International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 3, ISSN No. 2455-2143, Pages 229-235
Published Online July 2022 in IJEAST (http://www.ijeast.com)

DEEPFAKE DETECTION USING EYE-BLINKING PATTERN

Rithvika SanilM,S.Saathvik, Rithesh RaiK, Srinivas P M
Department of Computer Science and Engineering,
Sahyadri College of Engineering & Management, Adyar,
Mangaluru(575007), Karnataka,India

Abstract - **Deep learning algorithms have become so potent due to increased computing power that it is now relatively easy to produce human-like synthetic videos, sometimes known as & quot; deep fakes. & quot; It is simple to imagine scenarios in which these realistic face switched deep fakes are used to extort individuals, foment political unrest, and stage fake terrorist attacks. This paper provides a deep learning strategy novel for the efficient separation of fraudulent films produced by AI from actual ones. Automatically spotting replacement and recreation deep fakes is possible with our technology. To combat artificial intelligence, we are attempting to deploy artificial intelligence. The frame-level characteristics are extracted by our system using a Res-Next Convolution neural network, and later these features are applied to train an LSTM-based recurrent neural network to determine if those submitted video is being altered in any way or not, i.e. whether it is a deep fake or authentic video. We test our technique on a sizable quantity of balanced and mixed data sets created by combining the different accessible data sets, such as Face-Forensic++[1], Deep fake detection challenge[2], and Celeb-DF[3], in order to simulate real-time events and improve the model&#39;s performance on real-time data.**

Keywords **–Deepfake Detection , ResNext , LSTM.**

I. INTRODUCTION

Deepfake is an artificial intelligence-based method for synthesising human images. To create a video where the selected acts or says the same things as the source, already-existing photos and videos of the target person are combined and overlay over a video of the source person. The use of deepfake technology enables the generation of convincing but completely fake images. The use of deepfakes to impersonate politicians or celebrities might pose a severe risk to people&#39;s reputations and political stability. Most people identify deepfakes with false information. Deep fakes are very convincing and effectively persuade others to believe in an occurrence that never happened.

Multi-layer neural networks are used in the machine learning subfield known as deep learning. Deepfakes using machine learning and deep learning technologies. It is an important part of data science, including statistics and predictive modeling. Deep learning is currently being used in the most popular image recognition engines, natural language processing (NLP) and speech recognition software. Deep learning algorithms like a Generic Adversarial Network (GAN) that can be used to create deep games. The GAN can create an image that looks like a photograph of a human face, even if the face does not belong to any real person.

Machine learning&#39;s deep learning field makes use of multiple neural networks. deep learning technologies are used by Deep fakes. It is a crucial component of data science, which also includes statistical study and predictive modeling. The majority of widely used voice recognition, natural language processing (NLP), and picture recognition programmes currently employ deep

learning. Deep fakes may be produced using deep learning methods like a generative adversarial network (GAN). Although the faces in the photos don&#39;t actually represent any actual people, GANs may create images that resemble photographs of human faces.

A novel method to detect Deep fakes that utilizes an algorithm called Deep Vision and the generative adversarial network model examines a substantial difference in the blinking pattern. fluttering is a purposeful, unconscious motion. Human eye blinking patterns have been seen to vary greatly depending on a person&#39;s general physical health, cognitive function, biological makeup, and amount of information processing. For instance, the pattern can be influenced by a human&#39;s gender or age, the moment of day, one‘s mental reactions, or their state of consciousness. As a result, Deepfakes can be identified through data

integrity by using a heuristic based on the findings of health care, physiological, and brain engineering studies as well as machine learning and learning algorithm based on engineering and mathematical knowledge to supervise radical changes in the eye movement patterns of deepfakes. This implies that we may carry out integrity verification by detecting major changes in a subject&#39;s eye blinking pattern while they are being recorded. When eye blinks are repeatedly made in a very short amount of

time, In order to validate an anomaly in terms of frequency, repetitions, and duration, the recommended approach known as Deep Vision is applied.

II. LITERATURE SURVEY

David Guera et al., [1] In this study, a temporal-aware pipeline is suggested for automatically identifying deepfake films. Their approach collects frame level information using convolutional neural networks (CNNs). Then, to assess whether or not a video has been edited, a recurrent neural network (RNN) is trained using these traits. We compare our approach against a sizable sample of deepfake films gathered from various video sources. They demonstrate how, by using a simple architecture, our system can do this task with results that are competitive..

Yuezun Li et al ,[2] The effectiveness and quality of producing fake face videos with a realistic appearance have substantially improved thanks to recent advancements in deep generative networks In this study, they provide a unique method for revealing fraudulent face films produced by deep neural network models.. Our approach is based on identifying eye blinking, a physiological signal that is poorly displayed in artificial phoney videos, in the recordings. Our method is tested against benchmark datasets for eye- blinking detection and exhibits good results when it comes to identifying films produced by the DeepFake application, which is based on DNN.

Xin Yang et al,[3] In this article, we suggest a novel technique for exposing Deep Fakes, or AI-generated phoney face photos or videos. Our approach is subjected to the fact that Fake Videos are made through merging a synthetic countenance area with the real picture, producing faults in the process that can be seen when three dimensional head postures is calculated through the face photos. They do tests to demonstrate this occurrence and further develop a classification method based on this signal. An Support Vector Machine is assessed using a pair of genuine countenance photos and Fake Videos while applying this.

Yuezun Li et al,[4] In this research, we provide a deep learning method for effectively distinguishing fake AI-generated videos from real ones. Our approach is grounded on the knowledge that the DeepFake algorithm as it stands can only produce pictures with finite resolve, which must then to match the actual faces in the original clip, be more warped. We illustrate how such changes leave different artefacts in the resulting DeepFake films and how convolutional neural networks can successfully capture them (CNNs). Contrary to previous techniques, our method does not require DeepFake created photographs as negative training examples because we concentrate on the artefacts. Affine uses face warping to differentiate between authentic and fake images.. Our approach has two key benefits: (1) By performing basic image processing operations on a picture to create a negative example, such artefacts may be directly replicated. Our technique saves a tonne of time and money by eliminating the need to train a DeepFake model to produce negative instances; (2) since these artefacts are frequently present in DeepFake movies from many sources, our method is more reliable than others.

Ekraam Sabir and others, [5] Robust manipulation detection techniques are required because the dissemination of false information through realistic- looking but artificially created photos and videos has grown to be a serious issue. The identification of tampered faces in videos by utilising the temporal information in the stream has received less attention despite the focus on identifying face modification in still photos.A subset of deep learning models called recurrent convolutional models has excelled in using the temporal information from picture streams in a variety of applications. We then conduct extensive research to determine the optimum method for fusing these models underwent alterations with domain-specific face data pre - processing techniques, producing cutting-edge results on publicly available multimedia facial manipulation benchmarks.

Afchar, Darius, et al.,[6] The approach for detecting face tampering in videos is presented in this study, with an emphasis on Deepfake and Face2Face, two contemporary methods for producing faked films that look incredibly realistic. Due to the compression, which severely degrades the data, traditional image forensics techniques are typically not well suited to videos. In order to concentrate on the mesoscopic characteristics of pictures, this research uses a deep learning technique and provides two networks, each with a limited number of layers. On both an existing dataset and a dataset we created from web videos, we evaluate such rapid networks. With more than 98 percent for Deepfake and 95 percent for Face2Face, the tests show a very high detection rate.

Minha Kim et al,[7] With the development and availability of GAN-based video and image modification tools, efficient deepfake detection techniques are urgently needed.. Additionally, other deepfake generating methods have surfaced in recent years. Although several deepfake detection techniques have been presented, their effectiveness is hindered by emerging deepfake. By extracting 3 Transfer learning enables our student model to easily adapt to new deepfake types, doing away with the need for input datasets during domain adaptation. We show that, on the data augmentation task, FReTAL exceeds all benchmarks, obtaining accuracy rates of up to 86.97 percentage on poor deep fakes. by testing with Face Forensics++ datasets.

Davide Cozzolino.,[8] Now that synthetic picture synthesis and manipulation have advanced so quickly, there are serious questions about how this may affect society.At best, this results in a decline in confidence in digital material, but it may also have negative effects by disseminating misleading information or fake news. This study looks at how realistic modern picture modifications are and how challenging it is for people or machines to spot them. It provides a benchmark for face modification detection in

order to standardise the assessment of detection techniques.
On the basis of this information, we thoroughly examined data-driven fraud detection tools. We demonstrate the introduction of extra Even in the face of extreme compression, domain-specific information improves fraud detecting to virtually unknown levels of accuracy and greatly exceeds human observers.
Ammar Elhassan et al,[9] In this post, we use a cutting-edge technique to identify bogus films created using Deep Learning technology. The technique relies on using the mouth and teeth as differentiating characteristics, which are exceedingly challenging to do while fabricating movies. The suggested approach is more effective and accurate in identifying fraudulent videos. The research presented in this article builds on earlier work that introduced the key ideas by applying additional strategies for multiple transfer learning, Xception, CNN, Dense Net121, Dense Net169, Efficient NetB0, EfficientNetB7, InceptionV3, Mobile Net, MobileNetV2, ResNet50, and other networks to enhance the capacity to identify and classify deep fake films utilising characteristics obtained from the teeth and mouth as a biological signal..

Kandiga, S., Kini, U. N., Kini, U. K., & Mamatha, G,[10] This paper shows that the LBIQ takes into account only a single input which is an image and it gives out a set of attributes which can be further processed for better convenience.

Shetty, V., Vishwakarma, S., & Agrawal, A, [11].It uses inpainting algorithm which is proposed to create video synopsis.Simple processing of the video where frame segmentation and background replacement method is being followed and acts as a productive tool for indexing and browsing of the videos.

Airbail, H., Mamatha, G., Hedge, R. V., Sushmika, P. R., Kumari, R., & Sandeep, K [12].This study employs a method for categorising malware that depicts the threats as grayscale images. It uses a common photo highlight grouping technique.

Raja, K., Venkatesh, S., & Christoph Busch [13], R. B. to identify altered facial images from both digital and paper scans. This study suggests a novel method that uses the warped face picture database to fine-tune the initial fully connected layers of two D-CNNs at the feature level.

A, Harisha et al[14]. This research suggests using a Siamese Network to identify relationships between user-provided face photos. It has two convolutional neural networks, and the linked linear layers use the difference vector between them to detect if the input pictures are related to one another..

In [15], Falko Matern,Christian Riess et al. The problem of professional face editing in videos is important and undermines viewers' faith in the content. A number of face editing algorithms, however, exhibit artefacts that closely resemble well-known computer vision issues brought on by face tracking and editing when they are closely inspected. Thus, we query the difficulty of revealing false faces produced by current generators. To do this, we examine many unique processing artefacts as well as modern facial editing techniques. We also showhow visual artefacts that are really simple to use—like Deepfakes and Face2Face—can already be highly helpful in exposing such manipulations. The methods' reliance on visual attributes makes them easy to comprehend even for non-technical professionals.
In [16], Agarwal S et al. Thanks to recent advancements in machine learning and computer graphics, making convincing video and audio modifications is now easier than ever. These so-called deep-fake movies use a number of different approaches, such as full mouth and audio synthesis and replacement, complete full-face synthesis and replacement, and partial word-based audio and mouth synthesis and replacement (face-swap). It can be challenging to identify deep fakes with the least degree of spatial and temporal distortion. We describe a technique to identify such changed videos by making advantage of the fact that the dynamics of the mouth shape, or viseme, occasionally deviates from a spoken phoneme. The visemes associated with words containing the phonemes M (mama), B (baba), or P (papa), which require complete jaw closure to pronounce, are the focus of our attention.

In [17] , Chintha A et al. Digitally created audiovisual Deepfakes, or deceptive representations, are employed to delegitimize people and influence the public. Using a desktop computer fitted with an off-the-shelf graphics processing unit, an attacker may easily deceive a human observer thanks to the recent discovery of generative adversarial networks. To stop deepfakes from spreading, journalists, social media platforms, and the general public must all be able to identify them. In this paper, we provide very successful digital forensic methods for spotting audio spoofs and visual deepfakes. The latent representations for both audio and video have been meticulously built in order to make it easier to extract semantically rich information from the recordings.We use the FaceForensics++ and Celeb-DF video datasets as well as the ASVSpoof 2019 Logical Access audio datasets to illustrate our methodologies and create new benchmarks globally..

In [18] , Li L Bao et al. In this work, they introduce the face Xray image representation in order to recognise false faces in pictures. A greyscale image known as the face X-ray of an input face image is used to assess if the image may be divided into the blending of two distinct photographs from different sources. This is accomplished by merging the border around a fake photo and revealing no barrier around a real one. We observe that most of contemporary Fusing the altered face with the original is a common step in face modification procedures. a background image Face X-ray provides a trustworthy way to recognise possible hazards as a result. the bulk of face modification algorithms in use today. Indeed, Without employing fake images made using any of the most advanced methods for face alteration, the face X-ray algorithm may be trained. Numerous tests show that using Face X-ray on forgeries created by unseen faces still works.

H. H. Nguyen et al [19]. A capsule network is used in this study's methodology to identify several spoofs, including replay assaults that use printed pictures or recorded movies and computer-generated videos that employ deep convolutional neural networks..

Tariq, S., Lee, S., & Woo, S.S[20]. This book uses a video's sequence of frames as input to a transfer learning-based method and a convolutional LSTM-based residual network (CLRNet) to detect unnatural-looking artefacts in deep-fake movies. really helpful for modern deepfake detection.

III. PROPOSED TECHNIQUE

Although there are several tools for developing DeepFake , there are hardly any methods for their detection. With the aid of our techniques for discovering it, we would contribute immensely to halting their online proliferation. We could provide clients with an internet software platform so they may submit films and designate them as fake or real. From creating a content infrastructure, this initiative may be expanded to include the creation of a browser plugin for automatic Fake visual detections. Particularly big apps like Facebook and Instagram frequently integrate this study into their systems for a rapid preview of this technology prior to being able to distribute to another subscriber. Our model aims to identify all sorts of DeepFake, such as interpersonal Technology, substitution DeepFake, and retrenchment fake Video.

The suggested system & #39; s basic system architecture is shown in figure:

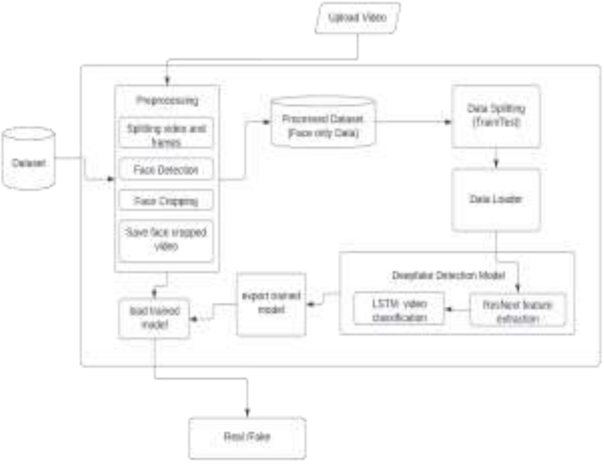


figure1: System Architecture

A. Dataset:

We are leveraging a composite dataset composed of an equal number of motion pictures from various collection sources, such as YouTube, Face Forensics++, and the Deep fake recognition challenge dataset. Our freshly developed dataset contains halves of the actual footage and halves of such changed photo shopped footage. Seventy percent of the sample is the train set, and 1/3 of the dataset is the test set.

B. Preprocessing:

The division of the clip into pixels is an aspect of dataset preparation. Then, a face is spotted, and the frame is cropped to include the face. In order to maintain consistency across the series of frames, the average of the recorded video is assessed, and a changes that were occuring facial expression focused collection is created using the frames that match the average. During pre-processing stage, the frames without faces are disregarded It would take a lot of system resources to process the 300 total pixels of the Ten seconds motion image at Thirty pixels per sec. Therefore, for experimental purposes, we advise training the model utilising just the first Hundred pixels.

C. Model:

The system is comprised of one LSTM layer followed byresnext50 32x4d. The preprocessed face-cropped clipsare imported by the data loader, which further divides the clips into a train set and a test set. Additionally, the system captures pixels from the edited films in small batches during ongoing training and testing.

D.Res Next CNN for Feature Extraction:

For extracting the features and precisely determining the frame level properties, we suggest using the ResNext CNN classifier instead of developing the classifier from scratch. In order to appropriately optimize the gradient descent of the model, we then will fine-tune the infrastructure by selecting an appropriate learning rate and adding any additional necessary layers. Following the pooling layers, 2048-dimensional feature vectors are supplied into the sequential LSTM as input.

E. Sequence Processing

Consider a two-node neurological networks using a sequence of ResNext CNN derived characteristics of input pixels as input and the likelihoods that the pattern is a deep fake video or an unmodified clips..We suggest using a 2048 LSTM unit with a 0.4 probability of failure for this issue, because it can help us accomplish our goal. Sequential analysis of the pixels using LSTM is used to do a time series assessment of the video is performed by contrasting the pixels at seconds "t" and "t-n" seconds. where any number of pixels before t can be represented by n.

F.Predict:

A sample clip is accepted by the trained algorithm for prediction. A brand-new video is standardized to incorporate the style of training sample. The footage is divided into pixels , faces are cropped, and instead of keeping the photo shopped pixels locally, they are immediately forwarded to training sample for spotting.

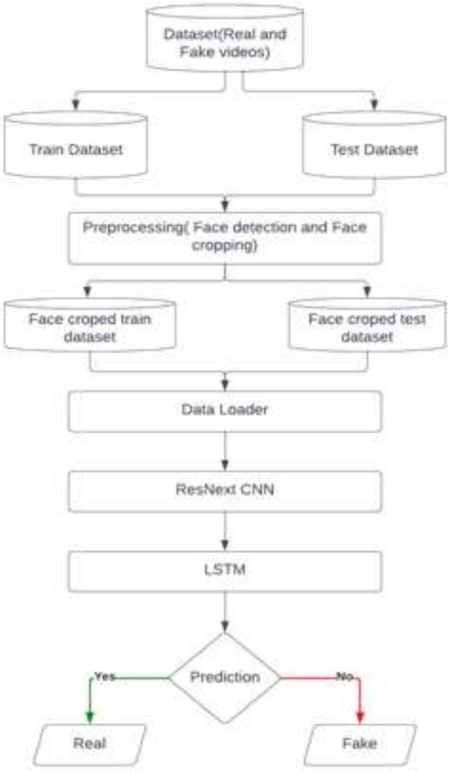
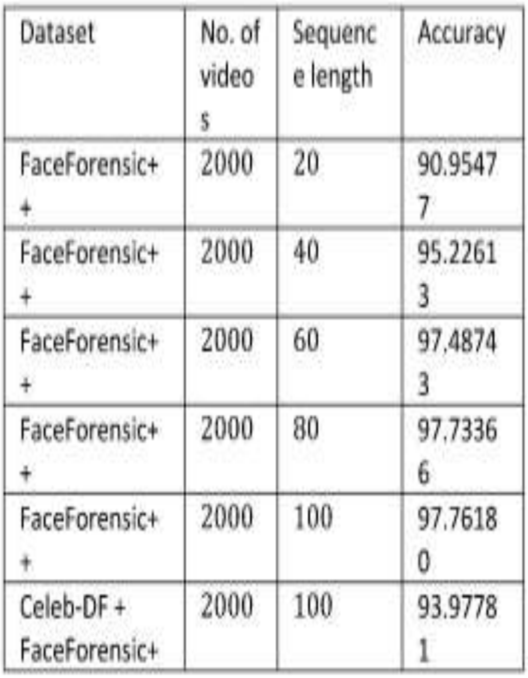

Dataset(Real and Fake videos)
Train Dataset
Test Dataset
Preprocessing(Face detection and Face cropping)
Face croped train dataset
Face croped test dataset
Data Loader
ResNext CNN
LSTM
Prediction
Yes
No
Real
Fake

Fig.2:Workflow

IV. RESULTS AND DISCUSSIONS



Dataset	No. of videos	Sequence length	Accuracy
FaceForensic++	2000	20	90.95477
FaceForensic++	2000	40	95.22613
FaceForensic++	2000	60	97.48743
FaceForensic++	2000	80	97.73366
FaceForensic++	2000	100	97.76180
Celeb-DF + FaceForensic+	2000	100	93.97781

We provided a neural network-based method for determining if a video is a deep fake or the real thing, along with the model’s level of confidence. Our technique can accurately anticipate the result after analysing one frame of video (10 frames per second). We built the model with an LSTM for temporal sequence processing to identify changes between the t and t-1 frame and a pre-trained ResNext CNN model to extract frame-level features. We can process movies with 10, 20, 40, 60, 80, and 100 frames per second using our model.

V. CONCLUSION

A neural network-based approach for classifying the video as deep fake or real was given, including the recommended model&#39;s level of confidence. The Deepfakes generated by Generative adversarial models with the assistance of Auto encoders serve as the model for the proposed methods. With the support of ResNext CNN and LSTM, our methodology classifies clips and detects pixels at the frame level. In accordance with the parameters mentioned in the study, the proposed methodology may detect if a video is a deep fake or real. We think it&#39;ll offer real-time data with an extremely high degree of precision.

VI. ACKNOWLEDGEMENT

This project has been supported by VGST, Department of ITBT & ST, Government of Karnataka. Project has been developed as a part of Project CoE Digital Forensics Intelligence GRD 853.

VII. REFERENCES

[1]. Güera, D., & 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.

[2]. Li, Y., Chang, M. C., & Lyu, S. (2018, December). In ictu oculi: Exposing ai created fake videos by detecting eye blinking. In 2018 IEEE International workshop on information forensics and security (WIFS) (pp. 1-7). IEEE.

[3]. Yang, X., Li, Y., & Lyu, S. (2019, May). Exposing deep fakes using inconsistent head poses. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8261-8265). IEEE.

[4]. Li, Y., & Lyu, S. (2018). Exposing deepfake videos by detecting face warping artifacts. arXiv preprint arXiv:1811.00656.

[5]. Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., & Natarajan, P. (2019). Recurrent convolutional strategies for face manipulation detection in videos. Interfaces (GUI), 3(1), 80-87.

[6]. Afchar, D., Nozick, V., Yamagishi, J., & 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.

[7]. Kim, M., Tariq, S., & Woo, S. S. (2021). Fretal: Generalizing deepfake detection using knowledge distillation and representation learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 1001-1012).

[8]. 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).

[9]. Elhassan, A., Al-Fawa'reh, M., Jafar, M. T., Ababneh, M., & Jafar, S. T. Enhanced Deepfake Detection Using Mouth Movement and Transfer Learning. Available at SSRN 3979595

[10]. Kandiga, S., Kini, U. N., Kini, U. K., & Mamatha, G. (2020, August). Image Processing and Location based Image Querier (LBIQ). In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 856-861). IEEE .

[11]. Shetty, V., Vishwakarma, S., & Agrawal, A. (2017, May). Design and implementation of video synopsis using online video inpainting. In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (pp. 1208-1212). IEEE.

[12]. Airbail, H., Mamatha, G., Hedge, R. V., Sushmika, P. R., Kumari, R., & Sandeep, K. (2021). Deep learning-based approach for malware classification. International Journal of Intelligent Defence Support Systems, 6(2), 61-80

[13]. Raghavendra, R., Raja, K., Venkatesh, S., & Busch, C. (2017, October). Face morphing versus face averaging: Vulnerability and detection. In 2017 IEEE International Joint Conference on Biometrics (IJCB) (pp. 555-563). IEEE..

[14]. A, Harisha et al. (2022), A Performance Evaluation of Convolution Neural Networks for Kinship Discernment: An Application in Digital Forensics'. Intelligent Decision Technologies, vol. 16, no. 2, pp. 379-386 DOI: 10.3233/IDT-210132.

[15]. Matern, F., Riess, C., & 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.

[16]. Agarwal, S., Farid, H., Fried, O., & 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).

[17]. Chintha, A., Thai, B., Sohrawardi, S. J., Bhatt, K., Hickerson, A., Wright, M., & Ptucha, R. (2020). Recurrent convolutional structures for audio spoof and video deepfake detection. IEEE Journal of Selected Topics in Signal Processing, 14(5), 1024-1037.

[18]. 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).

[19]. Nguyen, H. H., Yamagishi, J., & Echizen, I. (2019, May). Capsule-forensics: Using capsule networks to detect forged images and videos. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 2307-2311). IEEE.

[20]. Tariq, S., Lee, S., & Woo, S. S. (2020). A convolutional LSTM based residual network for deepfake video detection. arXiv preprint arXiv:2009.07480.