Page 2001 - 2019 **DOI:** *https://doij.org/10.10000/IJLMH.1111693*

INTERNATIONAL JOURNAL OF LAW MANAGEMENT & HUMANITIES

[ISSN 2581-5369]

Volume 9 | Issue 2

2026

Deepfake Manipulation and Criminal Responsibility: A Cyber Analysis of Creation and Sharing of Deepfakes

TANU YADAV[1]

ABSTRACT

Deepfake technology is an artificial intelligence (AI) and machine learning (ML) technology used to create hyper-realistic social imagery, videos, and audio records. This technology will be used to create materials that sound and appear to be produced by an actual human being through its use of face-swapping algorithms and voice-cloning methods. Although deepfakes were created as an entertainment and creative tool, they have also been used more often in a more fraudulent and unlawful way, such as identity theft, non-consent pornography, and spreading misinformation. These innovations have been a cause of concern in the sphere of cyber law.

As a paper that examines criminal responsibility, it discusses the creation and distribution of deepfake content through the lens of the current laws on cyber law. Along with the continually growing sophistication of deepfake technology and its accessibility, this threat is growing to the personal privacy, honour, and credibility, becoming a powerful tool of blackmail, defamation, and reputation damage, which is not inherently limited to prominent individuals but also to ordinary citizens.

The research takes three important dimensions. First, it discusses the legal ambiguity on the questions of jurisdiction, as the deepfake technology is reported to work well across the borders of countries but the information technology laws do not provide substantial measures to determine the cross-border cyber disputes. Second, it assesses the sufficiency of the current laws, which do not introduce a clear definition of deepfakes and a full-scale regulatory system. Third, it investigates the necessity of controlling the platforms, software that allow creating deepfakes, and not only the content that is produced, and suggests possible legislative enhancements.

Keywords: *Deep fake, Cyber Law, Technological Creativity, Cyber Security, Jurisdiction.*

I. INTRODUCTION

The word deepfake is a combination of two terms, the first of them is the notion of deep learning, which is a subdivision of machine learning, and the second is the word fake, which

[1] Author is a Student at Lloyd Law College, Greater Noida, Uttar Pradesh, India.

 [ISSN 2581-5369]

carries the meaning of being created by an AI. Deepfakes are artificially created media that are created in virtual space using artificial intelligence. They represent the results of computational creativity, and not documented reality.

The technology works on the principles of two competing algorithms which work simultaneously. The former algorithm creates a fake copy of the original image or video, whereas the latter algorithm compares the result against how similar one can detect it to be to the original content. Once some similarities are detected, feedback is sent to the initial algorithm whereby the algorithm is used to refine its output until the synthetic content is no longer distinguishable to the original media. In the case of modern AI-based face-swapping technologies, it is possible to replace the appearance of a person in an original video or image with that of another one and make it seem natural. Moreover, voice cloning is attained through feeding real audio information in the AI system, whereby the technology has the ability to change the whole context and the perceived intent of a speaker in the synthesised voice output.

Issues that Deepfakes raise

The abuse of the deepfake technology is of great concern in several fronts. Deepfakes are also applied to misinform, execute phishing attacks, create fake advertisements, infringe on personal privacy, and create and distribute pornographic materials, which happens not only to celebrities but also to a normal person.

In India, one of the most famous instances was that of 24-year-old Indian digital marketer Eemani Naveen, who was apprehended after creating and distributing a deepfake video of Indian actress Rashmika Mandanna through the digitally superimposition of her face onto video of British-Indian influencer Zara Patel entering a lift. The altered video was obscene in character and constituted defamation and violation of the modesty of the actress.

Deeperfakes in the United States have also been utilized to seek to manipulate the electoral process; an instance of this occurred with a fake audio message purported to be by Joe Biden sent via automated robocalls in an effort to demoralize Democrats to be involved in a primary election. In spite of these dangers, the deepfake technology is not necessarily an evil one as it can also be used in the form of creative expression and positive message spreading. Hence, care should be taken to identify the difference between bad malicious usage and justifiable or legitimate use of this technology.

Legal and Regulatory Issues

The blistering technological progress in the area of deepfakes often outcompetes the creation of proper legal regulations. The problem of jurisdiction is a crucial threat facing the law

 [ISSN 2581-5369]

enforcement agencies. The most important one is the question of jurisdiction: there is no established clarity over the scope of different countries and the scope of its jurisdiction to resolve cases involving deepfakes, especially when the creation of content in one country, and distribution in another. There are no digital territories, making it even more difficult to enforce the laws.

India does not have a single law to regulate the use of deepfakes even prior to addressing the issue of jurisdiction. Information Technology Act of 2000 that regulates digital transactions, cyber security, and e-commerce fails to assign the definition of deepfake as a specific type of cyber offence. The nearest relevant clauses are the 66D, 66E, and in the 67 of the IT Act, 2000 that provide penalties in relation to impersonation, transmission of private pictures without authorization, and dissemination of obscene information electronically. Nevertheless, it is argued that such provisions are not all-inclusive in order to deal with the multifaceted challenges of deepfakes.

Lack of strong legal regulations reduces the trust of the population in social platforms as it is extremely hard to draw the line between genuine and fake content when the latter is incredibly accurate.

The fact that there is no appropriate law that regulates the usage of such technology reduces the confidence that people place on social platforms since it is very challenging to differentiate realistic and deepfake content since it is pitch perfect. The paper is directed at discussing these three key issues.

1. India Lack of suitable legislations in the regulation of deepfake.
2. Jurisdictional ambiguity
3. Separating the effects on legitimate uses and those of illegal purposes.

II. LITERATURE REVIEW

The study Analysis and Conceptualization of Deepfake Technology as Cyber Threat, conducted by L. Dami[2] (2022), is based on the literature and is thorough. The paper uses scientific and grey literature, applicable wedges, and professional views to discuss the risks deepfakes create to cyber security. Dami notes that even simple smartphone filters and face-editing technology are trivial types of deepfake technology. More sophisticated deepfakes, when done legally are used in good purposes such as a parody, spoof, satire, and entertainment: CGI recreation of

[2] L Dami, 'Analysis and Conceptualization of Deepfake Technology as Cyber Threat' (June 2022) ResearchGate https://doi.org/10.13140/RG.2.2.21862.50246 accessed 7 March 2026.

 [ISSN 2581-5369]

characters in a film such as Rogue One.

The paper outlines some of the examples of risks linked to deep fakes as sextortion, defamation, biometric hacking, revenge pornography, cyberbullying, identity theft, threats to national security, and threats to the justice system. Dami stresses that users in social media are highly susceptible since they are usually believed what is shared to them. Importantly, the paper emphasizes that the work of a single photograph is already enough to produce a false deepfake material, and proving that an individual featured in a fake image is not a human being is a strenuous challenge.

Among the recommendations Dami makes, there are enhancing detection technologies, fostering digital education and awareness on AI professionals, regulating AI practices, using legal sanctions, and making international agreements to avoid geopolitical tensions created by the spread of misinformation through deepfakes. The paper ends by reiterating that the countermeasures of legislature are not enough as long as an average user is not able to tell the difference between genuine and fake content.

In their article Subjective and Objective Evaluation of Deepfake Videos, Korshunov and Marcel[3] (2021) analyzed subjective and objective deepfake detection evaluation. Participants of the empirical research were 60 naive participants (people who had little information about the deepfake technology) and were asked to state whether the video was authentic or fake. Out of the 120 Facebook videos (60 AI-synthesised and 60 original videos), the videos were classified into five levels of difficulty, starting with easy and ending with difficult.

The research found that there are a few important facts concerning the deepfake detection. The percentage of videos that were classified as very difficult and identified by the participants as deepfakes was only 24.5 per cent, proving the fact that the given type of technology has reached quite high sophistication. Within the category, which is considered as easy, the participants were quite struggling, and 75.5 percent of the cases remain undetected. Moreover, automated detection models, such as Xception and EfficientNet Variant B4 models, which were trained on datasets like Google and Celeb-DF, achieved lower results by humans in detecting manipulated material. Other factors that were identified to affect the results of the objective study were the quality of training data and the thresholds applied in detection. On the whole, the paper has come to the conclusion that the deepfake technology has advanced to the stage of its realism when it is incredibly hard and even impossible to tell a fake content and a real media.

[3] P Korshunov and S Marcel, 'Subjective and Objective Evaluation of Deepfake Videos' (June 2021) ResearchGate https://doi.org/10.1109/ICASSP39728.2021.9414258 accessed 6 March 2026.

The article The Future of Deepfakes: Need for Regulation by K. Gupta[4] (2023) examined the regulatory situation through the analysis of secondary data, literature review, and interviews with experts. The paper made a comparison of legislative systems in China, the United States and the European Union.

Gupta does not deny the valid and imaginative uses of the deepfake technology, such as artistic expression and medical testing (e.g. brain tumour detection). A medical industry, specifically, can find deepfake technology useful to create fake data of patients to be used in research without harm. However, the paper also defines considerable detrimental effects: the corruption of the litigation process by means of the fake evidence, the misinformation propagated in the name of the freedom of speech, and the distribution of the deepfakes pornography. Existing legislations in most jurisdictions are being determined insufficiently prepared to deal with non-consent usage of individual data, deep fake pornography, and violation of privacy. The paper suggests the global collaboration, development of AI detection devices, and improved awareness among the population.

The article by S. Kothari and S. Tibrewala[5] (2024) focuses on the threats that deepfakes have brought to the criminal justice system of India. Manipulation of the visual and auditory evidences can compromise the credibility of witnesses and result in wrongful convictions since any fake material can seem genuine to both the juries and the judges.

The paper examines some of the Indian laws such as the Information Technology Act of 2000, Digital Personal Data Protection Act of 2023, and Bharatiya Sakshya Adhiniyam of 2023 and finds a massive gap in each of them. As an example, the Bharatiya Sakshya Adhiniyam increases the area of digital evidence, but does not include particular recommendations on the detection and processing of the deep fake content. The authors suggest to enrich legal education of the law enforcement and judiciary, to create more sophisticated detection systems to be used, as well as to criminalize the malicious use of deep fakes.

In the article, S. Vig[6] (2024) discussed the Indian regulatory gap in the article Regulating Deepfakes: An Indian Perspective, using the qualitative approach and secondary sources of data. This research paper will address three main research questions: what are the problems of

[4] K Gupta, 'The Future of Deepfakes: Need for Regulation' (2023) HeinOnline https://heinonline.org/hol-cgi- bin/get_pdf.cgi?handle=hein.journals/nludslj2023&section=9 accessed 6 March 2026.
[5] S Kothari and S Tibrewala, 'AI's Trojan Horse: The Deepfake Conundrum Under The Criminal Justice System' (2024) 4(3) GLS KALP Journal of Multidisciplinary Studies 45 https://www.glskalp.in/index.php/GLSKALP/article/view/75/75 accessed 6 March 2026.
[6] S Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70 https://doi.org/10.5038/1944- 0472.17.3.2245 accessed 6 March 2026.

 [ISSN 2581-5369]

deepfakes, what are the laws and regulations that address this problem in India, and what can be done to enhance legislative policies.

The discussion finds that the current Indian laws, such as the intellectual property laws (such as Copyright Act of 1957), criminal law, and privacy rights laws, are not sufficient to deal with deepfake particular issues. Vig supports certain laws that define deepfakes in a clear way and specify legal consequences of its abuse, as well as technological means of detecting and digital literacy courses. In the paper, the author notes such international initiatives as the Digital Trust Initiative of the World Economic Forum (2022), the Global Coalition of Digital Safety of the Forum, and the EU programme Horizon 2020.

In the case of Legal Review of Liability: Deepfake Artificial Intelligence Contains Pornography, Abidin[7] (2023) examined the issue of legal liability framework, specifically within the Indonesian context. The research classifies the uses of MyHeritage, FaceApp, and DeepFake Studio as Private Scope Electronic System Operators (PSEs) under particular Terms and Conditions. The Moot indicates a legal gap in the Indonesian framework on accountability of deep fake abuse through normative juridical research methods. Abidin proposes extensive rules explaining the roles of individuals and platform developers as well as the civic educational campaign on responsible use of deepfake technology.

Research Problem

The production and distribution of deepfakes are associated with a high level of jurisdictional uncertainties in the cross-border context. At the same time, the emergence of easy access websites and technologies has made the production and dissemination of deepfakes easier than ever before, and the current legal systems have failed to effectively respond to the issues.

Scope and Limitations

The area of the research is limited to a few important dimensions of the regulation and the effects of deepfake technology. It explores legal gaps in the framework of the existing Indian legal framework, especially the lack of a particular legislation related to the production, consumption, and dissemination of deepfakes and the way such material takes advantage of loopholes in the current legislation, such as the Information Technology Act, 2000. It was also found that cross-border jurisdiction challenges are also examined as it has been determined that the enforcement issues arise when the deepfake material is produced in one jurisdiction and

[7] MI Abidin, 'Legal Review of Liability from Deepfake Artificial Intelligence That Contains Pornography' (December 2023) 39(2) Jurnal Sosial dan Pembangunan 344 https://doi.org/10.29313/mimbar.v39i2.2965 accessed 6 March 2026.

transmitted elsewhere. Moreover, it performs a comparative regulatory analysis and studies and compares the legal strategies of the United States, the European Union, and China with the purpose of extracting helpful information to be used to implement certain changes in the Indian legislation. The study also examines the technological and social awareness, such as the detection rate of deepfake technologies and whether people can recognize the manipulated media, with the significance of raising awareness and educational efforts.

In spite of its contributions, the study is prone to some limitations. The first weakness is the lack of data since the research depends more on secondary data like case studies and literature reviews that might restrict its capacity of capturing fast changing trends or the possibility of empirically validating the research using primary data. The jurisdictional focus of the study on the regulatory framework in India also implies that its findings and suggestions might not be generalizable especially in nations that have quite different legal systems. Also, there are still enforcement issues because of the untrustworthy detection technologies and low global collaboration in combating offences related to deepfakes. The second weakness is the reactive aspect of the existing law, which tends to act on the effects of deepfakes after they have taken place and not stopping their production. Lastly, and the least important are the gaps in implementation, where although the paper suggests legislature changes, it does not include all the practical issues of implementing new laws, training law enforcement officers, and successfully controlling digital platforms.

III. DEFINITION AND EVOLUTION OF DEEPFAKES

Deepfakes may be simply defined as reality manipulations made by means of artificial intelligence. They entail creation or manipulation of audio, images, or videos in a manner that voice, face or activities of one individual are substituted with those of another. Put simply, deepfakes are a more sophisticated type of photo manipulation, since they alter both visual and audio aspects of text. In an example, they can make it look like someone said something that he/she did not say or did not do.

This technology is created based on deep learning which is a field of artificial intelligence and machine learning. Deep learning works in a human-like fashion because the computer process depends and uses neural networks which make computers handle extensive data. The computers acquire to understand intricate patterns and variations in images, sounds, and texts by feeding big amounts of labelled information into these systems. With this acquired knowledge, neural networks are able to create artificial images, videos, or sounds that seem to be very real, despite them being completely fake.

 [ISSN 2581-5369]

One of the technological techniques involved in the generation of deepfakes is the Generative Adversarial Network (GAN) model. The GANs are comprised of two neural networks that act in competition with each other: the generator and the discriminator. The generator will generate artificial data considering the training datasets that uploaded to the system. It aims at producing data that is persuasive to seem factual. The discriminator however considers the produced data and contrasts it with the original data and then decides whether it is genuine or is a fabric. The generator is refined through this competitive process by learning the feedback of the discriminator to make the content generated more realistic. This is a cycle which repeats until the produced image, video, or audio is so realistic that it can hardly be determined that it was not a real piece of content.

Emergence of Deepfakes

In 2017, a user in the social media site Reddit posted a set of pornographic videos in which faces were superimposed onto other bodies using an artificial intelligence system, coining the name deepfake. Before 2014, realistic manipulated images were created by small, tedious, manual means and sophisticated computer graphic algorithms like 3D modelling and morphing. Nonetheless, there have been forms of manipulated imagery that were in existence earlier. Among the earliest, one can speak of the portrait of the politician John Calhoun, which was modified way in 1860 by changing his head with the one of a U.S. president with the purpose of propaganda[8]. This manipulation was done manually by methods like splicing together images, deleting or repainting parts of photographs and duplication of objects in images. Changes in colours, magnitude and proportions were also done to increase the realism of the manipulated portrait.

After 2017, the deepfake technology became viral worldwide. This technology became accessible to even those who lack sophisticated technical skills since it was developed into easy to use applications like FakeApp, FaceSwap, and ZAO. Moreover, a large number of guides that are found on websites like YouTube made the process of creating deepfakes even easier. Consequently, every person that has access to proper computer capabilities could create really compelling manipulated videos.

Deepfakes do not only go in the visual media, they also reach into the audio manipulation. Artificially intelligent devices are able to duplicate the sounds of an individual with great precision, which means production of synthesized audio messages that can be perceived as true.

[8] K Nagumotu, 'Deepfakes Are Taking Over Social Media: Can the Law Keep Up?' (2022) 62(2) Journal of the Franklin Pierce Center for Intellectual Property 102 https://search.ebscohost.com accessed 20 October 2024

 [ISSN 2581-5369]

Due to this ability, cybercriminals have also manipulated it to do fraudulent actions like scams and phishing. Deepfakes with audio are common to disseminate fake news, especially in relation to politicians or other people of influence. As an example, in March 2019, hackers exploited artificial intelligence software to impersonate the voice of the chief executive officer of one of the parent companies in Germany. They were able to use this imitation and conquer the CEO of one of the subsidiaries in the United Kingdom to part with thousands of pounds. The money was channelled[9] through a supplier in Hungary and later on to other different places, including Mexico.

Notable technological advances in between 2020-2022 boosted the believability of deepfakes. Improvements in technologies like better lip-synchronization greatly enhanced the level of match-up of the facial movements to speech. Moreover, more sophisticated AI tools like DeepFakeLab and StyleGAN also allowed one to make higher-resolution deepfake videos and images and thus make them even more persuasive. Despite the valid use of deepfakes in the entertainment, parody and digital graphics fields, it has also been employed in crimes like cyber-kidnapping, revenge porn, blackmail, sextortion and defamation. Though the cases of people who have found themselves as victims have mainly been those who are in positions of authority, there has been an increasing number of ordinary people being victimized. The fact that deepfake content is widely spread on the social media platforms also compromises the trust in the online information because people find it more challenging to distinguish the real content and the manipulated content. Such a case poses a threat of establishing a zero-trust environment where digital media are highly doubted.

Weakness of Deepfake Detection Technology.

The other significant issue is the ever-increasing access and affordability of deepfake technology. The effectiveness of the current detection tools is likely to decline over time due to the ease of the technology. A majority of deepfake detection systems can work by guessing the possibility that some content is being manipulated. As an example, a video can be reported by a detection tool that it is 65 percent real. Nevertheless, these findings tend to suggest which parts of the video were manipulated. This puts in doubt the manner in which authenticity thresholds are supposed to be understood. It also leaves questions on whether internet service providers ought to block content automatically when a certain degree of manipulation is detected and the level of percentage that ought to be used as enough to term content as fake.

[9] C Stupp, 'Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case' (Wall Street Journal Pro, August 2019) https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 accessed 20 October 2024.

 [ISSN 2581-5369]

Other detection programs also have the limitation of time of video it can analyse meaning that only a portion of a video could be analysed and the rest of the video is not analysed. In addition to this, most of the detection technologies use past trained datasets and thus cannot easily detect manipulations whenever new entirely novel or real-time data is presented. Therefore, even with the continuous progress, the technology of deepfake detection is not a perfect one. Research institutions, governments and technology firms worldwide are still striving to advance these systems so as to increase their accuracy and reliability in detecting manipulated online content.

IV. CURRENT LAWS THAT REGULATE DEEPFAKE TECHNOLOGY IN INDIA

There is no specific regulatory system of deepfakes in India. Nevertheless, some of the clauses provided in the Information Technology Act of 2000 and the Constitution of India provide partial yet incomplete coverage.

Right to Privacy

The right to privacy is captured in the Constitution of India, Article 21[10], which provides the right to privacy as part and parcel of the individual liberty and dignity. The case K.S. Puttaswamy v. Union[11] of India determined that the right to privacy could be considered as three-fold, namely:

1. Confidentiality of personal data.
2. Privacy of choice
3. Shielding against the intrusion of a physical body of a person.

The illegal production of non-consentual deepfakes with the aim to damage the reputation of an individual or compromise the physical autonomy of a person is a breach of Article 21. The fact that Deepfake technology is using the voice, face, and personal information of a person without their permission is a definite violation of privacy through providing fake information.

Nonetheless, not every content that has been fabricated with regards to personal information infringes the rights to privacy. The negative and negative aspects of the deepfakes also have their positive righteous uses and the right to privacy should be interpreted in relation to the right to freedom of expression. As an example, a deepfake video of Mark Zuckerberg designed as an art exhibition to show how social media controls individuals[12], tracks them without violating the privacy or dignity of anyone was a social good and, therefore, did not harm anyone.

[10] Constitution of India, Article 21.
[11] K S Puttaswamy (Retd) vs Union Of India (2019) 1 SCC 1.
[12] Facebook Lets Deepfake Zuckerberg Video Stay on Instagram' (*BBC News*,12June 2019) <https://www.bbc.com/news/technology-48607673> accessed 8 March 2026

 [ISSN 2581-5369]

The right to the privacy is related in Article 8 of the European Convention on Human Rights (ECHR)[13], whereas the freedom of expression is the subject of Article 10[14]. The ECHR has declared that the right to privacy also enshrines the right to protection of one honour and reputation whereas it has concurrently declared that the freedom of expression is rich and that it encompasses the right to offend, shock, criticise, express personal views, satirise. The main issue engagement is the border between freedom of expression and the right to privacy, which needs the examination of the intent and determination of clear legal definitions that will restrict the cases of evil use of deepfakes but not valid expression.

Information Technology Act, 2000

A number of the provisions of the IT Act, 2000 apply to the regulation of deepfake, even though none directly concern the technology:

Fraud using computer resources by way of impersonation falls in section 66D[15] of the Information Technology Act, 2000, which is a criminal offence as outlined in the law. Despite the lack of mention of deepfakes in the provision, the provision can be construed to include instances of impersonation fraud executed with the use of deepfakes technology.

The Information Technology Act, 2000, section 66E[16], is the law that gives the punishment of taking pictures of the personal or intimate part of an individual without their consent. But this does not specifically refer to the word deepfake, or even to videos or audio materials. As stated in this paper, deepfakes can be applied to not only images, but can also be done with manipulated videos and audio recordings. Moreover, the production of deepfake materials through artificial intelligence is quite different compared to the fact that it is possible to capture real-time images of the personal or even intimate parts of a person.

Section 67[17] of the Information Technology Act, 2000 punishes persons that publish or transmit obscene material electronically. Nevertheless, the clause does not criminalize the process of downloading, sharing or viewing such content except when it refers to a child. Because deepfakes can easily be confused with authentic content, people can easily post, download, or watch obscene videos, images, or sound without knowing that it is a fake image or fake video. When the spread of such content goes viral there is no longer any control over its distribution and the damage to the reputation of the person might already be done.

[13] European Convention on Human Rights, Article 8 (1950).
[14] European Convention on Human Rights, Article 10 (1950).
[15] Information Technology Act 2000, s 66D.
[16] Information Technology Act 2000, s 66E.
[17] Information Technology Act 2000, s 67

 [ISSN 2581-5369]

Section 67A[18] of the Information technology Act 2000 spells out a punishment against publishing or transmitting sexually explicit material in electronic format whereas Section 67B[19] is about punishing publishing or transmitting sexually explicit material relating to children. These clauses deal with severe crimes of explicit digital content, yet they fall short of considering the problematic issues of deepfake technology that is emerging.

Copyright Act, 1957

Indian copyright act Section 52 of the Indian Copyright Act, 1957[20] deals with the concept of fair dealing in that it gives a list of works that are considered as being non infringement. This is in accordance to Article 13[21] of the Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement that states: Members shall restrict limitations or exceptions to exclusive rights to only such special cases that do not conflict with a normal exploitation of the work or that do not unreasonably prejudice the legitimate interests of the owner of rights.

The Indian doctrine of fair dealing which is usually criticised on the basis of being too rigid to be compared to the broad fair use doctrine in the United States may become a means of fighting deepfakes. Nevertheless, its present formulation would define all deepfakes, such as the ones produced to entertain or serve any other honest purpose, as copyright violations. This position needs to be changed to allow legitimately created deepfakes without the arbitrary classification of all technologies in the technology as illegal.

Each of the legal provisions described above provides a shield against the impact of deepfakes but does not cover the creation of the latter. This basic fault in the system requires the implementation of new or additional laws.

V. JURISDICTIONAL CHALLENGES

Jurisdictional ambiguity of the current legislation in India is one of the major problems, which are explored in this paper. The act under Section 1(2) codified with Section 75[22] that provides extraterritorial application of the Act, in that, it concerns itself with offences committed abroad by the any individual. Such extraterritorial jurisdiction is based on the fact that nationality of the offender does not matter, but rather punishing the offence committed.

Nevertheless, although there is a list of different offenses, including cyber terrorism, that is

[18] Information Technology Act 2000, s 67A
[19] Information Technology Act 2000, s 67B
[20] Copyright Act 1957, s 52.
[21] TRIPS Agreement, Article 13 (1994).
[22] Information Technology Act 2000, s 75.

 [ISSN 2581-5369]

shown in Section 66-F[23], and special penalties are provided according to the legislation, there is no explicit regulation of the creation and distribution of deepfakes. Such offences have an extremely broad scopes, and deepfakes pose a special set of problems that cannot be compared to that of other types of digital manipulation.

There is a serious difference that must be made between morphing technology and deepfake technology, with the former being a straightforward fusion of two images to form a composite, and the latter being a complex series of artificial intelligence that seeks to produce completely new photographs and videos on things that may not have ever happened or which portray people that do not exist at all.

The Act of 2000, Section 66-C of the IT Act under identity theft punishes those who commit the crime, the offence is: "Whoever commits the offence of identity theft by fraudulently or dishonestly using the electronic signature, password or any other unique identification feature of any other person... Nevertheless, deepfakes are beyond the context of identity theft. This is not only an issue of utilizing the features of another person but of creating something completely non-existent either a description of events that did not happen or people who do not exist. One of the provisions meant to deal with identity theft cannot thus deal with the entire scope of the deepfake technology.

IT Act, sections 66-E, 67, 67-A and 67-B address the issue of pornography in terms of cybercrime. Section 67 however, exclusively punishes transmission and publication of obscene content; it is not punishable if such content is viewed, downloaded and even possessed unless the victims are children. Although the idea to criminalise the act of transmission and publication is praiseworthy, there is a huge gap namely that the real makers of the deepfake materials are not sufficiently punished.

This flaw is compounded by the fact that there are no trusted detection programs that would be able to differentiate between good content that could have been morphed and pure AI-generated deepfake content.

The Intermediary Guidelines and Digital Media Ethics Code, 2021, provides the exemption of intermediaries regarding the liability of third-party information, data, or links of communication placed on the medium owned by the intermediaries. Although it would be unfair to penalize the intermediaries who do not knowingly store the deepfake information, this exemption begs a critical question: who should the liability be when such websites end up serving the distribution

[23] Information Technology Act 2000, s 66F.

 [ISSN 2581-5369]

of harmful information willingly or unwillingly?

VI. CONSEQUENCES OF DEEP FAKE THAT REQUIRE REGULATION

Deepfakes were first created and applied in legitimate ways like, entertainment, parody, satire, advertising, and even in some cases in healthcare. But with these positive applications, the technology has been misused and been used to do negative and unlawful activities. As mentioned above deepfakes are highly exploited to pass misinformation, especially during elections, with the aim of deceiving the masses. They are also prone to abuse to make pornographic materials by overlaying images of celebrities and even of normal people, whereby their intentions are usually to tarnish reputations, blackmail, or even perform sextortion.

This becomes more complicated when such bogus content is used as evidence in a court of law. When a deepfake content is provided in a court of law as real evidence, it becomes very difficult to prove that it is not credible by the other party. This is a challenge since a deepfake media can look very realistic and difficult to determine whether it is a fake or original work. When such distorted material is admitted in courts as genuine, the verdicts made may bring grave injustices, and this may lead to conviction of innocent people. Although it may be later discovered that such evidence is the fake one, the harm to the reputation of a person, his/her personal life, and career can be, in most instances, irreversible[24] and serious.

Deepfakes contain realistic characteristics and fabricated elements to the point that even an average individual may find it hard to spot the differences as being manipulated. Assuming that such content will be used as evidence in court, it is possible that the process of legal actions will become considerably slower because of the time spent on checking authenticity. Besides, the lack of clear definitions and specific legal regulations that govern deepfakes contribute to an even greater difficulty in prosecuting and compromising the effectiveness of the legal system to handle such situations.

VII. REGULATION OF WEBSITES AND TECHNOLOGIES THAT FACILITATE THE PRODUCTION OF DEEPFAKES

The control of the creators and those involved in disseminating such materials is of utmost importance since the consequences of progress in the field of deepfakes are very high. Although there is a significant amount of regulation on the subject, the most important question is whether these efforts are sufficient to keep up with the fast development of technology. The existing

[24] H Mudgal, 'The Deepfake Dilemma: Detection and Decree' (Bar and Bench, 18November 2023) https://www.barandbench.com/columns/deepfake-dilemma-detection-and-desirability accessed 8 March 2026.

laws on IT, which are applicable in India, have not been developed with the deepfakes in mind. A special regulatory framework that limits the misuse of the deepfake technology is thus required.

At the current time, no legal means of regulating the technologies and platforms, which can be used to create deepfakes, are available. The current regulatory strategies are aimed at regulating and compensating the outcome of the harm instead of avoiding the production of harmful material in the former.

Why Are Deepfakes So difficult to regulate?

Although a variety of detection tools are available, it is still hard to distinguish between the images that have been edited with the help of the regular photoshopping applications and the fake news produced on the basis of the latest technologies of artificial intelligence and networks. Moreover, the deep fake technology is still advancing and the rate of change is very high and it may be hard to come up with the laws to effectively regulate the technology since it is ever changing. Even in the case of national regulation, the deepfake technology is not limited by geographical locations. It is an international phenomenon, which enables people in any part of the world to produce deep fake content against people living in totally different regions. Such manipulated material is further spread at a very high pace due to the global nature of the internet. Consequently, resolving such cases is becoming more complex, especially since at present there is no common law and international collaboration that is specifically aimed at controlling the deepfake technology[25].

Section 69A of the IT Act, 2000[26] within India, gives the Central Government the power to censor the access to information that it deems to be threatening to the sovereignty, security or the civil peace of India. But such a provision is in itself reactive not preventative. Before the content is done the prescribed procedure, it may have been sent out to millions. Moreover, blocking access will not eliminate future access, as copies already in circulation will be transferred.There is no federal law regulating AI or deepfakes in the United States.

Nevertheless, in a number of states specific laws have been introduced:

GLOBAL APPROACHES

The next section considers the laws that various jurisdictions worldwide have taken concerning the issues of deepfake technology.

[25] N Ughade, 'Are Deepfakes Illegal? Deepfake Laws & Regulations to Know' (HyperVerge) https://hyperverge.co/blog/are- deepfakes-illegal/ accessed 20 October 2024
[26] Information Technology Act 2000, s 69A.

 [ISSN 2581-5369]

United States of America

The United States does not have a federal law that specifically governs the use of artificial intelligence or deepfake technology. Nonetheless, some states have proposed legislation in an attempt to deal with deepfake abuse. The Colorado AI Act can be regarded as one of the brightest examples since it is regarded as one of the first broad legislative frameworks that impose obligations on both artificial intelligence systems developers and deployers. The Act limits the production and distribution of the deepfake content without consent. It was passed on May 17, 2024 and is likely to take effect in 2026.[27]

The next important legal action is the Texas Deepfake Pornography Law. Within this provision, deepfake video is referred to as a video that is made with the aim of deceiving by portraying a real-life person doing something that never took place. The legislation outlaws the creation and sharing of pornographic content that includes a human being without their effective consent, especially that depicting the person as having sex or showing their nooky parts.

On the federal level, the U.S. government has also presented the Deepfakes Accountability Act. The proposed law will also compel users who create the content related to deepfakes to provide the labels to such content in an obvious manner by using such solutions as digital watermarks and notifications about the alterations that took place to the original content. The inability to meet these requirements can lead to criminal punishment.[28]

China

In China, the use of the deepfake technology is controlled by the Deep Synthesis Provisions that are declared by the Cyberspace Administration of China. Such laws mandate creators and websites that feature manipulated or synthetic media to identify such material and have systems that enable the authorities to track the source of the fake content.

Europe

Some of the regulatory measures that the European Union has implemented to curb misinformation and deepfake technology are as follows. First, there was the introduction of the Code of Practice on Disinformation as a voluntary system to combat the use of misinformation online. This effort has since been strengthened by the Digital Services Act which imposes responsibilities on large technology firms, like Google and Meta, to detect and label deepfakes

[27] H Anderson, I Nunes and J Oltean, 'Newly Passed Colorado AI Act Will Impose Obligations on Developers and Deployers of High-Risk AI Systems' (White & Case LLP, 20 June 2024) https://www.whitecase.com/insight-alert/newly-passed-colorado-ai-act accessed 8 March 2026.
[28] H.R. 5586 (2023) Congress.gov https://www.congress.gov/118/bills/hr5586/BILLS-118hr5586ih.pdf accessed 9 March 2026.

 [ISSN 2581-5369]

or manipulated material on their sites. Non-conformity to such requirements may result into huge monetary fines.

Also, the European Union Artificial Intelligence Act, Article 50,[29] creates the duty of deployers of AI systems [30]that generate synthetic or manipulated media. It mandates such content to be properly labeled and imposes regulatory duties on a number of players within the AI ecosystem, such as providers, deployers, importers, distributors, and manufacturers of AI-based systems. The European Union has equally been spending huge sums in research programs that would enhance detection technologies. As an illustration, there was a funding programme called the Horizon 2020 that was active between 2014 and 2020 and had a budget of about 80 billion euros to fund research on misinformation, cybersecurity, and technologies to detect deepfakes.

However, notwithstanding these legislative and policy efforts, much of the current regulatory efforts are focused on the effects of deepfake sharing once the information has been shared on the Internet. Consequently, they serve as mostly remedial as opposed to preventive measures as they do not directly limit or inhibit the original production of manipulated media.

An Indian PIL to prohibit AI Technologies

In Chaitanya Rohilla v. Union of India[31], a Public interest litigation Union of India was initiated by Advocate Chaitanya Rohilla requesting the Central Government and the Ministry of Electronics and Information Technology to detect and prevent access to websites that offer access to AI software that can be used to create deepfakes. The PIL considered the threats of uncontrollable AI, the misleading character of deepfakes, the inexistence of any legal definition of deepfakes, and the insufficiency of the laws applicable to deepfakes in India. The petition also requested development of regulations that control AI technology in relation to the fundamental rights as provided by the Constitution.

Manmohan, an Hon'ble Acting chief justice, and Justice Tushar Rao Gedela advised the centre to deal with deepfakes urgently, describing it as a grave threat to society.[32] The court instructed the Government of India to present their response on the matter and the Centre is on the way to handle the matter.

---

[29] EU AI Act, Article 50 (2024).
[30] Hickman, DS Lorenz, A Jha, and DC Teetzman, 'Long Awaited EU AI Act Becomes Law After Publication in the EU's Official Journal' (White & Case LLP, 16 July 2024) https://www.whitecase.com/insight-alert/long-awaited-eu-ai-act accessed 20 October 2024.
[31] Chaitanya Rohilla v Union of India (28 August 2024).
[32] 'Start Working Against Deepfakes, Delhi HC Tells Centre' (The Hindu, 28 August 2024) https://www.thehindu.com/news/cities/Delhi/start-working-against-deepfakes-delhi-hc-tells-centre/article68576870.ece accessed 20 October 2024.

Major Technology Company Approaches

Large technology firms have started taking action to counter the increasing threat of deepfakes. As an example, Meta (previously Facebook) has been working on artificial intelligence detection systems that reverse-engineer to detect fake media and can trace its potential source. Furthermore, the company has established the policies that are supposed to limit or prohibit some data generated by AI on its websites, although the content produced with other purposes, like parody, satire, and entertainment, would be allowed. In the same manner Microsoft has also launched one tool called Microsoft Video Authenticator that analyzes the videos or images and gives a percentage result on how high the chances are that the material has been tampered with.

Considering the capacity of the social media sites to disseminate content to the millions of people just about once they are posted, the question of how to regulate the content posted on the media sites has gained relevance over the past few years. This puts the dire necessity to focus on preventive actions as opposed to fixing damages once they had been done, which the existing legal system in India leaves wanting at present. As a result, India needs specific changes to its current legislation or a new law that would be specifically related to the issue of deepfakes. The same should be applied to the software and online platforms facilitating the generation of deepfakes, where the focus should be on preventing the abuse, as opposed to handling the elimination of the toxic content once it has become viral.

VIII. RECOMMENDATIONS

This paper suggests the following. India ought to propose certain legal solutions to govern the process of making and sharing deepfakes, making the responsibility and responsibility transparent, as well as holing the loopholes that exist in the Information Technology Act, 2000.These measures should also enhance the jurisdictional clarity, especially when cross-border problems arise that are connected to deepfake messages.

Moreover, the development of detection abilities should be increased by investing into the sophisticated AI-driven technologies that would detect deepfakes and this effort could be accomplished through the collaboration of governmental agencies, commercial enterprises, and educational establishments.

In addition, the digital platforms and intermediaries, such as larger social media companies, should be held more accountable and they should put in place stringent measures to detect, eliminate and curb the dissemination of malicious deepfake messages.

 [ISSN 2581-5369]

Lastly, awareness campaigns would be implemented to teach the citizens about the type of deepfakes, as it consists of enhancing the capacity of the population to identify the misleading media and reporting of misleading or falsifying information.

IX. CONCLUSION

The issues of deepfake technology are quite serious ethical, legal, and social that require immediate and extensive solutions. This paper points out the flaws of the Indian legal framework on the malicious intent application of AI-generated content and the vagueness of jurisdiction in cross-border application. An international comparison implies what constitutes the possible ways of reforming the legislation; yet, India is in dire need of enacting a law that will classify deepfakes, host AI-driven websites, and develop effective detection solutions.

The research proposes a moderated solution to avoid biased perspectives on the problem that would allow safeguarding both legitimate and positive uses of deepfake technology and effectively reducing the threats it presents to personal rights, democracy, and trust in society. This balance will necessitate a concerted action of the legislators, technology developers, law enforcement agencies, the judiciary, and the civil society to build a sound and dynamic regulatory framework that will help in responding to the dynamic nature of AI-generated media.

*****

 [ISSN 2581-5369]