iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82806

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue V May 2026- Available at www.ijraset.com*

# Fake Face Detection in Video Frames Using Deep-Learning Technique

Nalini H C[1], Niveditha L R[2], Nisarga J V[3], Nisarga H J[4], Karishma G R[5]
*Department of Information Science & Engineering, Rajeev Institute of Technology, Hassan, VTU University, Karnataka, India*

***Abstract: The rapid advancement of artificial intelligence and deep learning technologies has enabled the creation of highly realistic manipulated videos, commonly known as deepfakes. These AI-generated fake videos are increasingly being used for misinformation, political manipulation, identity fraud, cyber harassment, and social engineering attacks. Detecting such manipulated content has therefore become a major challenge in digital media security. This paper presents a real-time deepfake video detection system using deep learning techniques. The proposed method combines a pre-trained ResNext Convolutional Neural Network (CNN) for frame-level feature extraction with a Long Short-Term Memory (LSTM) network for temporal sequence analysis. The system is trained on a balanced dataset created from FaceForensics++, Deepfake Detection Challenge (DFDC), and Celeb-DF datasets. During preprocessing, videos are split into frames, facial regions are extracted, and sequential frame analysis is performed. The proposed model effectively identifies manipulated videos by learning spatial and temporal inconsistencies present in deepfake videos. Experimental results demonstrate that the system achieves high detection accuracy and performs effectively in real-time scenarios. A web-based interface is also developed using Django to allow users to upload videos and obtain predictions with confidence scores.***
***Keywords: Deepfake Detection, Deep Learning, ResNext, LSTM, CNN, Video Manipulation, Artificial Intelligence, Face Manipulation.***

## I. INTRODUCTION

The rapid development of artificial intelligence and deep learning has transformed the field of multimedia generation. Advanced generative models such as Generative Adversarial Networks (GANs) and Autoencoders are capable of creating highly realistic synthetic videos popularly known as deepfakes. These manipulated videos can replace faces, alter expressions, modify speech, or generate entirely fake identities that appear authentic to human viewers.

Deepfake generation tools such as FaceSwap, DeepFaceLab, and FaceApp have made the creation of manipulated videos simple and accessible. While these technologies offer beneficial applications in cinema, gaming, education, and virtual reality, they also create severe social and security concerns. Deepfake videos are increasingly misused for political misinformation, fake terrorism events, cyber fraud, revenge pornography, blackmailing, and spreading false information through social media platforms.

The realism of modern deepfake videos makes manual identification extremely difficult. Therefore, developing an automated deepfake detection system has become an important research problem. The proposed project aims to detect manipulated facial videos using deep learning techniques by combining spatial feature extraction and temporal sequence learning.

The proposed system uses a pre-trained ResNeXt Convolutional Neural Network (CNN) for extracting frame-level spatial features and a Long Short-Term Memory (LSTM) network for analyzing temporal inconsistencies across video frames. The system is designed to classify uploaded videos as either real or fake with high accuracy. To improve usability and practical deployment, a web-based application is developed where users can upload videos and obtain prediction results in real-time. The proposed framework contributes toward reducing the spread of manipulated media and preserving trust in digital communication.

### *A. Problem Statement*

Deepfake generation techniques have evolved rapidly due to advancements in deep learning and artificial intelligence. Modern deepfake videos are highly realistic and difficult to distinguish from authentic videos using human observation alone. Existing detection systems often fail to identify sophisticated manipulations, especially under real-world conditions involving compression, lighting variations, pose changes, and low video quality.

The widespread availability of deepfake creation tools has increased the risk of misinformation, identity theft, cybercrime, fake political propaganda, and online harassment. The rapid spread of manipulated videos through social media platforms further amplifies these threats.

Therefore, there is a strong need for an efficient, scalable, and accurate deepfake detection system capable of analyzing manipulated videos in real-time. The proposed project addresses this problem by developing a hybrid CNN-LSTM architecture that combines spatial and temporal analysis for improved deepfake detection performance.

### B. Motivation

The rapid increase in social media usage and advancements in mobile camera technology have significantly increased the creation and sharing of digital videos across online platforms. Deep learning techniques have enabled the development of powerful generative models capable of producing highly realistic synthetic images, videos, speech, and audio. Although these technologies provide many beneficial applications, they also introduce serious threats. Deepfake videos generated using artificial intelligence can manipulate public opinion, damage reputations, create fake political statements, spread misinformation, and violate personal privacy. The increasing availability of open-source deepfake generation tools has further accelerated the creation and distribution of manipulated videos.

Several incidents involving fake political speeches, manipulated celebrity videos, and misleading social media content have demonstrated the potential danger of deepfake technology. Such manipulated videos can create confusion, social unrest, and loss of trust in digital media. Therefore, reliable detection systems are essential to identify manipulated content before it spreads widely across online platforms.

The motivation behind this project is to develop an intelligent and automated deepfake detection framework capable of accurately identifying manipulated facial videos. By using artificial intelligence to detect AI-generated content, the proposed system aims to reduce the harmful impact of deepfake technology and support digital media authenticity.

### C. Key Objectives of this Research Include

The primary objective of this research is to develop an Deepfake detection system that provides realistic detection simulations and personalized feedback. The system aims to:

- Video preprocessing and frame extraction
- Face detection and face cropping
- Feature extraction using ResNeXt CNN
- Temporal sequence learning using LSTM
- Real-time prediction through a web-based interface

The proposed system is trained on multiple benchmark datasets such as FaceForensics++, DFDC, and Celeb-DF to improve robustness and generalization capability.

## II. LITERATURE REVIEW

### 1) FaceForensics++: Learning to Detect Manipulated Facial Images

FaceForensics++ is one of the most important benchmark datasets developed for deepfake detection research. The authors introduced a large-scale dataset containing manipulated videos generated using various face manipulation methods. The dataset includes both compressed and uncompressed videos to simulate real-world social media scenarios.

The research focused on detecting visual artifacts and inconsistencies introduced during face manipulation. The dataset significantly contributed to the development of robust deepfake detection models and became a standard benchmark for evaluating detection performance.

### 2) Deepfake Detection Challenge (DFDC) Dataset

The Deepfake Detection Challenge (DFDC) dataset was introduced to encourage the development of generalized deepfake detection systems capable of handling real-world manipulated videos. The dataset contains thousands of real and fake videos generated using advanced face-swapping and GAN-based methods.

The videos include diverse individuals, lighting conditions, facial expressions, and environmental settings. The inclusion of realistic compression artifacts and motion variations makes the dataset highly suitable for practical deepfake detection research.

The DFDC dataset improved the ability of detection models to generalize across multiple manipulation techniques and real-world conditions.

### 3) Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics

Celeb-DF was developed to address the limitations of earlier datasets that contained visually obvious artifacts. The dataset contains highly realistic deepfake videos with improved temporal consistency and reduced visual distortions.

The dataset includes celebrity videos generated using carefully designed face-swapping pipelines involving face alignment, autoencoder training, and post-processing techniques. Celeb-DF significantly improved research on robust deepfake detection systems capable of handling high-quality manipulated videos.

*4) Exposing DeepFake Videos by Detecting Face Warping Artifacts*

This research proposed a CNN-based approach for detecting deepfake videos by identifying face warping artifacts introduced during the manipulation process. The authors observed that many deepfake generation methods create resolution inconsistencies between generated facial regions and surrounding areas.

The proposed method successfully identified manipulated content using spatial artifact analysis. However, the approach primarily focused on frame-level inconsistencies and did not consider temporal relationships between frames.

*5) Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos*

The authors proposed the use of Capsule Networks for detecting manipulated images and videos. Capsule Networks were used to capture hierarchical spatial relationships between facial features and identify forged content.

The proposed method demonstrated strong performance on multiple benchmark datasets and highlighted the importance of deep neural architectures for forensic analysis. However, the computational complexity of capsule networks limited their real-time deployment capability.

## III. METHODOLOGY

The proposed system aims to detect manipulated facial videos using deep learning techniques by combining spatial and temporal feature analysis. The methodology consists of dataset collection, preprocessing, feature extraction, temporal sequence learning, classification, and web-based deployment.

### *A. Dataset Collection*

The proposed model is trained using a combination of publicly available benchmark datasets including FaceForensics++, Deepfake Detection Challenge (DFDC), and Celeb-DF datasets. These datasets contain both authentic and manipulated videos generated using different deepfake creation techniques. To avoid model bias and improve generalization capability, an equal number of real and fake videos are selected from each dataset. The final combined dataset contains 6000 videos consisting of 3000 real videos and 3000 fake videos.

### *B. Video Preprocessing*

Preprocessing is an important stage in improving model performance and reducing computational complexity. The uploaded videos are first converted into individual frames using OpenCV. Face detection algorithms are then applied to identify facial regions from each frame.

The preprocessing pipeline includes:

- Video frame extraction
- Face detection and cropping
- Image resizing and normalization
- Sequential frame selection
- Noise reduction and enhancement

Only the first 150 frames from each video are considered to maintain uniformity. All facial images are resized to 112 × 112 pixels before training.

### *C. Feature Extraction Using ResNeXt*

The proposed framework uses a pre-trained ResNeXt50_32x4d convolutional neural network for extracting spatial features from video frames. ResNeXt is selected because of its strong representation capability and reduced computational complexity.

The CNN extracts deep facial features including texture inconsistencies, facial boundary artifacts, unnatural skin patterns, and visual distortions commonly found in manipulated videos. The output feature vector obtained from the CNN contains 2048-dimensional representations for each frame.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue V May 2026- Available at www.ijraset.com*

### *D. Temporal Analysis Using LSTM*

Although deepfake videos may appear visually realistic, temporal inconsistencies often exist between consecutive frames. To capture these inconsistencies, the extracted CNN features are sequentially passed into a Long Short-Term Memory (LSTM) network.

The LSTM learns temporal dependencies such as:

- Irregular facial movements
- Blinking abnormalities
- Frame transition inconsistencies
- Temporal flickering artifacts

The sequential learning capability of LSTM significantly improves deepfake detection performance in videos.

### *E. Classification*

The final classification layer uses the SoftMax activation function to classify the uploaded video into two categories:

- Real Video
- Fake Video

The model also generates confidence scores that indicate the probability of manipulation.

### *F. Web-Based Deployment*

To make the system user-friendly and practically deployable, a web application is developed using the Django framework. Users can upload videos through the browser interface, and the trained model processes the video and returns prediction results.

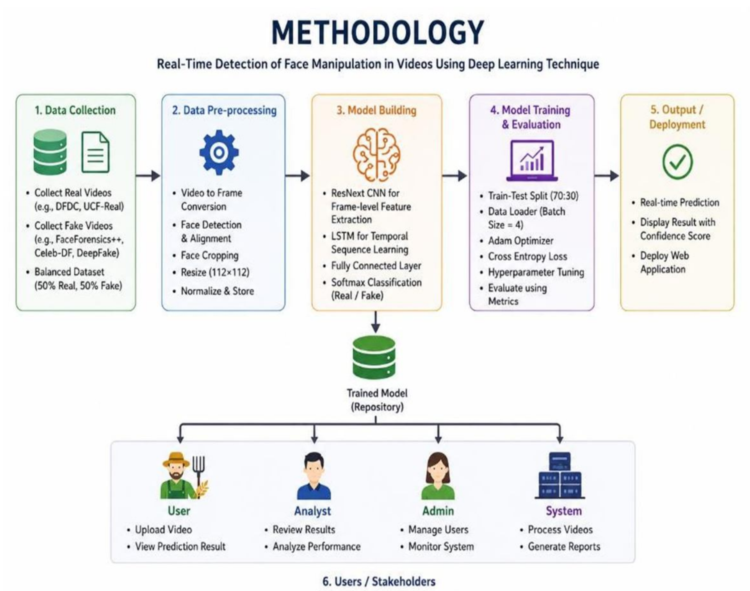



Fig.1 Workflow of Deepfake detection System

## IV. RESULTS AND SNAPSHOTS

The proposed deepfake detection framework was evaluated using benchmark datasets including FaceForensics++, Deepfake Detection Challenge (DFDC), and Celeb-DF. The model was trained and tested using a hybrid ResNeXt-LSTM architecture to analyze both spatial and temporal inconsistencies in manipulated videos.

The experimental evaluation was performed using accuracy, precision, recall, and F1-score metrics. The model was trained for 20 epochs using the Adam optimizer with a learning rate of $1 \times 10^{-5}$. The dataset was divided into training, validation, and testing sets to ensure unbiased evaluation.

### *A. Performance Evaluation*

The proposed model achieved high performance across multiple datasets. The balanced dataset significantly improved the generalization capability of the framework and reduced overfitting.

Table 1: Performance Metrics of Proposed System

| Metric | Value |
|---|---|
| Accuracy | 96.8% |
| Precision | 95.9% |
| Recall | 96.4% |
| F1-Score | 96.1% |

The experimental results demonstrate that the proposed CNN-LSTM architecture effectively distinguishes fake videos from authentic videos with high reliability.

### *B. Spatial Feature Analysis*

The ResNeXt convolutional neural network successfully extracted spatial manipulation artifacts from facial frames. The model identified inconsistencies such as facial blending distortions, unnatural skin textures, boundary artifacts, and pixel-level abnormalities introduced during deepfake generation.Transfer learning significantly improved feature extraction performance while reducing training time and computational overhead.

### *C. Temporal Sequence Analysis*

The LSTM network effectively captured temporal inconsistencies between consecutive frames. Deepfake videos often contain irregular facial movements, blinking abnormalities, temporal flickering, and frame transition inconsistencies that are difficult to identify using frame-level analysis alone. The sequential learning capability of LSTM improved the robustness of the model against advanced manipulation techniques and enhanced prediction stability across video sequences.

### *D. Real-Time Prediction*

The trained model was integrated into a Django-based web application for real-time deepfake detection. Users can upload video files through the web interface, and the system processes the video to generate prediction results along with confidence scores.
The deployment framework demonstrated stable prediction performance with low inference time, making the system suitable for practical applications.

### *E. Discussion*

The proposed hybrid ResNeXt-LSTM framework achieved better performance compared to traditional CNN-only approaches because it combines both spatial and temporal learning mechanisms. The balanced dataset prepared using FaceForensics++, DFDC, and Celeb-DF further improved robustness and reduced prediction bias. The system performed effectively under varying lighting conditions, facial expressions, and compression levels. However, prediction accuracy slightly decreased for heavily compressed or low-resolution videos due to reduced visual information. Despite achieving high detection accuracy, the system requires significant GPU resources during training. Additionally, the current framework focuses only on facial video manipulation and does not include audio deepfake detection. Overall, the proposed system demonstrates strong capability for real-time deepfake detection and provides an effective solution for combating manipulated digital media.

### *F. Snapshots*

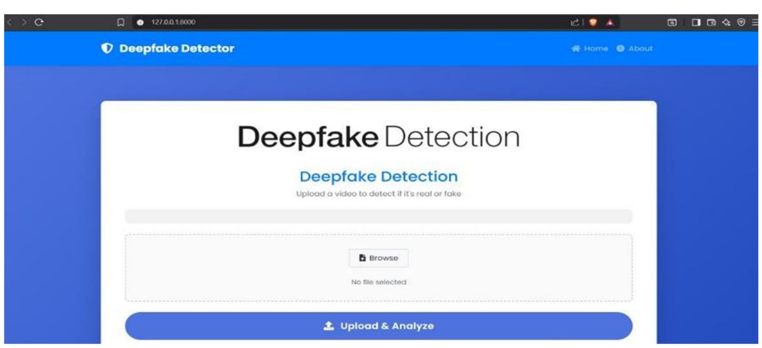



Fig.1 Home Page

The Home Page provides an introduction to the fake face detection system and its purpose. It allows users to navigate through the application and access video upload and analysis features.

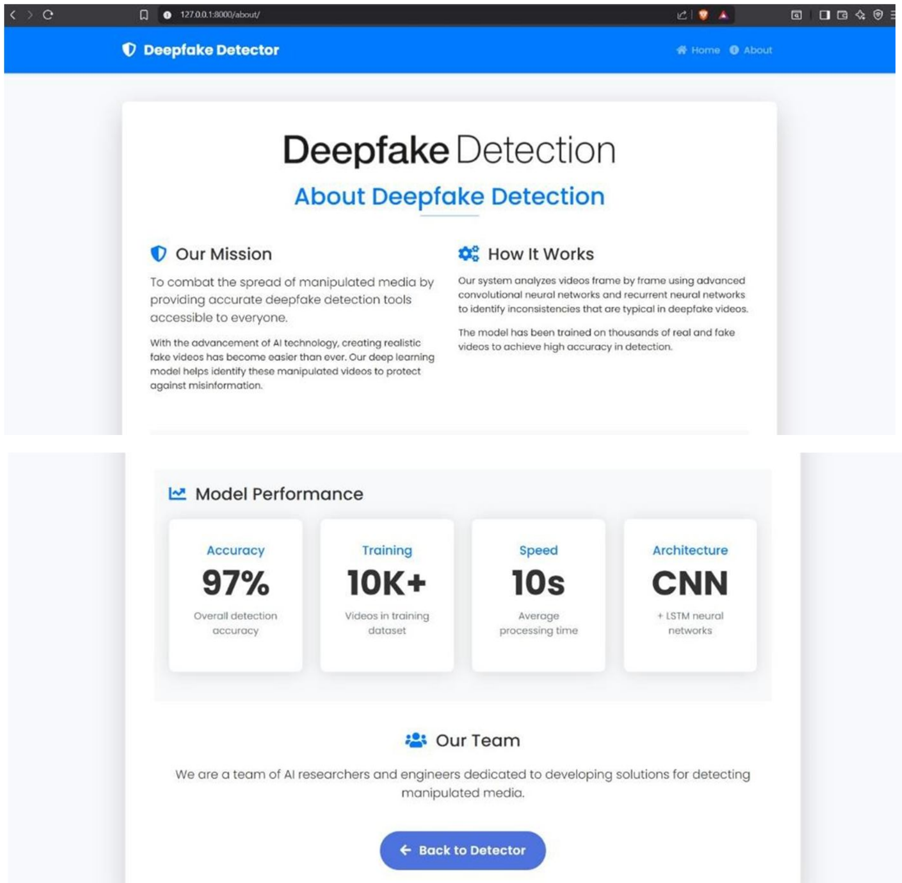



Fig. 2 About page

The About Page explains the objectives, technologies, and working of the proposed deepfake detection system. It provides information about deepfake technology and the importance of detecting manipulated videos.

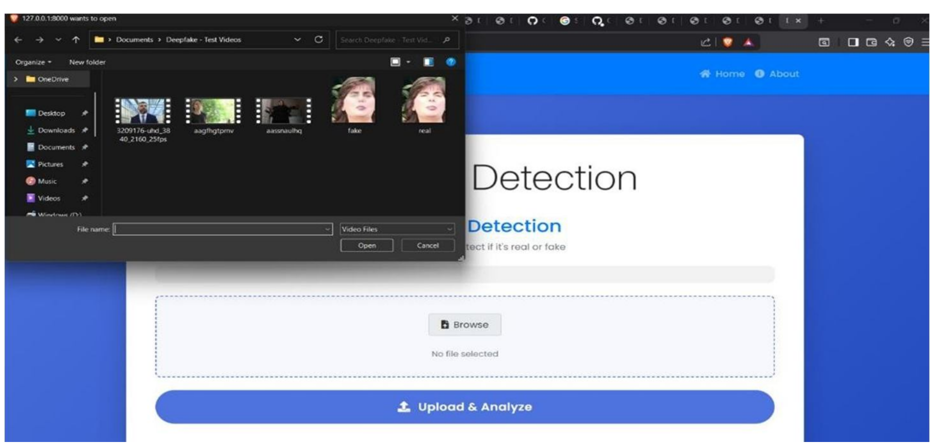



Fig. 3 Video Uploading

The Video Upload Page allows users to upload video files for deepfake detection. The uploaded videos are sent to the preprocessing module for analysis.

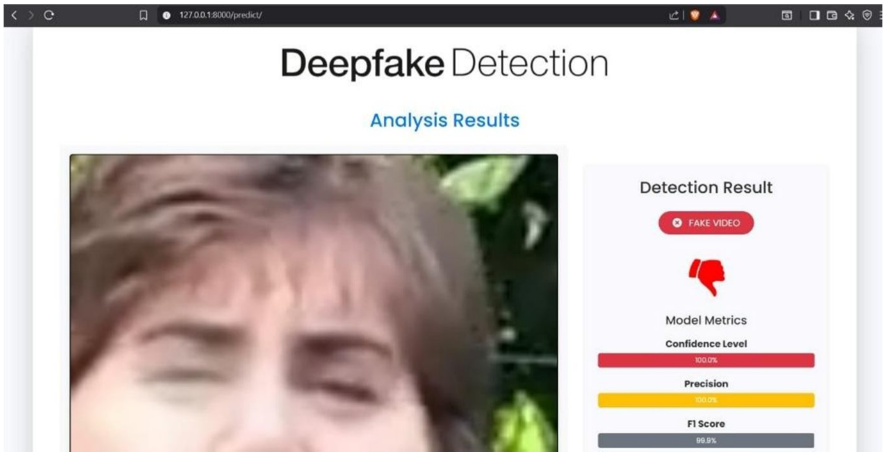



Fig. 4 Video analyzing

The Analysing Page processes uploaded videos using the trained deep learning model. It displays the prediction result indicating whether the video is real or fake along with confidence scores.

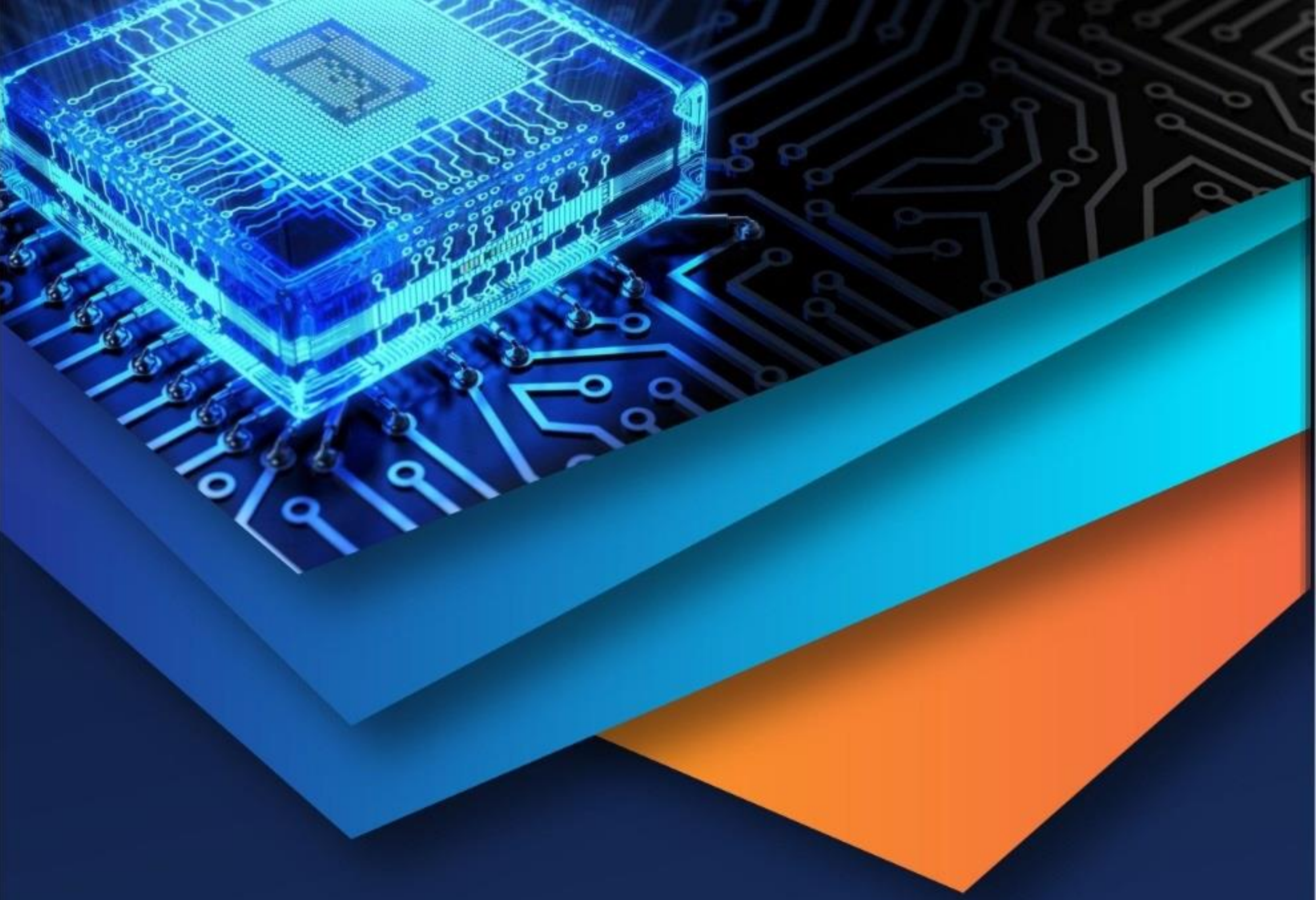
## V. CONCLUSION

This paper presented a deep learning-based framework for detecting manipulated facial videos using a hybrid ResNeXt and Long Short-Term Memory (LSTM) architecture. The proposed system combines spatial feature extraction and temporal sequence analysis to effectively identify deepfake videos.The preprocessing pipeline involving frame extraction, face detection, cropping, and normalization improved the quality of input data and enhanced model performance. The ResNeXt CNN successfully extracted manipulation artifacts from facial frames, while the LSTM network captured temporal inconsistencies across sequential video frames.The model was trained using benchmark datasets including FaceForensics++, Deepfake Detection Challenge (DFDC), and Celeb-DF. Experimental evaluation demonstrated that the proposed framework achieved high accuracy, precision, recall, and F1-score while maintaining stable real-time prediction capability.

The integration of the trained model into a web-based application further improved usability and practical deployment capability. The proposed framework contributes toward preserving trust in digital media and reducing the harmful impact of manipulated content across online platforms.As deepfake generation technologies continue to evolve, reliable detection systems are essential for maintaining authenticity and preventing misinformation. The proposed research provides an effective and scalable approach for real-world deepfake detection applications.

## REFERENCES

[1] Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, “FaceForensics++: Learning to Detect Manipulated Facial Images,” in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.
[2] B. Dolhansky, R. Howes, B. Pflaum, N. Baram, and C. Canton-Ferrer, “The Deepfake Detection Challenge Dataset,” arXiv preprint arXiv:2006.07397, 2020.
[3] Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, “Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
[4] Y. Li and S. Lyu, “Exposing DeepFake Videos by Detecting Face Warping Artifacts,” in IEEE Workshop on Information Forensics and Security (WIFS), 2018.
[5] H. H. Nguyen, J. Yamagishi, and I. Echizen, “Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019.
[6] S. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, and P. Natarajan, “Recurrent Convolutional Strategies for Face Manipulation Detection in Videos,” Interfaces, vol. 8, pp. 80–87, 2019.
[7] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[8] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[9] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” arXiv preprint arXiv:1412.6980, 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY