IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

THE EMERGING THREAT: DEEPFAKE AND WOMEN IN INDIA

Dr.Sarigama.R.Nair
Independent Research Scholar
University of Kerala ,India

Abstract

Artificial Intelligence (AI) programs are a common occurrence, and the use of deepfake software is steadily increasing. However, new security flaws and issues arise with every new piece of technology. Deepfakes are a new tool that the disinformation threat is using. There are obvious advantages to synthetic media or deepfakes in a number of fields, including criminal forensics, filmmaking, accessibility, education, and artistic expression. But as synthetic media technology becomes more widely available, so does the possibility of exploitation. Particularly women have suffered as a result of deepfake technology's advancement. They have become targets of revenge porn and have been harassed and bullied as a result. Deepfakes are videos that intentionally mislead people by using Artificial Intelligence (AI), deep learning, and photoshop to create images of real events that propagate false information. Deepfake technology is becoming more widely used, which puts women in India at serious risk and exacerbates already-existing issues with safety, privacy, and reputation. This article explores how deepfakes affect women, with particular attention to issues like misinformation, defamation, and the production of non-consensual pornography. Concerns regarding the need for more robust legal frameworks to address issues like privacy, data protection, and cybercrime have been raised by the growing use of fake technology. Although India has laws that can be used to counteract deepfake technology, more specialized legislation is required to address the particular problems that deepfakes present. The necessity for the Indian government to address this matter and create a regulatory framework is emphasized in this paper. Additionally, it examines how well the Indian legal system handles the violation of people's rights caused by deepfakes and makes the case that, even though the laws as they stand implicitly forbid the dissemination of deepfakes, there are certain issues with the system that could be resolved by proposed legislative changes.

Key words- Artificial intelligence,Deepfake,Pornography,Criptography,Revenge porn

Introduction

From being a minor annoyance, disinformation and hoaxes have developed into a kind of warfare that has the power to polarize society, cause social unrest, in certain situations, even affect the outcome of elections. Nation-state actors with unprecedented ease and reach can manipulate social media narratives on behalf of violent extremists, ideological adherents, geopolitical aspirations, and commercial interests.Deepfake technology is a potent instrument that can be applied to both good and bad ends. It is critical to understand the possible risks associated with deepfakes and to take precautions against falling for a scam. In India,

home to the world's largest population, may see a disproportionate impact on women from deepfakes due to a lack of regulation in this area. Simultaneously, this regulatory framework ought to take into account the larger context of media regulation and the evolution of concepts related to sexual privacy in India. This essay highlights the risks and decisions involved in addressing the threat of deepfake pornography by drawing on regulatory theory on the governance of technology and media as well as critical approaches to technology. On a larger scale, deepfake pornography mainly targets women, which contributes to the objectification and dehumanization of women in society. Most of the videos on deepfake pornographic websites feature Indian women, particularly those who work in the entertainment business. Furthermore, in an effort to silence and discredit their voices, politicians and journalists have frequently been the target of sexual harassment and abuse through the use of such non-consensual deepfake pornography. The people who have been the target of these deepfakes must endure financial, psychological, and emotional suffering on a micro level.

These people not only experience reputational harm that hinders their ability to pursue careers, but they also experience intimidation and bullying in the form of sextortion and the issue of "revenge porn." Finally, the victims' families and professional connections may also suffer harm from this violation of their safety and dignity. Deepfakes thus present a significant barrier to the realization of gender equality in our technologically advanced society.

What is Deepfake Technology?

Deepfakes are collections of synthetic photos and sounds assembled by machine learning algorithms with the intention of disseminating false information and substituting comparable synthetic voices or likenesses for the appearance or voice of real people. In 2017, an anonymous Reddit user going by the handle "Deepfakes" coined the phrase "deep fake." This person created and uploaded pornographic films by manipulating Google's deep-learning, open-source technology. Deepfake technology uses advanced algorithms to alter and superimpose face characteristics onto pre-existing footage, allowing the production of phony photos and movies that are incredibly convincing. When used improperly, this potent instrument can be used to create deceptive and explicit content, raising a number of privacy issues for women. The consequences are serious and wide-ranging, ranging from deliberate character defamation to non-consensual sexual content. More accurate deepfakes have been produced thanks to the usage of a technique known as Generative Adversarial Networks (GAN), which combines two AI algorithms. One algorithm creates the fake material, while the other scores the system's efforts and helps it get better. A deep learning architecture called a Generative Adversarial Network (GAN) pits two neural networks against one another in a framework akin to a zero-sum game.

The makers of deepfakes employ a big library of source photos to make sure the final product is as authentic as possible. For this reason, there are more deepfake videos featuring politicians, celebrities, and public figures produced. Next, one pieces of software uses the dataset to make a false film, while another uses it to look for evidence of forgery. The two software programs work together to generate the bogus video until the second program is unable to recognize the counterfeit. When machine-language models educate themselves, this is referred to as "unsupervised learning." It is challenging for other software to recognize deepfakes using this strategy. There are three types of deepfake technology incudes deepfake images and videos, deepfake audio and texual deepfakes. Deepfake Videos and Images refer to content that has been modified or created to provide information or behavior that differs from that of the original source. Deepfake Audio is a danger to speech-based authentication systems, particularly for those whose voice samples are readily available, such as politicians and celebrities. Textual Deepfakes refers articles that seem to have been written by genuine people are referred to as such.

Deepfake technology and women

Men and women are equally at risk from deepfakes, although women are more frequently the targets of this harmful information. A deepfake is an artificial intelligence (AI) abuse that predominantly targets women by editing their images and videos to produce non-consensual pornography. The most recent instances are the deepfakes that went viral and targeted actresses Alia Bhatt and Rashmika Mandanna. According to a recent analysis, India is among the nation's most vulnerable to this new digital menace, and politicians and celebrities are especially vulnerable. However, its victims are not only well-known people. The danger of such injury to any woman increases with the increasing popularity and ease of use of AI tools. Gender discrimination is not inherent in technology; rather, mishandled technology often reflects cultural prejudices and gender power dynamics. Misogyny, sexism, objectification, and gas lighting are among the many reasons why women are disproportionately targeted, starting with Photoshop-like tools and continuing with deepfake technology.

The fact that women are frequently the targets of deepfake abuse exacerbates its effects. Research has indicated that deepfakes are far more likely to target women than males. The discrepancy draws attention to the systemic gender prejudices and inequalities that permeate society as well as the ways in which technology may be used as a weapon to uphold this damaging practice. A holistic strategy is needed to address the problem of deepfake abuse, whether it is against women or anybody else. This includes cultural, technological, and legal initiatives such as raising awareness.

The possibility for exploitation and suffering associated with revenge porn and deepfake technology has led to increased worries in India. Sharing graphic or private photos or films of someone without their permission is known as revenge porn, and it usually happens with the intention of defaming, harassing, or exacting retribution. This problem is made worse by deepfake technology, which allows for the production of incredibly lifelike fake content, making it harder to tell what is real and what is edited deepfake and revenge porn have become more popular in India, posing a number of social and legal issues. The absence of strong regulations to expressly address deepfake technology and revenge porn has made it difficult for victims to pursue justice. There is a need for additional specific laws to handle these contemporary types of exploitation, even though the Indian Penal Code (IPC) and the Information Technology (IT) Act have provisions that can be enforced in some situations. For victims of revenge porn and deepfake content in India, the shame and social views around concerns of sexuality and privacy can make matters worse.

When their privacy is invaded in such a conspicuous and dehumanizing way, a great number of victims may experience emotional anguish, harassment, and social exclusion. The victims of this kind of abuse may experience serious problems as a result, such as "post-traumatic stress disorder, suicidal thoughts, depression, among other symptoms of poor mental health."This topic has gained attention because of the severe trauma that the victims are experiencing. The incidence of misuse based on deepfake images has been rising recently.

Highest risk factors

In India, the highest risk factors associated with deep fake technology include:

- Cyber bullying and Harassment: Deepfakes may be used as a weapon for cyber bullying and harassment, with a focus on underrepresented populations including activists, women, and minorities.
- Financial Fraud: By producing phony audio or video recordings of people approving financial transactions or submitting false claims, deep fakes can be used to cause financial losses
- National Security Threats: Malicious actors may utilize deepfakes to fabricate films purporting to show members of the armed forces or government officials, raising the possibility of security breaches or diplomatic disputes.
- Privacy Violations: Intimate video or photo spoofing can be produced utilizing deepfake technology, infringing upon people's privacy and perhaps causing them mental anguish or even physical injury.
- Misinformation and Fake News: By using deep fakes to produce convincingly fake audio or video recordings of famous persons, disinformation and phony news may be widely disseminated, which can have serious social and political repercussions.
- Election Interference: Videos of political candidates can be manipulated through the use of deep fakes, which might sway election results and jeopardize the democratic process.
- Reputation Damage: People are susceptible to having their reputations tarnished by deepfake films that show them acting in an improper or scandalous manner, especially public figures and celebrities.
- Legal and Ethical Concerns: Deepfake technology is widely used, which presents difficult moral and legal issues pertaining to digital material validity, freedom of speech, and privacy rights.

Spotting deepfakes

The term "deepfake" describes material, usually videos, that have been edited with artificial intelligence to make someone appear to be talking or acting in a way they never did. How to recognize a deepfake?

- Keep an eye out for irregularities: Take note of any changes in lighting, lip movements, or facial expressions that deviate from the subject's customary behavior or the environment.
- Keep an eye out for odd artifacts: Deepfakes frequently leave behind odd blurs or distortions surrounding the subject's face or torso.
- Check the source: If the video appears shady, look for the original source or supporting documentation from trustworthy sources.
- Pay close attention: Occasionally, the audio may not fully sync with lip motions, or the speech tone may appear strange.
- Sensational material should be treated with caution because deepfakes are frequently used to disseminate false information or make eye-catching headlines.
- Utilize technology: Although they might not be perfect yet, a number of instruments and programs are being created to identify deepfakes.In general, in today's media environment, skepticism and critical thinking are essential for spotting deepfakes.

- Mismatches in lighting and color: Inspect the subject's face and surroundings for any discrepancies in lighting.
- Compare Audio Quality: The video's visual content and the deepfake's possibly problematic audio should be compared.
- Strange Body Shape or Motion: Pay attention to any strange body shapes or movements, especially when exercising.
- Artificial Facial Movements: Draw attention to exaggerated or jerky facial expressions that clash with the video's overall tone.
- Unnatural Facial Feature Positioning: Look for any distortions or misalignments in your facial features.
- Uncomfortable Posture or Physique: Pay attention to any unusual actions, postures, or body proportions.

Combating deepfake

- Policies of Social Media Platforms: Press social media companies to combat deepfakes. Many sites currently have appropriate usage guidelines or regulations in place for deepfakes. To avoid deepfakes from spreading throughout their networks, these platforms could take action by implementing dissemination restrictions or using alternate promotional strategies like downranking or limiting sharing.
- Improved Media Literacy: In order to develop a perceptive public, media literacy initiatives must be strengthened. Consumer media literacy is the best defence against misinformation and deepfakes.
- Technological Solutions: Provide easily comprehensible and intuitive technological solutions to identify deepfakes, verify media, and endorse reliable sources.
- Individual Responsibility: It is the duty of each person to exercise critical judgment while consuming material online. Give material a moment of thought before posting anything on social media. It is crucial to participate in the fight against the "infodemic" by acting responsibly when using the internet.
- Create a Research and Development Wing: India should think about creating a specialized research and development organization akin to DARPA, which has led the way in developing deepfake detection technology.
- Through two overlapping programs, Media Forensics (MediFor), which terminated in 2021, and Semantic Forensics (SemaFor), the US Department of Defence (DoD)'s avant-garde research and development arm, made significant investments in detection technology. The goal of these projects was to create cutting-edge technology that could identify deepfake pictures and videos.

Laws regulating Deepfakes in India

- The Ministry of Electronics and Information Technology directed the major social media intermediaries in its most recent advisory, dated November 07, 2023, to make sure that reasonable efforts are made to identify misinformation and deepfakes, and that due diligence is exercised in order to identify information that violates user agreements, rules and regulations, and other relevant laws. Such cases will be dealt with promptly, well within the timeframes specified under the IT Rules 2021.
- Make sure that prompt action is taken, well within the deadlines set in the IT Rules 2021, and that access to the material and information is disabled.
- Remove any such content when reported within 36 hours of such reporting
- The Information Technology Act, 2000

- It also reiterated that failure to comply with relevant provisions of the Information Technology Act, 2000 (hereinafter referred to as the "IT Act") and Rules, Rule 7 of the Information Technology Rules (Intermediary Guidelines and Digital Media Ethics) Code, 2021 (hereinafter referred to as the "IT Rules") would result in organizations losing the protection afforded by Section 79(1) of the IT Act.

- Online intermediaries are not liable for any third-party information, data, or communication connection that they host or make available under Section 79 (1) of the IT Act. The Indian Penal Code's provisions may be invoked in court by anyone who feels wronged, according to Rule 7 of the IT Rules.

- According to Section 66E of the Information Technology Act, a person who violates their right to privacy by publishing or sending a photograph of their private region without their agreement faces a three-year jail sentence and a fine of INR two lakh.

- Specifically, posting or distributing obscene content, material involving sexually explicit acts, and minors represented in sexually explicit acts in electronic form are prohibited and punishable under Sections 67, 67A, and 67 B of the IT Act.

- Social media firms have been encouraged to take action within 24 hours of receiving a complaint about any content involving electronic impersonation, including the use of electronically modified photos of individuals. Because of this, Section 66D of the IT Act punishes anybody who uses a communication device or computer resource to deceive someone by impersonating them with a fine of up to one lakh rupees and a prison sentence of three years.

Several other provisions are available to use against the offender, including forgery (section 453), criminal intimidation (sections 503 of the IPC and 3(1) (r) of the Scheduled Castes and Scheduled Tribes (Prevention of Atrocities) Act, 1989), and impersonation (section 416), particularly when the purpose of the deepfake is to harm the victim's reputation or extract money from them. Given their distinct goals and purview, these laws would need a liberal interpretation of their constituents as well as a gender-sensitive and survivor-centric viewpoint. Rather from being motivated by a concern for sexual privacy, these measures are fundamentally predicated on acknowledging the unique aim of the offense. Thus, on their own, these intent-based offenses fall short of fully addressing all potential driving forces behind the offense, including the desire for personal fulfilment as well as motives those overlaps and interlock. Re-evaluating criminal law provisions and judicial techniques is necessary in India as well, as it recognizes the fundamental harm caused by the issue of deepfakes and allows for appropriate adaptations for the victim.

Global Regulation

The Bletchly Declaration – A group endeavour with a cooperative mind-set

The US, Canada, Australia, China, Germany, India, and the European Union are among the twenty-nine nations that have thrown their support behind the effort to stop the "catastrophic harm, either deliberate or unintentional" that comes with the growing use of artificial intelligence. The Declaration outlines a path forward for international cooperation on the current and future challenges posed by artificial intelligence.

China, the Ministry of Industry and Information Technology, the Ministry of Public Security, and the Cyberspace Administration of China emphasized in January 2023 that in order to avoid public confusion, the deepfakes must be identified explicitly.

The United States has pushed for the creation of a task group by the Department of Homeland Security (DHS) to handle "deepfakes," or digital content forgeries. Numerous states have passed laws of their own to stop deepfakes.

The government of the United Kingdom intends to present national guidelines for the artificial intelligence sector, which will assess the introduction of laws requiring explicit labelling for images and films created by AI.

The Digital Services Act, enforced by the European Union, requires social media platforms to comply with labelling requirements, improving transparency and assisting users in ascertaining the legitimacy of media.

Legislation was approved in South Korea that prohibits the distribution of deepfakes that might endanger public safety. Violators may face up to five years in prison or fines of up to 50 million won, or around 43,000 USD.

Conclusion

Women's privacy protection has to continue to be a primary concern as we navigate the rapidly changing technological world. If deepfake technology is allowed to go unchecked, it might ruin the lives of many women. It is our joint duties to push for legislation, speak out against injustice, and seek to establish a digital space where people's rights to privacy and dignity, especially that of women are protected. We can only expect to protect the integrity of privacy in the face of this enormous technical threat by working together.
It is possible to combat the negative effects of disruptive technology, such abusive deepfakes, reactively rather than proactively. Therefore, it is essential that industry players have a voice in the regulatory process, including SMIs and developers of deep learning technologies. These organizations are most qualified to offer a correct comprehension of the relevant technology at the specified moment. As a result, it will be easier to recognize, report, and determine how sceptical to approach deepfaked information. It will also help take action against abusive and unlawful deepfaked content. The digital world appears to be an ongoing arms race between security and innovation, with the bad guys, regrettably, usually coming out on top.

We need to keep improving our security procedures, making significant investments in security equipment, and spreading the word about these complex frauds. Together, we can create a more secure online environment where people can trust one another without worrying about dishonesty.The existing legal framework in India does not adequately handle cyber offenses resulting from the usage of deepfakes. It is challenging to adequately govern the use of artificial intelligence, machine learning, and deepfakes due to the absence of explicit regulations on these topics in the IT Act, 2000. It could be essential to amend the IT Act, 2000 to include sections that expressly address the use of deepfakes and the penalties for their abuse in order to properly control offenses caused by them. This would entail tougher legal safeguards for people whose likenesses or photos are exploited without their permission, as well as harsher punishments for those who produce or disseminate deepfakes with nefarious intent.

It is also critical to remember that the creation and use of deepfakes is a global problem, and that effective international coordination and collaboration will probably be needed to control their usage and stop privacy infractions.While this is going on, it is critical that people and organizations understand the possible dangers of deepfakes and exercise caution when confirming the legitimacy of content they come across online. In

order to effectively regulate deepfakes in India, methodical changes must be implemented in order to manage this potent and dangerous technology. It is critically necessary to adopt a strong, comprehensive deepfake legislation that addresses consent, privacy, redress, electoral, and criminal concerns. Important measures include mandatory labeling, restrictions, and fines for malevolent deepfakes. Guidelines for verifying video evidence are necessary for law enforcement, courts, and the media to prevent injustices from ccurring.

Social media companies should be required to quickly identify and eliminate dangerous deepfakes. To confirm the legitimacy of material, they can use digital fingerprinting based on blockchain. On the other hand, high liabilities can encourage over-censorship, necessitating safety measures.Government and media public awareness initiatives are essential for educating the public about the dangers of deepfake information. Education in media literacy, particularly for young people, may produce critical, educated digital citizens. To combat misinformation risks, fact-checking networks and cybersecurity need to be improved.

References

Bhaumik, Aarathrika (2023, December, 04). Regulating deepfakes and generative AI in India –Explained. The Hindu

Narrain, Siddharth (2020).The Information Technology Act: A User's Guide to India's Digital Legislation. Point of View, Issue XXXIX

Saxena ,Tanisha (2023 November,26).The deepfake gender paradox in AI's moral maze. Deccan Herald

Rana,Vikranth, Gandhi Anuradha and Thakur Rachita(2024 November,23). Deepfakes and breach of personal data - a bigger picture.Live Law

Chouhan Ashish(2021 September,8). Ahmedabad:Deepfakes replace women on sextortion calls.The Times of India

Pant Bhumika (2024 November,23).Gendered Deepfakes:The Precarity of Women's Embodiment in a Partriarchal World. Feminism in India

Citron, Danielle (2018 July ,14). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security.California Law Review of Texas Law, Public Law Research Paper No. 692

Dasgupta, Binayak (2020 February, 20). BJP's deepfake videos trigger new worry over AI use in political campaigns .Hindustan Times

Porup, JM (2019 April, 10). How and why deepfake videos work — and what is at risk.CSO Online

Jain, Simran and Jha, Piyush (2020 May ,21).Deepfakes in India: Regulation and Privacy. South Asia @ London School of Economics

Bajaj, Vikas (2019 November ,09). Revenge Porn' Crime Is Escalating Dramatically in India. The New York Times

The Information Technology Act, 2000, Section 66E.

The Information Technology Act, 2000, Section 67.

The Indian Penal Code, 1860, Sections 499-501.

The Indian Penal Code, 1860, Section 505.

The Indian Penal Code, 1860, Section 354C.