International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

Face V/S Fake: Real-Time Deep fake Detection Application for Android Devices

Hima G[1], Nandana B Nair[2], Rozana Beegum R[3], Sachin S Kumar[4], Sidhik A[5]

[2,3,4,5]UG - Computer Science Engineering, College of Engineering Perumon, Kollam, Kerala

[1]Assistant Professor - Computer Science Engineering, College of Engineering Perumon, Kollam, Kerala. ***Emails:****ghima0105@gmail.com[1],nandanab003@gmail.com[2],beegumrozana@gmail.com[3],sachinskumaratwork@gmail.com[4],sidhika@perumonec.ac.in[5].*

Abstract

Parkinson's disease (PD) is a progressive neuro- logical disorder that primarily impacts movement, resulting in symptoms such as tremors, stiffness,slowness and balance issues. Since early interventionmay still be effective in delaying disease progression and improving care and quality of life, timely diagnosis is crucial. Diagnosis Despite being aware of the early symptoms, routine diagnosis relies on neurologists' clinical examinations – something which can belengthy and subjective, also requiring patients to visit hospital multiple times. Strangely enough, slight changes in tone and tremor or raspy delivery that you'll often overlook when you're chatting are some of the earliest signs of Parkinson's disease.For the early diagnosis of Parkinson's disease, this work explores a voice-enabled machine learning system. The system identifies patterns indicative of Parkinson's disease through analysis of acoustic features in speech samples. Parameters such as pitch, jitter, shimmer, harmonics-to-noise ratio and other percepts that measure the lack of stability in a person's voice due to disease are key features. To tackle these inconsistencies, the system leverages a dataset of voice recordings taken both from people with Parkinson's and those without, in order to normalise values and prepare it for training strong machine learning models.These voice features are then used to train a number of machine learning algorithms, such as Random Forest, Support Vector Machines (SVM), Logistic Regression. The robustness of our models is guaranteed by extensive optimization via cross- validation and hyperparameter tuning. Among them, Random Forest stands out by its accuracy as well as interpretability as we are able to know which part of the voice contributes the most for predicting. The high recognition performance of the system to discriminate PD patients from healthy subjects is confirmed according to evaluation metrics such as accuracy, precision, recallF1-score and ROC-AUC.Streamlit is employed to integrate the trained model into an accessible web interface. The users have three alternative ways to engage with the system, by: manually entering features,uploading a pre-extracted feature CSV file, or directly recording their voice. Besides informing about the relevant voice features that influence the choice,such system also provides an immediate prediction as to whether or not some has Parkinson's. Doctors could have an additional, noninvasive tool to help them diagnose patients early in an easy and simple way through this method that allows screening and following the patient without having to makenumerous visits to the hospital. In conclusion, this project offers an approachable, compre- hensible, and efficient Parkinson's disease detection solution by fusing digital signal processing, artificial intelligence, and web technology.

Keywords: *Deepfake Detection; Android Application; Real Time Video Analysis; Efficient-Net; Tensor-Flow Lite; Media Projection API; Facial Analysis*

1. Introduction

Detecting deepfakes has important implications, as recent advances in AI have democratised the technology supporting their creation, enabling the production of synthetic media (i.e., altered images,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

videos and audio) that is increasingly difficult to detect. While they can be harnessed for legitimate uses in entertainment, digital media and education, they are also a global concern. Deepfakes have been used for a huge variety of online harms, such as impersonation, disinformation, and social engineering attacks, as well as harming trust in online content and social media [1]. Further, with the tools for generating deepfakes becoming more advanced and easier to access, it is increasingly difficult to spot them.Early detection methods relied on hand-engineered features, visual artefacts, and statistical irregularities in the manipulated video. However, customary detection methods are also not strong enough to some more advanced generation methods, such as video compression, addition of noise, and modification of resolution. More recent methods for deepfake detection are based on deep learning, and in particular convolutional neural networks (CNNs), which have been shown to detect deepfakes by identifying artefacts on a spatial and temporal level [1].Despite these advancements, most existing deepfake detection neural networks are computationally expensive, and utilise deep CNNs or transformer-based models with millions of parameters. Even though these models achieve a high accuracy on their respective datasets, they cannot be relied upon to be used for real-time detection or on edge devices such as mobile phones. This situation is problematic as social media and the web are the main distribution platforms of deepfakes and therefore fast and scalable detection methods are necessary. Recently, lighter, more efficient deep learning architectures have been proposed to achieve realtime performance for deepfake detection. Binary neural networks (BNNs) and models optimized for mobile devices have been shown to be effective at reducing computation while achieving similar accuracy [3], [4]. They result in faster inference and lower latency while reducing energy and memory footprint. This makes them ideal for resource-constrained devices such as smartphones and edge devices. However, the trade-off between efficiency and robustness, and their generalisation to unseen manipulation methods are open.Another limitation faced by many deepfake detection methods is that they are only based on a single modality (usually visual). As deepfakes contain multimodal artefacts (for example, audiovisual-image), thus, detection in a multimodal way is needed. Further iterations have sought to unify audio, visual and temporal cues in a holistic system design, improving robustness and accuracy for advanced deepfake detection [2]; however, such systems typically come with added complexity of architecture and design, making it difficult to scale for practical applications.This paper critically reviews common state-of-theart deepfake detection models based on deep learning methods and analyses them through the lenses of efficiency, robustness, and usability. Our objectives are to (i) understand the state-of-the-art in deepfake detection models using CNNs, transformers, and lightweight models, (ii) identify gaps in current research in real-time performance, computational efficiency, and generalizability to unseen datasets, and (iii) propose research directions to develop scalable deepfake detection models that can be deployed to combat synthetic media.

2. Related Work

Detection of deep fakes has become an important area of research due to the tremendous advancements in generative models that are capable of creating very realistic images and videos to be used for forging. Most of these approaches rely on deep learning-based computer vision methods for detection to find artifacts and inconsistencies introduced during the manipulation process. This section is meant to provide a thematic as well as critical analysis of recent deep fake detection methodologies, specifically architectural trends, detection capabilities, and practical deployment considerations. For structured comparison, the examined literature can be arranged into three major groups:

- CNN-based spatial detection methods,
- CNN-based spatial detection methods,
- Temporal and hybrid deep learning approaches, and Lightweight and Real-Time

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

Detection Frameworks

2.1.CNN- Based Spatial Detection Methods

CNN-based spatial detection methods represent the earliest and most extensively studied category of deepfake detection approaches. These methods operate by analyzing individual video frames or facial images to identify visual artifacts introduced during synthesis and manipulation. Common indicators include unnatural texture patterns, blending inconsistencies, colour distortions, and residual convolutional traces left by generative models. Since these approaches rely exclusively on static spatial features, they emphasize frame-level discrimination rather than temporal coherence. Several studies employ deep CNN architectures with transfer learning to learn discriminative facial features from manipulated content [1], [6], [7]. These models demonstrate strong detection capability on benchmark datasets, validating the effectiveness of spatial feature learning for identifying manipulation artifacts. Their architectural simplicity and compatibility with established image classification backbones make them attractive as baseline deepfake detectors.However, spatial-only detection approaches exhibit inherent limitations. By ignoring temporal relationships across frames, they struggle to identify high-quality deepfakes that maintain visual realism over time. Furthermore, crossdataset evaluations reveal performance degradation, indicating sensitivity to dataset bias and manipulation-specific artifacts [2], [6]. Consequently, while spatial CNN-based methods provide a strong foundation, their limited generalization capability restricts their effectiveness in real-world video-level detection scenarios.

Table 1 Comparison of CNN-Based Spatial Deep fake Detection Methods Data Collection and Voice Recording

Methodology	Input Type		Performance Metrics	Advantages	Limitations
CNN with transfer learning [1]	Video frames		High benchmark accuracy	Strong spatial feature learning	No temporal modeling
Convolutional trace analysis [6]	Images		Robust artifact detection	Forensic relevance	Dataset dependency
Visual artifact-based CNNs [7]	Images		Competitive detection rates	Simple Architecture	Limited generalization

2.2.Temporal and Hybrid Deep Learning Approaches

The temporal and hybrid methods of deep learning for spatial analyses also extend their ability to model relationships between frames and inconsistency in behaviour across manipulated video frames. The rationale for developing these new approaches is that, even though many individual frames have a very similar visual quality, there are significant differences in how deepfake videos appear over time with regards to facial motion, eye movement and transition of expressions. Temporal cues are used to try and capture information about manipulation in a video that cannot be captured through analysis of individual frames.Temporal-based detection methods exploit behavioural irregularities such as abnormal eye blinking and inconsistent motion patterns to enhance video-level detection reliability [5]. Hybrid architectures further integrate CNN-based spatial feature extractors with temporal learning mechanisms, enabling joint modelling of visual details and sequential dependencies [8]. In parallel, frequency-domain and texture-based representations combined with compact neural architectures demonstrate that temporal cues can be exploited efficiently without excessive computational overhead [3].Although temporal and hybrid approaches significantly improve robustness compared to spatial-only models, they introduce additional challenges.Increased architectural complexity leads to higher training cost and sensitivity to video quality variations, frame rate changes, and noise. Moreover,many such systems rely on curated

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

datasets and controlled evaluation settings, raising concerns regarding scalability and consistent performance in unconstrained real-world environments [3], [8].

Table 2 Comparison of Temporal and Hybrid Deep fake Detection Methods

Methodology	Model Architecture	Key Features	Advantages	Limitations
CNN-based temporal analysis [5]	CNN + motion cues	Behavioral inconsistencies	Improved video detection	Sensitive to video quality
Binary neural network approach [3]	BNN + FFT + LBP	Frequency & texture cues	High efficiency	Accuracy trade-off
CNN–LSTM hybrid [8]	Mobile Ne t + LSTM	Spatial–temporal fusion	Robust detection	Higher complexity

2.3.Lightweight and Real-Time Deep fake Detection Frameworks

Deep fake detection frameworks that are lightweight and work in time are focused on the things that matter when we use them. These deepfake detection frameworks have to deal with the limits of what we can do when we deploy them. These frameworks are made to be used in the real world. So, they have to be able to work with the constraints that come with that. These methods are trying to reduce the amount of work that computers have to do with detection complexity, inference latency, and energy consumption to enable largescale monitoring and edge device deployment. Such considerations are particularly important for online platforms and real-time content moderation systems where low-latency decision-making is essential.Compact CNN architectures, mobile-optimized models, and binary neural networks have been explored to achieve a balance between detection performance and efficiency [3], [8]. These systems significantly reduce resource requirements while maintaining competitive detection capability,making them suitable for real-time and embedded applications. Survey-based analyses further highlight the necessity of lightweight detection mechanisms for practical adoption beyond laboratory environments [2],[4].Despite their deployment advantages, lightweight detection frameworks often sacrifice robustness and adaptability. Many such systems exhibit reduced resilience to advanced manipulation techniques and lack mechanisms for continual learning or adaptation to evolving deepfake generation strategies[4], [7]. This trade-off between efficiency and long-term reliability is a major challenge, mainly in dynamic real-world scenarios where manipulation methods evolve rapidly.

Table 3 Comparison of Lightweight and Real-Time Deep fake Detection Systems

Methodology	Deployment Target	Performance Metrics	Advantages	Limitations
Mobileoptimized CNNs [8]	Edge / real time	Low latency, moderate accuracy	Resource efficient	Limited adaptability
Binary neural networks [3]	Real-time systems	Reduced FLOPs	Fast inference	Lower accuracy ceiling
Lightweight CNN frameworks [7]	Online platforms	Moderate detection rates	Easy deployment	Vulnerable to advanced fakes

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

3. Comparative Analysis And Discussion

In this section, the deepfake detection approaches discussed in the related works are comparatively analyzed. The discussion does not evaluate the studies individually; instead, it synthesizes key trends, architectural trade-offs, and system-level limitations across CNN-based models, transformer frameworks, lightweight mobile systems, and hybrid spatial–temporal architectures. The objective is to assess how effectively current detection strategies address real-world deepfake challenges and to identify persistent gaps that may affect scalability and long-term robustness.

3.1.Performance and Trade-offs of Detection

The reviewed studies indicate that convolutional neural network (CNN)-based systems form the foundational approach in deepfake detection. These models effectively extract spatial features and identify manipulation artifacts such as texture inconsistencies and blending errors. Their performance in benchmark datasets demonstrates strong classification capability and reliable detection accuracy in controlled environments [1]. However, CNN models primarily focus on local spatial information and may not fully capture contextual or sequential dependencies in complex video manipulations. Transformer-based and attention-enhanced architectures improve global feature modeling and contextual reasoning, thereby enhancing robustness against sophisticated deepfake techniques. While these advanced models achieve improved detection performance, they introduce increased computational complexity and training demands, highlighting a trade-off between accuracy and efficiency [5], [9].

3.2.Deployment Feasibility and Systems Integration

Deployment feasibility remains a critical concern in deepfake detection systems, particularly for real time and mobile applications. High-capacity architectures often require substantial computational resources, which may limit their scalability in practical environments. In contrast, lightweight and mobile-oriented frameworks aim to reduce model size and optimize inference speed while maintaining acceptable detection performance [4].Mobile-focused approaches further enhance deployment practicality by integrating efficient feature extraction and domain-specific cues suitable for edge devices. Although these systems improve accessibility and real-time usability, their simplified structures may reduce robustness when exposed to highly sophisticated or previously unseen manipulations. Unified and multimodal frameworks expand detection versatility but also introduce architectural complexity that may challenge efficient system integration [2], [6].

3.3.Robustness and Architectural Evolution

An observable architectural evolution is clear in the reviewed works, moving from traditional convolutional models to hybrid and transformerbased systems. Hybrid spatial-temporal frameworks improve detection reliability by combining frame-level feature extraction with sequential learning, which boosts performance in video-based deepfake scenarios [8].Additionally, optimization-driven strategies like binary neural networks aim to improve inference speed while trying to maintain detection accuracy. Despite these improvements, cross-dataset generalization and resilience against new manipulation techniques continue to be persistent challenges. Many models show high performance under certain evaluation conditions but may face drops in performance in varied real-world environments [3].

3.4.Adaptability and Long-term Effectiveness

A significant limitation observed across current detection systems is their dependence on supervised training using predefined datasets. While this approach yields strong experimental accuracy, it restricts adaptability to rapidly evolving deepfake generation methods. As generative models continue to improve in realism, detection systems must evolve to maintain effectiveness[7].Furthermore, efficiency-oriented models emphasize computational optimization but often lack adaptive or continuously updating mechanisms. Without dynamic learning strategies and cross domain robustness, long-term

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

reliability may be compromised. Therefore, future deepfake detection frameworks must integrate adaptability, scalability, and resilience to sustain effectiveness in rapidly changing digital ecosystems [4].

3.5.Summary of Key Observations

Overall, the comparative evaluation reveals that deepfake detection research has advanced significantly through improvements in spatial feature extraction, contextual modeling, temporal learning, and computational optimization. CNN based systems provide reliable baseline detection, transformer-based models enhance contextual reasoning, and hybrid frameworks strengthen video level consistency analysis [1], [5], [8]. However, trade-offs between detection accuracy, computational efficiency, and generalization remain evident across all categories. Lightweight deployment-oriented models improve scalability but may compromise robustness, whereas advanced architectures require higher resource investment. Limited adaptability to evolving manipulation techniques continues to present a major challenge, emphasizing the need for integrated, efficient, and adaptive detection frameworks capable of sustaining reliable performance in real-world applications.

4. Research Gaps

Despite camera-based deepfake detection using CNNs, transformers, and hybrid architectures making large steps forward, existing deepfake detection systems are typically designed and developed in high-performance computer environments rather than mobile or real-world environments. While many studies highlight advances in accuracy by utilizing computationally heavy models (e.g., LSTMs, vision transformers), the challenges of deploying these architectures tolow resource, on-device environments have received limited attention. Additionally, real-time smartphone-based detection has yet to be adequately examined, particularly with scenarios requiring continuous screen monitoring, low latency inference, and minimal battery consumption from the device. This represents a clear disconnect between theoretical or researchbased detection capability and standard or practical user level performance on devices.A significant area of concern is in the end-to-end system integration for real-world mobile environments. The majority of the current literature relates to either model development or dataset evaluation yet issues such as managing the screen capture process of real-time video, buffering frames, matching preprocessing (training and deployment), optimizing TF-Lite, and user notifications through overlay mechanisms are largely unaddressed in a cohesive manner.Furthermore, fully on-device processing avoiding dependency on the cloud to protect user privacy adds additional design constraints that are not well investigated in current literature. Therefore, lightweight, scalable and privacy-preserving detection of deepfakes will be greatly needed.

5. Proposed Ai-Based Real-Time Deepfake Detection System

As a result of the research gaps identified, the proposed research seeks to develop a deepfake detection framework for Android smartphones. Unlike the conventional methods, the proposed method focuses on the applicability of the developed model in terms of efficiency and the preservation of privacy. The proposed method focuses on the development of a screen-based acquisition, intelligent inference, decision-making, and alerting system for Android smartphones. Unlike the conventional methods, the proposed method focuses on the development of a structured end-to-end pipeline for Android smartphones, as depicted in Fig. 1.

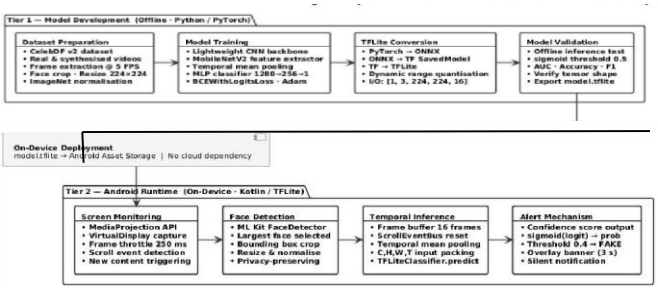

Figure1. Proposed Framework for AI-Based Real-Time Deepfake Detection System

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

At the system level, the framework incorporates the direct acquisition of visual content from the smartphone display at predefined time intervals and processes it using a lightweight convolution neural network optimized for mobile platforms. Subsequently, the acquired video frames are passed through the detection pipeline for further processing. This allows the system to track the visual content on the smartphone display, ensuring efficient utilization of resources for mobile platforms. Face detection is initially employed to identify relevant facial regions, which enhances the precision of the classification and eliminates redundant computation. The trained deep learning model is converted to Tensor Flow Lite to guarantee efficient execution with limited memory, CPU, and battery constraints. Since all the computation is executed locally, the system offers the lowest latency while maintaining the privacy of the users without relying on the cloud. In order to improve the robustness of the framework in video-based situations, a computationally efficient temporal pooling approach is included. Scroll detection is incorporated to detect changes in on-screen content and prevent conflicts between frames from previously processed video and newly displayed video. When a scrolling event occurs, the system captures the updated frame and refreshes the Tensor Flow Lite input buffer, ensuring that only the current visual content is analyzed. Instead of using complex recurrent models, a buffer-based approach is used to aggregate consecutive facial frames using a computationally efficient temporal pooling approach. This improves the consistency of the prediction results. The inference engine produces a confidence score indicating the probability of manipulated content. If the score exceeds a threshold, an alert system using an overlay-based approach is triggered, instantly alerting the user while the suspicious content is being viewed. In addition, modularity and scalability give way to being able to add improved lightweight models and optimization techniques in the future without requiring the entire system to be rebuilt. Additionally, there are stringent controls over background services so they do not compete for resources or deplete battery life, and there is complete transparency on whether apps have permission to use resources. The framework in Figure 1 demonstrates that incorporating edge intelligence results in real-time response capabilities, efficient mobile deployment, and safe but private computing, enabling smartphone-based deep fake detection via a reliable and scalable implementation.

Conclusion

This paper reviewed existing deepfake detection techniques, their limitations, and identified critical gaps in their ability to support real-time detection, integration with mobile devices, and preservation of user privacy. Although remarkable milestones have been achieved in enhancing deepfake detection accuracy through advanced deep learning techniques, their integration with resourceconstrained mobile devices for real-time detection of deepfake content still remains an understudied area. Most of the existing techniques focus more on model performance in controlled environments as opposed to integration with mobile devices that can support their practical application. As part of addressing these limitations, a conceptual framework was developed to support real-time deepfake detection on mobile devices, with particular emphasis placed on integration, temporal robustness, and user-centric alerting. By leveraging the power of edge computing with scalable and privacy-preserved design principles, this paper provides a structured approach to support the practical development of deepfake detection solutions. This paper, therefore, underscores the need to move from accuracy-centric deepfake detection techniques to more practical, real-world solutions that can support deepfake detection and mitigation, thus reducing the menace of deepfake content.

Acknowledgements

The completion of this project, “Face V/S Fake: Real-Time Deepfake Detection Application for Android,” is due to the guidance and support provided by many individuals. We would like to express our sincere appreciation to our guide Prof. Sidhik A, Department of Computer Science and Engineering, College of Engineering Perumon, for his valuable guidance,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

constructive suggestions, and encouragement in the development of this work.We would also like to express our sincere appreciation to the Department of Computer Science and Engineering, College of Engineering Perumon, for providing the required infrastructure and academic environment that helped us successfully complete this research work.We would also like to express our sincere appreciation to our friends for their valuable discussions and cooperation during the development of this work.

References

[1]. [1]A Karandikar, V. Deshpande, S. Singh, S.Nagbhidkar, and S.Agrawal,"Deepfake Video Detection Using Convolutional Neural Network", International Journal of Advanced Trends in Computer Science and Engineering, 9(2), pp. 1311–1315, 2020. https://doi.org/10.30534/ijatcse/2020/629220 20

[2]. Sar, S. Sati, T. Choudhury, P. Joshi, R. Sille, K. Srihari, and K. Bansal, "A Unified Neural Framework for Real-Time Deepfake Detection Across Multimedia Modalities to Combat MisleadingContent," IEEE Access, vol. 13, pp. 48683–48699, 2025. https://doi.org/10.1109/ACCESS.2025.3550770

[3]. R. Lanzino, F. Fontana, A. Diko, M. R. Marini, and L. Cinque, "Faster Than Lies: Real-Time Deepfake Detection Using Binary Neural Networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.

[4]. "Lightweight Deepfake Detection on Mobile Devices Using Attention-Enhanced MobileNet and Frequency Domain Analysis," Journal of Technology Informatics and Engineering (JTIE), vol. 4,no. 1, pp. 95–114, Apr. 2025

[5]. V. L. L. Thing, "Deepfake Detection with Deep Learning: Convolutional Neural Networks versus Transformers," pp. 1–8.

[6]. M. Mohzary, K. J. Almalki, B. Y. Choi, and S. Song, "MobiDeep: Mobile DeepFake Detection through Machine Learningbased Corneal-Specular Backscattering," IEEE Conference Paper, 2023.

[7]. Thakur, Sharma, and Barthwal, "AI Against AI: Deep Learning for Deepfake Detection," in Proceedings of the International Conference on Advances in Computing, Communication and Materials (ICACCM), 2024.

[8]. S. Maheswari, V. Dhilip Kumar, D. Ajith Kumar, R. Jahnavi, and C. Laharee, "RealTime Deepfake Detection Using a Hybrid MobileNet-LSTM Model for Enhanced Media Integrity," in Proceedings of the International Conference on Intelligent Systems and Digital Transformation (ICISD), 2025

[9]. H. Soudy, O. Sayed, H. Tag-Elser, R. Ragab, S. Mohsen, T. Mostafa, A. A. Abohany, and S. O. Slim, "Deepfake Detection Using Convolutional Vision Transformers and Convolutional Neural Networks," Neural Computing and Applications, 2024.

[10]. S. B. N. R. V. V. N. S. Vamsi, K. R. Kiran, and K. V. D. Kiran, "Deepfake Detection through Deep Learning," in Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp.1118–1122, doi: 10.1109/ICICCS48265.2020.9121085.

[11]. O. A. H. H. Al-Dulaimi and S. Kurnaz, "Deep Fake Image Detection Based on Deep Learning Using a Hybrid CNN-LSTM with Machine Learning Architectures as Classifier," in Proceedings of the 2024 6th International Congress on Human-Computer Interaction, Optimization and RoboticApplications(HORA),2024,pp.16,doi :10.1109/HORA61326.2024.10550728.

[12]. M. S. Raj, S. Suryaraman, S. Saravanan, and M. Muthulakshmi, "DeepFake Detection in Real-Time: A Hybrid LSTM-CNN

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2454-2462
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0328

Approach," Int. J. Res. Trends Innov.,vol.10,no.4,pp.538543,Apr.2025.[Online].Available:http://www.ijrti.org

[13]. R. S. Khudeyer and N.M.Almoosawi, "Fake Image Detection Using Deep Learning," Informatica, vol.47,no.7,Aug.2023.DOI:https://doi.org/10.31449/inf.v47i7.4741

[14]. Awotunde, J. B., Jimoh, R. G., Imoize, A. L., Abdulrazaq, A. T., Li, C. T., & Lee, C. C. (2023).An Enhanced Deep Learning-Based DeepFake Video Detection and Classification System.Electronics,12(1),87.https://doi.org/10.3390/electronics12010087

[15]. J. Singh, M. Lal, and K. P. S. Attwal, "A hybrid deep learning approach for deepfake detection using spatial and temporal features with attention mechanisms," Int. J. Eng. Comput. Sci., vol. 7, no. 2 Part A, pp. 30–37, 2025.Available:https://www.computerscienc ejournals.com/ijecs/archives/2025/vol7issue 2/PartA/7-2-7-940.pdf