

DEEPFAKE DETECTION AS FORENSIC EVIDENCE: MINIMUM TECHNICAL STANDARDS AND ADMISSIBILITY TESTS IN COURT

AUTHOR – DHANU* & MR. SUGITH KUMAR**

* STUDENT AT SCHOOL OF EXCELLENCE IN LAW, THE TAMILNADU DR. AMBEDKAR LAW UNIVERSITY, CHENNAI

* PROFESSOR AT SCHOOL OF EXCELLENCE IN LAW, THE TAMILNADU DR. AMBEDKAR LAW UNIVERSITY, CHENNAI

BEST CITATION – DHANU & MR. SUGITH KUMAR, DEEPFAKE DETECTION AS FORENSIC EVIDENCE: MINIMUM TECHNICAL STANDARDS AND ADMISSIBILITY TESTS IN COURT, *INDIAN JOURNAL OF LEGAL REVIEW (IJLR)*, 6 (3) OF 2026, PG. 684-701, APIS – 3920 – 0001 & ISSN – 2583-2344.

ABSTRACT:

The rapid development of artificial intelligence has enabled the creation of highly realistic synthetic media, commonly known as deepfakes. These manipulated audio, video, and image files pose significant challenges to the integrity of digital evidence and the administration of justice. As deepfake technology becomes more sophisticated, courts increasingly face difficulties in determining the authenticity and reliability of digital content presented as evidence. This study examines the role of deepfake detection in digital forensics and proposes minimum technical standards for identifying manipulated media in legal proceedings.

The paper analyses the technological methods used in deepfake detection, including machine learning-based forensic tools, metadata analysis, artifact detection, and biometric inconsistencies. It emphasizes the importance of establishing standardized forensic procedures, including proper chain of custody, validation of detection tools, reproducibility of results, and expert verification. Without such standards, the risk of wrongful admission or rejection of digital evidence may undermine the fairness of trials.

In addition, the study explores the admissibility tests applied by courts when evaluating digital evidence. These include relevance, authenticity, reliability, and compliance with evidentiary rules governing electronic records. The paper discusses how existing legal frameworks for electronic evidence can be adapted to address deepfake-related challenges, highlighting the need for clear judicial guidelines and expert testimony in evaluating AI-generated content.

The possibility of fabricated visual or audio material being presented as genuine evidence raises significant challenges for courts, investigators, and forensic experts. Therefore, establishing reliable deepfake detection methods and defining minimum technical standards for forensic examination have become essential for ensuring the credibility of digital evidence.

This study examines the role of deepfake detection as forensic evidence and explores the minimum technical requirements necessary for identifying manipulated media. It discusses the use of advanced forensic tools, including artificial intelligence-based detection algorithms, metadata analysis, frame-level examination, and biometric inconsistencies, which help experts determine the authenticity of digital content. The research also highlights the importance of maintaining the chain of custody, proper documentation, and verification procedures during the forensic investigation process.

INTRODUCTION:

The increasing use of digital technology in modern society has significantly influenced the nature of evidence presented in courts of law. Audio recordings, video footage, photographs, and other electronic records are now frequently relied upon in criminal and civil proceedings. With the advancement of artificial intelligence, particularly in the field of deep learning, a new challenge has emerged in the form of deepfake technology. Deepfakes refer to highly realistic synthetic media created using artificial intelligence techniques that can manipulate or fabricate audio, video, and images to make it appear as though a person said or did something that never actually occurred.

Deepfake technology, often developed using techniques such as Generative Adversarial Networks (GANs) and other machine learning algorithms, has rapidly improved in quality and accessibility. While these technologies have legitimate uses in entertainment, education, and digital media production, they also pose serious risks when used maliciously. Deepfakes can be employed to spread misinformation, commit fraud, manipulate public opinion, or falsely implicate individuals in criminal activities. As a result, the authenticity of digital evidence is increasingly questioned in judicial proceedings.

In the field of digital forensics, ensuring the reliability and authenticity of electronic evidence has become a critical concern. Courts traditionally rely on forensic experts to verify the integrity of digital materials through established procedures such as metadata examination, error level analysis, and biometric consistency checks. However, the sophistication of deepfake technology has made traditional verification methods less effective, creating a need for more advanced detection techniques and standardized forensic protocols.

Another important challenge lies in the legal admissibility of such evidence. Courts must

determine whether digital content suspected of manipulation meets the required evidentiary standards, including authenticity, reliability, and relevance. In many jurisdictions, including India, the admissibility of electronic records is governed by provisions such as Indian Evidence Act, 1872, particularly Section 65B, which lays down the requirements for electronic evidence. However, the law does not yet provide specific guidelines for handling AI-generated or deepfake content, leaving courts to rely heavily on expert testimony and evolving forensic methods.

Therefore, the emergence of deepfakes has created an urgent need to develop minimum technical standards for deepfake detection and to establish clear admissibility tests for such evidence in court. This research aims to examine the forensic techniques used to detect deep fakes, analyse the legal challenges associated with their admissibility, and propose a framework that ensures both technological reliability and legal credibility. By integrating forensic science with evidentiary law, courts can better address the growing threat posed by manipulated digital media and safeguard the integrity of the justice system.

CONCEPT AND EVOLUTION OF DEEP FAKE TECHNOLOGY:

Deepfake technology has evolved rapidly with the advancement of artificial intelligence and machine learning techniques. Early forms of digital manipulation involved simple photo editing or video splicing, which could often be detected through visual inspection or basic forensic analysis. However, modern deepfake technology uses advanced neural networks capable of generating extremely realistic synthetic media.

One of the key technologies behind deepfake creation is Generative Adversarial Networks, which consists of two neural networks: a generator and a discriminator. The generator creates synthetic images or videos, while the discriminator evaluates their authenticity. Through repeated training, the generator

improves its ability to produce highly realistic outputs that are difficult to distinguish from genuine media.

The accessibility of deepfake software has also increased significantly. Open-source tools and mobile applications now allow individuals with limited technical expertise to create manipulated content. This widespread accessibility raises concerns regarding the misuse of deepfake technology for fraud, political manipulation, defamation, and the fabrication of digital evidence.

IMPACT OF DEEPAKES ON DIGITAL EVIDENCE AND CRIMINAL JUSTICE:

The justice system increasingly relies on digital evidence such as CCTV footage, mobile phone recordings, and social media content. Traditionally, such evidence was considered highly reliable because it captured events directly. However, the emergence of deepfakes has undermined this assumption.

Deepfake technology can be used to fabricate incriminating videos or alter legitimate recordings, thereby misleading investigators and courts. For example, a fabricated video may show a person participating in a criminal act, delivering threatening statements, or engaging in fraudulent activity. If such manipulated evidence is accepted without proper verification, it may lead to wrongful convictions or unfair legal outcomes.

Another major concern is the phenomenon known as the “liar’s dividend.” This occurs when individuals accused of wrongdoing claim that genuine video or audio evidence is actually a deepfake. As deepfake technology becomes more widely known, it may be used as a defense strategy to deny authentic evidence. This creates additional challenges for prosecutors, investigators, and judges in determining the truth.

ROLE OF DIGITAL FORENSICS IN DETECTING DEEPAKE MEDIA:

Digital forensics involves the scientific analysis of electronic data to identify, preserve, and

present evidence in legal proceedings. In the context of deepfakes, digital forensic experts play a crucial role in examining media files to determine whether they have been manipulated.

Forensic analysis may include examining pixel-level inconsistencies, compression artifacts, irregular lighting patterns, and abnormal facial movements. Experts may also compare the suspected deepfake with verified recordings of the individual involved. Advanced forensic software can detect subtle irregularities that are not visible to the human eye.

Digital forensic experts also analyse metadata associated with the file, including information about the device used to create the recording, editing history, timestamps, and file format. These details may reveal whether the media has been altered or artificially generated.

TECHNICAL METHODS USED IN DEEPAKE DETECTION:

Deepfake detection involves a combination of computational analysis and forensic expertise. Several technical approaches have been developed to identify manipulated media.

One method is biometric analysis, which studies facial expressions, blinking patterns, and lip movements. Deepfake videos sometimes produce unnatural facial behaviors or inconsistent synchronization between speech and lip movement. These irregularities can be detected through specialized algorithms.

Another approach is frequency domain analysis, which examines the mathematical patterns within images and videos. Deepfake generation often introduces subtle distortions in frequency patterns that can be detected through advanced analysis techniques.

Machine learning models are also trained to distinguish between authentic and synthetic media by analysing thousands of examples. These models learn to recognize specific artifacts associated with deepfake generation. However, as deepfake technology improves, detection tools must also be continuously

updated to remain effective.¹

MINIMUM TECHNICAL STANDARDS FOR FORENSIC EXAMINATION OF DEEPFAKES:

For deepfake detection to be accepted as reliable forensic evidence, certain minimum technical standards must be established. These standards ensure that forensic analysis is conducted using scientifically valid and transparent methods.

First, forensic tools used for deepfake detection should be tested and validated by independent research institutions. Validation ensures that the tools produce accurate and consistent results under different conditions. Second, forensic analysts must follow standardized procedures when examining digital evidence, including proper documentation of each step in the analysis process.

Third, the results of deepfake detection should be reproducible. This means that another qualified expert using the same methods should be able to obtain similar results. Reproducibility is essential for maintaining scientific credibility and ensuring fairness in judicial proceedings.

AUTHENTICATION AND INTEGRITY OF DIGITAL EVIDENCE:

Authentication of digital evidence involves proving that the evidence is genuine and has not been altered. In cases involving deepfake technology, authentication becomes particularly important because manipulated media can appear highly realistic.

Courts often require forensic experts to demonstrate the authenticity of digital media through technical analysis. This may include verifying the original source of the file, examining its metadata, and confirming that it has not been modified after its creation.

Maintaining the integrity of digital evidence also involves protecting it from tampering during investigation and storage. Digital evidence must be securely preserved using reliable storage systems and cryptographic verification techniques to ensure that its original content

remains unchanged.²

CHAIN OF CUSTODY IN DEEPFAKE-RELATED DIGITAL EVIDENCE:

The chain of custody refers to the chronological documentation of how evidence is collected, handled, stored, and transferred during an investigation. Proper chain of custody is essential for ensuring that digital evidence remains reliable and admissible in court.

In cases involving deepfakes, investigators must carefully document every stage of the evidence handling process. This includes recording when the evidence was collected, who handled it, the tools used for analysis, and where it was stored.

Any break or inconsistency in the chain of custody may raise doubts about the authenticity of the evidence. Therefore, strict documentation procedures must be followed to ensure transparency and credibility.

For deep fake evidence to be accepted in court, maintaining a proper chain of custody is essential. Chain of custody refers to the documented process that records the collection, handling, storage, and transfer of digital evidence from the time it is obtained until it is presented in court.

If the chain of custody is broken, the authenticity of the evidence can be questioned. Key requirements include:

Proper documentation of when and where the digital file was obtained. Identification of the person who collected and handled the evidence.

Secure storage of the digital media to prevent tampering.

Use of digital hash values (MD5 or SHA algorithms) to verify that the file has not been altered.

Maintaining a transparent chain of custody ensures that the evidence remains reliable and admissible.

EXPERT TESTIMONY AND FORENSIC ANALYSIS IN COURT:

Expert testimony plays a critical role in cases involving complex technological evidence such as deepfakes. Digital forensic experts help the court understand the technical processes used to analyse and detect manipulated media.

Experts may present detailed reports explaining the forensic techniques applied, the findings of their analysis, and the level of certainty regarding whether the media is authentic or manipulated. During cross-examination, the expert may also be required to explain the limitations of the detection methods used.³

The credibility of expert testimony depends on the expert's qualifications, experience, and the scientific reliability of the methods applied. Courts rely heavily on such testimony when evaluating technical evidence.

ADMISSIBILITY TESTS FOR DEEPFAKE EVIDENCE IN JUDICIAL PROCEEDINGS:

Before digital evidence is admitted in court, it must pass certain admissibility tests. These tests typically include relevance, authenticity, reliability, and compliance with procedural rules.

In cases involving deepfakes, courts must apply additional scrutiny to ensure that the evidence has not been manipulated. This may involve requiring forensic verification, expert testimony, and independent validation of the detection methods used.

The purpose of these admissibility tests is to prevent unreliable or misleading evidence from influencing judicial decisions. By applying strict evidentiary standards, courts can maintain the integrity of the legal process.

LEGAL FRAMEWORK GOVERNING ELECTRONIC EVIDENCE:

The legal framework governing electronic evidence varies across jurisdictions. In India, the admissibility of electronic records is primarily regulated by the Indian Evidence Act, 1872,

particularly Section 65B.

Section 65B requires that electronic evidence be accompanied by a certificate confirming the authenticity of the record and the proper functioning of the device used to produce it. This certificate must be issued by a person responsible for the operation of the device or the management of the relevant system.⁴

Although the law provides a mechanism for admitting electronic evidence, it does not specifically address the challenges posed by deepfake technology. As a result, courts often rely on forensic experts and evolving judicial interpretations to address such issues.

ROLE OF DIGITAL FORENSIC EXPERTS:

Deepfake detection generally requires the assistance of digital forensic experts. These experts analyse the media using specialized forensic software and scientific techniques. Their responsibilities include:

Conducting technical examination of the suspected deepfake media.

Identifying signs of manipulation such as inconsistent shadows, unnatural blinking patterns, or audio-video mismatch.

Using AI detection tools to identify synthetic content.

Preparing a forensic report explaining the methodology used and the conclusions reached.

Courts often rely on expert testimony to understand the complex technological aspects of deep fake detection.

DOCUMENTATION AND REPORTING STANDARDS:

Forensic laboratories must follow standardized documentation procedures when analysing deep fake evidence. The forensic report should clearly explain:

The source of the digital evidence.

The tools and software used for analysis.

The steps followed during the forensic

examination.

The results and interpretation of findings.

The report should be written in a manner that allows the court to understand the process and verify that accepted scientific methods were used.

CHALLENGES IN PROVING DEEFAKE EVIDENCE:

Despite advancements in forensic technology, several challenges remain in proving deepfake evidence in court:

Rapidly evolving AI technology – Deepfake generation tools are constantly improving, making detection more difficult.

Lack of universal technical standards – Different forensic laboratories may follow different methods.

Difficulty in identifying the creator of the deepfake – Even if manipulation is proven, tracing the origin of the content can be challenging.

Potential misuse of evidence – False claims of deepfakes may be used to deny genuine evidence.

These challenges highlight the need for clear legal and technical frameworks.

IMPORTANCE OF JUDICIAL AWARENESS:

Judges and legal practitioners must develop awareness of emerging technologies such as deepfakes. Courts need to understand the scientific principles behind digital manipulation and detection techniques in order to properly evaluate expert testimony.

Judicial training programs and collaboration with forensic experts can help ensure that courts are capable of handling cases involving advanced digital evidence.

UNDERSTANDING DEEFAKE TECHNOLOGY:

Deep fakes emerge from sophisticated artificial intelligence systems, specifically Generative Adversarial Networks. These neural network architectures comprise two components: a generator creating synthetic content and a

discriminator evaluating authenticity. Through iterative competition, the generator learns to produce increasingly realistic fabrications that deceive the discriminator, ultimately creating content indistinguishable from authentic recordings. Machine learning algorithms trained on extensive datasets enable these systems to replicate facial features, vocal patterns, and behavioral mannerisms with remarkable precision.

Contemporary deep fake methodologies encompass distinct techniques. Face-swap technology transplants one individual's facial characteristics onto another's body in video footage, creating false appearances of presence and participation. Lip-sync manipulation modifies mouth movements to synchronize with fabricated audio, suggesting statements never uttered. Voice synthesis employs neural networks to clone vocal characteristics, producing convincing audio impersonations. Puppet-master techniques enable real-time facial manipulation, allowing operators to control subjects' expressions instantaneously during live interactions. Real-world manifestations have materialized across jurisdictions. Courts have encountered fabricated confessions in criminal proceedings, manipulated recordings in matrimonial litigation, and synthetic evidence in contractual disputes. High-profile instances include fabricated political statements, non-consensual intimate imagery, and fraudulent corporate communications. Detection challenges are formidable: while forensic tools identify certain artifacts unnatural blinking patterns, lighting inconsistencies, pixel-level anomalies the technology evolves rapidly, consistently outpacing authentication capabilities and creating asymmetric technological competition wherein fabrication perpetually exceeds detection.

DEEFAKES AS DOCUMENTARY EVIDENCE UNDER THE INDIAN EVIDENCE ACT, 1872:

The threshold inquiry concerns whether deep fakes constitute admissible documentary

evidence. Section 3 defines evidence as statements permitted by the Court concerning factual matters under inquiry, made by witnesses, and documents including electronic records produced for inspection. This definition, predating digital technology by over a century, establishes broad parameters theoretically encompassing deep fakes yet lacking specificity for synthetic media.

Sections 65A and 65B, inserted through the Information Technology Act, 2000, govern electronic evidence admissibility. Section 65B mandates that electronic records are admissible only when accompanied by certificates identifying the record, describing production methodology, and specifying device particulars. Subsection (2) establishes conditions: computers must be regularly used for information storage, data must be fed during ordinary operations, computers must function properly, and records must originate from relevant devices. Subsection (4) permits evidence of improper operation or tampering.

These provisions reveal critical inadequacies when applied to deep fakes. Section 65B presumes computer-generated records' reliability absent tampering evidence a presumption reasonable for conventional digital files but problematic for AI-generated content. Sophisticated deep fakes may exhibit no discernible tampering because they constitute original synthetic creations rather than modified authentic recordings. Certificate requirements become meaningless when certifying authorities cannot distinguish genuine from fabricated content. The statutory framework addresses data storage and retrieval but not algorithmic content generation. Consequently, deep fakes may satisfy Section 65B's procedural requirements while being substantively fraudulent, exposing fundamental incompatibility between nineteenth-century evidentiary principles and twenty-first-century synthetic media technology.⁵

TESTS OF ADMISSIBILITY: RELEVANCY, AUTHENTICITY, AND RELIABILITY:

Deep fake evidence must satisfy fundamental admissibility criteria beyond statutory compliance: relevancy, authenticity, and reliability. Relevancy, governed by Sections 5

Through 55, requires logical connection between evidence and facts in issue. Deep fakes readily satisfy this threshold fabricated video depicting alleged criminal conduct is unquestionably relevant if authentic. The critical challenge lies in establishing authenticity and reliability.

Section 45 permits expert opinion on scientific or technical matters. Digital forensic experts theoretically authenticate electronic evidence by examining metadata, analysing compression artifacts, identifying pixel-level inconsistencies, and detecting temporal anomalies. However, deep fake authentication presents unprecedented difficulties. Traditional forensic markers timestamp discrepancies, editing software traces, file modification histories may be entirely absent in sophisticated deep fakes. Advanced systems generate synthetic content without modifying existing files, rendering conventional forensic methodologies ineffective. Sections 47 through 73 address authentication through handwriting comparison and signature verification, provisions designed for physical documents and inapplicable to digital fabrications. The Best Evidence Rule under Section 64 mandates original document production, but deep fakes challenge this concept fundamentally: what constitutes the 'original' of algorithmically generated content, the neural network output, training dataset, or algorithm itself?

Chain of custody requirements, though not explicitly codified but judicially recognized as essential, prove inadequate. These requirements ensure evidence remains unaltered from seizure to production. Deep fakes subvert this principle because tampering occurs at creation rather than transmission.

Properly maintained custody demonstrates only that the deep fake remained unchanged during investigation, not that it depicts genuine events.

The burden of proof distribution becomes problematic. Section 101 places evidentiary burdens on parties asserting facts. When prosecution presents video evidence, accused parties must prove fabrication, often insurmountable given technical complexity and resource disparities. Section 114's presumption of regularity permits courts to presume official acts were properly performed and electronic records correctly generated. Applied to deep fakes, this presumption dangerously privileges technological sophistication over factual accuracy.⁶

JUDICIAL APPROACH: CASE LAW ANALYSIS:

Indian jurisprudence on electronic evidence has evolved considerably, though specific deep fake adjudication remains nascent. The Supreme Court's landmark decision in *Anvar P.V.v.*

P.K. Basheer (2014) established that electronic records require Section 65B certificates for admissibility, rejecting oral evidence as substitute authentication. This ruling emphasized procedural compliance, mandating certificates identifying electronic records, describing production methods, and specifying device particulars. The Court recognized electronic evidence's unique characteristics requiring special authentication protocols beyond traditional witness testimony.

Subsequently, *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal (2020)* modified this stringent approach, holding Section 65B certificates unnecessary when original documents are produced and authenticated through oral evidence. The Court acknowledged practical difficulties obtaining certificates from third parties like email providers or social media platforms. While pragmatically addressing genuine obstacles, this relaxation inadvertently creates deep fake vulnerabilities by reducing authentication barriers, potentially

facilitating fabricated evidence admission.

In *Shafhi Mohammad v. State of Himachal Pradesh (2018)*, the Supreme Court permitted electronic evidence admission without Section 65B certificates, reasoning that the Evidence Act's objective is preventing technical exclusions of relevant evidence. This liberal interpretation, though commendable for avoiding procedural injustice, overlooks that authentication requirements exist precisely to ensure reliability considerations paramount with deep fakes.

Lower courts have addressed manipulated digital evidence inconsistently. Some judges mandate forensic examination, directing original device production and appointing expert committees. Others uncritically accept electronic evidence based solely on witness testimony. This inconsistency reflects broader challenges: lack of judicial technical literacy, absence of standardized authentication protocols, and inadequate forensic infrastructure.⁷

International precedents offer instructive perspectives. United States courts apply Federal Rules of Evidence requiring authentication through evidence sufficient to support findings that items are what proponents claim. The Daubert standard governing expert testimony mandates judicial gatekeeping ensuring scientific validity and reliability. European jurisdictions emphasize independent expert verification for contested digital evidence. These frameworks suggest India's relatively permissive electronic evidence standards may prove inadequate for deep fake challenges, requiring more rigorous authentication protocols and enhanced judicial scrutiny.

EVIDENTIARY CHALLENGES POSED BY DEEPPAKES:

Section 81A establishes presumptions regarding electronic records: courts shall presume electronic signatures were affixed with authenticating intent and electronic messages

were transmitted by purported senders. While primarily addressing digital signatures, this provision reflects broader judicial inclination toward presuming electronic records' authenticity unless challenged. With deep fakes, this presumption becomes vulnerability, placing unrealistic burdens on parties contesting fabricated evidence.

Section 65B (4) permits evidence of improper computer operation or tampering, theoretically enabling deep fake challenges. However, establishing tampering requires technical expertise and forensic resources often unavailable to litigants. Sophisticated deep fakes may exhibit no tampering evidence because they constitute original synthetic creations rather than modified authentic files. Proving negative propositions that content is fabricated rather than genuine presents inherent evidentiary difficulties. Cross-examination limitations compound these challenges. The adversarial system assumes parties can effectively challenge witnesses and evidence. With deep fakes, this assumption fails. How does one cross-examine algorithms? How can counsel without technical expertise challenge forensic testimony about neural networks or Generative Adversarial Networks? Technical complexity creates information asymmetries favoring technologically sophisticated parties. Judicial understanding presents fundamental obstacles. Most judges lack technical training in artificial intelligence, machine learning, or digital forensics.

Courts cannot meaningfully evaluate deep fake evidence without this literacy. Technical complexity barriers extend beyond judges to advocates and even expert witnesses, creating systemic incapacity to assess synthetic media authenticity.

Reverse burden problems merit particular concern. When prosecution presents video evidence, accused parties bear fabrication-proof burdens often insurmountable given resource constraints. Well-resourced parties can afford sophisticated forensic analysis;

individual litigants cannot. This creates systemic inequity wherein justice becomes contingent on technological access rather than factual truth, fundamentally undermining the presumption of innocence in criminal proceedings.

FORENSIC SOLUTIONS AND DETECTION MECHANISMS:

The technological arms race between deep fake creation and detection has produced increasingly sophisticated authentication tools. Forensic analysis examines multiple dimensions: pixel-level inconsistencies, lighting angle discrepancies, unnatural eye movements, biological implausibility in facial expressions, and audio-visual synchronization anomalies. Advanced techniques employ neural networks specifically trained to detect AI-generated content, paradoxically using artificial intelligence to identify artificial intelligence.

Metadata analysis examines file creation timestamps, modification histories, device fingerprints, and GPS coordinates, potentially revealing inconsistencies. However, metadata can be falsified or removed, and sophisticated deep fakes generate plausible metadata. Block chain technology offers promising authentication through immutable creation records and verifiable custody chains. While news organizations increasingly implement block chain-based authentication, widespread adoption remains distant and legal recognition uncertain.

Expert witnesses under Section 45 play crucial roles but face inherent limitations. Experts identify forensic markers suggesting fabrication but rarely provide absolute certainty. Deep fake detection's probabilistic nature '73 percent likelihood of fabrication' conflicts with legal standards requiring proof beyond reasonable doubt or preponderance of evidence. Courts accustomed to definitive expert opinions struggle with forensic uncertainty, sometimes rejecting appropriately cautious testimony as

insufficiently conclusive, creating disconnect between scientific methodology and judicial expectations.

COMPARATIVE LEGAL FRAMEWORK:

International jurisdictions employ varying strategies addressing deepfake evidence. The United States Federal Rules of Evidence, particularly Rule 901 concerning authentication and Rule 702 regarding expert testimony, provide more rigorous frameworks than Indian law. Rule 901 explicitly requires authentication through evidence sufficient to support findings that items are what proponents claim, placing clear burdens on introducing parties. The Daubert standard governing expert testimony mandates judicial gatekeeping ensuring scientific testimony's relevance and reliability based on methodological validity.

The European Union emphasizes algorithmic transparency through GDPR and proposed AI regulations, mandating disclosure of AI system operations. Some European jurisdictions have implemented specific legislation criminalizing malicious deepfake creation. These comparative frameworks suggest India's permissive electronic evidence approach requires substantial enhancement, necessitating mandatory expert evaluation for contested evidence, standardized forensic methodologies, and enhanced judicial gatekeeping for scientific testimony to adequately address deep fake challenges.

PROPOSED REFORMS AND RECOMMENDATIONS:

Comprehensive legislative reform is imperative. Section 65B requires amendment explicitly addressing AI-generated content, mandating heightened authentication for audio-visual evidence. Proposed amendments should require forensic examination by accredited laboratories, disclosure of technical provenance including creation software and device specifications, and mandatory expert affidavits addressing synthetic generation possibilities. These enhanced requirements would establish authentication standards commensurate with

deepfakes sophistication.

Establishing a specialized Digital Evidence Authentication Authority would address institutional capacity limitations. This body, comprising technical experts and forensic specialists, could provide court-appointed neutral opinions, develop standardized protocols, maintain accredited laboratory networks, and conduct judicial training. Rather than relying on party-appointed experts with inherent bias, courts could reference this authority for objective assessment.

Procedural reforms should include reverse onus provisions: when forensic examination suggests probable fabrication, evidentiary burdens should shift to introducing parties to prove authenticity. Courts should be empowered to exclude electronic evidence when forensic uncertainty exceeds defined thresholds. Judicial education represents critical necessity technical training programs should familiarize judges with artificial intelligence fundamentals and forensic authentication methods.

Substantive criminal law should explicitly criminalize malicious deepfake creation and dissemination. While existing forgery and defamation provisions theoretically encompass deepfakes, specific legislation would provide clarity and appropriate sentencing frameworks reflecting the technology's unique dangers. Legislation must balance criminalizing harmful conduct while protecting legitimate uses in entertainment and artistic expression through carefully crafted intent requirements and exceptions.

AI IN CRIMINAL INVESTIGATIONS:

Law enforcement bodies in India have already included AI tools in their practices. The Delhi Police has started using facial recognition technology to identify suspects during protests. The National Crime Records Bureau has shown inclination towards predictive policing systems. Forensic labs have used AI for handwriting and image analysis.

While these tools prove to be advantageous in investigations, but their admissibility still remains unsettled in court. Use of technology in limited context has been permitted by Indian courts, such as, video conferencing testimony in the case of State of Maharashtra v. Dr. Praful Desai. Courts have admitted CCTV footage as evidence where proper certification exists. But there is still no precedent directly addressing AI-generated evidence.

CONSTITUTIONAL CONCERNS:

Two safeguards have been provided in the Constitution. First, Article 20(3) gives protection against self-incrimination. In *Selvi v. State of Karnataka*, compulsory narco-analysis and brain mapping was prohibited, stressing dignity and autonomy. Similar issues may be raised due to AI-based lie detection or behavioral prediction. Second, Article 21 gives guarantee of right to fair trial. It has been repeatedly held by the Courts that unreliable evidences violate due process of law.

If AI evidence is admitted without proper scrutiny, it will result in undermining this guarantee. Thus, the present framework is inadequate. The Evidence Act includes electronic records but does not cover AI related concerns. Law enforcement agencies are adopting AI tools without any statutory oversight. Constitutional rights require careful safeguards. Therefore, there is an urgent need of reform.

THE INADEQUACY OF SECTION 65B FOR AI EVIDENCE:

Section 65B of the Indian Evidence Act is concerned with the authenticity of electronic records, particularly emails and computer printouts, and documentation stored on DVDs, CDs, and similar electronic devices. The requirement of a certificate under Section 65B is indicative of a pre-AI world where human authorship was assumed. The advent of AI-generated evidence, however, raises many issues that the drafters of Section 65B never contemplated. Unlike digital documents,

outputs of AI may be probabilistic rather than definite, shaped by blurred and opaque training data, or there can be possibility of biasness.

This mismatch poses significant challenges for courts because they are relying on an outdated framework to address regionally novel challenges. This tension is evident in Indian case law. For example, in *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal*, the Supreme Court affirmed that Section 65B certificates are mandatory for the admissibility of electronic documents, but offered no guidance for when evidence does not have any human author. The reluctance of the Delhi High Court to admit CCTV footage in *State v. Navjot Sandhu*, could not be extrapolated to cover the type of analytics that had AI generated outputs, which include facial recognition technologies or predictive policing algorithms.

Comparative perspectives sharpen the critique. In the US, the Daubert standard requires scientific evidence to have been tested, peer-reviewed, and generally accepted, before admissibility. The U.K. has created a relatively systemic approach by placing special emphasis regarding the responsibility of judges to screen expert evidence for reliability. The European Union has proposed substantial transparency and accountability requirements for high-risk systems through its draft AI Act.

These jurisdictions have room for improvement but at least they are iterating evidentiary rules in order to apply to the complexity of algorithms. Meanwhile, in India we have not yet moved the needle to recalibrate Section 65B.⁸

RECOMMENDATIONS FOR A PROPOSED SECTION 65C:

A practical solution would be to create a new section called Section 65C, relating specifically to AI-generated evidence. This new section could build on but also move beyond Section 65B by incorporating the following three conditions:

Disclosure obligations: If a party wants to rely on AI-generated evidence they would need to

disclose the system's architecture, training data (to the extent feasible), error rates, and results of validation studies.

Judicial gatekeeping: Judges would assess admissibility with a reliability assessment (aligned with a Daubert-style consideration) rather than solely be an assessment of certification.

Judicial bans and presumptions: Some AI systems should be banned based on using "black box" systems such as predictive policing, while by contrast current validated forensic systems would be able to rely on a rebuttable presumption of reliability. Such framework prevents blind reliance on algorithmic outputs and gives courts structured tools to assess admissibility.

Definitions (to be inserted in Section 3)

"Artificial Intelligence-generated evidence" means any information, conclusion, or output produced wholly or partly by a system using machine learning, neural networks, or algorithmic decision-making.

"Algorithmic system" includes any automated system designed to process data and generate conclusions, predictions, or classifications without continuous human oversight.

AMENDMENT TO SECTION 45 (EXPERT OPINION):

After "science or art," insert: "or by an artificial intelligence system validated in accordance with prescribed standards."

Insert a new Explanation: "For the purposes of this section, the court may admit outputs of algorithmic systems as expert opinion if accompanied by certification of reliability, validation studies, and error rates as prescribed."

AMENDMENT TO SECTION 65B (ELECTRONIC RECORDS):

Substitute sub-section (4): "An electronic record, including AI-generated evidence, shall be admissible if—

(a) It is accompanied by a certificate from the

person responsible for the operation or maintenance of the device or system generating or storing such evidence, specifying the process by which it was produced;

(b) in the case of AI systems, the certificate shall also include— (i) information on validation tests, accuracy rates, and known biases; (ii) a statement on whether the system complies with Technical standards notified by the Central Government; and (iii) a declaration of human oversight in the generation of the evidence."⁹

New Section 65C (AI Evidence Oversight) "65C."

(1) The Central Government may, by notification, designate a regulatory Indian Journal of Integrated Research in Law authority to prescribe standards for the admissibility of AI-generated evidence.

(2) The authority shall publish guidelines on validation, certification, and disclosure of AI systems used in forensic or evidentiary contexts.

(3) Courts shall have discretion to exclude AI evidence where the system's reliability, transparency, or compliance with due process cannot be reasonably assured."

AUTHORITATIVE REPORTS AND TECHNICAL GUIDELINES:

1. Law Commission of India – Report No. 185 (2003)²¹ a) Comprehensive review of the Indian Evidence Act, 1872, including detailed analysis of Section 65B on electronic record.

b) Recommended retention of certification requirements to safeguard authenticity and reliability.

c) Advocated expert assistance in technologically complex matters – a principle echoed in the Bharatiya Sakshya Adhinyam, 2023 (BSA). d) Core themes of authenticity, reliability, and procedural safeguards remain directly applicable to BSA interpretation.

2. Ministry of Electronics and Information Technology (MeitY) –

Digital Forensic Methodologies Guidelines

- a) Provides SOPs for the identification, acquisition, preservation, analysis, and documentation of electronic evidence.
- b) Stresses maintaining chain of custody, generation of hash values, and admissibility-ready record-keeping.
- c) Though non-binding, frequently relied upon in investigative and prosecutorial contexts as a benchmark for forensic integrity.

Key Issues

1. Statutory Silence on Deep fakes and AI-Generated Outputs

- a) Neither the BSA nor the IT Act defines synthetic media or specifically regulates AI-created evidence.
- b) Courts currently address such material via analogy to forgery or falsification offences, coupled with expert testimony.
- c) This gap generates interpretive uncertainty in admissibility and weight of evidence.

2. Certification Challenges for Decentralized/Foreign Cloud Sources

- a) BSA-compliant certification often unattainable when data resides on decentralized or foreign-controlled servers.
- b) Overseas service providers may fall outside Indian legal process, producing evidentiary lacunae.
- c) Persistent problem noted in case law, indicating a need for legislative or treaty-based intervention.

ADMISSIBILITY IN THE AGE OF DEEPFAKES, CLOUD FORENSICS, AND AI EVIDENCE:

The contemporary evidentiary framework faces unprecedented stress from the simultaneous rise of synthetic media, globally distributed cloud storage, and autonomously generated content. The Bharatiya Sakshya Adhiniyam (“BSA”) provides a technologically neutral definition of “document” and “electronic record” but lacks explicit provisions addressing these emergent formats, compelling courts to

interpret existing rules purposively to uphold procedural fairness and accuracy in fact-finding.

DEEP FAKE CHALLENGES:

Despite the socio-legal urgency, neither the BSA nor the ITA articulates an explicit statutory definition of “deep fake” or “synthetic media”. Current reliance is on general offences of forgery, fraud, and misrepresentation under the BNS (ss 336–341) and the ITA’s penal provisions (s 66D–66E). From a procedural justice perspective, the absence of definitional clarity risks inconsistent admissibility determinations, particularly when defence counsel challenges the reliability threshold under s136 BSA.

Authentication demands sophisticated forensic examination. Expert testimony now frequently involves AI forensic analysis— detecting generative adversarial network (GAN) artefacts, inconsistencies in eye-blinking patterns, micro-expression anomalies, and irregular compression signatures. Such evidence must meet both the relevance requirement (s 4 BSA) and the balancing test for unfair prejudice, with the judge’s gatekeeping role ensuring that fabricated yet technically impressive exhibits do not mislead the trier of fact.

CLOUD FORENSICS:

The increasing use of foreign-hosted cloud services for personal, corporate, and governmental data storage generates profound jurisdictional complexity. Section 75 ITA provides for extra-Territorial application, yet in practice, compelling disclosure from non-treaty jurisdictions often depends on de facto negotiation or diplomatic channels, rather than enforceable judicial writs. From an evidentiary chain-of-custody standpoint, cloud-stored data challenges the procedural requirement under s63 (2) BSA and s 65B ITA to demonstrate integrity from the moment of collection. Issues include multi-node replication, elastic storage re-allocation, and automated retention-policy deletions. Maintaining procedural fairness therefore requires forensically sound

acquisition—imaging data in situ, securing contemporaneous hash validations, and documenting each transfer or extraction step in an audit-compliant log.¹⁰

AI-GENERATED EVIDENCE:

AI-produced evidence occupies a liminal legal space between expert-assisted demonstratives and autonomous fabrications. The former such as algorithmic accident reconstructions or predictive crime-mapping outputs prepared under judicial or investigative mandate may assist fact-finding if their underlying models are transparent, datasets are validated, and methodology satisfies scientific reliability tests.

By contrast, manipulative falsifications intentionally distort the factual record. Here, the “liar’s dividend” phenomenon compounds the challenge: litigants may claim genuine footage or recordings are “AI-generated” to manufacture reasonable doubt. Procedural justice demands that courts develop provenance verification standards disclosing metadata lineage, AI model specifications, and, where possible, the cryptographic signing of outputs at source to distinguish lawful reconstructions from deceptive forgeries.

FORENSIC TOOLS AND METHODOLOGIES:

Judicial assessment of admissibility in such cases increasingly leverages technical integrity tools, including:

- Cryptographic hashing – SHA-256 as the preferred algorithm, with MD5 used only alongside corroborative checks given its vulnerability to collisions;
- Metadata analysis (e.g., EXIF timestamps, camera serial identifiers, and modification logs) to corroborate asserted timelines;
- Block chain-based evidence sealing—anchoring a hash of the artefact in a distributed ledger to provide immutable proof of unaltered state from the moment of sealing;
- AI-assisted detection suites employing multi-modal analytics (visual, acoustic, and

linguistic channels) to identify patterns characteristic of synthetic generation. From a statutory interpretation viewpoint, the BSA’s technologically neutral language accommodates such tools, but expert admissibility remains subject to the court’s satisfaction that the methodology has demonstrable reliability and that the audit trail is unbroken.

DOCTRINAL AND POLICY CONSIDERATIONS:

The doctrinal trajectory points toward purposive construction of existing evidentiary provisions, adapting them to technological realities without undermining the principles of natural justice. Legislative intervention may still be warranted—either to define emerging categories such as deepfakes or to codify minimum technical standards for admissibility. Policy measures could include:

- Judicial capacity-building through specialist training in digital forensics;
- Establishment of nationally accredited forensic laboratories capable of verifying AI-linked evidence;
- Integration of provenance tracking protocols into law-enforcement evidence-handling SOPs;
- Strengthening cross-border mutual legal assistance frameworks for timely access to cloud-stored data. Such steps would align evidentiary law with both the right to a fair trial and the constitutional mandate for procedural due process, ensuring that technological sophistication enhances rather than erodes the justice system.

ADMISSIBILITY PROCEDURE IN THE USA FOR SECONDARY ELECTRONIC EVIDENCE:

The United States employs a flexible, process-driven evidentiary framework for the authentication and admission of secondary electronic records, principally governed by the Federal Rules of Evidence (“FRE”). The FRE’s technologically neutral drafting permits adaptation to evolving data formats, including cloud-stored logs, AI-generated files, and

social- media-based material, without necessitating statutory overhaul.

RELEVANT CASE LAWS:

1. Anvar P.V. v. P.K. Basheer (2014) Citation: (2014) 10 SCC 473

Facts:

The case involved the use of electronic evidence such as CDs and digital records in an election dispute.

Judgment:

The Supreme Court held that electronic records are admissible only if they comply with the requirements of Section 65B of the Indian Evidence Act, 1872. A 65B certificate is mandatory for admissibility of electronic evidence.

Principle Established:

Electronic evidence must be authenticated.

A certificate under Section 65B (4) is required.

Oral evidence cannot replace the statutory requirement. Relevance to Deepfake Detection:

Deepfake videos or audio recordings presented in court must also satisfy Section 65B certification to prove authenticity.

2. Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal (2020) Citation: (2020) 7 SCC 1

Facts:

The case concerned the admissibility of electronic evidence in an election dispute.

Judgment:

The Supreme Court reaffirmed that Section 65B certificate is mandatory unless the original electronic device is produced in court.

Key Observations:

Electronic records must satisfy statutory requirements.

Failure to produce a 65B certificate makes the evidence inadmissible. Relevance to Deepfake Evidence:

This judgment is extremely important for

deepfake cases because courts must verify the authenticity and source of digital media before admitting it as evidence.

3. Tomaso Bruno v. State of Uttar Pradesh (2015) Citation: (2015) 7 SCC 178

Facts:

The case involved a murder trial where CCTV footage was not produced by the prosecution.

Judgment:

The Supreme Court emphasized the importance of electronic evidence such as CCTV footage in modern criminal trials.

Key Principle:

Electronic evidence plays a vital role in criminal investigations. Courts should not ignore digital evidence when it is available. Relevance to Deepfake Detection:

The judgment highlights the growing reliance on digital evidence, which also increases the risk of deepfake manipulation.

4. State (NCT of Delhi) v. Navjot Sandhu (2005) Citation: (2005) 11 SCC 600

Facts:

The case involved the 2001 Parliament attack where telephone records and electronic evidence were used during the trial.

Judgment:

The Supreme Court initially allowed electronic evidence even without a Section 65B certificate.

Legal Impact:

This decision was later overruled by the Anvar judgment, which made Section 65B certification mandatory.

Relevance to Deep fake Evidence:

This case illustrates the evolution of judicial thinking regarding electronic evidence and authentication.

5. Suhas Kati v. State of Tamil Nadu (2004)

Facts:

The accused posted obscene and defamatory messages online using a fake identity.
Judgment:

The court accepted electronic records obtained from internet servers as evidence and convicted the accused under cybercrime laws.

GOVERNMENT SOURCES:

1. Ministry of Electronics and Information Technology (MeitY)

The Government of India, through the Ministry of Electronics and Information Technology (MeitY), has been actively addressing the issue of deepfake technology. The ministry has funded research projects aimed at detecting fake speech, videos, and manipulated images. These projects include the development of forensic software tools capable of identifying deep fake media and verifying digital authenticity.

Hindustan Times

MeitY has also formed expert committees involving government agencies, technology companies, and academic institutions to examine the risks associated with AI-generated content and propose regulatory frameworks for deep fake detection.

Hindustan Times

2. Ministry of Electronics and Information Technology – Status Report on Deep fakes

In a status report submitted before the Delhi High Court, the ministry highlighted the growing misuse of deep fake technology for misinformation, election manipulation, and financial scams. The report also pointed out that the absence of a clear legal definition of deep fake creates challenges in regulating synthetic media.

India TV News

The report further emphasized the need for:

Standardized detection mechanisms

Development of indigenous datasets for deepfake detection
Public awareness initiatives

Cooperation between law enforcement agencies and technology experts. India TV News

3. Ministry of Electronics and Information Technology – AI and Deep fake Regulation

The Government of India has proposed amendments to the Information Technology Rules requiring mandatory labelling of AI-generated or synthetic media. These rules aim to improve transparency and prevent the misuse of deep fake videos, audio, and images on digital platforms.

India Today

Under the proposed regulations:

Platforms must clearly label AI-generated content

Social media companies must verify user declarations about synthetic media
Metadata or markers should identify AI-generated information.

India Today

4. IndiaAI Mission

Under the IndiaAI Mission, the government has launched several research projects aimed at strengthening deep fake detection technologies and improving AI governance frameworks. These initiatives involve collaboration between universities such as IIT Madras and IIT Jodhpur to develop advanced systems capable of detecting manipulated media.

The Economic Times

The initiative also focuses on:

Building safer AI systems improving detection accuracy

Creating national datasets for identifying synthetic media. The Economic Times

Example Reference Format (for research paper)

Government of India, Ministry of Electronics and Information Technology, Status Report on Deep fake Technology, submitted before the Delhi High Court, 2025.

Government of India, Ministry of Electronics and Information Technology, Information Technology Rules on AI-Generated Content and Deep fake Regulation, 2025.

Government of India, IndiaAI Mission, Safe and Trusted Artificial Intelligence Initiative, 2024.

CONCLUSION:

Deep fake technology has emerged as one of the most significant challenges to the reliability and integrity of digital evidence in the modern criminal justice system. With the rapid advancement of artificial intelligence and machine learning tools capable of creating highly realistic manipulated audio, video, and images, courts are increasingly faced with the difficult task of determining the authenticity of digital content presented as evidence. Therefore, establishing minimum technical standards for deep fake detection and ensuring clear admissibility tests in courts has become essential for maintaining fairness in judicial proceedings.

From a forensic perspective, deep fake detection requires the application of scientifically validated techniques such as AI-based detection algorithms, metadata analysis, pixel-level forensic examination, facial movement analysis, and block chain-based verification systems. These technical standards must ensure reliability, reproducibility, and transparency in the forensic process. Digital forensic laboratories must follow standardized protocols, maintain proper chain of custody, and use validated tools so that the findings can withstand judicial scrutiny. Proper documentation and expert testimony are also necessary to explain the technical findings to the court.

In the legal context, courts rely on established evidentiary principles to determine whether digital evidence is admissible. In India, provisions such as Section 65B of the Indian Evidence Act, 1872 play a crucial role in governing the admissibility of electronic records. Courts also assess factors such as

authenticity, reliability, relevance, and the integrity of the digital evidence before admitting it into trial. Judicial precedents like *Anvar P.V. v. P.K. Basheer* and *Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal* emphasize the necessity of proper certification and procedural compliance when presenting electronic evidence.

Moreover, the use of expert opinion is recognized under Section 45 of the Indian Evidence Act, 1872, which allows forensic experts to assist courts in understanding complex scientific and technical issues such as deep fake detection. Courts must evaluate whether the expert's methodology is scientifically reliable and widely accepted within the forensic community.

Ultimately, addressing deep fake threats requires a multidisciplinary approach involving law, forensic science, and technological innovation. Governments, legal institutions, and forensic agencies must collaborate to develop uniform technical guidelines, strengthen digital forensic infrastructure, and train judicial officers in understanding emerging digital manipulation techniques. International cooperation and regulatory frameworks may also be necessary to deal with cross-border digital manipulation and cybercrime.

In conclusion, while deep fake technology poses serious risks to evidentiary integrity and public trust, the implementation of robust forensic detection standards, strict evidentiary rules, and informed judicial evaluation can ensure that manipulated media is properly identified and excluded from misleading the court. Strengthening both the technical and legal mechanisms for evaluating digital evidence will play a vital role in safeguarding the credibility of the justice system in the digital age.

REFERENCES:

Books:

Avtar Singh, *Principles of the Law of Evidence*, Central Law Publications, Allahabad. Vepa P. Sarathi, *Law of Evidence*, Eastern Book

Company, Lucknow.

Eoghan Casey, Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet, Academic Press.

Journal Articles:

Chesney, R., & Citron, D. (2019). "Deep fakes and the New Disinformation War." Foreign Affairs Journal.

Agarwal, S., Farid, H., Gu, Y., He, M., Nagano, K., & Li, H. (2020). "Protecting World Leaders against Deep fakes." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Verdoliva, L. (2020). "Media Forensics and Deep Fake Detection." IEEE Signal Processing Magazine.

Statutes:

Indian Evidence Act, 1872 Information Technology Act, 2000

Case Laws:

Anvar P.V. v. P.K. Basheer, (2014) 10 SCC 473.

Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 SCC 1. Tomaso Bruno v. State of Uttar Pradesh, (2015) 7 SCC 178.

State (NCT of Delhi) v. Navjot Sandhu, (2005) 11 SCC 600.

Reports and Online Sources:

National Institute of Standards and Technology, Deep fake Detection Challenge Report. Europol, Malicious Uses and Abuses of Artificial Intelligence Report.

MIT Media Lab, Research on deep fake detection and synthetic media.

ENDNOTES

1 Avtar Singh, Principles of the Law of Evidence, Central Law Publications, Allahabad.

2 Vepa P. Sarathi, Law of Evidence, Eastern Book Company, Lucknow.

3 Avtar Singh, Principles of the Law of Evidence, Central Law Publications, Allahabad.

4 Indian Evidence Act, 1872

5 Indian Evidence Act, 1872

6 Indian Evidence Act, 1872

7 Vepa P. Sarathi, Law of Evidence, Eastern Book Company, Lucknow.

8 Chesney, R., & Citron, D. (2019). "Deepfakes and the New Disinformation War." Foreign Affairs Journal.

9 Information Technology Act, 2000

10 Chesney, R., & Citron, D. (2019). "Deepfakes and the New Disinformation War." Foreign Affairs Journal.