Discover Computing

Review

A survey on multimedia-enabled deepfake detection: state-of-the-art tools and techniques, emerging trends, current challenges & limitations, and future directions

Abdullah Ayub Khan[1] · Asif Ali Laghari[2] · Syed Azeem Inam[3] · Sajid Ullah[4] · Muhammad Shahzad[1] · Darakhshan Syed[1]

Received: 17 January 2025 / Accepted: 10 April 2025

Published online: 25 April 2025

Abstract

Rapid technological breakthroughs in recent years, like Deepfake, have made it feasible to produce synthetic media that is remarkably lifelike, but they also present significant hazards to public trust, privacy, and security. This survey paper reviews the latest techniques for detecting deepfakes, focussing on important components as image and video manipulation, audio spoofing, and multimodal synthesis. It features state-of-the-art methods including machine learning (ML), deep learning (DL), and multimodal architectures that are especially made to address the previously described deepfake criteria. The report provides a critical review of assessment measures used to assess detection model performance, including precision, accuracy, recall, computing effectiveness and efficiency, and fast responses to adversarial attacks. In order to assist direct future research, this highlights recent advancements in the subject, including explainable AI, federated learning, and self-supervised learning hierarchy. In order to examine the problems with adversarial attacks, scalability across different datasets, and the ethical implications of detection techniques, it is also vital to look into the technological and societal challenges surrounding multimedia-enabled deepfake detection. In particular, the usage of Blockchain Distributed Ledger Technology (BDLT) for traceability, lightweight modelling, and resilient systems forms for cross-model deepfake evaluation are discussed in this review study along with potential solutions to these limitations and areas for further research. This paper offers a comprehensive resource for future research, experts, and practitioners looking to combat the growing threat of deepfake, especially in the social media space, using innovative and useful detection tools.

Keywords Deepfakes · Multimedia Systems · Machine Learning (ML) · Deep Learning (DL) · Reinforcement Learning (RL) · Blockchain Distributed Ledger Technology (BDLT)

✉ Abdullah Ayub Khan, abdullah.khan00763@gmail.com; abdullahayub.bukc@bahria.edu.pk; ✉ Sajid Ullah, sajidjalwan@gmail.com; Asif Ali Laghari, asifalilaghari@gmail.com; Muhammad Shahzad, muhammadshahzad.bukc@bahria.edu.pk; Darakhshan Syed, darakshansyed.bukc@bahria.edu.pk | [1]Department of Computer Science, Bahria University Karachi Campus, Karachi 75260, Pakistan. [2]Software Collage, Shenyang Normal University, Shenyang, China. [3]Department of Artificial Intelligence and Mathematical Sciences, Sindh Madressatul Islam University, Karachi 74000, Pakistan. [4]Department of Water Resources and Environmental Engineering, Nangarhar University, Jalalabad 2600, Nangarhar, Afghanistan.

Discover Computing (2025) 28:48 | https://doi.org/10.1007/s10791-025-09550-0

1 Introduction

Recent developments in artificial intelligence (AI) and other information and communication technologies (ICT) have improved the current computing infrastructure's dependability [1, 2]. Deepfake technology has become one of the most innovative AI derivatives due to advancements in the Deep Learning (DL) hierarchy, namely in the field of Generative Adversarial Networks (GANs). It has changed the development of synthetic media, especially in social media environments. Videos, audio, and images may be easily altered, manipulated, and tempered thanks to this technology, producing realistic forgeries that are frequently indistinguishable from actual content [2–4]. As illustrated in Fig. 1, deepfakes offer fascinating prospects in the majority of fields, including accessibility, entertainment, and education. Two good examples are linguistic dabbing and virtual avatars, but their misuse has raised severe concerns. These include the spread of misleading information, fraud, identity theft, cyberbullying, and political propaganda, all of which jeopardise public trust, personal privacy, and global security [4, 5].

There are limits in detecting and addressing these growing risks since detection system development has not kept pace with the notable developments in deepfake generating technologies [6, 7]. In multimedia-enabled deepfake materials, sophisticated techniques such as face synthesis, multimodal forgeries, high-resolution lip-syncing, and voice cloning that combines visual and aural deception are employed. As such, rapid and efficient detection techniques are currently very important [7–9]. However, this survey study provides a comprehensive examination of state-of-the-art deepfake detection techniques, focussing on important components such [8–10], as illustrated in Fig. 1: (1) audio spoofing, (2) multimodal synthesis, and (3) picture and video modification and alteration.

- **Spoofing in Audio:** Assessing and identifying speech synthesis and artificial voices in cloned audio is known as "spoofing" in audio.
- **Multimodal Synthesis:** Assessing and analyzing integrated audio–video edits and modifications to guarantee the validity of the information.
- **Video and Image Alteration and Manipulation:** Assessing and spotting irregularities in texture, motion patterns, lighting, and facial features.

The evaluation measures that are utilised to test detection systems, including precision, accuracy, adversarial robustness, recalls, and computational efficiency and effectiveness, are also examined in this survey study. We showcase recent advancements that employ lightweight models for real-time applications and Blockchain Distributed Ledger Technology (BDLT) for data traceability, along with new topics including self-supervised learning, federated learning, and explainable AI. The application of detection technologies in a range of datasets, the results generalizability, and the resulting ethical and legal concerns are some of the significant disadvantages that are also covered in this work. It seeks to serve as a guide for researchers, practitioners, and policymakers committed to preserving the integrity of digital information in a time when synthetic media predominates by offering a roadmap for future study and development in the field of deepfake detection.

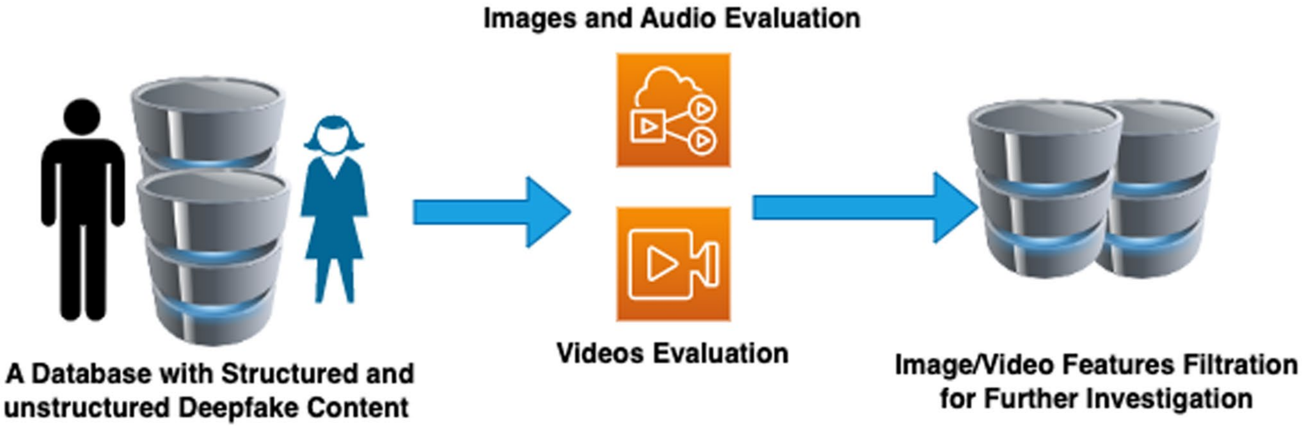
Images and Audio Evaluation
A Database with Structured and unstructured Deepfake Content
Videos Evaluation
Image/Video Features Filtration for Further Investigation

Fig. 1 Cycle of deepfake analysis and countermeasures

1.1 Security protocols in multimedia-enabled deepfakes

Digital security is now severely limited by the exponential growth of multimedia-enabled deepfakes, which calls for the creation of strong mechanisms to safeguard integrity and authenticity [11, 12]. These methods use cutting-edge technology including multimodal analysis, blockchain DLT, deep learning, and cryptographic watermarking to mitigate the hazards of multimedia-enabled deepfake alteration and manipulation. Nonetheless, digital watermarking is essential for incorporating imperceptible and unchangeable identifiers into multimedia files. Digital media authenticity and traceability are guaranteed by these markers, which may be metadata-based or cryptographic [12, 13]. Fragile watermarking is perfect for verifying the secrecy of sensitive media since it can detect even the smallest changes, whereas robust watermarking techniques are resistant to changes like compression. Watermarked content guarantees that news and evidence maintain their credibility while offering creators intellectual property protection.

An efficient distributed and immutable environment for content authentication is offered by Blockchain Distributed Ledger Technology (BDLT) [13, 14]. Any changes to the material are instantly identifiable because to this technology, which stores the cryptographic hashes of the original multimedia files on a distributed ledger. This idea creates a transparent and reliable environment for digital media verification in addition to preventing unwanted modifications. Blockchain-enabled protocols are especially beneficial for fields like digital rights management, journalism, and legal forensics. Conversely, AI approaches are used in Deep Learning-enabled detection systems to evaluate and identify abnormalities and detect altered information [15–17]. In order to detect small discrepancies specific to deepfakes, such as unnatural face movements and visual/audio mismatches, these techniques need a vast amount of data for training and authenticating the fake media. However, the sophisticated method of DL counts, which uses visual-audio sync detection and Convolutional Neural Networks (CNNs) to detect forgeries and tampering, allows for real-time detection on social media platforms and live streaming services [17, 18].

Moreover, hashing and metadata verification and validation offer extra layers of data protection and security by using the cryptographic hashing algorithm (SHA-256) to create unique identifiers for multimedia-enabled social media contents, which are then safely stored on the designated ledger and compared with newly generated hashes to identify any unauthorised changes [19, 20]. Metadata analysis supports hashing by confirming embedded details such as timestamps and source information, which guarantees that the media remains unaltered, unforged, and authentic in every way.

1.2 Motivation and contribution of this survey

Deepfake technology presents a serious threat to upholding public confidence, international security, and individual privacy, even while it has creative uses in fields like accessibility, entertainment, and education. As a result, deepfake technology is now more erratic. Inappropriate use of deepfakes has been linked to disinformation campaigns, financial fraud, political manipulation, cyber harassment, and personal theft. The integrity of digital content is jeopardised due to the possibility of very convincing synthetic media, which makes it more difficult to distinguish fact from fiction in important areas including news, social media communication, and legal evidence. However, the primary reason for the necessity to address the issue is the intricacy of modern deepfake manufacturing and related techniques. Highly realistic fake media can be produced with few noticeable artefacts using Generative Adversarial Networks (GANs) and other sophisticated algorithms. These developments have made traditional detection techniques outdated, requiring the creation of state-of-the-art systems that can function in real-time and across a variety of applications.

Deepfakes also have significant cultural repercussions. They could undermine democratic processes, undermine confidence in digital systems, and foster a climate of distrust and anxiety. In the current information era, where digital media is the main channel for stakeholder communication and aids in effective decision-making, this is especially important. Investigating a better detection hierarchy that can stop the abuse of multimedia-enabled deepfake technology is the main driving force behind this study. Through a thorough analysis of cutting-edge algorithms and methodologies, this research seeks to shed light on the advantages, disadvantages, and gaps of the current solutions. The goal is to provide professionals, young people, technologists, and legislators with the information and resources they need to reduce the risk of multimedia-enabled deepfakes and ensure a safe and reliable online environment.

However, the major contribution of this paper is highlighted as follows:

- This survey paper classifies multimedia-enabled deepfakes into DL-enabled techniques by providing a comprehensive analysis of the different detection procedures utilized to detect them.

- The Deepfake detection algorithms are evaluated using evaluation metrics, which are the primary topic of this work. It gives a detailed explanation of how well certain performance metrics, such as recall and precision, work.
- By providing a review of these new trends, particularly self-supervised learning, this work helps specialists investigate next generation deepfake detection mechanisms.
- This study draws attention to the main drawbacks in the domains, such as adversarial attacks, dataset generalisability, and ethical and legal issues pertaining to the deepfake detection cycle.
- This paper proposal of future directions for deepfake detection research, such as the incorporation of BDLT for deepfake detection in the near future, is one of its major contributions.
- Experts can use the paper as a comprehensive resource to comprehend the current state of deepfake detection and to gain practical ideas for reducing the risk connected with social media platforms.

1.3 Related works and related evaluation

By combining textual, visual, and aural data for a thorough assessment, multimodal analysis enhances detection capacities [21–23]. This technique detects mismatches across modalities, such as irregularities in the audio and lip movements, as well as artificial lighting in videos. Multimodal analysis provides a quick way to identify intricate deepfakes that take advantage of many aspects of multimedia content by combining data from multiple sources [22, 23]. Another possibility we consider in this study is biometric authentication, which validates the evaluation of people in multimedia and adds an extra degree of data security. While voice authentication cross-checks audio signatures, facial recognition technologies match a partial face match hierarchy with a certified database. This method is particularly helpful for digital identity authentication and virtual meetings, when guaranteeing the validity of the participants is critical [24–26].

The selecting criteria for literature, including:

- Web of Science, IEEE Xplore, Elsevier, ACM Library, Springer Link, MDPI Library, and Scopes are the databases that were used.
- Multimedia-enabled deepfake detection is one of the inclusion criteria used to choose studies.
- The last ten years are the time period covered.

We discovered ethical and legal framework difficulties by analysing deepfakes from many angles. These issues are crucial for regulating the development and spread of deepfake technology. Important steps to reduce the danger associated with this technology include requiring watermarks for all AL-enabled ML-generated content, enforcing sanctions for illicit deepfake usage, and encouraging openness in tool development [27, 28]. Public awareness efforts are necessary to support Suh frameworks by educating users about deepfake contents and the tools available to identify them [29–31].

Nonetheless, the following highlights this survey paper's outline distribution: The overview of deepfake creation and generative adversarial networks, as well as how GANs operate, are covered in Sect. 2. Section 3 discusses the use of AI-enabling approaches in conjunction with deepfakes to identify fake media material and its originality on social media. Section 4 lists the real-time deepfake tools, applications, and techniques that are used throughout. Section 5 discusses the limitations of the current design, which are the open challenges. Lastly, the conclusion and future prospects in Sect. 6 bring this paper to a close.

2 Overview of generative adversarial networks (GANs) and deepfake

Generative adversarial networks, sometimes known as GANs, are sophisticated models in the field of artificial intelligence, specifically generative AI [31–34]. GANs, in which neural networks act as the generator and discriminator in competition with one another, were first proposed by Ian Goodfellow in 2014. While the discriminatory provides a simulation that separates the real data from fraudulent data using both imitation and real samples, the generator mimics real data by creating fraud data. In exchange, the discriminator learns to recognise fraud data, and the generator learns to create fraud data that looks like the original image [35, 36].

GANs have been applied to a range of tasks in recent research, such as data enhancement, style transfer, video production, and image synthesis, as shown in Fig. 2. Their capacity to create naturally occurring, high-quality multimedia has sparked incredible advancements in the fields of design, entertainment, and medicine. However, this feature has also made it possible to use GANs to create irrational material, such as deepfakes [37, 38]. Media footage that has been

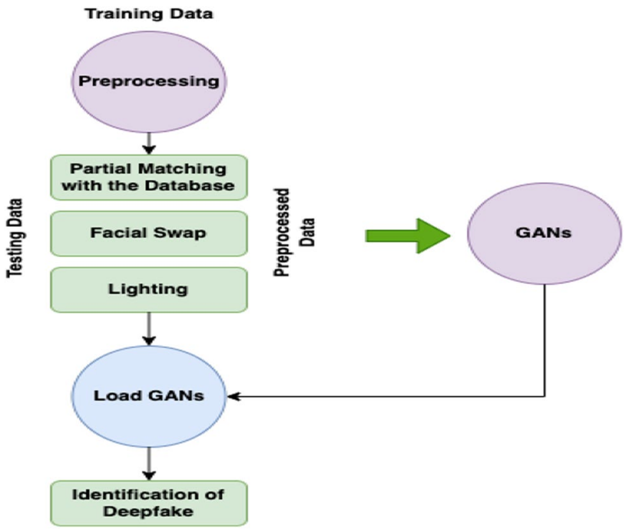

Fig. 2 Working cycle of GANs

artificially created or altered utilising a number of traditional techniques is known as a "deepfake", as shown in Fig. 2. These sophisticated tools have the ability to alter or replace any audio and video footage such that a person or event appears entirely fake but lifelike. For instance, although deepfake audio accurately mimics a person's voice, deepfake videos typically feature a different person's visage [39–41].

Although deepfake technology has generated controversy, there have also been certain benefits. For example, this has facilitated the creation of creative content such as virtual reality, video games, and movies [42–44]. But there has always been worried about these technologies being weaponised and used to spread misinformation or disparage people and organisations. Given their capacity to learn and replicate complex data distributions, GANs play a pivotal role in the production of deepfakes [45, 46]. More specialised edits are now possible thanks to other architectures, such as condition GANs, which have enhanced content creation capabilities and control methods, as shown in Fig. 2. Even the most sophisticated deepfakes, meanwhile, have identifying characteristics that can be exploited to identify them, including strange body motions or lighting [47, 48]. The development of detection algorithms, enhanced authentication procedures, and ethical concepts for the use of generative technologies were all undertaken in an attempt to tackle the issue of deepfake usage [48–50]. The true task for researchers, politicians, and technologists as GANs and deepfake technology advance will be to balance the potential risks and benefits of this technology.

3 Collaboration of deepfake with the real-time applications

To produce realistic fake content, deepfakes are created using advanced DL techniques, mainly GANs, autoencoders, and other DL-enabled approaches. Because of their capacity to withstand the posting of high-quality synthetic material on social media, the idea of technological integration has attracted a lot of interest [51, 52]. One of the most widely used deepfake production tools that we found in this study is DeepFaceLab, which enables users to produce high-quality face swapping videos using the DL approach, as shown in Figure 3. However, the main goal is to construct tools that can analyse deepfakes in real-time. The technology can be used for facial modification and alteration, including ageing and changes in facial expressions [53, 54]. It provides a variety of pre-trained models. We must assess which prospects are applying to make deepfakes before establishing a preventive atmosphere. It features face swapping, face replacement, learning fidelity for face swapping, complete control over DL-enabled models and their parameters for fine-tuning, and a training guide for creating deepfakes submitted by a big community.

A variety of tools have been developed to do various kinds of manipulations for amusement [55, 56], as shown in Figure 3. However, there is typically a fine line separating bullying from amusement. In this way, users can manipulate, and swap faces on photos and videos using Faceswap, an open source deepfake program that has features similar to DeepFaceLab [57–59]. It supports many neural networks for various kinds of deepfakes and offers a multitude of choices for fine-tuning the outcomes, as shown in Figure 3. Furthermore, a Chinese deepfake app gained popularity because to its ease of use and capability to instantly switch faces with well-known actors and movie characters. Consequently, it uses deep learning techniques to enable smooth face swaps in short video clips [59, 60].

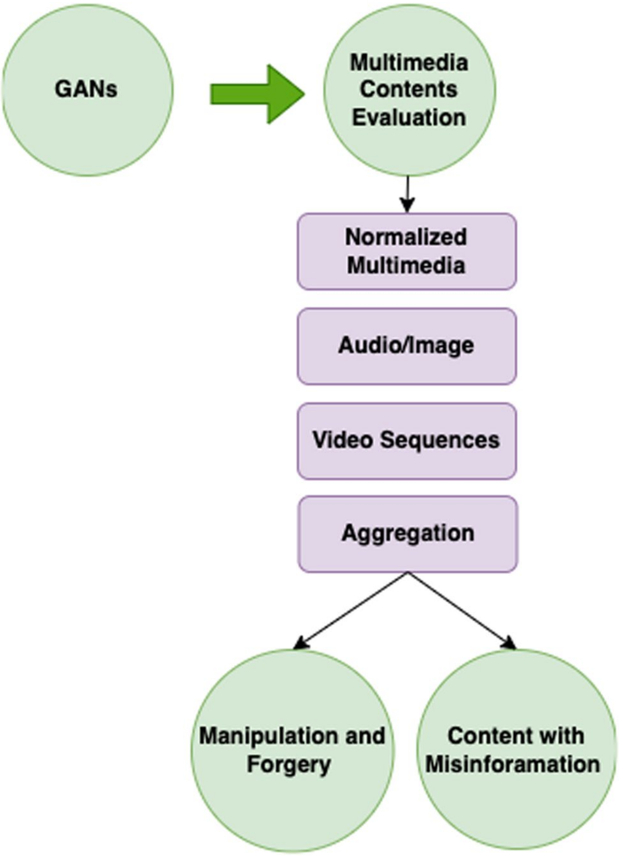

Fig. 3 Multimedia content evaluation for deepfake analysis

Reface is a smartphone app that lets users replace their faces with famous people in a range of video clips. It is very popular on social media since it employs DL techniques to do realistic face swaps in a second. DeepArt is not quite a deepfake tool, but it employs neural networks to create art by applying well-known artists' styles to pictures [61–63]. Styled movies and other artistic variations of deepfake content can be produced with it. Additionally, a real-time deepfake program that uses a still image to animate a person face. It can be used to substitute a dynamic avatar of well-known celebrities for a person's face in video conferencing apps like Zoom. Sensity AI does, however, provide tools for identifying deepfake concepts and pictures. It analyses image and video metadata and detects synthetic media using sophisticated machine learning techniques. Its core technology can be used to detect corrupted information and prevent the generation of deepfakes, even though its primary purpose is deepfake detection [64, 65]. Conversely, Lip Sync AI use deep learning to substitute fresh audio for speech in videos, synchronising it with the speaker's mouth movements [66, 67]. Making phoney videos of people saying things they never said is a typical usage for it.

3.1 Ethical considerations and limitations

The following significant ethical and legal issues might arise from the misuse of innovative tools that can be used for creative, educational, and entertainment reasons while taking ethical considerations into account: (1) falsely represented content; (2) impersonation fraud; and (3) invasions of privacy.

- Diverse tools have been proposed that can be used to create unauthorised content involving private individuals, violating their consent and privacy [68, 69].
- Deepfakes can be used to identify theft and frauds, impersonate forgery individuals for malicious purposes, and propagate false information that can damage reputations and undermine trust in digital media [70–72].

4 Real-time deepfake tools, techniques, and applicational evaluation

Deepfake detection has many applications, especially when it comes to using cutting-edge technologies and approaches to defend privacy, security, and trust in a variety of industries. The following is a mention of the discussion list:

4.1 Cybersecurity and fraud prevention

Social engineering attacks, identity theft, and impersonation can all be carried out using technologies such as deepfake [73]. In cybersecurity, deepfake detection technologies can be used to stop fraudulent activities like financial scams, phishing, and illegal access to secure ecosystems, as well as to verify identities in vital communications like voice-based authentication, audio synchronisation, and video conferences. However, in applications that use biometric systems for identity verification, including voice authentication and facial recognition, Deepfake detection is essential. By including deepfake detection in biometric systems, spoofing attempts can be avoided and access control mechanisms, such as payment systems, can be protected from sophisticated impersonation assaults.

4.2 Misinformation and fake news prevention

The propagation of false information is greatly aided by technologies such as deepfake and their present detection methods, especially when political campaigns, popular figures, and important events are involved. Detection systems can aid in stopping the viral spread of fake news that could sway public opinion, interfere with electronics, and threaten social stability by recognising and marking distorted media. Furthermore, deepfake detection systems can be integrated into social media handles, where deepfake content has the potential to spread rapidly, to automatically identify and eliminate audio, video, and image manipulation and alteration. This lessens the likelihood that deepfakes would be maliciously used to harass people, disseminate false information, and upset social order. Platform moderators can help remove hazardous information by using dynamic detection methods [74–76].

4.3 Digital forensics and evidence validation

Verifying the legitimacy of Chain of Custody (CoC) in legal and investigative contexts requires deepfake detection. This technology guarantees the authenticity of audio and video recordings that are used in crime investigations and are presented as a Chain of Evidence (CoE) [77, 78]. By verifying the accuracy of media used in court, detection systems help preserve the legitimacy of legal processes. In the context of the media sector, the importance of verifying the authenticity of media content has increased as a result of the digital media revolution. To stop the spread of intentionally altered content that could harm reputations and divert viewers, news organisations, social media platforms, and content can use deepfake detection techniques to confirm and validate that videos, photos, and audio shared with the public are real [79, 80].

5 Challenges and open research problems

In order to create reliable and scalable solutions, a number of open research concerns are brought to light by the quick development of deepfake technology. These problems illustrate how difficult it is to counteract deepfakes and cut across technical, ethical, and societal dimensions. The following is the discussion list:

5.1 Adversarial attacks in deepfake detection processes

Small, undetectable changes to the deepfake content can fool detection systems into identifying it as authentic, making deepfake detection models susceptible to adversarial attacks [81]. One significant constraint is the development of detection models that are resistant to such hostile modifications and manipulations. To create more robust models that can identify multimedia-enabled deepfake content even in hostile environments, more research is required. However, in addition to the novel forms of deepfakes prevention, the existing detection methods frequently work well on the datasets they are trained on but have trouble generalising to different datasets. Developing a lifecycle that can function well across many datasets, content sources, and deepfake generating algorithms is a major research challenge. Creating domain adaptability and transfer learning strategies may be crucial in this context.

5.2 Long-term viability and adaptability

Multimedia-capable new techniques are being created to make deepfakes even more difficult to identify as deepfake technology (MDT) continues to advance. It is crucial to create detection algorithms that are flexible enough to change with new deepfake methods. Investigations into self-updating systems and lifelong learning may be worthwhile [82–84]. However, deepfake detection models, particularly those based on DL, frequently operate as "black boxes," making it challenging to comprehend how they make their decisions [85–87]. Transparency and trust are hampered by explainability, especially in ethical and legal circumstances [88, 89]. To create algorithms that yield results that can be interpreted and allow for increased confidence in their outputs, explainable AI (XAI) research is essential for deepfake detection.

5.3 Blockchain integration for traceability

Because blockchain distribute ledger technology (BDLT) creates irreversible records of the provenance of content, it may be used to verify the legitimacy of media files [90–93]. To assist assure the authenticity of media content, research is required to determine how blockchain can be included into deepfake detection algorithms to provide transparent and safe tracking of the content's origin [94, 95]. The evaluation of numerous existing deepfake detection methods is also helped by this research. These methods are unsuited for real-time applications due to their high computational costs and processing power requirements. There is an urgent need to conduct research on efficient, lightweight gadgets, such as cell phones [96–101].

6 Conclusion and future directions

The advent of deepfake technology is discussed in this survey article, along with its notable drawbacks and promising future applications across a range of industries. There is conversation about social media, entertainment, cybersecurity, and both domestic and global security. On the other side, deepfake detection methods are now essential for safeguarding digital authentication, which prevents illegal content manipulation in addition to preserving media integrity. This survey provides a comprehensive review of the state-of-the-art deepfake detection techniques, while also highlighting evaluation measures, new trends, and limitations in the current infrastructure of deepfake detection research. Significant progress has been made in the creation of detection systems based on DL-enabled complex hierarchies. Reaching near-term technological maturity requires addressing several significant limitations, including real-time detection systems, cross-domain generalisation, and rapid protection against complicated deepfake manufacturing. Given how quickly deepfake generation techniques are developing, this study examines and identifies the areas that need more research and development: (1) enhancing detection robustness; (2) real-time systems; (3) cross-domain generalisation; (4) multimodal deepfake detection; (5) industry-academia collaboration; and (6) developing countermeasures and prevention strategies [102–105].

Acknowledgements Not applicable

Author contributions Abdullah Ayub Khan perform the Original Writing Part, Software, and Methodology; Abdullah Ayub Khan, Asif Ali Laghari, Syed Azeem Inam, Sajid Ullah, Muhammad Shahzad, Darakhshan Syed perform Rewriting, investigation, design Methodology, and Conceptualization; Abdullah Ayub Khan, Asif Ali Laghari, Syed Azeem Inam, Sajid Ullah performs related work part and manage results and discussions; Abdullah Ayub Khan, Muhammad Shahzad, Darakhshan Syed perform related work part and manage results and discussion; Abdullah Ayub Khan, Asif Ali Laghari, Syed Azeem, Darakhshan Syed perform Rewriting, design Methodology, and Visualization; Abdullah Ayub Khan performs Rewriting, design Methodology, and Visualization. The author accepts the online/published version of this manuscript.

The author accepts the online/published version of this manuscript.

Funding Not Applicable.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not Applicable.

Consent for publication Not Applicable.

Competing interest The authors declare no competing interests.

References

1. Heidari A, Jafari Navimipour N, Dag H, Unal M. Deepfake detection using deep learning methods: a systematic and comprehensive review. Wiley Interdis Rev. 2024;14(2): e1520.
2. Sun G, Zhang Y, Yu H, Du X, Guizani M. Intersection fog-based distributed routing for V2V communication in urban vehicular ad hoc networks. IEEE Trans Intell Transp Syst. 2020;21(6):2409–26. https://doi.org/10.1109/TITS.2019.2918255.
3. Sun G, Song L, Yu H, Chang V, Du X, Guizani M. V2V routing in a VANET based on the autoregressive integrated moving average model. IEEE Trans Vehicul Technol. 2019;68(1):908–22. https://doi.org/10.1109/TVT.2018.2884525.
4. Sun G, Zhang Y, Liao D, Yu H, Du X, Guizani M. Bus-trajectory-based street-centric routing for message delivery in urban vehicular ad hoc networks. IEEE Trans Veh Technol. 2018;67(8):7550–63. https://doi.org/10.1109/TVT.2018.282865.
5. Khan, A. A., Laghari, A. A., Alroobaea, R., Baqasah, A. M., Alsafyani, M., Bacarra, R., & Alsayaydeh, J. A. J. (2024). Secure Remote Sensing Data With Blockchain Distributed Ledger Technology: A Solution for Smart Cities. *IEEE Access*.
6. Ullah S, Qiao X, Abbas M. Addressing the impact of land use land cover changes on land surface temperature using machine learning algorithms. Sci Rep. 2024;14:18746. https://doi.org/10.1038/s41598-024-68492-7.
7. Khan AA, Wagan AA, Laghari AA, Gilal AR, Aziz IA, Talpur BA. BIoMT: A state-of-the-art consortium serverless network architecture for healthcare system using blockchain smart contracts. IEEE Access. 2022;10:78887–98.
8. Ullah S, Abbas M, Qiao X. Impact assessment of land-use alteration on land surface temperature in Kabul using machine learning algorithm. J Spat Sci. 2024;70:1–23. https://doi.org/10.1080/14498596.2024.2364283.
9. Xia J, Li S, Huang J, Yang Z, Jaimoukha IM, Gündüz D. Metalearning-Based Alternating Minimization Algorithm for Nonconvex Optimization. IEEE Trans Neu Net Learn Sys. 2023;34(9):5366–80. https://doi.org/10.1109/TNNLS.2022.3165627.
10. Khan AA, Dhabi S, Yang J, Alhakami W, Bourouis S, Yee L. B-LPoET: A middleware lightweight Proof-of-Elapsed Time (PoET) for efficient distributed transaction execution and security on Blockchain using multithreading technology. Comput Electr Eng. 2024;118:109343.
11. Ullah S, Qiao X, Tariq A. Impact assessment of planned and unplanned urbanization on land surface temperature in Afghanistan using machine learning algorithms: a path toward sustainability. Sci Rep. 2025;15:3092. https://doi.org/10.1038/s41598-025-87234-x.
12. Han F, Yang P, Du H, Li X. Accuth+: Accelerometer-Based Anti-Spoofing Voice Authentication on Wrist-Worn Wearables. IEEE Trans Mob Comput. 2024;23(5):5571–88. https://doi.org/10.1109/TMC.2023.3314837.
13. Zuo C, Zhang X, Yan L, Zhang Z. GUGEN: Global User Graph Enhanced Network for Next POI Recommendation. IEEE Trans Mob Comput. 2024;23(12):14975–86. https://doi.org/10.1109/TMC.2024.3455107.
14. Yang K. How to prevent deception: A study of digital deception in "visual poverty" livestream. New Media Soc. 2024. https://doi.org/10.1177/14614448241285443.
15. Khan AA, Laghari AA, Baqasah AM, Alroobaea R, Almadhor A, Sampedro GA, Kryvinska N. Blockchain-enabled infrastructural security solution for serverless consortium fog and edge computing. PeerJ ComputSci. 2024;10:e1933.
16. Li C, He A, Liu G, Wen Y, Chronopoulos AT, Giannakos A. RFL-APIA: a comprehensive framework for mitigating poisoning attacks and promoting model aggregation in IIoT federated learning. IEEE Trans Industr Inf. 2024;20(11):12935–44. https://doi.org/10.1109/TII.2024.3431020.
17. Lin, W., Xia, C., Wang, T., Zhao, Y., Xi, L., … Zhang, S. (2024). Input and Output Matter: Malicious Traffic Detection with Explainability. IEEE Networkhttps://doi.org/10.1109/MNET.2024.3481045
18. Khan AA, Chen YL, Hajjej F, Shaikh AA, Yang J, Ku CS, Por LY. Digital forensics for the socio-cyber world (DF-SCW): A novel framework for deepfake multimedia investigation on social media platforms. Egypt Informat J. 2024;27:100502.
19. Heidari A, Navimipour NJ, Dag H, Talebi S, Unal M. A novel blockchain-based deepfake detection method using federated and deep learning models. Cogn Comput. 2024;16:1073–91.
20. Haq IU, Malik KM, Muhammad K. Multimodal neurosymbolic approach for explainable deepfake detection. ACM Trans Multimed Comput Commun Appl. 2024;20(11):1–16.
21. Aghasanli, A., Kangin, D., & Angelov, P. (2023). Interpretable-through-prototypes deepfake detection for diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 467–474).

22. Nadimpalli AV, Rattani A. Proactive deepfake detection using gan-based visible watermarking. ACM Trans Multimed Comput Commun Appl. 2024;20(11):1–27.
23. Guarnera, L., Giudice, O., & Battiato, S. (2024). Mastering deepfake detection: A cutting-edge approach to distinguish gan and diffusion-model images. ACM Transactions on Multimedia Computing, Communications and Applications.
24. Khan AA, Laghari AA, Shafiq M, Cheikhrouhou O, Alhakami W, Hamam H, Shaikh ZA. Healthcare ledger management: A blockchain and machine learning-enabled novel and secure architecture for medical industry. Hum Cent Comput Inf Sci. 2022;12:55.
25. Li H, Xia C, Wang T, Wang Z, Cui P, Li X. GRASS: learning spatial-temporal properties from chainlike cascade data for microscopic diffusion prediction. IEEE Trans Neu Net Learn Sys. 2024;35(11):16313–27. https://doi.org/10.1109/TNNLS.2023.3293689.
26. Li D, Xing W. A comparative study on sustainable development of online education platforms at home and abroad since the twenty-first century based on big data analysis. Educ Inf Technol. 2025. https://doi.org/10.1007/s10639-025-13400-3.
27. Chen H, Bei Y, Huang W, Chen S, Huang F, Huang X. Graph Cross-Correlated Network for Recommendation. IEEE Trans Knowled Data Eng. 2025;37(2):710–23. https://doi.org/10.1109/TKDE.2024.3491778.
28. Jain, A., Gaur, A., Gupta, G., Mishra, S., Johari, R., & Vidyarthi, D. P. (2025). Securing Digital Integrity: Proposed Comprehensive Framework for Deepfake Detection and Blockchain Validation. In International Conference on Cognitive Computing and Cyber Physical Systems (pp. 579-589). Springer, Singapore
29. Cheng, G., Xia, J., Luo, L., Mi, H., Guo, D., … Ma, R. T. B. (2024). HyperPart: A Hypergraph-based Abstraction for Deduplicated Storage Systems. IEEE Transactions on Cloud Computing, 1–15. https://doi.org/10.1109/TCC.2024.3502464.
30. Liu X, Lou S, Dai W. Further results on "System identification of nonlinear state-space models." Automatica. 2023;148:110760. https://doi.org/10.1016/j.automatica.2022.110760.
31. Li, L., Cherouat, A., Snoussi, H., & Wang, T. (2024). Grasping With Occlusion-Aware Ally Method in Complex Scenes. IEEE Trans Automat Sci Eng, 1–11. https://doi.org/10.1109/TASE.2024.3434610.
32. Wang T, Hou B, Li J, Shi P, Zhang B, Snoussi H. TASTA: text-assisted spatial and temporal attention network for video question answering. Adv Intell Sys. 2023;5(4):2200131. https://doi.org/10.1002/aisy.202200131.
33. Khan AA, Laghari AA, Baqasah AM, Bacarra R, Alroobaea R, Alsafyani M, Alsayaydeh JAJ. BDLT-IoMT—a novel architecture: SVM machine learning for robust and secure data processing in Internet of Medical Things with blockchain cybersecurity. J Supercomput. 2025;81(1):1–22.
34. Alnaim NM, Almutairi ZM, Alsuwat MS, Alalawi HH, Alshobaili A, Alenezi FS. DFFMD: a deepfake face mask dataset for infectious disease era with deepfake detection algorithms. IEEE Access. 2023;11:16711–22.
35. Gao Y, Wang X, Zhang Y, Zeng P, Ma Y. Temporal feature prediction in audio-visual deepfake detection. Electronics. 2024;13(17):3433.
36. Saher M, Alsaedi M, Al Ibraheemi A. Automated grading system for breast cancer histopathological images using histogram of oriented gradients (HOG) Algorithm. Appl Data SciAnal. 2023;2023:78–87.
37. Najim FT. Effect of inlet velocity on the entrance length of laminar and turbulent flow in a circular pipe. Babylon J Mech Eng. 2024;2024:81–90.
38. Yaseen MG, Abed SA. Schizophrenia and the role of artificial intelligence in detecting and treating it: cognitive frontiers. Mesopot J Artif Int Healthcare. 2023;2023:61–5.
39. Mebarek-Oudina F. Exploring passive heat transfer enhancement techniques: applications, benefits, and challenges. Babylon J Mech Eng. 2024;2024:122–7.
40. Hussein A, Sallam ME, Abdalla MYA. Exploring new horizons: surgical robots supported by artificial intelligence. Mesopotam J Artif Intell Healthcare. 2023;2023:40–4.
41. Alaiwi Y, Ahmed T. Solar Air Heaters Classifications and Enhancement: A Review. Babylon J Mech Eng. 2024;2024:71–80.
42. Gopi RS, Suganthi R, Hephzipah JJ, Amirthayogam G, Sundararajan P, Pushparaj T. Elderly people health care monitoring system using internet of things (IOT) for exploratory data analysis. Babylon J Artific Intell. 2024;2024:54–63.
43. Shi G, Deng S, Wang B, Feng C, Zhuang Y, Wang X. One for All: A Unified Generative Framework for Image Emotion Classification. IEEE Trans Circ Syst Video Technol. 2024;34(8):7057–68. https://doi.org/10.1109/TCSVT.2023.3341840.
44. Chen Y, Li H, Song Y, Zhu X. Recoding hybrid stochastic numbers for preventing bit width accumulation and fault tolerance. IEEE Trans Circ Syst I: Reg Paper. 2024;72:1–13. https://doi.org/10.1109/TCSI.2024.3492054.
45. Peng Y, Chen X, Miao D, Qin X, Gu X, Lu P. Unveiling user identity across social media: a novel unsupervised gradient semantic model for accurate and efficient user alignment. Compl Intell Syst. 2024;11(1):24. https://doi.org/10.1007/s40747-024-01626-6.
46. Zhang M, Wei E, Berry R, Huang J. Age-dependent differential privacy. IEEE Trans Inf Theory. 2024;70(2):1300–19. https://doi.org/10.1109/TIT.2023.3340147.
47. Shi H, Hayat M, Cai J. Unified open-vocabulary dense visual prediction. IEEE Trans Multimedia. 2024;26:8704–16. https://doi.org/10.1109/TMM.2024.3381835.
48. Gong Y, Yao H, Xiong Z, Chen CLP, Niyato D. Blockchain-aided digital twin offloading mechanism in space-air-ground networks. IEEE Trans Mob Comput. 2025;24(1):183–97. https://doi.org/10.1109/TMC.2024.3455417.
49. Khan AA, Zhang X, Hajjej F, Yang J, Ku CS, Por LY. ASMF: Ambient social media forensics chain of custody with an intelligent digital investigation process using federated learning. Heliyon. 2024;10(1): e23254.
50. Zhang Y, Lin W, Xu J. Joint audio-visual attention with contrastive learning for more general deepfake detection. ACM Trans Multimed Comput Commun Appl. 2024;20(5):1–23.
51. Patel, Y., Tanwar, S., Gupta, R., Bhattacharya, P., Davidson, I. E., Nyameko, R., … & Vimal, V. (2023). Deepfake generation and detection: Case study and challenges. *IEEE Access*.
52. Gowrisankar B, Thing VL. An adversarial attack approach for eXplainable AI evaluation on deepfake detection models. Comput Secur. 2024;139:103684.
53. Xu, Y., Liang, J., Jia, G., Yang, Z., Zhang, Y., & He, R. (2023). Tall: Thumbnail layout for deepfake video detection. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 22658–22668).
54. Maheshwari, R. U., Kumarganesh, S., KVM, S., Gopalakrishnan, A., Selvi, K., Paulchamy, B., ... & Pandey, D. (2024). Advanced plasmonic resonance-enhanced biosensor for comprehensive real-time detection and analysis of deepfake content. *Plasmonics*, 1–18.

55. Gong Y, Yao H, Liu X, Bennis M, Nallanathan A, Han Z. Computation and Privacy Protection for Satellite-Ground Digital Twin Networks. IEEE Trans Commun. 2024;72(9):5532–46. https://doi.org/10.1109/TCOMM.2024.3392795.
56. Li T, Long Q, Chai H, Zhang S, Jiang F, Liu H, Li Y. Generative AI Empowered Network Digital Twins: Architecture, Technologies, and Applications. ACM Comput Surv. 2025;57(6):157. https://doi.org/10.1145/3711682.
57. Ding Y, Zhang W, Zhou X, Liao Q, Luo Q, Ni LM. FraudTrip: Taxi Fraudulent Trip Detection From Corresponding Trajectories. IEEE Intern ThingsJ. 2021;8(16):12505–17. https://doi.org/10.1109/JIOT.2020.3019398.
58. Raza MA, Malik KM, Haq IU. Holisticdfd: Infusing spatiotemporal transformer embeddings for deepfake detection. Inf Sci. 2023;645:119352.
59. Ahmed, S. R., & Sonuç, E. (2023). Evaluating the effectiveness of rationale-augmented convolutional neural networks for deepfake detection. *Soft Computing*, 1–12.
60. Gupta G, Raja K, Gupta M, Jan T, Whiteside ST, Prasad M. A comprehensive review of deepfake detection using advanced machine learning and fusion methods. Electronics. 2023;13(1):95.
61. Tan, C., Zhao, Y., Wei, S., Gu, G., Liu, P., & Wei, Y. (2024, March). Frequency-Aware Deepfake Detection: Improving Generalizability through Frequency Space Domain Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 5, pp. 5052–5060).
62. Xie, Y., Cheng, H., Wang, Y., & Ye, L. (2023). Domain generalization via aggregation and separation for audio deepfake detection. *IEEE Transactions on Information Forensics and Security*.
63. Khan AA, Yang J, Laghari AA, Baqasah AM, Alroobaea R, Ku CS, Por LY. BAIoT-EMS: Consortium network for small-medium enterprises management system with blockchain and augmented intelligence of things. Eng Applic Artif Intell. 2025;141:109838.
64. Nguyen, D., Mejri, N., Singh, I. P., Kuleshova, P., Astrid, M., Kacem, A., … Aouada, D. (2024). LAA-Net: Localized Artifact Attention Network for Quality-Agnostic and Generalizable Deepfake Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 17395–17405).
65. Narayan, K., Agarwal, H., Thakral, K., Mittal, S., Vatsa, M., & Singh, R. (2023). Df-platter: Multi-face heterogeneous deepfake dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9739–9748).
66. Almestekawy A, Zayed HH, Taha A. Deepfake detection: Enhancing performance with spatiotemporal texture and deep learning feature fusion. Egypt Informat J. 2024;27:100535.
67. Mittal, G., Hegde, C., & Memon, N. (2024, July). GOTCHA: Real-time video deepfake detection via challenge-response. In *2024 IEEE 9th European Symposium on Security and Privacy (EuroS&P)* (pp. 1–20). IEEE.
68. Becattini F, Bisogni C, Loia V, Pero C, Hao F. Head pose estimation patterns as deepfake detectors. ACM Trans Multimed Comput Commun Appl. 2024;20(11):1–24.
69. Lin H, Huang W, Luo W, Lu W. DeepFake detection with multi-scale convolution and vision transformer. Dig Sign Process. 2023;134:103895.
70. Wu J, Zhang B, Li Z, Pang G, Teng Z, Fan J. Interactive two-stream network across modalities for deepfake detection. IEEE Trans Circuits Syst Video Technol. 2023;33(11):6418–30.
71. Oorloff, T., Koppisetti, S., Bonettini, N., Solanki, D., Colman, B., Yacoob, Y., ... & Bharaj, G. (2024). AVFF: Audio-Visual Feature Fusion for Video Deepfake Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 27102–27112).
72. Khan, S. A., & Dang-Nguyen, D. T. (2023). Deepfake Detection: Analysing Model Generalisation Across Architectures, Datasets and Pre-Training Paradigms. *IEEE Access*.
73. Seraj, M. S., Singh, A., & Chakraborty, S. (2024). Semi-Supervised Deep Domain Adaptation for Deepfake Detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 1061–1071).
74. Xia, R., Zhou, D., Liu, D., Yuan, L., Wang, S., Li, J., ... & Gao, X. (2024, October). Advancing Generalized Deepfake Detector with Forgery Perception Guidance. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 6676–6685).
75. Ke J, Wang L. DF-UDetector: An effective method towards robust deepfake detection via feature restoration. Neural Netw. 2023;160:216–26.
76. Khan, A. A., & Por, L. Y. (2024). Special Issue on Information Security and Cryptography The Role of Advanced Digital Technology. *Appl Sci*, *14*(5), 2045.
77. Cai Z, Li M. Integrating frame-level boundary detection and deepfake detection for locating manipulated regions in partially spoofed audio forgery attacks. Comput Speech Lang. 2024;85:101597.
78. Xu, Y., Liang, J., Sheng, L., & Zhang, X. Y. (2024). Learning Spatiotemporal Inconsistency via Thumbnail Layout for Face Deepfake Detection. *Int J Comput Vis*, 1–18.
79. Kaddar, B., Fezza, S. A., Akhtar, Z., Hamidouche, W., Hadid, A., & Serra-Sagristà, J. (2024). Deepfake detection using spatiotemporal transformer. *ACM Transactions on Multimedia Computing, Communications and Applications*.
80. Tian, C., Luo, Z., Shi, G., & Li, S. (2023). Frequency-aware attentional feature fusion for deepfake detection. In *ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1–5). IEEE.
81. Vaidya, A. O., Dangore, M., Borate, V. K., Raut, N., Mali, Y. K., & Chaudhari, A. (2024). Deep Fake Detection for Preventing Audio and Video Frauds Using Advanced Deep Learning Techniques. In *2024 IEEE Recent Advances in Intelligent Computational Systems (RAICS)* (pp. 1–6). IEEE.
82. Liu, X., Yu, Y., Li, X., & Zhao, Y. (2023). Mcl: multimodal contrastive learning for deepfake detection. *IEEE Transactions on Circuits and Systems for Video Technology*.
83. Naitali A, Ridouani M, Salahdine F, Kaabouch N. Deepfake attacks: Generation, detection, datasets, challenges, and research directions. Computers. 2023;12(10):216.
84. Yadav A, Vishwakarma DK. AW-MSA: Adaptively weighted multi-scale attentional features for DeepFake detection. Eng Appl Artif Intell. 2024;127:107443.
85. Guo, Z., Jia, Z., Wang, L., Wang, D., Yang, G., & Kasabov, N. (2023). Constructing new backbone networks via space-frequency interactive convolution for deepfake detection. *IEEE Transactions on Information Forensics and Security*.
86. Lu Y, Ebrahimi T. Assessment framework for deepfake detection in real-world situations. EURASIP J Image Video Process. 2024;2024(1):6.
87. Xie, Y., Cheng, H., Wang, Y., & Ye, L. (2023). Learning A Self-Supervised Domain-Invariant Feature Representation for Generalized Audio Deepfake Detection. In *Proc. INTERSPEECH* (Vol. 2023, pp. 2808–12).

88. Coccomini DA, Caldelli R, Falchi F, Gennaro C. On the generalization of deep learning models in video deepfake detection. J Imag. 2023;9(5):89.
89. Zhang, Y., Miao, C., Luo, M., Li, J., Deng, W., Yao, W., ... & Chu, Q. (2024, October). MFMS: Learning Modality-Fused and Modality-Specific Features for Deepfake Detection and Localization Tasks. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 11365–69).
90. Huang, B., Wang, Z., Yang, J., Ai, J., Zou, Q., Wang, Q., & Ye, D. (2023). Implicit identity driven deepfake face swapping detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4490–9).
91. Kingra S, Aggarwal N, Kaur N. SFormer: An end-to-end spatio-temporal transformer architecture for deepfake detection. Forens Sci Int: Dig Invest. 2024;51:301817.
92. Zhu C, Zhang B, Yin Q, Yin C, Lu W. Deepfake detection via inter-frame inconsistency recomposition and enhancement. Pattern Recogn. 2024;147:110077.
93. Heo YJ, Yeo WH, Kim BG. Deepfake detection algorithm based on improved vision transformer. Appl Intell. 2023;53(7):7512–27.
94. Gong LY, Li XJ. A contemporary survey on deepfake detection: datasets, algorithms, and challenges. Electronics. 2024;13(3):585.
95. Maheshwari, R. U., Paulchamy, B., Pandey, B. K., & Pandey, D. (2024). Enhancing sensing and imaging capabilities through surface plasmon resonance for deepfake image detection. *Plasmonics*, 1–20.
96. Kumar N, Kundu A. SecureVision: advanced cybersecurity deepfake detection with big data analytics. Sensors. 2024;24(19):6300.
97. Khan, A. A., Laghari, A. A., Baqasah, A. M., Alroobaea, R., Gadekallu, T. R., Sampedro, G. A., & Zhu, Y. (2024). ORAN-B5G: A next generation open radio access network architecture with machine learning for beyond 5G in industrial 5.0. *IEEE Transactions on Green Communications and Networking*.
98. Stanciu, D. C., & Ionescu, B. (2023, June). Autoencoder-based data augmentation for deepfake detection. In *Proceedings of the 2nd ACM International Workshop on Multimedia AI against Disinformation* (pp. 19–27).
99. Passos, L. A., Jodas, D., Costa, K. A., Souza Júnior, L. A., Rodrigues, D., Del Ser, J., … & Papa, J. P. (2024). A review of deep learning-based approaches for deepfake content detection. *Expert Systems*, *41*(8), e13570.
100. Akhtar Z, Pendyala TL, Athmakuri VS. Video and audio deepfake datasets and open issues in deepfake technology: being ahead of the curve. Forens Sci. 2024;4(3):289–377.
101. Sunil R, Mer P, Diwan A, Mahadeva R, Sharma A. Exploring Autonomous Methods for Deepfake Detection: A Detailed Survey on Techniques and Evaluation. Heliyon. 2025.
102. Karim S, Liu X, Khan AA, Laghari AA, Qadir A, Bibi I. MCGAN—a cutting edge approach to real timeinvestigate of multimedia deepfake multi collaboration of deep generative adversarial networks with transfer learning. Sci Rep. 2024;14(1):29330.
103. Khan AA, Laghari AA, Inam SA, Ullah S, Nadeem L. A review on artificial intelligence thermal fluids and the integration of energy conservation with blockchain technology. Discov Sustain. 2025;6(1):1-18.
104. Xiong J, Yang J, Yan L, Awais M, Khan AA, Alizadehsani R, Acharya UR. Efficient reinforcement learning-based method for plagiarism detection boosted by a population-based algorithm for pretraining weights. Expert Syst Appl. 2024;238:122088.
105. Khan AA, Shaikh AA, Laghari AA, Rind MM. Cloud forensics and digital ledger investigation: a new era of forensics investigation. Int J Electron Secur. Digit Forensics. 2023;15(1):1-23.