Understanding Deepfake Technology In The Age Of Artificial Intelligence

Suraiya Parveen

School of Engineering Science and Technology Jamia Hamdard suraiya@jamiahamdard.ac.in

Abstract

Emerging technologies like Big Data, Artificial Intelligence, Data Analytics, Machine Learning, Artificial Neural Networks, and Deep learning are the vast point of apprehension for technocrats these days. These technologies are required to cope up with the challenges of the data generated by this generation. Deep fake is one such technology that is a combination of deep learning and synthetic media. Deep fakes are hyper-realistic videos digitally controlled to illustrate people saying and doing things that never actually happened. The modus operandi involves footage of two people into a deep learning algorithm to train it to swap faces. In other words, deep fakes use Facial Mapping Technology and Artificial Intelligence that swap the face of a source person with the target person on a video. In the past few years, Deep fake has become a problem that is a threat to public discourse, human society, and democracy. False information flows quickly through social media, where it can hit millions of users. This paper gives an insight into Deep Fake technology to evade its use to spread misinformation, damage reputations, and harm individuals.

Keywords: Deep fake, Artificial Intelligence, Deep Learning, Synthetic Media

1. INTRODUCTION

Deep fake is a technique that can superimpose face images of a target person to a video of a source person to create a video of the target person performing or saying things the source person does. The term ‘deep fakes’ first became popular in 2017 when a Reddit user, who went by the same name, began posting digitally altered videos on the website's forums. These are doctored such that the face of entertainers is switched with those of other people, typically celebrities. Deep fake algorithms generally require a large amount of image and video data to train models to create photo-realistic images and videos. Such as celebrities and politicians may have a large number of videos and images available online, they are the main targets of deep fakes. Deep fakes were used to swap faces of celebrities or politicians to bodies in fake images and videos. Deep fakes can be defined as visual and audio content that has been controlled using advanced software to change how a person, object, or environment is presented. Deep fakes are affecting the people around the globe that are using deep fakes for multiple reasons such as face-swapping, recreating videos and images with someone's face or body, and creating and disseminating fake news[1]. Deep fakes are heavily affecting democracy, privacy, security, religion, and the cultures of the people. Deep fakes are increasing over the clock, but there is no standard to evaluate deep fake detection techniques. The number of deep fake videos and images establish online has nearly doubled since 2018[2]. The journalism industry is probably going to face a massive consumer trust issue due to deep fakes[3].

Deep fakes pose a major threat to "traditional" fake news because they are harder to detect and people are inclined to believe the fake is real. The technology allows the creation of seemingly legitimate news videos that place the reputation of journalists and the media at risk. The intelligence community is much worried about deep fakes that will be used to threaten national security by disseminating political propaganda and disrupting election campaigns. Deep fakes are a significant threat to our society, political system, and business because they pressurize journalists struggling to filter real from fake news, threaten national security by disseminating propaganda and disruption in elections, hamper citizen trust toward information by authorities, and, raise cybersecurity problems for people and organizations[4].

Cybersecurity problems constitute another threat produced by deep fakes. The corporate world has already shown interest in protecting itself against viral frauds, as deep fakes could be used for market and stock control[5]. Deep fakes technology could be used for brand destruction, blackmail, or to embarrass management. Many deep fakes focus on celebrities, politicians, and corporate leaders because the internet is stuffed with source photos and videos of them. It is threatening to world security when deep fake methods can be used to create videos of world leaders with fake speeches for false purposes[6]. Deep fakes therefore can be abused to create political or religious conflicts between countries, to fool the public and affect results in election campaigns, or create chaos in financial markets by creating fake news.

There is positive usage of deep fakes such as creating voices of those who have lost theirs or editing scenes of movies without re-shooting them[7]. However, the number of poisonous uses of deep fakes largely dominates that of the positive ones. Fake videos and images created by deep fake techniques have become a great public issue recently.

Deep fake technology also has positive uses in many fields, including movies, educational media, and digital communications, games, and entertainment, social media and health protection, material science, and several business fields, such as fashion and e-commerce[8][9]. The movie-making industry can benefit from deep fake technology in multiple ways. For example, it can help in creating digital voices for actors who lost their lives due to disease, or for updating movie footage instead of reshooting it. Moviemakers will be able to recreate typical scenes in movies, create new movies casting long-dead actors, make use of special effects and advanced face editing in post-production, and enhance amateur videos to professional quality.

Deep fake technology can break the language barrier on video calls by translating speech and simultaneously altering facial and mouth movements to enhance eye contact and make everyone appear to be speaking the same language[10]. Deep fakes can help people deal with the loss of loved ones by digitally bringing an expired friend "back to life", and thereby potentially assisting a grieving loved one to say goodbye. Businesses are interested in the development of brand- applicable deep fake technology, as it can modify e-commerce and advertising in significant ways[11]. For example, brands can sign a contract with supermodels who are not supermodels and show fashion attire on a variety of models with different skin tones, heights, and weights.

Another factual benefit of deep fake is, it makes us familiar with such fake things and we should not believe in everything we see around us[12]. Once we find that it is fake we learn and next time when such things come through similar sources, we take time to believe or do some research to verify the news.

Photos and videos are regularly used as evidence in police investigations to resolve legal cases since they are considered to be reliable sources[13. However, cosmopolitan technology increases the development of fake videos, and photos that have potentially made these pieces of evidence unreliable. Generative Adversarial Networks (GANs) in the recent advanced image and video creating tool to create high quality exploited deep fake videos and images, and the media increases the fast distribution of these fake images and videos[14].

Deep fakes have become well known due to the quality of contaminated videos and also the easy-to-use ability of their applications to a wide range of users with various computer skills from professional to beginner[15]. These applications are mostly developed based on deep learning techniques. Deep learning is well known for its ability to represent complex and high-dimensional data. One variant of the deep networks with that capability is deep autoencoders, which have been widely used for dimensionality reduction and image compression. The autoencoder withdraws latent features of face images and the decoder is used to reconstruct the face images[4]. To swap faces of source images and target images, there is a need for two encoder-decoder pairs where each pair is used to work on an image set, and the encoder's limits are shared between two network pairs. This method enables the common encoder to find and learn the similarity between two sets of face images, which are relatively not challenging because faces normally have similar features such as eyes, nose, and mouth positions.

2. LITERATURE REVIEW

Fast development in AI technologies, especially in deep learning and computer vision, gave way to deepfakes—synthetic media that is created by image, video, and audio manipulation using very sophisticated machine learning models. Deepfake technology has improved so much since the introduction of GANs by Goodfellow et al. in 2014 to create very realistic, virtually undetectable media. This section surveys some of the important literature pertaining to deepfake technology, its underpinning mechanisms, and applications, along with some of the ethical, legal, and social challenges that have been forwarded[2].

Deepfake generation fundamentally builds on a class of deep learning models known as GANs, basically comprising a generator and a discriminator network. GANs learn to create synthetic outputs that resemble real data by iteratively improving the generator with feedback from the discriminator, whereby a discriminator has to decide whether an output is real or fake. This leads over time to ever more real looking synthetically generated media. Variants like StyleGAN and CycleGAN have further improved precision and quality of the media generated, allowing more sophisticated manipulations to be effected within videos, images, and audio[3].

Moreover, in application to face-swapping, beyond GANs, autoencoders have been broadly utilized. Methods like variational autoencoders and their derivatives provide the ability for facial features to be encoded and decoded into a format that can be reconstructed seamlessly into images (Kingma & Welling, 2014). With such developments, deepfakes became more accessible to the general public by allowing user-friendly tools and open-source software to help users create convincing fake media even if not being an expert in this area is present[4].

Deepfakes, however, apply in both benign and malicious ways. The entertainment industry used deepfakes to revive dead actors, enhance dubbing, or add special effects onto international movies. In the academic setting, researchers have been using deepfakes for educational purposes such as reenacting history, doing simulations, or recreating interactive learning experiences.

Deepfake technology, however, is even more infamous in the realm of their malicious applications. One of the most concerning usages includes politically motivated disinformation, fake news, and misleading propaganda[5]. Deepfake videos and audio clips are weaponized against political figures and for spreading lies to incite social unrest. Moreover, there have been rising cases of non-consensual pornography where deepfakes are applied to superimpose faces over explicit content. This has brought a devastating dimension to victims of this cybercrime (Paris & Donovan, 2019).

Deepfakes are enormous ethically, for this technology literally questions existing notions related to truth, consent, and privacy. Floridi argues that, with the blurring of real and fake content, general societal trust in the media is threatened. Further, deepfakes used for creating non-consensual pornography and slandering people give rise to major concerns over autonomy and dignity.

Deepfakes thus pose challenges to the present legal frameworks. The laws are always struggling to catch up with technological progress, thereby opening the possibility of open gaps for protection against identity theft, defamation, and invasion of privacy according to Kietzmann et al. in 2020. While some jurisdictions have enacted legislation to make the malfeasance associated with deepfakes criminal, such application proves quite difficult across jurisdictional borders in the light of the global spread of digital media.

Detection and Mitigation Strategies

Research in this regard is quickening up just like deepfake technology itself. Deepfake detection techniques are largely based on digital forensics, which involves algorithms that examine visual, auditory, and textual anomalies. Machine learning models have been developed recognizing patterns and artifacts that are typical of deepfakes, such as inconsistent lighting, unnatural facial movements, or even pixel-level irregularities. However, this "arms race" between deepfake creators and detection systems demands improvements on both fronts indefinitely[6].

Blockchain and digital watermarking are similarly a subject of recent research in verifying the authenticity of media. This approach contributes to tracing the origin of digital content, thereby making deepfakes' circulation without detection quite challenging.
Though deepfakes have made great strides in better understanding and countering these threats, several areas of research remain open. More effective and easily replicable mechanisms of detection that could work in real-time across platforms are needed. Greater attention needs to be paid to the psychological and social impact of deepfakes, especially in respect to media literacy and popular awareness. As Schwartz et al. say, legal and regulatory regimes should be updated to handle the new threats, so people do not go unshielded against deepfake injuries.

Table 1 Comparative Analysis of Related Work

Author(s) & Year	Title	Research Focus	Methodology	Key Findings	Research Gaps/Limitations
Goodfellow et al. (2014)	Generative Adversarial Networks	Introduction of GANs for image generation	Development of GAN framework	GANs enable realistic synthetic image generation	Initial GAN models faced instability during training
Karras et al. (2019)	StyleGAN: A Style-Based Generator Architecture for GANs	Improved GAN architecture for high-quality images	Style-based generator with adaptive instance normalization	StyleGAN produces photorealistic images with high detail	Limited application beyond image synthesis
Chesney & Citron (2019)	Deepfakes and the New Disinformation War	Ethical and legal implications of deepfakes	Theoretical analysis and case studies	Deepfakes pose severe risks for misinformation and privacy	Lack of comprehensive legal frameworks for deepfake regulation
Vaccari & Chadwick (2020)	Deepfakes and Disinformation in Politics	Use of deepfakes in political disinformation campaigns	Analysis of deepfake incidents in political contexts	Deepfakes are increasingly used to manipulate political narratives	Difficulty in detecting deepfakes in real-time settings
Verdoliva (2020)	Media Forensics and Deepfake Detection	Overview of detection techniques for deepfakes	Survey of forensic methods and machine learning models	Digital forensics can identify artifacts unique to deepfakes	Deepfake detection struggles against sophisticated new models
Li et al. (2020)	FaceForensics++: Learning to Detect Manipulated Facial Images	Deepfake detection techniques using facial analysis	Dataset creation and training of deepfake detection models	High accuracy achieved in detecting manipulated facial images	Models have reduced effectiveness on low-quality videos
Nguyen et al. (2021)	Blockchain and Watermarking	Methods for verifying	Integration of	Blockchain provides a	Scalability issues and high

	for Authenticity Verification	digital content authenticity	blockchain with digital watermarking	secure method for tracing media origins	computational costs for real-time use
Schwartz et al. (2021)	Deepfake Psychology: Understanding Public Perception	Societal impacts and public perception of deepfakes	Survey and experimental studies on media consumption	Deepfakes significantly influence trust in digital content	Need for more extensive research on media literacy interventions
Paris & Donovan (2019)	Deepfakes and the Threat to Truth	Ethical issues related to deepfake abuse	Theoretical analysis of privacy and consent issues	Deepfakes exacerbate concerns around privacy and dignity	Limited focus on legal solutions specific to non-consensual deepfakes
Kingma & Welling (2014)	Auto-Encoding Variational Bayes	Introduction of variational autoencoders (VAEs)	Development of VAEs for latent-space representation	VAEs are useful for encoding and decoding facial features	VAEs lack the precision of later GAN-based models

3. Creating Deepfakes:

Deepfakes are created using Deep Learning technology. This is done by using thousands of images in machine learning tools and train them to reconstruct other patterns.

Figure1:Deep Fake a simple illustration

3.1 Classification of Face creation Technology

A) Face replacement – Face replacement, also referred to as face swapping, is the process of meticulously "stitching" an image of one person's face (the source) onto another (the target). The source is the main focus, and the target's identity is concealed.

B) Face re-enactment – Face re-enactment, also referred to as puppetry, is controlling a target's facial features, such as their mouth, eyebrows, eyes, and head movement[16]. Reenactment is not about changing identities; rather, it's about manipulating facial expressions to make someone seem like they are saying something they are not.

C) Face generation-Face creation is the process of producing an entirely new face image. Generative Adversarial Networks, the newest type of deep learning, are used for this. It pits two neural networks against one another, with the first one creating an image and the second determining whether or not the output is realistic.

D) Speech synthesis – Speech synthesis is a relatively new subset of deep fakes that entails building a voice model that can mimic the target's mannerisms, tone, and speech synthesis[17]. Instead of reproducing a certain target, some speech synthesis companies, like Modulate.ai, let consumers select a voice of any age or gender.
Following are the steps involved:

a) **Extraction** – Gathering a large number of photos to train the face replacement model is the first task. A well-known technique involves beginning with recordings of the source and the target, cutting them into separate frames, and then cropping the pictures until the face is the only thing left in the image[18]. The two people would be an exact match in terms of head size and facial form. You can now use this software to assist with the extraction procedure.
b) **Training** – The next task is to use the gathered photos to train the face-swapping model. An autoencoder[19], a neural network composed of an encoder and a decoder, is used to do this. A facial image is input into the encoder, which then compresses it into a low-dimensional representation called the "latent face." After that, the decoder uses that representation to reconstruct the face in its original shape (refer to Figure 2).

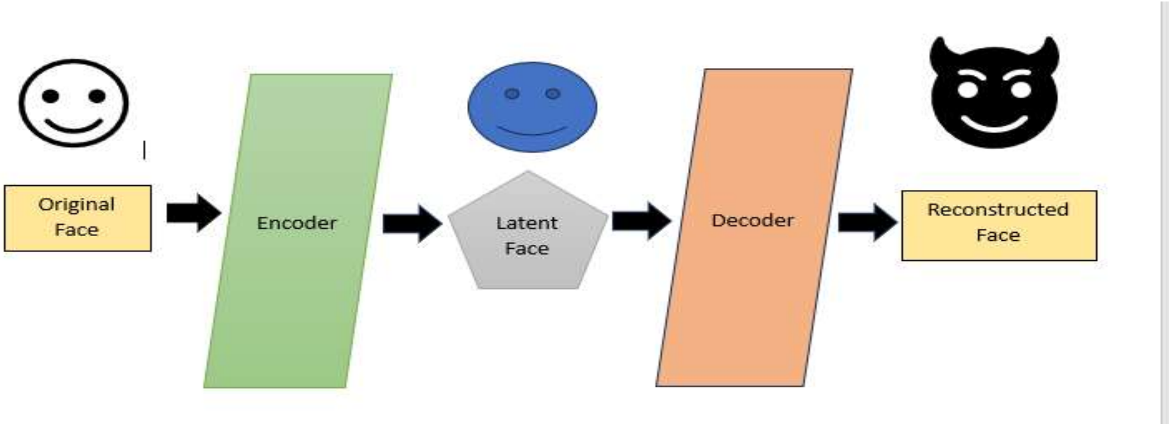


Figure2: Reconstruction using autoencoder

By training two of these neural networks, one for the source face and one for the target face, face replacement is theoretically possible. Since both networks use the same encoder, their decompressed forms have a baseline architecture that is comparable. They do, however, have different decoders (see Figure 2). Until the autoencoding procedure is able to recreate a picture of a face that is comparable to its original version, both neural networks are left to train.

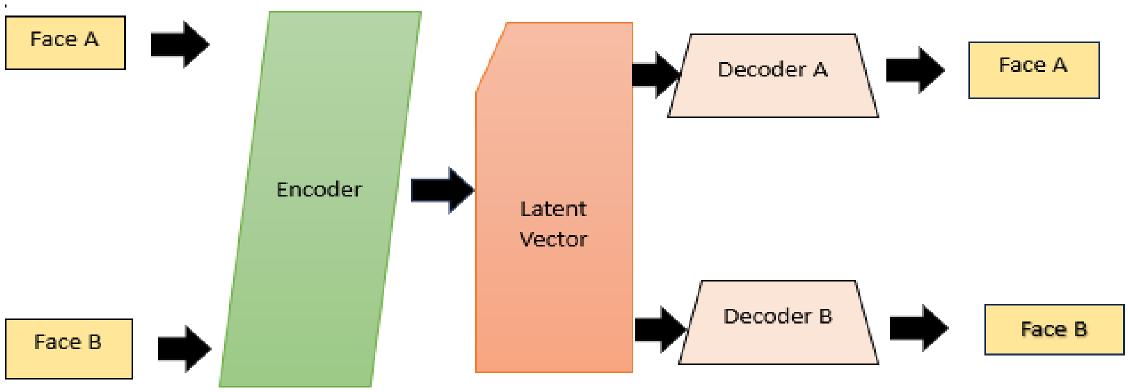
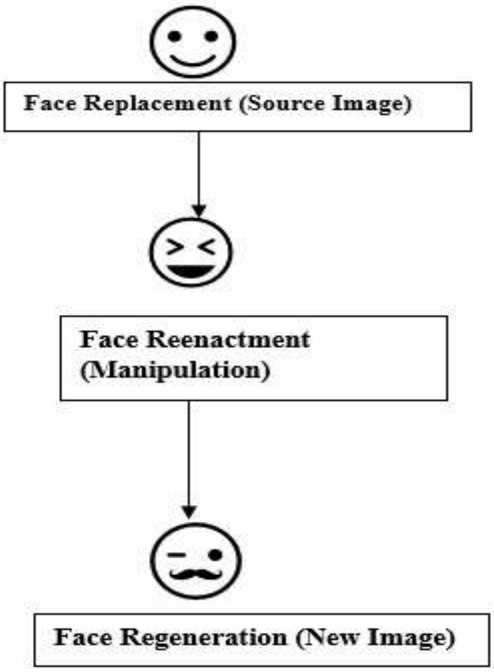


Figure3: Training Phase

Once the exercise is complete, the decoders are then swapped, in such a way that the source image is decompressed using its base encoder but reconstructed using the target image's decoder. The result is an image that accurately stitches the source's face onto the target's while staying true to the target's expressions (see Figure 3). Accordingly, when the object opens its mouth or moves their eyebrows, they do so but with the visual characteristics of the source person.

(c) Creation –

Inserting these deepfake images into the required video is the last and technically most difficult step. This indicates that the angle of the target person's head and the integrated face match for every frame in the video. This step of the process is more error-prone because it is the only one that uses hand-written code rather than machine learning algorithms.



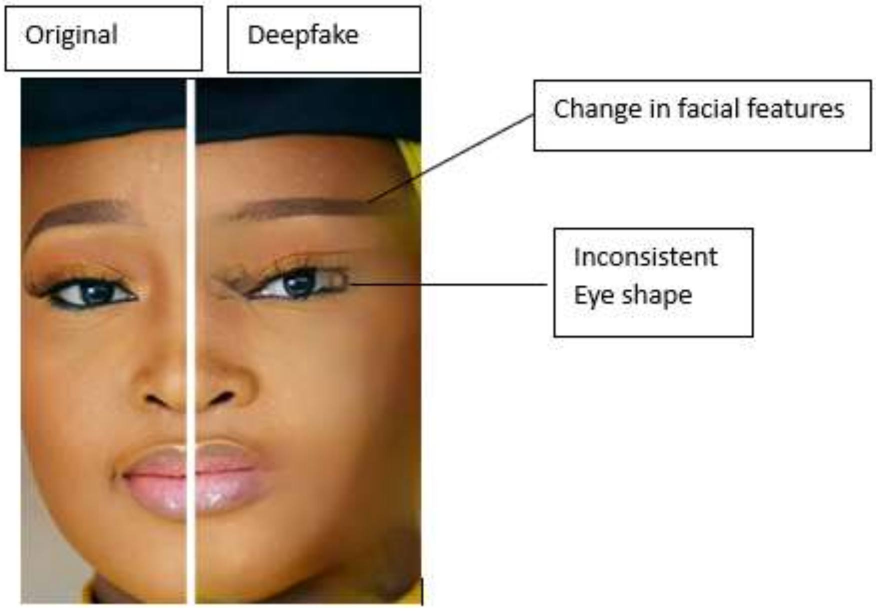
c)

Figure4: Creation Phase

d) Face re-enactment, the other familiar form of deep fake, also draws on a type of auto encoder. However, this method needs only a single video of the target rather than numerous individual images of both target and source, as is the case with the face replacement technique.

4. DEEP FAKE DETECTION

To detect deep fakes, various methods have been suggested after this threat was introduced. In a binary classification deep fake detection technique, the classifiers, classify the manipulated and real videos[4]. This type of detection technique requires a huge dataset of real and fake videos to train the machine. Many deep fake videos are available online on the Internet, but it is still limited to set a standard to evaluate different deep fake detection techniques. The reviewed articles proposed that there are ways to combat deep fakes such as using subsisting laws, legislation and regulation, additional action from social media companies, corporate policies and voluntary action, education, and training, establishing digital literacy curriculum in schools, enhance media literacy, and anti- deep fake technology, content authentication, deep fake prevention and develop deep fake detection methods.



Deep fakes are increasingly injurious to privacy, social security, and democracy. Methods for detecting deep fakes have been proposed as soon as this threat was introduced. Detection is another important confinement tool[5]. Media forensic methods have long been used in criminal courts to cross-question visual evidence, but they can also be implied to help identify deep fakes. One form of media forensics involves inspecting individuals in footage for physiological inconsistencies that emerge from the way doctored videos are constructed.

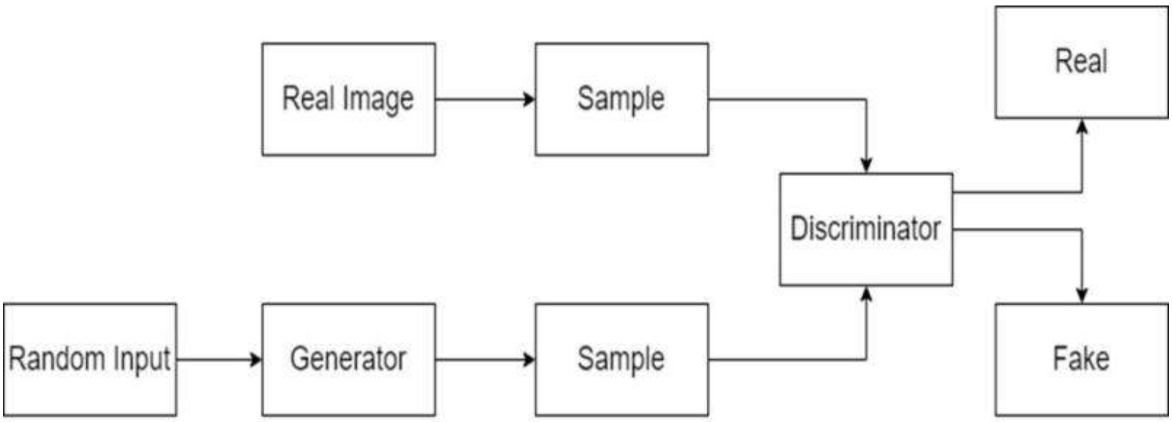


Figure 4 Architecture of GAN

This consists of looking at whether subjects blink during the footage, and whether the color and shadows on their skin appear to flicker[6]. Another approach is to check whether the acoustics of a video correlates with the scene being recorded, for example, the size of the room or the presence of individuals in the background.

Vid TIMIT's publicly available database consists of both low and high-quality deep fake videos, which can accurately imitate the lip movements, facial fallings, and eye blinking of the individual. These dataset videos are then used to test various deep fake detection methods. The result shows that the Face net and VGG face recognition systems can't detect deep fakes effectively. The image quality metrics and the lip-syncing method using Support Vector Machine (SVM) show an error when trying to detect deep fake videos[7]. This raises the critical need for more effective and efficient deep fake detection algorithms.

4.1 FAKE IMAGE DETECTION

Face swapping has many thrilling applications in video compositing, transfiguration in portraits, and specifically in identity protection as it can swap faces in photographs by ones from a collection of stock images. However, it is also one of the methods that cyber attackers employ to penetrate identification or authentication systems to gain illegitimate access. The use of deep learning such as CNN and GAN has made swapped face images more challenging for forensics models as it can conserve pose, facial expression, and lighting of the photographs Amid deep learning- generated images, those synthesized by GAN models are probably most difficult to spot as they are realistic and high-quality based on GAN's capability to learn the distribution of the complex input data and produce new outputs with similar input distribution[8].

Recently a two-phase deep learning method for detection of deep fake images. The first phase is a feature extractor based on the usual fake feature network (CFFN) where the Siamese network architecture is used[9]. The CFFN encompasses various dense units with each unit including different numbers of dense blocks to enhance the representative capability for the fake images[10]. The number of dense units is three or five depending on the validation data being the face or general images, and the number of channels in each unit is varied up to a few hundred. Discriminative features amid fake and real images, i.e. pairwise information, are extracted through the CFFN learning process. These features are then provided into the second phase, which is a small CNN concatenated to the last convolutional layer of CFFN to distinguish distorted images from genuine ones[19]. The suggested method is validated for both fake face and fake general image detection.

4.2 **FAKE VIDEO DETECTION**

Most image detection methods cannot be used for videos because of the strong humiliation of the frame data after video compression[20]. Furthermore, videos have material characteristics that are varied among sets of frames and thus challenging for methods designed to spot only still fake images. This subsection focuses on deep fake video detection methods and classifies them into two groups: methods that employ material features and those that explore visual artifacts within frames. The swapped face images produced by CNN and GAN deep learning are more challenging for forensics models to detect the changes[11]. The authors suggested a multitask learning method to perform the classification and segmentation of manipulated facial images simultaneously.

In the detection process, the material consistency of the video is not inflicted efficiently and enforced to use the Spatiotemporal contents of the video to detect deep fakes. The integration of conventional networks and recurrent unit manipulate the material inconsistencies of the frames. Using both Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) able to extract material features of the video, that are denoted through the order of frames[12]. The deep fake detection network takes the ordered frames and estimates the possibilities of the frame sequence that can be either original or deep fake. CNN is used to extract features of the frame and then fed them into the LSTM to generate a material sequence frame[13]. Using eye blinking, we can spot deep fakes in which a person in deep fakes has a lesser or higher eye blinking rate than in real videos. A normal person's eye will blink for 2 to 10 seconds, and each eye blink will take between 0.1 and 0.4 seconds[14]. In most cases, deep fake creation tools cannot create eyes that can blink like a normal individual. Most of the time eye blinking speed is manipulated videos are slower than in real ones[15]. The color of each eye is withdrawn using computer vision and the difference in eye color is also used to spot deep fakes.

Recurrent Convolutional Neural Networks (RCN) discovered material inconsistencies across the frame. It utilizes Spatiotemporal features of videos[16]. Deep recurrent network models mostly utilize material patterns across video frames to differentiate deep fake videos. Another approach that fragment videos into frames and recognize visual objects inside single frames to get discriminant features. These features are then split into low or high classifiers to detect fake and real videos. When deep fake videos are produced with low resolutions they need face rotation, zooming, and cutting to match with real ones. The resolution discrepancy between the nearby and the changing face area can be spotted by CNN models[17].

4.3 OTHER DEEP FAKE DETECTION TECHNIQUE

The familiar deep fake detection techniques such as Convolutional Neural Network (CNN) to extract frame feature, LRCN to capture the eye blinking material patterns, Recurrent Neural Network(RNN) to discover material discrepancies across frames, and Long Short Term Memory (LSTM) for temporal sequence analysis[18].

Long-Term Recurrent Convolutional Networks (LRCN) dynamically anticipate eye area sequences. LRCN captures temporal patterns of eye blinking in the videos since the eye blinking speed of deep fakes is slower or extremely faster than normal videos[19]. Also, it includes a feature extractor that extracts depend on CNN, an ordered learning through LSTM, and to anticipate the possibility of open and closed eye state. The eye blinking has material patterns and the LSTM can capture these temporal patterns effectively. The eye blinking rate is estimated as a peak above the threshold of 0.5 with a duration of lesser than 7 frames. The extremely frequent eye blinking may also be the criteria of modified videos[20]. CNN and LSTM CNN detects frame features then split to LSTM and LSTM detect temporal inconsistencies in videos to set frame order for classification.

CONCLUSION

Deep learning can be implemented in deep fake creation and detection methods. Deep fake creates manipulated images or videos that persons cannot differentiate from real images or videos. Deep fakes are created using generative adversarial networks (GAN), in which duo machine learning models exit. One model works on a dataset and the other model tries to detect the deep fakes. The forger creates fakes until the other model can't detect the falsification. Deep fake, creating fake news, videos, images, and terrorism events that can cause social and financial deception. It is increasingly affecting religions, organizations, individuals and communities, culture, security, and democracy. When deep fake videos and images rise on social media people will ignore to trust the truth. The available datasets, deep fake creation tools, deep fake challenges, fake video detection methods and detect fake video by using eye blinking are used.

REFERENCES:

1. P. Pérez, M. Gangnet, A. Blake, Poisson image editing. ACM Trans. Graph. (TOG). **22**(3), 313–318 (2003)
2. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014, June 10). Generative adversarial networks. arXiv.org https://arxiv.org/abs/1406.2661
3. Karras, T., Laine, S., & Aila, T. (2018, December 12). A Style-Based generator architecture for generative adversarial networks. arXiv.org. https://arxiv.org/abs/1812.04948
4. Kingma, D. P., & Welling, M. (2013, December 20). Auto-Encoding variational Bayes. arXiv.org. https://arxiv.org/abs/1312.6114
5. Suwajanakorn, S., Seitz, S. M., & Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama. ACM Transactions on Graphics, 36(4), 1–13. https://doi.org/10.1145/3072959.3073640
6. Paris, B., & Donovan, J. (2019, September 18). Deepfakes and cheap fakes. Data & Society. https://datasociety.net/library/deepfakes-and-cheap-fakes/
7. Citron, R. C. a. D. (2023, October 4). Deepfakes and the new Disinformation War: The coming Age of Post-Truth Geopolitics. Foreign Affairs. https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war
8. Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society, 6(1), 205630512090340. https://doi.org/10.1177/2056305120903408

9. Verdoliva, L. (2020). Media Forensics and DeepFakes: An Overview. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910–932. https://doi.org/10.1109/jstsp.2020.3002101
10. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019, January 25). FaceForensics++: Learning to Detect Manipulated Facial Images. arXiv.org. https://arxiv.org/abs/1901.08971
11. D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, S. K. Nayar, in ACM Transactions on Graphics (TOG), vol. 27. Face swapping: automatically replacing faces in photographs (ACM, 2008), p. 39
12. K. Dale, K. Sunkavalli, M. K. Johnson, D. Vlasic, W. Matusik, H. Pfister, in ACM Transactions on Graphics (TOG), vol. 30. Video face replacement (ACM, 2011), p. 130
13. H. -C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning. IEEE Trans. Med. Imaging. **35**(5), 1285–1298 (2016)
14. K. Guo, D. Tunnicliffe, H. Roebuck, Human spontaneous gaze patterns in viewing of faces of different species. Perception. **39**(4), 533–542 (2010)
15. Hunt ALLCOTT and Matthew GENTZKOW, "Social media and fake news in the 2016 election",Journal of economic perspectives, vol. 31, no. 2, pp. 211-36, 2017
16. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, M. Nießner, Faceforensics++: learning to detect manipulated facial images. arXiv preprint arXiv:1901.08971 (2019)
17. Y. Li and S. Lyu, Exposing DeepFake Videos By Detecting Face Warping Artifacts
18. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2016-Decem, pp. 770-778, 2016.
19. T. Karras, S. Laine and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks", 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4396-4405, 2019
20. I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., "Generative Adversarial Nets", NIPS, 2014.
21. Hadeer AHMED, Issa TRAORE and Sherif SAAD, "Detection of online fake news using n-gram analysis and machine learning techniques", International Conference on Intelligent Secure and Dependable Systems in Distributed and Cloud Environments, pp. 127-138, 2017
22. Y. Kim, "Convolutional neural networks for sentence classification", arXiv preprint arXiv:1408.5882, 2014.
23. Natali RUCHANSKY, Sungyong SEO and Yan LIU, "Csi: A hybrid deep model for fake news detection", Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 797-806, 2017
24. D. E. King, "Dlib-ml: A machine learning toolkit", J. Mach. Learn. Res., 2009
25. F. Pedregosa et al., "Scikit-learn: Machine learning in Python", J. Mach. Learn. Res., 2011
26. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, in Advances in Neural Information Processing Systems. Generative adversarial nets (Neural Information Processing SystemsMontreal, 2014), pp. 2672–2680
27. G. Wang, W. Li, M. Aertsen, J. Deprest, S. Ourselin, T. Vercauteren, Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing. 338:, 34–45 (2019
28. V. Blanz, K. Scherbaum, T. Vetter, H. -P. Seidel, Exchanging faces in images. Comput. Graph. Forum. 23:, 669–676 (2004). https://doi.org/10.1111/j.1467-8659.2004.00799.x
29. D. Chen, Q. Chen, J. Wu, X. Yu, T. Jia, Face swapping: realistic image synthesis based on facial landmarks alignment. Math. Probl. Eng.2019, 1–1 (2019). https://doi.org/10.1155/2019/8902701.