Polina Peron
CST 300
October 1, 2023

Deepfakes and Their Effects on the Web

Deepfake is a term that has gained popularity in the recent years. A portmanteau of “deep learning” and “fake”, they are manipulations of digital media through deep neural networks. Such technology allows for a person’s features, such as their face or voice, to be transferred to a destination video or audio. This process makes an impression that a person, whose features were transferred, is doing something that they have never done. This technology is achievable through machine learning algorithms that categorize words by mouth shapes (Hancock & Bailenson, 2021). Internet users create deepfakes for various motives, both good and bad. Some of the beneficial uses are entertainment, accessibility, privacy, education, and autonomy. Some of the malicious reasons are fraud, revenge porn, hoaxes, fake ransoms, and provocations. The polarizing nature of deepfakes has been the ground for controversy. Synthetic media can be very convincing, and many users have trouble recognizing an authentic video from a modified one, which might result in dangerous outcomes. Some people may question the legality of deepfake technology and whether such software should be limited to the general public.

Stakeholder Analysis

Stakeholder 1: Creators and Educators

Many people deeply value the integration of technological advancements in their daily lives. For some, it may make their jobs easier or save them money. Deepfakes can help creators with outreach, teachers can use them to bring historical figures to life, storytellers can visualize their message at low cost using tools like Midjourney, and people who have lost their voices due to medical conditions can hear them again. Aside from the more specific uses that pertain to work facilitation, internet users often see deepfakes used as a means of creating memes or comedic videos that feature the voices of famous people.

Deepfakes can be used as an aid to create immersive learning experiences. For example, in a museum dedicated to artist Salvador Dali in Florida, a deepfake of Dali appears on many interactive screens across the museum (Lalla et al., 2022). This helps engage the public with the museum material, and visitors also have the opportunity to take a selfie with the surrealist. People who could take advantage of this are educators. Bringing historical figures to life can be very engaging for young audiences and could also pique interest in tech in students.

For patients with amyotrophic lateral sclerosis (ALS), the ability of deepfake technology to emulate their voice can be invaluable. ALS negatively affects a person’s ability to speak. One famous person with ALS, Stephen Hawking, had an Intel computer that he would type words on, and then the machine would read the words using a speech synthesizer (Medeiros, 2015). While the technology that Hawking was using is now more available, deepfake software can make the experience more personalized for people with similar conditions (Jaiman, 2022). According to Jaiman, the company specializing in providing all the necessary equipment for people with ALS is developing solutions involving synthetic media.

Deepfake technology can benefit those who need their confidentiality. Since it is easy to swap faces and voices, users and creators can efficiently hide their identities if needed. This makes deepfakes a powerful tool for exercising one’s free speech. In addition, people can use deepfakes for outreach. In Jaiman’s (2022) article, they mentioned how David Beckham, a popular soccer player, recorded a video message that was translated into multiple languages with the use of AI, and his lips were superimposed to make it appear as if he was uttering those words himself.

Stakeholders in favor of deepfake technology use claims of value since they emphasize the reasons the software should stay accessible and how it can benefit the public. The points in favor of deepfakes mention positive outcomes, and they affect the lives of people who use them in a good way. Even though some use the technology perniciously, the benefits outperform the cons for the supporters.

Stakeholder 2: Victims and General Public

Those who act against deepfakes emphasize how this AI technology may impact the lives of innocent people in negative ways. There are many ways to use new AI features irresponsibly, and synthetic media is no exception, with a multitude of people falling victim to scams, misinformation, and shaming. The stakeholders value transparency and authenticity of information, along with the awareness of the dangers AI can bring.

One of such dangers is deepfake porn videos, which can be strikingly convincing. Such use is disproportionately harmful to women - according to Karen Hao of MIT Technology Review (2021), 90% of all nonconsensual deepfake porn featured women as the superimposed partaker. A user can superimpose someone’s face and voice on an existing pornographic video and use it to manipulate their victim.

Another use of deepfakes is in spreading political misinformation through social media. Some may superimpose the influential figures’ faces or lip movements to fit the words being spoken with their modified voice. Even if people are educated about the ways deepfakes can be identified, scientists found that superimposed videos can force false memories in tested individuals, especially if combined with the use of virtual reality (VR) (Hancock & Bailenson, 2021). False memories can affect the way a person thinks of certain figures, and, if viewed from a political perspective, could potentially affect how a person votes in elections.

As Faith Karimi of CNN wrote in her article (2023), on average, an American family loses $11,00 in fake kidnapping scams, and in 2022, Americans lost $2.6 billion in imposter scams (not exclusively related to kidnapping). In the article, Karimi writes about a mother who got a call asking for a ransom for her daughter worth a million dollars and suspected that the callers used AI to emulate her daughter's voice. "A mother knows her child," Jennifer DeStefano said later. The woman did not send the money since the police dispatcher recognized it as a scam, but not many people can tell a real ransom from a fake. It is suggested not to post much personal information on social media, especially about upcoming vacations to avoid becoming a victim. The scammers, supposedly, used software that drew samples of DeStefano's daughter's voice from sources like social media, then used it to create a recording of her asking for help.

Those who are not in favor of the wide availability of deepfake software are using claims of cause and policy. They use a claim of cause when arguing the effects of improper use of deepfakes such as ruined reputation, misleading, scamming, and more. They are using a claim of policy when arguing in favor of banning the software for general use as well as penalizing those who spread deepfake material intended for malicious use. The claim of policy applies here because some of the ways deepfakes can harm people are on the level of jeopardizing their safety.

Argument Question

The issue is polarizing, and the two sides arguing pro and against the availability of such powerful technology are quite convincing. But should it stay available for the public, or should it become privatized?

Stakeholder Arguments

Stakeholder 1 Argument: Creators and Educators

Those, who claim that deepfakes are beneficial and should stay available, use Kantian ethics to back their opinions. Immanuel Kant, a German philosopher of the 18th century, defined a good action as one that is backed by good intention and that is done because of the realization of what morally needs to be done (Bizarro, 2021). The idea behind deepfakes is artificial intelligence, which is made with humanity's technological advancement journey in mind. AI is used to accelerate problem-solving and a multitude of manual tasks so humans can be more productive. Deepfakes were initially used for entertaining purposes, but as the technology progressed, so did the potential for its misuse.

Things like assisting people with disabilities and enriching educational experiences are some of the outcomes that can apply positively to everybody. Therefore, deepfakes could be considered a moral tool that is widely misunderstood by the majority. Humans are generally driven by attempts to make another person's life better. Using deepfakes to nurture each other and help the generation grow is the only moral thing that it can be used for, and in this sense, the technology is good.

Keeping the deepfake technology available to the public will popularize it among internet users. This will encourage the skepticism necessary for recognizing and avoiding any dangers that synthetic media may bring. With availability being open, people in favor of deepfakes can enrich their educational and creative endeavors with little effort and expenses. Some will gain their independence with the ability to hide or bring back their personality online and in person.

Stakeholder 2 Argument: Victims and General Public

Stakeholders arguing the overall malicious results of deepfakes are backed by the utilitarian ethical framework, where an action is moral if it benefits the majority. The utilitarian approach emphasizes a net positive of action, and if the negative effects prevail, the act cannot be considered moral (Driver, 2014). With the overwhelming number of deepfakes produced being nonconsensual pornographic material, it is natural to deduct that the technology harms most of the people involved.

The number of victims of impostor call scams is very high, and the more sophisticated deepfakes become, the more difficult it will be for people to discern fakes from reality. And even if there was legislation in place, as Karimi (2023) mentioned, most of the calls originate from Mexico, and Americans in the southwestern United States will continue to fall victim.

The Department of Homeland Security (n.d.) has identified deepfakes as an emerging threat and has given scenarios in which they could be used to harm many people. Such scenarios include fabricating evidence in a criminal case, corporate fraud, stock market manipulation, and more. There are arguably more ways to harm with deepfakes than there are to draw benefit. By utilitarian approach, such conditions would make deepfakes a malicious technology.

By banning deepfake software for general use, stakeholders will gain a more authentic information flow on the internet. There will be little worrying about whether something someone said could be fabricated to look strikingly real. It will also become more difficult for scammers to extort money out of victims in phone and video scams.

Personal Position

Deepfakes have a lot of potential for making a difference. But the legislature is the only way to ensure the difference will be positive. Several states have put laws in place that limit the use of deepfakes based on certain conditions. For example, there is a ban in New York on creating deepfakes of deceased people for 40 years after their passing, and in California and Texas, there is a ban on creating deepfakes of politicians within 30 and 60 days of an election, respectively (Lalla et al., 2022). There are many resources online that teach users to identify deepfakes, and there should be training on how to access and use those resources to battle the spreading of misinformation. Companies can give comprehensive training to their employees, especially in positions that interact with social media, and schools should teach students the same skills in their informatics-related classes. Until appropriate practices that ensure preparedness are enforced, deepfakes will be harmful to people and the perpetrators will be able to freely utilize them for their advantage. This personal position is therefore in alignment with the victims of synthetic media.

Summary

Deepfakes are a revolutionary technology that can change people’s lives. But the changes could be both beneficial and detrimental. As of right now, the issue is dividing, and many people fear the range of possibilities that can be achieved through machine learning algorithms. Hence, certain regulations should be implemented to control improper use and maximize the beneficial effects of this innovation for all, which will come with challenges since some threats come from abroad.

References

Bizarro, S. (2021, September 21). *Kantian ethics-an introduction*. Medium. https://sarabizarro.medium.com/kantian-ethics-an-introduction-f4b95a8fae3d

Department of Homeland Security. (n.d.). *Increasing threat of Deepfake Identities - Homeland Security*. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf

Driver, J. (2014, September 22). *The history of Utilitarianism*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/utilitarianism-history/

Hancock, J. T., & Bailenson, J. N. (2021, March 17). *The social impact of Deepfakes | Cyberpsychology, behavior, and social ...* Mary Ann Liebert, Inc. https://www.liebertpub.com/doi/10.1089/cyber.2021.29208.jth

Hao, K. (2021, February 16). *Deepfake porn is ruining women's lives. now the law may finally ban it.* MIT Technology Review. https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/

Jaiman, A. (2022, August 2). *Positive use cases of deepfakes*. Medium. https://towardsdatascience.com/positive-use-cases-of-deepfakes-49f510056387

Karimi, F. (2023, April 29). *"Mom, these bad men have me": She believes scammers cloned her daughter's voice in a fake kidnapping*. CNN. https://www.cnn.com/2023/04/29/us/ai-scam-calls-kidnapping-cec/index.html

Lalla, V., Mitrani, A., & Harned, Z. (2022, June). *Artificial Intelligence: Deepfakes in the entertainment industry*. WIPO. https://www.wipo.int/wipo_magazine/en/2022/02/article_0003.html

Medeiros, J. (2015, January 13). *How intel gave Stephen Hawking a voice*. Wired. https://www.wired.com/2015/01/intel-gave-stephen-hawking-voice/