Deepfake: How deep does it go? (March 2021)

Author, E. Salt

***Abstract*—This paper discusses deepfaking, and the technology and ethics behind it. Deepfaking has a wide spectrum of uses and controversies, from mimicking political leaders, to de-aging or even resurrecting actors. This article will talk about how these feats can be achieved and how people feel about the use – or misuse – of this powerful technological tool. Going in-depth about the AI and deep learning neural networks that make it so prominent today.**

***Index Terms*—Deepfake, Artificial intelligence, Deep learning, Neural network.**

I. Introduction

As the world is becoming more digital, more and more complex technologies are becoming increasingly common and easily accessible to the general public [1]. Be it by simplified coding processes or, as it often happens these days, by an app. Facial editing has been one of the most recent trends to take off, first by the still popular face-swap, then by the more fully fledged deepfake.

But what is a deepfake, and how does it differ from other image modifiers like face-swapping or classic photoshopping?

II. Literature Survey

Well for starters, it's a good deal more complicated than simply copy and pasting a face between two images. A program like Photoshop (PS) is heavily reliant on a user's skills to determine how realistic an edited image will turn out and may require a lot of practice to get a convincing result.

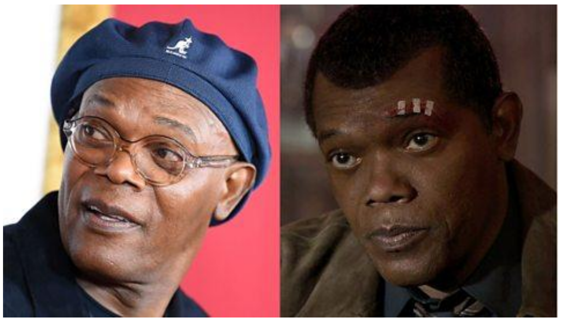

Fig. 1: Samuel L. Jackson de-aged with deepfake technology in Marvel's 'Captain Marvel'

This may start to sound like PS is more complicated than a deepfake, but this isn't the case. While a PS edit requires simply cutting out a face and moving onto another scene, a deepfake must look at all the individual features of a face and analyse its facial structure. For a person to do, this would take infinitely more time to do than a PS. This where computers and Artificial Intelligence (AI) come into play.

The AI in question comes in the form of Neural Networks (NN). These NNs function similarly to the neurons in the human brain. With user defined algorithms, they can learn to recognise patterns of raw data then group and classify it. Over time, they can continue to learn and improve their own efficiency [2]. Because of this, NNs are perfectly suited for solving complex computational problems, such as fraud detection, chemical compound identification, medical or disease diagnosis, and - of course - facial recognition.

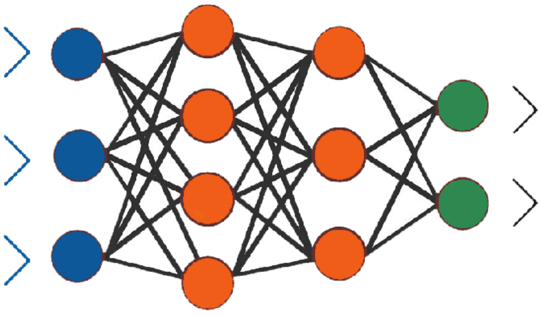

Fig. 2: A simplified example of a neural network

A usual image-orientated NN will work by providing one or multiple images for it to analyse – the more images that include the target object you wish to identify, the more accurate it can learn to be. So, by giving it a lot of images of different people's faces, it can eventually learn to identify the facial structure of theoretically anyone.

But it's not just faces that can be taught and modified. Using the same method, someone's entire body can be deepfaked onto a video of someone say, performing an intricate dance recital, giving the appearance of the person from the source image being a dance expert [3]. NN training also grants the option to deepfake a person's voice which, if done well enough, could be a dangerous tool coupled with a good quality visual deepfake.

While the possibilities of these tools can be scary if in the wrong hands, the likelihood they'll be used to create believable footage is very unlikely. Not to say it's impossible, but the current technology is far from perfect, and always contains some imperfections and telling signs to indicate it's not real. Full body faking is bound to have a lot more out of place artifacts than a simple face, and audio deepfaking is even more complex and hard to get right as, generally, people are much more sensitive to audible inconsistencies than visual ones.

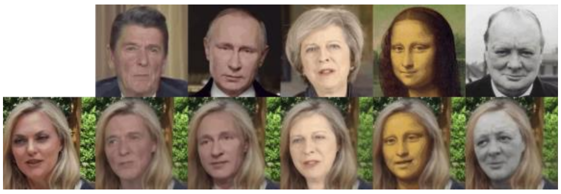

Fig. 3: Multiple deepfake examples, with each face on the top row faked onto the source image (left)

That's not the only reason they're not a threat either. Several technically minded people have found that it's possible to detect deepfaked media by reverse engineering algorithms used to create them. By recognizing patterns in how deepfakes are created these algorithms can pick up subtle errors that are typically found in deepfake content [4].

III. Discussion & Conclusion

Deepfake technology is not something that will be disappearing any time soon, and it will only continue to improve over time. It certainly has a lot of questionable uses, but there are no laws on how it's allowed to be used yet, though this could change in the future. It cannot be argued that it has it's uses, and it can certainly be a fun gimmick for people to play around with.

While deepfakes are an interesting tool and have many practical uses for entertainment, they are something that should be watched closely as they have the potential to be used for deceitful purposes when they finally advance to hyper-realistic levels.

References

[1] BBC Bitesize, "Deepfakes: What are they and why would I make one?", BBC, London, 2019 (according to the WayBackMachine's (WBM) earliest record: https://web.archive.org/web/20190802082211/https://www.bbc.co.uk/bitesize/articles/zfkwcqt)

[2] SAS Insights, "Neural Networks – What are they and why do they matter?", SAS, North Carolina, 2018 (according to WBM's earliest record: https://web.archive.org/web/20181116063805/https://www.sas.com/en_us/insights/analytics/neural-networks.html)

[3] Aliaksandr Siarohin , Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci and Nicu Sebe, "First Order Motion Model for Image Animation", NeurIPS, San Diego, 2019

[4] Kara Manke, "Researchers use facial quirks to unmask 'deepfakes'", Berkeley News, Berkeley, 2019

[x] Author, "Title", Company, City, State/County, Rep xxx, Year