Video Authenticity Verification in the Age of AI: A Cross Model Ensemble Against AI Engineered Visual Forgery Profiling

[1]Dr. P. S. Naveen Kumar, [2] Chebrolu Sailaja, [3]Ganji Venkata Sridevi, [4]Guggilam Venkata Naga Sai Kowsalya

[1]Assistant Professor, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

[2,3,4]U. G Student, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

ABSTRACT

Using cutting-edge computer vision and deep learning techniques, this study offers an AI-based system for detecting deepfake videos. Using CNNs, RNNs, and transfer learning models like InceptionV3 and ResNeXt, it uses GRU-based sequence analysis to extract spatiotemporal information from video frames. NLP is used to examine metadata for contextual comprehension, and preprocessing guarantees consistent input. Real and false videos are distinguished by a supervised classifier using softmax-based confidence rating. Real-time, comprehensible forecasts are made possible by the system's integration with a Flask web interface. By improving robustness, data augmentation achieves excellent accuracy and dependability on benchmark datasets for social media and forensic applications.

INTRODUCTION

Deepfake videos, which use artificial intelligence (AI) to produce extremely convincing fake footage, present significant ethical, societal, and legal concerns in the digital age. AI-based techniques are essential since traditional detection methods are inadequate. Facial and motion abnormalities are detected using CNNs, RNNs (particularly GRUs), and transfer learning models such as InceptionV3. Sequence models and Softmax classification improve accuracy, while preprocessing guarantees consistency. Real-time results and safe video uploads are possible with a Flask-based web application. The system, which was trained on balanced datasets, provides a scalable solution for forensics, research, and media monitoring by achieving high detection reliability that is proven by accuracy, precision, recall, and F1-score.

LITERATURE SURVEY

Using deep learning and ensemble-based models, recent research on video authenticity verification focuses on identifying AI-generated forgeries like deepfakes. To examine spatial and temporal irregularities in video frames, methods such as CNNs, RNNs, and transformers are used. Multiple architectures are combined in cross-model ensemble approaches to increase detection robustness against sophisticated forgery techniques. For improved generalization, researchers focus on feature fusion, frequency-domain analysis, and attention mechanisms. Despite advancements, there are still difficulties in identifying extremely realistic forgeries and guaranteeing scalability across many platforms and datasets.

EXISTING SYSTEM

Traditional digital forensics and metadata analysis methods are the mainstays of the current video authenticity verification system. To identify manipulation, these systems look for discrepancies in timestamps, file characteristics, and compression artifacts. To detect tampering, some methods employ pixel correlation and frame-level analysis. However, they frequently fall short against complex AI-generated material, like deepfakes. Conventional CNN-based classifiers are less effective across a variety of forgeries types since they were trained on small datasets. They frequently generate false positives and are not able to adjust to new AI models. Because of this, modern systems find it difficult to guarantee trustworthy authenticity verification in the age of sophisticated visual forgeries.

DRAWBACKS

- Less accuracy
- Feature analysis is less using LBP

PROPOSED SYSTEM

The procedure begins with gathering both actual and fake videos, which are then preprocessed using color conversion, resizing, and frame extraction. Pretrained CNNs (InceptionV3 or ResNeXt) are used to collect high-level spatial characteristics, which are then sequentially processed with GRU layers to capture temporal discrepancies. Overfitting is avoided and resilience is increased through data augmentation and dropout. A sigmoid layer is used to produce binary predictions from the model, which was trained using binary cross-entropy and the Adam optimizer. Confidence-based outcomes and real-time uploads are made possible with a Flask-based interface. To guarantee generalization and scalability, performance is assessed using accuracy, precision, recall, and F1-score on unseen films.

capabilities.

ADVANTAGES

- Lowers the cost of video campaigns.
- Deepfake technology can create better omnichannel campaigns.
- It can provide a hyper-personalised experience for customers.

SYSTEM ARCHITECTURE

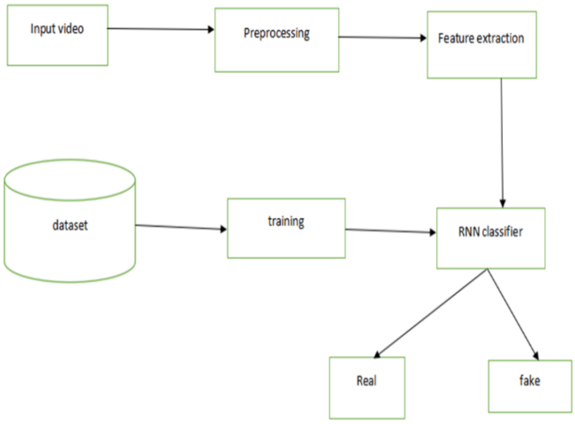
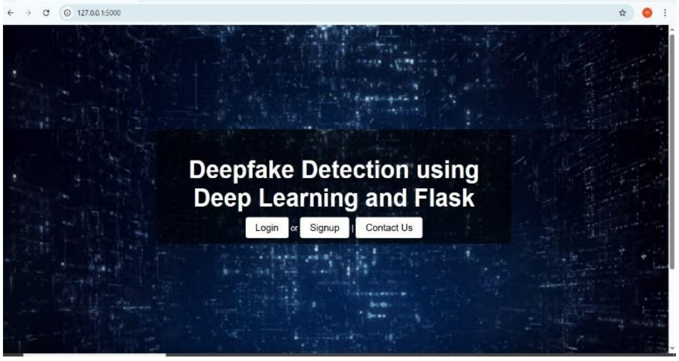


Figure:1 System Architecture

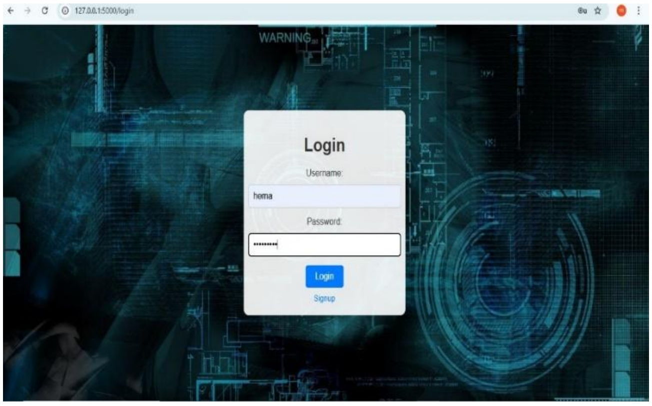
The first step in the process of detecting deepfake videos with an RNN classifier is preparing the input video, which involves extracting frames, identifying faces, and normalizing the data to guarantee consistency. In order to obtain spatial information, important aspects including motion dynamics, texture, and facial landmarks are subsequently retrieved, frequently utilizing CNNs. The RNN model learns the temporal irregularities typical of deepfakes from a tagged dataset of real and fake videos. By analyzing sequential frame attributes, the RNN may identify erratic motions or transitions. Lastly, the system uses deep learning frameworks like TensorFlow or PyTorch to classify videos as authentic or fraudulent, offering a dependable, scalable solution for media authenticity.

RESULTS



Deepfake Detection using Deep Learning and Flask

Figure2:Homepage

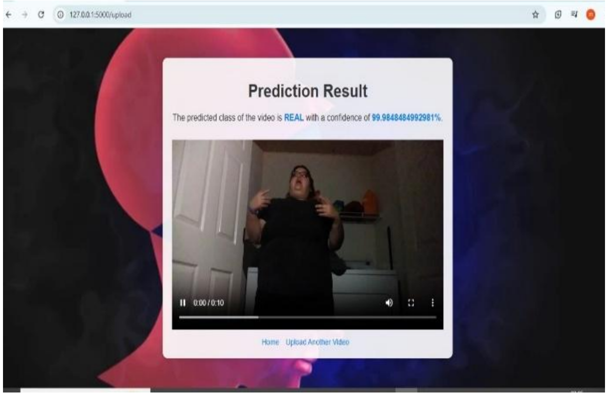


Login

Username: homa

Password: ••••••••

Figure 3: login page



Prediction Result

The predicted class of the video is REAL with a confidence of 99.9848484992981%.

Figure 4: Final result

CONCLUSION

In this paper we have presented a temporal-aware system to automatically detect deepfake videos. Our experimental results using a large collection of manipulated videos have shown that using a simple RNN we can accurately predict if a video has been subject to manipulation or not with as few as 2 seconds of video data. We believe that our work offers a powerful first line of defense to spot fake media created using the tools described in the paper. We show how our system can achieve competitive results in this task while using a simple pipeline architecture. In future work, we plan to explore how to increase the robustness of our system against manipulated videos using unseen techniques during training.

REFERENCES

[1] Naveen Kumar Polisetty, S., Sivaprakasam, T. & Sreeram, I. An efficient deep learning framework for occlusion face prediction system. *Knowl Inf Syst* **65**, 5043–5063 (2023). https://doi.org/10.1007/s10115-023-01896-5

[2] Chapala, H. (n.d.). Machine Learning based Bayesian Network Models for Reverse Engineering Data Optimization. International Conference on Edge Computing and Applications, ICECAA 2022 - Proceedings.

[3] Y. Li and S. Lyu, “Exposing DeepFake Videos By Detecting Face Warping Artifacts.”

[4] D. Guera and E. J. Delp, “Deepfake Video Detection Using Recurrent Neural Networks,” in Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2019.

[5] X. Yang, Y. Li, and S. Lyu, “Exposing Deep Fakes Using Inconsistent Head Poses,” in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2019.