East Journal of Applied Science
ISSN: 3079-9392

Volume 1, Issue 4
Publisher: East Publication & Technology
DOI: https://doi.org/10.63496/ejas.Vol1.Iss4.104

Addressing Deepfake Technologies Through Detection and Regulation: A Systematic Survey

Rania ElshiekhHamid Ibrahim* *[1]

[1]Management information system, Business Studies, Sudan University of Science and Technology, Khartoum, Sudan, raniaelshiekh@gmail.com.

Received: 15/05/2025, **Revised**: 13/08/2025, **Accepted**: 15/08/2025, **Published**: 25/08/2028

***Abstract*:**

Deepfake technology is growing in popularity because of the rapid growth of artificial intelligence (AI), which creates realistic looking but fake audio and video content. Although this significant development has revolutionary possibilities, it also raises serious ethical issues, such as concerns to public confidence, privacy, and security, in addition to the chance for manipulation and false information. Deepfake technology, powered by advances in artificial intelligence, particularly Generative Adversarial Networks (GANs), has introduced both groundbreaking opportunities and serious ethical and security concerns. This survey provides a comprehensive overview of the current state of deepfake detection methods and regulatory frameworks aimed at mitigating the risks associated with synthetic media. After scanning over 73 documents, we reduced the selection to 57 applying evaluation criteria such abstract, title, irrelevant focus, and duplication. The author analyzes recent technological approaches, including convolutional neural networks (CNNs), multimodal analysis, and biological signal detection, and evaluate global legislative responses to deepfakes across various jurisdictions.

***Keywords*: Deepfake, Detection, Regulation.**

1. Introduction

The emergence of deepfake technologies has transformed the landscape of digital content creation. While these tools offer creative and beneficial applications, their misuse poses significant threats to privacy, democracy, and personal identity. Deepfakes refer to artificially generated or manipulated audiovisual content that appears authentic but is fabricated using machine learning algorithms. The research problem or question that your study aims to address, explaining why it is significant

This paper presents a systematic review of existing literature on deepfake detection techniques and regulatory strategies. A deepfake is any of the content—whether it be audio, video, or otherwise—that has been artificially manipulated and that is completely or partially fabricated. It's difficult to embrace anything nowadays, while generals may download and utilize hundreds of apps. At present, anyone can produce a deepfake. Still images may be employed to create deepfake videos (Song 2019).

1.1 Background and Evolution of Deepfake Technology

Deepfakes emerged from advancements in artificial intelligence, particularly the development of Generative Adversarial Networks (GANs) in 2014. This architecture enabled the generation of realistic images, videos, and audio with minimal human intervention. While GANs were vital to producing deepfakes, their popularity was significantly influenced by current online traditions and customs.

1.2 Rise of Accessibility

By 2018, open-source implementations of GANs became widely available, allowing non-experts to generate convincing fake content. Social platforms like Reddit played a pivotal role in popularizing deepfake creation, especially in video editing and face-swapping.

Therefore, this emergence has caused a great deal of scientific research and publications. The strength and variety of GANs have been identified by academics and researchers around the world, leading to an extensive variety of studies that address various aspects of these neural network architectures. The scope of academic research in this area climbed between 2014 and 2025. A key driver driving deepfake technology advancement is the increasing openness of simple-to-use tools, software, and communities on the internet (Mukta et al. 2023).

1.3 Ethical and Societal Implications of Deepfakes

1. Misinformation and Disinformation

Deepfakes are increasingly used to spread false narratives, manipulate elections, and damage reputations. As AI advances, the ability to manipulate media content raises concerns about authenticity, consent, and the potential for misinformation. The emergence of deepfake technology has sparked widespread debate about its implications for society. This paper aims to highlight the necessity of detection methods and regulatory measures.

2. Privacy Violations

Creating deepfakes without consent violates individual rights and raises concerns about bodily autonomy and image ownership.

3. Bias and Discrimination

Deepfakes often target marginalized groups, amplifying social divisions and promoting hate speech.

2. Related Work

Previous research has made significant contributions to the field of fake audio detection in two primary ways. Initially, a dataset named H-Voice was created utilizing the imitation method, which involved extracting entropy features from both authentic and counterfeit audio samples. This dataset enabled the researchers to construct a machine learning model employing Logistic Regression (LR) to identify fake audio. The model demonstrated a remarkable success rate of 98% in detection tasks; however, it required manual preprocessing of the data to extract the pertinent features (Ballesteros, Rodriguez, and Renza 2020).

The research paper contains an in-depth review of modern deepfake (DF) detection approaches, highlighting their benefits, drawbacks, and appropriateness for use with various types of products. It measures several approaches, such as advanced deep learning models and traditional machine learning, showing the importance of feature extraction, dataset quality, and the ability to be general. The authors raise concerns involving overfitting specific datasets, the high computational demands of state-of-the-art models, and the lack of consistency in cross-dataset evaluations. The research has major shortcomings despite providing useful details about the advantages and disadvantages of current methods. It focuses primarily on image and video-based DFs, ignoring text and audio manipulation, and it skips over a thorough examination of practical deployment issues and ethical issues (Malik et al. 2022).

The use of deep learning algorithms for detecting deepfakes in a variety of formats for media, such as audio, video, and photos, is investigated in this research. By analyzing differences in characteristics such as facial landmarks, audio-visual synchronization, and spectral patterns, the authors present a unified framework that uses convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to detect altered information. The study investigates the proposed approach on several datasets and found that it is highly accurate in recognizing deepfakes, especially in videos. The paper does have limitations, though: it mainly concentrates on high-quality deepfakes, ignoring compressed or low-quality data, which are more frequent in real-world situations. Furthermore, the computational complexity of the model may make it harder to scale for real-time or large-scale applications (Borrelli et al. 2021)

By utilizing the idea of "volume of differences," which uses geometric representations to measure the differences between altered and original media, the research provides an innovative approach to detecting deepfakes. The

authors develop a machine learning system that accurately detects small differences in elements, including textures, facial dynamics, and audio-visual synchronization across a range of media categories. Tests show encouraging outcomes, especially when it comes to detecting superior deepfakes that frequently elude conventional detection techniques. Nevertheless, the study has several weaknesses. First, it skips other media kinds, such as text or audio modifications, instead focusing mostly on video-based deepfakes. The proposed method's scalability for real-time applications is constrained by its dependence on high-dimensional feature spaces, which increases information technology expenses (Atlam et al. 2025).

2.1 Key Detection Approaches

According to surveys, many papers on detecting deepfake content have applied many methodologies. Machine Learning Algorithms: Neural networks can analyze video and audio for inconsistencies that may indicate manipulation. The latest detection techniques are: first, deep learning detection, such as Convolutional Neural Networks (CNNs). CNNs are widely used to analyze visual artifacts in deepfake videos, such as inconsistencies in facial features, lighting, or textures [9]. Second, Multimodal Detection Audio-Visual Analysis

Combines visual and audio cues to detect deepfakes by identifying mismatches between lip movements and spoken words. Third, Biological Signal Analysis: Detects deepfakes by analyzing biological signals, such as heartbeat or breathing patterns, which are often missing or inconsistent in synthetic media (Comission 2016).

Existing Deep Fake Detection Methods

1. Image Detection Techniques

Numerous techniques are employed to detect image authenticity; for instance, utilizing Saturation Indicators. In this study, the architecture of a well-known GAN implementation's generating network is examined, revealing that the network's handling of exposure significantly differs from that of a genuine camera. Furthermore, it demonstrates that this indicator can effectively differentiate between GAN-produced visuals and those captured by a camera, including the successful distinction between GAN visuals and the authentic camera images utilized for training the GAN(Honeywell ACST, Golden Valley, MN 2019).

This study examined and evaluated deepfake videos employing Neural Networks. The binary classification of deepfakes utilized a mix of Dense and Convolutional neural network layers. It was found that an accuracy of 91% was achieved on a sample of authentic images, while 88% was reached using SGD (stochastic gradient descent) for categorical cross-entropy methodology, which also resulted in a commendable level of accuracy(Badale, Castelino, and Gomes 2021).

2. Audio Detection Techniques

Considering a focus on heart rate measurements taken from facial films, this research study presents a new method to recognise deepfake videos utilising multidimensional biological signals. The authors provide an approach utilising colour space dimensions for transforming video frames into maps (photoplethysmography maps), which are later categorised by a convolutional neural network (CNN). The technique they employ checks at the RGB and YUV colour spaces to capture modifications to brightness and colour, which are impacted by blood flow and may show inconsistencies in phony audio files. The results of the experiment show that our technique improves current approaches, reaching 98% accuracy on publicly accessible datasets. In contrast to existing detection strategies, their solution shows significant flexibility to various deepfake categories and isn't tied to distinct generation techniques(Jin, Ye, and Chen 2021a).

In this study, a synthetic voice identifier was created. This system takes audio input, extracts a collection of manually designed features inspired by the speech-processing field, and categorizes them as either closed-set or open-set. The suggested identifier is tested on a publicly accessible dataset featuring 17 synthetic voice generation methods, ranging from traditional vocoders to contemporary deep learning techniques. Findings indicate that the suggested approach surpasses recently introduced identifiers in the forensics domain(Borrelli et al. 2021).

3. Video Detection Techniques

The research project presents an extensive review of the available research on recognizing and comprehending the dissemination of deepfakes on social media networks. 286 primary studies are analyzed, classified by study importance, contribution type, and methodology. Most of these studies (71%) concentrate on DF detection via the employment of AI and machine learning methodologies. The study emphasizes the need for creative research in unstudied areas like ethical implications, proactive, and societal effects, even as it shows substantial shortcomings in areas like digital treatment methods. The study has several drawbacks, despite its methodical approach and helpful assessment of trends, difficulties, and restrictions. It departs from research on audio/text-based manipulation, builds a heavy focus on detection techniques rather than mitigation or prevention methods, and provides no practical suggestions for filling in the gaps that have been identified (Atlam et al. 2025).

Legislative and Regulatory Frameworks

Governments around the globe recognize the difficulties introduced by deepfakes, with several taking proactive measures to address this evolving issue. For instance, the United States has made a significant initial move by presenting the DEEPFAKES Accountability Act at the federal level in September 2023. This legislation suggests three primary actions: firstly, it mandates deepfake material to feature a text box clarifying that it is AI-created, along with a digital watermark for identifying the source; secondly, it forbids the impersonation of individuals in ways that they wouldn't acknowledge; and thirdly, it offers victims pathways for legal action.

Various legislative and regulatory measures aimed at safeguarding individuals from harassment have been introduced at the state level across the United States. Numerous states, such as California, Hawaii, New York, and Virginia, have established laws that prosecute AI-generated non-consensual explicit content (Farid 2022). Moving beyond protective strategies to address deepfake misinformation and disinformation, Texas prohibited the creation of deepfakes intended to influence electoral results in 2019. Likewise, California now permits candidates for public office to initiate legal proceedings against “individuals or entities who produce or disseminate election-related deepfakes within 60 days before an election” (Department of Homeland Security 2021, 29)(Ayata 2024)

In 2019, the Chinese government enacted legislation requiring both individuals and organizations to reveal their use of deepfake technology in videos and other forms of media. These regulations further forbid the dissemination of deepfakes unless accompanied by a clear disclaimer indicating that the content has been artificially created. Additionally, as of January 10, 2023, China has implemented regulations for deepfake service providers, overseen by the Cyberspace Administration.

Administration of China (CAC).

The Rule on Algorithmic Recommendation of Internet Information Services (RARIIS), the Rule on Deepfake of Internet Information Services (RDIIS), and the Interim Rule on Generative Artificial Intelligence Services (IRGAIS) are the three main administrative rules that together make up China's legal framework for regulating AI. China's approach to AI regulation is reactive and sector-specific, relying on administrative rules issued by State Council departments, compared to the European Union's horizontal, inclusive approach. Regarding their more straightforward legislative procedure, these regulations enable a prompt reaction to the quick advancement of AI technologies (Hu and Liu 2024).

In Canada, like California, the Canada Elections Act contains language that may apply to deepfakes. Canada has also made other efforts in the past to curb the negative impacts of deepfakes, including its "plan to safeguard Canada's 2019 election" and the Critical Election Incident Public Protocol, a panel investigation process for deepfakes (Kashif et al. 2024).

The EU has taken an active position on deepfake laws, asking for rules that would enforce the obvious labelling of artificially generated content as well as additional studies into deepfake detection and prevention. It has proposed laws requiring social media companies to remove deepfakes and other disinformation from their platforms. Updated in June 2022, the EU's Code of Practice on Disinformation addresses deepfakes through fines of up to 6 percent of global revenue for violators. The Digital Services Act now supports the code, which was first put forward in 2018 as a voluntary self-regulatory tool. Since it came into force in November 2022, the Digital

Services Act has raised the supervision of digital platforms for numerous forms of misuse. The proposed EU AI Act would impose disclosure and transparency rules on deepfake suppliers.

In 2020, South Korea passed a law that makes it illegal to distribute deepfakes that could "cause harm to public interest," with offenders facing up to five years in prison or fines of up to 50 million won, or approximately 43,000 USD.

Table 1: Abbreviation for Papers That Used the Above Detection Methods

Paper	Deepfake detection method	Technique	Detection Results
Paper(Malanowska et al. 2024)	Image	convolutional neural networks (CNNs)	achieve an accuracy of up to 76%
Paper(Malanowska et al. 2024)	Image and Video	Watermarking	perfect for image authentication and tamper recovery
paper(Borrelli et al. 2021)	Audio	deep learning	Introducing identifiers in the forensics domain
paper(emozioni automatico 2021)	audio and video	Neural network	In comparison with video-based procedures, audio-based techniques are more accurate at detecting changed media.
paper(Ayata 2024)	Video	Motion magnification+ deep learning	97.77% and 94.03% accuracy in detecting video sources
paper(Badale, Castelino, and Gomes 2021)	Image and video	mix of Dense and Convolutional neural network layers	91% accuracy was achieved by the image.
paper(Jin, Ye, and Chen 2021a)	Video	multidimensional biological signals	98% accuracy on public datasets
Paper(Jung, Kim, and Kim 2020)	Video	Eye blinking inconsistency analysis	87.5% accuracy rate
Paper (Li et al. 2024)	Image	Frequency domain analysis using ResNet-50	85.72
Paper(Module et al. 2022)	Video	MesoInception4 architecture trained on real/fake face crops.	More than 88%.
Paper(Alsolai et al. 2025)	Image	Guardian-AI	97% accuracy rate

3. Methodology

This study adopts a systematic literature review approach, following the PRISMA guidelines for conducting reviews in scientific research.

A Systematic Literature Review (SLR) covering the period from 2018 to 2025 was conducted. The author documents the rapid advancement of deepfake-generating methods together with associated detection measures

over the last six years. The three main steps of the SLR process were planning the review, performing the review, and releasing the results.

3.1 Organizing the Review

During the planning stage, the goals, inclusion/exclusion criteria, appropriate sources, and research questions that would drive the study were all determined.

Objectives:

To categories and identify current deepfake detection techniques.

To evaluate these approaches' efficiency using the metrics and datasets that are available.

To investigate structures, difficulties, and possible strategies for deepfake detection research.

Questions for Research (RQs):

To direct our analysis, we identified five main research questions:

RQ1: Which deepfake detection methods are frequently used?

RQ2: How are deepfakes discovered through experimental testing?

RQ3: Which methods of classification are suited to deepfake detection techniques?

RQ4: Based on experimental data, how effective are different deepfake detection techniques overall?

3.2 **Search Strategy**

To ensure full coverage of the literature, we used a multi-database search method. Boolean search terms were employed for searching the following electronic repositories:
We searched multiple databases, including IEEE Xplore, ScienceDirect, and Google Scholar, using keywords such as "deepfake," "detection, and "regulation".

3.3 Inclusion Criteria

Peer-reviewed journal articles and conference papers published between 2018 and 2025.
Papers discussing technical detection methods or regulatory frameworks.
Studies focusing on deepfake generation and detection in audio, video, or text formats.

3.4 Data Extraction

For each included study, we extracted:
- Type of deepfake (audio/video/image)
- Detection method used
- Dataset size and accuracy
- Regulatory implications (if discussed)

Modern techniques for detecting deepfakes
Various modern deepfake detection techniques have recently been deployed. With a focus on the function of the Coalition for Content Provenance and Authenticity (C2PA) to develop international standards for media authentication, the study covers the latest technological advances, including AI-powered detection systems, provenance tracking, and digital watermarking. By using advanced algorithmic analysis, these approaches attempt to detect altered media and verify the authenticity and origins of digital content(Boháček and Farid 2022).

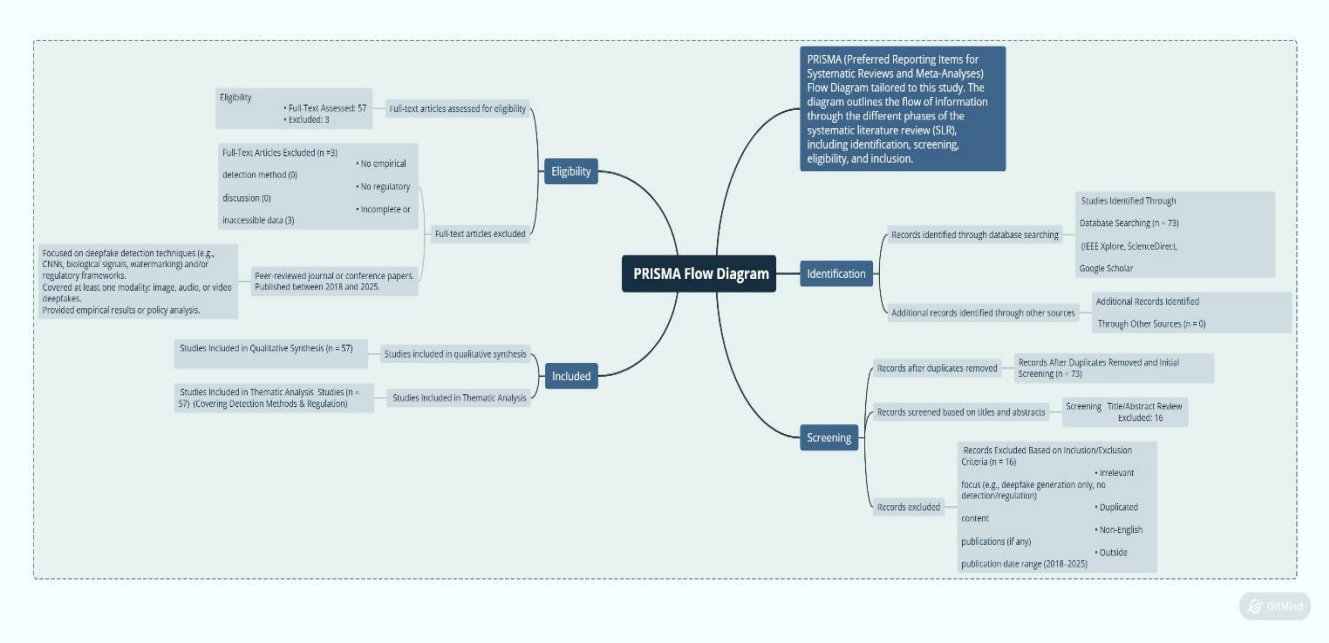

Figure 1: Prisma Flow Diagram

4. Result Discussion

The following is a summary of findings in response to the specified research questions (RQ1–RQ4), which rely on the systematic study of the literature presented in Addressing Deepfake Technologies Through Detection and Regulation: A Systematic Survey.

RQ1: Which deepfake detection methods are frequently used?

The most frequently used deepfake detection methods identified in the reviewed literature fall into four primary categories:

Four primary groups include the most employed deepfake detection techniques identified in the research literature:

Deep Learning-Based Detection (CNNs): The most common technique to identify visual inconsistencies in deepfake images and videos is by means of Convolutional Neural Networks (CNNs). These models investigate differences in lighting, shaping, face features, and spatial anomalies. MesoInception4 and ResNet-50-based models (like Preblended) are two examples; these kinds of models have been designed to detect facial changes.

Multimodal Analysis: This method takes advantage of the difficulty in coordinating both media in high-quality deepfakes by combining audio and visual components to detect inconsistencies, such as between speech and lip movements. To identify unusual expressions, multiple studies utilize emotion recognition.

Biological Signal Analysis: A highly effective and emerging method involves detecting physiological signals that are difficult to replicate in synthetic media, such as heart rate and breathing patterns. Techniques like remote photoplethysmography (rPPG) analyze subtle color changes in facial videos (in RGB/YUV color spaces) to identify inconsistencies in blood flow, achieving high accuracy (e.g., 98% in study (Jin, Ye, and Chen 2021b)).

Biological Signal Analysis: A novel yet highly successful technique seeks physiological signals, such as heart rate and breathing patterns, that are challenging to recreate in an artificial medium. For example, to recognize anomalies in blood flow, techniques like remote photoplethysmography (rPPG) assess small color changes in face videos (in RGB/YUV color spaces) with high precision (e.g., 98% in research (Jin, Ye, and Chen 2021b)).

More Notable Techniques:

Analysis of Eye Blinking Inconsistency: Older but fewer successful nowadays, deepfakes may replicate real blinking.

Frequency Domain Analysis: Approaches like FreqBlender show blending problems in GAN-generated images via frequency-level features.

Digital Watermarking and Provenance Tracking: To verify content authenticity in advance, organizations like the Coalition for Content Provenance and Authenticity (C2PA) propose the embedding of tamper-proof metadata or watermarks.

RQ2: How are deepfakes discovered through experimental testing?

Through a mix of controlled model training, benchmark datasets, and performance evaluation utilizing standardized criteria, deepfakes are experimentally found. Important elements consist of:

Datasets: For training and testing models, studies mostly employ publicly available deepfake datasets like Face Forensics++, Deepfake Detection Challenge (DFDC), and UADFV. These include videos that have been altered using methods like FaceSwap and Deepfakes.

Feature Extraction: Models are used to extract features, including spectral patterns, micro-textural artefacts, audio-visual sync, and facial landmarks.

Classification: Deep learning models (CNNs, RNNs) or conventional machine learning models (SVM, Logistic Regression) are frequently used for binary classification (actual vs. fake).

Evaluation Metrics: Performance is evaluated using cross-dataset generalization, accuracy, and AUROC (Area Under the Receiver Operating Characteristic). For example:

Study (Jin, Ye, and Chen 2021b) achieved 98% accuracy using biological signals.

Study (Ayata 2024) reported 97.77% accuracy using motion magnification and deep learning.

Study (Alsolai et al. 2025) (Guardian-AI) reached 97% accuracy in image-based detection.

However, a major limitation noted is overfitting to specific datasets, with poor performance when models are tested on data from different sources or under real-world conditions (e.g., compressed or low-quality videos).

RQ3: Which methods of classification are suited to deepfake detection techniques?

Deep learning-based models are among the most effective classification techniques, especially:

The best neural networks for spatial analysis of video and image frames are convolutional neural networks (CNNs).

Recurrent Neural Networks (RNNs): Effective in detecting unusual motion patterns in video sequences using temporal analysis.

Hybrid Models: By detecting both spatial and temporal anomalies, combining CNNs with RNNs (also known as dense layers) promotes detection.

Ensemble Methods: To enhance accuracy and resilience, some studies apply model ensembles.

When feature engineering is applied manually (e.g., entropy features in audio), traditional machine learning models like Support Vector Machines (SVM) and Logistic Regression are employed, but they are less effective on unprocessed data and demand a great deal of preprocessing.

Based on this review, deep learning models perform better than conventional techniques, especially when trained on huge, varied datasets. Their insufficient interpretability and computational complexity, however, continue to be problems.

RQ4: Based on experimental data, how effective are different deepfake detection techniques overall?

Detection techniques' performance depends heavily on their scalability, robustness, and practicality.

Table 2: Summary of Deepfake Techniques Performance

Technique	Summary of Performance
Multidimensional Biological Signals	★★★★★ – Since it depends on physiological signals (such as rPPG) that are not easy to falsify, it is quite successful. 98% accuracy was reached, and it demonstrates robust resilience to powerful GANs.
Motion Magnification + Deep Learning	★★★★☆ – Improves small movements (like blood flow); obtains high accuracy (97.77%) but is dependent on the quality of the camera and lighting.
Guardian-AI (Deep Learning)	★★★★☆ – 97% accuracy is reported.
Guardian-AI (Deep Learning)	★★★★☆ – 97% accuracy is reported.
Frequency Domain Analysis (ResNet-50)	★★★★ – 85.72% accuracy is obtained in detecting blending artefacts.
MesoInception4	★★★☆ – It's lightweight, productive, and has an accuracy of more than 88%.
Audio-Based Detection	★★★ –It is challenging to accurately replicate audio properties (like voice prosody and spectral traces); some models do better than video-based approaches.
Eye Blinking Analysis	★★☆ – Early approach; modern deepfakes that mimic blinking are simple to get around.
Digital Watermarking	★★☆ – Effective only if utilized prior to alteration; not applicable to unidentified or third-party content.

Since Deepfake is evolving quickly, it requires additional effort to solve its drawbacks. Advanced machine learning techniques, multimodal analysis, and innovative tools for addressing the increasing danger of synthetic media are some of the techniques used to detect deepfakes. such as

Machine Learning Models: Logistic Regression, SVM, Random Forest

Deep Learning Models: CNNs, RNNs, MesoNet

Multimodal Analysis: Combines visual and audio cues

Biological Signal Analysis: Detects inconsistencies in heartbeat or breathing patterns

Implementing Strict laws: to deter individuals, strict regulations that impose penalties and jail time on people who generate fake content, based on the scope of the harm. Prioritizing accountable

5. Conclusion

Deepfake technology presents both innovative possibilities and significant ethical dilemmas. This paper has reviewed current detection methods and regulatory efforts globally. An integrated framework combining technological solutions, legislation, education, and cooperation is essential to safeguard digital authenticity and trust. By fostering a collaborative approach among technologists, ethicists, lawmakers, and the public, we can mitigate the risks associated with deepfakes and promote a more trustworthy digital environment. Deepfakes can be addressed through suitable policies, regulations, personal initiatives, training, and educational efforts.

References

Alsolai, Hadeel, Khalid Mahmood, Asma Alshuhail, Achraf Ben Miled, Mohammed Alqahtani, Abdulrhman Alshareef, Fouad Shoie Alallah, and Bandar M. Alghamdi. 2025. 'Guardian-AI: A Novel Deep Learning Based Deepfake Detection Model in Images'. Alexandria Engineering Journal 126(April): 507–14. doi:10.1016/j.aej.2025.04.095.

Atlam, El Sayed, Malik Almaliki, Ghada Elmarhomy, Abdulqader M. Almars, Awatif M.A. Elsiddieg, and Rasha ElAgamy. 2025. 'SLM-DFS: A Systematic Literature Map of Deepfake Spread on Social Media'. Alexandria Engineering Journal 111(August 2024): 446–55. doi:10.1016/j.aej.2024.10.076.

'Audio-Video Deepfake Detection through Emotion Recognition Rilevamento Di Deepfake Audio-Video Tramite Riconoscimento Delle Emozioni Automatico'. 2021.

Ayata, Ozan. 2024. 'Artificial Realities : Mitigations against Deepfakes'.

Badale, Anuj, Lionel Castelino, and Joanne Gomes. 2021. 'Deep Fake Detection Using Neural Networks'. International Journal of Engineering Research & Technology (IJERT) 9(3): 349–54.

Ballesteros, Dora M., Yohanna Rodriguez, and Diego Renza. 2020. 'A Dataset of Histograms of Original and Fake Voice Recordings (H-Voice)'. Data in Brief 29: 105331. doi:10.1016/j.dib.2020.105331.

Boháček, Matyáš, and Hany Farid. 2022. 58 Computer Law & Security Review: The International Journal of Technology Law and Practice Protecting President Zelenskyy against Deep Fakes. Elsevier Ltd. doi:10.1016/j.clsr.2025.106162.

Borrelli, Clara, Paolo Bestagini, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro. 2021. 'Synthetic Speech Detection through Short-Term and Long-Term Prediction Traces'. Eurasip Journal on Information Security 2021(1). doi:10.1186/s13635-021-00116-3.

Comission, European. 2016. 'Detecting Handcrafted Facial Image Manipulations and GAN-Generated Facial Images Using Shallow-FakeFaceNet'. sciencedirect 4(1): 1–23. https://www.sciencedirect.com/science/article/abs/pii/S1568494621001794?via%3Dihub.

Honeywell ACST, Golden Valley, MN, USA. 2019. 'Https://Ieeexplore.Ieee.Org/Abstract/Document/8803661'. https://ieeexplore.ieee.org/Xplore/home.jsp. https://ieeexplore.ieee.org/abstract/document/8803661.

Hu, Qingle, and Wei Liu. 2024. The Regulation of Artificial Intelligence in China. Atlantis Press SARL. doi:10.2991/978-2-38476-259-0.

Jin, Xinlei, Dengpan Ye, and Chuanxi Chen. 2021a. 'Countering Spoof : Towards Detecting Deepfake with Multidimensional Biological Signals'. 2021(2). doi:10.1155/2021/6626974.

Jin, Xinlei, Dengpan Ye, and Chuanxi Chen. 2021b. 'Countering Spoof: Towards Detecting Deepfake with Multidimensional Biological Signals'. Security and Communication Networks 2021(2). doi:10.1155/2021/6626974.

Jung, Tackhyun, Sangwon Kim, and Keecheon Kim. 2020. 'DeepVision: Deepfakes Detection Using Human Eye Blinking Pattern'. IEEE Access 8: 83144–54. doi:10.1109/ACCESS.2020.2988660.

Kashif, Mohammad, Harshi Garg, Faizi Weqar, and Arokiaraj David. 2024. 'Regulatory Strategies and Innovative Solutions for Deepfake Technology'. Navigating the World of Deepfake Technology (July): 262–82. doi:10.4018/979-8-3693-5298-4.ch013.

Li, Hanzhe, Yuezun Li, Jiaran Zhou, Bin Li, and Junyu Dong. 2024. 'FreqBlender: Enhancing DeepFake Detection by Blending Frequency Knowledge'. (NeurIPS). http://arxiv.org/abs/2404.13872.

Malanowska, Agnieszka, Wojciech Mazurczyk, Senior Member, Minoru Kuribayashi, and Senior Member. 2024. 'Digital Watermarking — A Meta-Survey and Techniques for Fake News Detection'. 12(February). doi:10.1109/ACCESS.2024.3374201.

Malik, Asad, Minoru Kuribayashi, Sani M. Abdullahi, and Ahmad Neyaz Khan. 2022. 'DeepFake Detection for Human Face Images and Videos: A Survey'. IEEE Access 10: 18757–75. doi:10.1109/ACCESS.2022.3151186.

Module, Preprocessing, Zhiming Xia, Tong Qiao, Ming Xu, Xiaoshuai Wu, Li Han, and Yunzhi Chen. 2022. 'SS Symmetry Deepfake Video Detection Based on MesoNet With'. : 1–14.

Mukta, Md Saddam Hossain, Jubaer Ahmad, Mohaimenul Azam Khan Raiaan, Salekul Islam, Sami Azam, Mohammed Eunus Ali, and Mirjam Jonkman. 2023. 'An Investigation of the Effectiveness of Deepfake Models and Tools'. Journal of Sensor and Actuator Networks 12(4). doi:10.3390/jsan12040061.

Song, David. 2019. 'A Short History of Deepfakes'. Medium: 1–23. https://medium.com/search?q=A Short History of Deepfakes.