IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Deep Fake Model For Detecting Fake Videos And Audio In Branding And Misinformation Prevention

Dr. J. Sasidevi, Balasubramaniyam B, Kapilan S, Parthasarathi P

Head Of Department, Student, Student, Student

Department of Computer Science and Business Systems,
K.Ramakrishnan College of Engineering (Autonomous), Samayapuram, Trichy, India

Abstract: With the rapid advancement of artificial intelligence and generative technologies, the rise of deepfake audio and video has become a major concern for digital security, branding authenticity, and misinformation prevention. This project presents an automated system capable of detecting deepfake audio and video using advanced AI and Deep Learning techniques. The system analyzes facial expressions, lip synchronization, voice patterns, and spectral audio features to identify inconsistencies typical found in manipulated media. Deep Learning models such as CNNs and Transformer-based architectures are utilized for recognizing subtle manipulation artifacts.

Index Terms - Deepfake Detection, Artificial Intelligence, Deep Learning, CNN, LSTM, Misinformation Prevention, Digital Security.

I. INTRODUCTION

Artificial Intelligence (AI) is a powerful technology that enables machines to analyze complex patterns in audio and video content, making it essential for detecting deepfakes. In this project, AI plays a central role by examining facial expressions, speech characteristics, frame consistency, and visual–audio alignment to identify whether the media is real or manipulated. AI-driven algorithms process large amounts of multimedia data, detect anomalies, and flag suspicious content with high speed and accuracy. Deep Learning (DL), a specialized domain within AI, significantly enhances the system's capability to detect even the most sophisticated deepfake content. Using advanced neural networks like CNNs, LSTMs, and Transformer models, DL analyzes subtle cues such as lip–audio synchronization, micro-expressions, and frame-level irregularities.

II. LITERATURE SURVEY

The project reviews several methodologies. Alnabhan and Branco (2024) explored deep learning for fake news detection. Verma et al. (2023) investigated phishing website detection and deepfake audio analysis. Their research shows that SE-Enhanced 1D-CNN achieved an accuracy of 97.64% in audio detection. Wang and Long (2025) proposed a Global-Local Ensemble Detector. These studies form the foundation for our multimodal approach, which combines visual and auditory artifact analysis.

III. SYSTEM ANALYSIS

3.1 Existing System

Existing systems for identifying manipulated media are limited. Most current methods depend on traditional digital forensics like analyzing image noise or pixel distortions. These systems often focus on a single modality—either audio or video—and fail to detect complex, multi-layered fake media used in misinformation campaigns. Furthermore, they lack real-time processing and have high manual dependency.

3.2 Proposed System

The proposed system overcomes these limitations by utilizing deep learning and multimodal analysis. It processes both audio and video files through a streamlined pipeline: Input Data -> Feature Extraction (CNN/MFCC) -> Temporal Analysis (LSTM) -> Final Classification (Real/Fake). This approach improves adaptability and robustness against evolving deepfake techniques.

IV. SYSTEM REQUIREMENTS

The implementation requires an Intel processor (or AMD Ryzen 5), 8GB RAM, and 160GB Hard disk space. On the software side, Windows 11 OS, Python (programming), OpenCV (vision), TensorFlow/Keras (deep learning), and Flask (web deployment) are utilized.

V. MODULE DESCRIPTION

The project is divided into several key modules: (1) Data Collection & Dataset Preparation: Gathering genuine and manipulated samples. (2) Preprocessing & Feature Extraction: Resizing frames, normalizing audio, and extracting MFCCs/Visual features. (3) Deep Learning Model Development: Training CNN and GRU/LSTM networks. (4) Classification & Decision Fusion: Combining multi-modal outputs for a final verdict. (5) Reporting & Visualization: Displaying results through heatmaps and analysis reports.

VI. SYSTEM IMPLEMENTATION

Implementation focuses on real-time detection. Using 'cv2.VideoCapture', frames are extracted and processed. Facial landmarks are detected using Haar cascades. For audio, 'librosa' extracts spectral centroid and pitch variation. The backend, built on Flask/Streamlit, provides a user interface for uploading files and viewing analytics.

VII. CONCLUSION AND FUTURE ENHANCEMENT

7.1 Conclusion

The proposed Deep Fake Detection system provides a scalable solution for identifying manipulated media. By combining CNN-based video analysis and spectrogram-based audio classification, the system achieves an accuracy of approximately 85%. It successfully detects inconsistencies like facial distortions and irregular pitch variations.

7.2 Future Enhancements

Future work will focus on training models on larger datasets to enhance generalization. Integration with social media platforms for automatic detection and cloud deployment for large-scale access are also planned.

REFERENCES

[1] Babiker, N., 'Deepfake Audio Detection in Voice Authentication: A Spectral and CNN Based Comprehensive Review,' Engineering, Technology & Applied Science Research, vol. 15, 2025.
[2] Verma, K., et al., 'Advanced deep learning (DL) models for deepfake audio detection,' 2023.
[3] Tajrian, M., et al., 'A Review of Methodologies for Fake News Analysis,' IEEE, 2023.
[4] Wang, Y., and Long, W., 'Global-Local Ensemble Detector for AI Generated Fake News,' 2025.