A Systematic Review on Fake Image Creation Techniques

Ruby Chauhan
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
rubyyy.rana@gmail.com

RenuPopli
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
renu.popli@chitkara.edu.in

Isha Kansal
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
isha.kansal@chitkara.edu.in

***Abstract*—In the last few decades, Deep Learning has experienced remarkable growth, which has improved computer vision tasks. Deep fake is a technology that uses deep learning algorithms that manipulates the features of original content digitally to produce fake realistic-looking content that can be audio, video, photos, etc. There are various methods available in literature for creating deep fake such as GAN (General Adversarial Network), Auto Encoders, pix2pixGAN, Cycle GAN, Style GAN, Wave Net *etc.* and some of the open source digitally available tools are Deep Face Lab, Face Swap, *etc*. In the video game and film sectors, the adoption of the aforementioned techniques has expanded significantly. This study elaborates on a survey that was performed by several research organisations and focused on the feasibility gaps that need to be recovered for deep fakes. In this work, many contemporary strategies for creating false images are explained, along with the various dataset types that authors used. Finally, numerous research gaps and potential future directions are highlighted.**

Keywords— Deep Fake, GAN, Auto Encoders, StyleGAN, DeepFaceLab, Face Swap

I. INTRODUCTION

These days, the world is growing more and more computerized. Deep fakes (originates from the words "deep learning" and "fake," obviously) is a technique in which specifics of an original video or image, such as the face, body, audio, etc. are digitally altered using the features of another person so that it should appear to have been created by a different person and not be similar to the original video or image. The approaches used for fake image and video detection frequently demand a lot of photographs and videos of the target person to teach the models to originate pictures and videos that almost look natural[12-14]. In our daily lives, leaders and celebrities are usually the targets of abuse, but occasionally so are average people.

As is common knowledge, there are numerous pictures and videos of well-known people available online today. They therefore tend to be the first targets. GAN and Auto Encoders are two techniques that are often utilized in the production of deep fake content.

In terms of the degree of manipulation, there are four primary categories for facial alterations. Below is a description of each of them along with a timeline of changes over the last few years.

A. Entire Face Synthesis

This manipulation produces entirely fake facial pictures and often uses a GAN technique, such as Style GAN[23-24] or PGGAN[27]. With the use of these techniques, the most realistic and high-quality facial photographs can be created.

B. Identity swap

Due to the widespread public concern around Deep Fakes, this is one of the most popular face modification study areas nowadays. It involves swapping out one person's face for another person's in a clip. In identity swap, as opposed to entire face synthesis manipulation, where modifications are done at the picture level, the objective is to produce realistic fake videos.

C. Attribute Manipulation

A face's attributes, such as its skin color, gender, age, etc., can be altered using attribute manipulation, sometimes referred to as face editing. A GAN, such as the StarGAN[17], STGAN[15], AttGAN[16], *etc.,* is typically used to carry out this manipulation technique.

D. Expression Swap

Expression swap, commonly referred to as face recreation, alters a person's facial expression. This transfers the input expression to the target image. Various methods, including Face2Face[19], SVGAN[22], *etc.* can be used to do it.

There are various methods for creating deep fake such as: GAN, Auto Encoders, pix2pixGAN, Cycle GAN, Style GAN, Wave Net etc. and some of the open source digitally available tools like Deep Face Lab, Swap, etc. Two important methods are described below:

1) GAN(General Adversarial Network):

Generator and Discriminator are the two major components of the GAN. With the aid of noise signals (latent vector), The role of the generator is to create false visuals. It is the duty of the discriminator to classify incoming photos as authentic or fraudulent. The discriminator receives two different types of inputs: the genuine picture from the dataset and the image generated by the generator. In order to provide precise predictions, the discriminator tries to minimize discriminator loss. The generator, on the other hand, seeks to increase discriminator loss so that the discriminator can identify the bogus images as real images. In this process, the generator learns how to produce fake images that resemble real ones. Figure 1 illustrates the architecture of GAN.

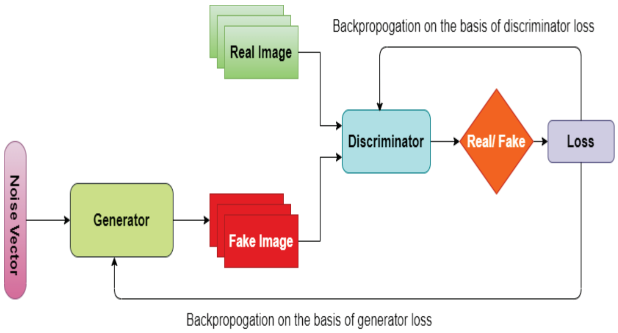

Fig. 1. GAN Architecture

2) Auto Encoders

The two components of an auto encoder are the encoder and decoder. The task of lowering image size and feature extraction from input images falls to encoders. And the decoder's role is to reverse the encoder's actions by reconstructing the face from the compressed face (extracted features) that it produces. For compression, the encoders use both linear and non-linear changes on the input images. The auto encoders find important features that are present in the data by lowering the reconstruction error between the source and final image. Neurons in the output layer of auto encoders are identical in number to those in the input layer. Figure 2: Illustrates the basic architecture of the Auto encoders used for face swapping.

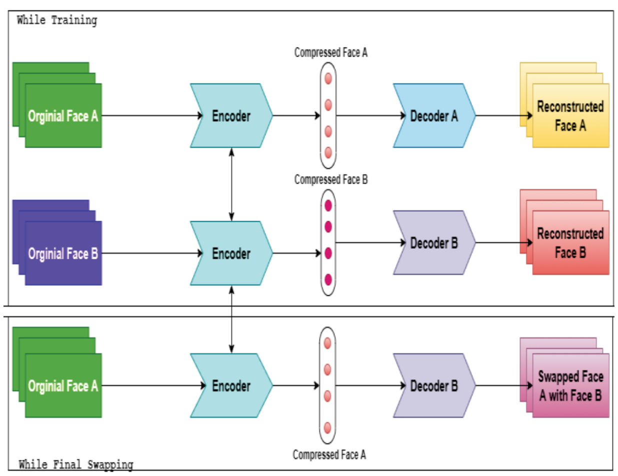

Fig. 2: Representation of the technology using Auto Encoders for face swapping.

II. LITERATURE REVIEW

The research of fake image production methods is contained in this section. An intriguing strategy in this area was depicted by Thanh Thi Nguyen *et al.,(2019)* [1] conjectured information on auto encoders, GAN, and Style GAN. They also includes a list of deep fake generating tools. Rahul Kataryet *al*., (2020) [11] proposed the fundamentals of the GAN and autoencoders, which are essential for deep fake generation. framework for fake image creation can be found by HarshalVyaset *al*., (2020) [12] where the authors discusses auto encoders, which are frequently employed in face swapping. More interestingly in the authorsBahar Uddin Mahmud *et al., (2021)* [10] *proposed* the basics of GAN, auto encoders and several open source available tools. Another application of DNN for fake image creation can be found by Swathi P *et al.(2021)* [8] the authors purposed the variations of GAN's like StarGAN, CycleGAN, *etc.* Andrey Kuznetsov *et al*., (2021) [9] summarizes how the model is trained on the prominent object area, by removing the original background from the input image. After the model is trained, it does interference of the kept background and face swap region. Sreeraj Ramachandran*et al.,(2021)* [6]. It covers the fundamentals of identity swapping, face swapping, expression swapping, *etc.* that are frequently used when making deep fakes. Nick Dunard*et al*., (2021) [7] introduces the basics of Style GAN and DCGAN, two upgraded forms of GANs.Asad Malik *et al., (*2022*)* [2]explained the principles of deepfake production, including whole face synthesis, *etc*., as well as numerous technologies involved like StyleGAN. It also includes several real datasets.Tao Zhang *et al.,* (2022) [4] performed on some deep fake creation techniques, such as Deep Generative Network, RNN, and GAN.Felix Juefei-Xu *et al., (2022)[5]* explained a thorough summary of the several publications that have been written about deepfake generating techniques including attribute manipulation, expression swap, *etc.*HrutujaSatpute*et al,(2022)* [3]authors proposed information about GAN and its uses in creating the driving video from the input image.

III. DESCRIPTION OF DATASETS

For deep fakes, datasets are essential, i.e., an abundance of photos and videos to feed the data to an ML system. This section's Table 1 gives an example of the many dataset types that have been used for deep fake creation by various writers. Forensics data can be divided into two basic categories: traditional datasets and deep fake datasets. Conventional datasets are built manually with a lot of work and under very particular guidelines, such as camera specifications, location-specific requirements, etc.

TABLE I. DESCRIPTION OF DATASETS USED IN FAKE CREATION

S.No	Dataset Description	Dataset Link
1.	It contains 10,533 training images and 5,019 test images. These sample images are gathered from ImageNet DET.	The DUTS
2.	It consists of real audios of 1,251 celebrities collected from youtube and the total number of available samples is more than 100,00.	The VoxCeleb1
3.	It consists of real face images of 2,622 different individual ones and the total images count is approximately 2.6 million.	VGG Face
4.	It consists of 494,414 face images of 10,575 different real individual ones scraped from the web.	CASIA-WebFace
5.	It consists of more than 200k celebrity images with 40 different comments i.e smiling, mustache, etc. added with them.	CelebA
6.	It consists of images of 9131 different individual ones and the total image count is approximately 3.31 million.	VGGFace2
7.	It includes 70,000 high quality images (1024px X 1024px) and these images have variations related to image background, etc. It was taken from Flickr.	FFHQ
8	It consists of 4,753,320 faces of 672,057 different individual ones.	MegaFace
9	It consists of images of 10 different categories eg.bedroom, etc. It includes train images in the range 120k to 300k. And also has 300 validation	LSUN

	images and 1000 train images of each category.	
10	It consists of more than 14000 real images of indoor and outdoor captured from 73 digital cameras which were taken from 25 different models to get also the device related information.	Dresden Image
11	It consists of more than 35000 real images and 1914 real videos captured from 35 devices which are of various brands.	Vision

IV. TOOLS FOR DEEPFAKE GENERATION

An overview of the most popular deepfake tools is shown in Table2. It provides information about the programme including its name, github repository URL, features, and the technology it employs to produce fake images.

TABLE II. VARIOUS TOOLS USED FOR DEEP FAKE GENERATION

Toolkit	Github Link	Characteristics
Deep FaceLab	https://github.com/iperov/DeepFaceLab	There are several face extraction techniques that are supported, including dlib, MTCNN, S3FD and manual (Deep Face Lab, 2016).
Faceswap	https://github.com/deepfakes/faceswap	The Autoencoders, which have an encoder and a decoder, are used. Additionally, a single encoder is utilized for all input images, but various decoders are used for various outputs.
Face Swap-GAN	https://github.com/shaoanlu/faceswap-GAN	Utilizing perceptual loss and adversarial loss, the auto-encoder architecture is improved.
DFaker	https://github.com/dfaker/df	The face is recreated using the DSSIM loss function. The Keras library was used to carry this out.
Avatar Me	https://github.com/lattas/AvatarMe	From a range of "in-the-wild" pictures, construct 3D faces. Realistic 3D faces with a resolution of 4K * 6K can be produced from a single low-resolution image.
Deep Fake tf	https://github.com/StromWine/DeepFake_tf	Identical to DFaker in implementation, though tensor flow
FaceShifter	https://lingzhili.com/FaceShifterPage	High-fidelity face swapping is possible by making use of and including the target attributes. The tool can be deployed for any new face pair without needing training in that particular field.
DiscoFaceGAN	https://github.com/microsoft/DiscoFaceGAN	Using separate latent characteristics including identification, emotion, position, and brightness, create virtual people's facial images. Adversarial learning ought to integrate 3D priors .
FSGAN	https://github.com/YuvalNirkin/fsgan	A model of face acting out and switching that can be used to pair of faces without any training. Adapt to alterations in posture and facial expression.
MarioNETte	https://hyperconnect.github.io/MarioNETte	A method that preserves the uniqueness of the target while reenacting faces in a few shots.
Few-Shot Face Translation	https://github.com/shaoanlu/fewshot-face-translation-GAN	Convictions from FUNIT and SPADE modules that are derived from semantics should be integrated.
Neural Voice Puppetry	https://github.com/keetsky/NeuralVoicePuppetry	A system for creating facial videos using aural input of a talking head using 3D face rendering with the sounds of another person.

V. CLASSIFICATION OF TECHNIQUES USED FOR FAKE IMAGE CREATION

Deep learning and machine learning are widely used in computer vision nowadays for the detection and prediction of deep fakes. The prediction of fraudulent picture detection has made use of similar preconceptions. Table 3 in this section provides an overview of the various contemporary techniques that various writers have utilised to create phoney images.

TABLE III. TAXONOMY OF CONTEMPORARY TECHNIQUES FOR FAKE IMAGE CREATION

Ref.	**AutoFn coders**	**GAN**	**RNN**	**CNN**	***U ̂ 2 Net***	**Motion Cosegmentation**	**DNN**	**Wave Net**	**VAE**	**Stule GAN**
Harshal Vyas [12]	✓									
Rahul Kataryet *al.,* [11]	✓	✓		✓						
Bahar Uddin Mahmud *et al.,* [10]	✓	✓	✓							
Andrey Kuznetsov [9]	✓	✓			✓	✓				
Swathi P *et al.,* [8]	✓	✓		✓			✓			✓
Nick Dunard [7]		✓		✓						✓
Sreeraj Ramachandran et al., [6]	✓	✓								
Felix Juefei-Xu *et al.,* [5]	✓	✓							✓	✓
Tao Zhang [4]		✓	✓	✓				✓		✓
HrutujaSatpute *et al.,* [3]	✓	✓						✓		
Asad Malik *et al.,* [2]		✓	✓	✓						✓
Thanh Thi Nguyen *et al.,* [1]	✓	✓								✓

VI. EVALUATION PARAMETERS

Many different types of parameters are employed by various authors to calculate and evaluate fake picture detection utilizing machine and deep learning, as mentioned in the present section [26, 27]. Comparing accuracy is one of the most popular performance dimension metrics in machine learning. The following written formula can determine accuracy:

A. Model accuracy

The following equation can be used to determine accuracy rate. Its formula is the proportion of all correctly identified picture sample counts to all evaluated sample counts. The equation for calculating accuracy rate is shown below. It is defined as the ratio of the total number of correctly identified image samples to the total number of samples evaluated.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

B. AUC

AUC (Area under the ROC Curve) AUC can be used to classify various thresholds as a composite feature. The likelihood that the model ranks a random example of a positive action higher than one of its negative counterparts is known as AUC.

C. Precision

The ratio of True Positives to All Positives is known as precision. It might be the proportion of cases among people who are actually afflicted with heart problems that we can actually identify as having them. True Positives (TP) and False Positives (FN).Mathematically,

$$Precision = \frac{(TP)}{(TP+FP)} \quad (2)$$

D. Recall

Recall is the rate of True Positives to the accumulation of True Positives and False Negatives.

Mathematically:

$$Recall = \frac{(TP)}{(TP+FN)} \quad (3)$$

VII. CONCLUSION

The increasing rate of applications in the field of fake face detection is the motivating factor behind this research. Also, a number of researchers have been researching in this topic using cutting-edge techniques like deep learning-based techniques, analytical techniques, and traditional machine learning-based techniques. This study classifies a variety of modern ways for fabricating images, assisting the research community in selecting the optimal strategy for their needs.

VIII. FUTURE SCOPE

Deep learning approaches have achieved the maximum accuracy, but additional work needs to be done to improve their performance, according to a detailed examination and assessment of numerous research techniques.Hence, there is still room for development in this area of research in the future.

REFERENCES

[1] Nguyen, Thanh Thi, Quoc Viet Hung Nguyen, Cuong M. Nguyen, Dung Nguyen, Duc Thanh Nguyen, and SaeidNahavandi, "Deep learning for deepfakes creation and detection,"Ar.Xiv preprint Ar.Xiv, vol.1909, pp. 11543-11573, April 2019.

[2] Malik, Asad, Minoru Kuribayashi, Sani M. Abdullahi, and Ahmad Neyaz Khan, "DeepFake detection for human face images and videos: a survey," IEEE Access, vol 10, pp.18757-18775, Sep 2022.

[3] Zhang, Tao, "Deepfake generation and detection, a survey," Multimedia Tools and Applications 81, vol. 5, pp 6259-6276, May 2022.

[4] Juefei-Xu, Felix, Run Wang, Yihao Huang, Qing Guo, Lei Ma, and Yang Liu, "Countering malicious deepfake: Survey, battleground, and horizon, " International Journal of Computer Vision, pp.1-57, April 2022.

[5] Ramachandran, Sreeraj, Aakash Varma Nadimpalli, and AjitaRattani, "An experimental evaluation on deepfake detection using deep face recognition," In 2021 International Carnahan Conference on Security Technology (ICCST) IEEE, pp. 1-6, August 2021.

[6] Dunard, Nick, "Deep Fakes: The Algorithms That Create and Detect Them and the National Security Risks They Pose," James Madison Undergraduate Research Journal (JMURJ),vol. 8, pp. 1- 5, 2020.

[7] Swathi, and Saritha Sk, "Deepfake creation and detection: A survey," In 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA) IEEE, pp. 584-588, April2021.

[8] Kuznetsov, Andrey, "A new deep fake method based on background removal," In 2021 International Conference on Information Technology and Nanotechnology (ITNT) IEEE, pp. 1-4, September 2021.

[9] Mahmud,Bahar Uddin, and AfsanaSharmin, "Deep insights of deepfake technology: A review, " arXiv preprint arXiv, vol.2105, pp.134-152, March 2021.

[10] Katarya, Rahul, and Anushka Lal, "A study on combating emerging threat of deepfake weaponization," In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) IEEE, pp. 485-490, 2020.

[11] Vyas, Harshal, "Deep Fake Creation by Deep Learning," Extraction 7, vol. 07 March 2020.

[12] Kamali Singla, Atul Garg, Vinay Kukreja and Deepali Gupta, “Rice Diseases Multi-Classification: An Image Resizing Deep Learning Approach”, International Conference on Decision Aid Sciences and Applications (DASA-2021), College of Business Administration at The University of Bahrain, Sakheer – Kingdom of Bahrain, from 7-8 December, 2021

[13] IshaKansal, Renu Popli, Geetanjali Kapoor, Dhawan Singh, "Fusion based fast de-fogging for foggy images," AIP Conference Proceedings. Vol. 2357. No. 1, 2022.

[14] Kansal, Isha, Renu Popli, and Monika Sethi, "Effect of non uniform illumination compensation on dehazing/de-fogging techniques," AIP Conference Proceedings, Vol. 2357, No. 1. 2022.

[15] Zhenliang, WangmengZuo, Meina Kan, Shiguang Shan, and Xilin Chen, "Attgan: Facial attribute editing by only changing what you want, " IEEE transactions on image processing 28, vol. no. 11, pp. 5464-5478,2019.

[16] Choi, Yunjey, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo, "Stargan: Unified generative adversarial networks for multi-domainimage-to-image translation, " In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8789-8797,2018.

[17] Perarnau, Guim, Joost Van De Weijer, Bogdan Raducanu, and Jose M. Álvarez, "Invertible conditional gans for image editing, " arXiv preprint arXiv:1611.06355 (2016).

[18] Thies, Justus, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Niebner, "Face2face: Real-time face capture and reenactment of rgb videos, " In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2387-2395, 2016.

[19] Suwajanakorn, Supasorn, Steven M. Seitz, and Ira Kemelmacher-Shlizerman, "Synthesizing obama: learning lip sync from audio, " ACM Transactions on Graphics (ToG) 36, vol. 4,pp. 1-13, 2017.

[20] Chauhan, Ruby, Renu Popli, and Isha Kansal. "A Comprehensive

[21] Review on Fake Images/Videos Detection Techniques." 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). IEEE, 2022

[22] Hyun, Sangeek, Jihwan Kim, and Jae-PilHeo, "Self-supervised video gans: Learning for appearance consistency and motion coherency, " In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10826-10835, 2021.

[23] Karras, Tero, Timo Aila, Samuli Laine, and Jaakko Lehtinen, "Progressive growing of gans for improved quality, stability, and variation, " arXiv preprint arXiv, vol. 1710-17196 , 2017.

[24] Karras, Tero, Samuli Laine, and Timo Aila, "A style-based generator architecture for generative adversarial networks, " In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4401-4410, 2019.

[25] Garg, A., R. Popli, and B. S. Sarao. "Growth of digitization and its impact on big data analytics." IOP Conference Series: Materials Science and Engineering. Vol. 1022. No. 1. IOP Publishing, 2021

[26] Li, Jia, Zhaoyang Li, Jie Cao, Xingguang Song, and Ran He, "Faceinpainter: High fidelity face adaptation to heterogeneous domains," In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 5089-5098,2021.

[27] Maninder Singh, Rajeev Kumar, Dinesh Tandon, Pallavi Sood, Manish Sharma, "Artificial intelligence and iot based monitoring of poultry health: A review." IEEE International Conference on Communication, Networks and Satellite (Comnetsat). IEEE, 2020..

[28] Shao, Chonglei, Preet Kaur, and Rajeev Kumar. "An improved adaptive weighted mean filtering approach for metallographic image processing." Journal of Intelligent Systems Vol. 30 No.1, 2021.

[29] Amin, Dar Masroof, and Atul Garg. "Performance analysis of data mining algorithms." Journal of Computational and Theoretical Nanoscience, Vol.16, No.9, 2019.

[30] Gupta, Reena, Raj Kamal, and Ugrasen Suman. "A QoS-supported approach using fault detection and tolerance for achieving reliability in dynamic orchestration of web services." International Journal of Information Technology,Vol. 10,pp. 71-81, 2018.

[31] Rahman, Abdul, and Mirza Mohammed Sufyan Beg. "Face sketch recognition: an application of Z-numbers." International Journal of Information Technology, Vol. 11, pp. 541-548, 2019.

[32] Mullai, A., and K. Mani. "Enhancing the security in RSA and elliptic curve cryptography based on addition chain using simplified Swarm Optimization and Particle Swarm Optimization for mobile devices." International Journal of Information Technology 13 (2021): 551-564.

[33] Kshetri, Nir. " Blockchain's roles in meeting key supply chain management objectives." International Journal of information management 39 (2018): 80-89.

[34] Chakraborty, S., Mandal, S.B. & Shaikh, S.H. Quantum image processing: challenges and future research issues. International Journal of Information Technology 14, 475–489 (2022). https://doi.org/10.1007/s41870-018-0227-8

[35] Chandra, M.A., Bedi, S.S. Survey on SVM and their application in image classification. International Journal of Information Technology, 13, 1–11 (2021). https://doi.org/10.1007/s41870-017-0080-1

[36] S.K.Muttoo1 and Sushil Kumar, "Data Hiding in JPEG Images", International Journal of Information Technology, vol. 1 no. 1, pp. 13-16

[37] Giri, K.J., Bashir, R. A block based watermarking approach for color images using discrete wavelet transformation. International Journal of Information Technology, 10, 139–146 (2018).

[38] Nandy, N., Banerjee, D. & Pradhan, C. Color image encryption using DNA based cryptography. International Journal of Information Technology,.13,533–540(2021).

[39] Yang, Jian, Amit Sharma, and Rajeev Kumar. "IoT-based framework for smart agriculture." International Journal of Agricultural and Environmental Information Systems (IJAEIS) 12.2 (2021): 1-14.