Proceedings of the International Conference on Sustainable Computing and Smart Systems (ICSCSS 2023)
IEEE Xplore Part Number: CFP23DJ3-ART; ISBN: 979-8-3503-3360-2

Deepfake: An Endanger to Cyber Security

1st Sudhakar K. N
Associate Professor
Dept. of ISE, CMR Institute of Technology, Bengaluru
sudhakar.kn@cmrit.ac.in

2nd Shanthi M.B
Associate Professor
Dept. of CSE, CMR Institute of Technology, Bengaluru
gowri.pc@gmail.com

2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) | 979-8-3503-3360-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSCSS57650.2023.10169246

***Abstract*—'Deepfake' got originated from the technology 'deep learning' working behind it and the type of information it generates 'fake' after manipulating the original information. It's an AI-based innovation used to make counterfeit recordings and sound that look and sound genuine. The inventions and implication of digital technologies in all spheres of mankind, has posed a great challenge to mankind to come up with secure solutions against a digital problem called Deepfake resulting from the application of deep learning thereby compromising authentication or originality. The said technology creates digital images and videos all new and totally fake. But the consequence it creates on society is totally a negative impact. With the aid of AI in developing hyper realistic videos, Deepfake has extended its giant wings to harm societal health and resulting in a critical challenge against the authenticity of the source. The Internet has become the platform to deliver these Deepfakes to unlimited destinations within no time. There are lot many researches have been carried on how to detect this deep fakes. Most of the research works have used deep learning models like Convolution Neural Network (CNN) for analyzing the convolution traces in deepfakes. Some of the research works have used Recurrent Neural Networks (RNN) by combining the Long Short-Term Memory (LSTM) with Blockchain. This research study has presented the comprehensive literature study, which highlights the various approaches used in generation and detection of deepfakes.**

***Index Terms*—Deepfake, Deep Learning, AI, Authenticity, Counter Techniques Algorithm, Digital Era**

I. Introduction

Deepfake is an intelligent technology used to generate false videos, images, or audio that looks and sounds real. An artificial Intelligence based technology called Deep Learning is used to manipulate the original content and to modify it as per requirement. The inventions and implications of digital technologies in all spheres of mankind have posed a great challenge to mankind to come up with secure solutions against a digital problem called Deepfake resulting from the application of deep learning thereby compromising authentication or originality. The said technology creates digital images, and videos in an all-new form and totally fake. But the consequence it creates on society is totally a negative impact. With the aid of AI in developing hyper-realistic videos, Deepfake has extended its giant wings to harm societal health and resulting in a critical challenge against the authenticity of the source [1]. The Internet has become the platform to deliver these Deepfakes to unlimited destinations within no time. The lack of techniques involved in verifying the received information at the receiver's end has made the receiver believe the information genuine. Social media for sharing information has made this information reach millions of people with a single click. The lack of awareness has made the common man believe the received data as in the received form. Thus, Deepfake has become the biggest threat to society. This must be combated through legislation and regulation, by introducing new corporate approaches, voluntary activity for authentication, and appropriate education and training by spreading awareness among people about Deepfake detection, and content authentication to prevent the spread of Deepfake. The research in digital technologies has made the world move forward with the digitisation of every field. This has made human life much simpler and easier. Every sphere of modern life is influenced by some or another flavor of digital access to all kinds of required resources. Starting from child games, education, work environments, communication, purchases, so forth and so on. In our day-to-day life we have strong connectivity with digital widgets. This has led us to face challenges in multitude. Deepfakes have taken active entries in threatening the security domains. It has awakened the world to grow stronger with the security aspects and have a strong defense against any kind of security threat. Deep fakes portray humans doing activities that they have never committed or expressed in their speech, just by ingesting thousands of models to fake target content by training them using Deepfake algorithms. The confirmation bias of the human being may be a tendency of a formed opinion to confirm our prejudice. The situations which are not perceived subjectively lead to misjudgment and specious information. Deepfake, therefore has become a greatest threat to the society. Current technologies in combating against the deepfake are limited to detect the fake content. There is still ample room for the inventions for alleviating the deepfake generation and protecting the originality of the source content.

The development of deep generative networks in the creation of deepfakes in different medias, have posed a challenge to the researchers to work on new technologies to detect the deepfakes. Many researchers have done intensive research on creation as well as detection of deepfakes.

II. Deepfake Creation

Because of how quickly rumors, conspiracies, and false information are spread on social media platforms and how consumers prefer to follow the crowd, deepfake brands are among the most common. The Reddit community released the

first Deepfake in 2017 on the internet. A class of algorithms known as auto-encoders, a subset of Artificial Neural Network (ANN) ideas have been used to alter films. They are employed in the unsupervised learning of effective data coding. A source face from one video is superimposed onto a target face from another in a face swap video. They later discovered that improvements in machine (deep) learning can be used to produce better Deepfakes [1]. From a technological standpoint, Deepfake is a result of the combination of two ANNs using generative adversarial networks (GAN). Figure 1 depicts how the unit and the discriminator unit are used in the manipulation of original content. The 'generator' attempts to make new samples that are good enough to deceive the 'discriminator', it determines the status of new media whether it looks real or fake. GAN looks into a considerable number of images of a human subject and generates fresh portraits that are approximately the same without being an accurate reflection of prior images. Initially, Machine Learning (ML) model would be fed with a considerable number of target images for pre-processing, by which the model can learn and recognize the print of the person's face. After training the model with an immense amount of training data, Deepfake algorithms can predict the appearance and the person's face while mimicking others. As GAN models will be trained by feeding thousands of images, generated videos can easily falsify the subject's identity.

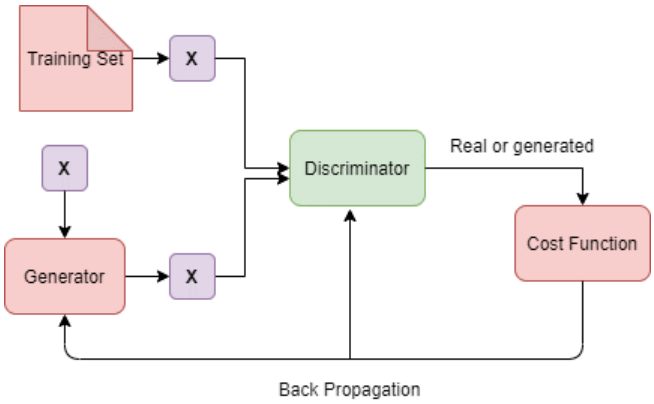

Fig. 1. GAN Model of Deep Fakes Creation

In each and every innovation, counter innovation need to be worked out to neutralize its negative impacts. In this regard, there is an immediate requirement for the implementation of Deepfake detectors. The latest updates related to research in Deepfake detection is mainly based on two major features of observation, that is either based on the hidden biological signals in the manipulated content or the pixel-level manipulations to differentiate the original content from the manipulated one.

Li, Y., and Lyu, S in [7], have implemented an AI-based deep fake generation algorithm to generate images with limited resolutions. These images are additionally distorted to coordinate them to the first essences of the source video. These changes leave unmistakable antiquities in the produced Deepfake recordings. Authors have developed an effective approach to capture these artifacts as the identities for detecting Deepfakes. They have used CNN for the detection of such footprints left in the manipulated images to detect the Deepfakes. Figure 6 depicts the production pipeline of Deepfakes. Once the residual signals of Deepfake are identified, they applied Boundary Smoothing techniques for generating the final version of Deepfake. In research work [9], authors have discussed about how to synthesize lip-sync Deepfakes. They have mainly discussed how to use Audio-to-Video and Text-to-Video synthesis techniques for generating the lip-sync Deepfakes. In the Audio-to-Video synthesis technique, it is supplied with the target video of the person speaking, and the audio record of the target speech. It generates the output video having the target video manipulated to give the supplied audio as the manipulated speech. In the Text-to Video generation technique, two inputs have the target video of manipulation and the target text to be manipulated. The output generated is a new falsified video having the words spoken from the target text content.

III. Detection of Deepfake

Deepfake finders indiscriminately using profound learning are not viable in getting phony substance, as generative models produce impressively reasonable outcomes. Hence, detection of Deepfakes needs to be done based on granular level changes in the fake data. The authors in [3], [5], [7] have used biological signals hidden in manipulated videos, as the descriptors to detect the Deepfake. The heartbeat signals are used to discover the fake content in the portrait videos. The recent research works found that Deepfake videos leave behind unique biological noise signals called as Deepfake heartbeats. Approaches used for detection involves identifying different spots in the falsified person's face and these cells are called as Photo Plethysmogram (PPG) cells. Some of the researches have given focus on the coordinated movement between the eye and neck in the falsified video to detect the fake content in it. Deepfake techniques typically require a huge number of pictures and video information to prepare models to make realistic images and recordings.

Yue zun Li et al, in [2], have implemented a Deepfake video detection model. The model is based on facial change and the eye blinking rate of the person in the original video footage. In their model, they have aligned the faces from each video frame to an equivalent reference frame and used the face movements and the orientation for detecting the Deepfake. From each reference frame, they have extracted the region of eye to form a stable sequence. By using Long-term Recurrent Convolution Networks (LRCN), they have captured the temporal dependencies across the sequence of frames during the blink of eyes.

LRCN model applies feature extraction as the first step. This is carried out to convert the input eye region into discriminative features. The Author utilized VGG16

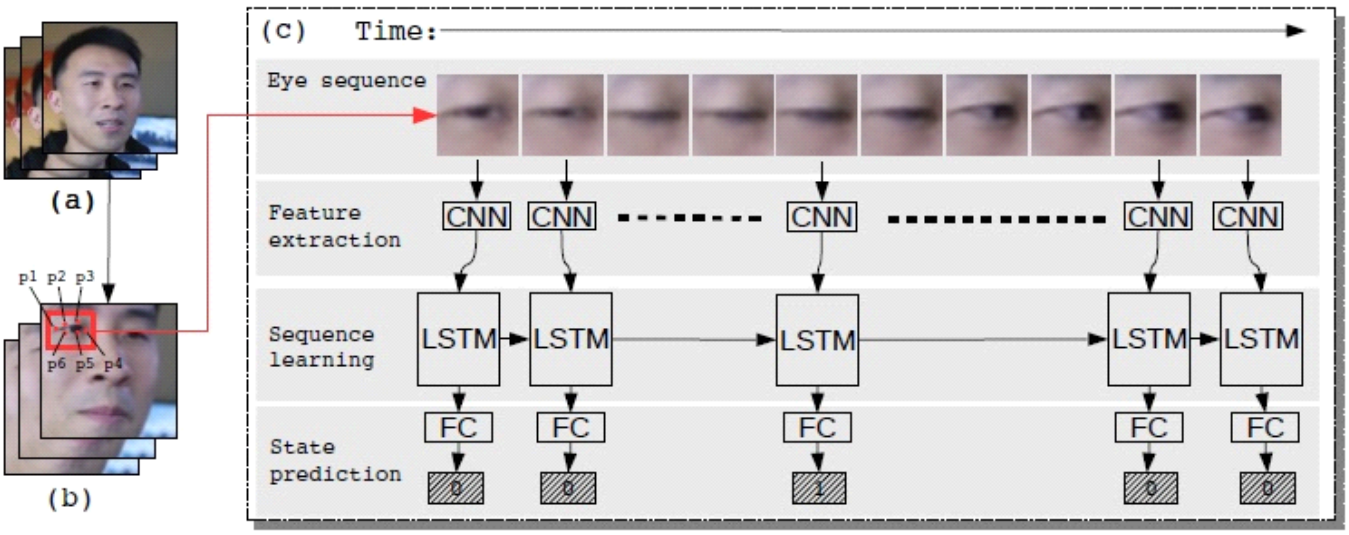

Fig. 2. Long-term Recurrent Convolution Networks capturing temporal dependencies.

framework [3] to achieve the mentioned results. The feature extraction output is fetched into sequence learning, which is in turn implemented with a Long Short Term Memory (LSTM) cells [4] and Recursive Neural Network (RNN).

Tian Qi Chen et al. [11] has proposed, Neural Ordinary Differential Equation (Neural- ODE). The proposed equation can detect Deepfake using the hidden heartbeat signals in falsified videos. In the proposed equation, they have integrated traditional numerically stable forward simulation and theory of differential equations. In the proposed model, each time series is represented by a latent trajectory. The variation auto encoder [12] was used to train latent model. The Authors experimented on the fake videos generated by an online portal deepfakeweb.com. The heartbeat of the actual video is estimated using Neural Ordinary Differential Equation (N-ODE) model. Later, the heartbeat of the fake video is estimated and compared with the original one. The difference in the values is mainly used for the detection of Deepfakes. Authors have used original videos taken from COHFACE which are made available at VIDTIMIT database and the generated Deepfake videos from the online portal deepfakeweb.com. The figure 3 depicts the observed Neuro-ODE predictions on Deepfake videos and the original videos taken from VIDTIMIT database.

In the domain for Deepfake, face manipulation in the videos is one of the critical and primary aspect to consider. Researchers had mainly focused on Face manipulation and identification of tampered faces in videos based on the temporal information hidden in the video streams. Sabir el al, [5], authors have used the combination of the recurrent convolution model and face alignment approach for experimentation to obtain the better performance in detection of Deepfakes. They have mainly used tools like Face2Face and FaceSwap for tampering faces in different video streams. Guera et al, [6], have proposed a temporal pipeline for the detection of Deepfakes in videos. Feature extraction has been carried out using CNN and extracted features are used to train RNN for detecting the Deepfake. Figure 5 shows the model used for the detection of Deepfake.

In the research work [10], Chintha A et al, Authors have introduced an effective digital forensic method for detecting the Deepfake audio spoof and visual Deepfake. They have designed a combined method of bidirectional recurrent structures and entropy function. They have chosen latent representations of audio and the visuals, to extract semantic information from audio and the video visuals. Extracted information is fed into the recurrent network to detect the temporal and the spatial signature of hidden Deepfakes. Demonstrated the working of the methods using FaceForensics++ and Celeb-DF video datasets.

Wang, L et al, in [11] proposed an approach called FakeSpotter, where he has monitored the neuron behavior at every layer to capture the ultra-fine features for detecting fake faces. The captured features are then amplified for differentiating the fake and the original faces. In the experiment, they have used GAN synthesized images for detection purpose. figure 5 depicts the overview of FakeSpotter using layer-wise neuron behavior as features to a binary classifier unit.

FakeSpotter has used the features extracted from each layer, rather getting it from the last layer for monitoring super grained behavior patterns for detecting the Deepfakes in videos.

F. Marra et al, [13] have proposed an incremental learning model for detection of Deepfakes in the GAN generated models. They have constructed an incremental model based on the object classification approach. They have used images generated from various GAN and the experimental results have shown considerable improvement without compromising on

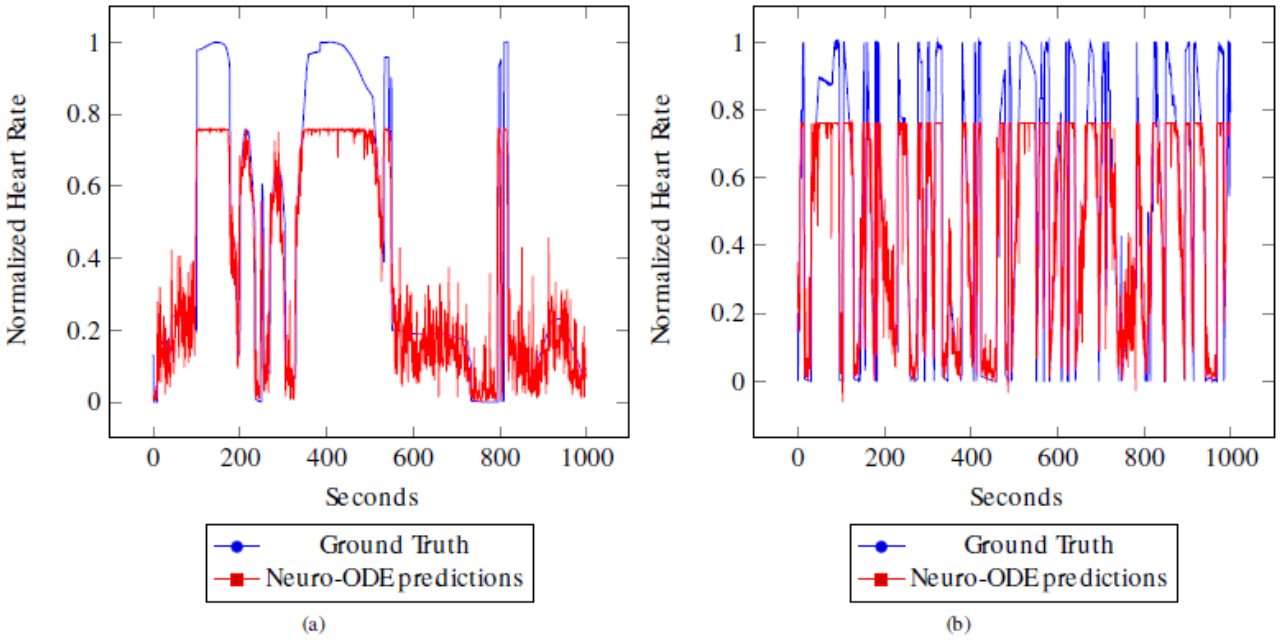

Fig. 3. The Min-max standardized heart rate acquired from skin shading variety (ground truth) and Neural-ODE (predictions) on: (a) Deepfake videos (b) DeepfakeTIMI database videos

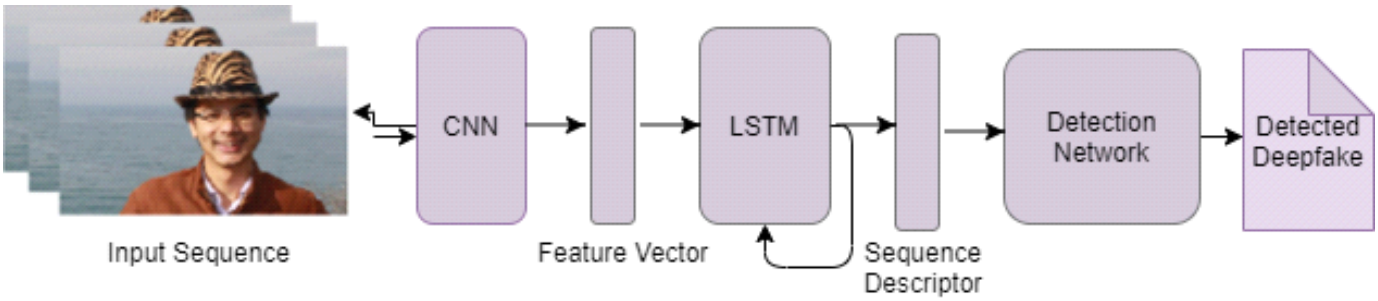

Fig. 4. Deepfake detection using the recurrent convolution model

the performance.

Amerini et al, [14] Authors proposed a method which exploits the optical flow field dissimilarities between original and fake videos. Authors have given the main focus on detecting the anomalies in the temporal dimension of the sequence. They have used motion vectors by representing as three-channel images as input to the neural network for processing. They have used FaceForensics++ dataset in their experiments. The results were found to be promising the field of video Deepfake detection. Table I gives the summary of the various methods and mechanisms discussed for the detection of Deepfakes.

IV. Pixel Level Irregularities

There is a greater breadth of research that extracts face features and uses different forms of deep learning to target intra-frame or inter-frame inconsistencies [11] [15]. While many of these methods perform well on specific types of manipulations, they fall short of being able to generalize several unknown types of Deepfakes, which is critical for the community.

Authors Zhang H [8]. et al. have introduced Conditioning Augmentation fin [7] in addition to Stacked Generative Adversarial Networks (StackGAN), for synthesizing realistic images. In the proposed technique, it breaks down the content to picture blend to a totally unique sketch-refinement measure. In the first stage of the image synthesis process, StackGAN sketches the basic color and shape of the image and in the stage-2; GAN corrects identified defects from the first image and adds more details about the quality of the image. Authors research work have specified that the algorithm called Facial Reenactment Manipulation (FAM) build with in Cyabra's detection tool can differentiate between an original image and a fake image created with GAN. This method utilizes many encoders to extract considerable amount of data from the target video. The mechanism opted extract as many parameters as possible out of the video, instead of just relying on frame by frame. Cyabra's Deepfake detection technology succeeded to achieve 91% fake Videos and Images detection in terms of accuracy.

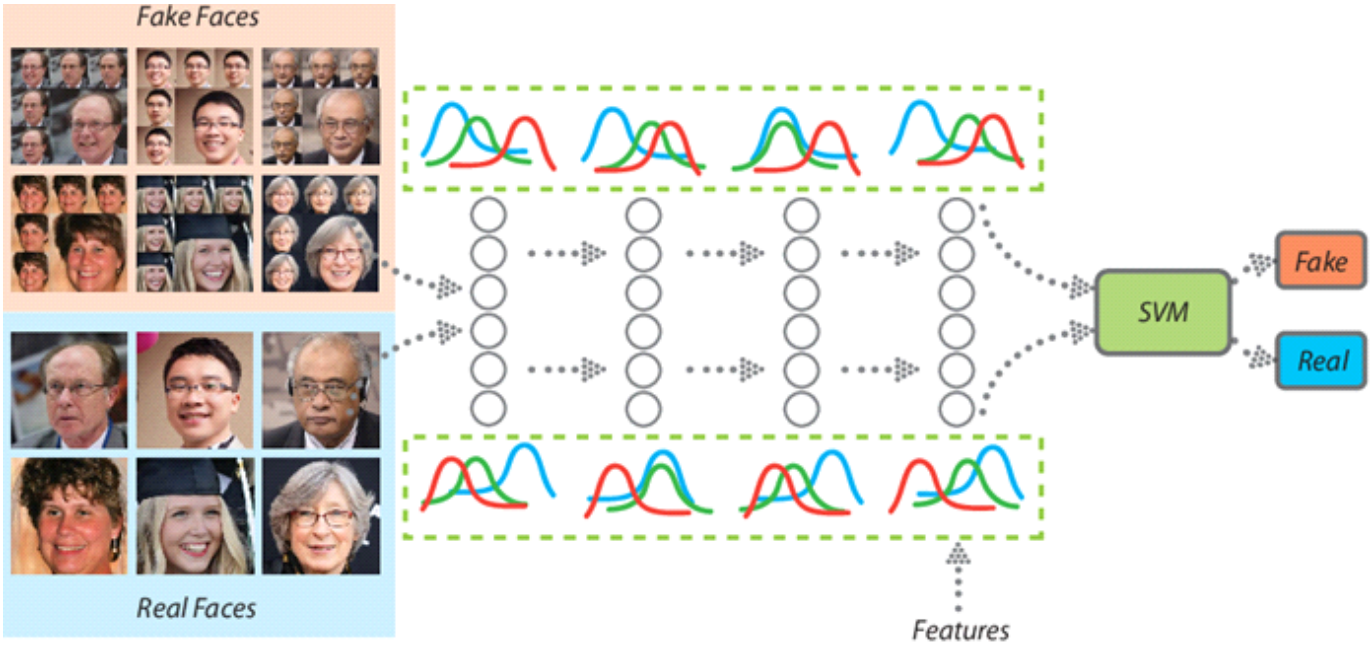

Fig. 5. FakeSpotter using layer-wise neuron behavior as features to a binary classifier unit

TABLE I
THE VARIOUS METHODS AND MECHANISMS DISCUSSED FOR THE DETECTION OF DEEPFAKES

Method	Dataset Used	Classifiers	Key Features
Temporal Inconsistencies in Video frames [6] [7]	Videos from different public websites	CNN LSTM	CNN – Extraction of Frame level features LSTM – Sequence descriptor for classification
Face-warping artifacts [15]	FaceForensics, Face2Face	CNN	MesoInception-4 and Meso-4 are introduced to examine Deepfake videos
Analysis of convolution traces [10]	GDWCT, StyleGAN, AttGAN, Style-GAN2 StarGAN	K-nearest neighbors, SVM, Linear Discriminate Analysis	Expectation Maximization Algorithm - GAN-Based image Deepfake generators
Spatio-temporal features [5]	FaceForensics	RCN	RCN – Temporal discrepancies across frames and the gated recurrent unit cells
Phoneme-Viseme mismatches [10] [11]	Four-in-the-wild lip-sync Deepfakes from Instagram and YouTube, Audio-to-Video (A2V), Test-to-Video (T2V)	CNN	Visemes - Explore the dissimilarities between the dynamics of the mouth shape Phonemes – Complete mouth closure is required or else Deepfakes can incorrectly synthesizeit.
Facial texture, Eye, and teeth [4] [8]	Video data set downloaded from YouTube	Logistic regression and neural networks	Exploit facial texture differences, and missing reflections in eye and teeth areas of Deepfakes
GAN-Pipeline Features [5]	Public dataset FaceForensics++, DFDC2, and Celeb-DF	SVM	InterFaceGAN, StyleGAN
GAN-Pipeline Features [5] [1]	CELEBA, RaFD dataset	KNN, SVM, LDA	GAN – Deepfake detection
Audio spoof detection [13]	FaceForensics++ and Celeb-DF video datasets	RCN	Digital forensic – Audio spoof and the visual Deepfake detection
Detecting GAN-Generated Fake Images [4] [5]	CycleGAN, StarGAN	CNN	Detection of GAN image- computation of co-occurrence matrices on the RGB channels of an image
Incremental Learning for GAN-generated images [13]	GDWCT, StyleGAN, AttGAN, Style-GAN2 StarGAN	CNN	GAN Image Classification

Current research in this field has given a focus on, event verification methods for determining the time, date and the physical origin of the content involved in Deepfake.
L. Nataraj et al, [12] have proposed a GAN based model to detect Deepfakes. Authors have used a combined approach by using co-occurrence matrices and deep learning. They have extracted the co-occurrence matrices on 3 different color channels in a pixel domain. Extracted features are supplied to the deep CNN framework for Deepfake detection process. Authors in the proposed research work has used Phonemes and Visemes of a subject or a person as indicators for detecting Deepfake in videos. phonemes are perceptually distinct units of sound in spoken language were as, Visemes, is the visual counterpart of a phoneme, corresponds to the mouth shape needed to enunciate a phoneme. In the research work [16], authors have used Visemes and Phonemes for detection of lip-sync Deepfakes. They have found that the fact that the mouth shape dynamics occasionally mismatches with the phoneme. The identified mismatched have been used for the detection of Deepfake. The figure 6 illustrates the steps involved in detection of lip-sync Deepfakes.

The input image will be first converted into gray scale. Profile with vertical intensity will be later extracted from the centre of the mouth. This value would be used for the differentiation of lip-sync Deepfake detection as depicted in figure 6

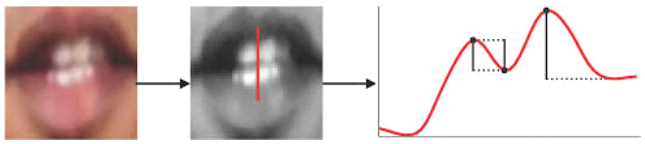

Fig. 6. Lip-Sync feature extraction based on phoneme and viseme

V. PROS AND CONS OF DEEPFAKE

Deepfakes in the current context pose a real threat to varied levels of society. Its use in the coming future will range from benign and novel to potentially ruinous and damaging effects. Existing detection methods struggle to combat evolving technology used to create Deepfake content. For the content consumers, it is hard to differentiate between original news from the fake one. This makes a robust case for an urgent need for specialized education and awareness among society from falling pray of such Deepfake content.

A. Potential Dangers of Deepfakes

Deepfakes could have an impact on all facets of society, from organised crime to threats. In order to make films and images appear authentic when they are not, Deepfake uses computers and machine learning methods to create them. According to experts, Deepfake can be used to start trouble and to make erroneous assumptions, particularly when it comes to a person's reputation, which is difficult to detect. By fabricating video of political figures saying or doing things they never said or did in an effort to sway public opinion, any political person can easily become the target. With the speed of Deepfake technology, movie stars, world leaders, corporate identities, presidential contenders, religious institutions, and other well-known authority are constantly assailed. The scenario got worse with the advent of Deepfake technology, including fake emergency forecasts, fake and inaccurate information during the election, misleading during election campaigns, terrorist advocacy, and many more.

B. Deepfake technology as a threat

Deepfake has also created new opportunities for threats in other technological fields. Deepfakes pose a real danger to facial recognition monitoring because facial recognition uses biometric algorithms that recognise faces and integrate access control with physical security. Enterprises that integrate operational technology, IT, and physical security internally connected create new IP-based infrastructures, implementing the usage of smartphones for biometrics and multifactor authentication. Smart phones can now grant access to secure locations. Deepfakes made it simple to disseminate bogus information and news. Creating memes can encourage someone to accept a collection of facts, whether they are true or false. When they review any facts, numbers, or hot takes, many people have biassed confirmation that evolved into true beliefs. When videos or photos are included, this biassed confirmation rises. So, it is up to us to determine what is true and what is fraudulent.

Deepfakes have become the goldmine for the criminals and virtual scams. They use Deepfake to mimic accent, annotation, tone and pattern of speech in creation of manipulated videos. Synthesized audio would be often utilized in kidnapping scams by targeting the victim via phone calls demanding the payment for releasing the kidnap. These scams often includes a virtual actor impersonating the kidnap within the background who screams for help. Victim fall pray believing the generated voice as the voice of the kidnapped and will immediately transfer a huge ransom to the criminal. In the corporate sector, the Deepfakes are often used in black-hat marketing. Criminals create manipulated videos of the organization's CEO making offensive statement and making it viral through the social media. Corporate sabotage with Deepfake can easily falsify the stock exchange information driving the organization to run under loss. Thus Deepfakes cause widespread civil unrest. On the other hand, the right use of Deepfake has created a technological solution for good cause, thereby creating ample space for the development of effective solutions in healthcare industries, entertainment, business and in E-Commerce.

C. Benefits of Deepfake technology

Deepfake has proven beneficial in its application, in different fields to shelter social media, movie industry, digital communications, games, entertainment, healthcare, and many business domains like fashion and E-commerce. It has raised the movie industries sooner. It can make use of digital sounds for the actors who face the problem in their voice. Filmmakers can recreate an impressive scene in movies, make fresh movies starring long-dead actors, create computer graphics and featured-face editing after post-production and improve video quality more professionally. In gaming applications, it enables the gamer to play multiplayer games with the assistance of digital twins. Deepfake superimposing technologies have simplified E-commerce and advertisement in significant ways. It helps in generating the targeted fashion ads which consistently changes with time, the trend and the taste of viewers. Healthcare industries are looking forward in applying Deepfake technology for creating deep generative models to for the invention of new possibilities in healthcare. Hence, artificial Intelligence must complement and augment human endeavor, not replace it. It is required to combine checks and balances that prevent inappropriate use of technology which creates the threat to the society. Rather there is a need for creating aright infrastructure which would connect the different experts that ensures the development and use of technology to thrive the society.

VI. Conclusion

With elevated technology, Deepfake are the fruits of deep generative modeling, and recent technology evolution that has enabled to produce a replica of original faces and builds new and elegant vivid Portraits of people who never exists. Moreover, Deepfake has put up a set of claiming policies, technology and legal concerns. Being a user we should always check originality of everything we observe, overhear or browse online. In the Deepfake world, to be a reliable user of technology, it is necessary to make sure the legitimacy of every tad of information spread via media, rather foolishly accepting everything. Deepfake technology is a bi-face tool having positive and negative implications. Certain activities and precautionary measures ought to be taken to attenuate the harm done by those that use Deepfakes with the odious plan. Protecting personal data against Deepfake traps is at most essential as Deepfake detection is not as power to combating Deepfake technology. Deepfake principal relies on human errors more specifically on the error of judgment, which is one of the principles of hacking techniques. To combat Deepfake, we have to spread adequate awareness within the society and develop detection and protection mechanisms by the programming community to safe guard humanity from falling prey.

REFERENCES

[1] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems (NeurIPS), pages 2672–2680, 2014.

[2] Yuezun Li, Ming-Ching Chang and Siwei Lyu ,In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking, arXiv:1806.02877v2 [cs.CV] 11 Jun 2018

[3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[4] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[5] Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., and Natarajan, P. (2019). Recurrent convolutional strategies for face manipulation detection in videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 80-87).

[6] Guera, D., and Delp, E. J. (2018, November). Deepfake video detection using recurrent neural networks. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (pp. 1-6). IEEE

[7] Li, Y., and Lyu, S. (2019). Exposing Deepfake videos by detecting face warping artifacts. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 46-52).

[8] Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., and Metaxas, D. N. (2019). StackGAN++: Realistic image synthesis with stacked generative adversarial networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(8), 1947-1962.

[9] Ohad Fried, Ayush Tewari, Michael Zollhofer, Adam Finkel-stein, Eli Shechtman, Dan B Goldman, Kyle Genova, Zeyu Jin, Christian Theobalt, and Maneesh Agrawala. Text-based editing of talking-head video. ACM Transactions on Graphics, 2019

[10] Chintha, A., Thai, B., Sohrawardi, S. J., Bhatt, K. M., Hickerson, A., Wright, M., and Ptucha, R. (2020). Recurrent convolutional structures for audio spoof and video Deepfake detection. IEEE Journal of Selected Topics in Signal Processing, doi: 10.1109/JSTSP.2020.2999185.

[11] R. Wang, L. Ma, F. Juefei-Xu, X. Xie, J. Wang, and Y. Liu, "FakeSpotter: A Simple Baseline for Spotting AI- Synthesized Fake Faces," arXiv preprint arXiv:1909.06122, 2019.

[12] L. Nataraj, T. Mohammed, B. Manjunath, S. Chandrasekaran, A. Flenner, J. Bappy, and A. Roy- Chowdhury, "Detecting GAN Generated Fake Images Using Co-Occurrence Matrices," Electronic Imaging, no. 5, pp. 1– 7, 2019.

[13] F. Marra, C. Saltori, G. Boato, and L. Verdoliva, "Incremental Learning for the Detection and Classification of GAN-Generated Images," in Proc. IEEE International Workshop on Information Forensics and Security, 2019.

[14] I. Amerini, L. Galteri, R. Caldelli, and A. Bimbo, "Deepfake Video Detection through Optical Flow based CNN," in Proc. IEEE/CVF International Conference on Computer Vision, 2019

[15] Thanh Thi Nguyen, Cuong M. Nguyen, Dung Tien Nguyen, Duc Thanh Nguyen, Saeid Nahavandi, Fellow, IEEE. "Deep Learning for Deepfakes Creation and Detection: A Survey", awrXiv: 1909.11573v2 28 Jul 2020.

[16] S. Agarwal, H. Farid, O. Fried and M. Agrawala, "Detecting Deep-Fake Videos from Phoneme-Viseme Mismatches," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 2814-2822, doi: 10.1109/CVPRW50498.2020.00338.