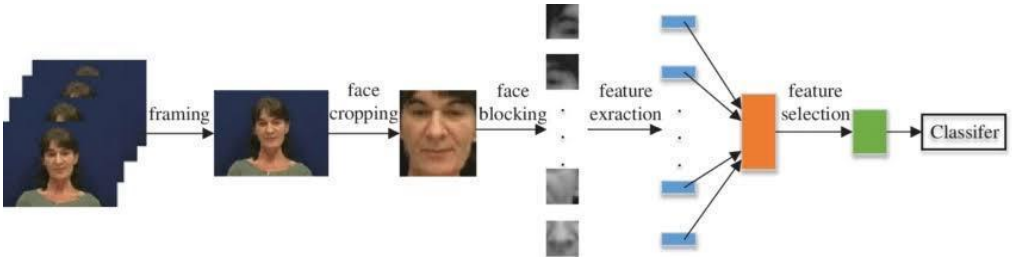
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

Research Paper on Introduction of Deepfake

Mr. Basant Yadav[1], Mr. Simranpreet Singh[2], Mr. Pankaj Yadav[3]
Department of Information Technology[1,2,3]
Currently Working as GET in Purchase Department in Dhoot Transmission PVT. LTD[1]
Dronacharya College of Engineering, Gurugram, India

Abstract: *Deepfake technology, powered by deep learning advancements and computational capabilities, is a rising challenge in the digital era. This paper presents a concise overview of deepfake detection research, methodologies, challenges, and ethical considerations. The paper first outlines state-of-the-art detection techniques, both traditional forensic and machine learning-based, that exploit subtle cues to differentiate between genuine and manipulated media.It then delves into the technicalities of deepfake generation, emphasizing the importance of understanding these techniques for developing robust detection mechanisms The paper also examines the societal impacts of deepfakes, including implications for privacy and free expression, and the role of various stakeholders in shaping deepfake development. In conclusion, the paper emphasizes the need for ongoing research, ethical considerations, and informed public discourse to harness the benefits of deep learning while mitigating societal risks.*



Keywords: Deepfake

I. INTRODUCTION

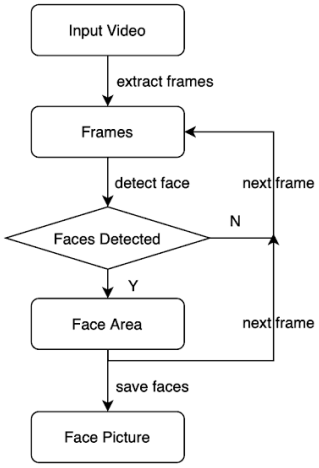
DEEPFAKE is the combination of deep learning and fake contents. It is a process that involves the swapping of a face from a person to another person in a video and making the face expression of a person similar to another person and acting like another person saying the same words that are said by another person. The swapping of faces especially on image and video or manipulation of facial expression is called Deepfake methods. Fake videos and images that get viral on the internet can easily exploit some individuals and it has become a public issue recently. Deepfake technique is also used to create fake news, do fake fraud and even spread hoaxes. Recently Deepfake has been a special concern of researchers . Deepfake can do pornography, political or blackmailing of a person by using his or her image and voice without his or her Permission. The main aims of deepfake are Political persons, public figures, and celebrities. Deepfake technology is used several times to spread the false messages of world leaders. It can also be used to provide fake images of maps to the military which could create serious damage.

Definition : Deepfakes are fake videos or pictures that are made by changing someone's face with another person's face. This is done by using computer programs that learn from data and use it to make realistic-looking fake images. These programs use advanced techniques like facial recognition and artificial intelligence. Photo editing started long ago, in the 1800s, and then people started using it in movies. As technology got better, especially with digital videos, editing became even easier. Now this manipulation can be done through Artificial intelligence and is now called Deepfake . Deepfake technology has been developed by researchers at academic institutions early in the 1990s, and later by amateurs in the online communities. More recently the methods have been adopted by industries for several purposes like entertainment, acting etc .

DOI: 10.48175/IJARSCT-17631

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

Deepfakes are getting a lot of attention because they can be used for bad things like making fake child abuse pictures, porn videos with famous people, spreading lies, and causing problems like bullying and scams. This kind of fake content can mess up how people make decisions in democratic systems because it tricks them and spreads hate and lies. Both businesses and governments are trying to stop this by finding and limiting deep fakes



On the other hand, deepfake tech has also improved a lot and is now used in fun things like movies and games. It's become so good that it's changing how we see entertainment and media.

Techniques

Deepfakes use a type of neural network called an autoencoder. This network has two parts: an encoder that simplifies an image into a smaller, hidden space, and a decoder that turns this simplified version back into the original image. In deep fakes, there's a special encoder that turns a person's appearance into this hidden space, capturing important details like their face and posture. Then, a specific decoder is used to apply this information to another video, merging the target's features onto the original video's underlying features. A more advanced version of this setup includes a generative adversarial network (GAN) connected to the decoder. In a GAN, there's a generator and a discriminator that work against each other. The generator
creates new images based on the hidden representation, while the discriminator tries to spot whether the image is fake. This competition pushes the generator to make images that look very real because any mistakes would be noticed by the discriminator. Both parts of the GAN keep improving, creating a constant back-and-forth to make better fakes. This makes deep fakes hard to detect and fight because they keep getting better and fixing any flaws that are found.

Applications

Deepfake technology has so many applications which could be used in two ways positively or negatively, however most of the time it is used to cause harm to someone.The unethical use of Deepfake technology has many harmful consequences in our society either in the short term or long term. People regularly using social media are at a huge risk of Deep Fake. However, proper use of this technology could bring many more positive results. Below are some negative and positive applications of Deepfake technology that are described in detail.

Positive Applications

Most of the time these technologies are used to cause harm to someone with bad intentions.But still it has some positive uses also in several sectors. To create new artwork , engage audiences and give them unique experiences this technology was used. Following positive application of deepfakes are given below.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

1. Entertainment

Deepfake technology has been widely used in entertainment for various purposes. It has been used to create hyper realistic videos of celebrities, such as making them sing or act in ways they never did in real life. For example, deepfake videos have been used to create virtual performances of famous musicians like Elvis Presley or to have multiple celebrities appear together in a single video. Additionally, deep face technology has been used to bring back deceased celebrities or to create entirely new characters for movies or TV shows. While these applications can be entertaining and innovative, they also raise ethical questions about
the authenticity of media and the potential misuse of such technology.

On June 8, 2022, Daniel Emmet, a former AGT contestant, worked with a company called Metaphysic AI to create a very realistic fake video of Simon Cowell, a judge known for being tough on contestants. Emmet sang a song on stage while a video of Simon Cowell singing along appeared perfectly in sync behind him.

Later, on August 30, 2022, Metaphysic AI made deep fake videos of Simon Cowell, Howie Mandel, and Terry Crews singing opera together on stage. Then, on September 13, 2022, they used a fake version of Elvis Presley for the finals of America's Got Talent.

Another project by MIT called 15.ai has been used to create content for different fan communities online, especially on social media.

In 2023, bands like ABBA and KISS worked with companies like Industrial Light & Magic and Pophouse Entertainment to create fake avatars that can perform virtual concerts.

2. Acting

Deepfake technology is increasingly being used in acting to bring deceased actors back to the screen or to digitally rejuvenate older actors. Digital clones of professional actors have appeared in films before, and progress in deepfake technology is expected to further the accessibility and effectiveness of such clones. The use of AI technology was a major issue in the 2023 SAG-AFTRA strike, as new techniques enabled the capability of generating and storing a digital likeness to use in place of actors.

3.Memes

Deep Face technology has been creatively used in memes to generate humorous and sometimes absurd content. One popular application is replacing faces in existing videos with those of well-known figures, celebrities, or fictional characters, creating unexpected and entertaining scenarios. For instance, deep fake memes may feature politicians delivering unusual speeches, movie scenes with characters swapped for unlikely alternatives, or singers performing songs in unexpected styles. These memes often rely on clever editing and realistic facial expressions to create a convincing illusion, leading to widespread sharing and engagement across social media platforms.In 2020, a funny internet trend started using deep face technology to make videos of people singing a song called "Baka Mitai" from the Yakuza 0 video game. In the game, this song is sung sadly by the player during a karaoke game. Many versions of this trend use a video from 2017 where a person lip-syncs the song, and then they swap the face with different people using deepfake technology.

Negative Applications

Deepfake and technology related to deep face are increasing rapidly in the current year. It has so many applications that are used to cause harm to someone working against human beings, especially against celebrities and political leaders. There are several reasons for making deep fake content that could be out of fun but sometimes it is used for taking revenge , blackmailing and the most common use of deepfake is to make pornography. Some negative applications of deepfake are explained below.

1.Blackmail -:

Deepfakes can be used to create fake evidence that makes someone look guilty of something they didn't do. A report from the American Congressional Research Service warned that deepfakes might be used to threaten politicians or people who know secret information, like spies do. Another way deepfakes can be used is if someone is actually being

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

blackmailed, they can say the proof against them is fake because deep fakes look so real. This makes the real blackmail evidence lose its power, and the person being blackmailed can break free from the blackmailer's control. This is called "blackmail inflation" because it makes real blackmail less
valuable. With just a simple program and a regular computer, people can make lots of fake blackmail material, making it harder to tell what's real and what's fake.

2. Fraud and Scams

Fraudsters and scammers use deepfakes to push people into fake investment schemes, financial fraud, cryptocurrency scams, sending money, and following fake endorsements. They often use the faces of famous people like celebrities and politicians to carry out these scams. This includes fake endorsements from stars like Taylor Swift, Tom Hanks, Oprah Winfrey, and Elon Musk, as well as news anchors and politicians. These fake videos appear in online ads on platforms like YouTube, Facebook, and TikTok, even though these platforms have rules against fake and manipulated content. These ads reach millions of viewers and can be used for various scams, such as fake giveaways, fraudulent weight-loss products, fake iPhone offers, and get-rich-quick schemes. Some scams also use AI to copy a celebrity's voice or use real-time deepfake technology to deceive people. Celebrities have warned their fans about these scams, but it can be hard to stop them completely due to the challenges of finding and using scammers online. Audio deepfakes have also been used to trick people into thinking they're talking to someone they trust, leading to scams like a CEO being scammed over the phone by someone using a fake voice. Overall, as deepfake technology improves, so do the scams targeting unsuspecting victims.

3.Politics

Deepfake technology has negative implications for politics as it can be used to spread misinformation, manipulate public opinion, and undermine trust in democratic processes.

During the 2020 Delhi Legislative Assembly election campaign, the Delhi Bharatiya Janata Party used special technology to create a version of a campaign ad in English by its leader, Manoj Tiwari. They changed it to Haryanvi language to appeal to voters in Haryana. They had an actor speak over the ad, and they used AI trained on Tiwari's speeches to make his mouth move along with the new voice. A party member said it was a "good" way to use deepfake tech because it helped them reach voters who speak a different language than the candidate.

4.Pornography

:In 2017, fake videos started showing up on the Internet, especially on Reddit, where people used deepfake technology to make fake porn videos. These videos often use the faces of female celebrities without their permission. A report in 2019 said that most of the deepfake videos online were pornographic, with only a small percentage being other types of content. Some famous examples include a fake video of Daisy Ridley in 2018. Most of these fake video feature British and American actors, but a significant number also use South Korean stars, especially K-pop idols.

In June 2019, a computer program called DeepNude was released, allowing users to create fake nude images of women using artificial intelligence. The program had both free and paid versions, with the paid version costing $50. However, the creators removed the program later that month and gave refunds to users.

Female celebrities are often targeted for these fake porn videos. For instance, in 2023, fake porn videos of Emma Watson and Scarlett Johansson were made using face-swapping technology. In 2024, fake nude images of Taylor Swift also appeared online

Concerns

Credibility and authenticity :

Although fake photos have been common for a while, making fake videos has been harder. Deepfakes make it even tougher to tell if a video is real or fake. AI researcher Alex Champandard says it's important to understand how quickly deep face technology can spread misinformation. The main issue isn't technical; it's about trusting information and news.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

Deepfakes can be used to harm, impersonate, and spread fake news. They can make it hard to know what's true anymore. This can create doubts and make people trust less in what they see and hear. It can also affect how societies work together, like discussing important issues and making decisions.

Hao Li, a computer science professor, warns that if we don't raise awareness about deep fakes, they could cause even more harm, like spreading fake news. He says that soon, real videos and deep fakes might look the same because of how fast technology is improving.

Experts like Shuman Ghosemajumder from Google also worry about deep fakes becoming a big problem. They think that eventually, anyone could make lots of deep fake videos easily, which could cause a lot of problems.

Deepfakes can be very damaging to individuals too. They can target one person or their relationships to create false stories that can change public opinions. They can even fake phone calls or conversations to trick people.

In September 2020, Microsoft announced that they are working on software to detect deep fakes.

II. CONCLUSION

As of 2024, the issue of deepfakes continues to be a complex and challenging area in the realm of artificial intelligence and digital media. The technology behind deepfakes has evolved significantly, allowing for increasingly convincing manipulations of audio, images, and video content. This has raised serious concerns about the potential misuse of deepfakes for various malicious purposes, including misinformation, fraud, and privacy violations.

One of the key challenges in addressing deepfakes is the ongoing cat-and-mouse game between creators of deepfake content and those developing detection and mitigation techniques. While there have been advancements in deepfake detection tools, such as machine learning algorithms that can identify inconsistencies in facial expressions or audio pattern, the rapid development of deepfake generation methods means that detection mechanisms must constantly adapt and improve.

Legal and ethical considerations surrounding deepfakes also remain complex. Questions about the responsibility of platforms in preventing the spread of malicious deepfake content, the potential impact on elections and public trust, and the implications for privacy and consent in the digital age are areas of ongoing debate and discussion.

In conclusion, while there have been strides in both the development of deepfake technology and efforts to combat its negative effects, the landscape remains dynamic and challenging. Continued research, collaboration between technology experts, policymakers, and stakeholders, as well as increased public awareness, are crucial in addressing the multifaceted issues posed by deepfakes in 2024 and beyond.

REFERENCES

[1]. Brandon, John (16 February 2018). "Terrifying high-tech porn: Creepy 'deepfake' videos are on the rise". Fox News. Archived from the original on 15 June 2018. Retrieved 20 February 2018.

[2]. "Deepfakes, explained". MIT Sloan. 7 March 2024.

[3]. Schwartz, Oscar (12 November 2018). "You thought fake news was bad? Deep fakes are where truth goes to die". The Guardian. Archived from the original on 16 June 2019. Retrieved 14 November 2018.

[4]. Charleer, Sven (17 May 2019). "Family fun with deepfakes. Or how I got my wife onto the Tonight Show". Medium. Archived from the original on 11 February 2018. Retrieved 8 November 2019.

[5]. Banks, Alec (20 February 2018). "What Are Deepfakes & Why the Future of Porn is Terrifying". Highsnobiety. Archived from the original on 14 July 2021. Retrieved 20 February 2018.

[6]. Christian, Jon. "Experts fear face swapping tech could start an international showdown". The Outline. Archived from the original on 16 January 2020. Retrieved 28 February 2018.

[7]. Roose, Kevin (4 March 2018). "Here Come the Fake Videos, Too". The New York Times. ISSN 0362-4331. Archived from the original on 18 June 2019. Retrieved 24 March 2018.

[8]. Ghoshal, Abhimanyu (7 February 2018). "Twitter, Pornhub and other platforms ban AI-generated celebrity porn". The Next Web. Archived from the original on 20 December 2019. Retrieved 9 November 2019.

[9]. Clarke, Yvette D. (28 June 2019). "H.R.3230 - 116th Congress (2019-2020): Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019". www.congress.gov. Archived from the original on 17 December 2019. Retrieved 16 October 2019.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

[10]. Lalla, Vejay; Mitrani, Adine; Harned, Zach. "Artificial Intelligence: Deepfakes in the Entertainment Industry". World Intellectual Property Organization. Retrieved 8 November 2022.

[11]. Sanchez, Julian (8 February 2018). "Thanks to AI, the future of 'fake news' is being pioneered in homemade porn". NBC News. Archived from the original on 9 November 2019. Retrieved 8 November 2019.

[12]. Bode, Lisa; Lees, Dominic; Golding, Dan (29 July 2021). "The Digital Face and Deepfakes on Screen". Convergence: The International Journal of Research into New Media Technologies. 27 (4): 849–854. doi:10.1177/13548565211034044. ISSN 1354-8565. S2CID 237402465

[13]. Fletcher, John (2018). "Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance". Theatre Journal. 70 (4): 455–471. doi:10.1353/tj.2018.0097. ISSN 1086-332X. S2CID 191988083

[14]. van der Nagel, Emily (1 October 2020). "Verifying images: deepfakes, control, and consent". Porn Studies. 7 (4): 424–429. doi:10.1080/23268743.2020.1741434. ISSN 2326-8743. S2CID 242891792.

[15]. Fallis, Don (1 December 2021). "The Epistemic Threat of Deepfakes". Philosophy & Technology. 34 (4): 623–643. doi:10.1007/s13347-020-00419-2. ISSN 2210-5433. PMC 7406872. PMID 32837868.

[16]. 16.Chesney, Robert; Citron, Danielle Keats (2018). "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security". SSRN Electronic Journal. doi:10.2139/ssrn.3213954. ISSN 1556-5068.

[17]. Hwang, Yoori; Ryu, Ji Youn; Jeong, Se-Hoon (1 March 2021). "Effects of Disinformation Using Deepfake: The Protective Effect of Media Literacy Education". Cyberpsychology, Behavior, and Social Networking. 24 (3): 188–193. doi:10.1089/cyber.2020.0174. ISSN 2152-2715. PMID 33646021. S2CID 232078561.

[18]. Hight, Craig (12 November 2021). "Deepfakes and documentary practice in an age of misinformation". Continuum. 36 (3): 393–410. doi:10.1080/10304312.2021.2003756. ISSN 1030-4312. S2CID 244092288.

[19]. Hancock, Jeffrey T.; Bailenson, Jeremy N. (1 March 2021). "The Social Impact of Deepfakes". Cyberpsychology, Behavior, and Social Networking. 24 (3): 149–152. doi:10.1089/cyber.2021.29208.jth. ISSN 2152-2715. PMID 33760669. S2CID 232356146

[20]. Suwajanakorn, Supasorn; Seitz, Steven M.; Kemelmacher-Shlizerman, Ira (July 2017). "Synthesizing Obama: Learning Lip Sync from Audio". ACM Trans. Graph. 36 (4): 95:1–95:13. doi:10.1145/3072959.3073640. S2CID 207586187

[21]. Mirsky, Yisroel; Mahler, Tom; Shelef, Ilan; Elovici, Yuval (2019). CT-GAN: Malicious Tampering of 3D Medical Imagery using Deep Learning. pp. 461–478. arXiv:1901.03597. ISBN 978-1-939133-06-9. Archived from the original on 20 June 2020. Retrieved 18 June 2020.

[22]. Cole, Samantha (24 January 2018). "We Are Truly Fucked: Everyone Is Making AI-Generated Fake Porn Now". Vice. Archived from the original on 7 September 2019. Retrieved 4 May 2019.

[23]. Hathaway, Jay (8 February 2018). "Here's where 'deepfakes,' the new fake celebrity porn, went after the Reddit ban". The Daily Dot. Archived from the original on 6 July 2019. Retrieved 22 December 2018