Deepfake Detection Using Deep learning

A project report,
submitted in partial fulfillment of the requirements for B. Tech project

by

Kammari Santhosh (2018IMT-043)

Under the Supervison of

Dr.Rajendra Sahu
Dr.Manoj Dash

विश्वजीवनामृतं ज्ञानम्

ABV INDIAN INSTITUTE OF INFORMATION TECHNOLOGY AND MANAGEMENT
GWALIOR-474 015
2021

CANDIDATES DECLARATION

I hereby certify that the work, which is being presented in the report, entitled **Deepfake Detection using Deep learning**, in partial fulfillment of the requirement for the award of the Degree of Bachelor of Technology and submitted to the institution is an authentic record of our own work carried out during the period *June 2021* to *october 2021* under the supervision of **Dr.Rajendra Sahu** and **Dr.Manoj Dash**. I also cited the reference about the text(s)/figure(s)/table(s) from where they have been taken.
Date:14-08-20201

K. Santhosh

Signature of the candidate

ABSTRACT

Ranging from big data analytics to human-level control and computer vision, we can successfully apply deep learning for solving complex problems. However, profound knowledge proceeding to generate software can create warnings to public and private security. Deep fake is one of these freshly emerged deep fake-powered applications. Using Deepfake algorithms, one can generate artificial images and videos that humans can't distinguish from actual ones.

The purpose of technologies that identify and estimate the honesty of digital media is therefore needed. Nowadays, people face AIsynthesized face-swapping video, which is an emerging difficulty known as DeepFakes, which will lead to creating threats to fraudulence and privacy. Sometimes regular and good quality DeepFake video recognition could be challenging to detect with a natural eye. We need to improve an algorithm or technology to choose whether a photo was replaced with DeepFake technology or not with an average accuracy score.Much study and analysis is dedicated to develop and improve detection methods to overcome the potential negative influence of deep fakes. Neural networks and deep learning's application is one way.

Index Terms: Deep Fakes, artificial intelligence, deep learning, computer vision, DeepFake detection

ACKNOWLEDGEMENTS

We are highly indebted to **Dr. Rajendra sahu** and **Dr. Manoj Dash**, and are obliged for giving us the autonomy of functioning and experimenting with ideas. We would like to take this opportunity to express our profound gratitude to them not only for their academic guidance but also for their personal interest in our project and constant support coupled with confidence boosting and motivating sessions which proved very fruitful and were instrumental in infusing self-assurance and trust within us. The nurturing and blossoming of the present work is mainly due to their valuable guidance, suggestions, astute judgment, constructive criticism and an eye for perfection. Our mentor always answered myriad of our doubts with smiling graciousness and prodigious patience, never letting us feel that we are novices by always lending an ear to our views, appreciating and improving them and by giving us a free hand in our project. It's only because of their overwhelming interest and helpful attitude, the present work has attained the stage it has.
Finally, we are grateful to our Institution and colleagues whose constant encouragement served to renew our spirit, refocus our attention and energy and helped us in carrying out this work.

TABLE OF CONTENTS

List of Figures

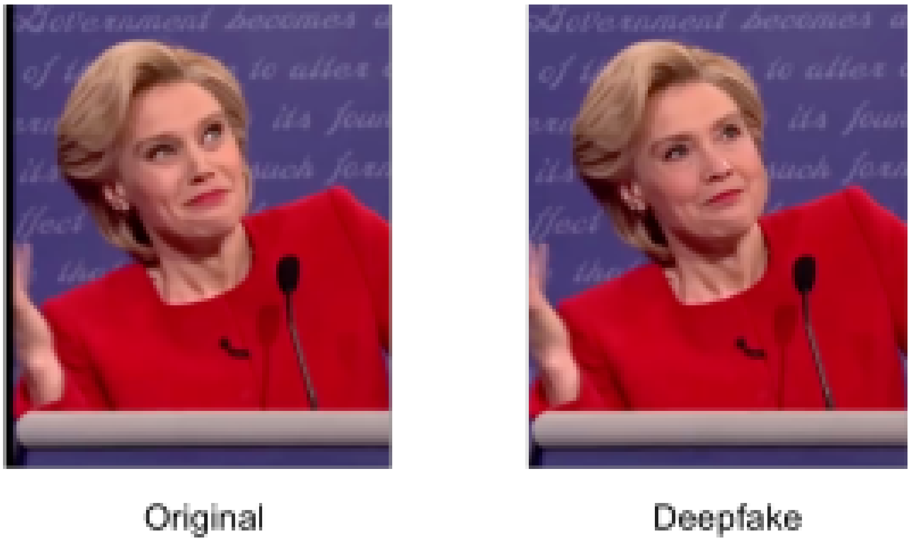
List of Tables

1 INTRODUCTION

Deepfake is a method that will map faces of a targetted person to a video of a related or actual person to generate a video of a targetted person making the same actions or speaking things the first person does. These models analyze a person's facial emotions or expressions and activities and combine facial pictures of different people performing the same terms and actions.

Deepfake systems typically need various photos and video data to develop models to create realistic videos and images. Higher deep networks and freely accessible data have made fake videos and images almost equal to humans and even advanced computer algorithms.

now a days the popularization of smartphones and increase in social networks have made digital images and videos become more common digital objects .according to several reports almost two million to billion pictures are uploaded everyday in internet this use of digital images has been followed by a rise of techniques to change image contents using some softwares like Photoshop for example. in order to regulate the circulation of falsified contents the field of digital image forensics research is dedicated for detection image forgeries. many of them are like analyse and construct inconsistencies relative to what a normal camera would be which will release on exception of specific image alternating in the resulting images. among others like image noise has been shown be a good indicator for detecting splicing which is a copy paste from an image to another form of image. to find out the image manipulation the detection of image compression is much needed.



Original

Deepfake

Figure 1: original vs Deepfakes

today the fake news is widely acknowledged and in a context aware II II video contents are watch daily e on social platforms, the spread of these fake videos have been raised more and more concern to everybody. while there is a significant improvement for image forgery detection and also digital Vidya falsification detection which still remains as a difficult task. most of the methods used with images cannot be directly used for videos because of the the strong and degradation of frames after video is compressed. the current video forensic studies focuses on video recapturing as well as re-encoding but the the challenge of video edition still remains.

over the past two decades deep learning methods has been successful e e deployed for image forensics. such as deep learning to locally detect JPG compress on images and network to detect image splicing and any image general falsification and distinguish computer graphics from photographic images so with this deep learning performs very well in all types of digital for ansys as well as traditional signal processing approaches.

as we all know that deep learning can also be used to fast five videos recently good and powerful tool which is named as deep fake has been designed for face capturing and also face re-enactment. normally this kind of methods are used for creation of some unusuall content,And this has not been presented in any academic Publications . Dip Sheikh follows face to face which is a non deep learning method introduced by tyres for all which which has a similar goal that is the target is for using more conventional real time computer vision techniques for false video generation or detection.

in this project we are going to discuss about the problem of detecting the video editing passes and finding whether the video is a real. sections 1.1 and 1.2 will discuss more about defect and also take video forgeries. In section 2 we will discuss about several deep learning networks. injection 3 willfully dive deeper into the evolution of this network as well as the data set assembled for the project.

1.1 Deepfake images:

deep fake,which is a technique that need to replace the face of a person which is targeted By the face of someone else in a video it is first appeared in autumn 2017 script used to generate a page contents. after work this technique is much more improved and by small complete communities which is notably created a user-friendly application which we now call as a fake app.

the main Idea lies in the training of two autoencoders. the architecture can vary according to the input and output size for a desired training time and also the quality of the video and also the available resources stop the main purpose of the encoder is to perform the the dimension reduction by encoding the data of the input video. In the case of defect original autoencoder image resolutions of 64 X 64 x 3 ,

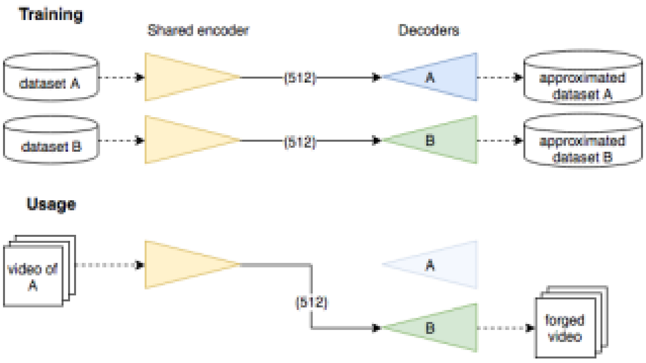


Figure 2: Encoding of dataset

the main process of generating this kind of d fake images is collecting the faces of different peoples say A and b then the with the help of auto encoder reconstruct the faces of from the data set of and with autoencoder to reconstruct faces of bi from dataset of b.

with implementation the results are quite good, resources how far how much popular the technique is,

1.2 Deepfake video:

the last and final step we have to do is to extract the images from the video with a certain frame rate and with the help of autoencoders will generate other faces with the same expressions and merged back to the video.

unfortunately with this technique we won't get a pretty good result. because the extraction of the images from the video with a certain frame rate can cause the distortion [6]in images also I will not able to collect the whole information of that picture. and also reintegration can fail. exam frames Canada with no face and with a large area are also some double faces. sorry are not following these techniques and where our to avoid this kind of we will use some advances networks more deeply this is a another application that will tend to a deep neural network.// However, when concept of neural networks grew big, and understanding improved its computational abilities, people started using this technology in their daily lives[1]. Example of excellent quality fraudulent photo with face swapping algorithm applied When it all got to that, researchers chose to find methods of deep face detection to help people

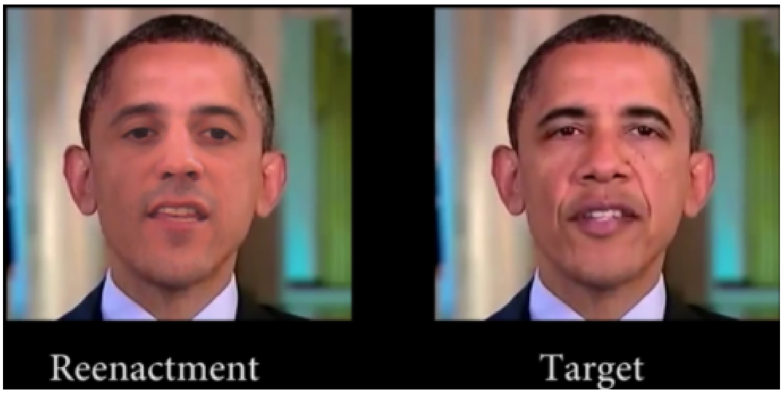


Figure 3: Reenactment of target image

from this danger. Video manipulation and Digital image technologies are developing quickly for many decades. The evolution of the model's evaluations confirms the high quality of precision in identifying fake and real videos

2 MOTIVATION

As deepfakes are fake videos generated using some digital software or face swapping ,in which computer creates an AI video[2] which lead to new video that shows the actions that wasn't happened.but the results looks like acceptable.they are so difficult to identify that they are false

the advancement in technology has not only filled media creation today but also provide a lot of opportunities for the people. however, the media technology is available which does the risk for exploiting the main such aspect is a deep fake. debates are evolving very unknowingly to a high range which created a social discord, increased polarization, and in some cases influence the outcome of elections inflicted damage to every single individual as well as institution, business, democracy.

Developing these techniques made a great impact on social and politics.The real danger is because of false information it creates mistrust or lack of interest in people about what they hear in online. these problems may leads to worst conditions.facial reanimation and using different audio tracks[4] are some of them this is why deepfake detection is needed.

3 LITERATURE SURVEY

Author	Title	Year	Publish	Work
Karnouskos, Stamatis	Digital Media: The Era of Deepfakes	2020	IEEE	considering the how deepfakes effects the media
Nguyen, Thanh and Nguyen, Cuong M. and Nguyen	Deep Learning for Deepfakes Creation and Detection: A Survey	2019	IEEE	the objective is survey of how to detect and create a deep fake
Meng, Cunyan and Zhang, Xinghui	Video Encryption Based on OpenCV	2010	IEEE	the objective here is to generate a dataset from input video using opencv
Morozov,Artem A.,Maksutov, and	Deepfake Detection Methods using Machine Learning	2020	IEEE	the objective here is to create fake video using gan and model experiments
Piotr Kawa and P. Syga	Deepfake Detection with Low-Resources	2020	IEEE	deep fake detection experimentst using mesonet-4 and mesonetinception -4
Pokroy, Artem A. and Egorov, Alexey D	DeepFake Detection: Comparing Pretrained Models	2021	IEEE	the study conducted is to compare the accuracies of various deep network models
Mahamkali, Naveenkumar and Ayyasamy	OpenCV for Computer Vision Applications	2015	IEEE	feature detection and object tracking using opencv tool

Table 1: Summary of Related Work

4 SETUP AND TOOLS

- Operating System: Windows 10
- Required Software
  - Python 3.0
  - Google Collaboratory: An Online Python IDE for Machine Learning
  - Creately: A website providing tools to build the System Architecture and other related
  - Lucidchart: A website providing tools to build Gantt Chart designs
- REQUIRED LIBRARIES:
  - Python
  - Numpy
  - Pandas
  - Seaborn
  - Sklearn
  - CPU or GPU(recommended)

5 OBJECTIVES

- In this system,a pretrained model is used on training set and model with best accuracy is used to determine wheather the video is fradulent or non-fraudulent.
- to train the model,dataset must be retrieved and data must be preprocessed in following steps:
  1. extract frames from video
  2. filtering the extracted frames
  3. splitting of data into train and test
- to determine the fradulent image using Meso model.

6 SYSTEM ARCHITECTURE

The final purpose of this work is to recognize whether a video obtained real or knowing whether it was created using some other deepfake methods[5]. As the input to the model or the system is a video
But, deep learning models do take pictures as input, so there needs to be a change in the input to model from the system[3]. This can be achieved through pre-processing.

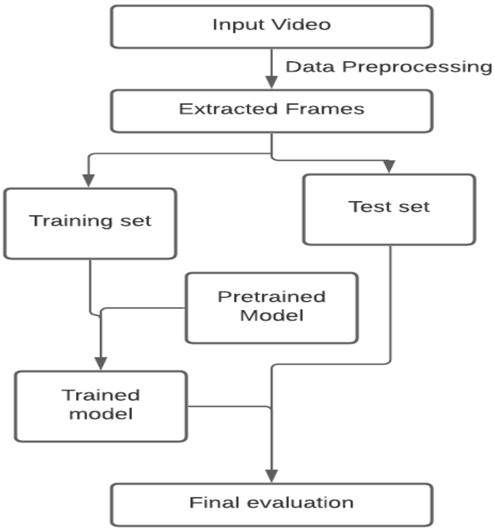


Figure 4: workflow of project

Preprocessing: In addition to the input data transformation,the preprocessing

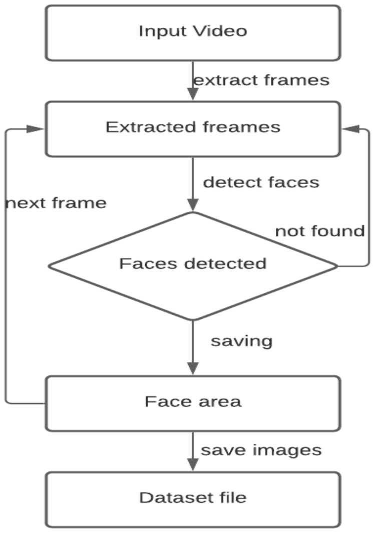


Figure 5: Data Preprocessing

module need to check how other factors impact on the model training.Most of the video frames not only contains faces but also person's body parts and background area of the image.Thus unrelated data can highly impact our model training.
So we need to focus on the face area and our preprocessing module need to capture the face as input.preprocessing consists of three steps:

- collect frames from the video
- detecting faces from the individual images.
- resizing and saving locally.

7 TASKS COMPLETED

- With the help of computer vision library successfully able to extract frames(images) from video.
- face detection,image size reduction able to fit the data for the model.
- implemented the code and visualized the output using mesonet model.

8 RESULTS

Importing the dataset and required libraries,the data is fetched using a opencv library which contains a video.the data is analyzed and preprocess in series of steps:

extraction of frames: with the help of opencv tool(imread,imwrite) we can extract the frames from the input video with certain frame rate depending how no of images we want to store.

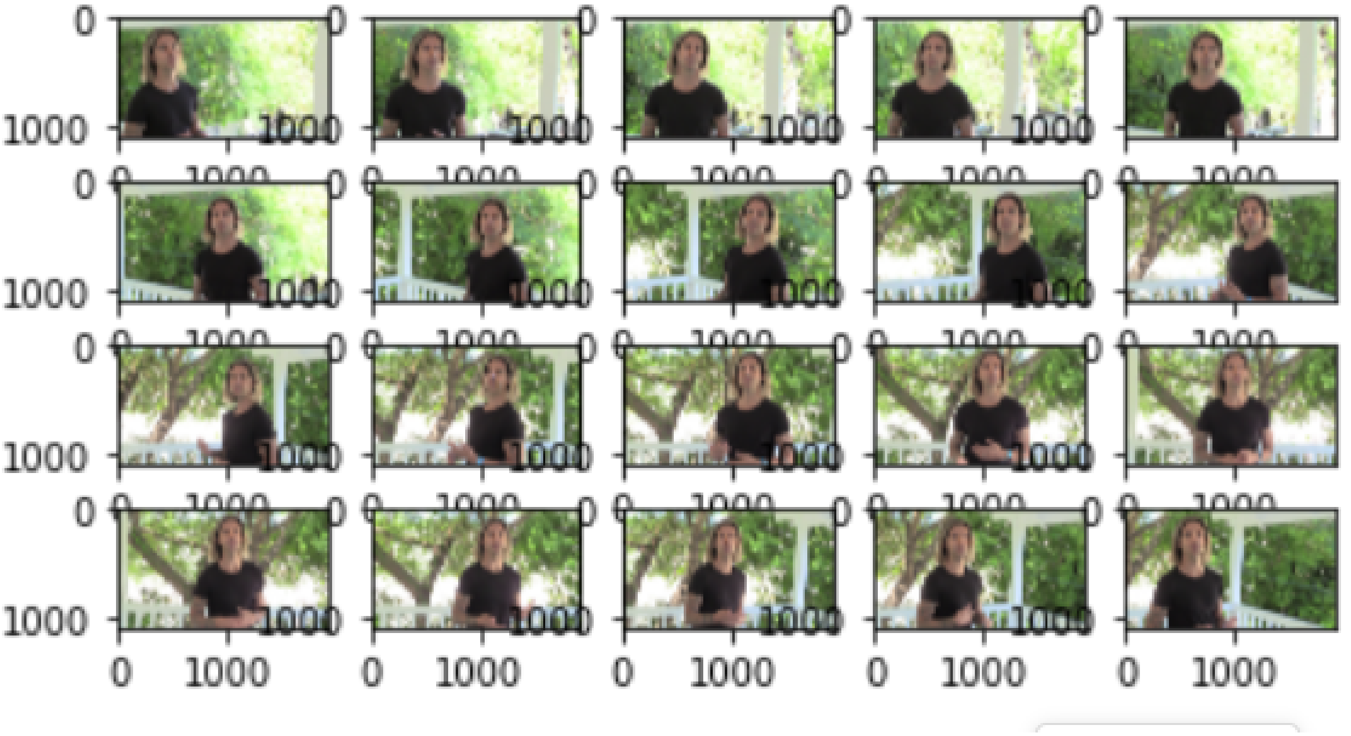
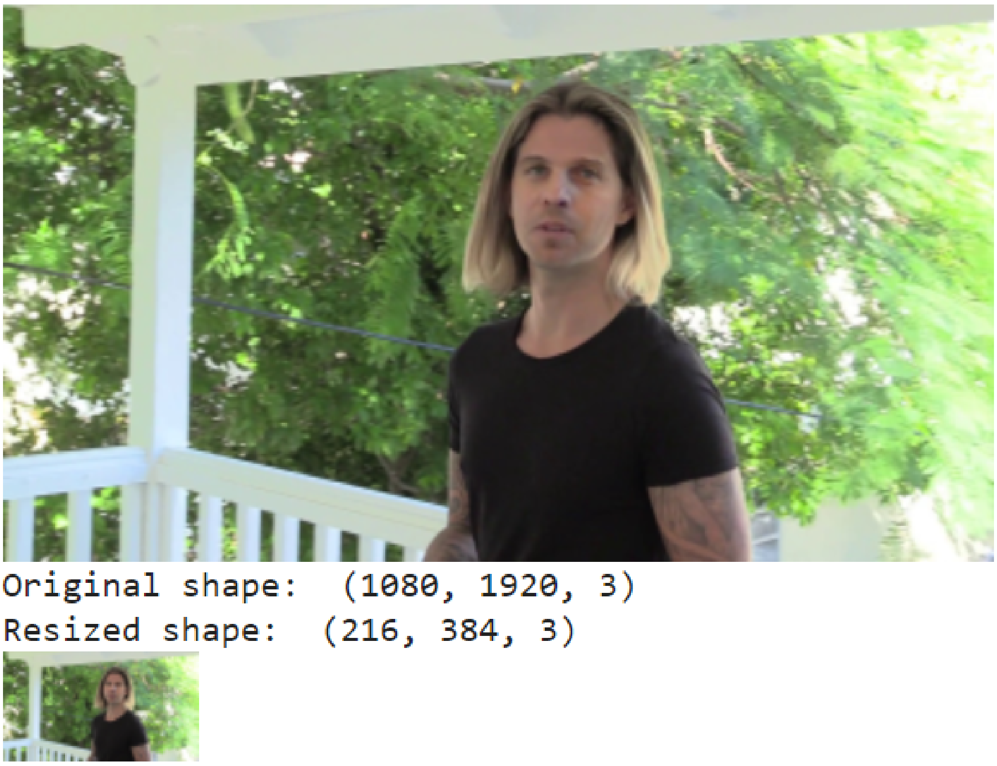


Figure 6: extracted images

Every extracted image(frame) is stored in a folder, which we can use for future processing.

size reduction: the dimensions of each of the image that we have extracted from video are much large, so for better output and to take less time we have scale down each image.



```
Original shape:  (1080, 1920, 3)
Resized shape:  (216, 384, 3)
```

Figure 7: size reduction

face identification: Now that we have resized our image, Apply the Cascade-Classifier function, an OpenCV function, to an XML file location for frontal face detection, i.e., haarcascade_frontalface_default.XML, in our project. I have used path in the machine that I have downloaded.
Generally the images are of RGB channel that we see. But OpenCv reads the image in BGR format, so we have to convert into gray channel because it is simple to process also takes less time for computing as it holds just one channel of white-black.

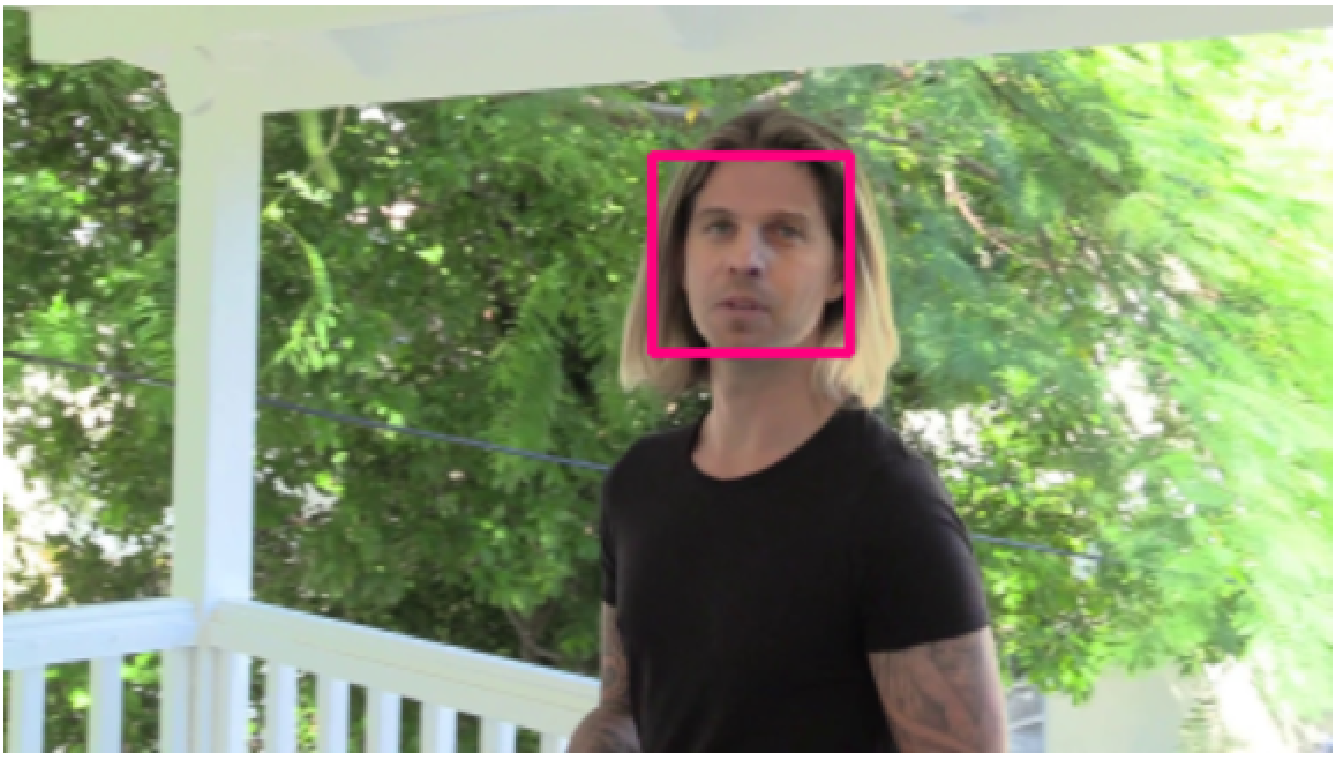
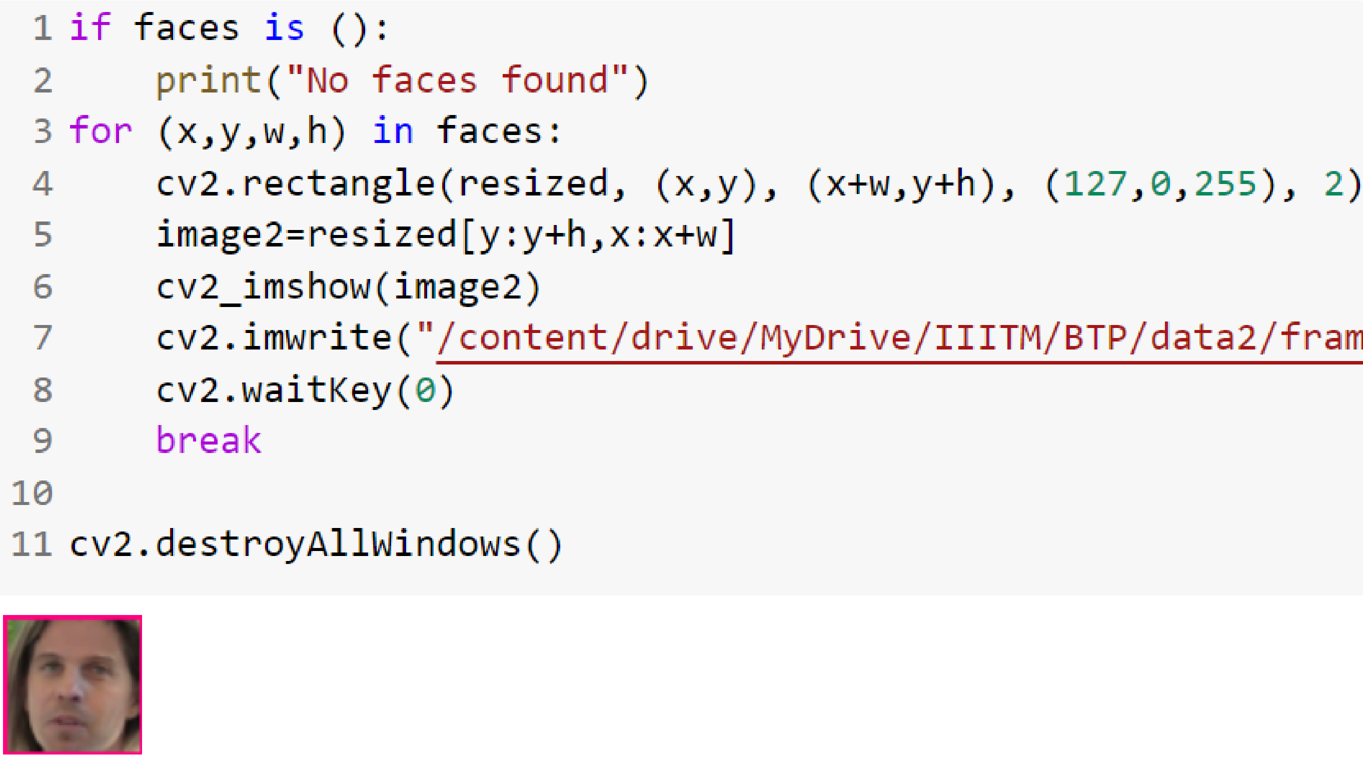


Figure 8: marking the face area

With the help of function detectMultiScale,we get the coordinates of face,as x-coordinate,y-coordinate,width and height.With the help of x and y coordinated a sqaure is drawn.

modifying the data: for an image if we found a face then we will replace it with updated image else remove the image from dataset which is initially created.



```
if faces is ():
    print("No faces found")
for (x,y,w,h) in faces:
    cv2.rectangle(resized, (x,y), (x+w,y+h), (127,0,255), 2)
    image2=resized[y:y+h,x:x+w]
    cv2_imshow(image2)
    cv2.imwrite("/content/drive/MyDrive/IIITM/BTP/data2/fram
    cv2.waitKey(0)
    break

cv2.destroyAllWindows()
```

Figure 9: data modification

With the mesonet network the prediction of image is to be done.so we know that we can extract images from a video but we have 120 images from a single video.so i have used a dataset file consist of 7000 images which have both deepfake and real.we used that dataset for training and testing the model.

Model confidence greater than 0.5 are said to be real and lower are deepfake according to model.
with the below real images every model confidence is satisfied.

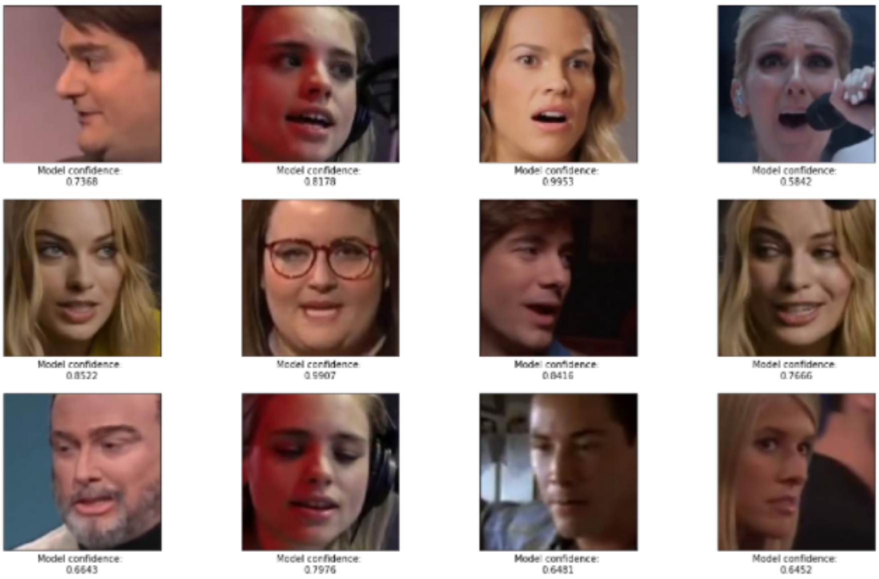


Figure 10: Model confidence

9 REFERENCES

[1] Stamatis Karnouskos. “Artificial Intelligence in Digital Media: The Era of Deepfakes”. In: *IEEE Transactions on Technology and Society* 1.3 (2020), pp. 138–147. DOI: 10.1109/TTS.2020.3001312.

[2] Piotr Kawa and P. Syga. “A Note on Deepfake Detection with Low-Resources”. In: *ArXiv* abs/2006.05183 (2020).

[3] Artem A. Maksutov et al. “Methods of Deepfake Detection Based on Machine Learning”. In: *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. 2020, pp. 408–411. DOI: 10.1109/EIConRus49466.2020.9039057.

[4] Cunyan Meng and Xinghui Zhang. “Video Encryption Based on OpenCV”. In: *2010 2nd International Workshop on Database Technology and Applications*. 2010, pp. 1–4. DOI: 10.1109/DBTA.2010.5658976.

[5] Thanh Nguyen et al. *Deep Learning for Deepfakes Creation and Detection: A Survey*. Sept. 2019.

[6] Artem A. Pokroy and Alexey D. Egorov. “EfficientNets for DeepFake Detection: Comparison of Pretrained Models”. In: *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*. 2021, pp. 598–600. DOI: 10.1109/ElConRus51938.2021.9396092.