WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Realtime Deep-fake Detection on Video Streams

Abhinav[1], Ms. Shivangi Singh[2], Ayush Bharadwaj[3]

[1]Computer Science and Engineering DepartmentGalgotias University, Greater Noida,Gautam Buddh Nagar, Uttar Pradesh, yabhinav597@gmail.com
[2]Computer Science and Engineering DepartmentGalgotias University, Greater Noida,Gautam Buddh Nagar, Uttar Pradesh, shivangisingh@galgotiasuniversity.edu.in
[3]Computer Science and Engineering DepartmentGalgotias University, Greater Noida,Gautam Buddh Nagar, Uttar Pradesh, ayushbharadwaj754@gmail.com

Abstract:

In recent years, the growing use of deepfake technology has introduced new risks to the credibility of video content, raising serious concerns around misinformation and digital deception. While many detection tools have been proposed, most struggle to operate effectively in real-time, especially under streaming constraints. This study sets out to improve the reliability and responsiveness of deep-fake detection in live video feeds. A key issue in current research is the lack of efficient and adaptable models that can maintain accuracy without sacrificing speed. Addressing this shortfall is vital for protecting real-time communication platforms and ensuring the integrity of visual media. The outcomes of this work are expected to contribute meaningfully to the development of faster, smarter, and more resilient detection systems capable of responding to rapidly evolving threats.

Keywords: *Deepfake, video forensics, real-time detection, neural networks, streaming, digital manipulation.*

I. Introduction

Recent progress in generative technologies has made it easier to fabricate realistic videos that can deceive both humans and automated systems. These videos, often called deepfakes, pose challenges to public trust, privacy, and digital security. While detection tools exist, most are suited for offline analysis. This paper concentrates on methods developed for real-time detection, which are essential in settings like live broadcasting, video conferencing, and surveillance. We explore key algorithms, real-time optimizations, and deployment concerns that define the effectiveness of these systems.

II. Motivation

Traditional detection techniques are often not fast enough for real-time use. As deepfakes become more advanced and accessible, there is a growing demand for systems that can flag manipulate content instantly during video playback or streaming. This review addresses the urgent need to identify fake content on-the-fly without relying on post-processing. By doing so, we aim to prevent misuse before it spreads and to maintain trust in video communication platforms.

III. Literature Review

In recent years, the rise of deepfake content has led to a surge of interest in methods to detect manipulated videos. Researchers have approached this challenge from multiple angles, using both spatial and temporal patterns to differentiate between real and fake content.

One commonly used architecture is the Xception model, known for its strong detection capabilities, especially when dealing with traditional face-swap manipulations. However, the model's complexity and size often make it unsuitable for low-latency or real-time scenarios [1]. To address this, lightweight models like Mobile Net and Efficient Net-Lite have been proposed. These models trade some accuracy for faster processing speeds, making them a better fit for devices with limited computational power [2].

Beyond static frame analysis, time-dependent architectures like RNNs and LSTMs have been applied to track subtle motion inconsistencies across frames—such as unnatural blinking or mouth movement—which are often overlooked in frame-based models [3]. By combining CNNs for feature extraction with temporal models like GRUs, hybrid frameworks have shown promise in real-time contexts, offering a good compromise between speed and performance [4].

In more recent studies, transformer-based models have also been introduced, allowing for better handling of long-range dependencies across video sequences. While powerful, these models typically require high-end hardware and are currently more suited to offline analysis [5].

Another strategy seen in modern research is the use of attention mechanisms that dynamically focus on potentially manipulated regions within a video frame. These mechanisms allow the model to give greater weight to areas like eyes, lips, or skin texture anomalies—areas commonly altered in deepfakes [6].

To further improve detection, optimization techniques like pruning, quantization, and knowledge distillation have been incorporated to shrink model size and speed up inference while maintaining accuracy [7]. Additionally, some techniques use Error Level Analysis (ELA) to reveal traces of editing by highlighting inconsistent pixel-level compression artifacts [8].

A few recent contributions have explored dual-branch architecture with adversarial training to help models generalize across unseen types of forgeries. This is especially important because many models perform well on one dataset but fail to detect new or emerging deepfake methods [9].

All these advancements rely heavily on training data. Public datasets like Face Forensics++, DFDC, Celeb-DF, and Deeper Forensics have become benchmarks in the field. However, models often struggle to generalize across them due to differences in quality, compression, actor diversity, and manipulation techniques [10].

Despite notable progress, many open issues remain, including handling deepfakes that target body movements or audio, ensuring cross-platform compatibility, and reducing false positives. These challenges suggest that future research must focus not only on improving detection accuracy but also on ensuring adaptability in real-world environments [11].

IV.Results and Insights

Evaluating deep-fake detection models involves understanding both their effectiveness and practicality, especially in real-time environments. While many approaches have achieved high accuracy under controlled conditions, their performance often varies when applied to live video streams, compressed footage, or unseen manipulations.

For instance, lightweight models such as Mobile Net and Efficient Net-Lite offer fast inference and are suitable for edge devices. However, they sometimes miss subtle forgery cues present in high-quality deepfakes. On the other hand, LSTM-based temporal models demonstrate a strong ability to detect inconsistencies across video frames, such as unnatural blinking or inconsistent facial expressions. Yet, these models typically require more computational resources, making them less feasible for real-time deployment without GPU support.

The following table provides a summarized comparison of different models evaluated in recent studies, based on their accuracy, speed, and real-time readiness:

Table1: Comparison of Real-Time Deepfake Detection Models

Model	Accuracy (%)	Real Time Capable
MobileNet + Attention	91.2	Yes
EfficientNet-Lite	93.3	Yes
CNN + LSTM Hybrid	96.5	Partial
XceptionNet	95.1	No
Transformer-Based Detector	94.7	No

Key Observations:

- MobileNet and EfficientNet-Lite are ideal for real-time environments but slightly sacrifice detection precision.
- CNN-LSTM hybrids perform better at spotting subtle manipulations but often lag in processing time.
- Transformers offer high robustness and accuracy across manipulation types but are too slow for real-time tasks unless heavily optimized.
- XceptionNet, although highly accurate, is generally unsuitable for live video detection without powerful hardware acceleration.

These findings highlight the persistent trade-off between model complexity and speed. In real-world applications, especially on mobile or embedded systems, lightweight models remain the most practical choice. However, combining speed with improved accuracy — possibly through hybrid or ensemble approaches — continues to be a promising direction for future research.

V. Conclusion

As deepfakes become more advanced, the demand for trustworthy, real-time detection systems grows stronger. This review has outlined the progress and gaps in existing methods, emphasizing the importance of lightweight architecture, intelligent optimization, and robust design. To stay ahead of

evolving threats, future models must be fast, generalizable, and adaptable to real-world constraints. By closing the latency gap without compromising reliability, we move closer to a future where live video integrity can be safeguarded in real time.

References

1. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). FaceForensics++: Learning to Detect Manipulated Facial Images. In ICCV.
2. Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). MesoNet: A Compact Facial Video Forgery Detection Network. In WIFS.
3. Guera, D., & Delp, E. J. (2018). Deepfake Video Detection Using Recurrent Neural Networks. In AVSS.
4. Sabir, E., AbdAlmageed, W., Masi, I., & Natarajan, P. (2019). Recurrent Convolutional Strategies for Face Manipulation Detection in Videos. In CVPR Workshops.
5. Dang, H., Liu, F., Stehouwer, J., Liu, X., & Jain, A. K. (2020). On the Detection of Digital Face Manipulation. In CVPR.
6. Wan, D., Cai, M., Peng, S., Qin, W., & Li, L. (2023). Deepfake Detection Based on Dual-Branch Data Augmentation and Attention Mechanism. Applied Sciences, 13(8313).
7. Korshunov, P., & Marcel, S. (2018). Deepfakes: A New Threat to Face Recognition? Assessment and Detection. arXiv.
8. Rafque, R., Gantassi, R., Amin, R., Frnda, J., Mustapha, A., & Alshehri, A. H. (2023). Deep Fake Detection and Classification Using Error-Level Analysis and Deep Learning. Scientific Reports.
9. Guarnera, L., Giudice, O., & Battiato, S. (2020). DeepFake Detection by Analyzing Convolutional Traces. In CVPR Workshops.
10. Li, Y., Chang, M. C., & Lyu, S. (2020). Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics. In CVPR.
11. Jiang, H., Zhang, Y., & Ma, X. (2020). DeeperForensics-1.0: A Large-Scale Dataset for Real-World Face Forgery Detection.