International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Extend Review on Deep Fake Creation and Detection Models

Mrs. Banu Priya[1], Jhosiah Felips Daniel[2]

[1]Assistant Professor, Department of Computer Science and Engineering, [2]Co-author, ME CSE, 1st Year - Sri Shakthi Institute of Engineering and Technology

ABSTRACT

Deep faking involves animating an image using a video. Current methods involve animating an image with respect to different videos. In recent days, self-supervised formulation approaches are designated to adapt for the training purpose. Mapping the motion of the item between video and image can be achieved with the help of generator network models. The result is a generation model that can generate any video of the specified classification with the desired audio. This paper explains the challenges, research trends and directions related to deep fake technology. This survey provides an overview of the deep fake creations and detections, the potential trends, and future directions by analyzing existing survey papers related to deep faking. The various deep faking generation and detection models are documented with key features and the datasets used. Such extended reviews provide an insight to the researchers with regards to the deep fake creation and detection models

Keywords: Image animation, Deep fake detection, Deep learning, Motion model.

INTRODUCTION

Image animation is a method which is manipulated to display the sequence of images in a 2D or 3D model or to generate a motion picture. The approach chosen relies on the deep faking of videos. Deep faking involves replacing someone or a thing in an existing video with another version of that someone or a thing. Basically the video is manipulated according to one's own needs. Deep faking is a modification of data or media using artificial intelligence algorithms and tools such as a generator and discriminator. Generator generates the multimedia content while the discriminator determines whether the content is real or not. The person in an existing video can be replaced with some other person. Deep faking manipulates visual and audio content with a high potential to deceive people that something is false is true. It is one of the most powerful techniques in artificial intelligence and machine learning - it involves deep fakes based on deep learning, generative neural network architecture such as auto encoders and generative adversarial networks (GANs). The facial morphing concept is implemented using deep faking technology. The tools used for deep faking technology are CrazyTalk, DeepFaceLab, dfaker, Dynamixyz, eSpeak, Face2Face, Face Crop Jet, Face Crop, Face Swap, Face Swap Live, Face Swap Online, Facerig, Faceswap, Lyrebird, Microsoft Garage Face Swap, Natural Front, Reflect, Voicer and Lab DJR Font. This paper discusses the various methods used for deep faking creation and detection.

LITERATURE REVIEW

They proposed the GPT-2 model [1], - this allows the tweet mapping from one user to another using OpenAI and transfer learning. [2] Suggested the survey of algorithms to create a deep fake, overview the deep fake and detect the deep fake method. Guidelines for the latest deep fake researchers were also provided. [3] Developed a model using the Generative Adversarial Network (GAN) for remote sensing of images through focusing on frequencies [4] Explained about the categories of deep fake and the technologies behind the development of deep fake techniques. Suggestions from the technology and regulatory point of view were given. [5] They used the Corridor's deep fake Keanu video technique to illuminate and analyze a range of manipulated videos. This was done with the help of a VFX platform. An Introduction to a hybrid deep learning approach was also given. [6] Depicts the use of a deep fake detection challenge dataset to differentiate between the original and fake videos. Detection of deep faking videos [7] using the Generative Adversarial Network (GAN) and usage of MesoNet to train GAN. Development of a model [8] to create and identify the flawless deep faking video which automatically detects the video over a large database. [9] Proposed a deep faking detection model based on NR-IQA technique to evaluate the video and image quality in social media. Proposed a SSTNet model [10] to detect the deep fake model. They discussed the field and development in relation to the deep faking techniques to deal with the increasing threats. Proposed a temporal-aware pipeline [11] to detect deep fake video which extracts frame level features by using the simple Convolutional Neural Network (CNN).

Analyzed a deep fake detection model [12] with the aid of an Expectation Maximization algorithm to extract local features and distinguish the different architecture and generation processes via the CELEBA dataset. Proposed an algorithm [13] which detects deep fake video manipulation by using photographic response from a non uniformity analysis. Designed a model [14] to generate deep faking videos by using a tuned parameter set based on the VGG and Facenet neural networks with the help of VidTIMIT database for face swapping technology development. Proposed a deep faking detection model [15] based on CNN - it obtains the FaceForensics++ dataset to compare real and fake videos. [16] Proposed a framework using the Ethereum smart contracts which traces and tracks the provenance of digital content to its original source and stores the digital content and its metadata. [17] Discussed the challenges and highlights of research opportunities which arises during the generation of high quality images and videos. [18] Explain the current and future capabilities of deep faking technology for identification and exposing of fake videos.

Proposed a system [19] which focuses on online frequency masking augmentation and large margin cosine loss function techniques to detect audio of the deep faking videos. Introduced the ID-Reveal [20] that learns to recognize facial features - such as how the person moves while talking - with the aid of the metric learning approach combined with versarial training strategy to detect the deep fake video detection. Developed the case study [21] explaining about the five attacks under the ROC curve. Trained generators were used to reduce the AUC classifiers by developing the black box attack. [22] Proposed a multi attentional deep fake detection network by using the multiple spatial attention, textural feature enhancement and aggregate low level textural features and high level semantic features with the aid of attention maps.

Developed a deep fake detection model [23] using an adversarial machine learning community. [24] Developed a deep fake detection method to perform the adversarial attack in a black box setting - designed for deep fake detectors to improve the transferability of adversarial examples. [25] Developed a deep fake detection model that is based on discrepancies between faces and their context which involves the face identification network, context recognition network in order to detect the fake image.

Proposed a deep learning network using the Siamese network architecture [26] and triplet loss. Comparison was done with the SOTA deep fake detection methods using the DF-TIMIT datasets for deep fake detection. Formulation of the deep fake detection methods [27] as a hypothesis - to testing the problem in classifying an image and bounding the error probability using the Euclidean approximation method. A relationship was established between the error probability and epidemic thresholds. [28] Developed a joint deep faking detection system which detects both audio and visual modules. Proposed deep fake detectors in both black-box and white-box settings using fast Gradient Sign Method [29] and Carlini and Wagner norm attack - used in the improvement of the Lipschitz regularization. [30] Designed a model to detect a deep fake text posted on social media platforms like twitter. Proposed a method based on the expectation-maximization algorithm [31] for the extraction of deep fake fingerprints from the given input images. [32] Developed an open source platform for deep fake detection known as DeepFake-o-meter which builds an interface between the users. [33] Developed a deep learning based on the image compression algorithm using the convolutional and deep learning based compression algorithms for considering both image fidelity errors and raw reconstruction errors to make improvement in PSNR. Presented an [34] energy compaction-based image compression architecture using a convolutional auto encoder (CAE) - This helps achieve higher coding efficiency. [35] Introduced a regularization scheme to compare a simple dimensionality reduction bottleneck (a Gaussian Variational Autoencoder (VAE)) and a discrete Vector Quantized VAE (VQ-VAE) for analyzing the quality of learned representations and to accurately reconstruct individual spectrogram frames using the VQ-VAE. The result is then used to measure the mapping to phonemes. [36] Proposed a framework to estimate the generative and discriminative models using an adversarial process trained with back propagation techniques through qualitative and quantitative samples.

DEEP FAKE CREATION

Deep fakes are becoming popular due to the quality of altered videos and also the user-friendly ability of the applications for a wide range of individuals with various computer skills (in the professional to novice range). Deep faking applications are developed based on deep learning techniques for representing its complex and high-dimensional data. Deep networks implemented using deep autoencoders are widely used for dimensionality reduction and image compression [33]–[35]. The first attempt of a deep fake creation was the development of FakeApp. This application was developed by a Reddit user using an auto encoder and a decoder pairing structure [37], [38]. In this method, the auto encoder extracts latent features of facial images while the decoder was used to reconstruct the facial images. In order to swap the faces between source images and target images, there is a need to use two encoder-decoder pairs. Each pair is used for training an image set, and the encoder's arguments are shared between two network pairs. In other words, both pairs have the same encoder network. This strategy enables the common encoder to find and learn the similarity between the two sets of facial images - this is relatively unchallenging because faces normally have such as eyes, nose, mouth positions which are common to all individuals. This approach is applied in several works such as DeepFaceLab [39], DFaker [40], DeepFake tf (Tensorflow based deep fakes) [41]

By adding the adversarial losses and perceptual losses implemented in VGGFace [42] to the encoder-decoder architecture, an improved version of deep fakes based on the generative adversarial network is created. (GAN) [36], i.e. faceswap-GAN, was proposed in [43]. The VGGFace perceptual loss is used to generate realistic and consistent eye movements with input faces- thus leading to higher quality output videos. This model facilitates the generation of outputs with 64x64, 128x128, and 256x256 resolutions. In addition, the multi-task convolutional neural network (CNN) from the FaceNet implementation [44] is introduced to make stable face detection and reliable face alignment. The CycleGAN [45] is utilized for generative network implementation. Popular deep fake tools and their typical features are summarized in Table II. Table I shows the number of published papers related to deep fakes in the year 2017 to 2021 and it was obtained from https://app.dimensions.ai on 1 November 2021 with the search keyword "deepfake" applied to full text of scholarly papers. Number of such papers in 2018, 2019 and 2020 are 67, 369 and 1337 respectively. From the beginning of 2021 to the end of 2021, there are 1210 papers about deep fake technology. Table II shows the summary of the notable deep fake tools. Figure I shows the graphical representation of the number of papers published in the years 2017 to 2021.

Table I: Number of published papers related to deep fakes in year 2017 to 2021

Year	Number of published papers in the year
2017	1
2018	67
2019	369
2020	1337
2021	1210

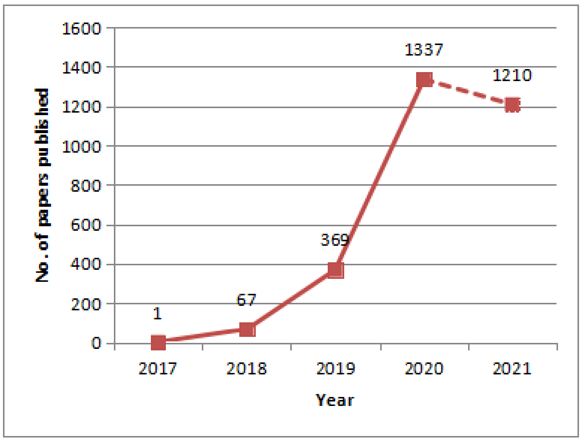

Figure I: Graphical representation of number of papers published in the year 2017 to 2021.

Tools	Key Features
FaceShifter [59]	- Face swapping in high-fidelity through the exploitation and integration of the target attributes. - Can be applied to new face pairs without subject specific training requirements [46].
DFaker [60]	- DSSIM function [47] is used to reconstruct the face. - Implemented based on the Keras library.
Few-Shot Face Translation [61]	- Use of a pre-trained face recognition model to extract the latent embeddings for GAN processing. - Incorporate the semantic priors obtained by modules from FUNIT [48] and SPADE [49].
DeepFake tf [62]	Similar to DFaker approach but implemented based on tensorflow.
DeepFaceLab [63]	- Update to the Face Swap method with new models, e.g. H64, H128, LIAEF128, SAE [50]. - Supports multiple face extraction modes, e.g. S3FD, MTCNN, dlib, or manual [50].
MarioNETte [64]	- A face reenactment framework that preserves the identity of the target. - No additional fine-tuning phase is required for identity adaptation [51].
Neural Voice Puppetry [65]	- A method for audio-driven facial video synthesis.
“Do as I Do” Motion Transfer [66]	- Automatically transfers the motion from a source to a target individual by learning video-to-video translation concept. - Creation of a motion-synchronized dancing video with multiple subjects [53].
Transformable Bottleneck Networks [67]	- A method for fine-grained 3D manipulation of image content. - Apply spatial transformations in CNN models using a transformable bottleneck framework [54].
Faceswap [68]	- Use of two encoder-decoder pairs. - Parameters of the encoder are shared.
DiscoFaceGAN [70]	- Generate facial images of virtual people with independent latent variables of identity, expression, pose, and illumination. - Embed 3D priors into adversarial learning [56].
AvatarMe [72]	- Reconstructs 3D faces from “in-the-wild” images. - Reconstructs authentic 4K by 6K-resolution 3D faces from a single low-resolution image [57].
FSGAN [73]	- A face swapping and reenactment model that can be applied to face pairs (no training required) - Adapts to both pose and expression variations [58].

DISCUSSIONS AND FUTURE RESEARCH

Directions: - Deep fakes have begun to corrode the trust people have in media contents since seeing them is no longer equal to believing in them .Sometimes there is no need to spread deep fakes to people and cause detrimental effects. People who create deep fakes with malicious purposes need to deliver them to target audiences as part of their interruption strategy. For example, this approach can be utilized by intelligence services trying to manipulate the decisions made by higher authorities such as politicians, - thus leading to national and international security threats [74]. Deepfakes' quality has been increasing - hence the performance of detection methods needs to be improved accordingly. Another suggestion is to integrate detection methods into distribution platforms such as social media to increase its effectiveness in. Screening and filtering mechanisms using effective detection methods can be implemented on such platforms to ease/simplify the deep fakes detection [74]. Watermarking tools are integrated into devices that people use everyday in order to make digital contents to create immutable data - - this helps in storage of original details such as time and location of multimedia contents [74]. This integration is difficult to implement but can be solved using disruptive blockchain technology. The blockchain has been used effectively in many areas and there are a few studies addressing the deep fake detection problems related to this technology. The development of machine learning and AI technologies could have been used in the modification of these digital contents. Presently images and videos are used as courtroom evidence due to open access to a wide range of digital manipulation methods [76]. There are some existing papers for deep faking detection while there are a lot of other existing papers related to the creation of deep faking methods. Table III shows the summary of prominent deep face detection methods.

Table III: Summary of prominent deep fake detection methods

Methods	Classifiers/ Techniques	Key Features	Dealing with	Datasets Used
Spatio-temporal features with LSTM [80]	Convolutional bidirectional recurrent LSTM network	- An XceptionNet CNN is used for facial feature extraction while audio embeddings are obtained by stacking multiple convolution modules. - Two loss functions, i.e. cross-entropy and Kullback-Leibler divergence, are used.	Videos	FaceForensics++ [81] and Celeb-DF (5,639 deep- fake videos) [82] datasets and the ASVSpoof 2019 Logical Access audio dataset [83].
Eye,teach and facial texture [85]	Logistic regression and neural network	- Exploitation of facial textural differences, missing reflections and details in eye and teeth areas of deep fakes. - Logistic regression and neural network are used for classifying	Videos	A video dataset downloaded from YouTube.
Intra-frame and temporal inconsistencies [88]	CNN and LSTM	- CNN is employed to extract the frame-level features, which are distributed to LSTM to construct sequences descriptor useful for classifications	Videos	A collection of over 600 videos obtained from multiple websites.
Face X-ray [89]	CNN	- Tries to locate the blending boundary between the target and original faces - Can be trained without use of fake images.	Images	FaceForensics++ [81], Deepfake Detection (DFD) [90], DFDC [91] and Celeb-DF [82].
FakeCatcher [92]	CNN	- Extracts biological signals in portrait videos and uses them as a descriptor of authenticity.	Videos	UADFV [93], FaceForensics [94], FaceForensics++ [81], Celeb-DF [82], and a new dataset of 142 videos, independent of the generative model, resolution, compression, content, and context.
Bag of words and shallow classifiers [95]	SVM, RF, MLP	- Extract discriminant features using the bag of words method. Feeds data to SVM	Images	The well-known LFW face database [96], containing 13,223 images with resolution of 250x250.
Pairwise learning [97]	CNN concate- nated to CFFN	- Two-phase procedure: feature extraction using CFFN based on the Siamese network architecture [78] and classification using CNN.	Images	- Face images: real ones from CelebA [99], and fake ones generated by DCGAN [100], WGAN [101], WGAN-GP [102], least squares GAN [85], and PGGAN [103].
Using attribution-based confidence	ResNet50 model [104], pre-trained on VGGFace2 [105]	- The ABC metric [106] is used for the detection of deep fake videos without accessing training data. - ABC values obtained for original videos are greater than 0.94 while those of deep fakes have low ABC values.	Videos	VidTIMIT and two other original datasets obtained from the COHFACE (https://www.idiap.ch/dataset/cohface)and from YouTube. datasets from COHFACE [107] and YouTube are used to generate two deepfake datasets by commercial website https://deepfakesweb.com and another deepfake dataset is DeepfakeTIMIT [108].

CONCLUSION

Not all deep fakes are of threat. However, since the deep faking technology makes it effortless to create realistic media, some users are misusing it to perform malicious attacks. These attacks are targeting individuals - causing psychological, monetary, and physical problems. In this survey, the focus was on creation and detection of deep fake images and videos. An in depth review in relation to how deep fake technologies were created, and what is being done to detect these deep fakes were given in this paper. We believe this information will be helpful to understand and prevent threatening deep fakes. Our future research focuses on one Method - The Monkey-Net method to deep fake in terms of animating a source image S with respect to the motions from a driving video D.

REFERENCES

1. Ressmeyer, R., Masling, S. and Liao, M., 2019. “Deep faking” political twitter using transfer learning and GPT-2.
2. Nguyen, T.T., Nguyen, C.M., Nguyen, D.T., Nguyen, D.T. and Nahavandi, S., 2019. Deep learning for deepfakes creation and detection: A survey. arXiv preprint arXiv:1909.11573.
3. Ren, C., Moore, J., Ziemann, A. and Theiler, J., 2021, April. Deep-faking it: generating and detecting synthetic remote sensing imagery. In Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVII (Vol. 11727, p. 117270M). International Society for Optics and Photonics.
4. Meskys, E., Kalpokiene, J., Jurcys, P. and Liaudanskas, A., 2019. Regulating deep fakes: legal and ethical considerations. Available at SSRN 3497144.
5. Bode, L., 2021. Deepfaking Keanu: YouTube deep fakes, platform visual effects, and the complexity of reception. Convergence, 27(4), pp.919-934.
6. Lewis, J.K., Toubal, I.E., Chen, H., Sandesera, V., Lomnitz, M., Hampel-Arias, Z., Prasad, C. and Palaniappan, K., 2020, October. Deepfake Video Detection Based on Spatial, Spectral, and Temporal Inconsistencies Using Multimodal Deep Learning. In 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (pp. 1-9). IEEE.
7. Aduwala, S.A., Arigala, M., Desai, S., Quan, H.J. and Eirinaki, M., 2021, August. Deepfake Detection using GAN Discriminators. In 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 69-77). IEEE.
8. Patidar, M., Nair, R.R. and Kanchana, M., 2021. Deepfake Video Generation & Detection: Analysis (pp. 659-666).
9. Yang, W.C. and Tsai, J.C., 2020. Deepfake Detection Based on No-Reference Image Quality Assessment (NR-IQA). Forensic Science Journal, 19(1), pp.29-38.
10. Katarya, R. and Lal, A., 2020, October. 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) (pp. 485-490). IEEE.
11. Güera, D. and Delp, E.J., 2018, November. Deepfake video detection using recurrent neural networks. In 2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS) (pp. 1-6). IEEE.
12. Guarnera, L., Giudice, O. and Battiato, S., 2020. Deepfake detection by analyzing convolutional traces. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 666-667).
13. Koopman, M., Rodriguez, A.M. and Geradts, Z., 2018, August. Detection of deepfake video manipulation. In The 20th Irish machine vision and image processing conference (IMVIP) (pp. 133-136).
14. Korshunov, P. and Marcel, S., 2019, June. Vulnerability assessment and detection of deepfake videos. In 2019 International Conference on Biometrics (ICB) (pp. 1-6). IEEE.
15. Amerini, I., Galteri, L., Caldelli, R. and Del Bimbo, A., 2019. Deepfake video detection through optical flow based cnn. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (pp. 0-0).
16. Hasan, H.R. and Salah, K., 2019. Combating deepfake videos using blockchain and smart contracts. Ieee Access, 7, pp.41596-41606.
17. Lyu, S., 2020, July. Deepfake detection: Current challenges and next steps. In 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (pp. 1-6). IEEE.
18. Maras, M.H. and Alexandrou, A., 2019. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. The International Journal of Evidence & Proof, 23(3), pp.255-262.
19. Chen, T., Kumar, A., Nagarsheth, P., Sivaraman, G. and Khoury, E., 2020, November. Generalization of audio deepfake detection. In Proc. Odyssey 2020 The Speaker and Language Recognition Workshop (pp. 132-137).
20. Cozzolino, D., Rossler, A., Thies, J., Nießner, M. and Verdoliva, L., 2021. Id-reveal: Identity-aware deepfake video detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 15108-15117).
21. Carlini, N. and Farid, H., 2020. Evading deepfake-image detectors with white-and black-box attacks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 658-659).
22. Zhao, H., Zhou, W., Chen, D., Wei, T., Zhang, W. and Yu, N., 2021. Multi-attentional deepfake detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2185-2194).
23. Cao, X. and Gong, N.Z., 2021. Understanding the Security of Deepfake Detection. arXiv preprint arXiv:2107.02045.
24. Neekhara, P., Dolhansky, B., Bitton, J. and Ferrer, C.C., 2021. Adversarial threats to deepfake detection: A practical perspective. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 923-932).
25. Nirkin, Y., Wolf, L., Keller, Y. and Hassner, T., 2021. DeepFake detection based on discrepancies between faces and their context. IEEE Transactions on Pattern Analysis and Machine Intelligence.
26. Mittal, T., Bhattacharya, U., Chandra, R., Bera, A. and Manocha, D., 2020, October. Emotions Don't Lie: An Audio-Visual Deepfake Detection Method using Affective Cues. In Proceedings of the 28th ACM international conference on multimedia (pp. 2823-2832).
27. Agarwal, S. and Varshney, L.R., 2019. Limits of deepfake detection: A robust estimation viewpoint. arXiv preprint arXiv:1905.03493.
28. Zhou, Y. and Lim, S.N., 2021. Joint Audio-Visual Deepfake Detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 14800-14809).
29. Gandhi, A. and Jain, S., 2020, July. Adversarial perturbations fool deepfake detectors. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
30. Fagni, T., Falchi, F., Gambini, M., Martella, A. and Tesconi, M., 2021. TweepFake: About detecting deepfake tweets. Plos one, 16(5), p.e0251415.
31. Guarnera, L., Giudice, O. and Battiato, S., 2020. Fighting deepfake by exposing the convolutional traces on images. IEEE Access, 8, pp.165085-165098.

32. Li, Y., Zhang, C., Sun, P., Ke, L., Ju, Y., Qi, H. and Lyu, S., 2021, May. DeepFake-o-meter: An Open Platform for DeepFake Detection. In 2021 IEEE Security and Privacy Workshops (SPW) (pp. 277-281). IEEE.
33. Punnappurath, A. and Brown, M.S., 2019. Learning raw image reconstruction-aware deep image compressors. *IEEE transactions on pattern analysis and machine intelligence*, *42*(4), pp.1013-1019.
34. Cheng, Z., Sun, H., Takeuchi, M. and Katto, J., 2019. Energy compaction-based image compression using convolutional autoencoder. *IEEE Transactions on Multimedia*, *22*(4), pp.860-873.
35. Chorowski, J., Weiss, R.J., Bengio, S. and van den Oord, A., 2019. Unsupervised speech representation learning using wavenet autoencoders. *IEEE/ACM transactions on audio, speech, and language processing*, *27*(12), pp.2041-2053.
36. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. *Advances in neural information processing systems*, *27*.
37. Faceswap: Deepfakes software for all. Available at https://github.com/deepfakes/faceswap
38. FakeApp 2.2.0. Available at https://www.malavida.com/en/soft/fakeapp/
39. DeepFaceLab. Available at https://github.com/iperov/DeepFaceLab
40. DFaker. Available at https://github.com/dfaker/df
41. DeepFake tf: Deepfake based on tensorflow. Available at https://github.com/StromWine/DeepFake tf
42. Keras-VGGFace: VGGFace implementation with Keras framework. Available at https://github.com/rcmalli/keras-vggface
43. Faceswap-GAN. Available at https://github.com/shaoanlu/faceswap-GAN.
44. FaceNet. Available at https://github.com/davidsandberg/facenet.
45. CycleGAN. Available at https://github.com/junyanz/pytorchCycleGAN-and-pix2pix.
46. Li, L., Bao, J., Yang, H., Chen, D., & Wen, F. (2019). FaceShifter: Towards high fidelity and occlusion aware face swapping. arXiv preprint arXiv:1912.13457.
47. DSSIM. Available at https://github.com/keras-team/kerascontrib/blob/master/keras contrib/losses/dssim.py.
48. Liu, M. Y., Huang, X., Mallya, A., Karras, T., Aila, T., Lehtinen, J., and Kautz, J. (2019). Few-shot unsupervised image-to-image translation. In Proceedings of the IEEE International Conference on Computer Vision (pp. 10551-10560).
49. Park, T., Liu, M. Y., Wang, T. C., and Zhu, J. Y. (2019). Semantic image synthesis with spatially-adaptive normalization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2337-2346).
50. [50] DeepFaceLab: Explained and usage tutorial. Available a https://mrdeepfakes.com/forums/thread-deepfacelabexplained-and-usage-tutorial.
51. Ha, S., Kersner, M., Kim, B., Seo, S., & Kim, D. (2020, April). MarioNETte: few-shot face reenactment preserving identity of unseen targets. In Proceedings of the AAAI Conference on Artificial Intelligence (vol. 34, no. 07, pp. 10893-10900).
52. Thies, J., Elgharib, M., Tewari, A., Theobalt, C., & Nießner, M. (2020, August). Neural voice puppetry: Audio-driven facial reenactment. In European Conference on Computer Vision (pp. 716-731). Springer, Cham.
53. Chan, C., Ginosar, S., Zhou, T., & Efros, A. A. (2019). Everybody dance now. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 5933-5942).
54. Olszewski, K., Tulyakov, S., Woodford, O., Li, H., & Luo, L. (2019). Transformable bottleneck networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 7648-7657).
55. Tewari, A., Elgharib, M., Bharaj, G., Bernard, F., Seidel, H. P., P´erez, P., ... & Theobalt, C. (2020). StyleRig: Rigging StyleGAN for 3D control over portrait images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6142-6151).
56. Deng, Y., Yang, J., Chen, D., Wen, F., & Tong, X. (2020). Disentangled and controllable face image generation via 3D imitative-contrastive learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5154-5163).
57. Lattas, A., Moschoglou, S., Gecer, B., Ploumpis, S., Triantafyllou, V., Ghosh, A., & Zafeiriou, S. (2020). AvatarMe: realistically renderable 3D facial reconstruction ”in-the-wild”. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 760-769).
58. Nirkin, Y., Keller, Y., & Hassner, T. (2019). FSGAN: subject agnostic face swapping and reenactment. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 7184-7193).
59. https://lingzhili.com/FaceShifterPage
60. https://github.com/dfaker/df
61. https://github.com/shaoanlu/fewshot-face
62. https://github.com/StromWine/DeepFake tf
63. https://github.com/iperov/DeepFaceLab
64. https://hyperconnect.github.io/MarioNETte
65. https://justusthies.github.io/posts/neural
66. https://github.com/carolineec/EverybodyDanceNow
67. https://github.com/kyleolsz/TB-Networks
68. https://github.com/deepfakes/faceswap
69. https://gvv.mpi-inf.mpg.de/projects/StyleRig
70. https://github.com/microsoft/DiscoFaceGAN
71. https://github.com/shaoanlu/faceswap-GAN
72. https://github.com/lattas/AvatarMe
73. https://github.com/YuvalNirkin/fsgan
74. Chesney, R. and Citron, D. K. (2018, October 16). Disinformation on steroids: The threat of deep fakes. Available at https://www.cfr.org/report/deep-fake-disinformation-steroids.
75. Hasan, H. R., and Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. IEEE Access, 7, 41596-41606.
76. Maras, M. H., and Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. The International Journal of Evidence and Proof, 23(3), 255-262.
77. Agarwal, S., Farid, H., Fried, O., and Agrawala, M. (2020). Detecting deep-fake videos from phoneme-viseme mismatches. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 660- 661).

78. Suwajanakorn, S., Seitz, S. M., and Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama: learning lip sync from audio. *ACM Transactions on Graphics (TOG)*, 36(4), 1–13.
79. Fried, O., Tewari, A., Zollho¨fer, M., Finkelstein, A., Shecht- man, E., Goldman, D. B., ... and Agrawala, M. (2019). Text- based editing of talking-head video. *ACM Transactions on Graphics (TOG)*, 38(4), 1-14.
80. Chintha, A., Thai, B., Sohrawardi, S. J., Bhatt, K. M., Hicker- son, A., Wright, M., and Ptucha, R. (2020). Recurrent convolutional structures for audio spoof and video deepfake detection. *IEEE Journal of Selected Topics in Signal Processing*, DOI: 10.1109/JSTSP.2020.2999185.
81. Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1-11).
82. Li, Y., Yang, X., Sun, P., Qi, H., and Lyu, S. (2020). Celeb-DF: A large-scale challenging dataset for deepfake forensics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3207-3216).
83. Todisco, M., Wang, X., Vestman, V., Sahidullah, M., Delgado, H., Nautsch, A., ... and Lee, K. A. (2019). ASVspoof 2019: Future horizons in spoofed and fake audio detection. *arXiv preprint* arXiv:1904.05441.
84. Matern, F., Riess, C., and Stamminger, M. (2019, January). Exploiting visual artifacts to expose deepfakes and face manipulations. In *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)* (pp. 83-92). IEEE.
85. Mao, X., Li, Q., Xie, H., Lau, R. Y., Wang, Z., and Paul Smolley, S. (2017). Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2794-2802).
86. Afchar, D., Nozick, V., Yamagishi, J., and Echizen, I. (2018, December). MesoNet: a compact facial video forgery detection network. In *2018 IEEE International Workshop on Information Forensics and Security (WIFS)* (pp. 1-7). IEEE.
87. Thies, J., Zollhofer, M., Stamminger, M., Theobalt, C., and Nießner, M. (2016). 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).
88. Guera, D., and Delp, E. J. (2018, November). Deepfake video detection using recurrent neural networks. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-6). IEEE.
89. Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., & Guo, B. (2020). Face X-ray for more general face forgery detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5001-5010).
90. Dufour, N., and Gully, A. (2019). Contributing Data to Deepfake Detection Research. Available at: https://ai.googleblog.com/2019/09/contributing-data-to- deepfake-detection.html.
91. Dolhansky, B., Howes, R., Pflaum, B., Baram, N., and Fer- rer, C. C. (2019). The deepfake detection challenge (DFDC) preview dataset. *arXiv preprint* arXiv:1910.08854.
92. Ciftci, U. A., Demir, I., & Yin, L. (2020). FakeCatcher: Detec- tion of synthetic portrait videos using biological signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, DOI: 10.1109/TPAMI.2020.3009287.
93. Yang, X., Li, Y., and Lyu, S. (2019, May). Exposing deep fakes using inconsistent head poses. In *2019 IEEE Interna- tional Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8261-8265). IEEE.
94. Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., and Nießner, M. (2018). FaceForensics: A large-scale video dataset for forgery detection in human faces. *arXiv preprint* arXiv:1803.09179.
95. Zhang, Y., Zheng, L., and Thing, V. L. (2017, August). Automated face swapping and its detection. In *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)* (pp. 15-19). IEEE.
96. Huang, G. B., Mattar, M., Berg, T., and Learned-Miller,E. (2007, October). Labelled faces in the wild: A database for studying face recognition in unconstrained environ- ments. Technical Report 07-49, University of Massachusetts, Amherst, http://vis-www.cs.umass.edu/lfw/.
97. Hsu, C. C., Zhuang, Y. X., and Lee, C. Y. (2020). Deep fake image detection based on pairwise learning. *Applied Sciences*, 10(1), 370.
98. Chopra, S. (2005). Learning a similarity metric discriminatively, with application to face verification. In IEEE Conference on Compter Vision and Pattern Recognition (pp. 539- 546).
99. Liu, Z., Luo, P., Wang, X., and Tang, X. (2015). Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3730-3738).
100. Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint* arXiv:1511.06434.
101. Arjovsky, M., Chintala, S., and Bottou, L. (2017, July). Wasserstein generative adversarial networks. In *International Conference on Machine Learning* (pp. 214-223).
102. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved training of Wasserstein GANs. In *Advances in Neural Information Processing Systems* (pp. 5767-5777).
103. Karras, T., Aila, T., Laine, S., and Lehtinen, J. (2017). Pro- gressive growing of GANs for improved quality, stability, and variation. *arXiv preprint* arXiv:1710.10196.
104. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770-778).
105. Cao, Q., Shen, L., Xie, W., Parkhi, O. M., and Zisserman, A. (2018, May). VGGFace2: A dataset for recognising faces across pose and age. In *2018 13th IEEE International Confer- ence on Automatic Face and Gesture Recognition* (pp. 67-74). IEEE.
106. Jha, S., Raj, S., Fernandes, S., Jha, S. K., Jha, S., Jalaian, B.,... and Swami, A. (2019). Attribution-based confidence metric for deep neural networks. In *Advances in Neural Information Processing Systems* (pp. 11826-11837).
107. Fernandes, S., Raj, S., Ortiz, E., Vintila, I., Salter, M., Urosevic, G., and Jha, S. (2019, October). Predicting heart rate variations of deepfake videos using neural ODE. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)* (pp. 1721-1729). IEEE.
108. Korshunov, P., and Marcel, S. (2018). Deepfakes: a new threat to face recognition? assessment and detection. *arXiv preprint* arXiv:1812.08685.