Debating the Use of Deepfakes

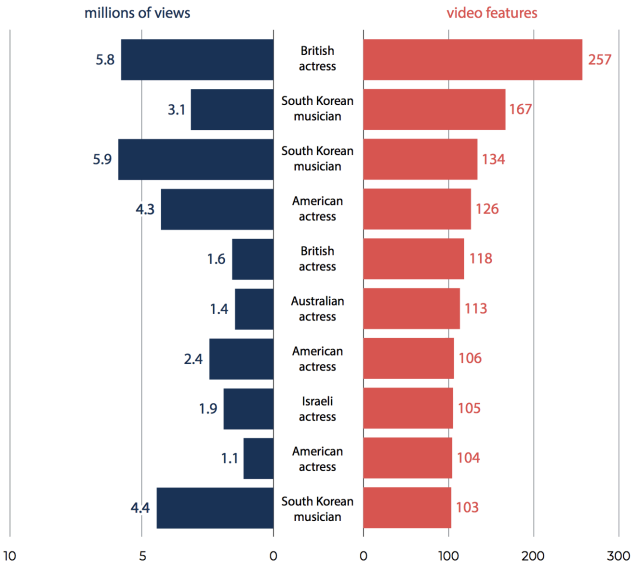
AI-generated synthetic media, a.k.a. ”deepfakes,” is a portmanteau of the words “deep learning” and “fake,” and was - of course - first coined in late 2017 by a Reddit user for pornographic video sharing purposes. Deepfake technology was invented by Ian Goodfellow in 2014, and most of the technologies are based on a branch of deep learning called generative adversarial networks (GANs). Nowadays, deepfakes have permeated many online communities, so people need to be vigilant and learn how to recognize them. They are already a notorious worldwide problem, and deepfakes can be a powerful weapon that jeopardize people’s careers and bring devastating consequences to their personal lives. People need to be concerned and know what they can do to protect themselves.

“ deepfakes can be a powerful weapon that jeopardize people’s careers and bring devastating consequences to their personal life.”

Known uses for deepfakes

The most common uses for deepfakes are generating images, videos, audio recordings or even writing in which people appear to do or say things they never did. With new AI algorithms, it’s not difficult to generate different fake faces. You can change the appearance of the fake person's image by choosing different ages, eyes, perspectives, moods, genders, races, and ethnicities.

In recent years, deepfakes are widely used for blackmail, sexual harassment, exploitation, and political purposes. They are most likely to target and harm females and celebrities to create pornography. Deepfake porn consists of taking a person’s face and swapping it with an adult performer’s body. A cybersecurity company, Deeptrace, states they have found 14698 videos involving deepfake pornography. The number of videos have nearly doubled since 2018, and 96% of videos are of the “non-consensual” genre —— a class that contains non-consensual sex acts.



Graph from The State of Deepfakes 2019. The report shows millions of views on videos targeting hundreds of female celebrities worldwide.
(source:https://medium.com/dataseries/what-strategy-does-europe-have-to-tackle-deepfakes-fb159040f0c)

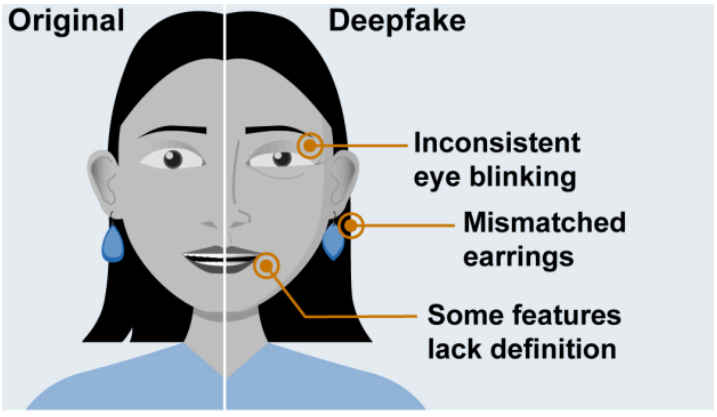
How to make deepfakes
With the continuous improvement of technology, anyone with an iPhone can make deepfakes easily. Deepfakes rely on artificial neural networks, face swapping and facial manipulation. A few years ago, people still needed to create deepfake videos on more powerful computers; now, they can create a video of any face with just a photo and app on their phone.

Avatarify and Wombo are popularly downloaded apps. Both of them can quickly turn a photo into a lip-sync video. A genealogy website called, MyHeritage, allows anyone to bring old still photos to life by simply uploading a picture of their lost friends or family members. Generated.Photos is a website that allows people to buy an imitation video of a person, and each video only costs $2.99. Rosebud.ai can even make a fake person talk.

The cyber threat
Since creating deepfakes does not require hacking skills, cybercriminals love it. Many clues show cybercriminals' ongoing targets are businesses and companies. One case shows cybercriminals successfully mimicking a chief executive's voice to transfer $243,000 to the wrong account.

Benefits of deep fakes
Despite the many malicious ways of using deepfakes, there are still some positive aspects of this technology. First, deepfakes can be used to change dialogue or video clips without reshooting and wasting money in feature film productions. Second, deepfakes can actually be used for privacy and identity protection. For example, AI-generated avatars are utilized in news reports so that interviewees can speak freely without feeling unsafe. Third, deepfake technology allows us to interact with historical figures and experience things that no longer exist.

How to protect yourself
According to a new UCL report, "Fake audio or video content has been ranked by experts as the most worrying use of artificial intelligence in terms of its potential applications for crime or terrorism." Thus, people should learn about deepfakes to protect themselves. The image below shows some tiny differences that can be detected from eyes, ears, or accessories between deepfakes and original images. Here is even more background on how to protect yourself from deepfakes.



Source: GAO; conceived from DARPA image at https://www.darpa.mil/news-events/2019-09-03a. | GAO-20-379SP

This image presents potential errors in deep fakes that fail to imitate people realistically. (source:https://www.gao.gov/products/gao-20-379sp)

Another important step in protecting yourself is to acknowledge confirmation bias. People tend to believe what they want to believe, however, this may lead to misjudgments and falling into the deepfake trap. One of the top concerns is that deepfakes will be used as a tool to influence elections, therefore, Texas and California passed laws to criminalize publishing and distributing deepfake videos to prevent campaigns using deep fakes to attack opponents.

Recently, a new deepfake spotting tool, reported by the University at Buffalo, proved to be 94% effective with portrait-like photos. Governments and some social media platforms have also begun to take deepfakes seriously. On December 20, 2019, former President Trump signed the nation's first federal law related to "deepfakes."

Remain vigilant. Deepfakes will continue to appear online, and there is no perfect solution to deal with its dangerous side so far.

-- Xuening Yang, xyang2@gwmail.gwu.edu