Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 2153–2162

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

A GAN-Based Model of Deepfake Detection in Social Media

Preeti [a*], Manoj Kumar [b*], Hitesh Kumar Sharma [a,b]

[a] *Research Scholar, School of Computer Science, University of Petroleum and Energy Studies (UPES),Dehradun, 248007 India.*
[b] *Associate Professor, Engineering and Information Sciences, University of Wollongong , Dubai Knowledge Park ,Dubai, UAE*
[a,b] *Associate Professor, School of Computer Science, University of Petroleum and Energy Studies (UPES),Dehradun, 248007 India.*

Abstract

DeepFake uses Generative + Adversarial Network for successfully switching the identities of two people. Large public databases and deep learning methods are now rapidly available because of the proliferation of easily accessible tools online. It has resulted in the emergence of very real appealing fake content that produced a bad impact and challenges for society to deal. Pre-trained generative adversarial networks (GANs) that can flawlessly substitute one person's face in a video or image for that other are proving supportive for implementing deepfake. This paper primarily presented a study of methods used to implement deepfake. Also, discuss the main deepfake's manipulation and detection techniques, and the implementation and detection of deepfake using Deep Convolution-based GAN models. A study of Comparative analyses of proposed GAN with other exiting GAN models using parameters Inception Score "IS" and Fréchet Inception Distance "FID" is also embedded. Along with the abovementioned, the paper discusses open issues and future trends that should be considered to advance in the field.

Keywords

Digital Forensics; Image Vision; Deep Learning; Generative adversarial network; Deep Fakes; Media Forensics, Face Manipulation; Face Recognition.

1. Introduction

Deepfake is related to the generation of fake photographs or videos consisting of face of a specific person which gets replaced with another person face utilizing convolutional neural networks especially Generative Adversarial Networks (GAN).Recent social media controversies involving celebrities' faces being replaced in pornographic movies are a huge disgrace to one's character and create permanent harm to famous and even ordinary people's identities that demand methods to detect Deepfake [1].

In addition to that, it is a great danger to the security of biometric information and rises to other counterfeiting and fraudulent activities [2]. Many faces swapping software Face2Face[3], mobile Apps like Snapchat pre-trained generative adversarial networks can very easily implement DeepFake.

* Corresponding author.
E-mail address: [a] preetiii.kashyup@gmail.com, [b] wss.manojkumar@gmail.com

1877-0509
10.1016/j.procs.2023.01.191

Ian J. Good fellow and his team created the first GAN model and introduced it in the year 2014[4]. According to their research, Minimax is a competition among Discriminator "D" and Generator "G". "D" tries increment the possibility that accurately distinguishes true(real) and phoney (log D(x)), while "G" tries to decrease the possibility that D will forecast that its outputs seem to be counterfeit "log(1-D(G(z))". In the aforementioned research paper, the GAN error rate estimated by calculation of loss can be noticed in equation 1 [4].

$$GminDmax\, V(D,G) \;=\; Ex \sim pdata(x)[lodD(x)] + Ez{\sim}pz(z)[log(1 - D(G(z)))] \qquad \textbf{(1)}$$

Where; G stands for Generator, D stands for Discriminator, Pdata(x) = real-world data distribution, P(z) = generator distribution, x = Pdata sample (x),z = a sample taken from P (z), D(x) denotes a discriminator network, G(z) denotes the generator network.

According to Mirza et al. [5], generative adversarial networks (GANs) may be used to produce phoney pictures and movies that are challenging for humans to identify from the genuine thing. Before using these algorithms to create phoney photos and videos, they are trained on a data set. A significant quantity of deepfake media training data is required for this kind of deepfake model.

Many other approaches have been proposed in previous research that help in the detection of Deep fakes. A paper entitled "Effective and Fast Deepfake detection approach" by Younus et al. [6] is based on the Haar wavelet form that provides a method for detecting deepfake movies using the haar wavelet transform. The method is entirely based on the defined logic that the deepfake algorithm construct fake faces of a specified size and resolution during video production. An additional function (blur) must be introduced to the phoney or artificial (faces) to correlate and adjust the particulars of the true identity on the videos "original".

One another paper by Ciftci et al. [7] proposed a method for deep fake source identification that relies on biological cues to understand residuals They believe that this is the first instance of deep fake source detection using biological signals. A 93.39 accuracy rate for source recognition from four deep fake generators was achieved on the Face Forensics++ dataset as part of their experimental validation of this system, which also included many ablation experiments.

Mirsky et al. [8] identify deepfakes, researchers focused on reenactment techniques (such as altering an expression, lips, posture, or entire body) and converting methods (such as swapping or transferring a target's face). Verdoliva et al. [9] provide an overview by distinguishing between traditional approaches (for instance blind approaches data used for training, depending on sensor-based(one-class) and model-based methods(other class), and supervised methods) and approaches based on deep learning technology (like vCNN models).Shobhit et al.[22] analyze the image and video modification kinds, popular methods, and forgery detection approaches. Syed Sadaf Ali et al.[23] develop advanced deep learning approaches for spotting double image compression forgeries. Huang et al.[24] use augmentation of partial data and single sample clustering to speedup FakeLocator's universality across DeepFake techniques.

The paper is assembled into various sections detailed in following manner: Section 1 defines the introduction about deepfake and various methods for its creation and detection. Section 2, reviews prior work in the domain of Deepfake techniques and GAN approaches to implement deep fakes. In Section 3, some background on the framework design and specifics regarding the architecture of DCGAN for implementing and detection of Deepfakes using GAN discriminator is discussed.

Section 4, provides results and analysis of the performance of proposed deep convolution GAN shown in Figure 4 and Figure 5 .There are also two values tables which shows the attained values of accuracy and loss function of generator and discriminator with 10 successive iterations. The table 3 elaborate the comparison values indicating "IS" Inception Distance and "FID" Fréchet Inception Distance with previous exiting models. Finally Section 5 presents the conclusion of the paper.

2. DeepFake (Detection and creation Techniques)

Facial alterations are classified into four groups based on the degree of transformation. The four types of manipulations include complete expression swap, face synthesis, identity swap, and attribute manipulation. These four basic methods of face modification are well-known in the scientific world and have gotten the greatest attention in recent years. Figure 1 [10] depicts a graphic summary of each face alteration group. In order of increasing manipulation complexity, each of them is explained below.

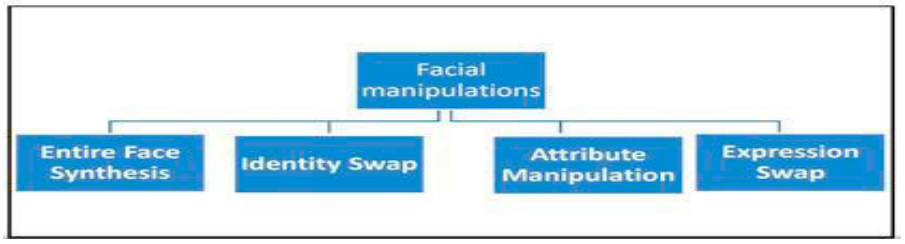

Figure 1 Different Fake manipulation techniques.

Auto-encoders "AE" and "GAN" Generative Adversarial Networks are two examples of deep learning approaches that make it easier to automatically construct artificial faces or edit a single actual identity mentioned referencing an image/ video by eliminating the need for labor-intensive human formatting. Concluding many open-source software and smartphone applications like "ZAO" and "FaceApp" smoothly accessible, allowing anybody to produce fake photos and videos with little to no training [11]. Media investigations have often included the use of in-camera fingerprinting, which examines the fundamental fingerprints embedded by the camera's equipment, including hardware (H/w) and software(S/w), like interpolation, lens "optical" one, compression, and a colour filter array.

2.1 Entire Face synthesis

These technologies produce excellent results, resulting in high-quality, highly realistic facial images. Recently proposed powerful GANs like StyleGAN [12] method that creates whole face photos that do not exist. The fast development of novel Generative Adversarial Network (GAN) architectures has advanced the methodology in image manipulation to the point that synthetic images are frequently regarded as genuine by humans. Face synthesis can be used in a variety of ways. People in the film business want to create virtual human characters that are indistinguishable from genuine human characters. People have been attempting to build interactive and lifelike human characters in video games. Facial synthesis techniques can also be used for face recognition. The major publicly datasets used for entire face synthesis include 100K-Generated Images, 100K Faces, iFakeFaceDB, DFFD, etc. Figure 2 [13] shows the Real and Fake images used in different Deepfake facial manipulation groups.

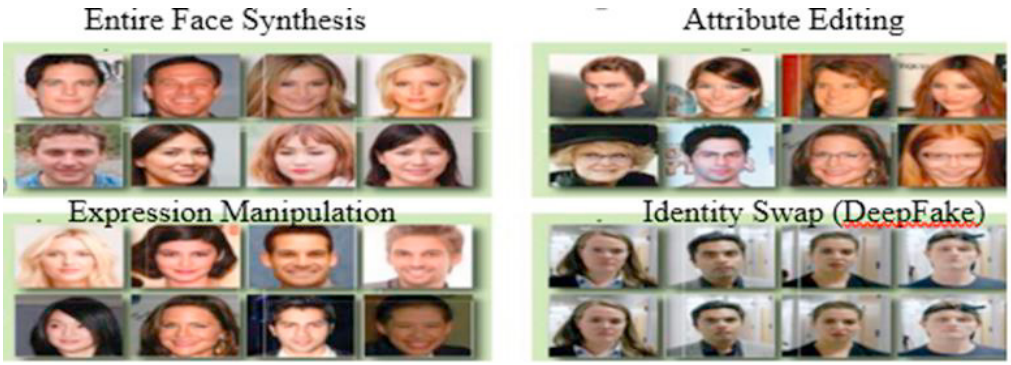

Figure 2: Four different types of DeepFake face manipulation techniques [13].

2.2 Attribute editing

Face-lifting is a cosmetic procedure that involves making minor adjustments to a person's face, such as changing their hair or skin tone, adding or removing facial hair, a beard or moustache, or even adding or subtracting years from their appearance. In order to identify artificial intelligence-generated fake faces, Wang et al. [14] propose Fake Spotter, a novel method based on monitoring neuron activity. Using neuron coverage and interaction as evaluation criteria for deep learning systems has been shown feasible by tests. Example: the widely used FaceApp app for smartphones, which is prone to this kind of manipulation. Virtually testing out new hairdos, eyewear, and makeup would be only the beginning of what this technology may provide. Face switching, relighting, and cosmetics transfer are only a few of the many uses for the method proposed by Xu et al. [15]. Separating texture and color from identity, expression, position, lighting, and region-wise style codes is achieved with the use of 3D priors.

2.3 Identity's Swap

This process of editing video files replaces one person's face with another. Recent developments in machine learning have made it possible to generate convincing fakes utilizing only a single picture and five seconds of a target's audio. It's possible that many other businesses, including the entertainment business, may benefit from this kind of manipulation. However, it might also be used maliciously, for example, in the creation of celebrity pornographic films, hoaxes, or financial fraud. One new approach proposed by Schroff [16] is to create Fake videos using weights in FaceNet. Multi-Task Cascaded Convolution Networks, as demonstrated by Li et al. [17], are used to improve the accuracy of detections and the consistency of face alignment.

2.4 Expression's Manipulation

Facial expression manipulation often called facial re-enactment, involves simulating or changing a person's facial expressions in order to create a new one. Many facial expression alteration techniques, such as computer graphics methods are essential for tasks such as "2D"/"3D" image rendering, flow mapping, and wrapping of images. Even though these strategies may yield photorealistic pictures with high resolution, the complicated and time-consuming processes involved are not ideal for everyday use. Recently, some research has used generative adversarial networks (GANs) to alter facial features such as expressions. Chen et al. [18] went on to simulate transitional zones across domains by modifying picture properties. In considerations of face attribute translation and expression synthesizing, the two methods are comparable to one another. However, only eight global expressions could be synthesized. Ding et al. [19] created an expression network using generative adversarial for alteration of expressions inculcating different "expression intensities" (ExprGAN). StarGAN was invented by Choi et al. [20] for multi-domain image translation.

2.5 Datasets

All the techniques described make use of some of the following available public datasets for deepfake implementation and detection One, DDFD; 2nd, 100K-Faces; 3rd, FFHQ ; 4th, Deepfake TIMIT; 5th, VGGFace2; 6th, the eye-blinking dataset; 7th, CASIA-WebFace. In this research, we use convolution neural GAN using the celeb A dataset, which is a publicly available dataset including more than 200K factual photos of celebrities annotated with 40 variables.

3. Implementation of Deepfakes using Deep Convolutional GAN

For the implementation of Deepfake various GAN-based techniques are used. With the advancement of GANs, it is easy to not only create swapped images or videos but can easily create non-existent faces. With the growth of GAN, it is tough to determine the originality of images, especially in the social media environment. This paper discuss an important GANs architecture called Deep Convolution GAN for the implementation of Deepfakes.The model is implemented using celebA dataset and implemented using foundation and idea from dcgan_faces_tutorial [21].

3.1 Architecture

Deep convolutional GAN is one of the most widely used and successful GAN network architectures. Many of the layers in the design are convolution layers, with fully connected layers present no max pooling. For both down sampling and up sampling, it employs convolutional stride and transposed convolution. Deep convolutional GAN employs neural layers of convolutional and convolutional-transpose networks for implementing generator as well as discriminator functions. As an illustration of a generator, Radford et al. [22] of "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks" offered the network architecture seen below.

- Use convolutional phase in lieu of all-max pooling.
- Up-sampling is accomplished by use of transposed convolution.
- If there are any layers that are totally linked, they should be discarded.
- To improve performance, it is recommended to use batch normalization on all levels except the out-put layer (generator) and the in-put layer (discriminator).
- Discriminator utilizes LeakyReLU, while the generator uses tanh (except for the output).

A 100x1 noise vector "z", is feeded to convolutional network"CNN", and output generates 64x64x3 vector of random values called G(Z). The network is reshaped by convolutional layers using the learned equation (N+P — F)/S + 1. The N parameter (Height/Width) in the figure advances from 4 to 8 to 16 to 32, there doesn't seem to be any padding, the gaussian filter factor F is 5x5, and the step is 2.

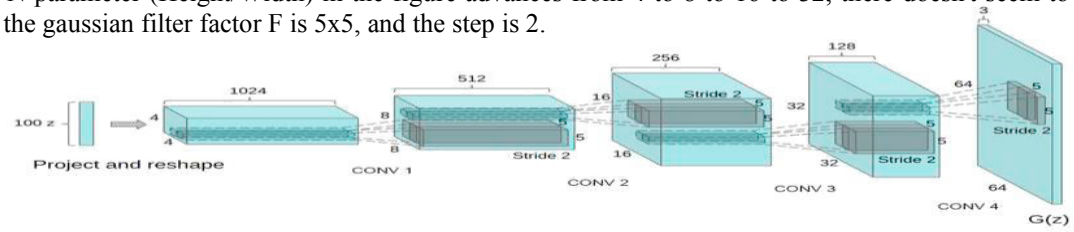

Figure 3: Deep Convolutional GAN generator's network architecture [22]

3.2 Details of DC Adversarial Network

- The Scaling to fit inside the [-1, 1] range of the tanh activation function was the sole preprocessing done to the training pictures. All models were trained via 128-batch mini-batch "stochastic gradient descent" (SGD). All weights were calculated using distribution "Normal" having mean of zero and "SD" standard deviation of 0.02. Leak's slope was set to 0.2 in every "LeakyReLU" model.

- In this example, the discriminator comprising stride layers "convolution", Normalisation layers"(Batch), and "Activation" function "LeakyRelu".

- An picture of 3x64x64 pixels can be uploaded. Layers of batch normalisation, ReLU activations, and convolution-transpose layers compose the generator. The resulting picture will be 3x64x64 RGB. The network extends to 1024 by 4 by 4 from 100 by 1! This layer is addressed as "Projection and reshaping layer".

3.2.1 Optimizers and Loss Functions

In BCE Loss i..e. Binary Cross-Entropy , is a loss generating function designated for utilized by the DC GAN's discriminator and generator. It's crucial to note that this function tabulates the objective function's "log (1-D(G(z))" and "log (D(x))" constituents. Adam optimization techniques are employed with a computed value of 0.0002 and an Attempt should be made of 0.5. Gaussian function commonly described as fixed noise is utilized to monitor the generator's evolution as it learns.

3.2.2 Training

Different mini-batches are utilized for real and fake photographs in order to maximize log D (G(z)), which is G's aim function. There are two parts to the training. The Discriminator is brought up to date in Section 1, and the Generator is brought up to date in Section 2.

Function "log (D(x)) + log (1-D) (G (z)))" should be maximized as a discriminator. A forward run "D" yields the "loss (log (D(x)))" while a backward pass yields the gradients; together, these steps provide a collection of "real" samples from the set (training). In order to produce more convincing forgeries, Generator must minimize log (1-D (G (z))). According to Good fellow, it doesn't supply sufficient gradients in the starting of the learning advance process. Rather, we aim to optimize log (D (G (z)) as a means of correction. This is done by first computing the loss of G with the help of real labels "GT", then executing the gradients of "G" pass(backward direction), and then incrementing the parameters of G using optimizer's step. The training pictures for DC GAN are shown in Figure 4. For the GAN, these stages constitute the training algorithm [4].

Training Algorithm:

Step1: For the available number of looping iterations do repeat
Step2: For q number of steps repeat

- Extract sample minibatch of 'r' examples {z1, z2… zr} from noise priors p (g (z)).
- Take a small subset of 'r' data points (x1, x2,..., xr) from the distribution pdata (x).
- Raising the discriminator's stochastic gradient represents an update:

$$\nabla\theta_d \ \frac{1}{r}\sum_{i=1}^{r}[\log D\ (x(i\ )) + \ \log\ [1 - D\ (\ G\ (\ z\ (\ i\ )))]. \quad \textbf{(2)}$$

End for.
Step 3: Take a small subset of 'r' samples (z1, z2, zr) and use the prior distribution of noise p (g (z)) for extraction.
Step 4: In order to implement a generator update, the stochastic gradient must be lowered.

$$\nabla\theta_d \ \frac{1}{r}\sum_{i=1}^{r} \log\ [1 - D\ \Big(\ G\ (\ z\ (\ i\ ))\Big). \quad \textbf{(3)}$$

End of For loop.

Figure 4: Images used by deep convolution GAN during Training

4. Results and Discussion

Here in result section we shows two alternative outcomes. Figure 4 shows the batch of images used by proposed GAN for its training. A set of G's new false photographs are generated by converting a set of real data from the celebA dataset.. Figure 5 discusses the performance of deep convolution GAN for 10 successive iterations. It is structured between three different plots in which first part (a) signifies the comparison study of Generator Loss with Discriminator Loss (%) in successive iterations. It is observed that with successive training epochs discriminator is getting better performance as compared to the generator with less loss percentage values. Part (b) analyses the performance of GAN Accuracy with respect to Generator and Discriminator Loss. It shows how D and G's losses varied over time and the accuracy achieved by GAN increases respectively during each iteration. The Loss values decrease with iterations signifies that the model converges to a significantly better optimum and decreases total training time, resulting in higher accuracy values. Part (c) produces the GAN Accuracy Cycle with successive iterations with increasing accuracy values reaching a maximum (100%) by end of 10 iterations. It shows that the GAN is working incredibly well with defined conditions and architecture. Tables 1, 2, and 3 contain a variety of model results to further highlight the point. In Table1 you can see the results of a 10-iteration deep convolution GAN discriminator's classification in percentage. This means that the deep convolution GAN discriminator improves its ability to tell fake from real as training iterations progress. In Table 2, we can see the GAN model's accuracy, generator and discriminator losses, and other metrics. Table 3 shows a comparative study of the deep convolution model with other existing models and shows optimized values of parameters IS and FID.

Merely described, the GAN version incorporates the Inception Score “IS” and the Fréchet Inception Distance (FID" to estimate the quality of the taken pictures (FID). In the FID, we look at how well the distribution of generated images resembles that of the real images we used to instruct the generator, whereas in the IS, we look at how consistently they follow a given distribution. If the FID is less, then the distribution of either the experimental results is more like the distribution of the actual statistics. A GAN's output is assigned a "Inception score" that measures its overall realism. It is advantageous to have a better score. It indicates that your GAN is capable of producing a wide variety of unique images. In this research both these parameters obtained optimized values 1.074 and 49.3 respectively. In a nutshell, it has been analyzed that losses of discriminator get decrease while of generator increases with each successive iteration which shows the effective fake detection capacity of deep convolutional GAN. The efficient adversarial training of both generator and discriminator without suffering from limitations like mode collapse and convergence help to achieve good accuracy values of proposed GAN with each successive iteration. It is also discovered that by grouping the classes, the values of the training parameters can be improved much more. This allows the discriminator to differentiate between sub-batches and evaluate whether or not a batch is authentic which helps to increase the quality of images generated by the generator as listed in Figure 6.

(a) (b) (c)

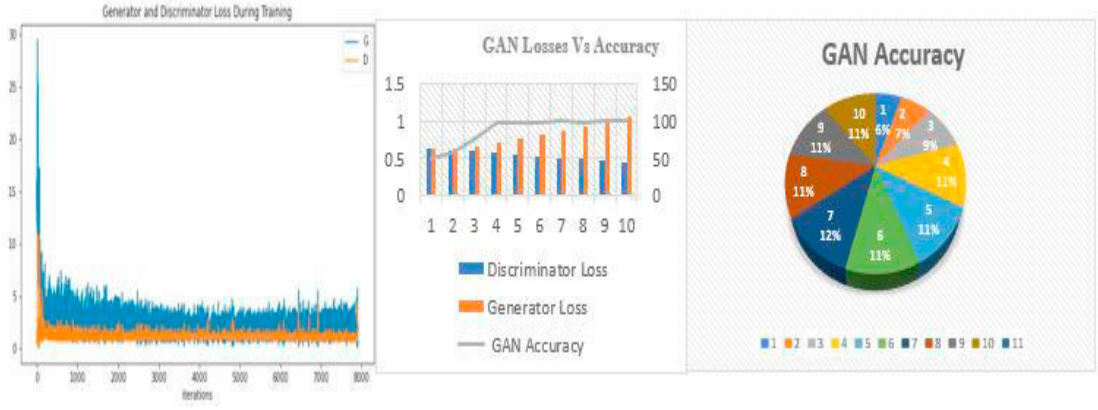

Figure 5: Deep Convolution GAN Performance Plots a: Generator Loss vs Discriminator Loss ; b: GAN Loss Vs GAN Accuracy; c : GAN Accuracy Cycle with successive iterations

Table 1: Deepfake detection percentages (%) of Discriminator

Successive Iterations	Detection Percentages (Real)	Detection Percentages(Fake)
1	37.5	6.25
2	92.1875	9.375
3	93.75	28.125
4	95.3125	48.4375
5	95.3125	75.0
6	87.5	96.875
7	92.1875	99.0324
8	85.9375	99.910
9	79.6875	100.0
10	81.25	100.0

Table 2: Deep Convolution GAN Losses of Generator and Discriminator (%) and Accuracy (%)

Successive Iterations	Discriminator Loss	Generator Loss	GAN Accuracy
1	0.623	0.637	50
2	0.61	0.637	57
3	0.596	0.67	76
4	0.58	0.708	99
5	0.558	0.754	98
6	0.535	0.806	97
7	0.509	0.866	100
8	0.497	0.935	99
9	0.474	1.001	100
10	0.45	1.071	100

Table 3: Comparative analyses of proposed GAN with other exiting GAN models using parameters IS and FID.

GAN-Variant	Year	Dataset	IS	FID
Proposed GAN (Deep Convolution GAN)	2022	CELEB A	1.074	49.3
DC-GAN [23]	2018	CIFAR10	6.69	35.6
		CIFAR100	6.20	41.8
DC-GAN [24]	2021	CIFAR10	6.69	42.5
DC-GAN [25]	2022	CIFAR10	7.06	42.23
		CIFAR100	6.87	44.18
C-GAN[26]	2021	CIFAR100	9.71	12.28
C-GAN[27]	2021	CIFAR10	7.10	4.31
		MNIST	9.87	36.67
S-GAN[28]	2019	CUB-200-2011	--	25.99 ± 4.26

Figure 6: Deepfake detection using Deep Convolutional GAN

5. Conclusion

Deep generative models, such as those employed by Deepfake, make it possible to generate synthetic data with a realistic appearance (such as photographs or movies) by merely presenting an algorithm in a huge number of different iterations. Deepfake is so hard to spot that sometimes not even humans can tell the difference. Deepfake Detection Challenge invites people to participate in discovering unique solutions for recognizing and avoiding falsified media. This paper suggested a deep convolution GAN detection model as a solution to the challenge. The proposed model is capable of working very well with relatively limited datasets by making use of noise for the diversity of data distribution to produce good (accuracy). Analysis has revealed that the suggested model's performance is excellent and consistent. The loss of discriminator is getting minimized compared to generator loss with successive iterations. Its fake detection strengthens with higher iterations. Adversarial training without mode collapse and convergence showed good predictive performance. It is also analysed that good accuracy can be achieved with fewer images under controlled conditions by optimizing the factors like a sufficient number of epoch cycles, normalized batch size of images, noise value, and effective model layers. In addition, the assessment parameters Inception Score “IS” and Fréchet Inception Distance “FID” obtained optimum condition having 1.074 and 49.3, respectively, indicating that the pictures that the proposed GAN model are of fine standards. In terms of future research, additional effort is still needed to improve the handling of small datasets and overcome GAN constraints like mode collapse, gradient descent and convergence issues. The generalization of GAN models is also a promising area to explore for further improvements.

References:

[1] Korshunov, P., & Marcel, S. (2018). Deepfakes: a new threat to face recognition? Assessment and detection. arXiv preprint arXiv:1812.08685.

[2] Xu, L. (2021, April). Face Manipulation with Generative Adversarial Network. In Journal of Physics: Conference Series (Vol. 1848, No. 1, p. 012081). IOP Publishing.

[3] Feng, D., Lu, X., & Lin, X. (2020, November). Deep detection for face manipulation. In International Conference on Neural Information Processing (pp. 316-323). Springer, Cham.

[4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Advances in neural information processing systems. Curran Associates, Inc, 27, 2672-2680.

[5] Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

[6] Younus, M. A., & Hasan, T. M. (2020, April). Effective and fast deepfake detection method based on haar wavelet transform. In 2020 International Conference on Computer Science and Software Engineering (CSASE) (pp. 186-190). IEEE.

[7] Ciftci, U. A., Demir, I., & Yin, L. (2020, September). How do the hearts of deep fakes beat? deep fake source detection via interpreting residuals with biological signals. In 2020 IEEE international joint conference on biometrics (IJCB) (pp. 1-10). IEEE.

[8] Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. ACM Computing Surveys (CSUR), 54(1), 1-41.

[9] Verdoliva, L. (2020). Media forensics and deepfakes: an overview. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910-932.

[10] Abdulredaa, A. S., & Obaida, A. J. (2022). A landscape view of deepfake techniques and detection methods. Int. J. Nonlinear Anal. Appl, 13(1), 745-755.

[11] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. D. beyond: A survey of face manipulation and fake detection. arXiv 2020. arXiv preprint arXiv:2001.00179.

[12] Kramberger, T., & Potočnik, B. (2020). LSUN-Stanford car dataset: enhancing large-scale car image datasets using deep learning for usage in GAN training. Applied Sciences, 10(14), 4913.

[13] Sabel, J., & Johansson, F. (2021). On the Robustness and Generalizability of Face Synthesis Detection Methods. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 962-971).

[14] Wang, R., Juefei-Xu, F., Ma, L., Xie, X., Huang, Y., Wang, J., & Liu, Y. (2019). Fakespotter: A simple yet robust baseline for spotting ai-synthesized fake faces. arXiv preprint arXiv:1909.06122.

[15] Xu, Z., Yu, X., Hong, Z., Zhu, Z., Han, J., Liu, J. & Bai, X. (2021). Facecontroller: Controllable attribute editing for face in the wild. arXiv preprint arXiv:2102.11464.

[16] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

[17] Li, M., Zuo, W., & Zhang, D. (2016). Deep identity-aware transfer of facial attributes. arXiv preprint arXiv:1610.05586.

[18] Chen, Y. C., Xu, X., Tian, Z., & Jia, J. (2019). Homomorphic latent space interpolation for unpaired image-to-image translation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2408-2416).

[19] H. Ding, K. Sricharan, and R. Chellappa,(2018). “ExprGAN: Facial expression editing with controllable expression intensity,” in 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, 2018, pp. 6781–6788.

[20] Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2018). Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 8789-8797).

[21] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

[22] Tyagi, S., Yadav, D. (2022) . A detailed analysis of image and video forgery detection techniques. Vis Comput . https://doi.org/10.1007/s00371-021-02347-4.

[23] Ali, S.S.; Ganapathi, I.I.; Vu, N.-S.; Ali, S.D.; Saxena, N.; Werghi, N.(2022). Image Forgery Detection Using Deep Learning by Recompressing Images. Electronics, 11, 403. https://doi.org/10.3390/electronics11030403.

[24] Y. Huang, F. Juefei-Xu, Q. Guo, Y. Liu and G. Pu,(2022). "FakeLocator: Robust localization of GAN-based face manipulations", IEEE Transactions on Information Forensics and Security.