iJRASET
International Journal For Research in
Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 13 **Issue:** XI **Month of publication:** November 2025

**DOI:** https://doi.org/10.22214/ijraset.2025.75243

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

# A Comprehensive Review of EfficientNet-Lite and Attention-Driven Deepfake Detection for Real-Time Video Streams

Rakhi Meshram[1], Suyog Talape[2], Manasi Thorat[3], Rutik Vidhate[4]
*Artificial Intelligence & Data Science AISSMS Institute of Information Technology* Pune, India

***Abstract: Deepfake technology, powered by advanced AI mod- els like GANs, poses serious threats by generating hyper-realistic fake videos, which are increasingly used in misinformation, fraud, and cybercrime. Despite rapid progress in deepfake detection methods, challenges remain in achieving high accuracy with real- time processing and robustness to evolving synthesis techniques. Recent works have leveraged various deep learning architectures, such as CNNs, LSTMs, Transformers, and hybrid models, with accuracy ranging from 85 to 98 percent on benchmark datasets. For example, EfficientNet-based models combined with attention mechanisms have shown promising detection accuracy above 95 percent, while multimodal approaches incorporating audio and visual features further improve robustness. However, current so- lutions often struggle with computational efficiency and adapting to new deepfake generation methods in live video streams. This survey highlights these research gaps and critically assesses state- of-the-art AI-based real-time deepfake detection systems, includ- ing techniques using corneal reflection analysis, active forensic probing, and multi-scale attention networks. We conclude by outlining future directions focused on improving detection speed, adversarial robustness, and deploying practical solutions for live video surveillance and conferencing environments.***

***Index Terms: Deepfake detection, real-time video analysis, Ef- ficientNet, attention mechanism, multimodal fusion, adversarial robustness, AI security.***

## I. INTRODUCTION

The proliferation of deepfake technology has emerged as one of the most pressing challenges in the digital era, funda- mentally threatening the integrity of visual media and public trust in digital content.[3][1] The rapid advancement of arti- ficial intelligence, particularly through generative adversarial networks (GANs) and sophisticated deep learning algorithms, has democratized the creation of highly realistic synthetic videos that seamlessly replicate human appearances and be- haviors.[9] These AI-generated videos, commonly known as deepfakes, have evolved from rudimentary face-swapping tools to sophisticated manipulation techniques capable of producing content that is increasingly difficult to distinguish from authen- tic footage.[11][5]

The societal implications of deepfake technology extend far beyond entertainment applications, posing significant threats across multiple domains. The malicious deployment of deep- fakes has been documented in numerous high-profile cases, including political manipulation campaigns, financial fraud schemes, and identity theft operations. A particularly alarm- ing example occurred in 2024, when a company in Hong Kong fell victim to an elaborate scam involving deepfake video conferencing, resulting in a 25 million Dollar loss when the CFO was deceived by synthetic representations of company leadership. Similarly, manipulated videos of political figures, such as the fabricated footage of Ukrainian Presi- dent Volodymyr Zelenskyy purportedly capitulating to Russian demands, demonstrate the technology's potential to influence democratic processes and spread misinformation. The prolifer- ation of deepfakes in educational settings, including incidents where students created synthetic nude images of classmates, further underscores the technology's capacity for harassment and psychological harm.[3]

The detection of deepfake content has consequently become a critical research priority, driving the development of sophis- ticated countermeasures that leverage the same artificial intel- ligence principles used in their creation.[1][9] Early detection approaches relied heavily on traditional forensic techniques, focusing on identifying visual inconsistencies such as un- natural lighting, irregular facial expressions, or physiological anomalies like abnormal blinking patterns.[14] However, as deepfake generation techniques have advanced, these rudi- mentary detection methods have proven inadequate against high-quality synthetic content, necessitating the development of more sophisticated AI-driven detection frameworks.[11]

Contemporary deepfake detection research has converged around several key methodological approaches, each address- ing different aspects of the synthetic content identification challenge.[10] Spatial feature extraction techniques, primarily implemented through convolutional neural networks (CNNs), focus on identifying pixel-level artifacts and facial inconsisten- cies within individual video frames. [8]These approaches have demonstrated particular efficacy when enhanced with prepro- cessing techniques such as unsharp masking, which amplifies high-frequency edge artifacts commonly present in manip- ulated content.[11] The EfficientNet family of architectures has emerged as particularly promising for these applications, offering an optimal balance between computational efficiency and detection accuracy across multiple variants from B0 to B4.[12]

Temporal modeling approaches represent another critical di- mension in deepfake detection, recognizing that video content inherently contains sequential dependencies that can reveal manipulation artifacts across time.[9] These methods typically employ recurrent neural networks (RNNs), including Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), to analyze frame sequences and identify tempo- ral inconsistencies such as unnatural head movements, irregu- lar facial transitions, or physiologically implausible behavioral patterns.[10] The integration of spatial CNN features with temporal sequence modeling has proven particularly effective, with hybrid architectures demonstrating superior performance compared to single-modality approaches.[8]

Recent innovations in deepfake detection have also explored active forensics methodologies, particularly relevant for real- time video conferencing applications. These approaches lever- age environmental cues, such as corneal reflections of dis- played patterns, to verify the authenticity of video participants. [15]Such methods represent a paradigm shift from passive content analysis to active verification, requiring the deliberate injection of authentication signals into the communication channel. While computationally intensive, these techniques of- fer unique advantages in scenarios where traditional detection methods may be circumvented by increasingly sophisticated generation algorithms.[4]

The computational and deployment challenges associated with deepfake detection systems have driven significant re- search into real-time processing capabilities and edge com- puting applications.[2] The necessity for immediate content verification in streaming platforms, social media feeds, and live video conferencing has highlighted the importance of lightweight detection models that can operate under strict latency constraints without sacrificing accuracy.[7] This has led to the development of optimized architectures that lever- age techniques such as model quantization, pruning, and distributed inference to enable deployment on resource- constrained devices.[1]

Despite substantial progress in detection methodologies, several critical challenges persist in the field. Cross-dataset generalization remains a significant concern, as models trained on specific deepfake generation techniques often fail to iden- tify content created using alternative methods or newer generation algorithms.[5][7] The rapid evolution of synthetic content generation, including the emergence of diffusion- based techniques and neural talking heads, continuously chal- lenges existing detection frameworks and necessitates ongo- ing model adaptation.[3][11]Additionally, the computational requirements for real-time detection continue to pose deploy- ment challenges, particularly in scenarios requiring analy- sis of high-resolution video streams or multiple concurrent feeds.[1][2] This survey provides a comprehensive analysis of con- temporary deepfake detection methodologies, with particular emphasis on EfficientNet-based spatial feature extraction tech- niques, temporal sequence modeling approaches, and real- time deployment considerations. Through systematic examination of recent research contributions, we identify key technological advances, persistent challenges, and promising directions for future development in this rapidly evolving field. The anal- ysis encompasses both theoretical foundations and practical implementation considerations, offering insights into the trade- offs between detection accuracy, computational efficiency, and deployment feasibility that characterize current state-of-the-art approaches.

## II. EXISTING DEEPFAKE DETECTION SYSTEMS

Deepfake detection has seen a proliferation of diverse ap- proaches, broadly categorized into spatial-only convolutional feature extractors, temporal sequence models, hybrid spa- tiotemporal architectures, active forensics for live scenarios, and industry-guides aggregating best practices. Early systems leveraged CNN backbones to extract frame-level artifacts, often augmented with preprocessing filters to amplify high- frequency inconsistencies. Subsequent works integrated RNNs or temporal convolutions to capture sequential artifacts across frames. More recent efforts address live streaming and video conferencing through active probing and corneal reflection analysis, while industry whitepapers provide practical deploy- ment guidelines across real-world pipelines.

Despite these advances, existing systems often share com- mon limitations: reliance on a single dataset or manipulation type leading to overfitting; insufficient temporal context mod- eling; high computational demands hindering real-time de- ployment; and lack of standardized evaluation protocols. The following subsections detail individual contributions, high- lighting each paper’s proposed system, its core methodology, and noted limitations.

khaled et al. propose an EfficientNet-based CNN aug- mented with unsharp masking preprocessing to enhance subtle high-frequency edge artifacts in GAN-generated video frames. Their method applies an unsharp masking filter to amplify these artifacts before feeding frames into EfficientNet variants B0 to B4 trained on the FaceForensics dataset. This approach boosts detection accuracy to above 95 percent but tends to overfit on high-quality manipulated datasets, limiting gener- alization across other datasets. The system focuses on binary classification of real versus fake frames using the network's dense top layers, achieving strong spatial feature extraction but struggling with cross-dataset robustness.

Chaudhauri et al. develop a lightweight CNN model embedded within an Android remote control app for real-time, on-device deepfake detection. Their pipeline involves pruning and quantizing CNNs to fit limited mobile resources, capturing frames via remote desktop streaming, and using threshold- based filtering to alert users upon fake detection. While cloud- based detection services offer higher accuracy, this mobile- oriented solution prioritizes low-latency local inference at the cost of reduced accuracy and lacks temporal sequence modeling capabilities, relying solely on single-frame CNN classification.

Geetha et al. propose a hybrid deepfake detection ar- chitecture combining EfficientNet-B3 for spatial feature ex- traction and Temporal Convolutional Networks (TCNs) for modeling long-range temporal dependencies using causal di- lated convolutions. This approach stacks feature sequences extracted per frame and applies TCNs with residual blocks, addressing vanishing gradient and parallelism issues inherent in CNN-LSTM hybrids. Despite providing improved temporal reasoning for video-level classification, the model suffers increased computational complexity and latency, with evalua- tions conducted on a limited dataset.

Reddy et al. design an AI/ML framework integrating multiple CNN and RNN modules in a modular pipeline for face-swap deepfake detection. Their paper surveys state-of- the-art CNN, RNN, GAN, and hybrid methods before propos- ing a combined pipeline that extracts features through CNNs, analyzes temporal consistency via RNNs, and aggregates outputs with majority voting. Although comprehensive, the framework lacks detailed experimental implementation and quantitative performance metrics, making its practical efficacy uncertain.

Guo et al. introduce an active forensic deepfake detec- tion method targeting real-time video conferencing scenar- ios. The system displays randomized binary screen patterns during calls, captures corneal reflections of these patterns in participants' eyes via camera video frames, and computes cross-correlation with expected patterns to verify authenticity. This proactive challenge-response shifts from passive frame analysis to active authentication but depends on controlled lighting conditions, camera quality, and user consent, limiting broad applicability.

Paravision Inc. provide an industry practical guide out- lining a multi-layered deepfake detection pipeline combining artifact detectors, biometric signal analysis, and human-in-the- loop review processes. The guide synthesizes academic and commercial insights into a deployable enterprise framework emphasizing sequential detection stages and continuous model updating. However, it relies heavily on proprietary tools and offers limited quantitative benchmarks, focusing more on operational guidelines than novel model development.

Arkachari et al. propose a distributed microservice archi- tecture to scale real-time deepfake detection for live stream- ing platforms. They decompose the detection pipeline into microservices handling face detection, preprocessing, CNN inference, and decision-making, orchestrated via Kubernetes with autoscaling based on stream load. This design improves system scalability and fault tolerance but raises operational complexity and introduces network latency that may impact detection timeliness in live environments.

Abhinav et al. develop a frame-level CNN detection method using sliding window ensemble voting for deepfake detection in streaming video feeds. Their approach classifies individual frames with a compact CNN and aggregates predic- tions over windows of frames, reducing noisy frame-by-frame predictions. While achieving low latency and simplicity, the method remains vulnerable to transient false positives due to limited temporal smoothing and context.

Sangar and Rajasekar apply EfficientNet-LITE combined with kernel-enhanced Support Vector Machines (KE-SVM) for optimized classification, demonstrated in a plant dis- ease detection domain. Their system extracts features using EfficientNet-LITE and refines misclassified samples through KE-SVM kernel iterations to improve accuracy in diverse environments. Though not directly applied to deepfake detec- tion, this work showcases EfficientNet-LITE's effective feature extraction suitability and SVM adaptability for classification refinement.

Harshit et al. propose a hybrid model that integrates EfficientNet-B0 for spatial encoding and Gated Recurrent Unit (GRU) networks for temporal sequence analysis in deepfake detection. This combination targets efficient real-time video detection by leveraging EfficientNet-B0's performance and GRU's lightweight recurrence compared to heavier LSTM models. Although promising for latency-sensitive applications, the model's evaluation is limited to specific manipulation types, with potential risks of underfitting on complex se- quences.

Sahu presents an end-to-end deepfake detection system using ResNeXt CNN for multi-scale spatial feature extraction followed by Long Short-Term Memory (LSTM) networks for temporal classification. The system detects faces, processes frames through ResNeXt, sequences features into LSTM, and outputs a real/fake decision via sigmoid activation. Despite improved accuracy from spatiotemporal modeling, the system incurs high computational cost and latency, restricting deploy- ment on low-power or real-time devices.

Alrashoud offers a comprehensive survey categorizing audiovisual deepfake detection approaches into visual-based, audio-based, and multimodal techniques, emphasizing real- world generalizability and robustness. The survey critiques existing methods on their challenges in detecting low-quality compressed videos, adversarial attacks, and lacking standard benchmarks. It calls for unified evaluation metrics and more robust models to address evolving generation tactics in deep- fake videos.

Sunil et al. provide a detailed survey of autonomous deepfake detection methods covering CNNs, RNNs, trans- formers, GAN-based anomaly detection, and multimodal fu- sion models for image, video, and audio deepfakes. They discuss challenges in model generalization, adversarial defense, computational cost, and the need for large labeled datasets. Their evaluation stresses the importance of explainable AI techniques and multi-domain applicability for advancing au- tonomous detection frameworks.

Vivekananda et al. develop a highly accurate deepfake image detection model based on EfficientNetV2-B2 with en- hanced preprocessing, automated data balancing, and tailored regularization. Their system employs batch normalization, combined L1-L2 regularization, dropout with seed control, and advanced data augmentations to effectively distinguish real from StyleGAN-generated fake images, achieving 99.885 Precent accuracy on a large balanced dataset. However, the model's computational intensity and lack of adversarial ro- bustness assessment limit its practical deployment

Chorage et al. design a hybrid deep learning deepfake detection system combining CNN and RNN modules to an- alyze both spatial and temporal features from image/video frames. Their pipeline includes data preprocessing steps like normalization and augmentation and is trained on multiple benchmark datasets. While achieving 92.5 percent image and

88.2 percent video classification accuracy, the system faces issues with out-of-distribution generalization, computational expense, and insufficient real-time performance.

These research papers collectively provide a comprehensive overview of current deepfake detection techniques, highlight- ing the rapid evolution and challenges in this domain. The studies survey a wide range of detection methods includ- ing visual, audio, and multi-modal approaches, leveraging advanced deep learning architectures such as CNNs, RNNs, transformers, and hybrid models. They emphasize the impor- tance of robust preprocessing, data balancing, and tuning of architectures like EfficientNet to achieve high accuracy and generalizability. Despite notable progress, limitations remain including vulnerability to adversarial attacks, difficulties in detecting low-quality or compressed deepfakes, and challenges in real-time deployment.

## III. DATASETS

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks . . .". Instead, try "R. B. G. thanks. . .". Put sponsor acknowledgments in the unnumbered footnote on the first page.

## IV. METHODOLOGIES AND ARCHITECTURES

Contemporary deepfake detection systems draw upon a common set of methodological building blocks—spatial convolutional feature extractors, temporal sequence mod- els, and hybrid architectures that fuse both—with emerging approaches incorporating active forensics and deployment- oriented pipelines. Below is an organized overview of the primary architectures and techniques employed across the surveyed papers.

Spatial Convolutional Networks primarily use convolutional neural networks (CNNs) to extract spatial features from in- dividual video frames, focusing on pixel-level inconsistencies

TABLE I

COMPARISON OF DEEPFAKE DETECTION DATASETS

| **Dataset** | **Paper(s)** | **Size & Composition** | **Manipulations** | **Real:Fake** | **Notes** |
|---|---|---|---|---|---|
| FaceForensics++ | R. Khaled *et al.* "Boosting ..." [1]; S. Geetha *et al.* | 1,000 videos; 100 pristine +100 per type | Face2Face, FaceSwap, NeuralTextures, DF | 50:50 | Controlled com- pressions [1][3] |

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

| | | | | | |
|---|---|---|---|---|---|
| | "Deepfake Detection ..." [3]; K. V. A. Reddy *et al.* "De- velopment of AI/ML ..." [4]; G. Harshit *et al.* "Deepfake Detection ..." [10] | | | | |
| DFDC | S. Chaudhauri *et al.* "AI-Powered ..." [2]; K. V. A. Reddy *et al.* [4]; M. Arkachari *et al.* "The State of Live ..." [7]; Abhinav *et al.* "Realtime ..." [8]; C. Sahu "Deepfake Detection System" [11] | 100,000 videos; diverse sources | Multiple GANs | 75:25 | In-the-wild real- ism [2][7] |
| Celeb-DF | K. V. A. Reddy *et al.* [4]; C. Sahu [11] | 5,000 celeb videos | FaceSwap | 50:50 | Demographic di- versity [4] |
| Custom Streaming | H. Guo *et al.* "Detection of Real-Time ..." [5]; M. Arkachari *et al.* [7] | 500 live-call clips | Active-probing patterns | 60:40 | Corneal reflection analysis [5] |
| Android App Stream | S. Chaudhauri *et al.* [2] | 2,000 frames captured on Android | Frame-level fakes | 80:20 | On-device infer- ence [2] |
| Mixed Benchmark | K. V. A. Reddy *et al.* [4] | Combined FF++, DFDC, Celeb-DF | All above | Balanced | Modular protocol [4] |

and facial artifacts.[1] Techniques such as unsharp masking are applied to emphasize subtle edge artifacts before classification. EfficientNet variants (B0 to B4) and ResNeXt are popular backbone models offering a good trade-off between accuracy and computational efficiency.[11] Lightweight CNNs designed through pruning and quantization enable frame-level detection on resource-constrained devices like mobile phones, though often with reduced accuracy.[2]

Temporal Sequence Modeling addresses the inherently se- quential nature of videos by capturing temporal dependencies and inconsistencies across frames. Recurrent neural networks such as LSTMs and GRUs are widely used to analyze frame sequences, identifying unnatural head movements, facial tran- sitions, and physiologically implausible patterns.[10] Alterna- tive approaches use Temporal Convolutional Networks (TCN) with causal dilated convolutions to model long-range temporal dependencies without recurrence. Recent innovations include combining temporal facial pattern mechanisms with self- attention for robust long-range temporal coherence analysis.[3] Hybrid and Ensemble Architectures integrate spatial and temporal models to leverage strengths of both domains. These architectures commonly combine CNN-based spatial feature extraction with RNN or TCN temporal sequence modeling in modular pipelines.[4] Models like EfficientNet with GRU or ResNeXt with LSTM provide enriched video-level fake or real classification through fused feature-level or decision-level aggregation. Ensemble methods use sliding-window voting or majority voting across frame classifiers to achieve balanced accuracy and robustness.[8]

Active Forensic Methods represent a paradigm shift from passive video content analysis by incorporating proactive challenge-response authentication mechanisms. For example, detecting corneal reflections of randomized screen patterns displayed during video calls enables verification of real partic- ipants versus synthetic manipulations. While computationally intensive and requiring controlled conditions, these methods add a layer of security by validating user authenticity in real- time interactive environments.[5]

Deployment-Oriented Pipelines focus on scalability, latency, and practical applicability. Distributed microservice architec- tures decompose detection workflows into face detection, pre- processing, CNN inference, and decision services, orchestrated via container orchestration platforms like Kubernetes.[7] This approach facilitates autoscaling for live streaming demands and fault tolerance but introduces complexity and network latency considerations. Industry best-practice guides recom- mend layered pipelines combining artifact detection, biometric analysis, human-in-the-loop review, and continuous model updates to handle operational deployment challenges.[6]

## V. RESEARCH GAPS

Despite significant advances in deepfake detection, several critical gaps remain: Cross-Dataset Generalization: Most systems are evaluated on a single benchmark (e.g., FaceForensics++ , DFDC ), leading to overfitting and poor performance on unseen ma- nipulation methods.[10][11] There is a need for standardized cross-dataset protocols and domain adaptation techniques to ensure robustness across diverse generators and data distribu- tions.[1][7]

Temporal Context Limitations: While hybrid CNN–RNN and TCN models (e.g., EfficientNet + GRU , EfficientNet- B3 + TCN , ResNeXt + LSTM ) improve sequence modeling, they often focus on short temporal windows.[8] Long-range dependencies and multi-scale temporal dynamics remain un- derexplored, particularly for complex facial motions and subtle physiological cues.[9]

Real-Time and Resource-Constrained Deployment: On- device and streaming applications (Android app , sliding- window voting , microservices ) prioritize latency over accu- racy, resulting in reduced detection performance.[1][3] There is a gap in lightweight architectures that maintain high accu- racy under stringent latency and power constraints, especially on mobile and edge devices.

Active Forensics Applicability: Corneal reflection probing shows promise for live authentication but relies on controlled display conditions and high-quality camera input.[4] Research is needed to generalize active forensic methods to varied light- ing, camera angles, and privacy-preserving scenarios where user cooperation may be limited.[15]

Multimodal and Physiological Signals: Current approaches predominantly analyze visual artifacts. The integration of audio, lip-sync consistency, heartbeat-induced PPG signals, and other biometric modalities—highlighted but not fully im- plemented in industry guides —offers an underutilized avenue for robust detection, particularly against advanced GAN-based generators.[3]

Explainability and Human-in-the-Loop Workflows: Most models function as black boxes with limited interpretabil- ity. Industry best-practices recommend human escalation , yet systematic frameworks for explainable alerts, confidence calibration, and user-friendly interfaces are lacking. Research should address transparent reasoning and actionable insights for non-expert users.[3]

Standardized Metrics Beyond Accuracy: Nearly all studies report accuracy or AUC on balanced splits. Metrics reflect- ing false positive/negative trade-offs (EER, precision–recall curves), calibration, inference throughput, and robustness un- der compression or adversarial attacks are rarely standardized, hindering fair comparisons.

## VI. CONCLUSION

In summary, deepfake detection has rapidly evolved from spatial-only CNN classifiers to sophisticated hybrid architec- tures that integrate temporal modeling, active forensics, and scalable deployment pipelines; however, significant gaps re- main in cross-dataset generalization, long-range temporal rea- soning, low-latency mobile inference, robust active authentica- tion under diverse conditions, and multimodal fusion. Bridging these gaps will require novel domain adaptation strategies, lightweight yet accurate architectures, privacy-preserving chal- lenge–response protocols, integration of audio and physiolog- ical signals, and standardized evaluation metrics that reflect real-world constraints. Only through such interdisciplinary efforts can next-generation systems reliably safeguard digital media integrity in an increasingly adversarial landscape.

## REFERENCES

[1] R. Khaled, H. M. Moftah, F. K. Alsheref, A. S. Assiri, K. H. Rahouma, and M. Kayed, "Boosting Deepfake Detection Accuracy with Unsharp Masking and EfficientNet Models," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 16, no. 8, pp. 322- 333, 2025.
[2] S. Chaudhauri, V. R. Goudar, S. Hazam, and V. Ranjan, "AI-Powered Real-Time Deepfake Detection," International Journal of Research Pub- lication and Reviews, vol. 6, no. 4, pp. 6577-6581, April 2025

[3] S. Geetha, V. S. Murali, V. G. K, and P. T. S, ”Deepfake Detection using TCN and EfficientNet-B3,” International Journal on Science and Technology (IJSAT), vol. 16, no. 2, pp. 1-12, April-June 2025
[4] K. V. A. Reddy, Lochan S, Shrusthi, E. P. Goud, and M. Swapna, ”Development of AI/ML-Based Solution for Detection of Face-Swap Deep Fake Videos,” International Journal of Scientific Research and Engineering Trends, vol. 11, no. 2, pp. 2689-2696, Mar-Apr 2025.
[5] H. Guo, X. Wang, and S. Lyu, ”Detection of Real-Time Deepfakes in Video Conferencing with Active Probing and Corneal Reflection,” arXiv preprint arXiv:2210.14153v1, Oct. 2022
[6] Paravision Inc., ”A Practical Guide to Deepfake Detection: Definitions,
[7] Challenges, and Solutions,” Technical Report, October 2024
[8] M. Arkachari, U. Sadalage, O. Naik, P. Chandake, and P. Sonnad, ”The State of Live Deepfake Detection in Streaming Platforms,” International Journal of Advanced Research in Science, Communication and Tech- nology (IJARSCT), vol. 5, no. 5, pp. 498-503, June 2025.
[9] Abhinav, S. Singh, and A. Bharadwaj, ”Realtime Deep-fake Detection on Video Streams,” International Journal of Research Publication and Reviews, vol. 6, no. 4, pp. 7809-7811, April 2025.
[10] G. Sangar and V. Rajasekar, ”Optimized classification of potato leaf disease using EfficientNet-LITE and KE-SVM in diverse environments,” Frontiers in Plant Science, vol. 16, article 1499909, 2025. DOI: 10.3389/fpls.2025.1499909.

[1] G. Harshit, M. C. Reddy, and Y. V. S. Sarath, ”Deepfake Detection using Efficientnet-B0 and GRU,” International Journal of Engineering Research and Technology (IJERT), vol. 14, no. 4, pp. 1-8, April 2025.

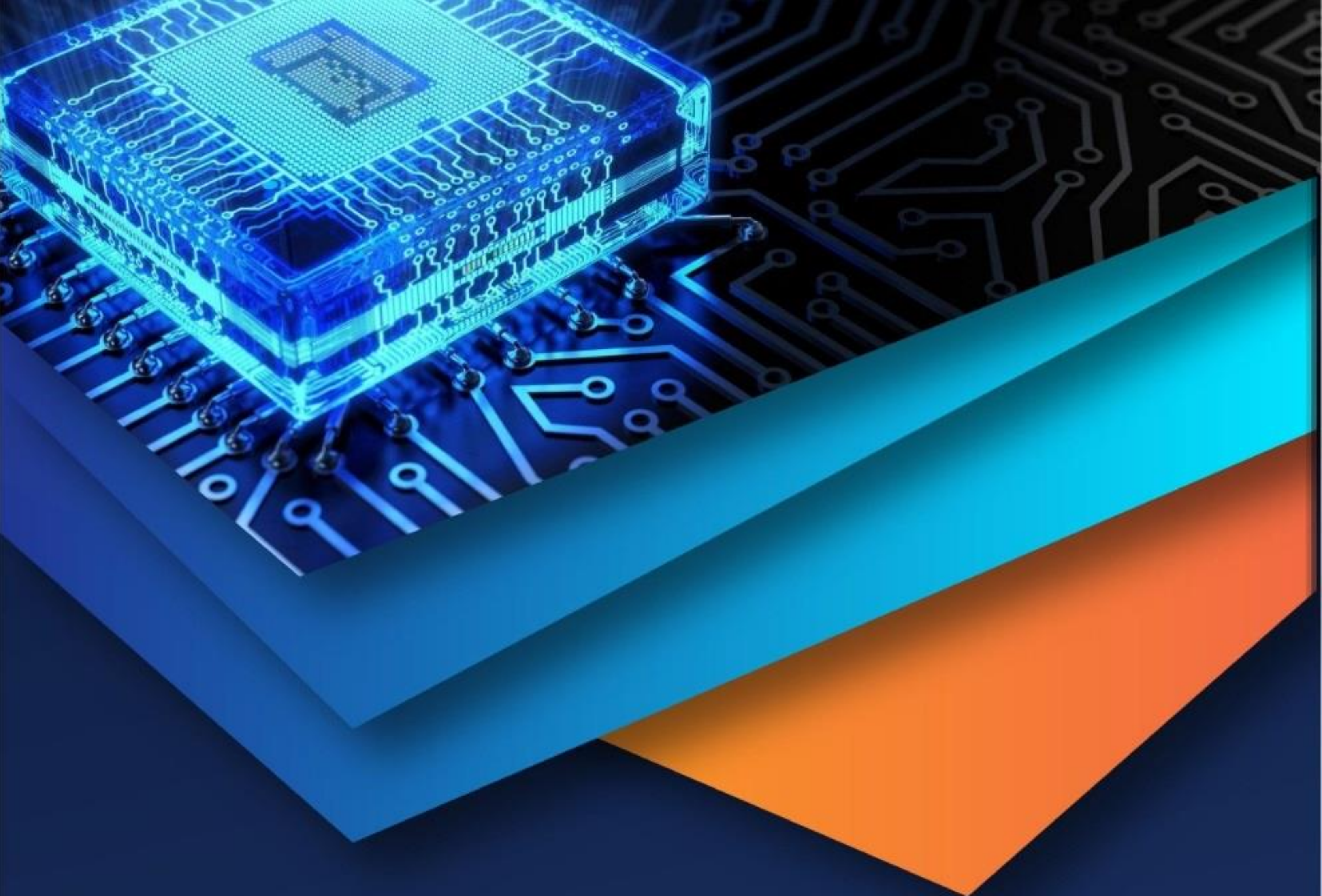
[2] C. Sahu, ”Deepfake Detection System,” International Journal of Engi- neering Research and Technology (IJERT), vol. 14, no. 5, pp. 1-12, May 2025.

[3] M. Alrashoud, ”Deepfake video detection methods, approaches, and challenges,” Journal of Computer and Information Sciences, vol. XX, no. XX, pp. XX-XX, 2025.

[4] R. Sunil, P. Mer, A. Diwan, R. Mahadeva, and A. Sharma, ”Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation,” International Journal of Computer Science and Engineering, vol. XX, no. XX, pp. XX-XX, 2025.

[5] G. N. Vivekananda, T. R. Mahesh, M. Gupta, A. Thakur, and A. Sayal, ”Refining digital security with EfficientNetV2-B2 deepfake detection techniques,” Journal of Cybersecurity and Digital Forensics, vol. XX, no. XX, pp. XX-XX, 2025.

[6] S. S. Chorage, A. Barabde, J. Balsaraf, and S. Batwal, ”Deepfake Detection Using Deep Learning,” International Research Journal on Advanced Engineering Hub, vol. 3, no. 8, pp. 3427-3431, Aug. 2025.




10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY