International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 51-53
Published Online June 2021 in IJEAST (http://www.ijeast.com)

DEEPFAKE DETECTION USING COMPUTER VISION

Ankit Mishra, Aman Verma, Arunav Dey, Abhay Singh
Department of CSE
IMSEC College, Ghaziabad, UP, India

***Abstract*— These days, a problem of AI-generated face swapping videos are rising rapidly, also known as the DeepFakes. These kinds of videos create threats to privacy of a specific person, peace and security of a country. On occasion perfect DeepFake clips authenticity could be hard to confirm with bare eyes. So scholars need to develop few algorithms to check the authenticity of videos. Here, we present overview of parameters that can tell us about the authenticity of videos. Main motive of this research is to develop a solution or methodology that can decide whether the face in the video was replace using DeepFake technology or not with some probability.**

Keywords— **DeepFakes, Haar Cascade Classifier, GAN's, CNN, Computer Vision**

I. INTRODUCTION

The idea of replacing face in video is not so new as, we can find examples of such photos made in 20th century. As an illustration, photo where the EX-U.S.A. President Obama's head was replaced with some random person head was produced. In the time of 2010 idea of neural networks was popular enough and humanity was on improving its computational power using GPU's. In these days, we can clone and execute such programs, these programs can help us to gain experience by observing without having great degree in math, computer theory, psychology, and many more.

Figure 1 Real(on left) & DeepFake(on right)

Today, there are many apps which help us to do so, like **FaceApp** that provides the facility to swap faces with high quality and make that funny. Luckily, thanks to the work of scholars, they brought together mystify and piece code of solid and firm face swapping technology and made it work on any personal PC. When it became open source, people started to use it to create vulgar and illegal visuals. For e.g., DeepFake celebrity pornographic videos/images to defame them. But such videos and photos are easy to detect the DeepFake with bare eyes. Massive threat of this technology content arose later, when it was difficult to recognize with bare eyes of people that whether video was manipulated or not, big threats are wrong informative videos or renowned person's manipulated video to defame him/her.

These days, there are many dangerous ways of implementing face swapping algorithms for e.g., like **faceswap**, in which the face of a person in the video is spontaneously replaced with another person's face without disturbing facial features of the face which was present earlier, **lip-sync**, which changes only the mouth region of person's face and on video that person is made to say something else that they had never said which was seen in the case of morphed video of Mr. Obama. The most dangerous part is in which specific person's face is replaced by another person, sitting in front of camera, to say something wrong or unethical to defame that person.

When the flood of Deepfake was seen on various platforms, research scholars chose to find solution to detect Deepfake visuals to save people from this privacy danger and country's peace. This is one of the reason why many datasets, containing AI-generated fake videos appeared.

In this paper we present algorithm which can decide whether photo was changed with DeepFake, face swapping, technology or not with high accuracy.

II. METHODS CREATING DEEPFAKES

There are 2 commons ways of creating Deepfake videos: autoencoders and GANs. Both of them are being applied now and we should talk about both of them if we want to create really good detect algorithm.

III. AUTOENCODERS

It is one of well-known deep learning techniques. Autoencoder insight is pair of encoder and decoder functions with shared weights, so 2 functions are training together.

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 51-53
Published Online June 2021 in IJEAST (http://www.ijeast.com)

Usually, we talk about autoencoders in cases of dimensional reduction and generative models learning. But autoencoders can be used for taking compressed representations of images to outdo existing image compressing standards. So we can use this representations (called autoencoder's latent vectors) from first encoder with decoder part of second autoencoder and get resulting mix of pictures - face swapped photo, for example. Open-source repositories using this technique are: Faceswap, DFaker, DeepFakeLab and etc.

IV. GANS

One of the most hard to train and use on computational hand deep learning technique is GAN. There we have 2 neural networks – generator and discriminator (judge, classifier, etc.). Generator net looks similar to autoencoder net, but we can achieve better results because discriminator net is rejecting
some bad examples. So GANs technique of Deepfake creation assumes that generator should fool discriminator (another machine) that makes such fakes more similar to real videos and also makes them harder recognizable by people's eyes. Some of the open-source projects are using this technique, for example, Faceswap-GAN.

V. TYPICAL SIGN OF DEEPFAKE VIDEOS

When we talk about Deepfake detection, obvious things that can tell us about video or photo "fakeness" can be found. There are some typical sign and we should check them out.

- Too smooth skin, lack of skin details
- Color mismatch between the synthesized face and the original face
- Visible parts of original face or temporal flickering
- Head position
- Artifacts on small moving parts
- Eye blinking rate
- Face warping artifacts
- Person's patterns of behavior

VI. MODULES FOR DETECTING TYPICAL SIGNS

We have developed following 7 modules :

A. Webpage for uploading media/input

A webpage using which a user can upload images and videos of which one wants to check whether the media is created or manipulated using Deepfake technology or not.

Figure 2 Module 1 SS

B. Face detection per frame

We are using **Haar cascade classifier** for the purpose of detection of a face in the video feed itself. Haar cascade classifier is based on the **Viola-Jones detection algorithm** which is trained in given some input faces and non-faces and training a classifier that identifies a face.

In each frame of the input video we are checking for the presence of a face, alongside of which we are also keeping a count of the total number of frames in the video.

$$\text{Presence of a face} = \frac{\text{Number of frames in which a face is detected successfully}}{\text{Total number of frames in the video}}$$

If the above value is above **0.75** then presence of a face is confirmed and the video is passed for further processing else the program terminates with appropriate error message.

C. Face recognition frame by frame

This module for detecting the following sign of Deepfake videos:

a) Too smooth skin, lack of skin details
b) Color mismatch between the synthesized face and the original face
c) Visible parts of original face due to temporal flickering

On the video we are performing face recognition frame by frame and adding up the probabilities of each frame and the average will be compared to the threshold.

PFN : Probability of Nth frame
N : Total No. of frames
Avg Probability = (PF1+PF2+PF3+.....+PFN)/N

Comparing Avg Probability with threshold probability, if greater then not a Deepfake video otherwise it's a deepfake video.

D. Landmarking of facial features

This module is used for facial features detection and landmarking. This module will detect and landmarking on the face as shown in figure on right side. It will plot following features :

a) Left eyebrow

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 2, ISSN No. 2455-2143, Pages 51-53
Published Online June 2021 in IJEAST (http://www.ijeast.com)

b) Right eyebrow
c) Left eye
d) Right eye
e) Lip
f) Nose
g) Jaw Line

This module is also be used in upcoming modules for further processing.

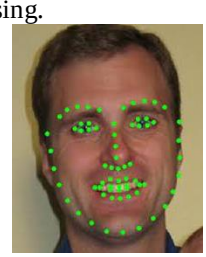

Figure 3 Module 4 SS

E. Flickering Detection

In this module, we are trying detect the flickering on every video. This module executes the following steps to detect the flickering:

- Access each frame of the video
- Detect whether face is present or not
- If face is present then extract the facial landmarks
- If any of the landmarks are missing then that frame will be count in variable x
- Probability that the video is fake = x/ tot. No. of frames

F. Eye blinking detection

In this module we trying to detect the eye blink of the person. This module uses the output of Landmarking of facial features.

Step 1: Extract facial landmark to to localize important regions of the face, including eyes, eyebrows, nose, ears, and mouth.

Step 2: Mark only those landmarks which are needed for the task, in this case eyes.

Figure 4 Contours on detected Eye

Step 3: Calculate Eye Aspect Ratio (EAR) using following formula:

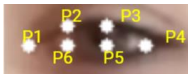

Figure 5 Landmarking points on detected eye

$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Step 4: Check EAR if it's greater than 0.25 then eyes are open, else blink detected.

G. Calculating eye blinking rate

From previous module consider 'x' blinks have been detected in 's' seconds. To calculate the blink rate we will use the following formula

$$blink\text{-}rate = \frac{x}{s} * 60$$

Where,
X is no. of blinks in video
s is length of video in seconds

VII. CONCLUSION

In this digital era everyone should always be very careful before sharing their personal data on any public platform as that data can be misused and it can one's reputation and image. Every technology has both the side a positive one and negative one we should always keep this in our and make any action on social media as it has become very hard to track source.

VIII. REFERENCES

[1]faceswap GitHub https://github.com/deepfakes/faceswap
[2]faceswap-ganGitHubhttps://github.com/shaoanlu/faceswap-GAN
[3]DeepfacelabGitHubhttps://github.com/iperov/DeepFaceLab
[4]DfakerGitHub https://github.com/dfaker/df
[5]DeepFake-tfGitHub
https://github.com/StromWine/DeepFake-tf
[6] Yuezun Li, Ming-Ching Chang, and Siwei Lyu. In ictu oculi: Exposing ai generated fake face videos by detecting eye blinking. In IEEE International Workshop on Information Forensics and Security (WIFS), 2018.
[7] Yuezun Li and Siwei Lyu. Exposing deepfake videos by detecting face warping artifacts. In IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019.
[8] Yuezun Li, Xin Yang, Pu Sun, Honggang Qi and Siwey Lyu. Celeb-DF: A New Dataset for Deepfake Forensics. arXiv preprint arXiv:1909.12962, 2019.
[9] Gao Huang, Zhuang Liu, Laurens van der Maaten and Kilian Q. Weinberger. Densely connected convolutional networks. In IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017.
[10] Maksutov A.A, Simonenko A.V., Shmakov I.S. Classifiers based on Bayesian neural networks. Proceedings of the 2017 IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conference, ElConRus 2017, 2017г.