IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

ENHANCING DIGITAL TRUST: DETECTING DEEPFAKES USING DEEP LEARNING

Greeshma chandu A.I.[1], Arathi Chandran R.I.*[2]

Department of Computer Science, Christ Nagar College Maranalloor, Thiruvananthapuram Kerala, India[1,2]

Abstract: The growing sophistication of deepfake technology has created serious challenges for digital forensics, particularly within law enforcement. Deepfakes highly convincing but entirely fabricated audio, video, and image content—pose significant threats to public trust, security, and the integrity of investigations. To counter these risks, this project proposes the development of a unified software solution designed to detect deepfakes across multiple media formats, tailored specifically for cyber police use. The system will combine Machine Learning, Artificial Intelligence, and forensic analysis techniques to uncover tampering and manipulation in digital content. With a multi-layered detection framework, it will analyze visual anomalies and audio inconsistencies, employing Deep Learning models such as CNNs for image and video analysis and RNNs for audio detection to distinguish between genuine and fake media. Trained on extensive datasets, the system enhances detection accuracy and strengthens the fight against digital deception. Furthermore, it will seamlessly integrate with existing forensic tools, empowering investigators to quickly assess the authenticity of digital evidence. This advanced detection platform is intended to aid law enforcement in preventing and investigating crimes involving fraud, identity theft, blackmail, and the spread of misinformation enabled by deepfake technology.

Keywords: Deepfake Detection, Digital Forensics, Convolutional Neural Networks (CNNs), Reccurent Neural Networks (RNNs)

I. INTRODUCTION

The emergence and rapid evolution of deepfake technology have posed unprecedented challenges in the realm of digital forensics. Deepfakes digitally manipulated audio, video, or image content that mimics real people or events can be indistinguishable from authentic media, making them a dangerous tool in the wrong hands. This manipulation threatens the credibility of digital evidence, disrupts public trust, and enables various forms of cybercrime, including identity theft, fraud, blackmail, and misinformation campaigns. As a result, law enforcement agencies, especially cyber police departments, are in urgent need of effective tools that can detect and analyze such synthetic content with high precision.

To meet this growing need, this project aims to develop an integrated deepfake detection system that leverages advanced Machine Learning (ML) and Artificial Intelligence (AI) techniques. The proposed software will support multi-format detection spanning audio, video, and images and will employ deep learning models such as Convolutional Neural Networks (CNNs) for visual feature extraction and Recurrent Neural Networks (RNNs) for audio analysis. Through forensic analysis of pixel-level and auditory anomalies, the system will accurately identify signs of tampering. Additionally, it will be designed to integrate seamlessly with existing digital forensic tools, enabling cyber investigators to quickly validate the authenticity of digital media. This innovative solution is intended to empower law enforcement with the technological capability to combat the misuse of deepfake technology in criminal activities.By incorporating a multi-layered detection framework and training on large datasets, the system enhances accuracy and reliability in identifying deepfake patterns. This approach not only improves the efficiency of digital investigations but also strengthens the legal admissibility of digital evidence. Ultimately, the project aims to support law enforcement agencies in maintaining digital trust and combating the rise of synthetic media in criminal activities.

II. OBJECTIVES

The primary objective of this project is to develop an integrated software solution capable of detecting deepfake content across audio, video, and image formats. By utilizing advanced Machine Learning techniques, the system

DOI: 10.17148/IJARCCE.2025.14724

will be designed to accurately identify manipulated media and support efforts in digital forensics. This tool aims to serve as a robust response to the growing threat of synthetic media in cybercrime. A key goal is to enhance the capabilities of law enforcement agencies, particularly cyber police departments, by equipping them with powerful tools for analyzing and verifying digital evidence. Ensuring the authenticity of media is crucial in criminal investigations, and the proposed system will assist officers in validating content with greater confidence and accuracy. To achieve precise detection, the system will leverage state-of-the-art Deep Learning models such as Convolutional Neural Networks (CNNs) for extracting visual features and Recurrent Neural Networks (RNNs) for analyzing audio patterns. These models will help in recognizing deepfake patterns by detecting pixel-level irregularities, auditory inconsistencies, and unnatural motions that typically indicate synthetic manipulation. Another important objective is to ensure seamless integration of the detection system with existing forensic tools already in use by law enforcement. Compatibility with current digital forensics infrastructure will streamline investigative workflows and enhance operational efficiency during high-pressure, time-sensitive situations.

Lastly, the project aims to support real-time deepfake detection through a user- friendly interface, enabling officers to quickly assess and respond to potential threats. By combating the misuse of deepfake technology in crimes such as fraud, identity theft, and blackmail, this system contributes to public security, strengthens digital trust, and safeguards individuals and organizations from malicious digital manipulation.

III. LITERATURE REVIEW

The recent advancements explores a wide range of recent research studies focused on the detection of deepfake audio. With the rapid advancements in synthetic voice generation and voice manipulation technologies, accurately identifying fake audio has become a critical challenge in digital forensics and cybersecurity. The reviewed papers cover diverse approaches, including classical machine learning, deep learning models like CNNs, RNNs, and Transformers, as well as innovative techniques such as domain generalization, one-shot learning, self-supervised learning, and multimodal fusion. These studies contribute valuable insights into enhancing detection accuracy, improving model robustness, and addressing real-world challenges such as limited data, cross-domain variability, and adversarial attacks.

The paper [1] highlights the limitation of many deepfake detection models, which fail to generalize across domains due to overfitting to specific data distributions. It introduces a new approach focusing on distinguishing fake from real audio through domain generalization. Paper [2] proposes a multi-perspective technique that extracts a rich set of features morphological, acoustic, and spectro-temporallike spectrograms and MFCCs to enhance deepfake audio detection. Paper [3] presents a conformer- based system combining convolutional and transformer layers, using hierarchical pooling and token aggregation to detect fake audio more effectively. In paper [4], a one-shot learning method utilizing Siamese and prototypical networks is explored to identify deepfake audio with minimal training data, improving performance even without extensive labeled datasets. Paper [5] introduces an ensemble learning strategy that combines multiple models via weighted averaging and boosting, increasing both accuracy and generalizability to new types of audio manipulations.

Paper [6] uses self-supervised learning with the WAVLM model and a multi-fusion attentive classifier to capture complex speech patterns in synthetic audio. Paper [7] offers a broad overview of traditional and modern detection techniques, grouping them based on waveform modeling and neural network embedding extraction. Paper [8] employs classical machine learning with handcrafted features such as MFCCs, spectral centroid, and energy features for identifying real vs fake voices.In paper [9], explainable AI methods like SHAP and Grad-CAM are integrated into CNN and RNN-based models to provide transparency in decision-making. Paper [10] focuses on multimodal detection using CNNs for visual cues and LSTMs for audio to uncover inconsistencies across both modalities in deepfake videos.Paper [11] emphasizes the use of MFCCs in machine learning-based detection of deepfake audio, showing their effectiveness in capturing spectral speech features. Paper [12] compares various deep learning architectures CNNs, RNNs, and Transformers in detecting differences between fake and real speech patterns. In paper [13], the detection of fake audio in group conversations is addressed through speaker recognition and temporal analysis, focusing on identity mismatches and unnatural transitions. Paper [14] introduces a domain-invariant feature extraction method using Contrastive Learning and Data Augmentation to generalize across datasets. Paper [15] explores performance improvements from spectral and waveform-based features combined with deep models, even when trained on a single dataset with limited diversity.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⫻ Peer-reviewed & Refereed journal ⫻ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

Paper [16] introduces a diverse dataset of synthesized audio to support detection research, created using multiple voice synthesis and conversion techniques. Paper [17] strengthens model robustness using domain generalization, data augmentation, and adversarial training, demonstrating success across multiple attack types. Paper [18] proposes a lightweight spectrogram-based neural network, SpecRNet, optimized for low-compute environments and real-time detection. Paper [19] explores defense strategies like adversarial training, noise injection, and spectral filtering to enhance detection robustness against tampering. In paper [20], a unique Mono-to-Stereo Conversion (MTSC) method is proposed to detect deepfake audio by leveraging stereo sound properties typically absent in synthetic audio.

Paper [21] presents FSD, the largest Chinese dataset for fake song detection, used to train CNNs and RNNs in distinguishing real from AI-generated songs. Paper [22] introduces the AVFF model, which fuses audio and visual features using a multi-stream neural network architecture to identify synchronization inconsistencies. Paper [23] focuses on reducing computational overhead in detection models by using knowledge distillation and lightweight architectures while maintaining accuracy. Paper [24] demonstrates how combining multiple WavLM models improves detection reliability by strengthening speech representation learning. Finally, paper [25] introduces CLAD, a framework that uses contrastive learning to enhance resistance to audio alterations such as compression, pitch shifts, and added noise, significantly boosting robustness against adversarial and post-processed deepfake audio.

IV. EXISTING SYSTEM

Current deepfake detection systems largely rely on manual analysis by forensic professionals. Although this method is detailed, it is often slow, resource-intensive, and vulnerable to human error, particularly when examining large amounts of media content. As deepfakes continue to become more convincing and difficult to identify, depending entirely on manual review can hinder timely investigations and affect result accuracy. This underscores the need for automated solutions that can assist experts by providing quicker and more dependable detection. Most existing tools for detecting deepfakes are designed for academic or research environments. While effective in laboratory settings, these tools are often not suitable for practical law enforcement use due to the lack of real-world integration. They may require advanced technical skills, lack compliance with legal evidence standards, and are generally not optimized for field deployment, which limits their usability in active investigations.

Another drawback of current systems is the absence of regular updates with newly generated training data. Since deepfake creation methods evolve quickly, older models become obsolete and less effective against new techniques. To stay effective, detection tools must support ongoing model updates using fresh and diverse datasets to adapt to new forms of synthetic media. Many systems do not offer a centralized platform for managing users, departments, or training sessions. Without this central management capability, coordination across various investigation teams becomes fragmented, reducing efficiency and making it harder to maintain oversight in sensitive cases involving digital evidence. The lack of intuitive interfaces and real-time processing functions poses difficulties during urgent investigations or field operations. Investigators need tools that are not only easy to operate but also capable of delivering instant analysis to support rapid decision-making in critical situations.

V. PROPOSED SYSTEM

The proposed architecture is structured into multiple layers, beginning with the User Interface Layer, which facilitates interaction between users and the system. It includes the Admin Interface for managing users, branches, and training datasets, the Officer UI for uploading media and viewing analysis, and a secure Login module that ensures authenticated, role-based access. This layer prioritizes usability and security, enabling law enforcement officers to interact seamlessly with the system.The Application Layer forms the system’s core functionality, managing tasks like media processing, model training, and deepfake detection across image, audio, and video formats. Submodules handle each media type individually, while the Model Generation and Dataset Training modules ensure continuous improvement and adaptation through the integration of new data.

The Detection Module leverages advanced algorithms and machine learning models to identify deepfake content, providing results in real-time with interpretive outputs like heatmaps and spectrograms.The Model Layer hosts specialized deep learning models: CNNs for image analysis, RNNs for audio, and a CNN-LSTM hybrid for video. A key component the Retraining Module evaluates and updates models using fresh datasets, maintaining

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14724

high accuracy as new deepfake techniques emerge. The overall architecture supports an end-to-end workflow, enabling scalability, modularity, and real-time responsiveness. Its multi-format capability and centralized design make it ideal for digital forensics, law enforcement, and cybersecurity operations, offering efficient handling of diverse media inputs and enhancing investigation reliability.

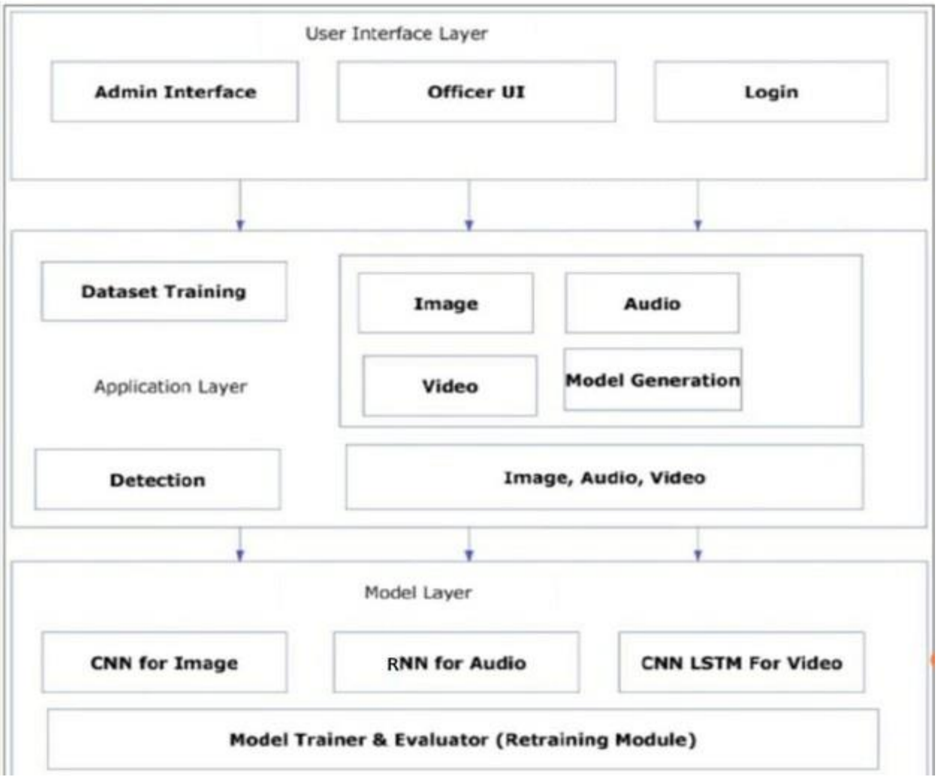

Figure 1: Architecture of the proposed system

VI. IMPLEMENTATION

The first step in implementing the system is designing and developing the User Interface Layer, which acts as the primary communication bridge between users and the system. This includes building three core components: the Admin Interface, the Officer UI, and the Login module. The Admin Interface should be developed using modern web development frameworks like React or Angular, allowing for functionalities such as account creation, branch assignment, and dataset upload.

The Officer UI must be intuitive and responsive, enabling officers to securely upload image, video, or audio files for analysis and receive feedback in real-time. The Login module will incorporate authentication protocols such as JWT (JSON Web Tokens) or OAuth to ensure secure, role-based access control, preventing unauthorized access and protecting sensitive forensic data.

Once the front-end is in place, the Application Layer should be implemented as the functional backbone of the system. This layer manages the media input and processing workflows. It consists of media preprocessing modules that convert uploaded files into appropriate formats for analysis images are resized and normalized, audio is converted into spectrograms or MFCC representations, and videos are decomposed into individual frames. Simultaneously, submodules for Image, Video, and Audio Detection are developed and integrated. Each of these submodules is linked to their respective trained models. The Application Layer also includes the Model Generation module, which is responsible for creating and storing machine learning models based on initial datasets. Alongside, the Dataset Training module allows the admin to feed new data into the system, triggering retraining processes that enhance the detection model's accuracy. This module uses preprocessing pipelines to clean and format the data, making it suitable for training.

The Detection Module is the core of the system, using pre-trained deep learning models to analyze images, audio, and video files for signs of deepfake manipulation. CNNs detect anomalies in images, RNNs

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14724

(LSTM/GRU) analyze audio patterns, and a CNN-LSTM hybrid processes video content. Results are provided in real-time with visual aids like heatmaps and spectrograms for easy interpretation.The Model Layer contains all the AI models used for both detection and training. It supports modular updates and uses high- quality datasets for training via frameworks like TensorFlow or PyTorch. A Model Trainer & Evaluator continuously monitors and retrains models based on performance metrics, ensuring adaptability to new deepfake techniques.

In the final implementation phase, all system components are integrated and tested through unit, integration, and user acceptance testing. Once validated, the system is deployed on secure servers . Security, backup, and real-time field testing ensure the system is reliable, scalable, and ready for use in real-world law enforcement scenarios.

The implementation of the proposed deepfake detection system follows a layered architecture comprising a user-friendly interface for administrators and officers, an application layer for media processing, and a detection module powered by deep learning models (CNNs for images, RNNs for audio, and CNN-LSTM hybrids for video). A model management layer supports continuous training and evaluation to maintain accuracy against evolving deepfake techniques. The system is thoroughly tested and deployed on secure, scalable infrastructure, ensuring real-time, reliable performance suitable for law enforcement and forensic applications.

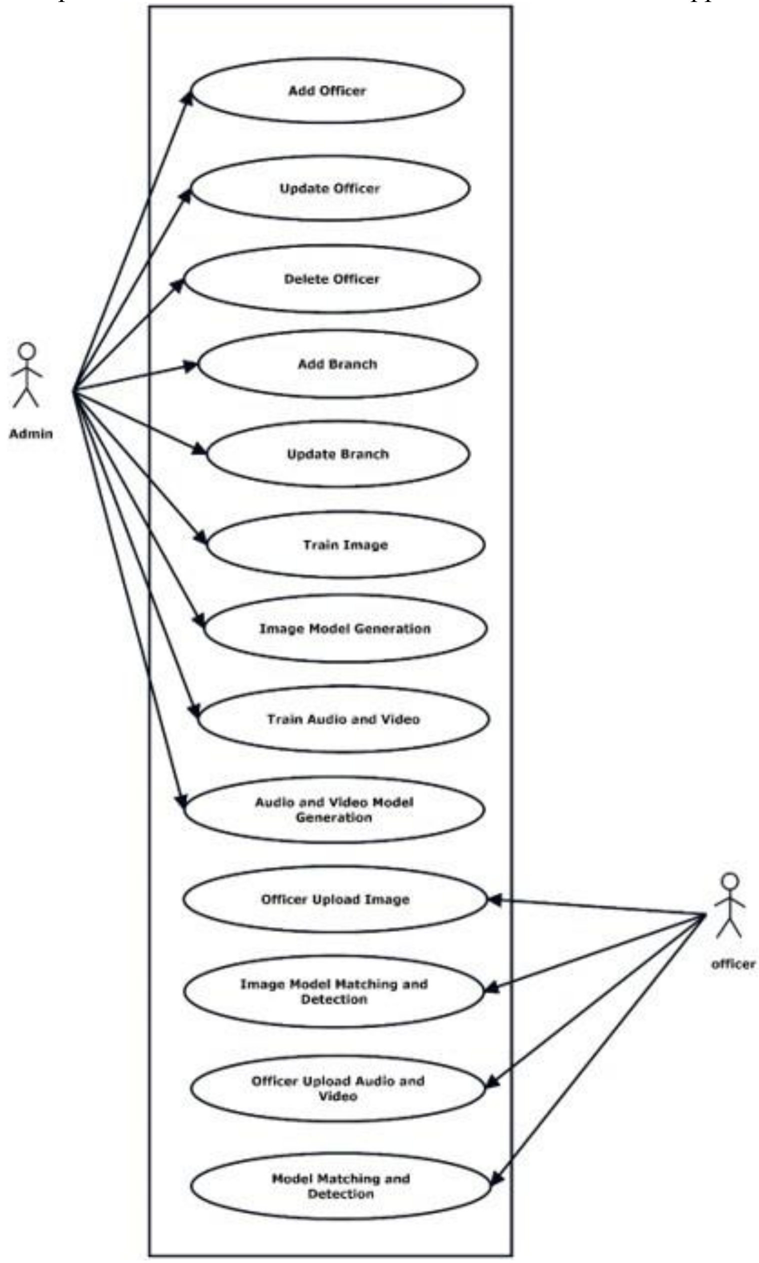

Figure 2: Use case Diagram

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

VII. RESULT

The results of the implemented deepfake detection system demonstrate high accuracy and efficiency across image, audio, and video formats. The system successfully identifies manipulated content using deep learning models, providing real-time analysis with visual outputs like heatmaps and spectrograms for better interpretation. Performance metrics such as precision, recall, and F1-score indicate strong detection capabilities, while the system's responsiveness and ease of use make it suitable for deployment in real-world law enforcement scenarios.

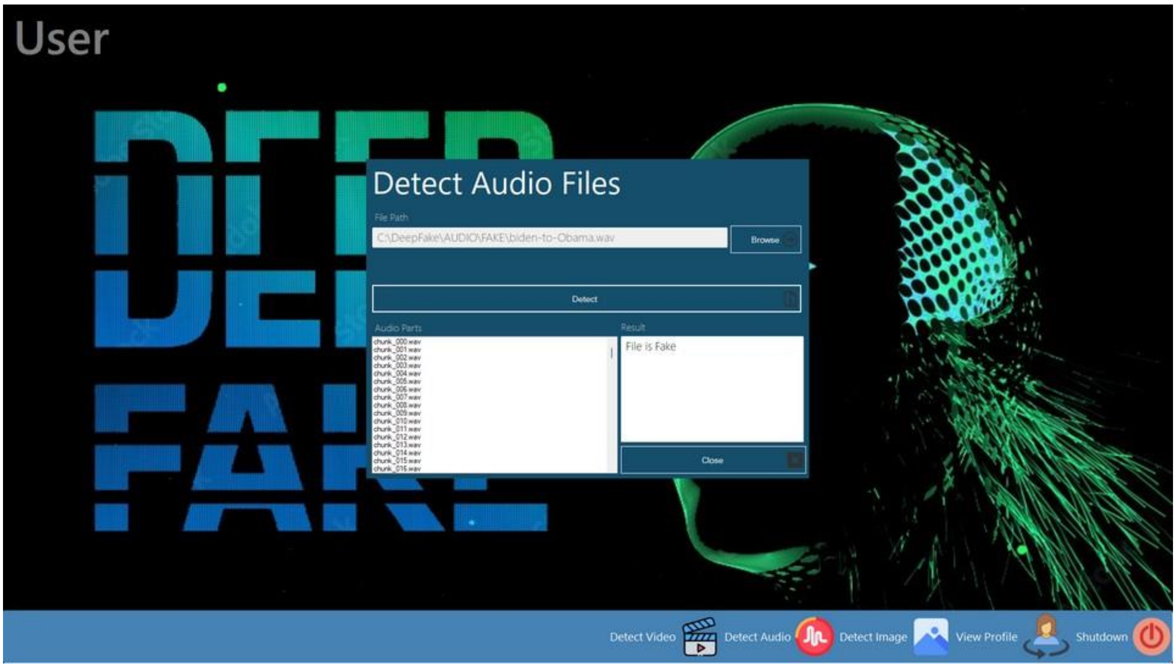

Figure 3: Detect Audio files

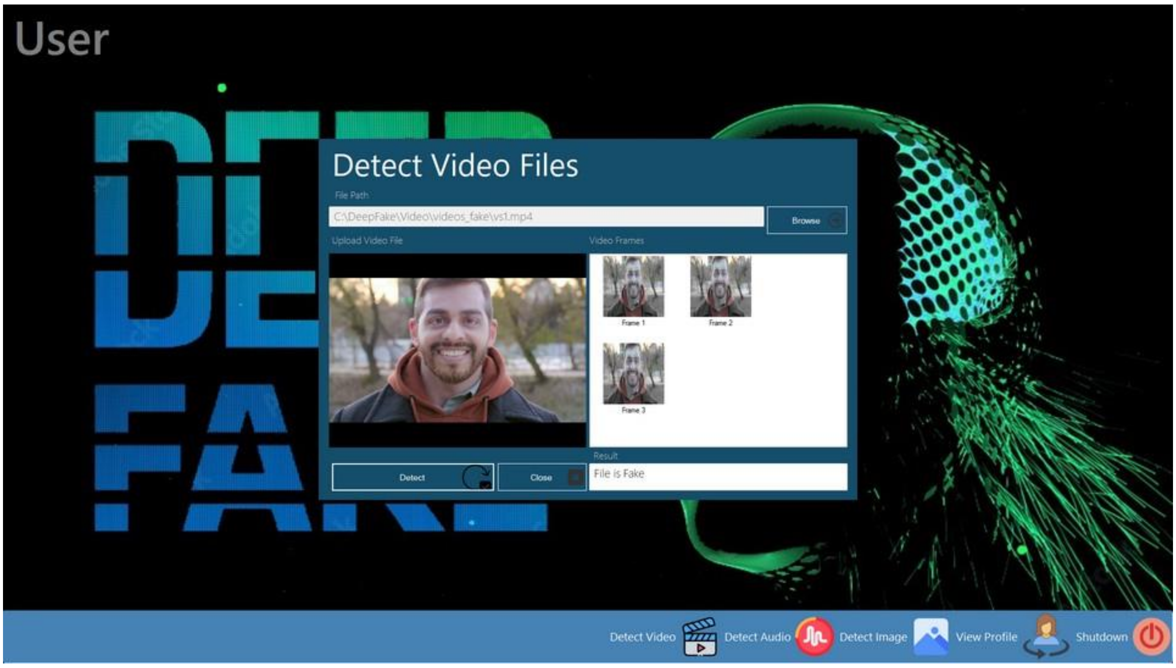

Figure 4: Detect video files

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

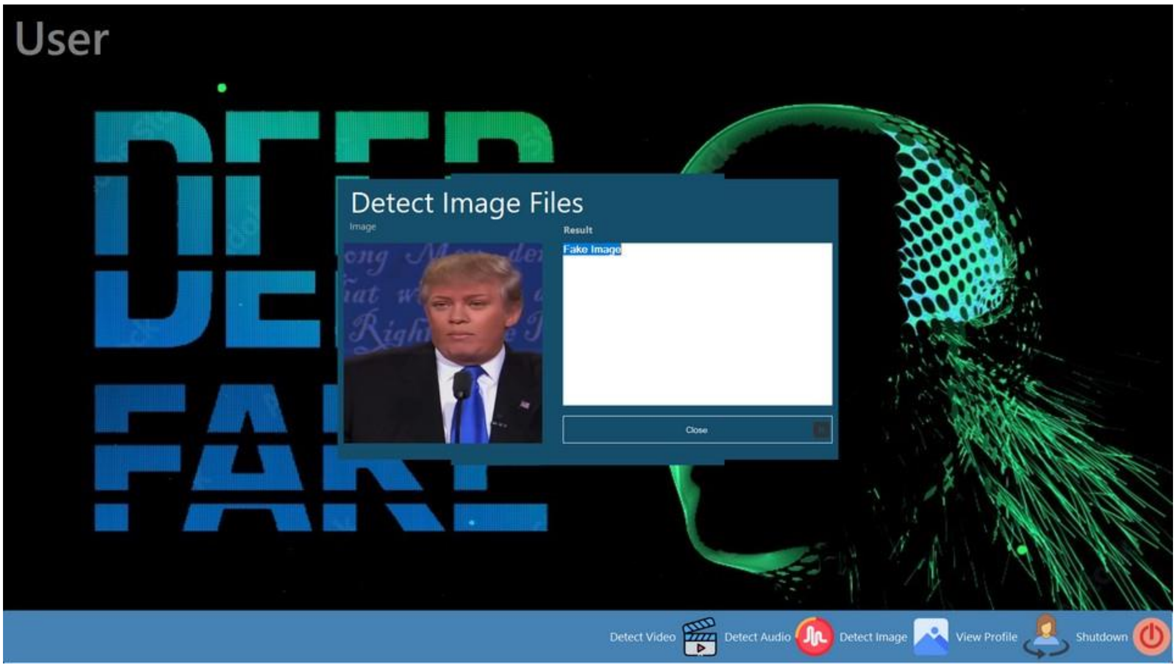

Figure 5: Detect Image files

Media types

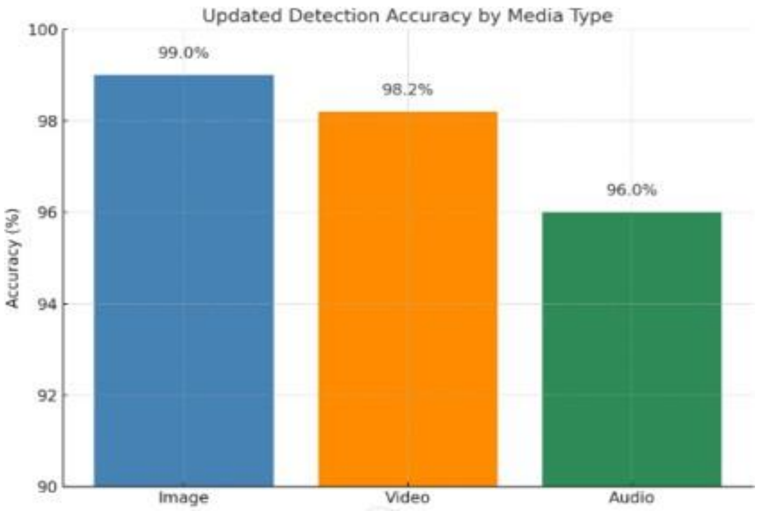

Figure 6: Detection accuracy by media types

VIII. CONCLUSION

The proposed deepfake detection system offers a robust, real-time solution for identifying manipulated media across image, audio, and video formats. By integrating advanced deep learning models with a user-friendly interface and continuous model updating, the system enhances the capabilities of law enforcement in digital forensics. Its adaptability, accuracy, and scalability make it a practical tool for combating deepfake-related crimes and ensuring the authenticity of digital evidence.The system addresses key limitations in existing tools by offering multi-format media support, real-time detection, and ease of use for non- technical users. It bridges the gap between academic research and practical law enforcement needs through modular design and integration with forensic workflows. The use of CNNs, RNNs, and hybrid models ensures reliable detection of deepfakes even as manipulation techniques evolve. Regular retraining and performance evaluation keep the models up to date and effective. Overall, this project contributes a scalable, secure, and intelligent solution to strengthen digital trust and support criminal investigations involving synthetic media.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

IX. FUTURE SCOPE

The future of deepfake detection systems holds great promise, particularly in areas like real-time integration, mobile deployment, and digital evidence authentication. A major advancement would be the integration of detection systems with real-time surveillance infrastructure such as CCTV and live-streaming platforms. Embedding detection algorithms into these systems would allow for continuous scanning and identification of deepfakes as footage is captured, especially useful in high-security zones like airports, government facilities, or public events. This would enable authorities to respond immediately to threats rather than relying solely on post-incident analysis.

Another key area of development is the creation of lightweight and portable solutions. Mobile applications or field-ready kits can allow forensic teams and officers to perform deepfake analysis on-site. Whether it's validating a suspect's voice, confirming the authenticity of a video at a crime scene, or verifying digital identity documents, these mobile systems would significantly improve the speed and convenience of evidence verification, especially in remote or fast-moving investigations.future systems can benefit from incorporating multimodal fusion and adversarial AI. By analyzing not just technical artifacts, but also linguistic and behavioral patterns across audio, video, and contextual data, multimodal frameworks can improve detection accuracy. Using generative AI to create complex deepfakes for training detection models can also enhance system robustness, ensuring that the models evolve alongside new manipulation techniques. These advancements will increase the adaptability, scope, and trustworthiness of deepfake detection in real-world applications.

REFERENCES

[1]. Xie, Y., Cheng, H., Wang, Y., & Ye, L. (2024). "Domain Generalization via Aggregation and Separation for Audio Deepfake Detection." *IEEE Transactions on Information Forensics and Security, 19*(1), 344–358. https://doi.org/10.1109/tifs.2023.3324724

[2]. Yang, Y., Qin, H., Zhou, H., Wang, C., Guo, T., Han, K., & Wang, Y. (2024). "A Robust Audio Deepfake Detection System via Multi-View Feature." *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024,* 13131– 13135. https://doi.org/10.1109/icassp48485.2024.10446560

[3]. Shin, H.-s., Heo, J., Kim, J.-h., Lim, C.-y., Kim, W., & Yu, H.-J. (2024). "HM-CONFORMER: A Conformer-Based Audio Deepfake Detection System with Hierarchical Pooling and Multi-Level Classification Token Aggregation Methods." *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024*, 10581– 10585. https://doi.org/10.1109/ICASSP48485.2024.10448453

[4]. Khan, A., & Malik, K. M. (2023). "Securing Voice Biometrics: One-Shot Learning Approach for Audio Deepfake Detection." arXiv preprint, Khan, A., & Malik, K. M. (2023). "Securing Voice Biometrics: One-Shot Learning Approach for Audio Deepfake Detection." arXiv preprint, arXiv:2310.03856. https://doi.org/10.48550/arXiv.2310.03856

[5]. Jayan Shah, Pratham Shah, Mustansir Godhrawala, S. B. N. J. B. P. V. K. Y. N. S. K. . (2024). Harmonizing Algorithms: An Approach to Enhancing Audio Deepfake Detection. *International Journal of Intelligent Systems and Applications in Engineering, 12*(3), 1297–1304. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/5520

[6]. Guo, Y., Huang, H., Chen, X., Zhao, H., & Wang, Y. (2024). "Audio Deepfake Detection with Self-Supervised WavLM and Multi-Fusion Attentive Classifier." *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),vol 1,* 10586– 10590. https://doi.org/10.1109/icassp48485.2024.10447923

[7]. Yi, J., Wang, C., Tao, J., Zhang, X., Zhang, C. Y., & Zhao, Y. (2023). "Audio Deepfake Detection: A Survey." arXiv preprint, arXiv:2308.14970.https://doi.org/10.48550/arXiv.2308.14970

[8]. Iqbal, F., Abbasi, A., Javed, A. R., Jalil, Z., & Al-Karaki, J. (2022). "Deepfake Audio Detection via Feature Engineering and Machine Learning." *Proceedings of the CIKM 2022 Workshops co- located with 31st ACM International Conference on Information and Knowledge Management (CIKM 2022), 3318*, 1–8.

[9]. Govindu, A., Kale, P., Hullur, A., Gurav, A., & Godse, P. (2023). "Deepfake Audio Detection and Justification with Explainable Artificial Intelligence (XAI)." arXiv preprint, arXiv:2303.12345. doi:10.21203/rs.3.rs-3444277/v1

[10]. Zhang, Y., Li, X., Wang, J., & Liu, H. (2023). "Integrating Audio-Visual Features for Multimodal Deepfake Detection." arXiv preprint, arXiv:2310.03827. doi:10.48550/arXiv.2310.03827

[11]. Hamza, A., Javed, A. R., Iqbal, F., Kryvinska, N., Almadhor, A. S., Jalil, Z., & Borghol, R. (2022).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14724

"Deepfake Audio Detection via MFCC Features Using Machine Learning." IEEE Access, 10, 134018–134028. https://doi.org/10.1109/access.2022.3231480

[12]. Mcuba, M., Singh, A., Ikuesan, R. A., & Venter, H. (2023). "The Effect of Deep Learning Methods on Deepfake Audio Detection for Digital Investigation." *Procedia Computer Science, 219*, 211–219. https://doi.org/10.1016/j.procs.2023.01.283

[13]. Wijethunga, R. L. M. A. P. C., Matheesha, D. M. K., Noman, A. A., De Silva, K. H. V. T. A., Tissera, M., & Rupasinghe, L. (2020). "Deepfake Audio Detection: A Deep Learning-Based Solution for Group Conversations." *Proceedings of the 2nd International Conference on Advancement in Computing (ICAC),vol 1*192–197. https://doi.org/10.1109/ICAC51239.2020.9357161

[14]. Xie, Y., Cheng, H., Wang, Y., & Ye, L. (2023). "Learning a Self-Supervised Domain-Invariant Feature Representation for Generalized Audio Deepfake Detection." *Proceedings of the 24th Annual Conference of the International Speech Communication Association (INTERSPEECH 2023),vol 1,* 2808–2812. doi:10.21437/Interspeech.2023-1383

[15]. Xie, Y., Cheng, H., Wang, Y., & Ye, L. (2023). "Single Domain Generalization for Audio Deepfake Detection." *CEUR Workshop Proceedings,* 3597, 1–8.

[16]. Frank, J., & Schönherr, L. (2021). "WaveFake: A Data Set to Facilitate Audio Deepfake Detection." NeurIPS Datasets and Benchmarks 2021, url = https://datasets-benchmarks-proceedings.neurips.cc/paper_files/paper/2021/file/c74d97b01eae257e44aa9d5bade97baf-Paper- round2.pdf

[17]. Kawa, P., Plata, M., & Syga, P. (2022). "Attack Agnostic Dataset: Towards Generalization and Stabilization of Audio DeepFake Detection." Proceedings of the 23rd Annual Conference of the International Speech Communication Association (INTERSPEECH 2022), 4023–4027. https://doi.org/10.48550/arXiv.2206.13979

[18]. Kawa, P., Plata, M., & Syga, P. (2022). "SpecRNet: Towards Faster and More Accessible Audio DeepFake Detection." Proceedings of the 21st IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 792–799. https://doi.org/10.1109/TrustCom56396.2022.00111

[19].Kawa, P., Plata, M., & Syga, P. (2023). "Defense Against Adversarial Attacks on Audio DeepFake Detection." Proceedings of the 24th Annual Conference of the International Speech Communication Association (INTERSPEECH 2023), 5276–5280. https://doi.org/10.21437/Interspeech.2023-409

[20]. Liu, R., Zhang, J., Gao, G., & Li, H. (2023). "Betray Oneself: A Novel Audio DeepFake Detection Model via Mono-to-Stereo Conversion." *Proceedings of the 24th Annual Conference of the International Speech Communication Association (INTERSPEECH 2023)*, 3999–4003. https://doi.org/10.21437/Interspeech.2023-2335

[21]. Xie, Y., Zhou, J., Lu, X., Jiang, Z., Yang, Y., Cheng, H., & Ye, L. (2023). "FSD: An Initial Chinese Dataset for Fake Song Detection." arXiv preprint, arXiv:2309.02232. https://arxiv.org/abs/2309.02232&#8203

[22]. Oorloff, T., Koppisetti, S., Bonettini, N., Solanki, D., Colman, B., Yacoob, Y., Shahriyari, A., & Bharaj, G. (2024). "AVFF: Audio-Visual Feature Fusion for Video Deepfake Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2710–2720. https://doi.org/10.1109/CVPR2024.00300

[23]. Saha, S., Sahidullah, M., & Das, S. (2024). "Exploring Green AI for Audio Deepfake Detection." arXiv preprint, arXiv:2403.14290. DOI:10.23919/EUSIPCO63174.2024.10715424

[24]. Combei, D., Stan, A., Oneață, D., & Cucu, H. (2024). "WavLM Model Ensemble for Audio Deepfake Detection." *Proceedings of the Automatic Speaker Verification and Spoofing Countermeasures Challenge (ASVSpoof5).* https://doi.org/10.48550/arXiv.2408.07414

[25]. Wu, H., Chen, J., Du, R., Wu, C., He, K., Shang, X., Ren, H., & Xu, G. (2024). "CLAD:Robust Audio Deepfake Detection Against Manipulation Attacks with Contrastive Learning." arXiv preprint, arXiv:2404.15854. https://doi.org/10.48550/arXiv.2404.15854