YourLawArticle

Open Access Law Journal, ISSN (O): 3049-0057

Editor-in-Chief – Prof. (Dr.) Amit Kashyap; Publisher – Reet Parihar

The Rise of Deepfakes: Navigating Legal Challenges and Ethical Dilemmas

Diksha Sharma, B.B.A.LL.B, Gitarattan International Business School

Published on: 02th May 2025

Abstract

Deepfakes—hyper-realistic synthetic media generated using artificial intelligence—pose a growing threat to individual privacy, public trust, and societal stability. Originally developed as tools for entertainment, deepfakes have evolved into powerful instruments of deception, capable of manipulating audio, video, and images with striking accuracy. This paper explores the alarming rise of deepfakes in India and abroad, highlighting cases of fraud, defamation, non-consensual pornography, and political misinformation. The analysis reveals that India's current legal framework, including provisions under the Indian Penal Code (IPC) and the Information Technology Act, 2000, is inadequate to fully address the complexity of deepfake-related crimes.

Sections such as 499, 500, 354A, and 354D of the IPC and 66C, 66D, 67, and 72 of the IT Act offer fragmented remedies, focusing on defamation, harassment, and privacy violations. However, these laws lack specificity, speed, and technological adaptability. The anonymity of online offenders, high burden of digital proof, and absence of admissibility guidelines for synthetic media compound the legal blind spots. In contrast, international jurisdictions like California, the European Union, and China have already enacted laws targeting deepfakes, offering India a policy roadmap.

The paper underscores the urgent need for dedicated legislation that defines deepfakes, criminalizes malicious intent, mandates swift takedowns, and establishes evidentiary standards. Without proactive regulation, deepfakes risk eroding public trust in media, damaging reputations, and undermining democratic processes. The document concludes by advocating for a future-ready legal response that matches the pace of technological advancement and protects citizens in a digitally manipulated world.

***Keywords**: Deepfake, Information Technology Act, Indian Penal Code, Digital Privacy, Cyber Law*

Introduction

Consider a hypothetical yet plausible scenario: A viral video circulates on social media, appearing to show the Indian Prime Minister announcing a full-scale war against China. The video is highly realistic, showing troops mobilising and diplomatic relations collapsing. Within minutes, financial markets crash and panic erupts on a massive scale. This has not happened, which is a relief. But a doctored video that triggers concrete consequences in the real world is becoming more workable.A single deepfake video posted on social media like this can trigger panic among people and an economic breakdown before the truth catches up. In today's world, where visual proof is authoritative, the distinction between what is real and what is not is not merely blurred but is being self-consciously rewritten.

Understanding Deepfake

Deepfake is an artificial intelligence that manipulates original content to make something fictional or fabricated appear true. These deep fake algorithms alter voice, face, and movements, creating an uncanny resemblance. Deepfakes are created using machine learning algorithms that produce strikingly realistic videos, images, and audio. This replication looks so crisp that it is hard to distinguish it from the fake. In the modern era of mass propagation and digital virality, these deepfake videos are widespread, reaching millions almost instantly without critical thinking. Deepfake videos, which are often shared under the tag of harmless entertainment, have triggered widespread concern across the globe.

Recent Incidents in India and Abroad

What started as an amusing internet trend has spiralled into a weapon of deception. Originally a method of digital entertainment, deepfakes have developed into instruments of deception, fraud, and reputation damage at an incredible pace. Some recent cases show the increased use of this technology. Earlier in December 2023, Ratan Tata, who served as the chairman of the Tata Group, unmasked a deepfake video of him giving investment advice. The manipulated clip falsely depicted him offering to invest in a project supposedly launched by him, claiming it to be 'risk-free'. Another notable case in early 2024, a morphed video of Indian actress Rashmika Mandanna that went viral on a social media platform, where her face was digitally superimposed onto footage originally featuring

another girl, depicting her obscene manner.[1]

The viral video was lifelike and plausible to such an extent that Mandanna experienced online criticism until it was declared a synthetic video. In a male-dominated society like India, deepfake technology has been weaponised to target and portray women in obscene and degrading ways. The use of deepfake technology isn't only confined to targeting celebrities online; it has also expanded into areas like phone scams, where artificial voices are crafted to mimic familiar individuals, leading to a sophisticated form of fraud.[2]

Cases of phone scams using deepfake technology have been on the rise, Scammers are now using deepfake technology to deceive people and even bypass security checks, allowing them access to sensitive information to commit crimes. No matter how educated or informed a person may be, the hyper-realistic nature of deepfakes makes almost anyone susceptible to such deception.

In a real instance in Hong Kong, a finance worker at a multinational firm was tricked into paying $25 million to a scammer using deepfake technology who impersonated the company's chief financial officer in a video conference call.[3]

Advancements in technology were meant to build a better world, but apparently, they also built a better scam.

Types of Deepfake Videos:

Deepfakes take various forms intended to manipulate perception and distort reality. These are not merely technological tricks but powerful tools capable of impersonation, fraud, and character destruction. To grasp the full extent of the danger, it is essential to understand the different ways in which deepfake content can be created and disseminated:

1**. Face-Swapping Deepfakes**

Perhaps the most widely recognised, face swap deepfakes involve digitally replacing one person's face with another's in a video. The result is an eerily convincing clip that shows someone appearing to do or say things they never actually did.

- Real-world risk: A public figure could be inserted into a controversial situation or a fake endorsement clip, instantly damaging credibility and public image.
- Deceptive strength: The accuracy of facial expressions, lighting, and movement makes these

[1] 'Ratan Tata Slams Deepfake Video That Features Him Giving "Risk-Free Investment" Advice' *The Economic Times* (27 November 2023) https://economictimes.indiatimes.com/magazines/panache/ratan-tata-slams-deepfake-video-that-features-him-giving-risk-free-investment-advice/articleshow/105805223.cms accessed 1 May 2025.

[2] 'Rashmika Mandanna Deepfake Video News' *NDTV* https://www.ndtv.com/topic/rashmika-mandanna-deepfake-video accessed 30 April 2025.

[3] 'How This Multinational Company Lost $25 Million to a Deepfake Conference Call' *The Times of India* (9 February 2024) https://timesofindia.indiatimes.com/gadgets-news/how-this-multinational-company-lost-25-million-to-a-deepfake-conference-call/articleshow/107430242.cms accessed 1 May 2025.

videos almost indistinguishable from genuine footage.

2. **Lip-Sync Deepfakes**

These involve manipulating the mouth movements of a person in a video so that their lips match the audio of a completely different dialogue. When paired with cloned audio, the result is a hyper-realistic clip of a person speaking words they never uttered.

- Example: Imagine a video of a world leader announcing a policy shift—or a war—that was never discussed in reality.
- Impact: It takes just a few seconds for such a video to go viral and create mass hysteria before the truth is revealed.

3. Audio Deepfakes (Voice Cloning)

Audio deepfakes use artificial intelligence to mimic a person's voice with remarkable accuracy. By analysing only a few seconds of genuine speech, AI models can recreate the voice's tone, pitch, and accent. These cloned voices are then utilised in phone calls, podcasts, or videos to make it sound as if someone said something they never actually did.

- Example: In 2023, cybercriminals used an AI-generated voice to impersonate the CEO of a UK-based energy firm. They called a senior executive and convinced him to transfer over €200,000, claiming it was for an urgent business deal. The voice was so accurate that the employee didn't question the command.[4]
- Implication: When a familiar voice speaks with authority, even trained professionals can be fooled, making voice deepfakes a powerful tool for financial and psychological manipulation.

4. **Motion-Capture Deepfakes**

These advanced deepfakes don't just mimic a person's face or voice—they replicate the entire body, gestures, posture, and expressions. Through sophisticated motion capture technology, a completely synthetic version of a person can be generated to perform fabricated actions on camera.

- Example: A fabricated video of a company CEO making a public statement that tanks stock prices within minutes.
- Danger level: When visual realism meets manipulated messaging, trust in video evidence itself is challenged.

5. **Non-Consensual Explicit Deepfakes**

This deeply disturbing category targets individuals, especially women, by inserting their likeness into sexually explicit or obscene content without consent. These videos are often shared maliciously,

[4] Jesse Damiani, 'A Voice Deepfake Was Used to Scam a CEO Out of $243,000' *Forbes* (3 September 2019) https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/ accessed 1 May 2025.

sometimes as a tool of revenge, blackmail, or public humiliation.

- Recent incident: The viral deepfake video of actress Rashmika Mandanna, where her face was overlaid onto an obscene clip, stirred nationwide outrage.
- Long-term damage: Victims suffer irreparable emotional, social, and professional harm, often with little legal recourse or digital remedy.

Despite their varied formats, all deepfakes share one common thread: the power to deceive at scale. In an era where visual and audio evidence was once the gold standard of truth, deepfakes are rewriting that rule, turning digital proof into a potential weapon.

The unfiltered side of Deepfake –

In today's digital age, where smartphones are in the hands of nearly everyone, from curious teenagers to tech-savvy grandparents, the line between reality and illusion is growing dangerously thin. With the rise of deepfake technology, the phrase "seeing is believing" no longer holds true. Seeing is not believing anymore. What once required sophisticated equipment and expert skills can now be done with frightening ease. Creating a deepfake video or image often takes nothing more than a couple of clicks and an internet connection—tools so accessible that even a 13-year-old can generate realistic manipulations. This unfiltered side of deepfakes, where the barrier to entry is virtually nonexistent, has profound consequences. If left unregulated, it threatens to upend the very foundations of trust, truth, and evidence. This isn't just a future concern, it's already unfolding.[5]

A recent ruling by the Delhi High Court underscored the growing judicial awareness of this threat. In a case involving allegations of adultery supported by photographic evidence, the court refused to accept the images at face value, emphasising that in the era of deepfakes, such content must be substantiated with reliable corroboration (Economic Times, 2024)[6]. As the line between fabricated and factual continues to blur, the unchecked spread of deepfakes could have devastating effects—not only on individual lives and reputations but on the legal system, journalism, and democracy itself. The urgency to act is no longer optional—it is imperative.

Legal Framework

Existing Indian Laws Related to Deepfakes -

Considering the devastating power of deepfakes, one would assume that laws to control them would

[5] James Cote, 'Deepfake and Fake News Pose a Growing Threat to Democracy, Experts Warn' *Northeastern University News* (1 April 2022) https://news.northeastern.edu/2022/04/01/deepfakes-fake-news-threat-democracy/ accessed 1 May 2025.

[6] 'In the Deepfake Era, Photos Claiming Adultery by Spouse to be Proved by Evidence: Delhi HC' *The Economic Times* (24 April 2024) https://economictimes.indiatimes.com/news/india/in-deepfake-era-photos-claiming-adultery-by-spouse-to-be-proved-by-evidence-delhi-hc/articleshow/110823310.cms accessed 1 May 2025.

be strict. But the truth is, India still has no effective legal measures to tackle the problem.

Disappointingly, India currently lacks a dedicated law specifically targeting deepfakes. However, there are provisions under the Information Technology Act, 2000 (IT Act) and the Indian Penal Code, 1860 (IPC) that can be applied to cases involving deepfakes, such as:

1) **Information Technology Act, 2000 (IT Act)**

- **Section 66C**

Under Section 66 (C)[7] of the Information Technology Act, 2000, any person who dishonestly or fraudulently uses another's electronic signature, password, or any other unique identification characteristic shall be punishable with imprisonment for a term of up to three years and a fine of up to one lakh rupees.

For deepfakes, one can use an individual's face or voice to generate a forged video or audio, pretending to be the person, tarnishing his or her reputation, or harming him or her. This section can be used in case of a deepfake being created using an individual's identity or image without his or her consent.

- **Section 67**

Section 67[8] of the Information Technology Act, 2000, prohibits individuals from publishing, transmitting, or making the publication in electronic form of any material that is labelled as lascivious, solicits prurient interests, or has the tendency to deprave and corrupt those likely to be exposed to it through reading, viewing, or listening.

On the first conviction, the person will be sentenced to imprisonment for a term, which may extend to three years, and a fine, which may extend to ₹5 lakh.

For the second consecutive conviction, the punishment is imprisonment for a term up to five years and a fine up to ₹10 lakh.

Deepfakes containing obscene content or sexually explicit content (often called "revenge porn" or other objectionable content) could fall under this provision. When someone creates and circulates a deepfake containing content that is not appropriate or explicit, it is a criminal offence under Section 67.[9]

- **Section 66D**

Section 66D[10] penalises any person who utilises any communication tool or computer facility to

[7] *The Information Technology Act 2000*, s 66C.
[8] *The Information Technology Act 2000*, s 67.
[9] Alexandra Canizales, 'Texas Court of Criminal Appeals Upholds "Revenge Porn" Law' *The Texas Tribune* (26 May 2021) https://www.texastribune.org/2021/05/26/texas-court-revenge-porn-law/ accessed 1 May 2025.
[10] *The Information Technology Act 2000*, s 66D.

deceive another person by impersonating another.

Any person convicted under Section 66D can face imprisonment for a term which may extend up to three years and a fine which may extend up to ₹1 lakh.

Deepfakes can be used as a form of personation, where a person's face, voice, or other characteristics are overlaid onto altered content. It can lead to instances of fraud, defamation, or impersonation, depending on whether the deepfake is used to cause a different individual to believe something untrue and suffer loss or harm as a consequence of impersonation.

- **Section 72**

It pertains to the invasion of privacy and confidentiality. It states that whoever has gained access to any electronic record, book, register, correspondence, information, document, or other matter without the permission of the concerned individual and makes it available to another individual shall be punished with imprisonment or a fine.

If a deepfake is a misuse of a person's data (e.g., intimate video capture, personal images, or voice recordings) and is the foundation of an imitation or false video, this would be a violation of privacy under Section 72[11]. Even if the content is not openly offensive, creating and distributing a deepfake using personal data without permission would be a breach of the person's privacy.

2) Indian Penal Code, 1860 (IPC)

- **Sections 499 & 500**

Whoever, orally or in writing, or by signs, or by visible representations, publishes any statement about a person with the intent to harm, or knowing or having reason to believe that the statement will harm, the reputation of such person, is said to defame such person.

Section 500[12] provides the punishment for defamation under Section 499[13], which is simple imprisonment for a term not exceeding 2 years, or fine, or both. A deepfake that defames an individual in a derogatory manner (e.g., putting them in a compromising or defamatory position) can be an offence under Section 499. A deepfake to disseminate defamatory and false news about an individual can harm their reputation, causing hurt in their personal, professional, or social life.

- **Section 354A & 354D**

Section 354A[14] covers sexual harassment and penalises acts like making sexually colored remarks, showing pornography with or without consent, or uninvited bodily advances, with imprisonment for three years or a fine.

If a deepfake is created involving explicit or sexually suggestive content of an individual,

[11] *The Information Technology Act 2000*, s 72.
[12] *The Indian Penal Code 1860*, s 500.
[13] *The Indian Penal Code 1860*, s 499.
[14] *The Indian Penal Code 1860*, s 354A.

particularly a woman, it could fall under Section 354A if it is done with the intent to harass, humiliate, or degrade the victim. Creating a sexually explicit deepfake and circulating it without consent could be seen as sexual harassment under this section, leading to criminal liability.

Section 354D[15] covers stalking, including following a woman physically or virtually, monitoring her online, or regular contact without consent, with a maximum imprisonment of three years for the first conviction and a maximum of five years for second or subsequent convictions.

If a person creates or posts deepfakes of another individual regularly, particularly in an attempt to harass, intimidate, or threaten them, this would be stalking under Section 354D. For instance, regular posting of fake videos or pictures that result in distress, fear, or emotional agony to the individual may amount to stalking, particularly if the victim perceives a sense of danger or harassment from the web posting.

The Legal Blind Spots for Deepfakes: Why Existing Laws Fall Short in Addressing Deepfakes

Looking at the intensity of this crime, these laws still fail to tackle the problem and address the harm caused by it, they do not fully capture the complex and evolving nature of deepfake technology.

Here are some of the key reasons why these provisions may not be enough to address the deepfake problem comprehensively:

1) **Lack of Specificity –**

The above-mentioned sections focus on specific types of offenses (defamation, sexual harassment, stalking) but do not specifically address the unique characteristics of deepfakes. The current legal framework suffers from ambiguity and vagueness, rendering it ineffective against the highly sophisticated and evolving threat of deepfakes.

2) Anonymity and Pseudonymity –

The anonymity of the internet and the ease with which deepfakes can be made without revealing themselves make it difficult to trace the culprits. Provisions such as Section 354D (stalking) or Section 499 (defamation) are based on the identification of the perpetrator, but with deepfakes, the perpetrator who made or shared the fake can remain anonymous using pseudonyms, fake identities, or VPNs, and it is difficult to trace and prosecute them.

3) Exemptions in existing law -

Deepfakes can be claimed as freedom of expression or satire in some cases, especially if they are not offensive. Section 66A of the Information Technology Act criminalizes offensive content online. It is

[15] *The Indian Penal Code 1860*, s 354D.

important to note that Section 66A of the Information Technology Act, which once criminalized offensive online content, was declared unconstitutional and struck down by the Supreme Court in Shreya Singhal v. Union of India (2015), and therefore can no longer be invoked.

4) **Burden of Proof in Digital Forgeries**

In cases involving deepfakes, one of the most significant challenges is the burden of proof. Establishing the authenticity of digital content and demonstrating its forgery requires specialized expertise, particularly in fields like digital forensics and AI analysis. However, Indian law does not currently provide clear guidelines on how to handle such evidence in courts. Traditional methods of proof, such as witness testimony or physical evidence, are insufficient in addressing the unique challenges of synthetic media. The burden on the victim to prove that content is fake, especially in the face of highly convincing deepfakes, places an unjust strain on individuals, making it even more difficult to seek legal recourse.

5) The Need for Admissibility Rules for Synthetic Media

Another significant gap in the legal framework is the lack of standardized rules for the admissibility of synthetic media as evidence in court. Deepfakes are often used as tools for defamation, fraud, and blackmail, and their impact can be devastating. However, there are no clear standards on how synthetic media should be handled in legal proceedings. Courts may struggle with determining whether to admit deepfake videos, audios, or images as evidence, as current rules do not address the specificities of digitally altered content. The absence of such rules makes it difficult for the justice system to respond to deepfake-related crimes efficiently and effectively.

6) **Lack of Rapid Takedown Mechanisms**

Deepfake videos and audios spread virally on social media, necessitating rapid response mechanisms for harmful content removal. However, India's legal framework lacks a swift process for victims to address deepfake content, allowing harmful material to linger for days or weeks, further damaging victims' reputations and causing financial loss. Implementing a fast-track removal mechanism is essential to mitigate the impact of deepfakes. Without defined frameworks for online platforms to act, victims remain at risk of ongoing harm.

7) The Constitutional Perspective

The Right to Privacy under Article 21, as incorporated in the case of K.S. Puttaswamy v. Union of India[16], is severely infringed by deepfakes, particularly in non-consensual explicit content. However, no constitutional right or law directly bestows authority upon individuals to address this type of cyber abuse.

[16] Justice KS Puttaswamy (Retd) v Union of India (2017) 10 SCC 1.

Comparative Legal Perspectives

The United States: California and Texas both passed laws criminalizing deepfake pornography and election tampering.[17]

European Union: The EU's Digital Services Act will hold platforms responsible for AI-generated content.

China: Enacted a law requiring watermarks and disclosure labels on synthetic media.

Conversely, **India is far behind in creating a systematic legal and regulatory framework.**

Need for new laws

In a rapidly digitizing world, where synthetic media can mimic reality with unnerving precision, India's legal apparatus relies on statutes drafted in an analog age. Deepfakes—AI-generated audio-visual forgeries—pose a novel threat far beyond defamation or obscenity. They manipulate not just images or voices, but the very foundation of trust in digital content.

The existing patchwork of laws, such as Sections 66C, 66D, and 67 of the Information Technology Act, along with provisions from the Indian Penal Code, offers only fragmented remedies. These sections may address identity theft, obscenity, or fraud, but they fall short of capturing the unique nature of deepfakes, where deception is often hyper-realistic, anonymous, and instantaneous.

The absence of a clear legal definition of "deepfake," coupled with outdated evidentiary standards, leaves both victims and law enforcement unequipped. Proving intent, establishing harm, and tracing anonymous perpetrators in cases of AI-generated content are legal challenges that current frameworks do not anticipate. Moreover, in a country where women are disproportionately targeted through sexually explicit deepfakes, the lack of focused legislation further silences victims by leaving their grievances in legal limbo.

A dedicated deepfake law must go beyond punishing the act—it should define the offense, establish mechanisms for swift takedown, recognize psychological harm, and impose liability on platforms that allow synthetic content to thrive unchecked. As technology evolves, so too must the law. India cannot afford to legislate yesterday's internet while living in tomorrow's reality.

Conclusion: The Need for Legislative Urgency

Deepfakes are no longer an emerging concern—they represent an immediate and tangible danger. In an era where manipulated visuals and cloned voices can mimic reality with frightening accuracy, the

[17] Office of the Governor of California, 'Governor Newsom Signs Bills to Combat Deepfake Election Content' (17 September 2024) https://www.gov.ca.gov/2024/09/17/governor-newsom-signs-bills-to-combat-deepfake-election-content/ accessed 1 May 2025.

age-old principle of "seeing is believing" has become dangerously unreliable. As this sophisticated form of digital deception continues to evolve, India cannot afford to remain tethered to outdated and fragmented legal provisions.

The urgency for a dedicated, well-defined, and technologically responsive legal framework cannot be overstated. Without timely legislative intervention, the legal vacuum surrounding deepfakes will only deepen, leaving citizens vulnerable, law enforcement incapacitated, and justice delayed. The law must not merely catch up with technology—it must anticipate its misuse, mitigate its dangers, and empower victims with swift redress mechanisms.

As synthetic media accelerates at the speed of innovation, our legal system must transition from a state of reactive patchwork to a forward-thinking, preventative architecture. The clock is ticking, and each moment of inaction increases the risk of irreversible damage. India must step into this new digital frontier with laws as advanced as the threats it aims to combat.

References

1. Alexandra Canizales, Texas Court of Criminal Appeals Upholds "Revenge Porn" Law, The Texas Tribune (May 26, 2021), https://www.texastribune.org/2021/05/26/texas-court-revenge-porn-law/.

2.James Cote, Deepfake and Fake News Pose a Growing Threat to Democracy, Experts Warn, Northeastern Univ. News (Apr. 1, 2022), https://news.northeastern.edu/2022/04/01/deepfakes-fake-news-threat-democracy/.

3. Jesse Damiani, A Voice Deepfake Was Used to Scam a CEO Out of $243,000, Forbes (Sept. 3, 2019), https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/.

4. In the deepfake era, photos claiming adultery by spouse to be proved by evidence: Delhi HC, The Economic Times (Apr. 24, 2024), https://economictimes.indiatimes.com/news/india/in-deepfake-era-photos-claiming-adultery-by-spouse-to-be-proved-by-evidence-delhi-hc/articleshow/110823310.cms.

5. Press Release, Office of the Governor of California, Governor Newsom Signs Bills to Combat Deepfake Election Content (Sept. 17, 2024), https://www.gov.ca.gov/2024/09/17/governor-newsom-signs-bills-to-combat-deepfake-election-content/.

6. Supreme Court Observer, Fundamental Right to Privacy – Justice K.S. Puttaswamy v. Union of India (Aug. 24, 2017), https://www.scobserver.in/cases/fundamental-right-to-privacy/.

7. Rashmika Mandanna Deepfake Video News, NDTV, https://www.ndtv.com/topic/rashmika-mandanna-deepfake-video (last visited Apr. 30, 2025).

8. Ratan Tata Slams Deepfake Video That Features Him Giving 'Risk-Free Investment' Advice, The Economic Times (Nov. 27, 2023), https://economictimes.indiatimes.com/magazines/panache/ratan-tata-slams-deepfake-video-that-features-him-giving-risk-free-investment-advice/articleshow/105805223.cms.

9. How This Multinational Company Lost $25 Million to a Deepfake Conference Call, The Times of India (Feb. 9, 2024), https://timesofindia.indiatimes.com/gadgets-news/how-this-multinational-company-lost-25-million-to-a-deepfake-conference-call/articleshow/107430242.cms