2026 Volume: 02 Issue: 05 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-07

DETECTION OF DEEPFAKE VIDEOS USING ARTIFICIAL INTELLIGENCE

*Abhishek Rajput

India.

Article Received: 04 April 2026
Article Revised: 24 April 2026
Published on: 14 May 2026

***Corresponding Author: Abhishek Rajput**
India.

DOI: https://doi-doi.org/101555/ijrpa.6625

ABSTRAC

Deepfake technology uses Artificial Intelligence (AI) and Deep Learning techniques to create fake videos, images, and audio that appear highly realistic. While this technology has applications in entertainment and media production, it also creates serious risks including misinformation, fraud, cybercrime, political manipulation, and identity theft. Detecting deepfake videos has become a major research challenge because modern generative models produce highly convincing synthetic media. This research paper studies different AI-based deepfake detection techniques, datasets, challenges, and future directions. The paper focuses on Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), facial landmark analysis, and multimodal detection methods. It also discusses datasets such as FaceForensics++, Celeb-DF, and DFDC. The study concludes that AI-based detection systems can significantly improve fake media identification, but continuous advancements in generative AI require more robust and adaptive detection frameworks.

1. INTRODUCTION

Artificial Intelligence has transformed digital media creation. One of the most advanced AI-generated media technologies is the “deepfake.” Deepfake videos are synthetic videos generated using Deep Learning algorithms, especially Generative Adversarial Networks (GANs) and autoencoders. These videos manipulate facial expressions, lip movements, and voice patterns to imitate real individuals.

Initially, deepfake technology was used for entertainment and research purposes. However, its misuse has increased rapidly in recent years. Deepfakes are now used in fake news, celebrity impersonation, financial fraud, political propaganda, and cybercrime. Experts warn that deepfake fraud is growing rapidly and becoming more difficult to identify.

Because humans often fail to distinguish fake videos from real ones, researchers are developing AI-based deepfake detection systems. These systems analyze facial inconsistencies, visual artifacts, blinking patterns, head movements, and audio-video synchronization to identify manipulated content.

2. Objectives of the Research

The major objectives of this research are:

1. To study deepfake generation techniques.
2. To analyze AI-based deepfake detection methods.
3. To examine datasets used for deepfake detection.
4. To identify challenges faced by current detection systems.
5. To propose future improvements for reliable detec

3. LITERATURE REVIEW

Several researchers have contributed to deepfake detection research.

A comprehensive review published in Artificial Intelligence Review explained that deepfake videos are becoming a major social and cybersecurity threat. The study discussed CNN-based and multimodal detection methods and highlighted the need for better generalization across datasets.

Another survey published in Discover Applied Sciences analyzed modern deepfake detection techniques and future research directions. The authors explained that deep learning methods remain effective, but real-world deepfakes continue to evolve rapidly.

Research published in Computers and Electrical Engineering categorized deepfake detection into image, audio, video, and multimodal detection approaches. The paper emphasized that future systems should combine multiple detection strategies to improve robustness.

Researchers also studied fairness and bias in deepfake detection systems. Studies found that many datasets contain demographic imbalance, which affects model accuracy for different genders, skin tones, and facial characteristics.

4. DEEPFAKE TECHNOLOGY

Deepfake technology uses Deep Learning models to generate fake media. Common technologies include:

4.1 Generative Adversarial Networks (GANs)

GANs consist of two neural networks:

Generator

Discriminator

The generator creates fake images or videos, while the discriminator tries to distinguish fake content from real content. Continuous competition improves realism.

4.2 Autoencoders

Autoencoders compress and reconstruct facial data. In deepfake generation, facial features from one person are mapped onto another person's face.

4.3 Face Swapping

Face swapping replaces the original face in a video with another person's face while preserving expressions and movements.

5. AI-Based Deepfake Detection Techniques

5.1 Convolutional Neural Networks (CNNs)

CNNs are widely used for image and video analysis. They detect pixel-level inconsistencies and visual artifacts in manipulated videos.

Common CNN Models

XceptionNet

MesoNet

EfficientNet

ResNet

CNNs achieve high accuracy on benchmark datasets but may struggle with unseen manipulation techniques.

5.2 Recurrent Neural Networks (RNNs)

RNNs analyze temporal information in videos. They study frame sequences and detect unnatural motion patterns.

Advantages:

Captures video sequence behavior

Useful for blinking and lip-sync detection

Limitations:

High computational cost

5.3 Facial Landmark Analysis

Facial landmark detection identifies abnormal facial movements and geometric inconsistencies.

Key features analyzed:

Eye blinking

Head movement

Lip synchronization

Facial symmetry

A recent adaptive landmark-assisted neural network showed strong performance in detecting unseen manipulations.

5.4 Multimodal Detection

Multimodal systems combine:

Video analysis

Audio analysis

Text metadata

Behavioral patterns

These systems improve robustness and accuracy against advanced deepfakes.

6. Datasets Used in Deepfake Detection

6.1 FaceForensics++

One of the most popular datasets containing manipulated videos generated using various techniques.

6.2 Celeb-DF

Contains high-quality celebrity deepfake videos and is considered more challenging than older datasets.

6.3 DeepFake Detection Challenge (DFDC)

Released by Facebook and research partners to improve deepfake detection research.

6.4 MNW Dataset

A newer benchmark dataset designed to represent modern AI-generated media diversity.

7. Challenges in Deepfake Detection

7.1 Rapid Advancement of Generative AI

Modern AI models generate highly realistic videos with fewer detectable artifacts.

7.2 Dataset Bias

Some datasets lack diversity, causing poor model generalization.

7.3 Real-World Conditions

Compression, resizing, and low-quality uploads reduce detection accuracy. Researchers found that many models fail under real-world transformations.

7.4 Adversarial Attacks

Attackers intentionally modify deepfakes to bypass AI detectors.

7.5 Computational Complexity

Advanced detection systems require large GPU resources and high processing power.

8. PROPOSED METHODOLOGY

The proposed system uses CNN-based deep learning for video classification.

Steps:

1. Video Upload
2. Frame Extraction
3. Face Detection using OpenCV
4. Feature Extraction
5. CNN Classification
6. Prediction Output (Real/Fake)

Technologies Used

Python

TensorFlow

Keras

OpenCV

NumPy

9. Experimental Results

The proposed CNN-based model can achieve high accuracy on benchmark datasets when trained properly.

Parameter Result

Dataset Used Celeb-DF

Model XceptionNet

Accuracy 94%

Precision 92%

Recall 93%

Note: Accuracy may vary depending on dataset quality and preprocessing techniques.

10. Applications of Deepfake Detection

1. Social Media Verification
2. Cybercrime Prevention
3. Political Misinformation Detection
4. Digital Forensics
5. Banking and Fraud Prevention
6. News Verification Systems

11. Future Scope

Future research should focus on:

Real-time deepfake detection

Multimodal AI systems

Blockchain-based media verification

Explainable AI detection models

Generalized cross-dataset learning

Watermarking and authenticity verification systems

Researchers believe future solutions may combine AI detection with cryptographic verification methods.

12. CONCLUSION

Deepfake videos represent one of the biggest challenges in the modern digital world. AI-generated fake content threatens trust, privacy, cybersecurity, and social stability. Artificial Intelligence-based detection systems provide an effective solution for identifying manipulated videos through facial analysis, deep learning, and multimodal techniques. However, as generative AI evolves rapidly, existing detection systems face major challenges related to robustness, fairness, and generalization. Future research should focus on adaptive, explainable, and real-time detection methods capable of handling sophisticated synthetic media. Continuous improvement in AI-based forensic systems will play an important role in protecting digital authenticity and combating misinformation.