IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

Exploring Deepfake Generation and Detection: A Comparative Study

Yash J. Ingole[1], Dhawal P. Ramdham[2], Pranjal P. Hejib[3], Pratik S. Rohankar[4], Roshan R. Karwa[5]

Department of Computer Science and Engineering, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra – 444701, India

Abstract: Deep learning has been effectively used to address a variety of challenging issues, from computer vision to big data analytics. Advances in deep learning have also been used to develop software that poses risks to national security, privacy, and democracy. Deepfake is one of those recently developed deep learning-based applications. The recent entry of deepfakes has marked a turning point in the development of fake material, even if manipulations of visual and auditory media are as old as media itself. Deepfakes offer automated methods for producing fake content that is getting more difficult for human observers to spot thanks to the most recent developments in artificial intelligence and machine learning. Therefore, the development of systems that can instantly identify and evaluate the integrity of digital visual media is essential. The discipline of computer vision, a branch of computer science, has developed methods for producing and identifying deepfakes. Humanities and social science approaches have focused on the social and ethical ramifications of deepfakes. This study examines the algorithms used to produce deep fakes.

Keywords: Deepfakes, Generation and Detection of Deepfake, Generative Adversarial Network, Autoencoders.

I. INTRODUCTION

A deepfake is a convincingly realistic-looking yet artificial intelligence-altered image, audio, or video. The underlying technology enables the replacement of faces, editing of facial expressions, creation of synthetic faces, and voice synthesis. Deepfakes can show someone saying or doing something that they have never actually said or done. Deepfakes are frequently employed for exploitation, despite the fact that they have positive and acceptable applications in fields like entertainment and business. Deepfakes could also be used for other objectives, notably disinformation, which is another worry. Deepfakes could be used as a tool of psychological warfare, to sway elections, to stir up social disturbance, or both. Moreover, they might cause people to ignore valid proof of wrongdoing, undermining public confidence in audio-visual content more broadly. The term "Deepfake" is a combination of "Deep Learning (DL)" and "Fake," and it refers to photo-realistic video or image content created with DL's assistance. Deepfakes are the outcome of face swapping, facial expression enactment/animation, and/or digitally generated sounds or non-existing human faces. Face manipulation, on the other hand, entails changing facial characteristics such as age, gender, ethnicity, morphing, attractiveness, skin colour or texture, hair colour, style or length, eyeglass, makeup, moustache, emotion, beard, pose, gaze, mouth open or closed, eye colour, injury, and drug effects [11,12], as well as adding imperceptible perturbations (i.e., adversarial examples). To create such fake videos, two neural networks were used: (i) a generative network and (ii) a discriminative network with a FaceSwap technique [3], [4]. Using an encoder and a decoder, the generative network generates fictitious images. The discriminative network determines the legitimacy of newly generated images. Ian Goodfellow [5] proposed Generative Adversarial Networks (GANs) as a combination of these two networks. It is critical to have systems in place to detect, fight, and defeat deepfake digital content such as fake movies, photos, paintings, audios, and so on. It is not difficult to achieve this goal if there is a legitimate, safe, and trusted means to trace the history of digital content. Consumers should have access to trusted data provenance of digital information and should be able to trace an item back in time to establish its originality and authenticity [1]. This method can help people avoid being duped into believing in fraudulent digital content.

II. RELATED WORK

In this section, we review and discuss related work found in the literature on the approaches and algorithms used in creation and detection of deepfake content.

Li and Lyu [6] proposed an Artificial Intelligence-based approach for detecting deepfake videos. The proposed solution is based on one AI algorithm vs. another AI algorithm. Their method is based on training convolutional neural networks

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

(CNN) utilising modified and real-world data. Four different CNN networks were used in the testing, with accuracy ranging from 84% to 99%. Their findings appear encouraging, but the authors acknowledge that there are numerous problems that must be addressed.

Mirsky Y. et al. [7] examine visual deepfake production techniques in detail, although deepfake detection approaches are only briefly covered. This paper is the first to attempt to provide a comprehensive study and assessment of both audio and visual deepfake detection algorithms, as well as generative approaches. It contributes to the research community by providing insight into many forms of deepfake production and detection methods based on video and audio. It informs the reader on the most recent advancements, trends, limitations, and issues in audio-visual deepfakes.

Kingma, D.P., Welling, M. [8] in their paper mentioned that GANs, variational autoencoders, and fully-visible belief nets are perhaps the most prominent approaches to generative modelling (e.g., Frey[9],[10]). Because none of these approaches rely on Markov chains, the reason for the current interest in GANs is not that they achieved their original goal of generative modelling without Markov chains, but rather that they have succeeded in generating high-quality images and have proven useful for a variety of tasks other than straightforward generation.

Vougioukas et al. [13] created a photorealistic film immediately from a still image and speech data using a temporal GAN composed of an RNN. Without relying on manually produced audio-visual characteristics, the finished film had coordinated lip motions, eye-blinking, and natural face emotion. To manage frame quality, audio-visual synchronisation, and overall video quality, many discriminators were used.

According to a yearly report [14] in Deepfake, DL researchers accomplished numerous related generative modelling advancements. For example, computer vision researchers proposed the Face2Face [15] approach for facial re-enactment. This technology translates facial expressions from a single person to a true digital 'avatar' in real time. CycleGAN [16] was developed by UC Berkeley researchers in 2017 to change photos and videos into various styles.

Rossler et al. [25] presented a new dataset and methodology for detecting manipulated facial images, which they call FaceForensics++. The dataset consists of over 1,000 original videos and more than 500,000 manipulated videos, covering four different manipulation methods: Deepfakes, Face2Face, FaceSwap, and NeuralTextures. The authors then propose a two-stage approach for detecting manipulated facial images. In the first stage, they train a binary classifier to distinguish between original and manipulated videos using a set of handcrafted features. In the second stage, they use a convolutional neural network (CNN) to classify the type of manipulation used in the video.

Kim H. et al. [18] evaluated their FaceSwap algorithm on a dataset of videos and compare its performance to other state-of-the-art methods. They report the precision, recall, and F1-score for detecting FaceSwap videos, as well as the AUC-ROC curve and the detection rate for different levels of false positive rates. The authors also analyze the failure cases and limitations of their algorithm, providing valuable insights into the challenges of detecting FaceSwap videos.

Akhtar et al. [11] provided an overview of face manipulation techniques, specifically deepfake generation, detection, and recognition. It discusses the importance of face authenticity in various fields, including law enforcement, security, and media. The paper also covers different types of face manipulation techniques such as image splicing, video manipulation, face swapping, and deepfakes. It provides a brief description of the working principles of these techniques and their potential implications.

Tariq et al. [34] proposed a new method for detecting deepfake videos using a convolutional long short-term memory (ConvLSTM) based residual network. According to the authors' literature assessment, the rapid development of deep learning and computer vision techniques has resulted in an increase in deepfake films, which can be exploited for malevolent reasons such as spreading misinformation and propaganda. Several deepfake detection methods have been developed, however they frequently have flaws, such as being computationally expensive or ineffective versus newer deepfake algorithms.

S. Khan et al. [19] proposed a deep learning-based approach to detect deepfake images and videos. The authors utilize a fused convolutional neural network (CNN) approach that incorporates three different CNN architectures to improve the robustness of deepfake detection. The authors also introduce a new dataset for deepfake detection, which includes various types of manipulations such as face swapping, expression modification, and attribute modification.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

III. DEEPFAKE GENERATION AND DETECTION

Deepfake/face manipulation can be categorized into four main groups: identity swap, face re-enactment, attribute manipulation, and entire face synthesis [17].

3.1 Re-enactment of a Face

An overview of earlier face re-enactment (changing the individual's facial expression) production and detection techniques is offered here.

Face Re-enactment Generation: This entails substituting one person's facial expression in the target image/video with another person's facial expression in the source image/video [17]. It is often referred to as expression swapping or puppet mastering. Thies et al. [15], for example, created real-time face recreation RGB video streams. Kim et al. [18], Zhang et al. [20], Doukas et al. [21], and Cao et al. [22] designed encoder-decoder, RNN, unified landmark converter with geometry-aware generator, GANs, and task-agnostic GANs-based schemes, respectively.

Face Re-enactment Detection : Cozzolino et al. [23] used CNNs to design face re-enactment detection methods; Matern et al. [24] used visual features with logistic regression and MLP; Rossler et al. [25] used mesoscopic, steganalysis, and CNN features; Sabir et al. [26] used RNN; Amerini et al. [27] used Optical Flow + CNNs; Kumar et al. In contrast, Zhao et al. [28] created a spatiotemporal network that can use complimentary global and local information. The system employs a spatial module for global information, while the local information module pulls features from patches chosen by attention layers.

3.2 Attribute Manipulation:

This area provides an overview of existing attribute manipulation or face retouching or face editing generation and detection techniques (i.e., changing facial attributes such as skin tone, age, and gender).

Attribute Manipulation Generation: This entails changing various facial characteristics, such as hair/skin colour, gender, age, and the addition of glasses. It is also referred to as face editing or face retouching. Xiao et al. [2] demonstrated a multi-attribute manipulation GANs-based system. Spatial attention in GANs, variational autoencoder (VAE) + GANs, multi-domain GANs, geometry-aware GANs, mask-guided GANs, 3D face-morphable model, and GIMP animation-based approaches have also been developed.

Attribute Manipulation Detection: The efficacies of different deep learning models were concluded for attribute manipulation detection which includes the deep Boltzmann machine, CNN, Adaptive manipulation traces, encoder-decoder, facial boundary features and PRNU were exploited.

3.3 Identity Swap

Here, an overview of existing identity swap or face swap (i.e., replacing a person’s face with another person’s face) generation and detection methods is presented.

Identity Swap Generation: This consists of replacing the face of a person in the target image/video with the face of another person in the source image/video [29]. For example, Korshunova et al. [30] developed a face-swapping method using Convolutional Neural Networks (CNNs). While Nirkin et al. [31] proposed a technique using a standard fully convolutional network in unconstrained settings. Mahajan et al. [32] presented a face swap procedure for privacy protection.

Identity Swap Detection: Several experiments on identity swap deepfake detection have been undertaken. Koopman et al. [33] investigated photo response non-uniformity (PRNU) for detection. Additionally employed were warping artefacts, eye blinking, optical flow with CNNs [27], heart rate, picture quality, local image textures, long short-term memory (LSTM) and recurrent neural network (RNN) [34], multi-LSTM and blockchain.

3.4 Entire Face Synthesis:

The process of entire face synthesis typically involves collecting a large dataset of images of real human faces, which is then used to train a deep neural network. The network learns to identify the key features of a human face, such as the shape of the eyes, nose, and mouth, as well as the texture and colour of the skin. Once the network is trained, it can be used to generate new images of human faces that are not present in the original dataset. This can be done by providing the network with a set of random input values, which are then used to generate a new image.

IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.124107

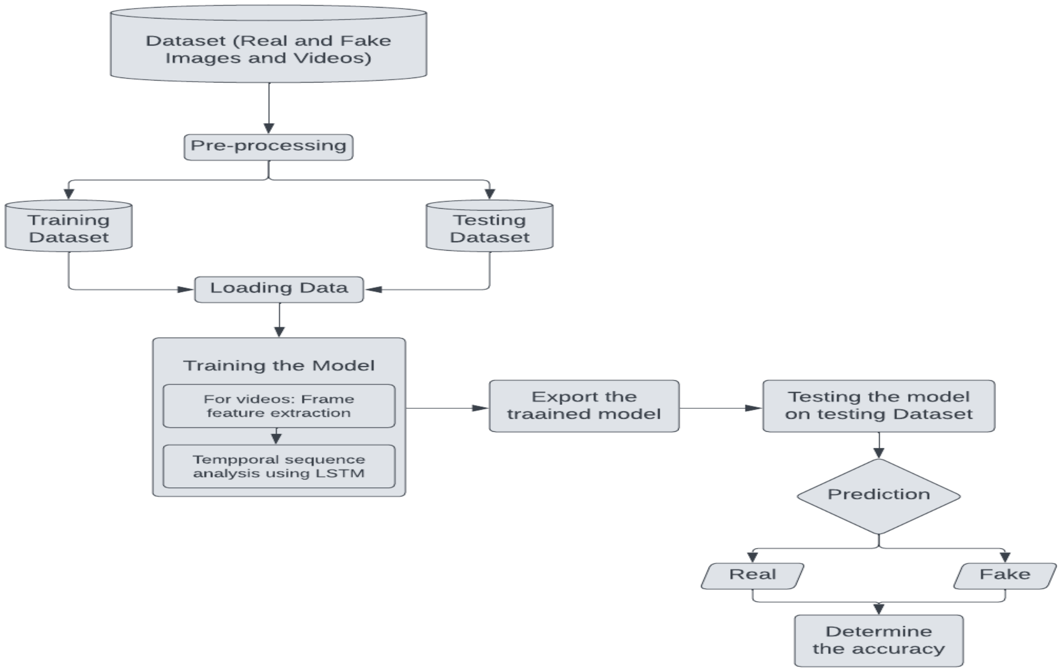

Figure 3.1: Flowchart for Generation and Detection of Deepfakes

IV. ALGORITHMS USED IN DEEPFAKE TECHNOLOGY.

Some of the algorithms used in deepfake technology include GANs, Autoencoders, CNNs, DNNs, RNNs. These algorithms and approaches mainly consist of deep neural network architecture that are well suited for processing visual data.

4.1 Generative Adversarial Network

The Generative Adversarial Network (GAN) has shown tremendous capability and potential in the machine learning world to create realistic-looking images and videos. A GAN is a system that consists of two models: a generator and a discriminator.

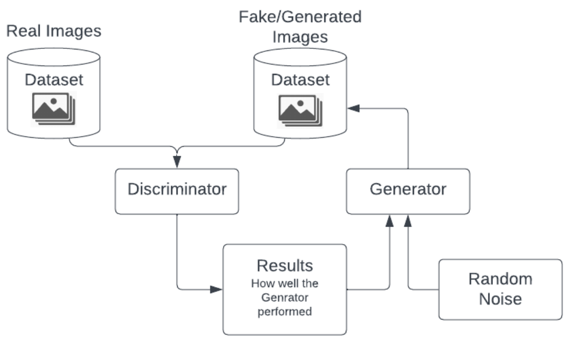

Fig. 4.1 GAN Flowchart

The discriminator is basically a classifier that identifies whether a particular image is from the dataset or was made artificially by the generator. A convolutional neural network will be used to implement this binary classifier. The purpose of the generator is to accept random input values (noise) and generate a picture from them using a deconvolutional neural network. Consider it like seeding a random number generator: the same input noise will produce the same output. The generator creates an image by using the random noise as a sort of seed. The purpose of the two-model system is for the

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁐ Impact Factor 8.102 ⁐ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

generator to trick the discriminator while still allowing the discriminator to classify the generator's images as accurately as possible. This continual conflict between two generative adversarial network implies that both models improve by attempting to outperform the other. The discriminator provides input to the generator on how convincing its images are, and the generator provides more data for the discriminator to train on.

4.2 Autoencoder

The autoencoder is composed of an encoder with seven convolutional layers followed by a fully connected layer and a decoder composed of a fully connected layer followed by seven deconvolution layers. After each of the convolution and deconvolution layer, there is a batch normalization layer used to speed up training and reduce the influence of initial network weights. The ReLu is used as the activation function for all the batch normalization layers. The Sigmoid function is used as the activation function for the decoder's output layer.

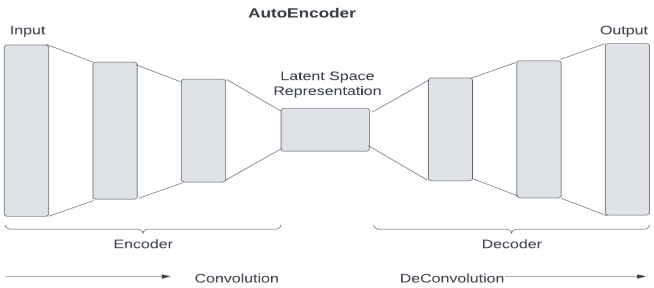

Fig. 4.2 Autoencoder

4.3 Convolutional Neural Network

The CNN consists of multiple layers, including Convolutional layers, Pooling layers, and fully connected layers. The Convolutional layer applies filters to the input image to extract features, the Pooling layer down samples the image to reduce computation, and the fully connected layer makes the final prediction. The network learns the optimal filters through backpropagation and gradient descent.
The convolution generalizes an image into one of the lower dimensions.

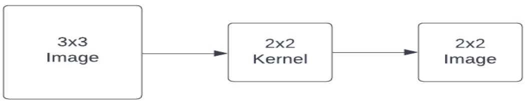

Fig. 4.3 Convolution of image by CNN

4.4 Deconvolutional Neural Network

The images are generated by the deconvolutional neural network (DNN) technique. Deconvolutional neural networks, as the name implies, can be thought of as "running a CNN backward," although the mechanics are far more sophisticated. DNNs, also known as deconvs or transposed convolutional neural networks, use layers identical to those found in CNNs — but backward — to up-sample (rather than down-sample) images, allowing them to be larger.

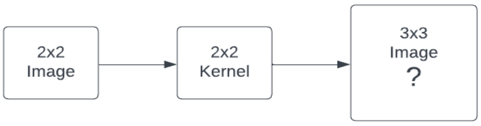

Fig. 4.4 Deconvolution

This is more challenging than a CNN. Getting an image smaller by compressing or generalising its information is significantly easier than making an image larger without blurring or losing detail. The transposed convolution solves this problem (deconvolution).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⪤ Impact Factor 8.102 ⪤ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

V. STEPS TO TRAIN A DEEPFAKE DETECTION MODEL.

Collect a dataset of real videos and deepfakes. The FaceForensics dataset is a good starting point, but it may need supplementary data to improve the model's performance.

- Pre-process the data. This involves cropping the videos, resizing them to a standard resolution, and converting them to a consistent format.
- Train a deep learning model. Various types of deep learning models for deepfake detection can be used, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or a combination of the two. You may also want to experiment with different architectures and hyperparameters to optimize your model's performance.
- Augment the data. You can use data augmentation techniques such as random cropping, flipping, or rotation to increase the diversity of your training data and improve your model's generalization ability.
- Validate your model. Use a separate validation set to evaluate your model's performance and fine-tune it as necessary.
- Test your model. Use a separate test set to evaluate your model's performance on unseen data and compare it to other deepfake detection algorithms.
- Deployment: Integrate the model into a system that can analyse new videos or images in real-time or batch mode, and provide a confidence score indicating the likelihood that the media is a deepfake.
- Continuous monitoring: Deepfake technology is constantly evolving, so the detection model needs to be updated regularly to keep up with new techniques and methods used by deepfake creators.

VI. COMPARISION OF PERFORMANCE METRICS FOR DUFFFERENT METHODS

FaceForensics++ [25]: This is a dataset of videos that have been manipulated using various deepfake methods. Researchers have developed several deepfake detection algorithms based on this dataset, including a method that uses a combination of CNNs and LSTMs to analyze the frames of a video. Experimental evidence suggests that these methods can be effective against some types of deepfakes, although they may not work as well against more advanced deepfakes.

FaceSwap [18]: This is a deepfake method that involves swapping the faces of two individuals in a video. Researchers have developed several deepfake detection methods specifically for FaceSwap, including a method that uses a combination of CNNs and LSTMs to analyze the frames of a video. Experimental evidence suggests that these methods can be effective against FaceSwap deepfakes.

Xception: This is a type of CNN that has been used for deepfake detection. Researchers have developed Xception-based deepfake detection methods that analyze the frames of a video for anomalies that might indicate the presence of a deepfake. Experimental evidence suggests that these methods can be effective against some types of deepfakes.

Mesonet [23], [25]: This is a deepfake detection method that uses a combination of CNNs and LSTMs to analyze the frames of a video for anomalies that might indicate the presence of a deepfake. Experimental evidence suggests that Mesonet can be effective against some types of deepfakes, although it may not work as well against more advanced deepfakes.

MesoInception-4 [12]: This is a deepfake detection method that uses a combination of CNNs and Inception blocks to analyze the frames of a video for anomalies that might indicate the presence of a deepfake. Experimental evidence suggests that MesoInception-4 can be effective against some types of deepfakes.

CFS-NET [21]: This is a deepfake detection method that uses a combination of CNNs and attention mechanisms to analyze the frames of a video for anomalies that might indicate the presence of a deepfake. Experimental evidence suggests that CFS-NET can be effective against some types of deepfakes.

RIFE-Flow: This is a deepfake detection method that uses optical flow analysis to detect inconsistencies in the motion of objects in a video. Experimental evidence suggests that RIFE-Flow can be effective against some types of deepfakes, although it may not work as well against more advanced deepfakes.

CelebA-HQ [19] is a large-scale face attributes dataset, which is an extended version of the CelebA dataset. It consists of 30,000 high-resolution images of celebrity faces and includes annotations for 40 different facial attributes such as gender, age, and facial expression. ResNet50 is a deep convolutional neural network architecture commonly used for image classification tasks. It has 50 layers and was introduced in 2015 by Microsoft Research. In the context of CelebA-HQ,

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

ResNet50 can be used for tasks such as facial attribute recognition, where it can be trained on the dataset to classify the various facial attributes present in the images.

Table 6.1 Comparing various factors of DeepFake Detection Algorithms:

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC	Detection Rate	False Positive Rate	False Negative Rate
FaceForensics++ Rossler et al. [25]	0.924	0.903	0.936	0.919	0.973	0.897	0.076	0.064
FaceSwap, Kim H. et al. [18]	0.973	0.977	0.970	0.973	0.995	0.960	0.015	0.030
XceptionNet	0.942	0.927	0.956	0.941	0.974	0.902	0.067	0.044
MesoNet Cozzolino, D.; et al. [23],[25]	0.921	0.903	0.935	0.919	0.970	0.897	0.076	0.064
MesoInception-4 [12] Mirsky, Y	0.974	0.977	0.970	0.973	0.991	0.960	0.022	0.030
CFS-NET [21] Doukas, M.	0.981	0.985	0.977	0.981	0.994	0.970	0.010	0.023
Head2Head++ [21] Doukas, M	0.975	0.965	0.985	0.975	0.990	0.980	0.025	0.015
CelebA-HQ, ResNet-50 S.A.Khan et al [19]	0.94	0.91	0.96	0.93	0.97	0.89	0.07	0.11

Depending on the dataset utilised and other experimental details, the values for each factor may vary. This table is merely an example for comparison.

VII. CONCLUSION AND FUTURE SCOPE

In this research paper, the current state-of-the-art deepfake detection and generation algorithms are explored. First a comprehensive review of the various techniques used in deepfake generation and detection is presented. A comparative analysis of several popular deepfake detection algorithms, including feature-based, CNN-based, and GAN-based models is performed, based on their performance on benchmark datasets. Results show that while some algorithms perform better than others, there is still room for improvement in deepfake detection. In addition, some of the latest deepfake generation techniques, including GANs and neural networks are reviewed. These methods have shown great potential for generating highly realistic deepfakes, which can be used for malicious purposes. The limitations of current detection techniques, which have not yet caught up with the rapidly advancing technology of deepfake generation are discussed.

Overall, this research shows that deepfake detection and generation are rapidly evolving fields, with ongoing developments in algorithms and techniques. However, there is still much work to be done to improve the accuracy and effectiveness of deepfake detection, as well as to prevent the malicious use of deepfake technology. As deepfake technology becomes increasingly sophisticated, it is essential to continue to develop and refine detection methods to ensure that we can identify and mitigate the risks posed by this technology.

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁒ Impact Factor 8.102 ⁒ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124107

REFERENCES

[1] L. Floridi, ‘‘Artificial intelligence, deepfakes and a future of ectypes,’’ Philosophy Technol., vol. 31, no. 3, pp. 317–321, 2018.
[2] Xiao, T.; Hong, J.; Ma, J. Elegant: Exchanging latent encodings with gan for transferring multiple face attributes. In Proceedings of the ECCV, Munich, Germany, 8–14 September 2018; pp. 168–184
[3] G. Oberoi, Exploring DeepFakes, Jan. 2021, [online] Available: https://goberoi.com/exploring-deepfakes-20c9947c22d9.
[4] J. Hui, How Deep Learning Fakes Videos (Deepfake) and How to Detect it, Jan. 2021, [online] Available: https://medium.com/how-deep-learning-fakes-videos-deepfakes-and-how-to-detect-it-c0b50fbf7cb9.
[5] I. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, et al., "Generative adversarial nets", Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS), vol. 2, pp. 2672-2680, 2014.
[6] Y. Li and S. Lyu. (2018). ‘‘Exposing deepfake videos by detecting face warping artifacts.’’ [Online]. Available: https://arxiv.org/abs/1811.00656
[7] Y. Mirsky and W. Lee, ”The Creation and Detection of Deepfakes: A Survey,” arXiv preprint arXiv:2004.11138, 2020
[8] Kingma, D.P., Welling, M. Autoencoding variational bayes. In Proceedings of the International Conference on Learning Representations (ICLR) (2014).
[9] Frey, B.J. Graphical Models for Machine Learning and Digital Communication. MIT Press, Boston, 1998.
[10] Oord, A. v. d., Li, Y., Babuschkin, I., Simonyan, K., Vinyals, O., Kavukcuoglu, K., Driessche, G. v. d., Lockhart, E., Cobo, L.C., Stimberg, F., et al. Parallel wavenet: Fast high-fidelity speech synthesis. arXiv preprint arXiv:1711.10433 (2017).
[11] Akhtar, Z.; Dasgupta, D.; Banerjee, B. Face Authenticity: An Overview of Face Manipulation Generation, Detection and Recognition. In Proceedings of the ICCIP, Talegaon-Pune, India; pp. 1–8.
[12] Mirsky, Y.; Lee, W. The creation and detection of deepfakes: A survey. ACM Comput. Surv. 2021,54, 1–41.
[13] K. Vougioukas, S. Petridis, and M. Pantic, ”End-to-End Speech-Driven Realistic Facial Animation with Temporal GANs,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 37-40
[14] G. Patrini, F. Cavalli, and H. Ajder, ‘‘The state of deepfakes: Reality under attack,’’ Deeptrace B.V., Amsterdam, The Netherlands, Annu. Rep. v.2.3., 2018.
[15] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt, and M. Niessner, ‘‘Face2Face: Real-time face capture and reenactment of RGB videos,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2387–2395, doi: 10.1109/CVPR.2016.262.
[16] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, ‘‘Unpaired image-to-image translation using cycle-consistent adversarial networks,’’ in Proc. ICCV, Venice, Oct. 2017, pp. 2242–2251, doi: 10.1109/ICCV.2017.244.
[17] Juefei-Xu, F.; Wang, R.; Huang, Y.; Guo, Q.; Ma, L.; Liu, Y. Countering Malicious DeepFakes: Survey, Battleground, and Horizon. Int. J. Comput. Vis. 2022, 130, 1678–1734. [CrossRef] [PubMed]
[18] Kim, H.; Garrido, P.; Tewari, A.; Xu, W.; Thies, J.; Niessner, M.; Pérez, P.; Richardt, C.; Zollhofer, M.; Theobalt, C. Deep video portraits. ACM Trans. Graph. (TOG) 2018, 37, 1–4.
[19] Sohail Ahmed Khan, Alessandro Artusi, Hang Dai. Adversarially robust deepfake media detection using fused convolutional neural network predictions.
[20] Zhang, J.; Zeng, X.; Wang, M.; Pan, Y.; Liu, L.; Liu, Y.; Ding, Y.; Fan, C. Freenet: Multi-identity face reenactment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5326–5335.
[21] Doukas, M.; Koujan, M.; Sharmanska, V.; Roussos, A.; Zafeiriou, S. Head2Head++: Deep Facial Attributes Re-Targeting. IEEE Trans. Biom. Behav. Identity Sci. 2021, 3, 31–43.
[22] Cao, M.; Huang, H.; Wang, H.; Wang, X.; Shen, L.; Wang, S.; Bao, L.; Li, L.; Luo, J. Task-agnostic Temporally Consistent Facial Video Editing. arXiv 2020, arXiv:2007.01466.
[23] Cozzolino, D.; Thies, J.; Rossler, A.; Riess, C.; Niener, M.; Verdoliva, L. Forensictransfer: Weakly-supervised domain adaptation for forgery detection. arXiv 2018, arXiv:1812.02510.
[24] Matern, F.; Riess, C.; Stamminger, M. Exploiting Visual Artifacts to Expose DeepFakes and Face Manipulations. In Proceedings of the IEEE Winter Applications of Computer Vision Workshops, Waikoloa Village, HI, USA, 7–11 January 2019; pp. 1–10.
[25] Rossler, A.; Cozzolino, D.; Verdoliva, L.; Riess, C.; Thies, J.; Nießner, M. Faceforensics++: Learning to detect manipulated facial images. pp. 1–11.
[26] Sabir, E.; Cheng, J.; Jaiswal, A.; AbdAlmageed, W.; Masi, I.; Natarajan, P. Recurrent Convolutional Strategies for Face Manipulation Detection in Videos. pp. 1–8.

DOI: 10.17148/IJARCCE.2023.124107

[27] Amerini, I.; Galteri, L.; Caldelli, R.; Del Bimbo, A. Deepfake Video Detection through Optical Flow Based CNN. In Proceedings of the ICCVW, Seoul, Republic of Korea, 27–28 October 2019; pp. 1205–1207.

[28] Zhao, X.; Yu, Y.; Ni, R.; Zhao, Y. Exploring Complementarity of Global and Local Spatiotemporal Information for Fake Face Video Detection. In Proceedings of the 2022 (ICASSP), Singapore, 22–27 May 2022; pp. 2884–2888.

[29] Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Morales, A. Ortega-Garcia. Deepfakes and beyond: A survey of face manipulation and fake detection. Inf. Fusion 2020, 64, 131–148.

[30] Korshunova, I.; Shi, W.; Dambre, J.; Theis, L. Fast Face-Swap Using Convolutional Neural Networks. In Proceedings of the ICCV, Venice, Italy, 22–27 October 2017; pp. 3697–3705.

[31] Nirkin, Y.; Masi, I.; Tuan, A.T.; Hassner, T.; Medioni, G. On Face Segmentation, Face Swapping, and Face Perception. In Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition, Xi'an, China, 15–19 May 2018; pp. 98–105.

[32] Mahajan, S.; Chen, L.; Tsai, T. SwapItUp: A Face Swap Application for Privacy Protection. pp. 46–50.

[33] Koopman, M.; Rodriguez, A.; Geradts, Z. Detection of deepfake video manipulation. In Proceedings of the 20th Irish Machine Vision and Image Processing Conference (IMVIP), Coleraine, UK, 29–31 August 2018; pp. 133–136.

[34] Tariq, S.; Lee, S.; Woo, S. A Convolutional LSTM based Residual Network for Deepfake Video Detection. arXiv 2020, arXiv:2009.07480.

[35] Xueyu Wang, Jiajun Huang, Siqi Ma, Surya Nepal, Chang Xu; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 14920-14929

[36] Juan Hu, Xin Liao, Difei Gao, Satoshi Tsutsui, Zheng Qin, Mike Zheng Shou, DeepfakeMAE: Facial Part Consistency Aware Masked Autoencoder for Deepfake Video Detection.

