



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Human Based Deep Fake Video Detection and Estimation using Region Convolutional Neural Networks

Mrs.Sree Lakshmi ,S.Likitha, P.Venu Sai Ram, Ch.Lakshmi Janardhan, M.Chaitanya

Department of ECE, Sri Vasavi Institute of Engineering & Technology (Autonomous), Nandamuru, Pedana, A.P, India

ABSTRACT: This Deepfake technology has become a significant threat to digital media integrity, enabling the creation of highly realistic yet fraudulent videos. To combat this, we propose a human-based deepfake video detection and estimation system using a Region-based Convolutional Neural Network (R-CNN) for precise facial manipulation detection. The system leverages Raspberry Pi Pico for lightweight processing and ESP32-CAM for real-time video capture and analysis. Our approach focuses on facial region anomalies, such as unnatural eye blinking, lip-sync mismatches, and texture inconsistencies, which are common in deepfake videos. The R-CNN model is trained on a dataset of real and deepfake videos to classify manipulated content accurately. The embedded hardware setup ensures low-cost, energy-efficient, and portable deepfake detection, making it suitable for edge computing applications. Experimental results demonstrate that our system achieves high detection accuracy while maintaining low computational overhead, providing a practical solution for real-world deployment in security, social media moderation, and forensic analysis.

KEYWORDS : Fake video detection, Res-Next CNN, LSTM

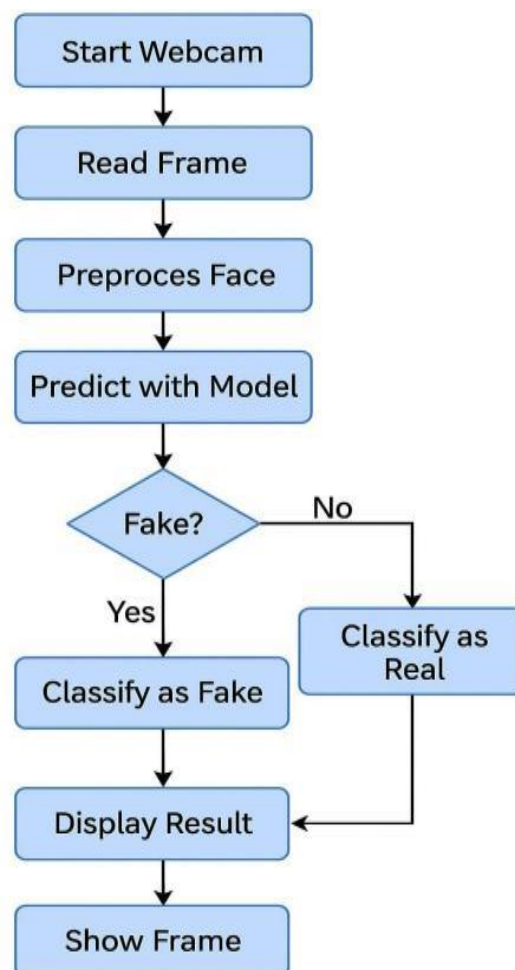
I. INTRODUCTION

The rise of deepfake technology has introduced significant challenges in digital media authenticity, enabling the creation of highly realistic yet fraudulent videos that can spread misinformation, manipulate public opinion, and compromise security. Traditional detection methods often rely on computationally intensive deep learning models that require high-end hardware, making them impractical for real-time, edge-based applications. To address these limitations, this project proposes a human-based deepfake video detection and estimation system using a Region-based Convolutional Neural Network (R-CNN) optimized for deployment on low-power embedded devices, specifically the Raspberry Pi Pico and ESP32-CAM. The system focuses on identifying subtle facial anomalies—such as unnatural eye movements, inconsistent lip-syncing, and texture irregularities—that are common in deepfake videos. By leveraging region-specific analysis, our approach improves detection accuracy while maintaining computational efficiency, making it suitable for real-world applications like social media moderation, forensic analysis, and secure authentication. The integration of Raspberry Pi Pico ensures energy-efficient processing, while the ESP32-CAM facilitates real-time video capture and analysis without relying on cloud-based solutions. This project bridges the gap between advanced deepfake detection and practical, low-cost hardware implementation, offering a scalable solution to combat the growing threat of synthetic media manipulation. Future enhancements include expanding the model's robustness with larger datasets and incorporating multimodal detection (audio-visual analysis) for improved reliability.

II. PROPOSED SYSTEM

The proposed system aims to detect and estimate the authenticity of human-based deepfake videos using a Region-Based Convolutional Neural Network (R-CNN) implemented on a Raspberry Pi pico setup with a camera, SD card, and display. This approach leverages deep learning techniques to analyze facial distortions, inconsistencies, and synthetic features commonly found in deep fake videos. The first stage of the system involves capturing or uploading video footage using the esp32 camera or an external source. The video frames are extracted and pre-processed, where each frame undergoes noise reduction, contrast enhancement, and face region detection using OpenCV and Haar cascades. The R-CNN model is then applied to these frames

to detect manipulated regions by analyzing fine-grained facial details, such as abnormal skin textures, unnatural eye movements, and misaligned facial expressions. In the second stage, the system extracts feature maps from the detected face regions using a pre-trained deep learning model (e.g., ResNet, VGG-16, or EfficientNet). These extracted features are compared against a dataset of authentic and synthetic videos to determine the likelihood of the video being manipulated. The Region Proposal Network (RPN) within the R-CNN generates multiple bounding boxes around possible deepfake regions, refining the detection accuracy. The third stage involves real-time classification of the analyzed video. The system assigns a deepfake probability score to each frame and aggregates the results to classify the entire video as real or fake. If the probability exceeds a predefined threshold, the system triggers an alert message on the display, notifying the user about the potential deepfake. The Raspberry Pi Pico serves as an energy-efficient processing unit, while the ESP32-CAM captures and streams video for real-time analysis, eliminating dependency on cloud infrastructure. This embedded hardware integration ensures low-cost, portable, and scalable deepfake detection, making it suitable for applications such as social media moderation, forensic investigations, and secure authentication systems. Future enhancements include adversarial training to improve robustness against evolving deepfake techniques and multimodal detection combining visual and audio analysis for higher reliability. This combination of region-specific deep learning and embedded system optimization provides a novel, efficient, and practical solution for real-world deepfake detection



III. EXPERIMENTAL RESULTS

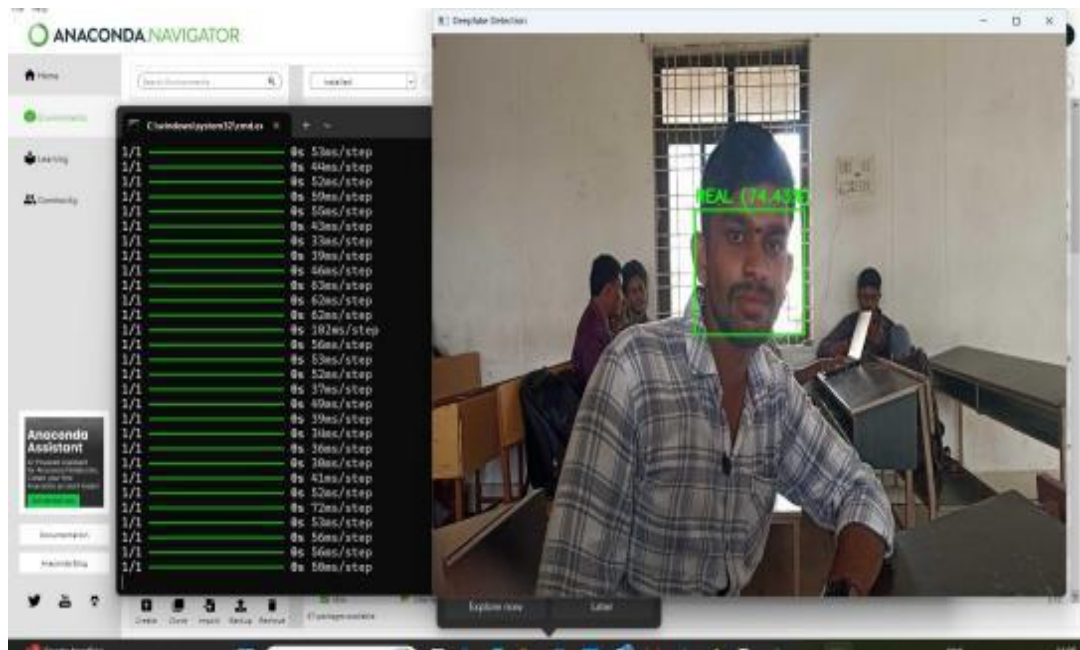


Fig.1.REAL VIDEO ANNALYSATION



Fig.2.FAKE VIDEO ANNALYSATION

IV. CONCLUSION

The human-based deepfake video detection and estimation system using Region-based Convolutional Neural Network (R-CNN) and Raspberry Pi provides an efficient and automated approach to identifying manipulated videos in real-time. By leveraging AI-driven analysis, the system detects subtle facial inconsistencies, unnatural movements, and pixel anomalies, which are commonly associated with deepfake videos. The integration of a camera, SD card, and display ensures seamless data processing and visualization, making it accessible for practical applications.

This solution plays a crucial role in digital forensics, media authentication, and cybersecurity, as deepfake videos pose a serious threat to misinformation, identity fraud, and online safety. By implementing real-time deepfake detection on an edge device like Raspberry Pi, the system provides a cost-effective and portable solution that can be deployed in security agencies, social media platforms, and digital content verification systems. Despite its advantages, challenges such as computational limitations, real-time processing efficiency, and evolving deepfake techniques remain. However, future advancements in optimized AI models, enhanced GPU acceleration, and cloud-based integration can further improve accuracy and performance. In conclusion, the R-CNN-based deepfake detection system represents a significant step forward in combating digital manipulation, ensuring trust and authenticity in visual content across multiple domains.

REFERENCES

1. *FaceApp*, Jan. 2021, [online] Available: <https://www.faceapp.com/>.
2. *FakeApp*, Jan. 2021, [online] Available: <https://www.fakeapp.org/>.
3. G. Oberoi, Exploring DeepFakes, Jan. 2021, [online] Available: <https://goberoi.com/exploring-deepfakes-20c9947c22d9>.
4. J. Hui, How Deep Learning Fakes Videos (Deepfake) and How to Detect it, Jan. 2021, [online] Available: <https://medium.com/how-deep-learning-fakes-videos-deepfakes-and-how-to-detect-it-c0b50fbf7cb9>.
5. I. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, et al., "Generative adversarial nets", *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 2, pp. 2672-2680, 2014.
6. G. Patrini, F. Cavalli and H. Ajder, "The state of deepfakes: Reality under attack", 2018, [online] Available: <https://s3.eu-west-2.amazonaws.com/rep2018/2018-the-state-of-deepfakes.pdf>.
7. J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt and M. Niessner, "Face2Face: Real-time face capture and reenactment of RGB videos", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 2387-2395, Jun. 2016.
8. J.-Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks", *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 2242-2251, Oct. 2017.
9. S. Suwajanakorn, S. M. Seitz and I. K. Shlizerman, "Synthesizing Obama: Learning lip sync from audio", *ACM Trans. Graph.*, vol. 36, no. 4, pp. 95, 2017.
10. L. Matsakis, Artificial Intelligence is Now Fighting Fake Porn, Jan. 2021, [online] Available: <https://www.wired.com/story/gfycat-artificial-intelligence-deepfakes/>.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details