Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

7

An Exploration of Various Image and Video Deepfake Detection Methods and Models

Miss. Ashwini S. Kaware

Research Scholar,

Department of Computer Application, Sant. Gadge Baba Amravati University, Amravati.

Dr. S. S. Sherekar

Associate Professor,

Department of Computer Science, Sant. Gadge Baba Amravati University, Amravati.

Abstract:

Deepfake has become the most challenging attack in this digital age. It has become source of spreading misinformation and forging people through social media content. Deepfake attacks like face synthesis, attribute manipulation, identity swap, face reenactment and lips syncing are discussed in this paper. After deepfake is acquainted as a threat to human beings, various detection methods and models have developed based on machine learning, deep learning, transfer learning and digital forensics. In this research paper survey of deepfake detection techniques based on CNN, RNN, XceptionNet, EfficientNet, ResNext, MesoNet and, CLRNet, LSTM and other methods are provided. Different ensemble models are also developed by combining various algorithms. This paper provides an in-depth analysis of evolution, types, impacts and deepfake detection methods by considering visual artifacts, temporal features and

handcrafted features. This track record of deep-fake and evolving methods in the deepfake detection process could be helpful for researchers in their future research.

Keywords: Deepfake, Face Synthesis, Attribute manipulation, Identity Swap, Face Reenactment, Lip-Syncing.

I. Introduction:

The "Deepfake" is one of the trending topics in computer science and is a challenging threat to cyber security systems. It has evolved from the ocean for machine learning and deep learning technology. The generative adversarial network (GAN) technology is a base through which deepfake algorithms are generated. The deepfake could be in any form as forged image, video, audio or textual data. This paper mainly focuses on image or video-based deepfake.

The term "DeepFake" was coined by a Reddit user in 2017. Some people shared videos with celebrities' faces swapped onto the bodies of other persons. In the beginning, it was just a fun part to swap faces or manipulate the original to a synthetic one. Some proprietary software was also developed and launched. These apps include FakeApp, DeepFaceLab and web-based apps that provide access to the end user to alter image and video. [1] [2]

Later on there is a large increase in number of deepfake found. Also deepfakes of Nancy Pelosi in 2019 [3], Donald Trump in 2020 and Tom Cruise in 2021 were found. [4] The intention behind them was to spoil the reputation of well known personalities.

The Deepfake is capable to alter, manipulate or create the originality to partial or full synthetic or computerized content. Due to this ability, it is used for advertisements, AI effects in movies, education with simulation and also in medical sector. [5] It allured public to use itself in life, enjoy it and also share among other individuals. As each coin has two sides, deepfake looks too real to differentiate by people. This result in the unethical use of deepfake has risen by cyber-criminals. They can create deepfakes for passing misinformation, mental harassment or defamation, political chaos, financial fraud and so on. [6]

Since models get tested over multiple training datasets, the quality of machine-generated media contents is mounting. Also, such advanced technology is open to all. Thus, deepfake detection systems struggle to separate real and forged content. Even the APIs of Microsoft and Amazon could not easily differentiate between forged and real [7].

The deepfake is an outcome of artificial intelligence technology, and mostly generative adversarial network (GAN) based algorithms are used in its development. Generation of deepfake includes manipulation of expression, facial identification manipulation, changing the movements of the mouth to synchronize with random voice and also swapping faces or even whole body [8].

To prevent of spreading of deepfakes among communities, researchers have used various approaches for identification and manipulations in images, videos and audio content. The deeepfake identification approaches/methods are developed such that they can detect forgery based on inconsistencies that remained during its generation. XceptionNet, FaceForensic++, FakeSpotter, and GANomaly are some of the deepfake detection models that have been developed [9].

II. Related Work:

The deepfake is an emerging technology that has attracted the attention of common people, researchers and also the intruders. Many deepfakes are drifting over the internet. Some of them could be dangerous and could have the intention of spreading misinformation [10]. Some of them look too realistic to be identified by the human eye. To capture them, a robust mechanism is needed. Thus, researchers are focusing on developing new techniques and improving existing ones. Some of the methods that are invented for deepfake detection are discussed below.

Samuel Henrique Sliva et al. [11] have proposed a model that is based on the hierarchical grouping of different methods with attention-based data. This model is trained over the DFDC dataset. For obtained results, training and testing are performed over 100,000 videos. Even though the proposed architecture is for deepfake video detection, the process is performed over multiple frames (i.e. still images) from videos. The feature extraction is done by CNN methods

Xception and EfficientNet-b3. The weekly supervised attention-based data is augmented by zooming and ignoring the attention regions for the robustness of the detection model.

The model which combines the features of ResNext and LSTM is proposed by Vurimi Veera Venkata Naga Sai Vamsi et al. [12]. Here frame-by- frame-by-frame video is analyzed and then the facial part is cropped for further processing. After training and testing confusion matrix is obtained which helps in accuracy detection, in this model, accuracy is directly proportional to a number of input images. That means the model learns with each increase in frames. Celeb-DF v2 and Youtube videos have played important roles in the model's performance evolution.

The deepfake videos or images were get posted over the internet through social media platforms like Facebook, instagram, whatsapp, YouTube and so on. It is very crucial to develop some framework to identify forgery on these platforms. With this kind of idea, Abdullah Ayub Khan et al. [13] have developed a novel framework which was named DF-SCW. Pixels and neighbouring pixel values are analyzed and if any forgery is detected, then flags are applied to indicate risk. Kaggle's Deepfake Detection Challenges Dataset has been chosen for the training and testing of the model.

Ruben Tolsana et al. [14] have proposed the model, which is a collaboration of considering whole faces or considering only specific parts for forgery detection. UADFV, FaceForensics++, Celeb-DF V2 and DFDC datasets are used for performance analysis and evaluation of results. The eyes, nose, mouth and remaining parts are analyzed respectively for face analysis. For the further process of segregation between real and fake, the proposed model uses Xception, Capsule Network and DSP-FWA methods for deepfake detection. This model is mainly developed for identity swap deepfake detection.

Most research methods are developed for the detection of specific deepfake types. Abdul Qadir et al [15] have described the method that works on three types of deepfakes. This method is named as Resnet-Swish-BiLSTM. Even this method works over video forgery, that video is divided into frames. The proposed architecture applies DFDC and FF++ datasets. Facial landmarks are acquired by the face detector toolkit "OpenFace 2.0". The training videos are 70% real and 30% fake.

In 2021, Ahmed Sedik et al. [16] developed a multimodal architecture for deepfake detection. Both CNN and ConvLSTM methods are combined to detect fraud. The results obtained are good, but some drawbacks in intra-frame detection and frame forgery localization.

Zhiqing Guo et al. [17] have develop a image based model for deep manipulated features, which was named as an adaptive manipulation traces extraction network (AMTEN). It is efficient in pre-processing delicate features and slight manipulation in images. By combining the efficiency of AMTEN with CNN, they develop a detector known as ANTENet.

Face swap or identity swap is mostly found in deepfake. For the detection of this type, Yuval Nirkin et al. [18] proposed a model in a single image using remaining traces. In this approach, there are face identification and context recognition networks which identify face regions bounded by a tight semantic segmentation and facial parts, respectively. FaceForensics++, Celeb-DF-v2 and DFDC datasets are used.

In many researches, it is found that artifacts play an important role. But if there is an unknown artefact used of deepfake generation the performance decreases. Xiaoyi Dong et al [19] thought about Identity-Driven DeepFake Detection. In this approach image or video is taken as an input and it is matched with target identity. Instead of image artifacts it generally focuses on differentiating between suspected image/video with original identity of target. They developed OuterFace algorithm for deepfake detection.

Li Zhang et al [20] have developed deepfake based on swarm optimization technique. It consists of the combination of CNN-RNN, I3D and MC3 networks. The new version PSO algorithm is created adding numerical analysis based leader enhancement, Q learning based optimal search operation selection, petal helix search intensification and cross- reed elite signal generation

Daniel Xie et al [21] have enhanced AlexNet model by increasing six layer organizations. These six layers are for Convolution2d, max pooling, dense, flatten, activation and dropout layers. The experimentation is performed over commonly available datasets UADFV, FaceForensics++ and Celeb-DF.

Aminollah Khormali and Jiann-Shiun Yuan [22] has proposed Attention Based Deepfake Detection (ADD) model. In this model, face localization, preprocessing and localized discriminative feature extraction are the steps performed for deepfake detection. The Face Close-Up and Face Shut-Off data augmentation approach is employed. In the Face Close-Up approach, facial details observed. Whereas, in Second approach i.e Face Shut-Off approach ignores previous details and look for discriminative features from other parts. Celeb-DF(V2), WildDeepfake datasets are used in this model.

Shahroz Tariq et. al. [23] suggested three different learning models were used, namely single domain learning configuration, merge learning configuration and transfer learning configuration. These learning models were used for training and then they are tested across various deep fake detection methods (i.e. CNN+LSTM, DBiRNN, ShallowNet, Xception, MesoNet and CLRNet). CLRNet is proposed model which can capture real and deep fake frame based on convolutions inside a residual network, and inter-frame inconsistencies in real video and deep fake video using Convolutional LSTM Cell.

Three different deep fake detection techniques are proposed for training the model by Anubhav Jain et. al. [24], which includes typical binary classification approach, an attribution based approach, and attribution based on triplet-loss using Siamese network. These approaches are used across FaceForensic++ and Celeb-DF dataset are used.

Chih-Chung Hsu et. al. [25] have proposed novel deep forgery discriminator (DeepFD) based on embedding the contrastive loss. In this method many fake and real images are collected to learn contrastive loss based jointly discriminative features. Then classifier will concatenated to discriminative features for identifying fake images. CelebA dataset is used for creation of training samples. Also DCGAM, WGAP, LSGAN, PGGAN, WGAN-GP are some kinds of GANs that are used for fake images.

Nhu-Tai Do et. al. [26] has presented the GAN based face forensic detection method. For training Celeb-A is used for real faces and DC-GAN and PG-GAN are used for generating fake faces. Also these images help in training the model. Face Processing, deep feature extraction

and face matching are the main modules that plays important role in Deep face recognition system. After face extraction, the fine-tuning is used for fake face classification.

Shilpa Pant et. al suggests the deep fake detection using LSTM.[27]. In this model, EfficientNet B0 is used for feature extraction and vision transformer examines them. HOHA dataset is used for gathering real videos for training the model. CNN model is used for feature extraction and LSTM model for feature analysis. The genuineness of video can be determined by fully connected layer which is added after LSTM unit.

The CNN model is used for deep fake video detection by Muhammad Mussadiq Rafiqee et. al. [28]. Videos from Kaggle dataset repository are used for training and testing. Then videos are pre-processed which includes the extraction of facial image and region of interest. Also during pre-processing frames from videos are obtained for further process of real and fake identification. Then, CNN model is applied to that dataset for deep fake detection.

The ResNet architecture is presented by Sreeraj Ramachandran et. al. [29]. Also the model is tested over high quality deep fake videos that are gathered through Celeb-DF and FaceForensic++. ArcFace, Combined Margine, CosFace, SphereFace, SoftMax and Triplet loss are the six different loss functions that are used during the training of the model. Intra-class correspondence and Inter-class variety have used for efficiency.

Ipek Ganiyusufoglu et. al. [30] has proposed spatio- temporal model for generalized detection of deepfake videos. 3D CNNs can be used for development of such model that learns combined depiction of spatial and temporal features. FF++, DFDC and Deeper Forensics datasets are used in their research. The model is examined to identify whether it can detect the deepfake which was not included during training. For that the proposed method represented by 3D Convnets R3D and I3D are tested with XceptionNet, EfficientNet and RNN. It is found that R3D can detect unknown deepfakes.

III. Types of deep fake:

Deepfakes are appeared in people's attention as serious threat to cyber security and privacy of individuals. Recently large numbers of researchers are working over this trending topic. Various kinds of deepfakes are evolving with change in time. Mainly they can be categorized as follows:

Face Synthesis:

The Face synthesis is the process of creation of an artificial image of human face that doesn't exist in real life. [31] This types of synthesized deepfake can be used for creation of realistic looking animated objects in advertisements and movies. But at the same time it is evolve at extent to fool people to believe as real. Even such person doesn't exist naive people interacts with it. That means evil sources could use it as a way for digital forgery. Algorithms like StyleGAN are used for creation if such kinds of deepfake creation.[32]

Attribute Manipulation:

Attribute manipulation refers alteration of particular part of original image related to attribute that needs to change. [33] Instead of creating whole new structure as in face synthesis, attribute manipulation is easier to design.

In this process attributes like age, gender, hairstyle, eyebrow shape or facial parts can be modified. Otherwise in some situation glasses and tattoos can be altered. VGG16, SqueezeNet, DenseNet, ResNet are some techniques to identification of such attack. [34]

Identity Swap/ Face Swapping:

In this type, there is union of facial attributes, expressions of one person and remaining part of another person. It changes whole structure of genuine person, thus its name is identity swap or face swap [35].

Zao mobile application, FaceSwap and DeepFaceLab are some of the application through such kind of contents can be created. For creation of such content face detection and cropping is first step. Later intermediate condition extraction, creation of new face by mixing another face, matching created face into source video are some of the main tasks which are performed in this process.[36]

Face Reenactment / Puppet- Master:

The face reenactment involves manipulation of expressions of legitimate person in image or video with another person. The face reenactment is also called as puppet-master because, though the face of target remains unchanged. Still, the movements and utterance is controlled by the person behind the curtain (i.e prowler) [37]. In many cases around the world, it is generally used for spoiling the reputation of well-known personalities.

Ana Pantelić and Ana Gavrovska[38] have provided information about the Meso-4 model for puppet master deepfake detection in their research paper.

Lip-Syncing:

Deepfake can be created with Lip-Syncing by focusing on the lip movement of a legitimate person and by manipulating it to match fake audio. [39] This can be achieved by training the model over multiple videos of person and analyzing the position and movements of lips for a particular tone. This type of deepfake spreads rumours and misleading information among people.

IV. Deep Fake Detection methods:

Various methods were developed for the detection of deepfakes. Some Deepfake detection techniques are given as follows:

Visual Artifacts:

The visual artefacts generally include flaws in images or videos which are easy to identify. The deepfake generator sometimes leaves some clues which later on used as artifacts for detection.

Rough edges and improper position of pixels are some visual artifacts. Mostly synthetically created video content have non-natural pixel patterns around neck, mouth and borders. Also color inconsistencies and irregular lighting pattern are some of the commonly detected artifacts which could be found from media contents. [40][41]

Temporal features:

The temporal features are the features which are obtained through changes in elements with change in video sequence. Commonly identified temporal features include eye blinking patterns, improper mouth movements. Generally the deepfake videos generator could not create realistic hairs and also blinking movement is slow as compared to real blinking. [42] Similarly mouth movements in deepfake do not match with real audio. The mouth position for pronouncing each alphabet is different. Based on them inconsistencies can be obtained. Also facial expression is one of the source for temporal inconsistencies.[43] Zhihao Gu et al [44] have proposed deepfake detection model based on spatio-temporal inconsistency learning.

Handcrafted Features:

The handcrafted features satisfy its name “Handcrafted” as researchers can manually define the. Such characteristics have been found by identifying the distance between facial parts, inconsistent skin textures and by observing blinking patterns. For this technique, researchers need to have deep knowledge about their domain. Generally, handcrafted features are used over known deepfake types. [45][46]

Deep Feature:

The deep features are the intricate features which are too complex to be identified by the human eye. The neural networks process raw data and retrieve important deep features without human interference. The unnatural deep features later on are useful in deepfake detection.

Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM) are vital algorithms for deep feature extraction. Based on these algorithms many models have been developed for deepfake detection.

Xception and EfficientNet-B7 are CNN models. [47]

V. Analysis & Discussion:

Deepfake has become most challenging attack for security system. As incidents of deepfake attacks are increasing, various methods are also developing. Mostly researchers have created deepfake detection models based on deep learning algorithms, deep artifacts, handcrafted features and so on.

Generally, the performance of detection models decreases due to the low quality of video and unnecessary noise. Extra smooth or Extra sharp images/video could be synthetically created. But in the case of low-quality image/video chosen for testing, detection model could not identify the smooth features and artefacts.

Currently different ensemble models are trending. Ensemble models have developed by combining two or more different algorithms and methods. CNN, RNN and LSTM algorithms are used in such models along with analyzing facial attributes. Still full accuracy is not provided.

During the training and testing of models different datasets are used. Celeb-DF, Face Forensic++ and DFDC are commonly used datasets. Accuracy of different models goes high for them. But some detection models fails on higher version of same datasets. Also some detection models could not work properly for unknown datasets. All these aspects need to be focused during further research.

VI. Conclusion:

The rapid growth of technology has led to major threat of deepfake for everyone, from ordinary people to well known personalities. The misuse of deepfake technology is humiliating people and terrifying the public. In order to discover if a given video is real or fake various methods are created based on artificial intelligence, machine learning and deep learning.

In this paper the analysis is provided for many models for deepfake detection among which many techniques and algorithms are found to be useful. From this analysis, it is found that some models are tested across varied datasets for proper differentiation of real and fake content. Also models have been trained through multiple learning approaches. Many methods consume more time for detection of forgery. The models developed for the generalization of deepfake does not have good performance while detecting new form of forgery. As new deepfakes are building day by day, an efficient model based on generalization feature is needed to be developed.

References:

[1] Rachael Brooks, Yefeng Yuan, Yuhong Liu and Haiquan Chen, "DeepFake and its Enabling Techniques: A Review", APSIPA Transactions on Signal and Information Processing, 2022, Received 24 April 2022; Revised 23 June 2022, ISSN 2048-7703; DOI 10.1561/116.00000024, 2022.

[2] Teng Zhang, Lirui Deng, Liang Zhang, Xianglei Dang, "Deep Learning in Face Synthesis: A Survey on Deepfakes", 2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology, 978-1-7281-8811-9/20, ©2020 IEEE, 2020.

[3] S. Awah Buo, ''The emerging threats of deepfake attacks and countermeasures,'' 2020, arXiv:2012.07989.

[4] Yogesh Patel, Sudeep Tanwar, Rajesh Gupta, Pronaya Bhattacharya, Innocent Ewean Davidson, Royi Nyameko, Srinivas Aluvala, And Vrince Vimal, "Deepfake Generation and Detection: Case Study and Challenges", Digital Object Identifier 10.1109/ACCESS.2023.3342107.

[5] Sami Alanazi, Seemal Asif, and Irene Moulitsas, "Examining the Societal Impact and Legislative Requirements of Deepfake Technology: A Comprehensive Study", International Journal of Social Science and Humanity, Vol. 14, No. 2, 2024, DOI: 10.18178/ijssh.2024.14.2.1194, https://www.researchgate.net/publication/379615642, 2024.

[6] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales and Javier Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection", arXiv:2001.00179v3 [cs.CV] 18 Jun 2020.

[7] Fakhar Abbas, Araz Taeihagh, "Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence", Expert Systems With Applications 252 (2024) 124260, www.elsevier.com/locate/eswa, https://doi.org/10.1016/j.eswa.2024.124260, Accepted 14 May 2024, 2024.

[8] Rachel Selva Dhanaraj, M Sridevi, "Face Warping Deepfake Detection and Localization in a Digital Video using Transfer Learning Approach", Journal of Metaverse. 4 (1), 11-20. Doi: 10.57019/jmv.1338907, 2024.

[9] Yogesh Patel, Sudeep Tanwar, Rajesh Gupta, Pronaya Bhattacharya, Innocent Ewean Davidson, Royi Nyameko, Srinivas Aluvala, And Vrince Vimal, "Deepfake Generation and Detection: Case Study and Challenges", Digital Object Identifier 10.1109/ACCESS.2023.3342107, IEEE, Received 21 October 2023, accepted 4 December 2023, date of publication 12 December 2023.

[10] Alexander Godulla, Christian P. Hoffmann, & Daniel Seibert, "Dealing with deepfakes – an interdisciplinary examination of the state of research and implications for communication studies", https://doi.org/10.5771/2192-4007-2021-1-72, am 20.02.2022, 04:03:42, 2022.

[11] Samuel Henrique Silva, Mazal Bethany, Alexis Megan Votto, Ian Henry Scarff, Nicole Beebe, Peyman Najafirad, "Deepfake forensics analysis: An explainable hierarchical ensemble of weakly supervised models", Published by Elsevier B.V., https://doi.org/10.1016/j.fsisyn.2022.100217

[12] Vurimi Veera Venkata Naga Sai Vamsi, Sukanya S. Shet, Sodum Sai Mohan Reddy, Sharon S. Rose, Sona R. Shetty, S. Sathvika, Supriya M. S., Sahana P. Shankar, "Deepfake detection in digital media forensics", Publishing Services by Elsevier B.V., https://doi.org/10.1016/j.gltp.2022.04.017

[13] Abdullah Ayub Khan, Yen-Lin Chen, Fahima Hajjej, Aftab Ahmed Shaikh, Jing Yang, Chin Soon Ku, Lip Yee Por, "Digital forensics for the socio-cyber world (DF-SCW): A novel framework for deepfake multimedia investigation on social media platforms", Published by Elsevier BV, https://doi.org/10.1016/j.eij.2024.100502

[14] Ruben Tolosana, Sergio Romero-Tapiador, Ruben Vera-Rodriguez, Ester Gonzalez-Sosa, Julian Fierrez, "DeepFakes detection across generations: Analysis of facial regions, fusion, and performance evaluation", Published by Elsevier Ltd., https://doi.org/10.1016/j.engappai.2022.104673

[15] Abdul Qadir, Rabbia Mahum, Mohammed A. El-Meligy, Adham E. Ragab, Abdulmalik AlSalman, Muhammad Awais, "An efficient deepfake video detection using robust deep learning", Published by Elsevier Ltd., https://doi.org/10.1016/j.heliyon.2024.e25757

[16] Ahmed Sedik, Osama S. Faragallah, Hala S. El-sayed, Ghada M. El-Banby, Fathi E. Abd El-Samie, Ashraf A. M. Khalaf, Walid El-Shafai, "An efficient cybersecurity framework for facial video forensics detection based on multimodal deep learning", Neural Computing and Applications, Springer Nature 2021, https://doi.org/10.1007/s00521-021-06416-6

[17] Zhiqing Guo, Gaobo Yang, Jiyou Chen, Xingming Sun, "Fake face detection via adaptive manipulation traces extraction network", arXiv:2005.04945v2 [cs.CV] 16 Dec 2020.

[18] Yuval Nirkin, Lior Wolf, Yosi Keller, and Tal Hassner, "DeepFake Detection Based on Discrepancies Between Faces and their Context", arXiv:2008.12262v1 [cs.CV] 27 Aug 2020.

[19] Xiaoyi Dong, Jianmin Bao, Dongdong Chen, Weiming Zhang, Nenghai Yu, Dong Chen, Fang Wen, Baining Guo, "Identity-Driven DeepFake Detection", arXiv:2012.03930v1 [cs.CV] 7 Dec 2020.

[20] Li Zhang, Dezong Zhao, Chee Peng Lim, Houshyar Asadi, Haoqian Huang, Yonghong Yu, Rong Gao, "Video Deepfake classification using particle swarm optimization-based evolving ensemble models", Published by Elsevier B.V., https://doi.org/10.1016/j.knosys.2024.11141

[21] Daniel Xie, Prosenjit Chatterjee, Zhipeng Liu, Kaushik Roy, Edoh Kossi, "DeepFake Detection on Publicly Available Datasets using Modified AlexNet", 2020 IEEE Symposium Series on Computational Intelligence (SSCI), 978-1-7281-2547-3/20/ ©2020 IEEE

[22] Aminollah Khormali and Jiann-Shiun Yuan, "ADD: Attention-Based DeepFake Detection Approach", Big Data Cognitive Computing 2021, 5, 49., https:// doi.org/10.3390/bdcc5040049, 2021.

[23] Shahroz Tariq, Sangyup Lee, Simon S. Woo, "One Detector to Rule Them All Towards a General Deepfake Attack Detection Framework", arXiv:2105.00187v1 [cs.CV] 1 May 2021, 2021.

[24] Anubhav Jain, Pavel Korshunov, and S´ebastien Marcel, "Improving Generalization of Deepfake Detection by Training for Attribution", 978-1-6654-3288-7/21/ 2021 IEEE, 2021.

[25] Chih-Chung Hsu, Chia-Yen Lee, Yi-Xiu Zhuang, "Learning to Detect Fake Face Images in the Wild," 2018 International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 2018, pp. 388-391, doi: 10.1109/IS3C.2018.00104, 2018.

[26] Nhu-Tai Do, In-Seop Na, Soo-Hyung Kim," Forensics Face Detection From GANs Using Convolutional Neural Network", publication at: https://www.researchgate.net/-publication/327905310 , 2018.

[27] Shilpa Pant, Chhaya Gosavi, Sheetal Barekar, " Deep Fake Detection using LSTM and Survey of Deep Fake Creation Technologies", International Journal of Intelligent Systems And Applications In Engineering (IJISE), ISSN:2147-6799, Submitted: 28/09/2023 Revised: 14/11/2023 Accepted: 26/11/2023, 2023.

[28] Muhammad Mussadiq Rafiqee, Zahid Hussain Qaiser, Muhammad Fuzail, Naeem Aslam, and Muhammad Sajid Maqbool, " Implementation of Efficient Deep Fake Detection Technique on Videos Dataset Using Deep Learning Method ", Journal of Computing & Biomedical Informatics, https://doi.org/10.56979/501/2023, ISSN: 2710 – 1606, Volume 05 Issue 01, 2023.

[29] Sreeraj Ramachandran, Aakash Varma Nadimpalli, Ajita Rattani, " An Experimental Evaluation on Deepfake Detection using Deep Face Recognition", arXiv:2110.01640v1 [cs.CV] 4 Oct 2021.

[30] Ipek Ganiyusufoglu, L. Minh Ngˆo, Nedko Savov, Sezer Karaoglu, Theo Gevers, "Spatio-temporal Features for Generalized Detection of Deepfake Videos", arXiv:2010.11844v1 [cs.CV] 22 Oct 2020, 2020.

[31] Amal Naitali, Mohammed Ridouani, Fatima Salahdine and Naima Kaabouch, "Deepfake Attacks: Generation, Detection, Datasets, Challenges, and Research Directions", Computers 2023, 12, 216., https://doi.org/10.3390/ computers12100216.

[32] Joao C. Neves, Ruben Tolosana, Ruben Vera-Rodriguez, Vasco Lopes and Hugo Proenc¸"Real or Fake? Spoofing State-Of-The-Art Face Synthesis Detection Systems", JOURNAL OF LATEX CLASS FILES, VOL. 13, NO. 9, MARCH 2016.

[33] Xiangtao Meng, Li Wang, Shanqing Guo, Lei Ju, and Qingchuan Zhao, "AVA: Inconspicuous Attribute Variation-based Adversarial Attack bypassing DeepFake Detection", arXiv:2312.08675v1 [cs.CV] 14 Dec 2023.

[34] Zahid Akhtar, Murshida Rahman Mouree , Dipankar Dasgupta, "Utility of Deep Learning Features for Facial Attributes Manipulation Detection", In 2020 IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI) (pp. 55-60). IEEE.

[35] Walczyna, Tomasz, and Zbigniew Piotrowski, "Quick Overview of Face Swap Deep Fakes", Applied Sciences 13, no. 11: 6711. https://doi.org/10.3390/app13116711, 2023.

[36] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia, "An Introduction to Digital Face Manipulation", Handbook of Digital Face Manipulation and Detection, Advances in Computer Vision and Pattern Recognition, 2022. https://doi.org/10.1007/978-3-030-87664-7_1.

[37] Ana Pantelić and Ana Gavrovska, "From puppet-master creation to false detection", PROCEEDINGS, IX INTERNATIONAL CONFERENCE IcETRAN, Novi Pazar, Serbia, 6 - 9. june 2022. ISBN 978-86-7466-930-3, 2022.

[38] Ana Pantelić and Ana Gavrovska, "From puppet-master creation to false detection", PROCEEDINGS, IX INTERNATIONAL CONFERENCE IcETRAN, Novi Pazar, Serbia, 6 - 9. june 2022., ISBN 978-86-7466-930-3.

[39] Soumyya Kanti Datta, Shan Jia, Siwei Lyu, "EXPOSING LIP-SYNCING DEEPFAKES FROM MOUTH INCONSISTENCIES", arXiv: 2401.10113v1 [cs.CV] 18 Jan 2024, 2024.

[40] Jie Gao, Marco Micheletto, Giulia Orrù, Sara Concas, Xiaoyi Feng, Gian Luca Marcialis, Fabio Roli, "Texture and artifact decomposition for improving generalization in deep-learning-based deepfake detection", Engineering Applications of Artificial Intelligence 133 (2024) 108450, https://doi.org/10.1016/j.engappai.2024.108450.

[41] Falko Matern Christian Riess Marc Stamminger, "Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations", 2019 IEEE Winter Applications of Computer Vision Workshops, 978-1-7281-1392-0/19/ ©2019 IEEE, DOI 10.1109/WACVW.2019.00020

[42] Ashifur Rahman, Md. Mazharul Islam, Mohasina Jannat Moon, Tahera Tasnim, Nipo Siddique, Md. Shahiduzzaman, and Samsuddin Ahmed, "A Qualitative Survey on Deep Learning Based Deep Fake Video Creation and Detection Method", Australian Journal of Engineering and Innovative Technology, https://doi.org/10.34104/ajeit.022.013026, 2022.

[43] HaiweiWu, Jiantao Zhou, Shile Zhang, Jinyu Tian, "Exploring Spatial-Temporal Features for Deepfake Detection and Localization", arXiv:2210.15872v1 [cs.CV] 28 Oct 2022.

[44] Zhihao Gu, Yang Chen, Taiping Yao, Shouhong Ding, Jilin Li, Feiyue Huang, Lizhuang Ma, "Spatiotemporal Inconsistency Learning for DeepFake Video Detection", arXiv:2109.01860v3 [cs.CV] 11 Oct 2021.

[45] Jialiang Li, Haoyue Wang, Sheng Li, Zhenxing Qian, Xinpeng Zhang, Athanasios V. Vasilakos, "Are handcrafted filters helpful for attributing AI-generated images?", ACM ISBN 979-8-4007-0686-8/24/10 https://doi.org/10.1145/3664647.3680945.

[46] Ying Xu, Sule Yildirim Yayilgan, "When Handcrafted Features and Deep Features Meet Mismatched Training and Test Sets for Deepfake Detection", arXiv:2209.13289v1 [cs.CV] 27 Sep 2022.

[47] Gourab Naskar, Sk Mohiuddin, Samir Malakar, Erik Cuevas, Ram Sarkar, "Deepfake detection using deep feature stacking and meta-learning", https://doi.org/10.1016/-j.heliyon.2024.e25933, Published by Elsevier Ltd