2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 979-8-3503-8247-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/UPCON59197.2023.10434896

Deepfake Generation and Detection – An Exploratory Study

Diya Garg
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
garg.diya@chitkara.edu.in

Rupali Gill
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
rupali.gill@chitkara.edu.in

***Abstract*—Deepfakes generated through algorithms based on deep learning have obtained a lot of interest recently. Deepfakes are utilized to manipulate content (audio, video and image) with high realism. Deepfakes have been influenced using artificial intelligence to make it look like someone is saying or doing something that they never actually said or did. Deepfakes can be used for malicious purposes, such as to tarnish someone's reputation or to influence public opinion. Researchers are developing new methods to detect deepfakes, as it is a challenging to distinguish between real and fake content. Deep learning is a powerful tool that is usable to develop both deepfake generation and detection methods. This article provides an extensive review of the existing state of research in creation and detection of deepfakes. It covers the diverse techniques for creation and detection, and existing benchmark datasets.**

***Index Terms*—Deepfakes, Deep Learning, Forgery, Audio Deepfake, Video Deepfake, Image Deepfake, Face swap**

I. Introduction

The fast rate of spreading information through various platforms such as social media, television, and other online services provides an ease to timely communication. This has initiated concerns regarding the technology's potential for ethical and unethical usage by using an enormous amount of data [1]. Deepfakes are a sort of synthetic media that deploys or modifies pre-existing visual and audio content. These contents look to be real but are fully produced or altered. This can be done with advanced machine learning algorithms, particularly deep learning. The amalgamation of the words "deep learning" and "fake" makes the term deepfake. In deepfake, the source person's face is swapped by a targeted person mimicking their expressions, and acting like they are speaking the words that are being said by the source person as shown in Figure 1 [2].

The inception of this phenomenon dates back to 2017, marked by a Reddit user's utilization of deep learning methods to substitute an individual's face in explicit videos, giving rise to remarkably realistic fabricated content [4]. According to the Sensity survey, a company that monitors deepfake films online claims that, since 2018 the number has doubled every six months reaching 85,047 videos at the end of December 2020 [5]. Because of the rapid growth of technology, it is extremely simple to produce realistic forgery content by swapping faces which makes it exceedingly difficult to find the replaced traces. Deepfakes can be extremely harmful while being used to create fake news. Politicians, including US presidents Barack Obama and Donald Trump, have also fallen victim to deepfakes in the past [6]. In February 2020, two fake videos featuring Delhi BJP president Manoj Tiwari demonstrated the impact of deepfakes. These films showed him making a bilingual election offer in Haryanvi and English before the assembly elections [7]. These technologies pose a threat to both privacy and national security because they could be used in cyberattacks. Various apps have been commonly used for faceswap such as "DeepFaceLab" and "FakeApp" which make it simple for users to utilize for malicious causes by generating deepfakes. Although this technology is frequently employed for malevolent purposes with negative intentions, there are some good uses for it as well. This technology is heavily used to generate new works of art, fascinate audiences, and provide them with distinctive experiences. The instances of deep fakes and their uses are shown in Figure 2. A short time ago the Museum Salvador Dali in St. Petersburg, Florida offered visitors a unique opportunity to connect closely with the artist's life and get to know him by using artificial intelligence [8]. There are several more instances, such as the well-known player David Beckham, who addressed in nine different languages to support a campaign in anticipation of malaria [1].

Fig. 1. Example of Deepfake [3]

Fig. 2. Instances and Uses of Deepfake

This paper offers an extensive review of the existing research on deepfake creation audio) where Section 2, discusses the literature on deepfake. The database is explained in Section 3. Techniques used in deepfake creation and detection are introduced in Section 4. Further, the Conclusion is outlined in Section 5.

II. Literature Review

Deepfakes derive their name from their utilization of advanced deep learning technology to fabricate counterfeit visuals and audio content. There has been a rapid evolution of deepfake methodologies in recent years, granting malicious attackers the capability to manipulate videos across various dimensions and modalities. As an example, the authors [9] discussed the dissemination of fake media streams created by generative adversarial networks (GANs). They proposed an improved D-CNN for the detection of forged content, which demonstrated reasonable accuracy and extensive generalizability. Employing a binary-cross entropy and Adam optimizer, the model enhanced the learning frequency of the D-CNN. The proposed model by authors [10], trained on images from multiple sources, showed promising accuracy across several deepfake datasets. The proposed model is based on an Inception-based network, which is employed to get frame-level features of images to detect forged content.

The authors [11] have formulated various detection techniques for a single modality and could face several challenges in the actual world where combined modalities are forged. To overcome this challenge, this paper proposed the AVoiD- DF technique to exploit audio-visual disparity. It is initiated through integrating temporal-spatial information inside the encoder, followed by the creation of a multi-modal joint decoder to amalgamate multi-modal features. Afterward, a classifier is used to discriminate between real and forged. Since limited datasets for the detection of deepfakes are available that work on multiple modalities, in this work author created a benchmark dataset named DefakeAVMiT for multi-modal deepfake detection.

The authors [12] highlight the need to detect deepfakes in a world that relies on digital communication. In this work, the rising danger and potential fallout of malicious uses of deepfakes have been discussed. The author acknowledges existing public tools for deepfake detection but highlights the pressing need for advancements, specifically in the areas of transferability and cost-effectiveness. The study observes that many deepfake detection models are not proficient at applying techniques they have learned to deepfakes created by different methods, resulting in a decrease in accuracy. Additionally, the author emphasizes that for these models to be practical in real-life scenarios, they need to be affordable and operable on standard user devices. The paper concludes the need for further research to develop models that perform effectively across various databases, to better equip them for real-world applications and help people worldwide detect deepfakes.

In [13], a generalized detection technique was suggested to identify various types of deepfake methods. The authors focused on common traces generated by the deepfake process, such as warping artifacts, residual noise, and blur effects. The application of these traces to the proposed network resulted in superior detection performance. The study suggested the future expansion of the method to deepfake video detection and explored its robustness against signal- and time-based attacks. The authors [14] focus on costing and processing time rather than accuracy for detecting deepfakes. A facial sparse optical flow approach is generated to obtain facial features with a light CNN model and compare it with XceptionNet and R3D to distinguish among original and fake content. Comparing the three models, the CNN model is around 68 to 110 times more compact. Although the generated approach accuracy is lower, this method is effective for training time and GPU memory. In the paper [15], the authors identified the critical issue of deepfake detection in society, businesses, and politics, due to the realistic nature of deepfake videos. The problem of generalizability was discussed to create networks that are resilient against all forms of attacks. The study ended by indicating the importance of future research in exploring more deep learning models and diverse datasets for video-based deepfake detection.

In [16], The authors evaluated the efficacy of deep fake recognition for the identification of deepfakes. Two class CNNs are implemented on the two datasets Celeb-DF, Face forensics++, and corresponding AUC calculations are performed. As a result, this technique for deepfake detection is affected by face- swapping methods rather than expression-swapping methods that change only expressions, not identities. The authors [17] conducted a comprehensive investigation for the detection using multiple methods. Eight distinct CNN models are used to genuinely distinguish between original and deep fake pictures. Both datasets consisted of 70,000 images and were sourced from Kaggle and StyleGAN. The designed CNN models worked with appropriate values of dropout and padding.

The continuously evolving nature of deepfake techniques poses a significant challenge to detection algorithms. In this paper [18], an approach that employs deep reinforcement learning to get exploited images and videos was used. The proposed method uses a policy gradient approach to continuously adapt to evolving manipulation techniques. The experimental results demonstrate that the proposed deep reinforcement learning-based detection method consistently outperforms traditional deepfake detection methods, even when faced with new and unseen manipulation methods.

A lightweight 3-D CNN model consisting of two modules spatial-temporal and channel transformation is designed for better detection of deepfakes that use fewer parameters than a normal 3-D CNN model. It also overcomes the drawback of heavy storage consumption and is used to fuse spatial features in time dimension [19]. The proposed method also used a 3-D CNN model that works on adjacent frames in a video rather than a single executive frame and used it to extract spatial-temporal features from consecutively framed sequences. This method performed better as compared to other state-of-art methods [20].

III. Databases

Databases play a vital role in the training of the model used to detect deepfakes. Multiple databases are publically available for all the modalities such as image, video, and audio. Celeb-

DF is one of the databases which contains real and forged videos of celebrities. The forged videos were created using different facial manipulation techniques such as DeepFakes, Face2Face, and NeuralTextures. Another database named FaceForensics++ consists of real and forged videos used for testing the performance of deepfake detection algorithms. DFDC database includes several variations of deepfake videos, including those that have been created using face swaps, lip syncs, and other manipulations and are commonly used for deepfake detection research. ForgeryNet is the most extensive publicly accessible database for deep-face forgery in terms of data volume. The database images have been manipulated using a variety of techniques, including copy- move, splicing, and object removal, and are intended to be used for image forgery detection. Table I shows a brief description of all the databases that have been analyzed during this study.

IV. Techniques Used In Deep Fake Creation And Detection

Advancements in deep learning have significantly enhanced the development of technology for facial modification. Some deep learning techniques can be used for the creation and detection of images and videos [21].

A. ***Convolutional Neural Network***: CNNs are deep learning classification neural networks that consist of multiple layers, each contributing unique features to the network. The input layer is the first layer of the model which receives the input image. The convolutional layer is the second layer that applies filters to the input image to extract feature maps that capture relevant patterns. Then activation functions introduce non-linearity into the model by applying a transformation to the output of each neuron in the convolutional layers. Pooling Layer reduces the dimensionality of the feature maps by down-sampling them. The fully Connected and the final layers process the output of the above layers to make a final prediction [22] [23].

B. ***Recurrent Neural Network***: Another use of an artificial neural network that can learn characteristics from sequence data is the recurrent neural network (RNN). RNNs are constructed similarly to neural networks using several invisible layers, each of which has a bias and a weight. The direct cycle graph's node relationships execute sequentially in RNN. RRN has the benefit of enabling the discovery of temporal dynamic behavior. RNN stores in- formation sequences from previous inputs in internal memory, making it useful in many fields like speech recognition and natural language processing [22].

C. ***Auto Encoders***: Autoencoder is one of the widely used deep learning methods. Autoencoder insight is a pair of shared-weight encoder and decoder functions that train together as a function. Typically, we discuss autoencoders in the context of dimensionality reduction and the training of generative models [24]. Nevertheless, autoencoders can be employed to improve upon the requirements for image compression by taking compressed representations of images. The combination of the decoder part of the second image and the encoder part of first image is used for face swapping. This method is employed by open-source repositories like DeepFakeLab, DFaker, and FaceSwap [25].

D. ***Generative Adversarial Networks (GAN)***: GAN has proved to be a most formidable challenging computational deep learning algorithm to utilize and train. There are two neural networks: a discriminator (judge, classifier, etc.) and a generator. With the discriminator net, better results can be achieved as it rejects certain poor examples even if the autoencoder net looks the same as the generator net. Hence, GAN's deep-fake production method presupposes that the generator's goal is to outwit the discriminator, which is another machine; as a result, deepfakes are created that closely resemble real videos, making it harder for humans to recognize them visually. This method is employed by some open-source projects, such as Faceswap-GAN [25].

The work done by the researchers for the detection of deepfakes is shown in Table II and the creation or generation of deepfakes is shown in Table III. Table II states that both machine learning and deep learning algorithms are used for the detection of the Deepfake [42]. The accuracy achieved by the deep learning algorithm that is Hybrid CNN is 95.75% and by the machine learning algorithm i.e., SVM is 88.33% in the current year 2023. So, Deep Learning algorithms are more accurately detecting the deepfake than machine learning algorithms. The modalities used by the researchers are image, video, and audio for the generation as well as detection of deepfakes. The majority of the researchers work on a single modality as per Table II and Table III. Either they work on image or video modality, not on both. The main deep learning algorithms used for the detection of deepfakes are CNN, LSTM, and Transfer Learning techniques.

The algorithms used for the creation of deepfakes are autoencoders and GAN. The maximum SSIM (Structural Similarity Index) is calculated in the year 2019 by using the STGAN technique. The value of SSIM is between 0 and 1. The value 1 means the generated image is similar to the original image. The more is the value of SSIM, the more is the similarity of the generated image w.r.t the original Image. The second metric used to assess the quality of the generated image is FID (Frechet inception distance). Lower the value of FID the higher the quality of the generated image and vice-versa. The value of FID shown in Table III is in a percentage format. The minimum value of FID is in the year 2022 by using the Lightweight Identity-aware Dynamic Network i.e., 0.067. From the literature studied maximum of the work is done for the detection of deepfakes. The commonly used datasets for the detection of deepfakes are DeepfakeTIMIT, FaceForensics++, Celeb-DF, and DFDC.

TABLE I. DEEPFAKE DATABASES

S.No.	Dataset Name	Image / Video	Ratio (Real: Fake)	Total Videos/ Images	Publically Available	Links
1	UADFV	Video	1:1	98	No	-
2	DeepFake-TIMIT	Video	Only Fake	620	Yes	https://www.idiap.ch/en/dataset/deepfaketimit
3	Celeb-DF	Video	1:9.5	6229	Yes	(on request) https://github.com/yuezunli/celeb-deepfakeforensics
4	FaceForensics++	Video	1:5	6000	No	-
5	DFDC	Video	1:5	128064	Yes	https://ai.facebook.com/datasets/dfdc/
6	ForgeryNet	Image/Video	1:1	2.9 million / 221247	Yes	https://yinanhe.github.io/projects/forgerynet.html#download
7	DFFD	Image	1:5	299039	Yes	http://cvlab.cse.msu.edu/category/downloads.html
8	FFHQ	Image	-	70000	Yes	https://github.com/NVlabs/ffhq-dataset
9	iFakeFaceDB	Image	-	87000	Yes	https://github.com/socialabubi/iFakeFaceDB
10	VGGFace	Image	-	3.31 million	Yes	https://github.com/ox-vgg/vgg_face2
11	DeepfakeAVMiT	Audio/Video	1:11	6000	No	-
12	FakeAVCeleb	Audio/Video	1:19	20,000	yes	FakeAVCeleb (google.com)

TABLE II. TECHNIQUES AND DATASETS USED BY VARIOUS RESEARCHERS FOR THE DETECTION

Techniques	Reference / Year	Modality		Method	Evaluation Dataset				Performance	
		Image	Video	Model	Face Forensics++	Celeb-DF	DFDC	Other / Custom	ACC	AUC
Classic ML	[26] 2023	✔		SVM	-	-	-	-	88.33	-
	[27] 2021	-	✔	SVM			-	-	-	FF++-56.8
	[28] 2021	✔	-	Canny Edge Detection, Hough Transform	✔	✔	-	-	-	94
	[29] 2020	✔	-	K-NN	-	-	-	✔	90.22	-
	[30] 2019	✔	✔	SVM	-	-	-	✔	-	89
	[31] 2023	-	✔	Hybrid CNN	-	-	✔	-	95.75	-
	[32] 2022	-	✔	CNN- LSTM	✔	✔	✔	-	FF++-91.21 Celab-DF-79.49 DFDC-66.26	-
	[33] 2021	-	✔	Convolutional LSTM-based Residual Network (CLRNet)	✔	-	-	✔	93.86	-
	[34] 2020	✔		Xception Networks	✔	✔	✔	-	-	75
	[35] 2020	-	✔	LSTM	✔	-	-	-	94.29	-
	[36] 2020	-	✔	CNN	✔	-	-	-	-	99
Deep Learning	[37] 2020	✔	-	ResNet, Gram-Net	-	-	-	-	ResNet-80.55	-
	[38] 2020	✔	✔	FakeSpotter	-	✔	-	-	-	66.8
	[39] 2019	✔	-	DCGAN	-	-	-	✔	95.45	-
	[40] 2019	-	✔	VGG16, ResNet50	✔	-	-	-	VGG16-81.61 ResNet50-75.46	-
	[41] 2019	✔	-	GAN	-	-	-	✔	-	99

TABLE III. TECHNIQUES AND DATASETS USED BY VARIOUS RESEARCHERS FOR THE CREATION

Reference / Year	Modality		Method	Database			Performance	
	Image	Video	Model	CelebA-HQ	CelebA	Other	FID	SSIM
[43] 2023	✔	✔	Deep Convolution GAN	-	✔	-	49.3	-
[44] 2022	✔	-	Lightweight Identity-aware Dynamic Network	-	-	✔	6.79	-
[45] 2020	✔	-	GAN	-	-	✔	29.31	-
[46] 2020	✔	-	StarGAN v2	✔	-	-	23.9	-
[47] 2020	✔	-	Encoder-Decoder	-	-	✔	12.17	0.1717
[48] 2019	✔	-	Feed Forward Network	-	-	✔	29.5	0.74
[49] 2019	-	✔	FSGAN	-	-	✔	-	0.51
[50] 2019	✔	-	STGAN	-	-	✔	-	0.948

V. Conclusion

Deepfake technology is a new way to alter digital content and create videos that look very real. However, this technology can be misused for spreading false information or deceiving people, making it crucial to develop an effective method for the detection of deepfakes. The responsible use of deepfake technology is essential, as its inappropriate application can lead to significant consequences, from harming individuals' reputations to influencing public opinion. In this paper, various aspects i.e., datasets, techniques, and models used in deepfakes have been discussed. The study concludes that the existing models used for the detection of deepfakes face several challenges. There is no single model that can work well with different types of datasets. Also, the methods used to create deepfakes are changing quickly, making it even more difficult for existing detection models to keep up. In the future, the author will aim to develop a highly accurate robust detection model and also propose a way to overcome the challenges caused by deepfakes.

References

[1] L. A. Passos, D. Jodas, K. A. da Costa, L. A. S. Júnior, D. Colombo, and J. P. Papa, "A review of deep learning-based approaches for deepfake content detection," *arXiv preprint arXiv:2202.06095*, 2022.

[2] P. Korshunov and S. Marcel, "Deepfakes: a new threat to face recognition? assessment and detection," *arXiv preprint arXiv:1812.08685*, 2018.

[3] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "Faceforensics++: Learning to detect manipulated facial images," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1–11.

[4] C. Öhman, "Introducing the pervert's dilemma: a contribution to the critique of deepfake pornography," *Ethics and Information Technology*, vol. 22, no. 2, pp. 133–140, 2020.

[5] (2021) How deepfakes are a problem for us all and why the law needs to change. https://informationmatters.net/ deepfakes-problem-why-law-needs-to-change/. [Online; accessed 19-July-2023].

[6] H. F. Shahzad, F. Rustam, E. S. Flores, J. Luís Vidal Mazón, I. de la Torre Diez, and I. Ashraf, "A review of image processing techniques for deepfakes," *Sensors*, vol. 22, no. 12, p. 4556, 2022.

[7] A. Ray, "Disinformation, deepfakes and democracies: The need for legislative reform," *TheUNIVERSITY OF NEW SOUTH WALES LAW JOURNAL*, vol. 44, no. 3, pp. 983–1013, 2021.

[8] M. Mihailova, "To dally with dal´ı: Deepfake (inter) faces in the art museum," *Convergence*, vol. 27, no. 4, pp. 882–898, 2021.

[9] Y. Patel, S. Tanwar, P. Bhattacharya, R. Gupta, T. Alsuwian, I. E. Davidson, and T. F. Mazibuko, "An improved dense cnn architecture for deepfake image detection," *IEEE Access*, vol. 11, pp. 22 081–22 095, 2023.

[10] S. Mathews, S. Trivedi, A. House, S. Povolny, and C. Fralick, "An explainable deepfake detection framework on a novel unconstrained dataset," *Complex & Intelligent Systems*, pp. 1–13, 2023.

[11] W. Yang, X. Zhou, Z. Chen, B. Guo, Z. Ba, Z. Xia, X. Cao, and K. Ren, "Avoid-df: Audio-visual joint learning for detecting deepfake," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 2015–2029, 2023.

[12] M. Masood, M. Nawaz, K. M. Malik, A. Javed, A. Irtaza, and H. Malik, "Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward," *Applied intelligence*, vol. 53, no. 4, pp. 3974–4026, 2023.

[13] J. Kang, S.-K. Ji, S. Lee, D. Jang, and J.-U. Hou, "Detection enhancement for various deepfake types based on residual noise and manipulation traces," *IEEE Access*, vol. 10, pp. 69 031–69 040, 2022.

[14] S. Fang, S. Wang, and R. Ye, "Deepfake video detection through facial sparse optical flow based light cnn," in *Journal of Physics: Conference Series*, vol. 2224, no. 1. IOP Publishing, 2022, p. 012014.

[15] S. Suratkar and F. Kazi, "Deep fake video detection using transfer learning approach," *Arabian Journal for Science and Engineering*, vol. 48, no. 8, pp. 9727–9737, 2023.

[16] S. Ramachandran, A. V. Nadimpalli, and A. Rattani, "An experimental evaluation on deepfake detection using deep face recognition," in *2021 International Carnahan Conference on Security Technology (ICCST)*. IEEE, 2021, pp. 1–6.

[17] H. S. Shad, M. M. Rizvee, N. T. Roza, S. Hoq, M. Monirujjaman Khan, A. Singh, A. Zaguia, S. Bourouis *et al.*, "Comparative analysis of deepfake image detection method using convolutional neural network," *Computational Intelligence and Neuroscience*, vol. 2021, 2021.

[18] Y. Lu, Y. Liu, J. Fei, and Z. Xia, "Channel-wise spatiotemporal aggregation technology for face video forensics," *Security and Communication Networks*, vol. 2021, pp. 1–13, 2021.

[19] J. Liu, K. Zhu, W. Lu, X. Luo, and X. Zhao, "A lightweight 3d convolutional neural network for deepfake detection," *International Journal of Intelligent Systems*, vol. 36, no. 9, pp. 4990–5004, 2021.

[20] X. H. Nguyen, T. S. Tran, K. D. Nguyen, D.-T. Truong *et al.*, "Learning spatio-temporal features to detect manipulated facial videos created by the deepfake techniques," *Forensic Science International: Digital Investigation*, vol. 36, p. 301108, 2021.

[21] C. Kaushal, S. Bhat, D. Koundal, and A. Singla, "Recent trends in computer assisted diagnosis (cad) system for breast cancer diagnosis using histopathological images," *Irbm*, vol. 40, no. 4, pp. 211–227, 2019.

[22] A. Malik, M. Kuribayashi, S. M. Abdullahi, and A. N. Khan, "Deepfake detection for human face images and videos: A survey," *Ieee Access*, vol. 10, pp. 18 757–18 775, 2022.

[23] I. Kansal, R. Popli, and C. Singla, "Comparative analysis of various machine and deep learning models for face mask detection using digital images," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2021, pp. 1–5.

[24] P. K. Shukla, J. K. Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary,and P. K. Shukla, "Multiobjective genetic algorithm and convolutional neural network based covid-19 identification in chest x-ray images," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–9, 2021.

[25] A. A. Maksutov, V. O. Morozov, A. A. Lavrenov, and A. S. Smirnov, "Methods of deepfake detection based on machine learning," in *2020 IEEE conference of russian young researchers in electrical and electronic engineering (EIConRus)*. IEEE, 2020, pp. 408–411.

[26] L. Pryor, R. Dave, M. Vanamala *et al.*, "Deepfake detection analyzing hybrid dataset utilizing cnn and svm," *arXiv preprint arXiv:2302.10280*, 2023.

[27] S. Fung, X. Lu, C. Zhang, and C.-T. Li, "Deepfakeucl: Deepfake detection via unsupervised contrastive learning," in *2021 international joint conference on neural networks (IJCNN)*. IEEE, 2021, pp. 1–8.

[28] S. Hu, Y. Li, and S. Lyu, "Exposing gan-generated faces using inconsistent corneal specular highlights," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 2500–2504.

[29] L. Guarnera, O. Giudice, and S. Battiato, "Deepfake detection by analyzing convolutional traces," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 666– 667.

[30] X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes using inconsistent head poses," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 8261–8265.

[31] S. T. Ikram, S. Chambial, D. Sood *et al.*, "A performance enhancement of deepfake video detection through the use of a hybrid cnn deep learning model," *International journal of electrical and computer engineeringsystems*, vol. 14, no. 2, pp. 169–178, 2023.

[32] P. Saikia, D. Dholaria, P. Yadav, V. Patel, and M. Roy, "A hybrid cnn-lstm model for video deepfake detection by leveraging optical flow features," in *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–7.

[33] S. Tariq, S. Lee, and S. Woo, "One detector to rule them all: Towardsa general deepfake attack detection framework," in *Proceedings of the web conference 2021*, 2021, pp. 3625–3637.

[34] Y. Nirkin, L. Wolf, Y. Keller, and T. Hassner, "Deepfake detection based on discrepancies between faces and their context," *IEEE Transactions*

on Pattern Analysis and Machine Intelligence, vol. 44, no. 10, pp. 6111–6121, 2021.

[35] I. Amerini and R. Caldelli, "Exploiting prediction error inconsistencies through lstm-based classifiers to detect deepfake videos," in *Proceedings of the 2020 ACM workshop on information hiding and multimedia security*, 2020, pp. 97–102.

[36] S. Agarwal, H. Farid, T. El-Gaaly, and S.-N. Lim, "Detecting deep-fake videos from appearance and behavior," in *2020 IEEE international workshop on information forensics and security (WIFS)*. IEEE, 2020, pp. 1–6.

[37] Z. Liu, X. Qi, and P. H. Torr, "Global texture enhancement for fake face detection in the wild," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8060–8069.

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

[39] X. Xuan, B. Peng, W. Wang, and J. Dong, "On the generalization of gan image forensics," in *Chinese conference on biometric recognition*. Springer, 2019, pp. 134–141.

[40] I. Amerini, L. Galteri, R. Caldelli, and A. Del Bimbo, "Deepfake video detection through optical flow based cnn," in *Proceedings of the IEEE/CVF international conference on computer vision workshops*, 2019, pp. 0–0.

[41] F. Marra, D. Gragnaniello, L. Verdoliva, and G. Poggi, "Do gans leave artificial fingerprints?" in *2019 IEEE conference on multimedia information processing and retrieval (MIPR)*. IEEE, 2019, pp. 506–511.

[42] C. Singla, R. Kaur, J. Singh, N. Nisha, A. K. Singh, and T. Singh, "Fine- tuned pre-trained deep neural network for automatic detection of diabetic retinopathy using fundus images," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 9s, pp. 735–742, 2023.

[43] M. Kumar, H. K. Sharma *et al.*, "A gan-based model of deepfake detection in social media," *Procedia Computer Science*, vol. 218, pp. 2153–2162, 2023.

[44] Z. Xu, Z. Hong, C. Ding, Z. Zhu, J. Han, J. Liu, and E. Ding, "Mobilefaceswap: A lightweight framework for video face swapping," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 3, 2022, pp. 2973–2981.

[45] S. Singh, R. Sharma, and A. F. Smeaton, "Using gans to synthesise minimum training data for deepfake generation," *arXiv preprint arXiv:2011.05421*, 2020.

[46] Y. Choi, Y. Uh, J. Yoo, and J.-W. Ha, "Stargan v2: Diverse image synthesis for multiple domains," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8188–8197.

[47] J. Zhang, X. Zeng, M. Wang, Y. Pan, L. Liu, Y. Liu, Y. Ding, and C. Fan, "Freenet: Multi-identity face reenactment," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 5326–5335.

[48] E. Zakharov, A. Shysheya, E. Burkov, and V. Lempitsky, "Few-shot adversarial learning of realistic neural talking head models," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp.9459–9468.

[49] Y. Nirkin, Y. Keller, and T. Hassner, "Fsgan: Subject agnostic face swapping and reenactment," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 7184–7193.

[50] M. Liu, Y. Ding, M. Xia, X. Liu, E. Ding, W. Zuo, and S. Wen, "Stgan: A unified selective transfer network for arbitrary image attribute editing," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 3673–3682.