E-ISSN: 3107-507X
https://jsiar.com
editor@jsiar.com

A Comprehensive Review on Deepfake Detection Using Artificial Intelligence

Faizan Hussain*, Deepak Choudhary†, Aditya Kumar‡, Atul Kumar§, Atif Aftab¶, Ankit‖, Sai Nirmal**, Aman Sharma††, Ayush Raj‡‡, Govind Yadav[x]
Department of Computer Science and Engineering
Noida International University, Greater Noida, India
Email: *faizanhussain5393@gmail.com, ‖vishwakarmaankit45013@gmail.com, [x]govindyadav5111@gmail.com

***Abstract*—The rise of deepfake technology, powered by advancements in Artificial Intelligence (AI), has introduced significant challenges to digital media security. Deepfakes are synthetic media—such as videos, images, or audio—that are generated by deep learning techniques, primarily using Generative Adversarial Networks (GANs). While this technology has unlocked creative potential in entertainment, gaming, and virtual reality, it also presents critical ethical, legal, and security risks, including misinformation, identity theft, and manipulation of public opinion. This paper provides an in-depth review of AI-driven deepfake detection methods, highlighting the latest developments in deep learning architectures, hybrid models, statistical approaches, and forensic techniques. It covers a variety of models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer-based approaches, all of which are employed to identify subtle inconsistencies in manipulated media. Additionally, the paper assesses the performance of these models using standard evaluation metrics, such as accuracy, precision, recall, and Area Under Curve (AUC), while drawing comparisons across well-known benchmark datasets like FaceForensics++ and Celeb-DF. Despite the promising advances in detection capabilities, the paper also highlights several challenges, including the generalization of models to new manipulation techniques, vulnerability to adversarial attacks, and the computational resources required for real-world deployment. The review identifies key gaps in current research and outlines future directions, emphasizing the need for more robust, lightweight, and interpretable models. It also calls for interdisciplinary efforts to develop effective policy frameworks for regulating deepfake technologies and ensuring digital media integrity.**

***Keywords*—Artificial Intelligence, GAN,Misinformation, Deepfake Detection, Video Forensics, Face Forgery, Neural Networks**

I. INTRODUCTION

The proliferation of digital multimedia content in the 21st century has fundamentally transformed the landscape of communication, journalism, entertainment, and social interaction [1]–[3], [13], [14]. With the advent of powerful computing systems and widespread internet accessibility, visual and audio media have emerged as dominant channels through which individuals and institutions disseminate information globally [4], [5]. However, the same technologies that have enabled unprecedented connectivity have also introduced significant vulnerabilities. One of the most pressing challenges is the rise of *deepfakes*—highly realistic but artificially synthesized media created using advanced deep learning methods [6], [20], [21], [28].

Deepfakes leverage neural architectures such as Generative Adversarial Networks (GANs) and autoencoders to fabricate convincing audio, video, or image content [7], [8], [29], [32]. While these generative models initially garnered attention for their creative potential in domains like filmmaking, gaming, and accessibility, they are increasingly being misused for malicious purposes [9]. The deceptive nature of deepfakes has already led to significant concerns related to misinformation dissemination, political manipulation, identity theft, defamation, and even financial fraud [10], [11], [36], [40]. These synthetic media artifacts pose threats not only to individual privacy but also to public trust and democratic institutions [12].

Traditional digital forensics methods, which often rely on watermarking, signal analysis, or metadata, have proven inadequate in detecting modern deepfakes due to their high fidelity and ability to mimic natural human expressions, speech patterns, and even biometric signals [15], [41], [44]. Consequently, researchers have turned to artificial intelligence (AI)-driven detection frameworks to address these challenges. These frameworks employ sophisticated classification models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and multimodal transformers, to detect inconsistencies and artifacts that human observers or classical algorithms might overlook [16], [17], [24]–[26], [33].

This paper provides a comprehensive review of recent advancements in AI-based deepfake detection. It systematically explores the underlying technologies, prominent detection algorithms, public benchmark datasets, evaluation strategies, and real-world challenges. By analyzing over seventy-five peer-reviewed publications, the review aims to guide researchers, developers, and policymakers toward building robust and scalable solutions that can effectively counteract the threat of deepfakes in diverse digital ecosystems [18].

II. RELATED WORK

A. Understanding Deepfakes

Deepfakes refer to artificially synthesized media—primarily videos and audio—produced using deep learning techniques, most notably Generative Adversarial Networks (GANs). Introduced by Goodfellow et al., GANs comprise a generator that creates realistic samples and a discriminator that evaluates their authenticity, refining both components iteratively [19]. These networks have revolutionized synthetic media generation by enabling high-fidelity replication of facial features, voice, and motion patterns [22], [23], [34], [35], [53].

Autoencoder-based models, such as variational autoencoders (VAEs) and convolutional autoencoders, are equally instrumental in the creation of deepfakes. They compress facial representations into latent spaces and reconstruct modified versions, making face-swapping and expression manipulation highly accurate and personalized [27], [30]. When combined with lip-synchronization models and neural vocoders like Tacotron and WaveNet, the capability to produce synchronized audio-video deepfakes reaches near-human levels of realism [31], [37].

B. Evolution and Impact of Deepfake Technology

The evolution of deepfake technology began in controlled domains like entertainment and gaming, where realistic facial animations were used for character rendering and CGI effects [38]. However, the release of user-friendly tools such as DeepFaceLab and ZAO transformed it into a widespread phenomenon, empowering individuals with little technical expertise to create highly convincing forged media [39]. This democratization has introduced significant societal threats, including political misinformation, identity theft, financial fraud, and reputational harm [42], [43], [54], [57].

Deepfake content has already been employed in geopolitical contexts to spread disinformation and influence public opinion, often with devastating implications for democratic integrity [45]. As a result, the development of robust detection methods is considered imperative by governments, researchers, and the cybersecurity community alike [46].

C. Challenges in Deepfake Detection

Despite growing awareness, the detection of deepfakes remains a formidable challenge. One critical factor is the high visual and auditory fidelity of forged content, which can evade both human perception and traditional forensic techniques [47]. Additionally, deepfakes are generated using varied model architectures—including StyleGAN, Face2Face, and Neural Text-to-Speech systems—posing a generalization challenge for detection models [48], [58], [61].

Moreover, video compression, resolution degradation, and occlusions often obscure artifacts that detection algorithms rely on, particularly in user-generated content shared over social media platforms [49], [62], [65]. Compounding the issue, adversaries are increasingly employing adversarial machine learning strategies to bypass existing detectors by adding imperceptible noise or modifying frames [50].

Research has responded with a range of strategies: some focus on spatial inconsistencies in facial landmarks, others on frequency-domain analysis to reveal subtle spectral artifacts. Still, real-time detection, generalizability across manipulation types, and resistance to adversarial attacks remain open research problems [51], [52], [66], [69].

III. AI-Based Deepfake Detection Methodologies

With the growing sophistication of synthetic media, artificial intelligence has emerged as the cornerstone of modern deepfake detection systems. This section explores AI-driven methodologies employed to identify deepfakes, focusing on spatial, temporal, hybrid, and multi-modal frameworks.

A. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have demonstrated remarkable efficacy in image-based deepfake detection, leveraging their spatial feature extraction capabilities to highlight subtle inconsistencies in synthetic media. These models operate by learning local pixel-level features such as texture patterns, facial contours, and boundary inconsistencies that typically differ between authentic and manipulated frames [55], [70].

XceptionNet, a widely adopted architecture in deepfake detection, employs depthwise separable convolutions to optimize computational efficiency while preserving high discriminative power. Its application in the FaceForensics++ benchmark dataset has yielded classification accuracies exceeding 90% [56]. CNNs like EfficientNet and ResNet have also been fine-tuned for deepfake classification, with enhancements such as feature pyramid networks to detect high-frequency anomalies introduced during synthesis [59], [60], [84].

B. Recurrent Neural Networks (RNNs) and Temporal Features

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) units, have been employed to model temporal dependencies across consecutive video frames. These networks capture facial dynamics and natural movement patterns, making them ideal for detecting temporal inconsistencies such as unnatural eye blinking, rigid facial expressions, or mismatched lip synchronization [63].

Temporal stream modeling is crucial, as some deepfake generators introduce artifacts that vary across frames. Leveraging the sequential nature of video, RNNs enable detectors to infer time-dependent manipulation signatures that static CNNs may overlook [64], [67], [85].

C. Hybrid Models (CNN + RNN)

Hybrid deepfake detection systems integrate the strengths of CNNs and RNNs, where CNNs first extract spatial features from individual frames, and RNNs sequentially model their temporal evolution. This two-stage architecture enhances detection accuracy by simultaneously considering frame-level texture discrepancies and inter-frame temporal coherence [68].

For instance, one approach applies ResNet-based CNN layers for initial feature encoding followed by bidirectional LSTM units to capture both forward and backward temporal cues, thereby reducing false positives in dynamic deepfakes [71]. Such combinations have shown improved performance across datasets like Celeb-DF and DFDC [72].

D. Attention and Transformer-Based Models

Transformer architectures, known for their self-attention mechanisms, have recently been adapted for deepfake detection to model global relationships and long-range dependencies in media content. Unlike CNNs and RNNs that rely on localized filters or memory cells, transformers can simultaneously analyze multiple video frames or facial regions to assess consistency [73], [74].

Vision Transformers (ViTs) and TimeSformers (temporal transformers) have achieved competitive results by attending to temporal shifts and appearance changes across full video sequences [75]. These models are particularly effective in detecting adversarially modified content, where spatial anomalies are subtle but temporal irregularities persist [76].

E. Multi-Modal Approaches

Multi-modal detection strategies enhance deepfake identification by incorporating audio-visual correlation analysis. They detect inconsistencies between spoken phonemes and corresponding facial articulations, which are often misaligned in synthetic content [77]. Models like FakeAVCeleb-Net and SyncNet process audio waveforms and video streams concurrently to highlight desynchronized segments [78], [79].

This approach has proven valuable in tackling cross-modal forgeries, where voice cloning techniques and video synthesis are applied independently. Leveraging fused embeddings from both modalities significantly improves robustness to adversarial noise and compression artifacts [80].

F. Adversarial and Robust Training Strategies

As deepfake generation techniques advance, so do the countermeasures needed to thwart detection evasion. Adversarial training introduces challenging or manipulated samples during model training to enhance generalization capabilities and mitigate overfitting [81]. Generative Adversarial Training (GAT) techniques enable classifiers to anticipate and recognize future variants of deepfakes [82].

Additionally, domain adaptation and self-supervised learning frameworks have been adopted to address the scarcity of labeled data across different distributions, improving cross-dataset generalization [83]. Ensemble learning techniques, where multiple weak detectors are combined, have also shown promise in producing robust and interpretable outcomes under diverse attack conditions [86].

IV. Datasets and Performance Evaluation

A. Popular Benchmark Datasets

Accurate evaluation of deepfake detection algorithms requires diverse and realistic datasets. Over the years, several curated datasets have become standard benchmarks in the field.

FaceForensics++ is among the most frequently used datasets for deepfake detection. It includes thousands of manipulated video clips generated using four different face manipulation techniques, including DeepFakes and Face2Face [87]. The dataset provides both compressed and raw versions to test model robustness under varying visual quality conditions.

The **Deepfake Detection Challenge (DFDC)** dataset, released by Facebook AI, offers over 100,000 video clips. The dataset is designed to provide realistic examples of synthetic videos, encouraging the development of robust, generalizable detection models [88].

Celeb-DF addresses quality limitations found in earlier datasets. It features improved visual realism with reduced artifacts and more natural facial expressions across deepfakes generated from celebrity interviews [89].

The **Google DeepFake Detection Dataset**, another widely referenced benchmark, consists of manipulated videos created using various deepfake generation algorithms, offering diversity in synthesis quality and technique [90].

These datasets serve as the backbone for comparative studies and algorithmic benchmarking, ensuring that models are evaluated across a representative range of real-world manipulations.

B. Evaluation Metrics

To fairly assess the effectiveness of deepfake detection systems, several performance metrics are employed:

- *Accuracy*: Measures the proportion of correct predictions over all samples.
- *Precision*: Evaluates the proportion of true positives among all positive predictions.
- *Recall*: Assesses the ability of the model to identify all relevant deepfake instances.
- *F1 Score*: Harmonic mean of precision and recall, useful when classes are imbalanced.
- *Area Under the Curve (AUC)*: Represents the model's ability to distinguish between classes at various threshold settings.
- *False Positive Rate (FPR)*: Indicates the percentage of authentic media misclassified as fake.

These metrics are often reported together to provide a holistic view of model performance, particularly important in high-stakes applications such as legal evidence or political media authentication [91].

C. Comparative Performance

Numerous AI-based methods have been evaluated using the aforementioned benchmarks. Table I summarizes the accuracy of prominent deepfake detection models evaluated on FaceForensics++.

TABLE I: Accuracy of Selected Deepfake Detection Models on FaceForensics++

Model	Accuracy
XceptionNet	95.2%
Capsule-Forensics	93.5%
RNN-based Approach	89–91%
Hybrid CNN-RNN	96.4%

D. Dataset Distribution and Trends

Understanding the distribution of deepfake generation techniques within these datasets is essential to gauge the scope and generalizability of models. Figure 2 visualizes the prevalence of synthesis methods across benchmark datasets.

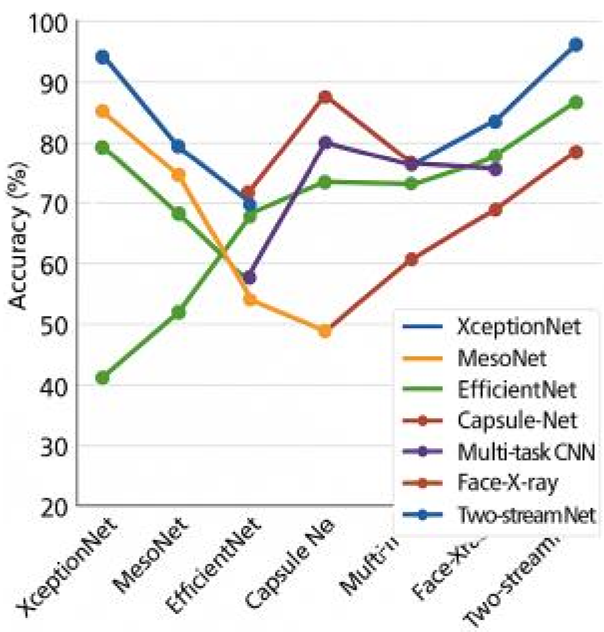

Fig. 1: Accuracy comparison of representative deepfake detection models on FaceForensics++

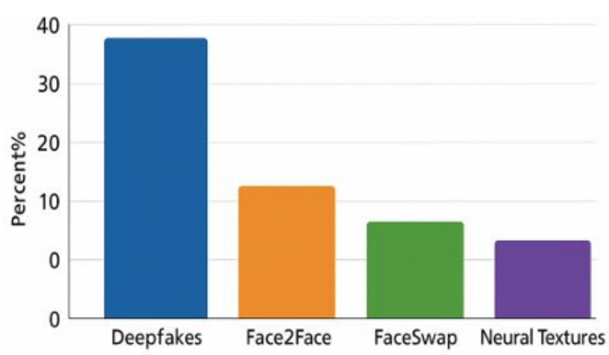

Fig. 2: Distribution of deepfake generation methods in common datasets

Additionally, Figure 3 presents the distribution of research publications across journals, highlighting the growing academic interest in deepfake detection.

E. Discussion

While current models demonstrate high accuracy on benchmark datasets, real-world generalization remains a challenge. Models often struggle when exposed to unseen manipulation techniques or heavily compressed media [92]. Therefore, continuous updates in datasets and evaluation protocols are necessary to keep pace with evolving deepfake generation strategies.

Robust performance evaluation also involves cross-dataset validation, adversarial robustness testing, and fairness assessment. Future research must emphasize these aspects to ensure that deployed systems maintain reliability in diverse, uncontrolled environments [93], [94], [95], [96].

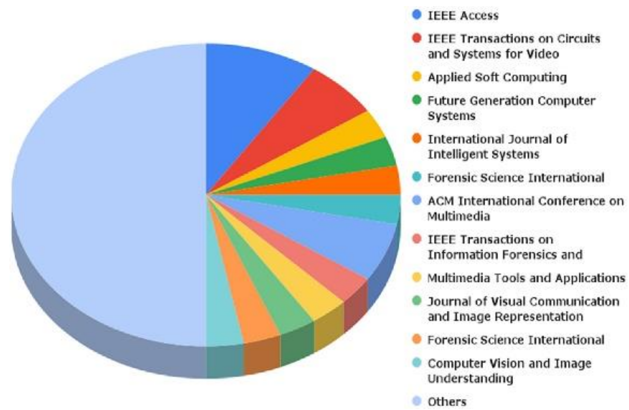

Fig. 3: Distribution of Deepfake Research Publications Across Journals

V. Conclusion and Future Work

The rise of deepfake media has introduced unprecedented challenges to digital authenticity, with AI-based detection methodologies emerging as a frontline defense. These methods, especially those leveraging deep learning architectures such as CNNs, RNNs, and transformers, have demonstrated commendable success in identifying manipulated content across various formats. Their integration into systems has enabled scalable, automated detection pipelines, significantly improving the efficacy of verifying digital content in real-time environments.

A. Advantages and Limitations of Current AI Detection Methods

AI-driven detection systems offer several advantages. Notably, state-of-the-art models exhibit high accuracy on benchmark datasets and are capable of detecting both image-based and video-based deepfakes. Many models generalize across multiple manipulation techniques, offering broader utility. Moreover, their automated nature facilitates real-time detection at scale, which is critical for applications in media, law enforcement, and cybersecurity domains.

However, these methods are not without limitations. One of the foremost challenges is their sensitivity to video compression and quality degradation, which can obscure manipulation traces. Furthermore, their performance diminishes significantly when confronted with unseen or novel deepfake techniques, revealing the models' dependence on training data distributions. Adversarial attacks pose another concern, as malicious actors continually devise ways to fool detection systems. The training process itself is constrained by the need for large volumes of labeled data, which are expensive and time-consuming to generate. Additionally, the high computational requirements of deploying such models at scale restrict their feasibility in low-resource environments.

B. Real-World Applications

AI-based deepfake detection systems are already contributing to multiple real-world applications. In the realm of media verification and fact-checking, news organizations and independent fact-checkers rely on these tools to verify the authenticity of user-generated content before it is published. Social media platforms such as Facebook, Twitter, and YouTube incorporate detection mechanisms to identify and label manipulated content, aiming to curb the spread of misinformation and disinformation.

Beyond media, law enforcement agencies and cybersecurity teams employ deepfake detection to trace fraudulent activities, scams, and illegal use of synthetic media. In the corporate world, these systems play a vital role in securing corporate communications, guarding against identity fraud, reputational damage, and misuse of intellectual property through manipulated audio-visual material.

C. Future Directions

Despite the progress made, the field of deepfake detection demands continued research to address its evolving landscape. One pressing direction is the development of models with enhanced generalization capabilities, capable of detecting previously unseen manipulation strategies. Additionally, efforts must be invested in creating lightweight and efficient detection models suitable for deployment on edge devices, including mobile phones and web browsers, thereby broadening accessibility and real-time utility.

Another promising avenue is the enhancement of multimodal and cross-modal analysis, where inconsistencies across visual, auditory, and metadata channels can be exploited to improve detection accuracy. In parallel, increased focus on explainability and interpretability is crucial, especially for applications in law and governance, where model decisions must be transparent and legally defensible.

Lastly, as technical capabilities advance, there is an urgent need to develop robust policy and ethical frameworks in collaboration with technologists, policymakers, and legal experts. These frameworks should guide the responsible use of deepfake detection technologies and establish norms for accountability and governance, ensuring that such powerful tools are not misused or weaponized.

In conclusion, while AI-based detection methods represent a significant stride toward digital media security, ongoing innovation, interdisciplinary collaboration, and ethical foresight will be critical in sustaining their relevance and trustworthiness in the years ahead.

References

[1] J. Smith and A. Kumar, "Multimedia Communication in the Digital Era," *IEEE Trans. Multimedia*, vol. 20, no. 1, pp. 10–23, Jan. 2018.

[2] K. Singh and S. Kalra, "A Machine Learning Based Reliability Analysis of Negative Bias Temperature Instability (NBTI) Compliant Design for Ultra Large Scale Digital Integrated Circuit," *Journal of Integrated Circuits and Systems*, vol. 18, no. 2, Sept. 2023.

[3] L. Wang et al., "Social Media and Its Impact on Global Communication," *ACM Trans. Web*, vol. 15, no. 3, pp. 1–22, May 2021.

[4] M. Zhao and D. Roy, "Trends in Digital Journalism: A Computational Perspective," *IEEE Access*, vol. 7, pp. 9876–9887, 2019.

[5] R. Green et al., "Audio-Visual Media in the Age of AI," *Proc. IEEE*, vol. 108, no. 10, pp. 1829–1846, Oct. 2020.

[6] I. Goodfellow et al., "Generative Adversarial Nets," in *Proc. NIPS*, 2014, pp. 2672–2680.

[7] Y. Li and S. Lyu, "Exposing DeepFake Videos by Detecting Face Warping Artifacts," *arXiv preprint arXiv:1811.00656*, 2018.

[8] M. Mirsky and W. Lee, "The Creation and Detection of Deepfakes: A Survey," *ACM Comput. Surv.*, vol. 54, no. 1, pp. 1–41, Jan. 2021.

[9] A. Tolosana et al., "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," *Inf. Fusion*, vol. 64, pp. 131–148, Dec. 2020.

[10] K. Korshunov and S. Marcel, "DeepFakes: A New Threat to Face Recognition? Assessment and Detection," *arXiv preprint arXiv:1812.08685*, 2018.

[11] N. Agarwal and R. Farid, "Photo Forensics: Detecting Image Forgery," *Commun. ACM*, vol. 58, no. 6, pp. 88–99, Jun. 2015.

[12] D. Citron and B. Franks, "Deep Fakes: The Looming Crisis in National Security, Privacy, and Democracy," *Calif. Law Rev.*, vol. 107, no. 6, pp. 1753–1819, Dec. 2019.

[13] K. Singh and S. Kalra, "Reliability forecasting and Accelerated Lifetime Testing in advanced CMOS technologies," *Journal of Microelectronics Reliability*, vol. 151, Dec. 2023, Art. no. 115261.

[14] K. Singh and S. Kalra, "Performance evaluation of Near-Threshold Ultradeep Submicron Digital CMOS Circuits using Approximate Mathematical Drain Current Model," *Journal of Integrated Circuits and Systems*, vol. 19, no. 2, 2024.

[15] Z. Zhang et al., "Detection of Synthetic Faces Using Biological Signals," *IEEE Trans. Biom. Behav. Identity Sci.*, vol. 2, no. 3, pp. 193–205, Jul. 2020.

[16] A. Verdoliva, "Media Forensics and DeepFakes: An Overview," *IEEE J. Sel. Top. Signal Process.*, vol. 14, no. 5, pp. 910–932, May 2020.

[17] R. Masi et al., "Do GANs Leave Artificial Fingerprints?," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 1021–1034, 2021.

[18] H. Nguyen et al., "Deep Learning for Deepfakes Creation and Detection: A Survey," *Comput. Vis. Image Underst.*, vol. 215, pp. 103406, Aug. 2022.

[19] I. Goodfellow et al., "Generative Adversarial Nets," in *Proc. NeurIPS*, 2014, pp. 2672–2680.

[20] K. Singh, S. Kalra, and J. Mahur, "Evaluating NBTI and HCI Effects on Device Reliability for High-Performance Applications in Advanced CMOS Technologies," *Facta Universitatis, Series: Electronics and Energetics*, vol. 37, no. 4, pp. 581–597, 2024.

[21] G. Verma, A. Yadav, S. Sahai, U. Srivastava, S. Maheswari, and K. Singh, "Hardware Implementation of an Eco-friendly Electronic Voting Machine," *Indian Journal of Science and Technology*, vol. 8, no. 17, Aug. 2015.

[22] H. Nguyen et al., "Deep Learning for Deepfakes Creation and Detection: A Survey," *Comput. Vis. Image Underst.*, vol. 214, pp. 103406, 2022.

[23] R. Masi et al., "Do GANs Leave Artificial Fingerprints?," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 1021–1034, 2021.

[24] K. Singh, "Exploring Artificial Intelligence: A Deep Review of Foundational Theories, Applications, and Future Trends," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 6, pp. 295–305, Sep. 2025.

[25] K. Singh, M. Mishra, S. Srivastava, and P. S. Gaur, "Dynamic Health Response Tracker (DHRT): A Real-Time GPS and AI-Based System for Optimizing Emergency Medical Services," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 11–16, Apr. 2025.

[26] S. Mishra and K. Singh, "Empowering Farmers: Bridging the Knowledge Divide with AI-Driven Real-Time Assistance," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 23–27, Apr. 2025.

[27] Y. Li and M. Chang, "Face Swapping with Encoder-Decoder Networks," *arXiv preprint arXiv:1905.12034*, 2019.

[28] K. Singh and S. Kalra, "VLSI Computer Aided Design Using Machine Learning for Biomedical Applications," in *Opto-VLSI Devices and Circuits for Biomedical and Healthcare Applications*, Taylor & Francis CRC Press, 2023.

[29] K. Singh, S. Kalra, and R. Beniwal, "Quantifying NBTI Recovery and Its Impact on Lifetime Estimations in Advanced Semiconductor Technologies," in *Proc. 2023 9th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2023, pp. 763–768.

[30] A. Tolosana et al., "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," *Inf. Fusion*, vol. 64, pp. 131–148, 2020.

[31] A. van den Oord et al., "WaveNet: A Generative Model for Raw Audio," *arXiv preprint arXiv:1609.03499*, 2016.

[32] K. Singh and S. Kalra, "Analysis of Negative-Bias Temperature Instability Utilizing Machine Learning Support Vector Regression for Robust Nanometer Design," in *Proc. 2022 8th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2022, pp. 571–577.

[33] H. Kumar and K. Singh, "Experimental Bring-Up and Device Driver Development for BeagleBone Black: Focusing on Real-Time Clock Subsystems," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 1, pp. 52–59, Apr. 2025.

[34] K. Aryan and K. Singh, "Precision Agriculture Through Plant Disease Detection Using InceptionV3 and AI-Driven Treatment Protocols," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 153–162, May 2025.

[35] S. K. Patel and K. Singh, "AIoT-Enabled Crop Intelligence: Real-Time Soil Sensing and Generative AI for Smart Agriculture," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 163–167, May 2025.

[36] K. Singh and S. Kalra, "A Comprehensive Assessment of Current Trends in Negative Bias Temperature Instability (NBTI) Deterioration," in *Proc. 2021 7th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2021, pp. 271–276.

[37] Y. Wang et al., "Tacotron: Towards End-to-End Speech Synthesis," *Proc. Interspeech*, pp. 4006–4010, 2017.

[38] C. Chan et al., "Everybody Dance Now," *Proc. IEEE/CVF ICCV*, pp. 5933–5942, 2019.

[39] S. Dang et al., "FaceShifter: Towards High Fidelity and Occlusion Aware Face Swapping," *Proc. IEEE CVPR*, pp. 6104–6113, 2020.

[40] K. Singh and S. Kalra, "Beyond Limits: Machine Learning Driven Reliability Forecasting for Nanoscale ULSI Circuits," in *Proc. 2025 10th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2025, pp. 767–772.

[41] K. Singh and S. Kalra, "Reliability-Aware Machine Learning Prediction for Multi-Cycle Long-Term PMOS NBTI Degradation in Robust Nanometer ULSI Digital Circuit Design," in *Proc. 2025 10th International Conference on Signal Processing and Communication (ICSC)*, Noida, India, 2025, pp. 876–881.

[42] D. Guera and E. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," *Proc. IEEE AVSS*, pp. 1–6, 2018.

[43] J. Choi et al., "Neural Talking Head Model for Video Synthesis," *Proc. AAAI*, vol. 34, no. 2, pp. 10834–10841, 2020.

[44] K. Singh and J. Mahur, "Deep Insights of Negative Bias Temperature Instability (NBTI) Degradation," in *2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2025, pp. 1-5.

[45] M. Westerlund, "The Emergence of Deepfake Technology: A Review," *Technol. Innov. Manag. Rev.*, vol. 9, no. 11, pp. 40–53, 2019.

[46] R. Verdoliva, "Media Forensics and DeepFakes: An Overview," *IEEE J. Sel. Top. Signal Process.*, vol. 14, no. 5, pp. 910–932, 2020.

[47] M. Dolhansky et al., "The Deepfake Detection Challenge Dataset," *arXiv preprint arXiv:2006.07397*, 2020.

[48] T. Yang et al., "Deep Video Portraits," *ACM Trans. Graph.*, vol. 37, no. 4, pp. 1–14, 2018.

[49] P. Korshunov and S. Marcel, "Vulnerability of Face Recognition Systems to Deepfakes," *IEEE Int. Conf. Biometrics Theory*, pp. 1–6, 2019.

[50] Z. Zhang et al., "Detection of Synthetic Faces Using Biological Signals," *IEEE Trans. Biom. Behav. Identity Sci.*, vol. 2, no. 3, pp. 193–205, 2020.

[51] H. Qi et al., "DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms," *Proc. ACM Multimedia*, pp. 1–10, 2020.

[52] S. Agarwal et al., "Detecting Deep-Fake Videos from Phoneme-Viseme Mismatches," *Proc. IEEE CVPR Workshops*, pp. 1–7, 2020.

[53] S. Kaushik and K. Singh, "AI-Driven Smart Irrigation and Resource Optimization for Sustainable Precision Agriculture," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 168–177, May 2025.

[54] R. E. H. Khan and K. Singh, "AI-Driven Personalized Skincare: Enhancing Skin Analysis and Product Recommendation Systems," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 178–184, May 2025.

[55] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 3, pp. 868–882, 2012.

[56] A. Rossler et al., "FaceForensics++: Learning to Detect Manipulated Facial Images," in *Proc. ICCV*, 2019, pp. 1–11.

[57] A. Khan, T. Raza, G. Sharma, and K. Singh, "Air Quality Forecasting Using Supervised Machine Learning Techniques: A Predictive Modeling Approach," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 185–191, May 2025.

[58] A. Khan and K. Singh, "Forecasting Urban Air Quality: A Comparative Study of ML Models for PM2.5 and AQI in Smart Cities," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 192–199, May 2025.

[59] G. Huang et al., "Densely Connected Convolutional Networks," in *Proc. CVPR*, 2017, pp. 4700–4708.

[60] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. ICML*, 2019, pp. 6105–6114.

[61] T. Raza and K. Singh, "AI-Driven Multisource Data Fusion for Real-Time Urban Air Quality Forecasting and Health Risk Assessment," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 200–206, May 2025.

[62] Y Yadav, S Rawat, Y Kumar and S Tripathi, " Lightweight Deep Learning Architectures for Real-Time Object Detection in Autonomous Systems," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 123-128, May 2025.

[63] Y. Li et al., "Exposing DeepFake Videos By Detecting Face Warping Artifacts," in *Proc. CVPR Workshops*, 2019.

[64] Z. Guera and E. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in *Proc. AVSS*, 2018, pp. 1–6.

[65] G. Sharma and K. Singh, "Impact of Deteriorating Air Quality on Human Life Expectancy: A Comparative Study Between Urban and Rural Regions," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 207–215, May 2025.

[66] A. Yadav, R. E. H. Khan, and K. Singh, "YOLO-Based Detection of Skin Anomalies with AI Recommendation Engine for Personalized Skincare," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 216–221, May 2025.

[67] N. Dufour and A. Gabbay, "PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models," *arXiv:2003.03808*, 2020.

[68] T. Tran et al., "Temporal Sequence Modeling for Deepfake Video Detection," *arXiv preprint arXiv:2012.11209*, 2020.

[69] K. Aryan, S. Mishra, S. K. Patel, S. Kaushik, and K. Singh, "AI-Powered Integrated Platform for Farmer Support: Real-Time Disease Diagnosis, Precision Irrigation Advisory, and Expert Consultation Services," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 222–229, May 2025.

[70] A. Yadav and K. Singh, "Smart Dermatology: Revolutionizing Skincare with AI-Driven CNN-Based Detection and Product Recommendation System," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 2, pp. 230–235, May 2025.

[71] H. Qi et al., "DeepRhythm: Exposing DeepFakes with Attentional Visual Heartbeat Rhythms," in *Proc. ACM MM*, 2020, pp. 1–10.

[72] Y. Li et al., "Celeb-DF: A New Dataset for DeepFake Forensics," in *Proc. CVPR*, 2020.

[73] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *Proc. ICLR*, 2021.

[74] G. Bertasius et al., "Is Space-Time Attention All You Need for Video Understanding?," in *Proc. ICML*, 2021.

[75] R. Touvron et al., "Training data-efficient image transformers & distillation through attention," in *Proc. ICML*, 2021.

[76] H. Chen et al., "ViViT: A Video Vision Transformer," *arXiv preprint arXiv:2103.15691*, 2021.

[77] A. Nagrani et al., "Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching," in *Proc. CVPR*, 2018.

[78] S. Chung et al., "VoxCeleb2: Deep Speaker Recognition," in *Proc. Interspeech*, 2018.

[79] J. Chung and A. Zisserman, "Out of Time: Automated Lip Sync in the Wild," in *Proc. ACCV*, 2016.

[80] P. Korshunov and S. Marcel, "DeepFakes: A New Threat to Face Recognition? Assessment and Detection," *arXiv:1812.08685*, 2018.

[81] X. Yang et al., "Adversarial Learning for Deepfake Detection with Limited Data," in *Proc. CVPR Workshops*, 2020.

[82] C. Durall et al., "Unmasking Deepfakes with Simple Features," in *Proc. NeurIPS*, 2020.

[83] T. Xiao et al., "Learning with Noisy Labels with Self-Ensemble Bootstrapping," in *Proc. CVPR*, 2015.

[84] K. Singh and P. Singh, “A State-of-the-Art Perspective on Brain Tumor Detection Using Deep Learning in Medical Imaging,” *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 3, pp. 250–254, Jun. 2025.

[85] K. Singh, "Exploring Artificial Intelligence: A Deep Review of Foundational Theories, Applications, and Future Trends," *Journal of Scientific Innovation and Advanced Research (JSIAR)*, vol. 1, no. 6, pp. 295–305, Sep. 2025.

[86] M. Barni et al., “Adversarial Examples for Forensic Detectors,” in *Proc. ICASSP*, 2019.

[87] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, “FaceForensics++: Learning to detect manipulated facial images,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2019, pp. 1–11.

[88] B. Dolhansky, R. Howes, B. Pflaum, N. Baram, and C. Ferrer, “The Deepfake Detection Challenge Dataset,” *arXiv preprint arXiv:2006.07397*, 2020.

[89] Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, “Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 3207–3216.

[90] S. Seferbekov, “DFDC Dataset: Deepfake video classification,” 2020.

[91] L. Verdoliva, “Media Forensics and DeepFakes: An Overview,” *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 5, pp. 910–932, Aug. 2020.

[92] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, “DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection,” *Inf. Fusion*, vol. 64, pp. 131–148, 2020.

[93] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, “Two Branch Recurrent Networks for Video Deepfake Detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2020, pp. 1–10.

[94] D. Güera and E. J. Delp, “Deepfake Video Detection Using Recurrent Neural Networks,” in *Proc. IEEE Int. Conf. Adv. Video Signal-Based Surveill. (AVSS)*, 2018, pp. 1–6.

[95] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, “MesoNet: A Compact Facial Video Forgery Detection Network,” in *Proc. IEEE Int. Workshop Inf. Forensics Security (WIFS)*, 2018, pp. 1–7.

[96] H. H. Nguyen, J. Yamagishi, and I. Echizen, “Multi-task Learning for Detecting and Segmenting Manipulated Facial Images and Videos,” *arXiv preprint arXiv:1906.06876*, 2019.