Protecting Leaders against DeepFake

Shubham Gupta
M190718CS
Guided by: Dr Lijiya A, Dr Vasudevan AR

November 11, 2020

Keywords: Deepfake, Computer Vision, Deep Learning, Artificial Neural Network, OpenFace2,

Abstract

Creating a sophisticated Fake video of any person is not a big task in the current technology. Creating a DeepFake video of a Political leader can leads to an election manipulation or constitutional crisis. It can create a civil war between two countries or can be used for making a profit using an illegal way. Hence this paper Describes the DeepFake Technique's concerns to society. We have prepared a "Forensic Technique is to solve this issue." We design a Forensic Technique which is able to predict the given leader video is real or a DeepFake. We have given a conclusion till current progress.

1 Introduction

With the recent advancement in deep learning within the last few years, creating a sophisticated Fake video is not a big task. But using New technology Of Deep learning is such on a huge level that we created Deep-Fake looks realistic. Suppose a situation where a Deepfake Video can lead to election manipulation, a civil war between two countries, and much. There are a huge number of Side effects of Deep-Fake. There fore we discuss these issues in our problem formulation. We need a Forensic technique that is able to detect that the given video is real or a Deep-Fake but also works as an authentication system.

Fig 1 : These two Figure contains Face-swap Deepfake images[6]

2 Literature Survey

2.1 What is Deep-Fake Technology

DeepFake is a Synthetic media in which a person in an existing image or video is replaced with someone else's likeness[1]." Deep" word comes from "Deep learning". So using Deep learning to get Fake videos by replacing the Lipsing movement, Face Swapping, changing expression is a way to create deep fake. There are a huge number of Applications of Deep-Fake.Using deep Fake we can manipulate or generate visual and audio content with a high potential to deceive[1].

2.2 Type of Deep-Fake

DeepFake is a classified into the of the following catagories–

Lijiya A

2.2.1 Face-swap

Face-swap is a very famous type of Deep-Fake.In this, the face of a real person is replaced with the desired persons' face. This type of deep fake is used is inserting the face of an actor in movies and also widely used in Pornography by replacing the face of famous Celebrities.

2.2.2 Lip-Sync

In Lips sync, DeepFake's whole face remains the same and only the region of mouth is changed. Hence we will make them to say what we want by moving the lips according to us. This type of Deep-fake is created mostly to change the audio message and add their own message. Such type of deep Fake is created by Jordan Peele, Peele's voice id DeepFake with Obama video, to generate awareness in this issue.

2.2.3 Puppet-master

Puppet-master, in which a target person is animated (head movements, eye movements, facial expressions) by a performer sitting in front of a camera and acting out what they want their puppet to say and do.

2.3 How to create Deep-Fake

It's an automatic technique, no manpower is required. We start with an image or video that we want to make deep fake. Then we have a Generator, it makes changes to image/video and passes to the Discriminator. The discriminator has access to all real images. It will check that the given modified image/video is equal to real ones or not. If equal then it goes to generator again in a loop (almost millions of times), if no than the given deep fake is created. This type of Deep Fake generation also called a generative adversarial network (GAN). This technology is fully automatic hence a huge number of deep fake are create and will increase in the future.

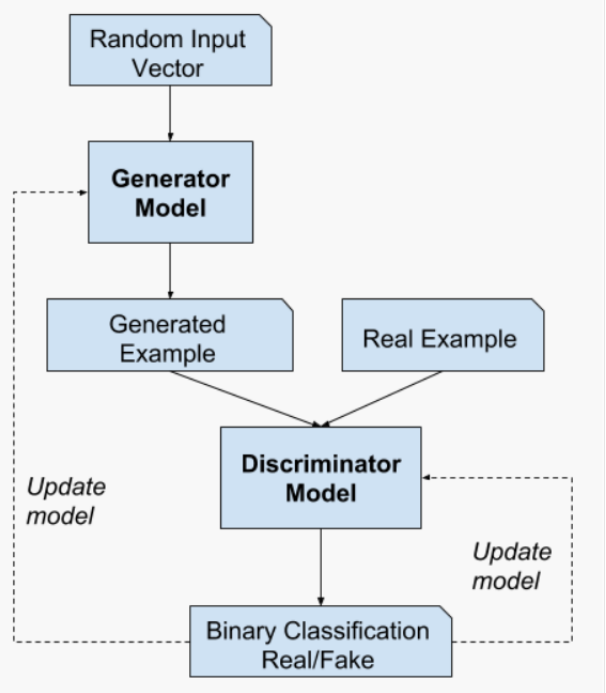

Fig 2: Representing Block daigram to generate DeepFake (generative adversarial network (GAN))[9]

2.4 Concerns of Deep Fake

Deep Fake technology is having a huge number of good applications like in movies difficult scenes can be done by trained professionals instead of real actors. But deep Fake technology is having more concerns which are as follows—

2.4.1 Political Concerns

Deefakes have been used to misrepresent well knows politicians in videos. In April 2018, Jordan Peele collaborated with Buzzfeed to create a Deep-fake of Barack Obama with Peele's voice to increase the awareness of DeepFake[2].So an election can be manipulated by creating a fake video of a world leader and spread the wrong message and benefit the oppositions.

2.4.2 Pornography Concerns

Creating a DeepFake video of Celebrities for Pornography is in a huge number. Almost 96% of online available Deep fakes are coming in this category [3].

2.4.3 Fraud Concerns

Audio deepfakes have been used as part of social engineering scams, fooling people into thinking they are receiving instructions from a trusted individual[1].

2.5 Motivation

Today's world, making fake videos is not a big task. Creating a video of a world Leader confessing to illegal activity leading to a constitution crisis, or a military leader saying something racially insensitive leading to civil unrest in the area of military activity. So, fake videos can lead to a massive disaster of society. To solve this problem, we must have a forensic technique to find that the given video is real one or a deep fake. Such a system will give an assurance that our world leader will save from deep fake. So, the main motivation of this project is to made a model that can be use used for the authenticity of a given world leader.

2.6 Literature methods

2.6.1 Detection using Artifacts

Whenever a deep fake is created, the frame goes through affine warping or manipulation. Superimposition of a source image on target image leaves resolution inconsistencies in the deep fake which can be used for the detection of a deep fake. Convolution neural networks like VGG16 can be used for such type of deep fake detection However, this approach fails for a deep fake which is created using the expression swap method.[7]

2.6.2 Detection using Appearance and Behaviour

In expression swap deepfake, the source's expressions are mounted onto the target video.However head movements, facial expressions are unique for a person.This uniqueness can be used for confirming real video. Temporal, behavioral biometric-based movements can be learned using ResNet101.Also, static facial biometric is obtained using the VGG network.[8]

2.6.3 Detection using facial and speaking Pattern

Everyone on this earth has unique facial features, these features include the shape of eyebrows, shape of lips when closed/open, relative location of landmarks on the face like nose, eyes, lips, etc. These unique features can be used for the identification of the leader in the real videos. If any video that shows deviation from such features then that video has to be fake. Three feature extraction networks can be used for higher accuracy. we can use GoogleNet,XceptionNet,DenseNet for extraction of features.[10] These networks extract and learn the features of a leader and can predict if a person's facial features in a video are genuine or fake.

Every person has a unique way of speaking with facial expression. When individuals speak, they exhibit relatively distinct patterns of facial and head movements. This method takes patterns of a given person using his speaking and facial expression and trains them to detect.This paper used SVM classifer to train it[4]
Both the first two method cannot be used for authentication while the last once gives.

3 Problem Statement

"Design a Forensic Technique that is able to Predict the given world leader video is Real or Deep-fake" if the video is real than it should be a real video of given world leader".The problem is just not only to identify that is Deepfake but can be used for authentication purposes. So That is can be used for the Verification Process.

Dataset contains videos 598 videos, having real, lip-sync deep fake , faceswap and Imposter deep

fake.
Fakes: lip-sync deep fakes and face-swap deep fakes
Imposter: comedic impersonations
Real: sample original videos (most scraped from YouTube)[4]

4 Proposed Method

Method Works on the principle that every person has a unique way of speaking with facial expression. When individuals speak, they exhibit relatively distinct patterns of facial and head movements. Also, the creation of all three types of deep fakes tends to disrupt these patterns because the expressions are being controlled by an impersonator (face-swap and puppet-master) or the mouth is decoupled from the rest of the face (lip-sync)[4].
Hence this property is can be used for detection of Deep Fake

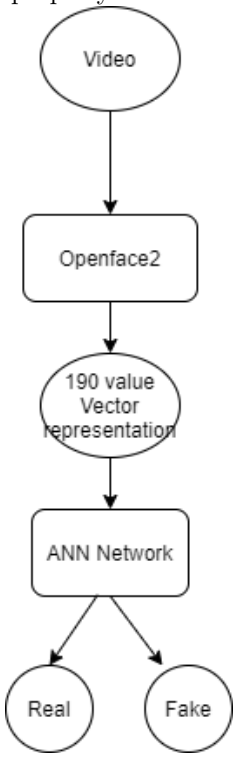

Fig 3:Representing High Level Block diagram

We will use the open-source facial behavior analysis toolkit OpenFace2 [5] to extract facial and head movements in a video.OpenFace2 provides the library provides 2-D and 3-D facial landmark positions, head pose, eye gaze, and facial action units for each frame in a given video.
The OpenFace2 toolkitprovides the intensity and occurrence for 17 AUs: inner brow raiser (AU01), outer brow raiser (AU02), brow lowered (AU04), upper lid raiser (AU05), cheek raiser (AU06), id tightener (AU07), nose wrinkler (AU09), upper lip raiser (AU10), lip corner puller (AU12), dimpler (AU14), lip corner depressor (AU15), chin raiser (AU17), lip stretcher (AU20), lip tightener (AU23), lip part (AU25), jaw drop (AU26), and eye blink (AU45). Eyeblink is not useful so we will element it. We will add some extra AU like head rotation around the x-axis, Head rotation about the z-axis, etc. Now we will make 20 AU are formed for the given video.
We will use the Pearson correlation to find a linearity between them as shown in figure 4. like inner brow raiser with lip tightener etc . so total 20 AU and chossing two at a time while gives 190 vector values for each video. As we know that These 190 features are so much powerful for Authentication and Deep Fake Detection.
In real video of Obama the brow is saying and in similar way lips will move but in the deep fake these properties get voided, the brow will not move in similar ways as lips will move hence these 190 vector representation are important for the deepfake detection.

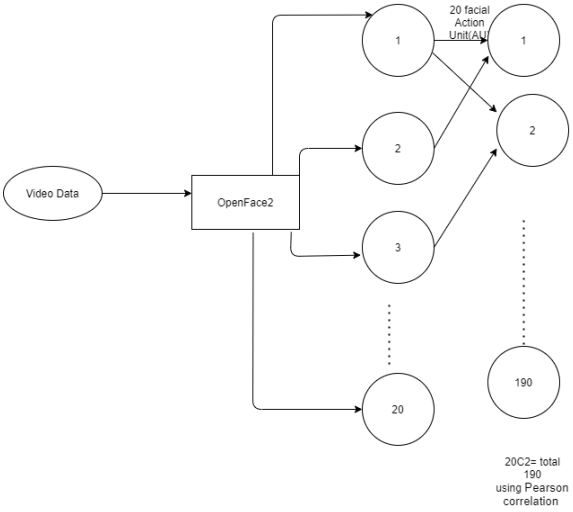

Fig 4: Representing Basic Block daigram to generate vector data from video

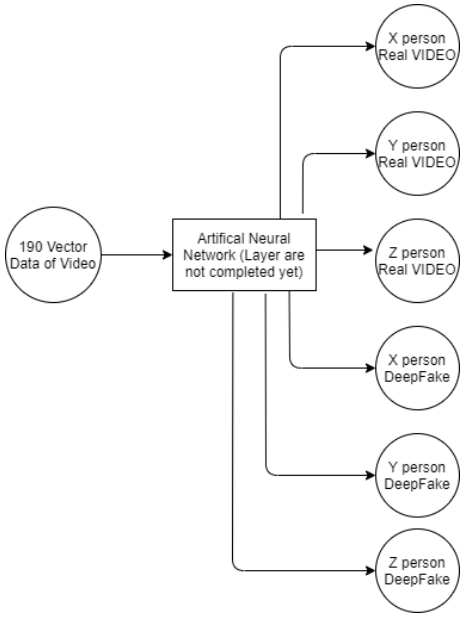

Fig 5: Vector data to classifier output

After this, we will make an Artificial Neural Network Multiclass Classifier that will differentiate between a real or a fake video. The purpose is not only to detect Real or deepFake but given authentication that it's a real video of a given world leader. The detail of Artificial network are not completed and work is keep going on.

5 Conclusion

Currently, we are in the designing phase of the model, We are converting a video into 190 vector information used for Detecting deep fake. We are currently working on this conversion. The method will be useful for detecting deep fake for leaders and celebrities who are having a large number of videos available. It can be used for authentication.

References

[1] *DeepfakeFrom Wikipedia, the free encyclopedia.* From Wikipedia.

[2] Aja Romano. *"Jordan Peele's simulated Obama PSA is a double-edged warning against fake news"*. 18 April 2018.

[3] Francesco Cavalli Henry Ajder, Giorgio Patrini and Laurence Cullen. *"The State of Deepfake - Landscape, Threats, and Impact"*. . Deeptrace., first edition, 1 October 2019. Retrieved 7 July 2020.

[4] S. Agarwal H. Farid Y. Gu M. He K. Nagano and H. Li. Protecting world leaders against deep fakes. *Workshop on Media Forensics at CVPR, Long Beach*, 1(1):2–2, Nov 2019.

[5] Tadas Baltrusaitis Peter Robinson and Louis-Philippe Morency. Openface: an open source facial behavior analysis toolkit. *In IEEE Winter Conference on Applications of Computer Vision*, I(I):1–10, 2016.

[6] Tadas Baltrusaitis Peter Robinson and Louis-Philippe Morency. Deepfake video detection using recurrent neural networks. *Video and Image Processing Laboratory (VIPER), Purdue University*, 2016.

[7] Yuezun Li and Siwei Lyu. Exposing deepfake videos by detecting face warping artifacts. *Computer Science Department University at Albany State University of New York USA*, 2016.

[8] Deep learning for deepfakes creation and detection: A survey.

[9] Jason Brownlee. A gentle introduction to generative adversarial networks (gans). *Jason Brownlee on Generative Adversarial Networks*, 2017.

[10] Emil Johansson. Detecting deepfakes and forged videos using deep learning.