ARC Journal of Forensic Science
Volume 9, Issue 2, 2025, PP 36-39
ISSN No. (Online) 2456-0049
DOI: https://doi.org/10.20431/2456-0049.0902005
www.arcjournals.org

Deepfake Detection and Multimedia Forensics: Investigating Synthetic Media, Image Forgery, and Video Manipulation in Cybercrime Cases

Nishchal Soni

Department of Forensic Science, Lovely Professional University, Phagwara, Punjab, 144411

****Corresponding Author:*** ***Nishchal Soni,*** *Department of Forensic science, Lovely Professional University, Phagwara, Punjab, India. Email: Nishchalresearch@gmail.com*

Abstract:

The growing sophistication of synthetic media technologies, more so deepfakes, has created new challenges for cybercrime investigations and digital forensics. Previously, a tool for creative uses and entertainment, deepfakes are now extensively misused for harmful ends such as identity theft, fraud, disinformation campaigns, political manipulation, and harassment. This review collates existing literature on deepfake detection and multimedia forensics, with focus on image forgery, video manipulation, and multimodal analysis methods. Critical forensic technologies and uses from legacy metadata analysis and pixel-by-pixel analysis to AI-based solutions like Microsoft Video Authenticator and DARPA's MediFor initiative are investigated for their potential in verifying digital evidence. In spite of these advancements, investigators continue to grapple with ongoing challenges such as quick improvements in generative AI models, anti-forensic techniques, and issues of admissibility of AI-generated evidence in courts. Future directions identify the unification of explainable AI, blockchain-supported provenance systems, IoT-supported real-time detection, and global collaboration as key strategies for enhancing forensic resilience. Through bridging technological innovation and legal and ethical considerations, multimedia forensics will continue to be the cornerstone to protecting digital trust and preventing the abuse of synthetic media in cybercrime.

Keywords: *Deepfake detection; multimedia forensics; synthetic media; cybercrime investigation; image forgery; video manipulation; explainable AI; blockchain provenance; digital evidence*

1. Introduction

The rapid evolution of artificial intelligence, especially in the area of generative adversarial networks (GANs) and diffusion models, has accelerated the spread of deepfakes and multimedia forgeries. These artificial media objects from doctored photographs and falsified audio to hyper-realistic video simulation have become an increasingly significant threat to digital trust, information integrity, and security. Initially explored for entertainment and creative applications, deepfakes are now increasingly exploited in cybercrime, including identity theft, financial fraud, misinformation campaigns, political manipulation, and online harassment (Singh & Dhumane, 2025).

In a forensic context, detection and attribution of such manipulated media have become an immediate priority. Multimedia forensics is the scientific examination of digital artifacts to identify evidence of tampering, confirm authenticity of content, and determine culprits. Synthetic media makes forensic investigation difficult compared to conventional digital evidence because of its realism and rapid advancement of generative models. Scientists have shown that multimedia forensics is capable of detecting pixel distribution anomalies, compression artifacts, and spatio-temporal inconsistencies that indicate the existence of synthetic manipulation (Nastasi, 2021; Ferreira et al., 2021).

Deepfake detection goes beyond visual inspection to encompass audio-visual cross-validation and machine learning-based techniques, which utilize convolutional neural networks (CNNs), transformers, and temporal modeling to detect minute anomalies imperceptible to the naked eye (Qureshi et al., 2024). Further, the merging of IoT-based surveillance evidence and digital forensic systems has been suggested to improve the validity of deepfake investigations in social media and cybercrime environments (Khan et al., 2025).

While cybercriminals turn to more advanced anti-forensic techniques to cover their traces, multimedia forensics also needs to keep up. This review thus reviews existing methods, tools, and issues involved in deepfake detection and explains new approaches like explainable AI and blockchain-based authentication to enhance digital investigations.

2. Deepfake and Forgery Detection Techniques

Identification of forged multimedia content is still the core problem in digital forensics because GANs and diffusion models generate ever more real-looking fakes. Detection methods fall broadly into image forgery, video editing, multimodal verification, and anti-forensic detection.

Image forgery detection has conventionally depended on pixel-level and statistical processing, such as Error Level Analysis, copy-move detection, and compression anomalies. The more sophisticated techniques employ blind forgery detection and noise pattern analysis, whereas deep learning models like CNNs and vision transformers now dominate, providing scalable solutions to face swapping and morphing (Shukla et al., 2024; Verdoliva, 2020).

Video forgery detection is more involved because of temporal dynamics. Methods detect anomalies in expression, eye blinking, illumination, and head orientation, whereas temporal CNNs and recurrent models detect frame-level anomalies. Benchmarking datasets such as FaceForensics++ and DeepFakeDetection are still essential for training detectors (Tyagi & Yadav, 2023; El-Shafai et al., 2024).

Multimodal analysis becomes more reliable through cross-validation of auditory and visual signals, e.g., lip reading against the speech or spectrograms of voices. Blending with biometric systems makes forensic authentication even more robust (Qureshi et al., 2024; Amerini et al., 2025).

Lastly, researchers have to deal with anti-forensic strategies such as post-processing, adversarial perturbations, and noise injections that aim to hide forensic evidence. Detection thus needs ongoing updating of the dataset and models that can adjust to new manipulations (Nastasi, 2021).

3. Forensic Tools and Applications

The emergence of deepfaked media has accelerated the creation of advanced forensic instruments that merge conventional techniques with AI-based methodologies to verify and analyze deepfakes.

Video and image verification software like Amped Authenticate and Forensically identify inconsistencies in compression artifacts, pixel distributions, and exif metadata (Soni, 2025), while sophisticated systems like Deepware Scanner and Microsoft Video Authenticator use deep learning to detect faint indications of tampering, such as texture and lighting discrepancies (Nastasi, 2021; Verdoliva, 2020).

Forensic frameworks powered by AI continue to improve the accuracy of detection. Software such as DARPA's MediFor seeks to establish real-time authentication pipelines that are automated, whereas the Digital Forensics Socio-Cyber World (DF-SCW) combines video, audio, and biometric information for multimodal investigations (Amerini et al., 2025).

Multimodal methods compare visual and auditory cues, like lip-sync discrepancy and voice spectrogram examination, with a guarantee of resisting single-modality attacks and working well in political misinformation, fraud, and impersonation evidence (Qureshi et al., 2024).

Practically, the tools are being used more and more in cybercrime investigations ranging from sextortion and financial scams to disinformation campaigns. Their combination with IoT surveillance and cloud provides mass-scale, remote tracking of altered content (Khan et al., 2025). These emerging solutions constitute the technological foundation of multimedia forensics needed to respond to the increasing sophistication of synthetic media.

Challenges In Deepfake and Multimedia Forensics

Even with developments, deepfake and multimedia forensics are hampered by significant challenges that restrict their utility in actual cybercrime cases. These are technical and socio-legal challenges. Technically, the fast pace of developments in GANs and diffusion models creates very realistic forgeries that lower the validity of conventional cues such as pixel anomalies or compression marks. Detection systems are also dataset-dependent, working well on benchmarked sets but not in uncontrolled real-world scenarios a "generalization gap" that detracts from deployment on social media platforms (Amerini et al., 2025; El-Shafai et al., 2024).

Cybercriminals additionally take advantage of anti-forensic measures of re-compression, blurring, noise injection, and adversarial perturbations, all of which are compromising to detector effectiveness. Open-source deepfake software enables such abilities to be used by everyone, which brings down the bar for attackers (Lyu, 2022; Singh & Dhumane, 2025).

Legally, issues emanate from AI-based evidence admissibility. Most detection systems are "black boxes" that restrict transparency and cause concern in court. Meanwhile, investigators need to balance forensic access and privacy rights, particularly where cross-border data sharing is subject to jurisdictional conflicts (Khan et al., 2025; Nastasi, 2021).

4. Future Directions

The future of deepfake detection and multimedia forensics is in the development of robust and explainable systems that are able to cope with fast-evolving generative technologies.

Integration of explainable AI (XAI) in forensic pipelines will play a vital role to make the process transparent and admissible as evidence in court (Devi, 2025). Researchers also highlight cross-model evaluation frameworks, which are generalizing well to previously unseen manipulation strategies (Khan et al., 2025).

Improvements to multimodal fusion combining audio, visual, and biometric signals are likely to enhance resistance against adversarial attacks and disinformation (Qureshi et al., 2024). Concurrently, blockchain-supported provenance systems can authenticate digital evidence through tamper-evident verification, and IoT-supported infrastructures might facilitate real-time observation and early warning (Amerini et al., 2025; Khan et al., 2025).

Lastly, countering the international dissemination of synthetic media calls for global coordination and harmonized forensic practices to assist in cross-border investigations (Masood et al., 2023). Together, these solutions underscore the need for infusing technological innovation with legal and policy responses to enhance society's resistance to synthetic media threats.

5. Conclusion

The emergence of deepfakes and multimedia forgeries poses severe challenges for cybercrime investigations and digital forensics. These manipulations are becoming more prevalent in identity theft, fraud, disinformation, and harassment, which compels distrust in digital evidence. This review identified the existing detection methods, forensic tools, and their implementation and discussed the technical, legal, and societal challenges investigators are confronted with. While anti-forensic techniques and generative models keep developing, innovative solutions like explainable AI, blockchain provenance, multimodal analysis, and global cooperation provide promising directions to pursue. Prolonging multimedia forensics will not only demand technological advancements but also harmonized policy and legal frameworks for maintaining the authenticity and reliability of digital evidence in the synthetic media era.

References

[1] Amerini, I., Barni, M., Battiato, S., Bestagini, P., & Boato, G. (2025). Deepfake media forensics: Status and future challenges. *Journal of Imaging, 11*(3), 73. https://www.mdpi.com/2313-433X/11/3/73

[2] Devi, D. S. (2025). Deepfake detection in the era of multimedia: Methods, gaps, and evolving research directions. *International Journal of Innovative Science and Research Technology, 10*(7). https://eprint.innovativepublication.org/id/eprint/2089/1/IJISRT25JUL1768.pdf

[3] El-Shafai, W., Fouda, M. A., & El-Rabaie, E. S. M. (2024). A comprehensive taxonomy on multimedia video forgery detection techniques: Challenges and novel trends. *Multimedia Tools and Applications, 83*(15), 43651–43690. https://link.springer.com/article/10.1007/s11042-023-15609-1

[4] Khan, A. A., Laghari, A. A., & Bacarra, R. (2025). *Cybersecurity, digital forensics, and the IoT for deepfake investigation on social media platforms: A review*. ResearchGate. https://www.researchgate.net/profile/Dr-Abdullah-Ayub-KhanPhd/publication/389693661_Cybersecurity-Digital-Forensics-and-the-IoT-for-Deepfake-Investigation-on-Social-Media-Platforms-A-Review/links/6850ce1626f43051a580f526/Cybersecurity-Digital-Forensics-and-the-IoT-for-Deepfake-Investigation-on-Social-Media-Platforms-A-Review.pdf

[5] Lyu, S. (2022). Deepfake detection. In R. Subramanian, J. Joshi, & H. Yu (Eds.), *Deep learning-based approaches for social computing* (pp. 319–336). Springer. https://library.oapen.org/bitstream/handle/20.500.12657/54043/978-981-16-7621-5.pdf?sequence=1#page=320

[6] Masood, M., Nawaz, M., Malik, K. M., Javed, A., & Irtaza, A. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence, 53*(6), 6575–6592. https://link.

springer.com/article/10.1007/s10489-022-03766-z

[7] Nastasi, C. (2021). *Multimedia forensics: From image manipulation to the deep fake. New threats in the social media era* [Doctoral dissertation, University of Naples]. University Repository. https://tesidottorato.depositolegale.it/handle/20.500.14242/124097

[8] Qureshi, S. M., Saeed, A., Almotiri, S. H., & Ahmad, F. (2024). Deepfake forensics: A survey of digital forensic methods for multimodal deepfake identification on social media. *PeerJ Computer Science, 10*, e2037. https://peerj.com/articles/cs-2037/

[9] Shukla, D. K., Bansal, A., & Singh, P. (2024). A survey on digital image forensic methods based on blind forgery detection. *Multimedia Tools and Applications, 83*(14), 41785–41814. https://link.springer.com/article/10.1007/s11042-023-18090-y

[10] Singh, S., & Dhumane, A. (2025). Unmasking digital deceptions: An integrative review of deepfake detection, multimedia forensics, and cybersecurity challenges. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.Cfm?abstract_id=5275910

[11] Soni, N. (2025). Forensic Value of Exif Data: An Analytical Evaluation of Metadata Integrity across Image Transfer Methods. Deleted Journal, 2(2), 10006. https://doi.org/10.70322/plfs.2025.10006

[12] Tyagi, S., & Yadav, D. (2023). A detailed analysis of image and video forgery detection techniques. *The Visual Computer, 39*(8), 2893–2910. https://link.springer.com/article/10.1007/s00371-021-02347-4

[13] Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing, 14*(5), 910–932. https://ieeexplore.ieee.org/abstract/document/9115874/

Citation: *Nishchal Soni, Deepfake Detection and Multimedia Forensics: Investigating Synthetic Media, Image Forgery, and Video Manipulation in Cybercrime Cases. ARC Journal of Forensic Science. 2025; 9(2):36-39. DOI: https://doi.org/10.20431/2456-0049.0902005.*

Copyright: *© 2025 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*