JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ERA OF DEEPFAKE TECHNOLOGY: THREAT OR AIMBLE

[1]Aishwarya Gupta, [2] Megha Gupta, [3] Garuna Chauhan
[1,3] Student, [2]Faculty
[1,2,3]Department of Computer Science,
[1,2,3]Mata Sundri College For Women, University of Delhi, Delhi, India

Abstract : Deepfakes are videos that use someone else's likeness to replace persons in existing footage. Currently, the majority of their influence is restricted to pornography, and they are also used to malign people. Although it appears to be a fun approach to create phoney movies or images of anything or some persons, it has the potential to propagate disinformation over the internet. Deepfake material might be harmful to people as well as our communities, companies, governments, faiths, and so on. Because Deepfake material is created with a high degree of competence and a mix of numerous deep learning algorithms, it appears virtually real and authentic and is difficult to distinguish. Face identification, multimedia forensics, watermarking, and convolutional neural networks are some of the approaches used to identify deepfakes (CNNs). To identify any type of modification in photographs and videos, each approach employs machine learning, a technology from the field of artificial intelligence. This paper presents the concept of Deepfake and how Deepfake is a threat to our society. This paper also discussed the measures and strict regulations that could prevent the adverse effect of it.

IndexTers **- DeepFake, Face Detection, Photos/Videos manipulation**

I. INTRODUCTION

Deepfakes are a new type of synthetic media that first appeared in the globe in 2017. False information travels swiftly on social media, affecting millions of people. Currently, YouTube is the second most popular source of news for one out of every five Internet users, trailing only Facebook. While spreading false information is straightforward, repairing the record and combating deepfakes is more complex. To combat deepfakes, we must first understand what they are, why they exist, and the technology that powers them. Deepfakes will very certainly be utilized in the future for revenge porn, bullying, forged video evidence in court, political sabotage, terrorist propaganda, blackmail, market manipulation, and false news. Deepfakes are synthetic media created by artificial intelligence in which a person in an existing image or video is replaced with someone else's likeness as shown in Fig 1. While the act of falsifying information is not novel, deepfakes employ advanced machine learning and artificial intelligence techniques to modify or synthesize visual and audio content with a high potential for deceit.
Deep learning is used to build deepfakes, and basic machine learning methodologies involve training generative neural network architecture.

]
Figure 1. Examples of Deep Fake [1]

While the capacity to automatically swap faces to generate genuine and realistic-looking synthetic video has some intriguing early uses, this is clearly a risky technology with potentially disturbing implications. In fact, one of the earliest real-world implementations of deepfakes was to generate synthetic pornography. It was invented in 2017 by a person with the same name on Reddit. This person established an area on the online news and aggregation site where they published pornographic movies that made advantage of open-source face-swapping technologies. The main purpose of deepfake is to influence viewers and listeners to believe something that did not happen. That is why it is mostly used in movies for a creative effect when the entertainers are not available themselves.

When applied correctly, this approach enables the creation of lifelike films at an extremely cheap cost. For example, the end of Rogue One contained a digital recreation of the irreplaceable Leia; a highly costly moment that required the knowledge of many individuals. Figure 1 depicts a comparison of the original scene and one produced using deep learning.

Deepfakes rely heavily on machine learning. It has facilitated the development of deepfakes at a considerably faster and lower cost. To create a deepfake video of someone, a developer must first train a neural network using hours of genuine video footage of the subject. It provides it with a realistic "knowledge" of how he or she appears from different angles and lighting circumstances. After that, the trained network is coupled with computer graphics methods. This technique superimposes a duplicate of the individual onto another actor as shown in Fig 2. Fig 3 is representing the flow-chart of the working of the Deepfake technology.

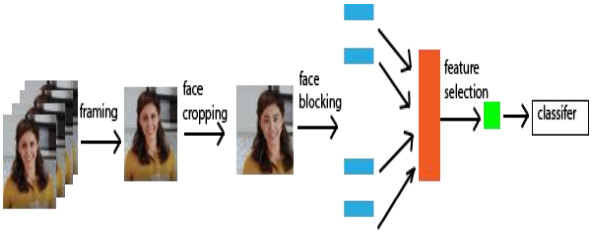

Figure 2. Superimposing of Pictures

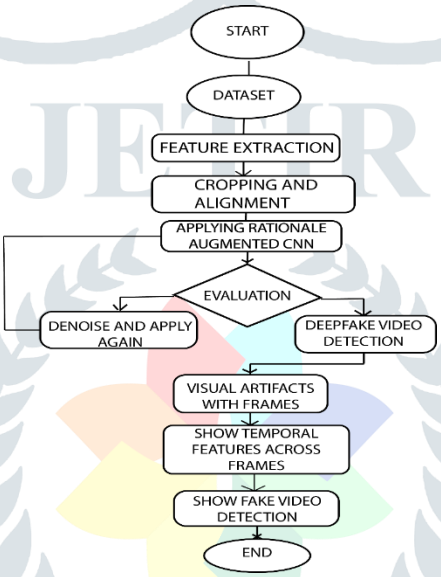

Fig 3. Working of Deep Fake

II. LITERATURE WORK

The authors in work [1] discussed Deepfake and its cyber issues. The author in work [2] investigates what deepfake is, who manufactures it, and what the pros and drawbacks of this technology are. They presented a deepfake evaluation with cybersecurity and AI entrepreneurs with economic potential in combating media forgeries and fake news.

In paper [3] authors discussed how this technology can help realize epistemic right and inflict epistemic wrong of any sort. Author has also discussed the various effects of deepfake. Author in work [6] presented the history and origin of the deepfake. They also discussed how deep fake videos and audio were produced. They also look at how deepfake films and audio are made, as well as the features of photos and videos that change during deepfaking.

III. EFFECT OF DEEPFAKE ON SOCIETY

This section presents various examples of Deepfake which shake the society. Deepfake are also a cause for concern at a time when WHO has started that the Covid-19 crisis has triggered an info Demic and there have been intentional attempts to distribute incorrect information in order to undermine public health responses and alternative objectives of groups or people.
Last year, before the Delhi assembly polls, videos of Delhi BJP President Manoj Tiwari speaking in English and Haryanvi went viral [4]. In these videos, Tiwari was seen criticizing Arvind Kejriwal and asking people to vote for BJP. The videos which were shared in over 5,000 WhatsApp groups, were later revealed to be deepfake. This is not a deepfake, but it demonstrates why their potential abuse has become so feared in politics.
Nancy Pelosi, the speaker of the United States House of Representatives, video was slowed down by 25% and the pitch was changed to make it appear as if she was mumbling her words [5]. "The pattern seems strange." Facebook first declined to remove the video, but later stated that it had been limited in circulation once it was found to be fraudulent. The message was later deleted;

however, it could not be known who deleted it. The example exemplifies the type of abuse that critics worry may be made using technology released by Stanford University in June that allows audio in a video to be changed as readily as text.

The very first instance of malevolent use of deepfake was found in pornography, causing emotional, reputational, and, in some cases, physical harm on the subject, primarily women. According to a Deepfake research, 96 percent of Deepfakes are pornographic movies, with over 135 million views on the pornographic.

Symantec reported that millions of dollars were taken from three businesses who were victims of a deepfake audio assault. During each attack, the AI-generated synthetic voice would contact a senior finance person and seek immediate cash transfer. Deepfakes models were trained on public statements given by the CEO. In February 2020, a Pennsylvania attorney was duped by his son's deepfake voice, which said he required $9,000 in bail money.

a. Good factors of deepfake

To safeguard people's privacy, Deepfake can be used to disguise the identities of their voices and appearances. Deepfakes may be used by individuals to generate avatar experiences for self-expression on the internet. Using a personal Avtar, individuals may obtain autonomy and grow their purpose, thoughts, and beliefs. This technology is contributing in various sectors e.g. entertainment, education. Deepfakes offer a wide range of possible applications in the film business. We have the ability to resurrect a deceased star or actress.

It can be contested from an ethical standpoint; however it is doable and really simple if we do not consider ethics! Also, it's probably a lot less expensive than other possibilities. Deepfakes may be a wonderful tool for authentically enacting the core themes of comedy or satire. These might be a reflection, stretching, twisting, or appropriation of real-world occurrences. Furthermore, synthetic media made by artificial intelligence has enormous promise. It has the potential to open doors in the entertainment sector. In addition, we're seeing a lot of independent producers and YouTubers take advantage of the opportunity.

Imagine a world in which you could take physics lessons from Albert Einstein whenever and wherever you wanted! Deepfake makes the unthinkable possible. Learning things from masters is a great motivator. You can improve efficiency, but there is still a long way to go.

b. Bad factors of Deepfake

Rather than helping everyone, this AI-based technology has negative consequences for many groups in our society. Deepfake is mostly used for revenge, defaming prominent personalities, in addition to manufacturing fake news and propaganda. When fraudulent films become viral, people initially trust them and continue to share them with others. This embarrasses the targeted individual, and until and until the targeted celebrity issues an official comment, many people begin to believe that they are making their life unpleasant, especially when their supporters are condemned on platforms such as social media..

- **Scamming-** Unfortunately, there are several ways to use this strong technology to commit fraud. We've lately seen several examples. Everything is accompanied with a price. This technology paves the way for scams like Scams like this are made possible by this technology. Financial frauds are another source of concern. Audio deepfakes have previously been used to clone voices and deceive people into thinking they are speaking with someone they trust. While everyone from DARPA to the FBI is working on techniques to detect and combat deepfakes, as well as legislation against them, it may be too late. That's because the fundamental problem with deepfakes is that they erode our faith in technology – our trust in TV and video news, our trust in online clips, our trust in live-streamed events, and even our trust in our own eyes.
- ***Spreading Misleading News Via Politicians*-** Nowadays, it is difficult to detect forgeries. We are so susceptible to trusting what we see or hear in the media. And if you see a public figure discussing an issue, do you generally wonder if the person is "genuine" or "fake"? When deepfake exists, however, someone with evil intents may simply influence the news. This can result in warfare, turmoil, and even famine; the usage area of the deepfake is entirely dependent on the individual.
- ***Generating Fake News*-** Have you heard the movie Spider-Man: No Way Home? Despite the fact that he has not been confirmed to be in the film, footage of actor Andrew Garfield on the set has emerged. Fans expected him to appear in the film, but there was no official confirmation from Sony or Marvel. The internet exploded once this clip was leaked.
  Everyone was debating whether or not this video was real. Some people suspected it was a deepfake, but there was no evidence of it. It was quite accurate.
- *Privacy Problems*: We've all got accounts on social media. Every one of us has a Facebook, Instagram, or Twitter account that generates a lot of data every day. We all share images from various perspectives and in various moods, which may cause some privacy issues. A deep faker with malicious intent can simply gain access to your images, capture them without your consent, and use them. You would not want to appear in a phoney video without your permission. However, it is conceivable. We require law enforcement because privacy is the most pressing issue of the twenty-first century.

IV. PREVENTION – TO STOP MISUSE OF DEEPFAKE

The amount of deepfake stuff on the internet is continually increasing. While impressive, today's AI-powered deep fake technology is still not quite on par with genuine video footage. In any case, innovation is advancing at an astounding rate. Deepfakes, according to experts, will be distinguishable from actual images in a short period of time. Deepfakes continue to have negative implications. The majority of them have obvious artifacts that separate them from their true essence. Even the more persuasive ones may be found if you search closely. In any event, it won't be long until the technology is advanced enough to trick even qualified specialists. Its deadly force will take on an entirely new dimension at that moment. As of now, a few initiatives are being made to address this looming digital media credibility issue, but they lack the clarion cry of criticality that is required to bring the matter to the forefront of public consciousness. The death of history and the collapse of trust jeopardize the advancement of civilization itself; nonetheless, the great majority of people are still hesitant to speak in such stark terms. Following are some measures that could be taken to stop misuse of Deepfake.

***Legal Measures**:* Legal safeguards are also important. There are currently no substantial measures in place to protect persons against deep faking. Imposing severe penalties on the activity will encourage specialists to create and distribute counterfeit goods, as well as function as a deterrent to improper uses of technology. In any event, these techniques may be feasible as long as humans can distinguish between phoney and authentic media. When the technology matures, it will be very impossible to prove that a particular video or audio clip was created by AI algorithms. Then, someone else may take advantage of the concerns and uncertainty around crime that was created by AI. That assurance, too, will be impossible to disprove.

***Use of Blockchain**:* The blockchain might be used as a potential solution. Blockchain is a distributed ledger technology that allows you to store data online without the need for centralized servers. Furthermore, blockchain is resistant to the vulnerabilities that centralized data storage suffers from. Distributed ledgers are not yet capable of storing a large amount of data, but they can be used for putting hashes and digital signatures.

***Train Computers to Spot Fakes**:* Using obvious antiquities analysis, it is now able to identify some of the current imperfect deepfakes. Microsoft has released a new method for detecting flaws in synthetic media. The Defense Advanced Research Projects Agency is working on a program known as SEMA, whose goal is to uncover semantic flaws in deepfakes, such as a snapshot of a guy generated with anatomically incorrect teeth or a figure wearing socially unusual jewelry.

***Don't Trust Everything You See Online*:** Remember that, just as anyone may steal images and create bogus accounts online, Deepfake technologies enable scammers to go one step farther. Before you believe anything you see on the internet, use your best judgment.

To recapitulate, there are four recognised methods for combating deepfakes: (1) law and regulation, (2) corporate policies and volunteer action, (3) education and training, and (4) anti–deepfake technology. Legislative action against some deepfake makers is possible, but it is ineffective against foreign powers. Rather, deepfake addressing content moderation policies and quick removal of user flagged content on social media platforms, as well as education and training aimed at improving digital media literacy, better online behaviour, and critical thinking, which create cognitive and concrete safeguards against digital content consumption and misuse, are likely to be more efficient. Government officials, businesses, educators, and journalists must raise citizen awareness of the challenges presented by AI to media confidence and prevent the fraudulent use of such technology for economic, political, or anti-social objectives. Technological solutions, such as automated systems for deepfake identification, content authentication, and deepfake prevention, are part of a dynamic area of security measures.

CONCLUSION

This paper discusses the ways in which deepfakes pose a significant threat to society, political systems, and businesses by putting pressure on journalists to distinguish between real and fake news, endanger national security by disseminating propaganda that interferes with elections and erodes citizens' trust in government and institutions. In this line, this paper mainly validates earlier research findings while further detailing these risks through instances of actual and hypothetical usage of deepfakes.

As a result, there are countless business prospects for technology entrepreneurs, particularly in the fields of cybersecurity and artificial intelligence. According to the report, deepfake technology is advancing at a rapid rate. Humans are rapidly losing the ability to discern between authentic and fraudulent videos. As a consequence, our findings include a plethora of indications for detecting deepfakes, and we propose using AI to detect AI-generated fakes as an effective battle approach. Having said that, the study does have certain drawbacks. First, while the empirical study included 84 online news items on deepfakes, there are many more accessible, and given the rapid growth of this technology, those pieces might potentially provide further information on deepfakes and suggest alternative techniques to combat them. Second, we collect our empirical data from publicly available sources, especially Internet news media sites. Other types of data, such as deepfake-focused online community discussions and interviews with GAN developers and deepfake artists, some of whom are well-known to the general public as a deepfake technology developers but also as anti-deepfake technology developers, could provide more information about the fight against deepfake strategies. Besides, the comment sections of some analyzed news articles contained a large number of opinions and views of readers; analyzing these comments may provide additional insights into how deepfakes are viewed by a wider audience, and thus education-focused combat methods. These constraints provide up several avenues for future study on deepfakes.

REFERENCES

1. Bahar Uddin Mahmud and Afsana Sharmin, Deep Insights of Deepfake Technology : A Review, DUJASE vol. 5(1 & 2) 13-23, 2020.
2. Westerlund, M. 2019. The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review, 9(11): 40-53. http://doi.org/10.22215/timreview/1282
3. Kerner, Catherine and Risse, Mathias. "Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology in Digital Lifeworlds" Moral Philosophy and Politics, vol. 8, no. 1, 2021, pp. 81-108. https://doi.org/10.1515/mopp-2020-0024
4. MIT Technology review, https://www.technologyreview.com/2020/02/19/868173/an-indian-politician-is-using-deepfakes-to-try-and-win-voters/
5. NDTV, Indo-Asian News Service,2019 https://www.ndtv.com/world-news/facebook-fact-checks-fake-viral-video-of-us-speaker-nancy-pelosi-2042867
6. Marwan Albahar, Jameel Almalk, Deepfakes: Threats and Countermeasures Systematic Review, Journal of Theoretical and Applied Information Technology, vol 97, no 22, November 2019.