Impact Factor-4.013 *e-ISSN: 2581-6667*
International Journal of Engineering and Creative Science, Vol. 8, No. 1, 2025
www.ijecs.net

An Empirical Review of Deepfake Detection Techniques Using Machine Learning Techniques and Performance Metrics

Ms. Pritee H Raut[1], Ms. Hemlata Dakhore[2]

[1] M. Tech CSE scholar, [2] Assistant Professor
Department of Computer Science and Engineering,
G H Raisoni College of Engineering and Management, Nagpur (M.S.), India

Abstract – *In view of the rapid proliferation which this deepfake technology has reached in creating highly realistic manipulated media, issues have been raised across major sectors: media integrity, security, and public trust. Although many reviews exist around deepfake detection, most of them limit their scope to either specific techniques or datasets, offering little or no comprehensive comparison of the many different methods now available for this process. Most of the reviews also lack discussions on computational efficiency or adaptability concerning real-world detection systems; hence, theoretical and practical understanding of the technologies could be potentially at a deficit. This work considers a critical review of state-of-the-art deepfake detection methodologies that offer comprehensive performance comparison based on efficiency and the scope of application. Advanced machine learning techniques reviewed in this work involve ensemble deep learning combined with optical flow, blockchain-based federated learning combined with CNN and SegCaps, and new real-time solutions such as plasmonic resonanceenhanced biosensors. Additionally, approaches that will merge human cognitive skills with machine learning and neurocognitive testing for the investigation of deep fake audio are examined in order to demonstrate the impact of human factors on detection performance. Methods with light-weighted architectures, such as shallow vision transformers, and schemes of multimodal detection, such as the cross-modal attention network, were also discussed for their efficiency in resource-constrained environments and detection along several dimensions of deepfakes. In fact, the proposed review would compare the accuracy, efficiency, and scalability of these methods; it would also critically evaluate their adaptability against evolving deepfake techniques and adversarial conditions. This work has great implications; it puts into perspective a unified framework for understanding the current capabilities and limitations of deepfake detection systems. It holistically presents the emerging trends to researchers and practitioners for making better decisions in developing robust, scalable, and efficient solutions toward securing the authenticity of digital media sets.*

Keywords: ***Deepfake Detection, Machine Learning, Federated Learning, Optical Flow, Real-Time Detection, Process***

I- INTRODUCTION

Within the last decade, advances in deep learning and artificial intelligence have gone a long way toward improving media creation, one very strong development being that of the so-called "deepfakes." Deepfakes are media-that is, images or videos-created with advanced deep learning algorithms which superimpose faces, morph audio, or generate completely fabricated content. While deepfake technology has opened up a whole new frontier in entertainment, art, and content creation, this technology has also been misused in a very serious way about privacy, security, misinformation, and digital content integrity. Deepfakes can also be maliciously used to deceive individuals, manipulate public opinion, or even threaten national security through the creation of political speeches or news events. While deepfake technology becomes increasingly accessible-through the

rise of user-friendly AI tools and platforms that require little to no prior technical knowledge-detection of these forgeries has turned into a field of crucial research. Classic detection methods using forensic image analysis or manual verification, for example, are far insufficient as modern deepfakes may become indistinguishable from genuine content by the human eye. The high fidelity of these synthetic media has rendered conventional detection methods obsolete; hence, the need for the development of advanced automated systems that can find the deepfakes at scale.

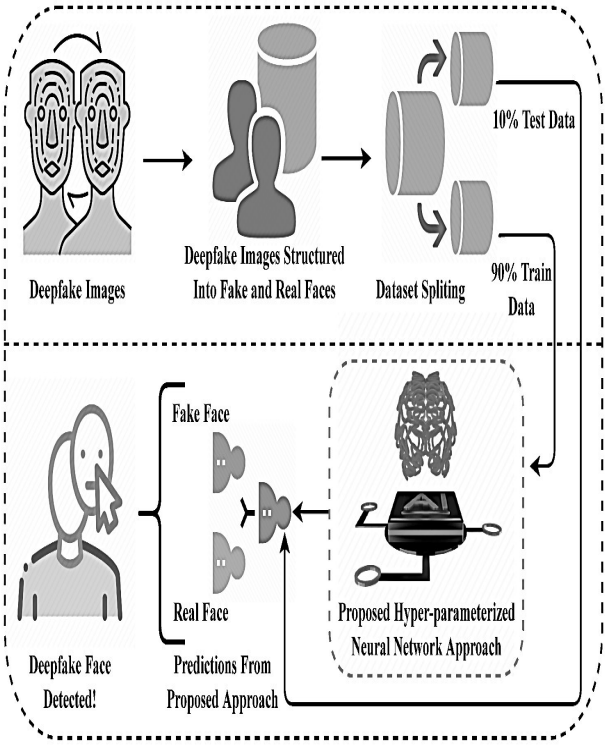

Fig. 1. Deepfake Detection Model Process

A number of machine learning-based methods that span from CNN models designed to operate at a frame level have so far been adopted for the addressed problem of deepfake detection. There are more sophisticated approaches that take into consideration temporal inconsistencies in the video sequences. Indeed, the availability of benchmark datasets like FaceForensics++ and that from the Deepfake Detection Challenge, DFDC, has fostered comparative analysis among them. Nevertheless, all these efforts have not yet resulted in any single technique emerging as a panacea for detecting these deepfakes of all kinds. Given the variability in methods for generating deepfakes, from image manipulation to audio alteration and video splicing, multi-modal approaches for detection from various forms of synthetic content have become quite critical. Therefore, most recent studies have considered innovative methodologies in the improvement of performance detection. This fact is demonstrated, for instance, by the performance of optical flow methods, which capture the motion of pixels between frames to reveal small inconsistencies in deepfake videos that are difficult to catch in frame-level analyses. FL based on blockchain has also been presented as a method for training deepfake detection models in a manner that ensures the privacy of decentralized datasets while still enhancing model robustness. On the other hand, plasmonic resonance-based biosensor technologies have been in a position to detect deepfakes in real time by detecting anomalies in optical signatures. Notwithstanding these advances, the area remains so dynamic, with day-in-and-day-out researchers coming up with ways of trying to tame the ever-increasing sophistication of algorithms used in the generation of deepfakes.

One of the major challenges with current deepfake detection is scalability. Many state-of-the-art deep models require heavy computational resources and are not suitable for real-time tasks or simple deployment on resource-constrained devices. Hence, lightweight architectures that provide a good balance between detection performance and low computational overhead shall be considered, including shallow vision transformers. These latter models reduce the number of parameters and FLOPS applied, finding their perfect fit for real-world applications without significant loss of accuracy. Another very active research area involves integrating human cognitive capabilities into deepfake detection systems. Although machines are able to process voluminous data and find patterns that are invisible to the human naked eye, the combination of algorithmic detection with human intuition has proved effective in certain cases. Indeed, it has been shown that when given some form of statistical metadata or other contextual clues, the rate of correct identification of deepfakes by humans increases significantly. This hybrid approach to deepfake detection-a fusion of human and machine intelligence-consigns a new frontier in the fight against deepfakes, especially in high-stakes scenarios such as elections or judicial processes.

The implications of detecting deepfakes span much more than academia. Deepfakes have critical cybersecurity consequences: from a direct hit on the credibility of digital content to journalism and financial markets, where entire industries can be disrupted. Fake news generation, impersonation of a person, and simulation of events stoke moral questions about the menace of

deepfake technology misapplication. Also, the jurisprudence of deepfakes has just started to take shape, as many countries have to grapple with the task of making laws and policies regarding the production and dissemination of this form of media. Urgency in preventing the spread of harmful content underlines the need for powerful, scalable solutions for detecting deepfakes at the platform level. While this comes with the fact that deepfake types range from visual to auditory forgeries, methods for the detection should also be multifaceted. Though most of the models for detection studies focus on manipulation of the face, recent trends have shown the growing prevalence of audio deepfakes where AI-generated voices impersonate real people. This adds another layer of complexity because detecting an audio deepfake requires completely different sets of features and techniques than image-based models. Spectral analysis, MFCC, and RNN have become key in recent times for analyzing deepfake audio content. The domain of deepfake detection has turned multivariate with the integration of expertise in computer vision, audio processing, and natural language understanding. Therefore, this review tries to present a comprehensive comparison of different detection methods corresponding to state-of-the-art due to the imminent threat of deepfakes. This paper aims to establish which techniques are more promising for further research and practical implementation by analyzing their performance on various datasets considering criteria such as accuracy, computational efficiency, scalability, and adaptability. Additionally, this review underlines some lacunas in the current literature and provides insight into unresolved challenges related to deepfake detection. The ultimate goal of this work is to ultimately give a fine understanding to the researchers and practitioners about the deepfake detection landscapes. This review contributes to the ongoing effort of securing the integrity of digital media and preventing malicious use of AI-generated content by showing both the strengths and limitations of the existing methods. While deepfake technology is only bound to get better, it is only by effective detection systems that will be increasingly necessary, hence going hand in glove with the cutting-edge research by way of interdisciplinary work, that this may be considered valid for the process. The paper is organized as follows: section II presents Review of Existing Models used for Parkinson Analysis Section III outlines the state of the art on Transmitarray covering in particular transmitarray for antenna beamsteering, polarization control and hybrid transmitarray, that enable both features simultaneously. Finally, the main conclusions are drawn in section IV.

II - MOTIVATION AND CONTRIBUTION

The malicious use of deepfakes within domains of politics, media, and personal privacy is an ever-growing concern, serving as the main motivation for this work. Deepfakes are not simply interesting technical minutiae; their impact can be quite serious when related to the erosion of public trust in the integrity of digital information. In turn, the more this technological advancement drives the generation of very realistic synthesized content, the more challenging it becomes to detect these deepfakes. Most of the existing detection systems are tailored to specific types of deepfakes, including the manipulation of faces and splicing of videos, hence limiting their adaptability to other forms which deepfake forgery may take, especially in the domains of audio, voice, and multimodality. Besides that, the computational cost of running many of such detection systems is hardly applicable in real time, especially in resource-constrained environments.

The paper reviews and makes a comparison of different deepfake detection methodologies in detail. Whereas other reviews focus on narrow areas, like a technique or dataset, this work presents a holistic view of the state-of-the-art by covering a wide range of approaches, starting from machine learning-based using CNNs and transformers to more innovative ones involving blockchain-based federated learning and integrated biosensors for real-time analysis. The review identifies the limitation of each method, emphasizing major trade-offs among some involving accuracy, computational efficiency, and scalability. It assists in the identification of lacunars in the literature and hence opens frontiers for development of more adaptive and robust detection systems.

The contribution of this paper goes beyond a simple comparative analysis; it frames both the technical and practical limitations of the current detection frameworks, hence offering significant insights for both researchers and practitioners. Novel methods included herein, such as the ensemble learning based on optical flow and the cross-modal attention networks, provide new light regarding how multi-modal deepfakes should be approached. This review also covers the real-world deployments of deepfake detection systems at scale with respect to concerns regarding privacy, heterogeneity in data, and their legal landscape. In general, the contribution provides a broad roadmap of future research in this area and hence makes useful suggestions for solving practically the challenges imposed by the rapid development of deepfake technology.

. III- REVIEW OF EXISTING MODELS USED FOR PARKINSON ANALYSIS

The rapid development of deepfake technologies has brought huge challenges across many industries; the main impact on digital security, misinformation, and personal privacy is very large. The literature review covers major developments in advances and detection techniques developed to fight deepfakes that use artificial intelligence for highly real synthetic media.

A. Emergence of Deepfakes and its Implication

The term "deepfake" symbolizes hyper-realistic media; this generally can take the form of images or videos created through deep learning algorithms in such a manner that these can easily masquerade as real people. Work in [1] identified the broad diffusion of such content across different media channels, creating disruptions in society with respect to identity theft, social engineering, and political disinformation. The increased proliferation of deepfakes challenges human cognition in that most of this manipulation of media information is not that easily detectable by the naked eye. This goes to show the urgent need for more sophisticated methods of detection that can establish whether some media information is real or fabricated.

B.Deepfake Detection Techniques

Many of them include detection methods, from deep learning-based to new hardware implementations. One of the major works in the deepfake detection domain is the ensemble deep learning-based system proposed in [1], using optical flow techniques for differentiating between real and fake images, which reported an accuracy of 86.02% on the DeepFake subset of the FaceForensics++ dataset. That brings out innovation in optical flow in apparent motion extraction, a breakthrough in bringing up the accuracy of detection. However, this study has pointed out the deficiency of such techniques for large-scale, practical detection of deepfakes. Another promising approach is deepfake analysis based on blockchain and federated learning. According to [2], one can combine FL with CNN and capsule networks for training global models much more robustly with preservation of anonymity of data sources. With the integration of blockchain, this model is more capable of handling heterogeneity and preserving data confidentiality, with 6.6% improvement in accuracy compared to six benchmark models. Indeed, this is a federated approach which mitigates the challenge of data privacy, very often a limitation in many centralized deepfake detection systems. Another frontier in the detection technology is the integration of biosensors, shown in [3]. They proposed a biosensor enhanced by plasmonic resonance, together with machine learning algorithms that capture subtle anomalies in digital content. This biosensor provides real-time detection at an accuracy of 98.7% and is one of the most effective tools within the deepfake detection landscape of today, with adaptive learning capabilities against evolving deepfake generation techniques. The authors, in this work, have investigated the interplay of human cognitive skills with machine learning for the detection of deepfakes. According to their experiments, human recognition of deepfakes, when combined with machine learning models, could achieve impressive detection accuracy up to 98.3%. It is indeed a hybrid approach: intuition by humans and precision by algorithms. This complements a purely machine-based detection systems.

C.Deep Learning Methods and Model Optimization

Most of the recent works emphasize deep learning-based approaches for deepfake detection and focus on CNNs and attention-based networks. The work in [5] looks deep into neurocognitive responses to audio deepfakes; it shows that the cortico-striatal network of the brain plays a critical role in decoding deepfake-level audio manipulations. This insight opens avenues toward neurocognitive-based detection frameworks, which may themselves be used to enhance further human resistance against deepfake identity spoofing. In [6], it is proposed that a CNN-based deepfake detection framework performs on frame-level analysis with the utilization of vision transformers to draw out features. They have proposed a model that performs detection accuracy at 97% and, thus established the efficiency of CNN-based architectures to handle big data sets such as FaceForensics++ and DFDC process. Simultaneously, attention mechanisms, cross-modal strategies are turning popular. In that, a novel architecture of cross-modal attention was proposed, which effectively captured the fake content of both audio and video modalities. The results showed that the Bidirectional Recurrent Convolutional Network used in the current study remarkably enhanced the detection performance, especially for multimodal deepfakes, which normally would be hard to detect because they need more sources of manipulated data samples. In resource-constrained environments, shallow learning models have also emerged as alternatives. Indeed, work in proposed a shallow vision transformer for deepfake detection, which leverages multi-head attention for the localization of manipulated sections within an image. Although it is a shallow model, it can achieve accuracy rates of 92.15% and 88.52% on Real Fake Face and RFFD datasets,

respectively. Thus, efficient lightweight models can execute comparably well with deep learning systems that are computationally intensive in certain scenarios.

Table 1. Empirical Review of Existing Methods

Reference	Method Used	Findings (in context of Deepfake Analysis)	S
[1]	Ensemble deep learning-based system using optical flow techniques.	Achieved accuracy of 86.02% on the DeepFake subset of FaceForensics++ dataset by detecting apparent motion in pixels, enhancing real vs. fake image differentiation.	N fl le a m e d
[2]	Blockchain-based federated learning combined with SegCaps and CNN, utilizing transfer learning and data normalization.	Enhanced global model training for deepfake detection with an accuracy increase of 6.6% over six benchmark models and a 5.1% AUC improvement in process.	P s a w a h p g c
[3]	Plasmonic resonance-enhanced biosensor integrated with CNN for real-time detection.	Achieved 98.7% detection accuracy with real-time analysis and low false positive/negative rates.	H s ra ti a e d te
[4]	Machine learning combined with human cognitive abilities to classify videos as fake or real.	Achieved 98.3% accuracy in predicting human evaluations of video authenticity, blending cognitive analysis with algorithmic detection	F t h c a d d
[5]	Neurocognitive sensitivity testing through audio deepfake identity spoofing.	Identified brain regions involved in deepfake voice detection, offering insights into human vulnerability and resilience	Expands understan of h deception deepfakes the neurocog level.
[6]	Deep learning-based detection using CNN and vision transformers on FaceForensics++ and DFDC datasets.	Achieved 97% accuracy in identifying fake images and videos using a multi-component CNN and vision transformer approach	Strong performan with m metrics su precision F1-score.
[7]	Evaluation of prior information's effect on deepfake speech recognition and deepfake audio quality metric.	Found that prior information and audio quality significantly influence human ability to detect deepfake speech.	Novel approach simulatin world sce of unpredict deepfake exposure.
[8]	Classification algorithms and ensemble model applied in the frequency domain to detect low-resolution deepfakes.	Achieved up to 99.97% detection accuracy for high-resolution deepfakes using random forest and other classification methods.	Extremely detection especially low-reso deepfakes robust validation
[9]	Cross-modal attention architecture with a bi-directional recurrent convolutional network.	Demonstrated promising performance in recognizing multi-modal deepfakes by capturing spatial-temporal deformations in audio and video.	Effective multi-m deepfake detection, addressin multiple content modalitie
[10]	Hybrid ResNext 50 + LSTM architecture using	Outperformed other deepfake detection models	Effective handling datasets

	CNN and Big Data techniques.	like 'FakeCatcher' and 'Face X-ray' with better resource-efficiency and detection accuracy.	te ir o r aj
[11]	Shallow vision transformer model for constrained resource environments.	Achieved accuracy of 92.15% on Real Fake Face dataset with reduced model parameters and FLOPS.	E u c e d re re
[12]	Ensemble-based D-Fence framework using uni-modal and cross-modal classifiers.	Achieved 92% detection accuracy across facial and vocal manipulations, with resilience to novel adversarial attacks.	R m d m a a c
[13]	Remote Photoplethysmography (rPPG) and DeepPhys model for biological signal detection in videos.	Identified biological signals in deepfake videos, demonstrating higher availability than traditional rPPG methods.	In p d d e d m b a
[14]	Unified framework using Mel-Frequency Cepstral Coefficients (MFCC) and spectral contrast features for audio deepfakes.	Achieved 98.52% detection accuracy using MFCC DeltaDelta features with spectral contrast in audio recordings.	H ir au d e a
[15]	DefakeHop++ model using layered detection approach for platform responsibility and government intervention process.	Proposed multi-layered platform policies and AI algorithms for protecting democratic processes against deepfakes.	E g p re ir te detection.
[16]	Fusion of hand-crafted and deep-learned features for exposing deepfakes in video datasets.	Demonstrated effective detection performance across Celeb-DF, DFDC, and FaceForensics++ datasets using feature fusion.	Combines strengths traditiona deep-lea methods improved accuracy.

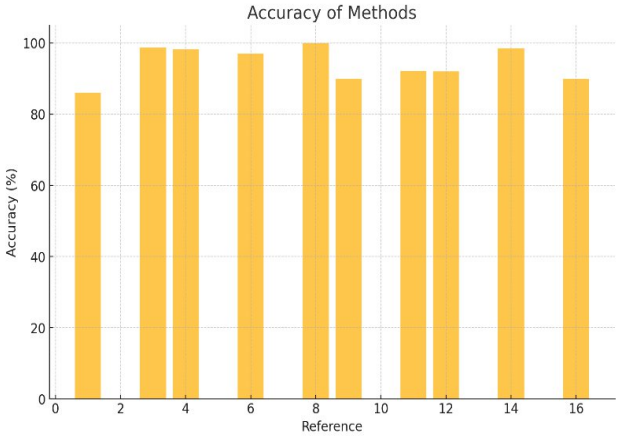

Fig. 2. Accuracy Analysis

D.Multi-Modal Deepfake Detection and Emerging Threats

Deepfake content spans many different media, including facial expressions, speech, and body movements. So, the research issue is very complex. The framework developed in [12] solved this challenge; hence, it integrated the classifiers of uni-modal and cross-modal for manipulation detection over the facial, vocal, and textual elements. For the D-Fence layer, the detection accuracy is as high as 92%, whereas the adversarial attack scenarios outperformed the existence of multi-modal detection frameworks. Another critical aspect of multi-modal deepfakes involves the audio domain. Work in [14] proposed a unified framework to detect deepfake audio by fusing MFCC and spectral contrast features. The model had optimized the feature extraction process and utilized a lightweight ANN architecture that achieved accuracy as high as 98.52% and proved the system robust enough for distinguishing between original and deepfake audio recordings. Despite advancements in detecting deepfakes, several challenges persist. Improved deepfake generation techniques keep

continuously raising the bar for detection models by forcing them to stay ahead of novel attacks. This work in [13] exposed the weakness of the current detection frameworks and suggested that more robust biometric signal analysis, such as Remote Photoplethysmography (rPPG), is required to detect physiological inconsistencies in manipulated videos. Therein, the need for more robust legal frameworks and platform-level interventions was pointed out in [15], on the role that government-industry cooperation could play in regulating deepfake dissemination during critical periods of election time, for example. Future research directions should now focus on developing hybrid detection systems by combining machine learning and cognitive science with sensor technologies in real time. While deepfake generation develops in a more sophisticated direction, especially in the low-resource environment, detection models have to grow hand-in-hand. The work in [16] demonstrated the potential of fusing hand-crafted and deep-learned features and thus provided a promising direction toward future multi-modal deepfake detection systems.

IV - STATISTICAL COMPARATIVE RESULT ANALYSIS

This section presents a comparison of various deepfake detection methods, considering their efficiency and performance for different metrics of detection. Several papers were chosen; several techniques are used in each: from deep learning models and hybrid models that use both machine learning and blockchain to new approaches, such as biosensors for real-time analysis. This review makes a comprehensive comparison among various methods developed for deepfake content detection concerning the effectiveness of detection based on key performance indicators like accuracy, false positive rates, and adaptability. Observations in the paper provide a view of the strengths and limitations of each technique, therefore showing their various applicative contexts in the deepfake analysis.

Table 2. Statistical Review of Existing Methods

Reference	**Method Used**	**Results (In Numerical Form)**	**Effi Dee Ana**
[1]	Ensemble deep learning-based system using optical flow	Accuracy: 86.02% (DeepFake subset), 85.7% (FaceSwap subset)	Moc suit ima anal
[2]	Blockchain-based federated learning with SegCaps and CNN	Accuracy: 6.6% improvement over benchmarks, AUC: 5.1% improvement	High; performs well i handling large dataset
[3]	Plasmonic resonance-enhanced biosensor integrated with CNN	Accuracy: 98.7%, FPR: 1.2%, FNR: 0.5%, Response time: 0.8s	Extremely high; excel in real-tim detection
[4]	Human cognitive abilities combined with machine learning	Accuracy: 98.3% in predicting video authenticity	High; effective when huma cognition included
[5]	Neurocognitive testing with deepfake voice matching	Intermediate deception rate; accuracy of brain signal decoding is inferred	Moderate; focuses o auditory deepfakes
[6]	CNN with vision transformer	Accuracy: 97% (CNN), 85% (CViT model)	High fo CNN; moderate fo CViT
[7]	Prior information impact and audio quality metric for speech deepfakes	Detection accuracy highly dependent on prior	Moderate; effective fo auditory deepfakes but lack

		knowledge; approximately 80-90% accuracy in different conditions	mul capa
[8]	Classification algorithms with ensemble model	Accuracy: 99.97% (high-resolution), 98.27% (low-resolution), 98.72% (mixed datasets)	Very exce and reso dete
[9]	Cross-modal attention architecture for audio and video	Accuracy: ~90% (approx.)	High hanc mul deep effe
[10]	ResNext 50 + LSTM with Big Data techniques	Accuracy: Outperforms competitors, resource-efficient	High exce bala effic and perf
[11]	Shallow vision transformer	Accuracy: 92.15% (RFF), 88.52% (RFFD)	High opti cons envi
[12]	D-Fence framework with uni-modal and cross-modal classifiers	Accuracy: 92% under adversarial attacks	High resil adve conc
[13]	Remote Photoplethysmography (rPPG) and DeepPhys	High availability for biological signal detection	High innc for physiologica deepfake detection
[14]	Unified framework using MFCC and spectral contrast for audio deepfakes	Accuracy: 98.52%	Very high excellent fo audio deepfake detection
[15]	DefakeHop++ layered model	Effective in platform-wide detection during elections	Moderate; applicable for platform level intervention
[16]	Hand-crafted and deep-learned feature fusion	Accuracy: ~90% across datasets	High; performs well wit feature fusion

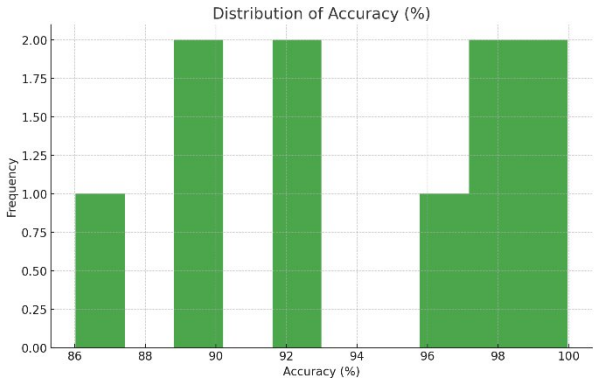

Fig. 3. Model's Distribution of Accuracy Levels

In this comparative review, various deepfake detection methods are discussed; all of them have better performance compared to others in certain scenarios. Those models that combine multiple data modalities-for example, cross-modal architectures or multi-feature ensembles-exhibit high accuracy with different datasets and samples. However, resource requirements and computational complexity are among certain counterbalancing factors, especially for much more

sophisticated technologies, which also include blockchain-based federated learning and real-time biosensors. Although lightweight models, like shallow vision transformers, do present workable solutions in constrained environments, they inherently have to give away some degree of accuracy for the benefit of speed compared to deep learning.

V- CONCLUSION & FUTURE SCOPES

This paper reviewed state-of-the-art deepfake detection methods, indicating a remarkable contribution toward fighting the evolution of synthetic media. Several models were reviewed ranging from classical deep learning methods like CNNs and vision transformers to more novelty approaches involving blockchain-based federated learning to plasmonic resonance-enhanced biosensors. Among these, it is quite fair to say that no single technique is applicable for a wide range of deepfake types, from facial manipulations to audio deepfakes, and even multimodal forgeries. However, some models perform better in context because they were intrinsically designed and optimized for those contexts. For example, CNN-based models have remained prevalent for image and video-based detection of deepfakes. That makes them very effective in detecting facial manipulations, especially for extracting intrinsic features from visual content. Notably, some methods, like ensemble deep learning with optical flow techniques, demonstrate superior performance for motion-based analysis and therefore would be well-suited for applications where temporal consistency in videos is one of the key indicators of deepfake content. Moreover, performance raises by combining the usage of CNNs with temporal networks like LSTMs or hybrid architectures such as ResNext 50 + LSTM [10]. The latter is very suitable for real-world applications in which a video is required to be analyzed fast enough, like immediate video recognition. Models such as these are very useful in media and social networks, where runtime verification of content is highly desired. In contrast, blockchain-based federated learning models allow a high amount of value to privacy-preserving and decentralized model training of a use case. Such techniques will work well in highly sensitive environments, like healthcare and government applications. Using federated learning, these models are enabled to train on distributed datasets with leakage of sensitive information. This method also enhances model robustness, considering many sources of data improve its generalizability via various kinds of deepfakes. Biosensor-based deepfake detection models, in turn, proposed by [3], allow excellent performance of deepfake detection in real time, especially in applications related to security and surveillance. Such models are performing very well in environments where detection needs to be done right away, as their embedding of plasmonic resonance technology provides the capacity for high-sensitive processing of visual anomalies at a fast pace in the process.

Such a shallow architecture, such as a shallow Vision Transformer, is helpful in resource-constrained setups where computational efficiency is prioritized with no significant compromise in the accuracy of detection. These light models reduce computational cost due to fewer parameters and FLOPS, hence ideal to be used for mobile or edge computing applications. While they may not achieve the same level of accuracy as more complex deep learning architectures, efficiency is certainly a value that they hold in real-world scenarios where scalability with low-latency detection is required. Models based on cross-modal attention mechanisms also bear increasing relevance. They provide robust multi-modal deepfake detection and handle audio and visual data together powerfully. These models play a crucial role in digital forensics, for example, where multi-model deepfakes are one of the most difficult challenges since the analysts have to analyze multiple streams of information concurrently. Although significant enhancement has been achieved so far, challenges persist, especially concerning generalization and adaptability. Considering the continuous evolution in techniques for generating deepfakes, one avenue related to future research deals with how effective models of detection can remain against new and more sophisticated forgeries. In addition, the increasing prevalence of adversarial attacks also poses a great challenge, since most of the existing detection models are much affected under adversarial conditions. Future research may lie in developing more attack-resilient models by incorporating methods of adversarial training or robust feature extraction process.

Besides that, there is great scope for multi-modal detection frameworks that integrate data from multiple sources, such as audio and video, alongside textual information. This would definitely be crucial in dealing with deepfakes spanning multiple domains, given the fact that single-modality detection models are already not good enough for complex forgeries. Another major future scope of the process is the investigation into real-time deepfake detection systems, mainly for live streaming platforms and social media. Real-time analysis, in particular, acquires much significance in view of fast dissemination of fake content online, which

requires speedy and efficient detection solutions to operate on live data samples. In a nutshell, deepfake detection model analysis underlines the importance of context-specific solutions. While CNN-based models continue to dominate in the landscape of visual deepfake detection, more specialized models involving federated learning and real-time biosensors offer different advantages with respect to privacy-centric and real-time applications, respectively. This includes lightweight models such as shallow transformers that may provide scalable solutions to resource-limited environments, while cross-modal approaches will be ever more necessary with multi-model deepfakes. The future of deepfake detection will involve integrating these various techniques and further underlines the need for adaptable, multi-modal, resilient detection systems that will evolve in concert with rapid improvements in deepfake generation technologies.

REFERENCES

[1] *Vashishtha, S., Gaur, H., Das, U. et al. Optifake: optical flow extraction for deepfake detection using ensemble learning technique. Multimed Tools Appl (2024). https://doi.org/10.1007/s11042-024-18641-x*

[2] *Heidari, A., Navimipour, N.J., Dag, H. et al. A Novel Blockchain-Based Deepfake Detection Method Using Federated and Deep Learning Models. CognComput **16**, 1073–1091 (2024). https://doi.org/10.1007/s12559-024-10255-7*

[3] *Maheshwari, R.U., Kumarganesh, S., K V M, S. et al. Advanced Plasmonic Resonance-enhanced Biosensor for Comprehensive Real-time Detection and Analysis of Deepfake Content. Plasmonics (2024). https://doi.org/10.1007/s11468-024-02407-0*

[4] *Salini, Y., HariKiran, J. DeepFake Videos Detection Using Crowd Computing. Int. j. inf. tecnol. (2023). https://doi.org/10.1007/s41870-023-01494-2*

[5] *Roswandowitz, C., Kathiresan, T., Pellegrino, E. et al. Cortical-striatal brain network distinguishes deepfake from real speaker identity. CommunBiol **7**, 711 (2024). https://doi.org/10.1038/s42003-024-06372-6*

[6] *Soudy, A.H., Sayed, O., Tag-Elser, H. et al. Deepfake detection using convolutional vision transformers and convolutional neural networks. Neural Comput&Applic (2024). https://doi.org/10.1007/s00521-024-10181-7*

[7] *Malinka, K., Firc, A., Šalko, M. et al. Comprehensive multiparametric analysis of human deepfake speech recognition. J Image Video Proc. **2024**, 24 (2024). https://doi.org/10.1186/s13640-024-00641-4*

[8] *Pandey, M., Singh, S., Malik, A. et al. Detecting low-resolution deepfakes: an exploration of machine learning techniques. Multimed Tools Appl **83**, 66283–66298 (2024). https://doi.org/10.1007/s11042-024-18235-7*

[9] *Asha, S., Vinod, P. & Menon, V.G. A defensive attention mechanism to detect deepfake content across multiple modalities. Multimedia Systems **30**, 56 (2024). https://doi.org/10.1007/s00530-023-01248-x*

[10] *Kumar, N., Kundu, A. Cyber Security Focused Deepfake Detection System Using Big Data samples. SN COMPUT. SCI. 5, 752 (2024). https://doi.org/10.1007/s42979-024-03105-8*

[11] *Usmani, S., Kumar, S. & Sadhya, D. Efficient deepfake detection using shallow vision transformer. Multimed Tools Appl **83**, 12339–12362 (2024). https://doi.org/10.1007/s11042-023-15910-z*

[12] *S, A., P, V., Amerini, I. et al. D-Fence layer: an ensemble framework for comprehensive deepfake detection. Multimed Tools Appl **83**, 68063–68086 (2024). https://doi.org/10.1007/s11042-024-18130-1*

[13] *Xu, Q., Qiao, H., Liu, S. et al. Deepfake detection based on remote photoplethysmography. Multimed Tools Appl **82**, 35439–35456 (2023). https://doi.org/10.1007/s11042-023-14744-z*

[14] *Jellali, A., Ben Fredj, I. & Ouni, K. Pushing the boundaries of deepfake audio detection with a hybrid MFCC and spectral contrast approach. Multimed Tools Appl (2024). https://doi.org/10.1007/s11042-024-19819-z*

[15] *Pranay Kumar, B., Shaheer Ahmed, M. & Sadanandam, M. Designing a Safe Ecosystem to Prevent Deepfake-Driven Misinformation on Elections. DISO **3**, 19 (2024). https://doi.org/10.1007/s44206-024-00107-0*

[16] *Megahed, A., Han, Q. & Fadl, S. Exposing deepfake using fusion of deep-learned and hand-crafted features. Multimed Tools Appl **83**, 26797–26817 (2024). https://doi.org/10.1007/s11042-023-16329-2.*