Technology: “Deepfakes”

Corey J. Wagner

Helena College University of Montana

WRIT 101: College Writing

Karen L. Henderson

November 1, 2020

Technology: "Deepfakes"

There is extensive discussion in both popular and scholarly platforms on the realm of new media configurations, contrasting opportunity to the potential risk factors of increasingly progressive technology (Kietzmann et al., 2020; Westerlund, 2019; Yaldin-Segal & Oppenheim, 2020). Today, it has been argued people live in a "post-truth" era where disinformation and misinformation are catapulted through society to manipulate and deceive public opinion. With the seemingly limitless and quick advancements we have had in technology over the past several decades, the future of the societal relationship with machines remains unclear and profound. The quick and surprising rise of what is known as "deepfakes" has shed light on new ways to alter media with less traceable factors to distinguish if the content is real or fake. This blur in the easily accepted dichotomy between credibility and fraud poses many questions related to the media's dynamic ecosystem (Yadlin-Segal & Oppenheim, 2020). "Deepfakes are the product of artificial intelligence (AI) applications that merge, combine, replace, and superimpose images and video clips to create fake videos that appear authentic" (Westerlund, 2019, p. 39). In essence, a complex neural network and algorithm structured autoencoder is initiated for deep machine learning of precise facial recognition characteristics, along with contributing factors such as mannerisms and voice, to create previously non-existent images or videos (Kietzmann et al., 2020). Since there is a huge influence from social media, deepfakes are mainly targeted at people there, allowing this type of false material to easily reach millions and millions of individuals within a short amount of time and appear genuine. This type of technology surfaced on the internet quite recently in late 2017, so scholarly research and resources to combat and monitor deepfakes are exiguous in scale (Westerlund, 2017). Although there are some potential beneficial and seemingly harmless applications for deepfakes, the seriously dangerous and malicious issues far outweigh any positives.

This development of deepfake technology, which is constantly further progressing and transforming into an almost unrecognizably accurate counterfeit, can be incorporated into many useful applications including, but not limited to, entertainment, educational media, business, and social and medical fields. For example, we have seen this type of technology grow in computer-generated imagery (CGI), mainly used in the entertainment industry where we see the artistic world of cinema come to life in a very real and engaging way. Westerlund (2019) argues, “movie makers will be able to recreate classic scenes in movies, create new movies starring long-dead actors, make use of special effects and advanced face editing in post-production, and improve amateur videos to professional quality” (p. 41). This type of labor-intensive work requires tremendous knowledge, funding, and expertise. However, it has been argued, “tools of today and certainly those of tomorrow allow anyone to create fakes that appear real without significant investment in training, data collection, hardware, and software” (Kietzmann et al., 2020, p. 136) opening the door for manipulation of existing content from relatively anyone. This allows easier access to harassment, defamation of character, and bullying. This poses a threat to the credibility of virtually any person who can get caught in the crossfire of malicious, non-consensual, or unwanted content generated to appear incredibly real.

Benefits of this deepfake technology have been demonstrated in a 2019 global malaria awareness ad, where well-known athlete David Beckham was able to disassemble language barriers through voice altering, thus appearing multilingual (Westerlund, 2019). “Similarly, deepfake technology can break the language barrier on video conference calls by translating speech and simultaneously altering facial and mouth movements to improve eye-contact and make everyone appear to be speaking the same language” (Westerlund, 2019, p. 41). In the medical field, there are strong possibilities to temporarily digitally recreate the life of the deceased, allowing closure for a grieving loved one who missed an opportunity to say goodbye or provide someone who is transgender to view an image of themselves as a preferred gender (Westerlund, 2019). These types of applications could improve our overall quality of life, as well as provide better-established connections to each other across the globe and help unify us as people through easier and more efficient communication. However, the flip side of using this deepfake technology, as we have seen overzealously utilized by people in a position of power, can be far more insidious and catastrophic in scale.

On a personal level, consider the trigger reactions of someone being the target of rancorous propaganda from receiving a deepfake of a spouse romantically involved with that spouse’s best friend (Yaldin-Segal & Oppenheim, 2020). Contemplate the various scenarios that could result from an action such as this. The reputation of a company could be put into question because the CEO was put in a compromising situation where stockholder agreements could be at risk or reports of false earnings caused stock values to unnecessarily go down (Kietzmann et al., 2020). This is just another one of the other numerous ways many people could also be victimized or experience unfair financial hardships. Moreover, this provides an opportunity for well-timed acts of sabotage, such as seeing a president of a country trying to cover up misdeeds, making racist or other questionable and inappropriate remarks (Kietzmann et al., 2020), or perhaps even declaring war on another country. The potential fallout, panic, and fear that could ensue solidify the argument that “deepfake propaganda and the disinformation they seed threaten efficient governance for all democracies if not democracy itself” (Kietzmann et al., 2020, p. 143). Furthermore, it has been argued:

> like ‘fake news,’ it will become a shield for liars and conspiracy theorists, used to dismiss any evidence that runs counter to their own beliefs. In the age of AI, the next ‘grab them by the pussy’ video will be even more easily shrugged off as a fake under a miasma of reasonable doubt. This breeds a sort of media nihilism, a belief that no audiovisual content can ever be definitively said to be ‘real’. (Yaldin-Segal & Oppenheim, 2020, p. 11)

Societal concerns over this issue are such that the seriousness for potential harm and mistrust given to the gatekeepers of influential institutions such as journalism make it impossible to adequately distinguish a shared reality and attacks people’s emotional security.

The impact of this issue is significant because people put a lot of trust into technology through photo, audio, and video evidence regarding any relevant topics of discussion. People’s faith in the brain’s visual system, despite its value and complexity, is flawed; this is confirmed from sleight-of-hand tricks or simple visual aids like “optical illusions and bistable figures such as the well-known Jastow rabbit-duck and Rubin vase-faces that can be viewed in two different ways” (Kietzmann et al., 2020, p. 136). Believability and accessibility of these deepfakes mark the salient and necessary turning point of societal and technological countermeasures to establish a concrete defense to monitor proper use of this technology. Undoubtedly, society is in the age of unbelievably powerful technological advancements in all categories, with the core of this shift being media evolving to an instantaneous global reach, having the capacity to distribute individually directed propaganda at alarming rates (Yaldin-Segal & Oppenheim, 2020). Successful strategies we have been utilizing to avoid this potential for disaster incurring from fraudulent media have been mainly technological advancements such as “Blockchain” and other automated tools for deepfake detection. It has been argued that these applications are “highly resistant to forgeries and can store data in an accruable, safe, transparent, and traceable way, but it can also track and certify the origins and history of the data” (Westerlund, 2019, p. 48). To prevent and mitigate potential malicious use of deepfake technology in society, we can also incorporate the idea of the “R.E.A.L. framework” to help manage fraud risks. It has been suggested, “for people in a position of power such as policymakers or social media platform leaders, it is encouraged to record original content, expose deepfakes early, advocate legal protection, and leverage trust” (Kietzmann et al., 2020, p. 144) to cohere to this agenda and help eradicate this issue. Technology is always going to be improving and forever changing, with new ways of thinking and living giving birth to new ways of life. Through a more public understanding of this technology, we can educate and inform the public to obtain the best positive outcome towards combatting the dark side of deepfakes, while also improving our lives with this interesting technology.

References

Kietzmann, J., Lee, L., McCarthy, I., & Kietzmann, T. (2020). Deepfakes: Trick or treat? *Business Horizons, 63*(2), 135-146. doi:10.1016/j.bushor.2019.11.006

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review, 9*(11), 39-52. doi:10.22215/timreview/1282

Yadlin-Segal, A., & Oppenheim, Y. (2020). Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence: The International Journal of Research into New Media Technology*, 1-16. doi:10.1177/1354856520923963.