System for Detecting Deepfake in Videos – A Survey

Akshay Ramachandra Bhat
akshay1ga19cs009@gmail.com

Ankush J
ankush31001@gmail.com

Bharath
Bharathasodu2001@gmail.com

Akash Deep Singh Sodhi
akashdeep329ar@gmail.com

---***---

Abstract - *nowadays, freely to be had software program grounded on device literacy methods has resulted inside the era of veritably realistic fake content that has counter accusations on society in an duration of faux news. Software including FaceApp are freely to be had and may be utilized by each person to create practical searching fake motion pictures. Such videos if used with a terrible rationale also the outcomes can be critical and may have an effect on society and people. alternatively, crucial exploration has been completed to be able to increase discovery styles to reduce the negative outcomes of deepfakes. This paper give a review of which are used to descry comparable manipulated d videos. We explore several techniques used to create face based totally manipulation videos and evaluate a number of deepfake discovery methods grounded on numerous parameters which incorporates generation styles, technique used, datasets and so on.*

Key Words: **Deepfake detection, Deep Learning, Generative Adversial Networks (GANs), Convolution Neural Networks (CNN).**

1. INTRODUCTION

Thanks to the progress in deep getting to know era in addition to laptop vision in current years, a surge has been seen in fa- ux face media. every day, a massive variety of DF pictures and films are shared on social media systems. DF movies are spreading, feeding faux information and endangering social, country wide, and international ties. Human beings are - worried that what they study at the internet or watch at the net is now not reliable and honest. On this backdrop, in January 2020, a popular social media platform introduced a brand new coverage prohibiting AI-manipulated videos that could mislead the viewers for the duration of elections. The trouble is that that is depending on the potential to tell the distinction among real and false videos. creation of Deepfake movies is based totally at the concept of changing a person's face with any individual else's face. The requirement to achieve this is that the sufficient wide variety of photos of each the humans have to be available. studies carried out these days has focused on how those deepfake motion pictures are crafted and a way to understand them with the aid of analysing different capabilities intently. But with the development in era, it has come to be an increasing number of challenging to inform apart fake movies from the real ones.

2. TECHNICAL BACKGROUND

A.CNN

A Deep Learning advances in computer vision have been constructed and improved over time, largely through one technique – a Convolutional Neural Net-based approach work[31]

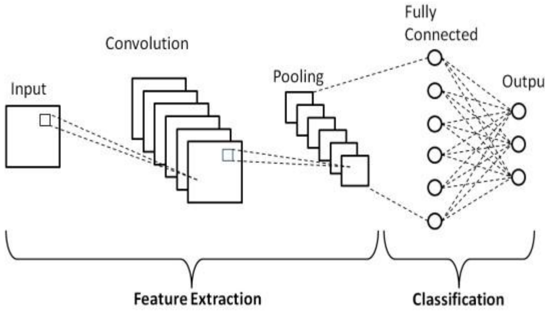


Fig. 1. Basic Architecture of CNN

CNNs are a type of deep learning network[31]. An algorithm that can take an image as input and give priority to distinct aspects/objects in the image (learnable biases and weights) while distinguishing between them. A CNN requires significantly less pre-processing than a neural network. In contrast to other classification systems, CNN has hand-engineered filters in its core techniques, which they can use with the proper training. The ability to detect these filters/features is determined by the design of the building. The study of Neurons in the Human Brain of the Organization of the Visual Cortex influenced the connectivity pattern of a CNN, which is analogous to the connectivity pattern.

The CNN's job is to condense the pictures into a format which is simpler to process while retaining important elements for accurate prediction. This is essential for designing an

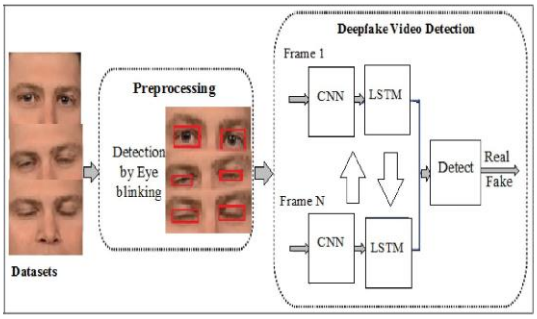


Fig. 2. DeepFake Detection using CNN and LSTM

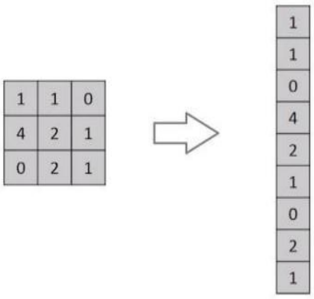


Fig. 3. Flattening of a 3x3 image matrix into a 9x1 vector

Architecture that can learn features while also being scalable to large datasets.

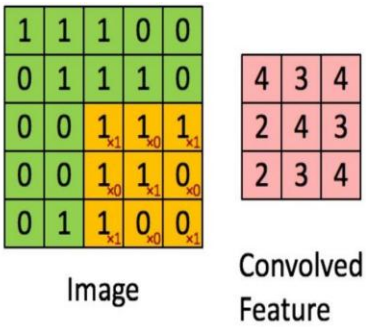


Fig. 4. Image to convoluted feature

In order to retrieve high level characteristics such as edges from the source images, CNN is used. CNN doesn't have to have a single Convolutional Layer. Traditionally, the first ConvLayer is in control of capturing Low Level information like edges, color, gradient direction, and so on. The architecture gets adjusted to the High-Level characteristics as layers are added, resulting in a network that comprehends the pictures in the dataset in the same way as we do.

B. RNN

An RNN[32] is a neural network wherein the result of the previous step is being utilised as input in the following step. This in contrast to regular neural networks where all the inputs and outputs are not dependent on each other. However, in few situations, for example trying to predict the following term of a phrase, the previous words are crucial, and thus the previous words must be remembered.As a result, RNNs were synthesised. They utilise the intermediate layers to solve a problem. The hidden state in RNN remembers information about the hidden sequence. RNNs have "memory" that stores all of the results of the calculations. The memory uses the same configurations for all inputs or hidden layers. Therefore, it achieves exact same outcome by performing identical work on all inputs or hidden layers.

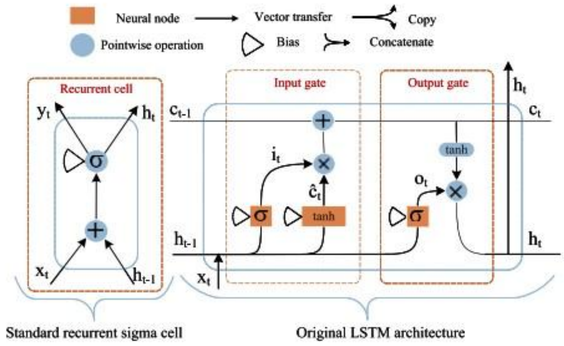


Fig. 5. RNN's fundamental architecture

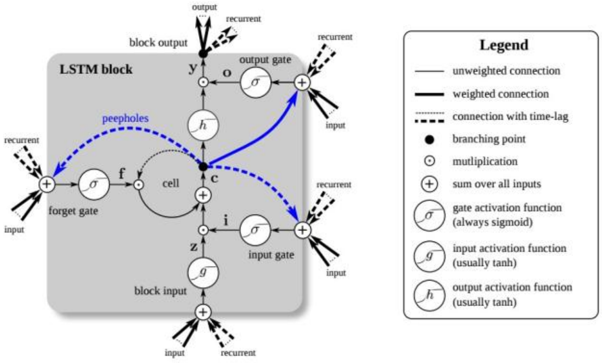


Fig. 6. LSTM Units

Hochreiter and Schmidhuber[33] LSTM was proposed to deal with long-term dependencies whenever the separation among relevant data input is great. It achieves all the intriguing result based on RNN and hence they have been the centre of deep learning. The recurrent hidden layers of RNN's,are composed of recurring cells whose conditions are controlled prior instances as well as the current sources via feedback networks.

C. GAN

Texts, images and video generation, drug discovery, and textto-image conversion have all been employed in real-world applications.

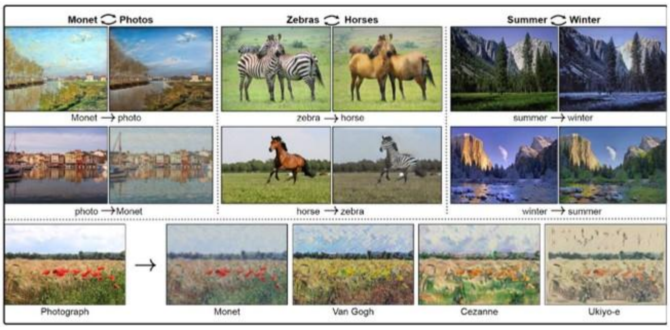


Fig. 7. Image to Image synthesis using GAN

GANs are among the most predominant Machine Learning techniques devised in recent times. In a nutshell, they are algorithms from the generative models group. These algorithms are a subcategory of the unsupervised learning discipline, that concentrates on algorithms which understand the fundamental structure of the data without defining a specific value. Generative models discover the inherent distribution of data input p(x), allowing them to generate

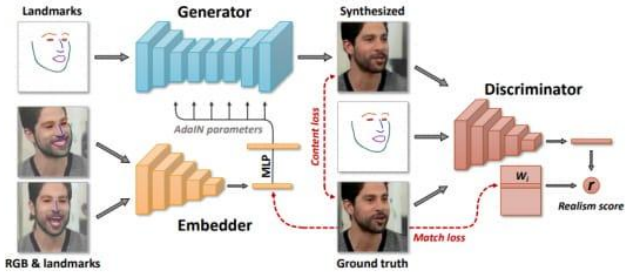


Fig. 8. DeepFake created using GAN Model

- The first model, referred to as a Generator, attempts to develop new information which is exactly equivalent to the simulated values. The Generator is like a person who creates forged artwork.

- The next model is the Discriminator. The primary objective of this model is to find whether an input data set is 'real,' indicating it relates to the original dataset, or 'fake,' meaning a forger created it was created by a forger. A Discriminator in this situation is similar to an art expert who attempts to determine not whether works of art are genuine.

The above generator is modeled using a neural network G(z,θ1). Its task is to convert the noise of the input variable z to the required data source x. (say pictures). On the other hand, the second neural network D (x, θ2) models discrimination and generates the probability that the data is from real data on a scale of 0% to 100%. (0,1).In both cases, the weight or index that defines each neural network is specified by thetai.

As a consequence of this training, the Discriminator can accurately identify input data as true or false. This implies that its weight values are optimised to optimise the likelihood of any real input data x being classified as a part of the genuine dataset whilst also reducing the likelihood of any false picture being categorized as the real dataset. In more technical jargon, with the help of loss/error function, D(x) is optimized and D(G(z)) is minimized.

Besides that, the Generator has been trained to misguide the Discriminator by producing data which is as real as possible, suggesting that Generator's weights have been optimized to increase the probability that any counterfeit

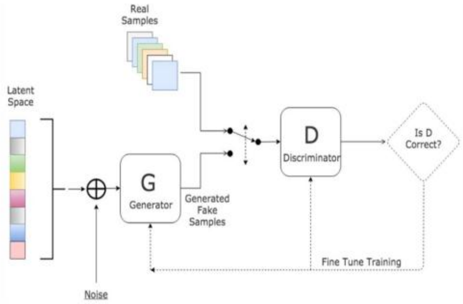


Fig. 9. Global concept of GAN

Since the neural networks represent both generator and dis- criminator, the GAN can be trained with the aid of a gradient based optimization technique picture will be categorized as belonging to the legitimate dataset. In mathematical terms, this means that the network's loss/error function maximizes D(G(z)).

3. FACE – BASED VIDEOS MANIPULATION METHODS

In the last twenty years, the prominence of simulated face tampering has shot up. Zollhofer et al. [16] provided a detailed report. Bregler et al. [17] specifically presented Video Rewrite,to create a new different video artificially of an individual with differing mouth movements by using imagebased method. With video face replacement, one of the very basic automatic face swap approaches [18] were proposed by Dale et al. They use single camera movies to recreate a three - dimensional model of those two faces and then use the resulting 3D model. The first facial reenactment expression transfer was accomplished by Thies et al. [19]. A consumerlevel RGB-D camera was used for recreation and tracking 3D representation of the source and target actors both. The analysed deformities in original faces are applied to model of face to be modified. To convert the graphically computed modification of faces back to their original form, Kim et al. [22] learned an image translation neural network. NeuralTextures [19] improves the texture generated using this method in agreement with the network to determine the restored output instead of a pure translated network of image from an image.

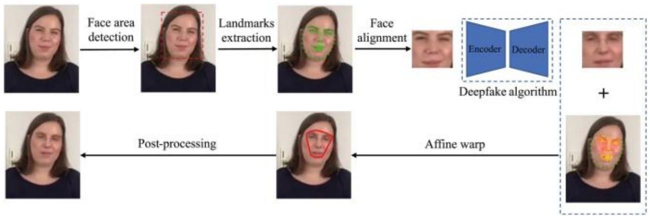


Fig. 10. Fake face Synthesis

The following table shows popular deepfake generation tools.

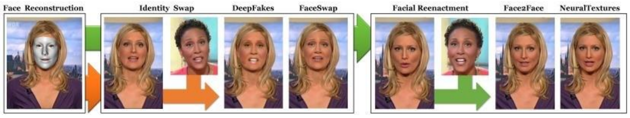


Fig. 11. Face Manipulation Methods

TABLE I

MANIPULATION METHODS

Methods	Deepfake Generation	Techniques used
Entire Face synthesis	This it creates nonexistent faces	GAN (ex:StyleGAN)
Identity Swap	Replacing of One person's face with another person's face	FaceSwap, ZAO mobile application
Attribute Manipulation	Modification of Specific Features such as face editing, adding spectacles to a face image, and so on	GAN (ex:StarGAN)
Expression swap	modification of a person's face expression	Face2Face, Neural Textures
Picture Animation	Regardless of the initial image, it is animated into a talking head, which is quite similar to the driving video	Fake Image Animation
Image to image generation	Input image and target image will be provided to cGANs networks, and anticipated picture will be accomplished	cGANs (conditi onal GAN)

Image to image generation	Input image and target image will be provided to cGANs networks, and anticipated picture will be accomplished	cGANs (conditi onal GAN)

A.FACESWAP

FaceSwap [9] is a method of transferring facial features from one video to another. Remove the facial area using the infrequently recognized facial markings. The method use scomposite images to fit these characters into 3D models. The textures from the source images are used to back-project this model to the intended image, lowering the discrepancy between both the predicted shape and the localized landmarks. eventually, the image is mixed with the rendered model, and color-variance correction is applied. Until one video ends, we repeat these processes for all source and target frames in pairs. With a lesser cost of computation the execution can be run effectively on the CPU.

B. DEEPFAKES

Deepfakes[9] is a term that has come to refer to deep learning-based face substitute, but the same name is used for a deception technique that has expanded via online platforms too. A face from the origin video or photo collection is used to replace a face in a target image or video. The method uses a shared encoder to train two auto encoders to create images for training of the source and faces to be manipulated. A facial recognition system is used to crop and align the photos. The trained encoder and decoder of the source face are applied to the target face in order to create a fake image. The output of the auto encoder is then blended into the image.

C. FACE2FACE

Face2Face [9,28] is a facial reconstructive mechanism that pass on the manifestations of a source video to an intended video without compromising the target's identity. The two video input streams were used by the initial implementation with a manually done keyframe selection. A dense face reconstruction made use of these frames to reintegrate the face under various lighting conditions and facial expressions. We use the Face2Face technique to fully automate the creation of reenactment alterations for our video database.

D. NEURALTEXTURES

For their Neural Textures-based rendering approach, Thies et al. [9] used face reenactment as an example. It learns a neural texture of the target individual, including a rendering network, from the original video data.

Then the models can be trained by making use of the combined losses occurred by an adversarial network and while photometric reconstruction. Tracked geometry is employed during training and testing period in the Neural Textures method. These figures are generated using Face2Face's tracking module. Only the facial emotions corresponding to the mouth region are modified; while the region of eye remains intact.

4. DATASETS

There are many useful online resources that can be used to monitor the progress of technology. New deep learning models such as GANs have been used in recent developments in humanbased video integration. The GAN model consists of two parts, both of which are deep neural networks trained in series. The first is that the mesh builder aims to create face images as close to real images as possible. Second, the discrimination of the mesh is designed to distinguish between them, making the synthetic image of the face look real.

This is the main idea behind creating deep objects. Therefore, a new video was created by superimposing the target's face on one of the viewers. Therefore, a new video is created from the beginning of the task.

TABLE II

DATASET USED

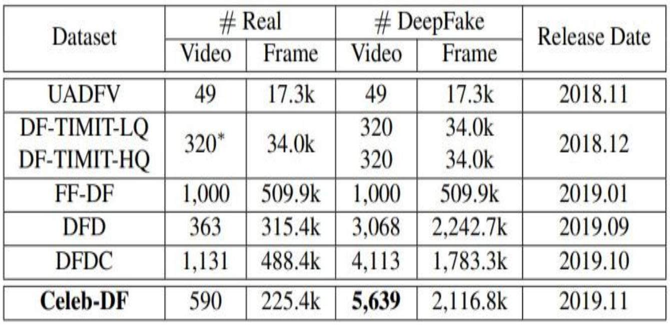
	Datasets	About
1	FaceForencic++	It is a forensics dataset made up of 1000 original video sequences which are modified using 4 different face modification techniques: Deepfakes, Face2Face, FaceSwap, and NeuralTextures. The input came from 977 YouTube videos, all of which included frontal face which was clearly visible, allowing tampering methods to create plausible forgerie
2	CelebA HQ	The CelebA-HQ dataset is an enhanced version of CelebA with 30,000 images at 1024x1024 pixels.
3	Flicker	They allow us to establish a standard for image localization of textual entity mentions.
4	UADF	This dataset contains 49 medium-quality real and fake videos.
5	VoxCeleb	It is free to download and install, and it is available worldwide. After many hours of work, the dataset was created by recording interviews with celebrities and well-known people on YouTube, one of the most popular websites. i)VoxCeleb1's database contains over 100,000 samples. VoxCeleb2 has nearly a million samples.
6	CelebDF	Celeb-DeepFake DF, the DF synthesis Algorithm is used to create the videos which is critical to improving visual quality. It is divided into sections that address various visual defects found in current datasets.

Face of a specific individual as an input which is swapped with an individual from the source. Zhu et al. [29] proposed CycleGAN as a strategy for improving GAN performance. Bansal et al. [30] developed Recycle-GAN, which extends previous work by incorporating spatial and temporal signals via conditional generative adversarial networks. The two primary types of forensics datasets. First one is Traditional and second is Deepfake. Classical DeepFake forensics datasets are the two main types. Traditional forensic datasets are handcrafted in less controlled environments such as camera artifacts, compositing, painting, resampling, and rotation detection. IFSTC hosted the first Forensic Medicine Photo Contest (2013), an international event where participants film thousands of indoor and outdoor scenes with 25 digital cameras. There are 82 occurrences of 92 unique and 101 unique mask variants in the Wild Web Dataset (WWD) [9]. WWD tries to bridge the measurement gap in native photography techniques. [3] Evaluate performance.The CelebFaces Attributes Collection (CelebA) includes over 200,000 celebrity photos and 40 feature notes. With a total of 10,177 people, 202,599 face images, 5 geographies, and 40 binary feature notes per image, CelebA has a total of 10,177 people, 202,599 face images and rich annotations.

DeepFake datasets are the second most common type of forensics dataset. GAN-based models, which are particularly famous because of their performance, are commonly used to generate these datasets. The UADFV [3] is made up of 49 real videos and 49 DeepFake videos created with FakeAPP and the DNN model. These films are about 11:14 seconds long on average, with a resolution of 294 x 500 pixels. The DeepFakeTIMIT (DF-TIMIT) dataset [3] was created by merging the VidTIMIT dataset [3] and FaceSwap-GAN; 16 similar-looking pairs of people from VidTIMIT [3] were chosen, and the database generated approximately 10 videos for each of the 32 people using low-quality of size 64 × 64, i.e., DF-TIMIT(LQ), and high-quality of size 128 × 128.

FaceForensics (FF) [8] is a DeepFake dataset aimed at performing forensic tasks such as facial detection and segmentation of falsified images. Over 500,000 frames, it is made up of 1004 videos (facial videos taken from YouTube). Sourceto-target manipulation, in which Face2Face reenacts the facial expressions of a source video, and self-reenactment manipulation, in which Face2Face reenacts the facial expressions of a source video, are the two types of manipulation.

The FaceForensics++ (FF++) [15] dataset has 1,000 actual YouTube videos were gathered, and 1,000 DeepFake videos were created using each of the four face methods: DeepFake, Face2Face, FaceSwap, and Neural Texture are some of the tools available.



Dataset	# Real		# DeepFake		Release Date
	Video	Frame	Video	Frame	
UADFV	49	17.3k	49	17.3k	2018.11
DF-TIMIT-LQ DF-TIMIT-HQ	320*	34.0k	320 320	34.0k 34.0k	2018.12
FF-DF	1,000	509.9k	1,000	509.9k	2019.01
DFD	363	315.4k	3,068	2,242.7k	2019.09
DFDC	1,131	488.4k	4,113	1,783.3k	2019.10
Celeb-DF	590	225.4k	**5,639**	2,116.8k	2019.11

Fig. 12. Basic Information of Various DeepFake Video Dataset

5. METHODOLGY

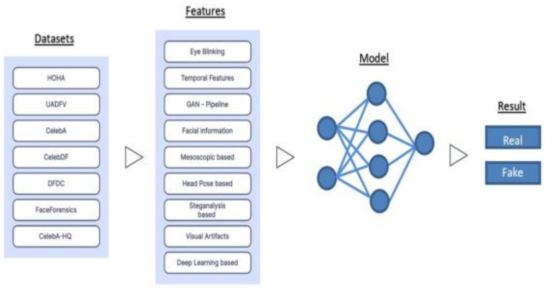


Fig. 13. Some important models and features used for detection

Methodology is a systematic, theoretical evaluation of the procedures utilized in an area of study. It involves a conceptual investigation of a collection of methods and principles related to a specific field of study. It typically includes terminology such as mode of thinking, theoretical model, stages, and quantitative and qualitative methodologies.

TABLE III

DEEP FAKE DETECTION METHODS

No	Title of the paper	Techniques used	Dataset used
1	Deepfake Video Detection Using Recurrent Neural Networks	employs LSTM and RNN, CNN for feature extraction, LSTM for sequence Processing	HOHA.
2	Detection DeepFakes in Videos Utilizing Feature engineering in Deep Learning CNN Frameworks	DWT, CNN+SIFT	False texts, voices, movies, and images are all examples of fraudulent media..
3	Deep Learning ar Super Resolutic Algorithms Combined for Deep Fake Detection	CNN Resnet! Model wi Super Resoluti Algorithms	CelebA and UADFV
4	Deepfake Detection Using Clusteringbased Embedding	Regularization Meso4 algorithm, FWA algorithm, EVA algorithm, Multitask algorithm, and the final Xception-c23	UADFV, CelebDF, and DeepFakeDetection algorithms
5	Deepfake Detectio n Using SVM	DeepFake, Ima Processing, SV GAN, DFT	CelebA dataset
6	DeepFakes and Beyond: A Survey of Fake and Detection and Face Manipulation Databases	CNN, RNN	DFFD
7	Fighting deepfak e with residual noise utilising CNN	InceptionResNet V2, CNN Deep Learning algorithm.	DFDC FaceForensics dataset
8	Deepfakecreation and detectio n using Deep Learning	MesoNet CNN,	Deepfake Detection Challenge

9	Detecting CNN Generated Facial Images in Real-World Sce- Narios	Using Generati Adversarial Networks (GANs) CelebA-H (CAHQ)	Flickr-FacesHQ (FFHQ)
10	FaceForensics++: Attempting to Detect Manipulated Faci Images	Dete Using Steganalysis Method	FaceSwap, DeepFakes, Face2Face,Neural Textures,Post Processing Video Quality

These criteria define the form or type of data collection or, in some cases, how a result will be calculated. The process does not describe a specific process, but most importantly the type of process that needs to be done to achieve the work and goals.

[1] proposed a timesensitive pipeline for automatic detection of deep fake videos. It uses a convolutional neural network (CNN) to extract phase features. These features are used to train a neural network (RNN) to determine whether the video has been manipulated. The main results of this study are: the analysis is performed in two stages, the CNN extracts the sample stage, and the timesensitive RNN network captures the physical difference in the face swapping process. 600 videos, half of which are deep fakes from various video hosting sites, were used to test the proposed method.

The result of the given method is shown visually in the equation, allowing it to decide whether the suspect video is correct or not.

Due to the large amount of data frames lost after video compression, most image detection methods cannot be used for video alone. Extracting video into frames is a great way to learn about this medium. DWT is used to filter in [2]: DWT splits the image into four parts. When using the Python library Py Wavelets (HH). Vertical Detail, often referred to as highlow, HL, is a special output to use in it.This decomposition image can be thought of as a frequency filter for the frame. The CNN feed will be a new filter. CNNs are used for classification. This will help detect errors or inaccuracies in deep fakes. The collected frames are split into 90% training frames and 10% test frames before being fed into the CNN model.This is an important stage of model training. To increase accuracy, the training process is programmed up to 50 times. Tuition fees vary for videos; in fact, the higher the resolution, the longer it takes to train the model.

In [5], a new method is proposed that includes group based embedding on regular basis. Open source techniques are used to make films that can reproduce special features in deep flims. Clusterbased embedding normalization is incorporated into objective classification to increase local smoothness of the representation space, resulting in a model that learns to avoid overlapping events. Our latest deep data is used to test the model.Experimental results show the effectiveness of the method. The Xception network uses positive and negative models for training classification. During the training phase, the number of classes was determined as 3 and the samples formed during the test were classified as negative samples.During training, a constant loss is also used to ensure the spacing between classes and the smoothness of the inclass placement area.

[8] proposed a method for using residual noise to be the difference between the original image and its noisefree version. Residual noise has been shown to be useful in deep sensing due to its specificity and discrimination, which can be achieved through neural networks with adaptive learning. The method was tested on two datasets: lowresolution FaceForensics++ videos and highresolution videos from the Kaggle Deepfake Detection Challenge (DFDC). In this article, we propose an adaptive learningbased classifier that uses convolutional neural networks to learn the noise of real and fake videos. The performance of the proposed method (DFDC) is demonstrated in two datasets, FaceForensics++ and DFDC.The aim is to determine whether the noise obtained from real video is different from the noise obtained from fake video. To remove the noise, remove the noisefree version of the frame from the frame itself. Frames are denoised using the wavelet transform function WF. Add residual noise to each frame. The backbone is the deep learning algorithm InceptionResNetv2, a 164layer CNN that trains over one million images from the ImageNet dataset.

[14] proposes a method for detecting the appearance of facial forgery, which is used at the level of mesoscopic analysis. In fact, microscopic research based on image noise becomes illegal in the case of video with image noise degradation after video compression. Similarly, it is difficult for the human eye to classify fake images at a higher level, especially when images show human faces. Therefore, it is recommended to use a deep neural network with a sufficient number of layers as an intermediate method. The two designs below, with the lowest representation and the least amount, yielded higher separation scores in all tests.They rely on image classification networks containing clusters for efficient convolution techniques and feature extraction, and use convolutional networks for classification.Meso4: This network starts with four layers of connectivity and integration, then adds a dense network with a hidden layer. All layers use the elimination function to increase their flexibility, while the layers use the ReLU function to provide nonlinearity and bulk normalization to match their generation and gradientinhibiting effects. The network has a total of 27,977 teaching parameters.MesoInception4: A different model using a different version of the initialization module from Szegedy et al. To replace the first two layers of Meso4.The purpose of this model is to release the output of several convolutional processes with different shapes to

expand the workspace where the model can be changed. Instead of the 5×5 convolution of the original Deepfake Detection module using neural networks, 3×3 extended convolution is used to avoid high semantics. It can be seen in the cross link between the idea of using extended convolutions with an initial module.

Deep fake detection algorithms Xception and MobileNet are discussed in [15] as two strategies for classification work to verify deep fake movies. FaceForensics++ training and analysis data, which includes four datasets using four different and wellknown deep recognition methods, were used. The results showed high accuracy across the data, with accuracy ranging from 91% to 98%. A voting machine was also developed that could detect fake videos by combining all four methods instead of one. This study discusses deep learning algorithms for implicit classification and hence detection of deep fake videos. FaceForensics++ was used as a raw video data source to train two neural networks using preliminary images: Xception and MobileNet. Each network is trained to provide four samples, one for each of the four most popular deepphishing platforms. These include Deepfakes, Face2Face, FaceSwap, and NeuralTextures. Evaluation of the model shows that it can distinguish real video from high quality video, but this performance depends on the depth mock platform used. To solve this problem, the voting process is proposed to use the results of various models to create better solutions.

6.CHALLENGES

Currently, there are many deepfake devices on the market with reasonable performance levels, but there are still many devices in development. In contrast, the development of deep fake creations, high demand for forensics and deep mastering expertise in conservation. GAN is a wellknown artificial intelligence algorithm that consists of two types of discriminators and generated models that compete to produce false positives. These parodies of real people are often famous and spread quickly on social media sites, making them super advertising tools. With this fact in mind, forensic experts who have a basic understanding of forensic tools can use minor branches of legal protection to avoid investigations. Therefore, the control model should be able to find the accuracy of the misinformation and reduce the error. Therefore, many methods of legal protection aimed at obscuring existing sensors are important for the development of multimedia forensics as it exposes gaps in already existing responses and encourages similar research for more robust solutions.

There are many models available today for creating or exploring deep holes, but they still lack details and certain limitations.

CHALLENGES IN DEEPFAKE CREATION

- Despite efforts to increase the visibility of DeepFakes cr eation, there are still some issues that need to be addres sed. Generalization, body instability, lighting, lack of ori ginality in eyes and lips, hand gestures and personal age are some of the problems with making DeepFakes.

- Generalization: The properties of generative models are determined by the dataset used to train them. As a result, after completing training on a certain dataset, the model's output reflects the learned properties (fingerprint). Furthermore, the output quality is influenced by the amount of the dataset used during training. As a result, in order for the model to provide high-quality output, it must be given a dataset large enough as input to attain a specific sort of feature.

- Temporal coherence: Other flaws involve visible fluttering and juddering among frames, as well as a dearth of temporal coherence. Deepfake creation methods which operate on every frame do not take into account temporal inconsistency which leads to several issues.

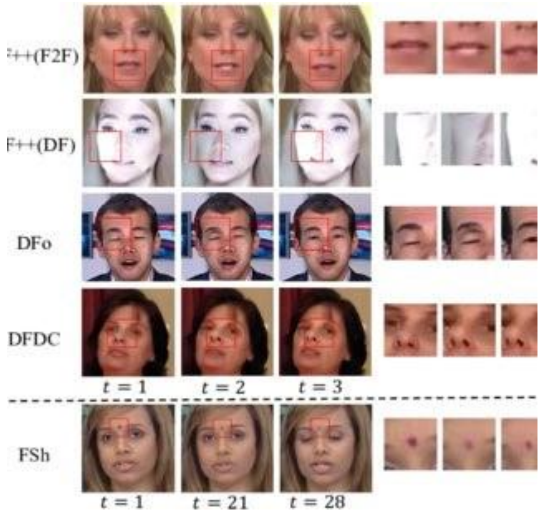


Fig. 14. Abnormalities of temporal coherence

- Differences in illumination: DeepFake datasets are created in a controlled environment with consistent lighting and background. In indoor/outdoor settings, however, a rapid change in lighting conditions results in colour inconsistencies and strange irregularities in the output.

- Lack of realistic emotions: The main challenges of DeepFake generation based on eye and lip synchronisation are an absence of emotions, disruptions, and the target's communication tempo.

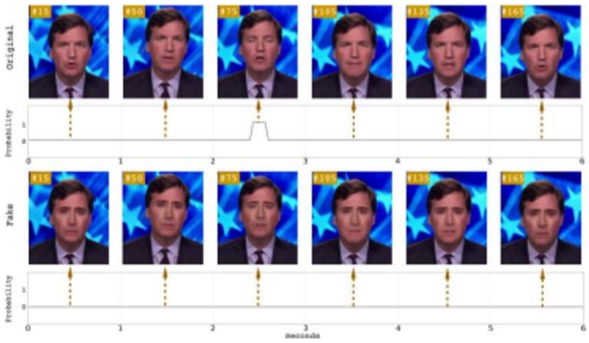


Fig. 15. Abnormalities of eye blinking

• Gestures: When the target expresses emotions with hand gestures, the developed DeepFake model has difficulty matching gestures with emotions expressed by gestures. Also, it is difficult to create such deepfakes due to the limited information of these articles.

CHALLENGES IN DEEP FAKE DETECTION

Although DeepFake detectors have improved significantly in performance, current detection algorithms still have some shortcomings that need to be addressed. Deepfake detection systems face many problems,the most important of which are described below.

- No DF dataset: DF model detection performance depends on the variety of dataset used during training. If testing is done on downloads with unknown features, it is difficult to build a detectable model without knowing the function. Due to the popularity of web platforms, post processing techniques are used for DF Multimedia to avoid misleading DeepFake detectors. Such operations can include removing artifacts from the body, blurring, smoothing, clipping, and more.

- Unfamiliar attack type:Another difficult task is developing a solid DF detection model against obscure sorts of assaults. These techniques are utilized to trick classifiers in their result.

- Temporal Aggregation: Current DF detection methods use binary frame-level classification to determine if a video frame is legitimate or fraudulent. These approaches, however, may encounter challenges like temporal anomalies and real/artificial frames appearing at regular intervals because they do not account for interframe temporal consistency.

- Unlabeled data:DeepFake detection methods are typically developed using massive datasets. At times, for example, in journalism or policing, only a limited number of dataset is available. Therefore, fostering a DeepFake detection model, unlabeled dataset is difficult.

7. CONCLUSION

This paper is about a new method, DeepFake, and this paperexamines a new and well-known method, DeepFake. Principles, advantages and disadvantages of DeepFake, GAN-based DeepFake applications are explained. A DeepFake detection model is also mentioned. Most modern deep learning search algorithms cannot be modified and extended, which means that the search space is still in its infancy. Many well-known organizations and experts working to improve the implementation process are willing to agree. However maintaining data integrity requires more effort and additional security measures.

REFERENCES

[1] David Guera, Edward J. Delp. Deepfake Video Detection Using Recurrent Neural Networks.

[2] Sonya J. Burroughs, Balakrishna Gokaraju, Kaushik Roy and Luu Khoa. DeepFakes Detection in Videos using Feature Engineering Techniques in Deep Learning Convolution Neural Network Frameworks. 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR).

[3] Yuezun Li , Xin Yang , Pu Sun , Honggang Qi and Siwei Lyu Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics. arXiv:1909.12962v4 [cs.CR] 16 Mar 2020.

[4] Nikita S. Ivanov, Anton V. Arzhskov, Vitaliy G. Ivanenko. Combining Deep Learning and Super-Resolution Algorithms for Deep Fake Detection.

[5] Kui Zhu, Bin Wu and Bai Wang. Deepfake Detection with Clusteringbased Embedding Regularization in 2020 IEEE Fifth International Conference on Data Science in Cyberspace (DSC).

[6] Harsh Agarwal, Ankur Singh and Rajeswari D. Deepfake Detection Using SVM in Proceedings of the Second International Conference on Electronics and Sustainable Communication Systems (ICESC-2021).

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

[8] Marwa Chendeb El Rai ,Hussain Al Ahmad, Omar Gouda, Dina Jamal,

Manar Abu Talib and Qassim Nasir. Fighting Deepfake By Residual Noise using Convolution Neural Networks in 2020 3rd International Conference on Signal Processing and Information Security(ICSPIS).

[9] Andreas Rossler, Davide Cozzolino, Justus Thies, Luisa Verdoliva, Matthias Nießner, Christian Riess. FaceForensics++: Learning to Detect Manipulated Facial Images in arXiv:1901.08971v3 [cs.CV] 26 Aug 2019.

[10] Nils Hulzebosch, Sarah Ibrahimi, Marcel Worring. Detecting CNNGenerated Facial Images in Real-World Scenarios.

[11] Hady A. Khalil, Shady A. Maged. Deepfakes Creation and Detection Using Deep Learning in 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC).

[12] Artem A. Maksutov, Viacheslav O. Morozov, Aleksander A. Lavrenov and Alexander S. Smirnov Methods of Deepfake Detection Based on Machine Learning.

[13] Luisa Verdoliva Media Forensics and DeepFakes: an overview in arXiv:2001.06564v1 [cs.CV] 18 Jan 2020.

[14] Darius Afchar, Vincent Nozick, Junichi Yamagishi and Isao Echizen. MesoNet: a Compact Facial Video Forgery Detection Network in arXiv:1809.00888v1 [cs.CV] 4 Sep 2018.

[15] Deng Pan, Lixian Sun, Rui Wang, Xingjian Zhang and Richard O. Sinnott. Deepfake Detection through Deep Learning in 2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT).

[16] Michael Zollhofer, Justus Thies, Darek Bradley, Pablo ¨ Garrido, Thabo Beeler, Patrick Peerez, Marc Stamminger, ´ Matthias Nießner, and Christian Theobalt. State of the art on monocular 3d face reconstruction, tracking, and applications. Computer Graphics Forum, 37(2):523–550, 2018.

[17] Christoph Bregler, Michele Covell, and Malcolm Slaney. Video rewrite: Driving visual speech with audio. In 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ’97, pages 353–360, 1997.

[18] Kevin Dale, Kalyan Sunkavalli, Micah K. Johnson, Daniel Vlasic, Wojciech Matusik, and Hanspeter Pfister. Video face replacement. ACM Trans. Graph., 30(6):130:1–130:10, Dec. 2011.

[19] Justus Thies, Michael Zollhofer, Matthias Nießner, Levi Val- ¨ gaerts, Marc Stamminger, and Christian Theobalt. Real-time expression transfer for facial reenactment. ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2015, 34(6):Art. No. 183, 2015.

[20] Justus Thies, Michael Zollhofer, Marc Stamminger, Chris- ¨ tian Theobalt, and Matthias Nießner. Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. In IEEE Conference on Computer Vision and Pattern Recognition, pages 2387–2395, June 2016.

[21] Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, and Kilian Weinberger. Deep feature interpolation for image content changes. In IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[22] Hyeongwoo Kim, Pablo Garrido, Ayush Tewari, Weipeng Xu, Justus Thies, Matthias Nießner, Patrick Perez, Christian Richardt, Michael Zollhofer, and Christian Theobalt. Deep Video Portraits. ACM Transactions on Graphics 2018 (TOG), 2018.

[23] Zhihe Lu, Zhihang Li, Jie Cao, Ran He, and Zhenan Sun. Recent progress of face image synthesis. In IAPR Asian Conference on Pattern Recognition, 2017.

[24] Grigory Antipov, Moez Baccouche, and Jean-Luc Dugelay. Face aging with conditional generative adversarial networks. In IEEE International Conference on Image Processing, 2017.

[25] David Guera and Edward J. Delp. Deepfake video detection ¨ using recurrent neural networks. In IEEE International Conference on Advanced Video and Signal Based Surveillance, 2018.

[26] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In International Conference on Learning Representations, 2018.

[27] Yongyi Lu, Yu-Wing Tai, and Chi-Keung Tang. Conditional cyclegan for attribute guided face image generation. In European Conference on Computer Vision, 2018.

[28] Justus Thies, Michael Zollhofer, Marc Stamminger, Chris- ¨ tian Theobalt, and Matthias Nießner. Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. In IEEE Conference on Computer Vision and Pattern Recognition, pages 2387–2395, June 2016.

[29] Mohammed Akram Younus, Taha Mohammed Hasan. Effective and Fast DeepFake Detection Method Based on Haar Wavelet Transform

in 2020 International Conference on Computer Science and Software Engineering (CSASE), Duhok, Kurdistan Region – Iraq.

[30] A. Bansal, S. Ma, D. Ramanan, and Y. Sheikh, ”Recycle-gan: Unsupervised video retargeting,” in Proceedings of the European Conference on Computer Vision (ECCV), Washington, 2018, pp. 119-135.

[31] https://en.wikipedia.org/wiki

[32] A. Malik, M. Kuribayashi, S. M. Abdullahi and A. N. Khan, ”DeepFake Detection for Human Face Images and Videos: A Survey,” in IEEE Access, vol. 10, pp. 18757-18775, 2022, doi: 10.1109/ACCESS.2022.3151186.
S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.