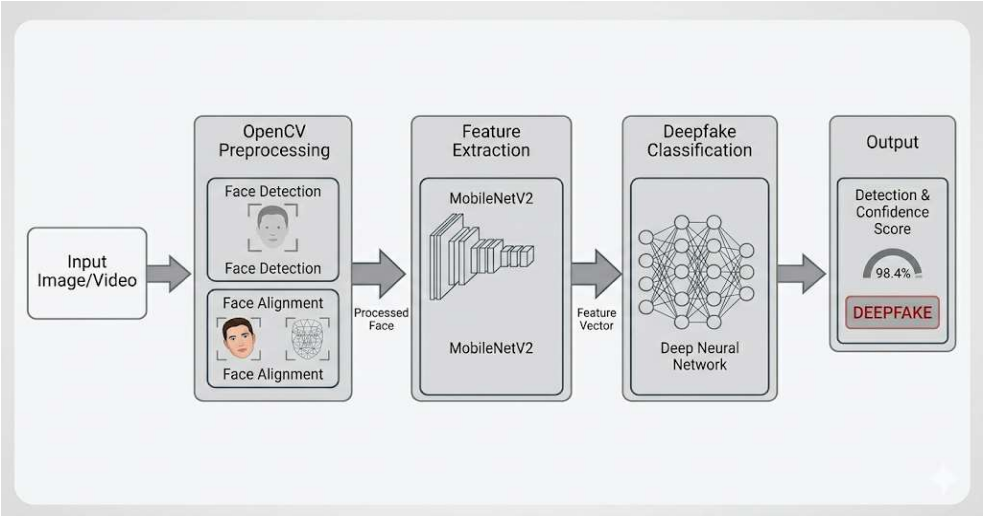
Model Calibration: Adjusts hyperparameters (learning rate, batch size, optimizer selection, and epochs) and utilizes cross-validation to avoid overfitting and underfitting.

Model Performance: Evaluates the model using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Confusion matrix analysis helps identify false positives and negatives.

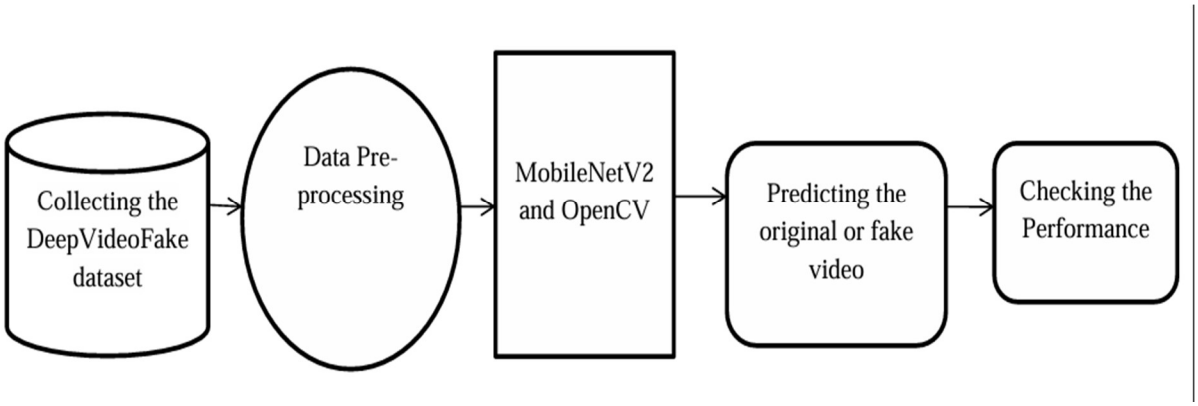
Outcome Prediction: Deploys the trained model to predict whether new input media is real or fake, generating a classification result with confidence scores.

Block Diagram

Input Image/Video → OpenCV Preprocessing (Face Detection & Alignment) → Feature Extraction via MobileNetV2 → Deepfake Classification → Output Detection & Confidence Score



System Architecture



Implementation (Algorithm / Flowchart)

Algorithm Steps

Capture input image or video stream.
Pass input to OpenCV for facial detection and bounding box generation.
Preprocess and align facial regions to standardize inputs.
Feed processed frames into the trained MobileNetV2 architecture.
Extract hierarchical spatial features to identify texture anomalies and blending artifacts.
Generate classification probabilities (Real vs. Fake).
Display bounding boxes with classification labels and confidence scores on the user interface.
Compute Grad-CAM heatmaps to visualize the regions influencing the model's decision.

Testing

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components and ensures the software system meets user expectations without failing unacceptably.

Types of Tests

Unit Testing: Validates that the internal program logic is functioning properly and that program inputs produce valid outputs.

Functional Test: Provides systematic demonstrations that functions tested are available as specified by the business and technical requirements. Validates inputs, functions, outputs, and interfacing systems.

System Test: Ensures that the entire integrated software system meets requirements, testing a configuration to ensure known and predictable results.

Performance Test: Ensures that the output is produced within the time limits and evaluates the time taken by the system for compiling and responding to users.

Integration Testing: Incremental testing of two or more integrated software components to verify they interact without error.

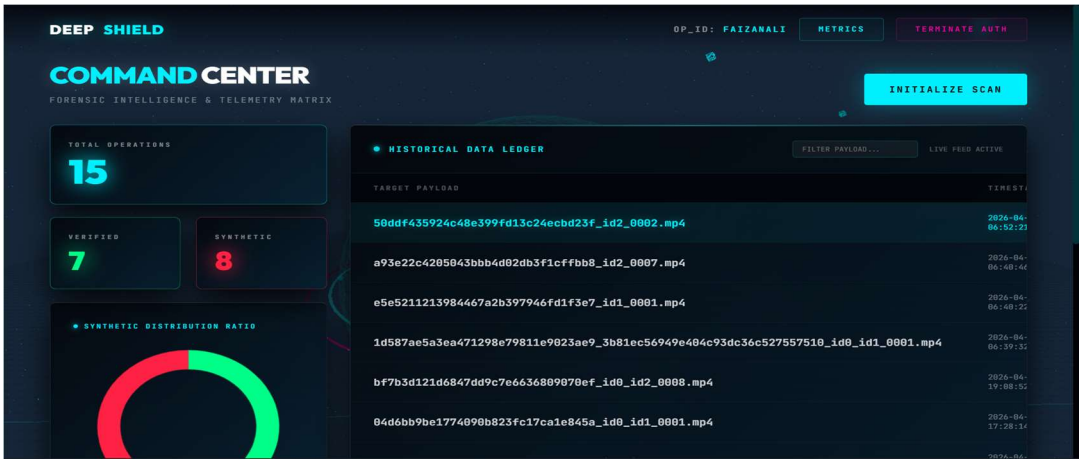
Acceptance Testing: Ensures that the system meets the functional requirements with significant participation by the end-user.

Results

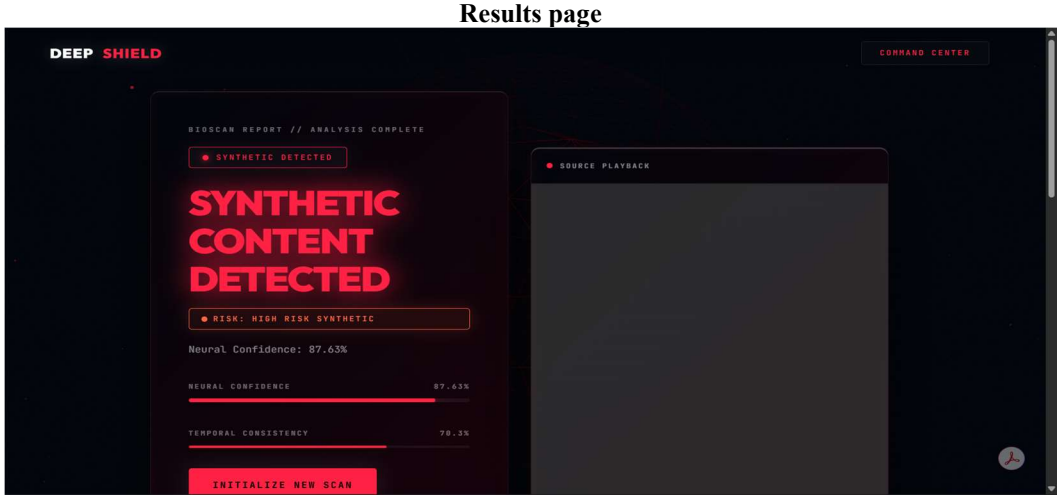
Model	Architecture Type	Accuracy	Speed / Inference Time	Scalability
Standard CNN	Spatial only	Moderate	Slow	Low
Hybrid Transformer	Spatial + Temporal	Very High	Very Slow	Low
MobileNetV2	Lightweight Spatial	High	Fast	High

Output Screenshots

Dashboard



Results page



Conclusion

The project on deepfake detection highlights the growing necessity of combating AI-generated misinformation in today's digital world. Deepfake technology, while innovative, poses serious threats to security, privacy, and trust in digital media. By leveraging advanced machine learning algorithms, specifically combining OpenCV preprocessing with the MobileNetV2 architecture, the system provides a robust framework to differentiate between real and manipulated content. The modular workflow ensures a systematic approach to model development, and experimental results confirm the potential of AI-based solutions in identifying subtle inconsistencies in manipulated media. Ultimately, this work lays a strong foundation for protecting digital integrity in an AI-driven era, empowering individuals and strengthening trust across online platforms.

Future Enhancements

In the future, the deepfake detection system can be enhanced by incorporating multimodal analysis, combining audio, video, and text cues for more reliable detection. Real-time detection mechanisms can be integrated into social media platforms to automatically flag or block harmful content. The use of blockchain technology can ensure the authenticity and traceability of digital media. Furthermore, federated learning can be adopted to train models collaboratively without compromising user privacy, and improvements in explainable AI (XAI) will allow users to better understand the reasoning behind detections. Continuous dataset updates and adversarial training will further improve robustness against evolving manipulation techniques.

References

[1] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, et al., "Deepfakes and beyond: A survey of face manipulation and fake detection," Information Fusion, vol. 64, pp. 131-148, 2020.
[2] Y. Z. Li, M. C. Chang, and S. W. Lyu, "In ictu oculi: Exposing AI created fake videos by detecting eye blinking," in Proc. IEEE Int. Workshop on

Information Forensics and Security (WIFS), pp. 1-7, 2018.
[3] X. Wu, Z. Xie, Y. T. Gao, et al., "SSTNet: Detecting manipulated faces through spatial, steganalysis and temporal features," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), pp. 2952-2956, 2020.
[4] A. Rossler, D. Cozzolino, L. Verdoliva, et al., "Faceforensics++: Learning to detect manipulated facial images," in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 1-11, 2019.
[5] Y. Y. Qian, G. J. Yin, L. Sheng, et al., "Thinking in frequency: Face forgery detection by mining frequency-aware clues," in Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 86-103, 2020.
[6] Y. C. Luo, Y. Zhang, J. C. Yan, et al., "Generalizing face forgery detection with high-frequency features," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 16312-16321, 2021.
[7] X. Y. Dong, J. M. Bao, D. D. Chen, et al., "Protecting celebrities from deepfake with identity consistency transformer," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 9458-9468, 2022.
[8] D. Afchar, V. Nozick, J. Yamagishi, et al., "MesoNet: A compact facial video forgery detection network," in Proc. IEEE Int. Workshop on Information Forensics and Security (WIFS), pp. 1-7, 2018.