A Comprehensive Review of the Legal Challenges Posed by Deepfake Technology

Neha Nataraj[1] and Aroon Manoharan[#]

[1]Lambert High School, USA
[#]Advisor

ABSTRACT

Deepfake technology has steadily evolved since its creation in 2017 - the term refers to a form of technology that creates realistic digital falsifications using deep learning. This paper examines the impact the inception of deepfake technology will have on a legal basis as well as future implications on society. Due to the nature of deepfake technology, it is becoming increasingly accessible, raising ethical and legal concerns. Specifically, the technology can and has been misused in the context of nonconsensual pornography, misinformation, and threats to political stability. Taking this into consideration, the implementation of legislation is essential and has been attempted at both the federal and state level, and yet substantial gaps remain in the legal framework. Much of this paper explores current legislation including the Communications Decency Act and various attempts to address deepfake-related consequences. In addition to this, it emphasizes the societal implications and proposes a collaborative approach to developing a policy that safeguards against many of the dangers posed by deepfakes.

Introduction to Deepfake Technology

The term “Deepfake" first arose in 2017 when a Reddit user named "deepfakes" created and published celebrity porn by superimposing the face of a celebrity on an adult video using “face swapping” technology. (Reddit has since banned deepfake porn). It is now a portmanteau of the two words “deep learning” and “fake”. It refers to a type of machine learning based on artificial neural networks. Multiple layers of processing are used to extract progressively higher-level features from data, eventually resulting in a video indistinguishable from its genuine counterpart. Both in terms of technological sophistication and ease of use, the software used to create deepfakes has improved tremendously. ‘How to’ tutorials for creating deepfakes are easily found on the internet. There is even an underground industry that offers to create deepfake videos on demand. According to the estimates by Kaspersky experts, prices per minute of a deepfake video may range from $300 to $20,000. (1) And Chinese Tech giant Tencent has announced a new service that lets anyone make a high-definition digital human of any individual with just three minutes of live-action video and 100 sentences of voice material. The Deepfake-as-a-Service (DFaaS) costs $145 (1,000 yuan) and takes just 24 hours to create. The service is available in both Chinese and English. (2)

Societal and Individual Impacts of Deepfakes

The proliferation of deepfakes has serious implications for individuals as well as societies and nations. The largest application of deepfake technology is “revenge porn”. (3) And women and LGBTQ persons are disproportionately victimized. In 2019, Democratic Representative Katie Hill of California resigned from office after the conservative website RedState published sexualized nude images without her consent. (4)

In addition to videos, even speech can be forged using deepfake technology. One top-quality audio fake even sufficiently fooled an employee of a multinational corporation that transferred $35 million to a scamster. (5) In addition to targeting individuals, deepfake technology could be used to spread disinformation to the detriment of politicians and governments. It is conceivable that deepfake technology could play a disruptive role in future elections.

Legal Recourses for the Victims of Deepfakes

State-Level Legal Protections

At the state level, as of April 2023, 48 states and Washington D.C. prohibit the distribution or production of nonconsensual pornography. Nonconsensual pornography refers to the distribution of sexual or pornographic images of individuals without their consent. The two remaining states without such a law are Massachusetts and South Carolina.

Federal-Level Legal Protections

In March 2022, as part of the Violence Against Women Act Reauthorization Act of 2022, Congress prohibited the production or distribution of nonconsensual pornography. Victims can file a federal lawsuit against a person who disclosed intimate images without the individual's consent. The victims of revenge porn can also seek relief by suing the perpetrator of the deepfakes. The suit may allege defamation, invasion of privacy, commercial exploitation of one's likeness, and violation of copyrights. The law will likely be tested by advocates seeking to protect the First Amendment rights (freedom of speech) of the offender.

Much is still unknown about how successful various causes of action will be in deepfake space. Courts have yet to decide how to deal with this growing problem or the recent federal and state legal prohibitions. In the meantime, a great deal of uncertainty exists as to how those affected by deepfakes can restore their damaged reputations. (7)

Challenges Posed by the Communications Decency Act

One complication is the Communications Decency Act, a federal law passed in 1996 regulating pornography on the internet. It protects websites and service providers from liability for content that they did not co-create but was posted by their users. According to Section 230 of the Act, operators of internet services and websites are not considered publishers of content their users post and are not liable for content that is posted by others.(6) As such, websites and service providers have no legal obligation to remove nonconsensual pornography unless it otherwise violates copyright or federal criminal laws. Hence, the platform providers (like Twitter, WhatsApp, etc.) are legally protected.

Legislative Efforts to Address Deepfake Threats

At the federal level, the Congress is certainly conscious of the fact that deepfakes have the potential to cause much harm. Efforts have been made by both the House of Representatives and the Senate to curb the use of deepfakes. In the 116th Congress, the House of representatives proposed a Bill called "The Deepfakes Accountability Act - Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019." The bill was referred to a committee but was not passed by the House. The same Bill was reintroduced in the 117th Congress in August of 2021.

This bill establishes requirements for advanced technology that can create false personation records i.e., deep fakes, and establishes criminal penalties for related violations. Specifically, it requires producers of deep fakes to generally comply with certain digital watermark and disclosure requirements. It establishes new criminal offenses related to (1) the production of deep fakes which do not comply with related watermark or disclosure requirements, and (2) the alteration of deep fakes to remove or meaningfully obscure such required disclosures. Violators are subject to a fine, up to five years in prison, or both. It also establishes civil penalties and permits individuals to bring civil actions for damages. The bill also directs the Department of Justice to take certain actions, such as publishing a report related to deep fakes that includes a description of the efforts of Russia and China to use technology to impact elections. Software manufacturers who reasonably believe software will be used to produce deep fakes must ensure it has the technical capability to insert watermarks and disclosures. Finally, the bill directs the Department of Homeland Security to establish a task force to, among other things, advance efforts of the federal government to combat the national security implications of deep fakes." (8)

The US Senate for its part saw the risks associated with deepfakes and introduced a bill. The U.S. Senate Committee on Homeland Security and Governmental Affairs voted unanimously on 29th July 2021 to advance the Deepfake Task Force Act, reporting it favorably to the Senate. It would establish a public-private team charged with investigating policy and technology strategies for curbing the harms of deceptively used deepfake technology. The proposed legislation would examine technology- and policy-based approaches to detect and combat maliciously deployed deepfakes. This marked yet another attempt to legislate against the controversial technology.

However, the bills introduced both in the House of Representatives as well as the Senate never become law. Although it is heartening to note that Congress is aware of the dangers posed by the deepfakes to individuals, such awareness has not yet resulted in the laws necessary to safeguard individual's privacy and reputation. It is hoped that we will see specific legislation enacted to protect individuals from the harmful effects of the deepfakes.

Potential Solutions and Ethical Considerations

With the establishment of adequate preventative regulations, however, the mass exploitation of this technology may be successfully curbed. Implementation of a deepfake containment policy framework should involve a collaborative effort among various stakeholders, including government entities, technology companies, researchers, civil society organizations, and the public. The policy framework should also be regularly evaluated and updated to keep pace with technological advancements and changing societal needs.

Many ethical considerations surround deepfake technology. One of the most reliable methodologies to ensure a deepfake is entirely ethical is to ensure each party involved has granted full consent to the creation of the video. However, this is not always possible. For example, when the subject of the video is deceased, there is no way to ensure they would have consented. In situations where consent is lacking, ensuring the subject is respected in the video is similarly quite crucial. Legislation that addressed these primary concerns as well as ensuring all deepfake technology is clearly labeled as deepfake technology would significantly reduce the risks of exploitation.

Conclusion

In sum, deepfakes themselves are not automatically negative or unethical. They can be used for entertainment, artistic expression, or educational purposes. The controversy lies in the potential for misuse and the need for responsible use, transparency, and safeguards to mitigate reputational and emotional harm. Therefore, with proper measures, the mass exploitation of this technology can be quite effectively curtailed with the implementation of proper prevention policies.

In our new digital age, “shocking revelations” must be met with skepticism. The old social and legal limits to public conduct have given way, one hopes, only temporarily. As wary consumers of this new media, we must seek the truth before judgment rather than remain convinced that our first judgments must be true.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Bandara, Pesala. “Chinese Company Lets You Make a Deepfake “Digital Human” for $145.” PetaPixel, 1 May 2023, petapixel.com/2023/05/01/chinese-company-lets-you-make-a-deepfake-digital-human-for-145/#:~:text=On%20Friday%2C%20Tencent%20confirmed%20that. Accessed 27 May 2024.

Brewster, Thomas. “Fraudsters Cloned Company Director’s Voice in $35 Million Heist, Police Find.” Forbes, 14 Oct. 2021, www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/?sh=601a23ca7559. Accessed 27 May 2024.

Kaspersky Team. “Deepfake Market Analysis.” Usa.kaspersky.com, 11 May 2023, usa.kaspersky.com/blog/deepfake-darknet-market/28308/. Accessed 27 May 2024.

“Nonconsensual Pornography (Revenge Porn) Laws in the United States.” Ballotpedia, ballotpedia.org/Nonconsensual_pornography_(revenge_porn)_laws_in_the_United_States#:~:text=As%20of%20April%202023%2C%2048. Accessed 27 May 2024.

November 07, et al. “Katie Hill, Deepfakes, and How “Political Risk” Is Defined - Women’s Media Center.” Womensmediacenter.com, 7 Nov. 2019, womensmediacenter.com/news-features/katie-hill-deepfakes-and-how-political-risk-is-defined.

Oliveri, Jessica. “Internet Providers Still Enjoy Broad Immunity from Liability for Content.” McLane Middleton, 1 June 2023, www.mclane.com/insights/internet-providers-still-enjoy-broad-immunity-from-liability-for-content/#:~:text=Congress%20codified%20the%20CDA%20in. Accessed 27 May 2024.

Pattison-Gordon, Jule. “Senate Committee Advances Bill to Create Deepfake Task Force.” GovTech, 6 Aug. 2021, www.govtech.com/security/senate-committee-advances-bill-to-create-deepfake-task-force.

Pepper, Carolyn, et al. “Reputation Management and the Growing Threat of Deepfakes.” News.bloomberglaw.com, 9 July 2021, news.bloomberglaw.com/us-law-week/reputation-management-and-the-growing-threat-of-deepfakes.

Wang, Chenxi. “Deepfakes, Revenge Porn, and the Impact on Women.” Forbes, 1 Nov. 2019, www.forbes.com/sites/chenxiwang/2019/11/01/deepfakes-revenge-porn-and-the-impact-on-women/?sh=18410c1c1f53. Accessed 27 May 2024.