DEEP FAKE

Detection

ABHIJITH T S

C0PSCS2102

Guided by Mrs. LISHA A

Table of contents

Introduction

- Deep fake is a technique for human image synthesis based on artificial intelligence
- Deep fakes are created by combing and superimposing existing images and videos onto source images or videos using a deep learning technique known as generative adversarial network.
- Deepfake detection system allow computers to identify these kind of fake media

Problem statement

- To Design and Develop a Deep Learning algorithm to classify the video as deepfake or pristine.

Objectives

- The main objective of this system is to predict the media novelty
- Current deepfake generation systems are so accurate that human naked I can’t detect a fake video
- The system must be able to detect these media with high confidence rate

Motivation

Real

Fake

Fake

Fake

Fake

Real

Existing methods

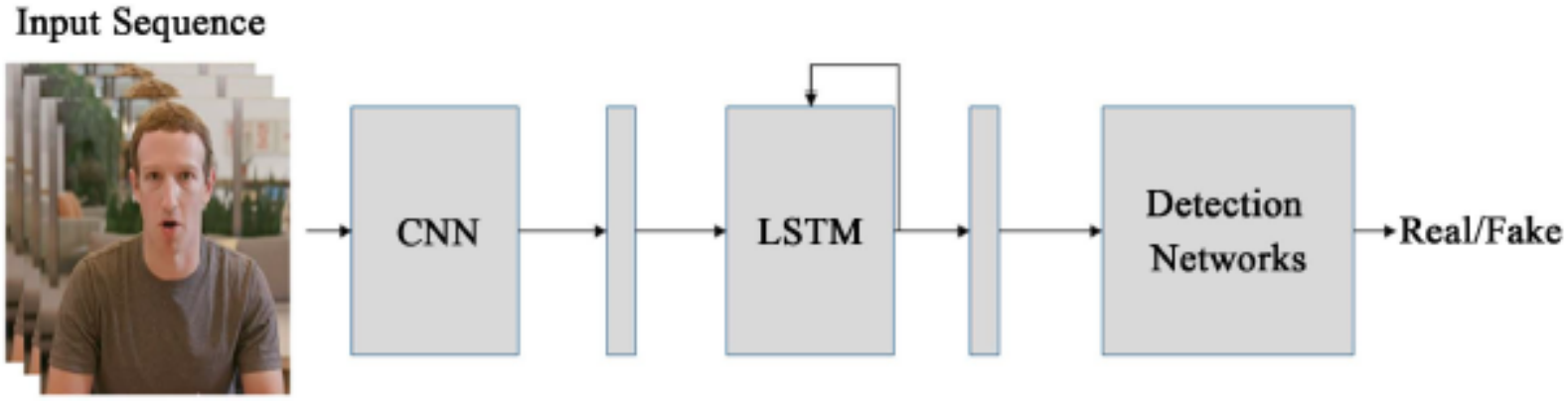
- **Detection using eye blinking**

[1]Yuezun Li presented a new approach based on natural network to detect Fake Face Videos. Compared with the previous work, this method considers eye blinking to detect fake videos, which is an important physical feature that can be used to distinguish the fake videos.

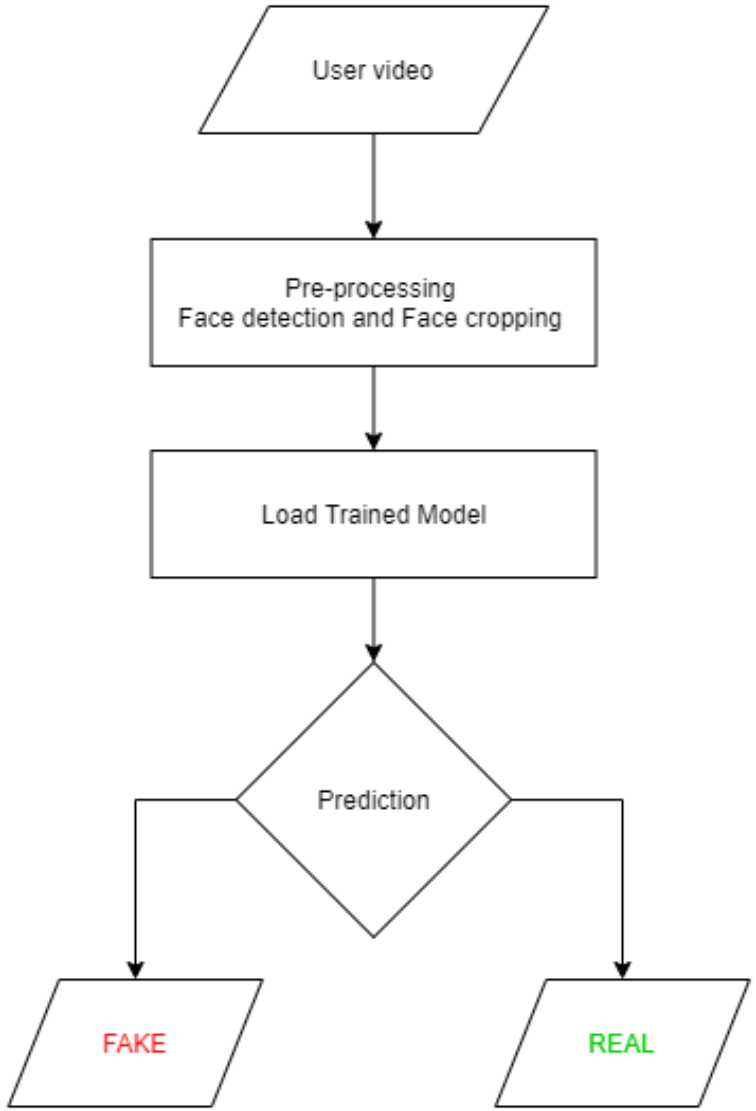
- **analyzing the “heartbeat” of deep fakes.**

[2]Ciftci et al.for example has designed a Generative Adversarial Network (GAN) based model that can detect the deepfake video source by analyzing the “heartbeat” of deep fakes. The proposed model starts by having several detector networks where the input to this model is the real video.

Design



- The design implements a CNN for features extraction and a LSTM network for prediction [3][4]


User video
Pre-processing
Face detection and Face cropping
Load Trained Model
Prediction
FAKE
REAL

Implementation steps

- Download dataset
- Preprocessing dataset and corresponding labels
- Loading the data and train the models
- Plotting test accuracy
- Uploading video file for prediction
- Getting classification result with confidence rate

Methodology

- **Dataset used**

The primary data set used was provided by Kaggle[5] which contain 400 GB of data, selected 400 train videos and 400 test videos

- **Algorithms used**

A CNN called resnext50 which is used to extract features from each frames. Then this data is fed in to a LSTM for prediction

- **Libraries used**

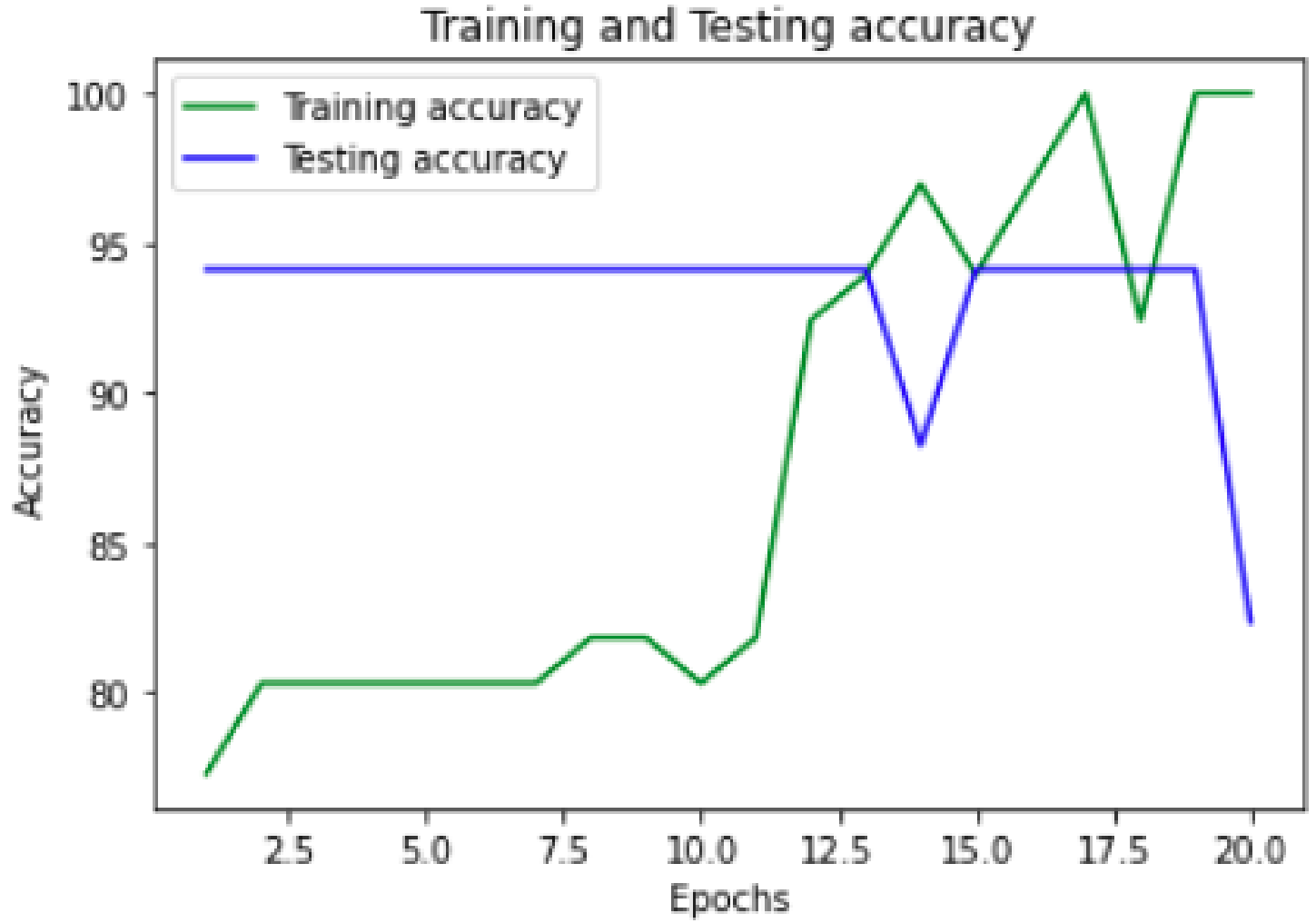
Pytorch is an open source machine learning (ML) framework based on the Python programming language and the Torch library. It is one of the preferred platforms for deep learning research. The framework is built to speed up the process between research prototyping and deployment.

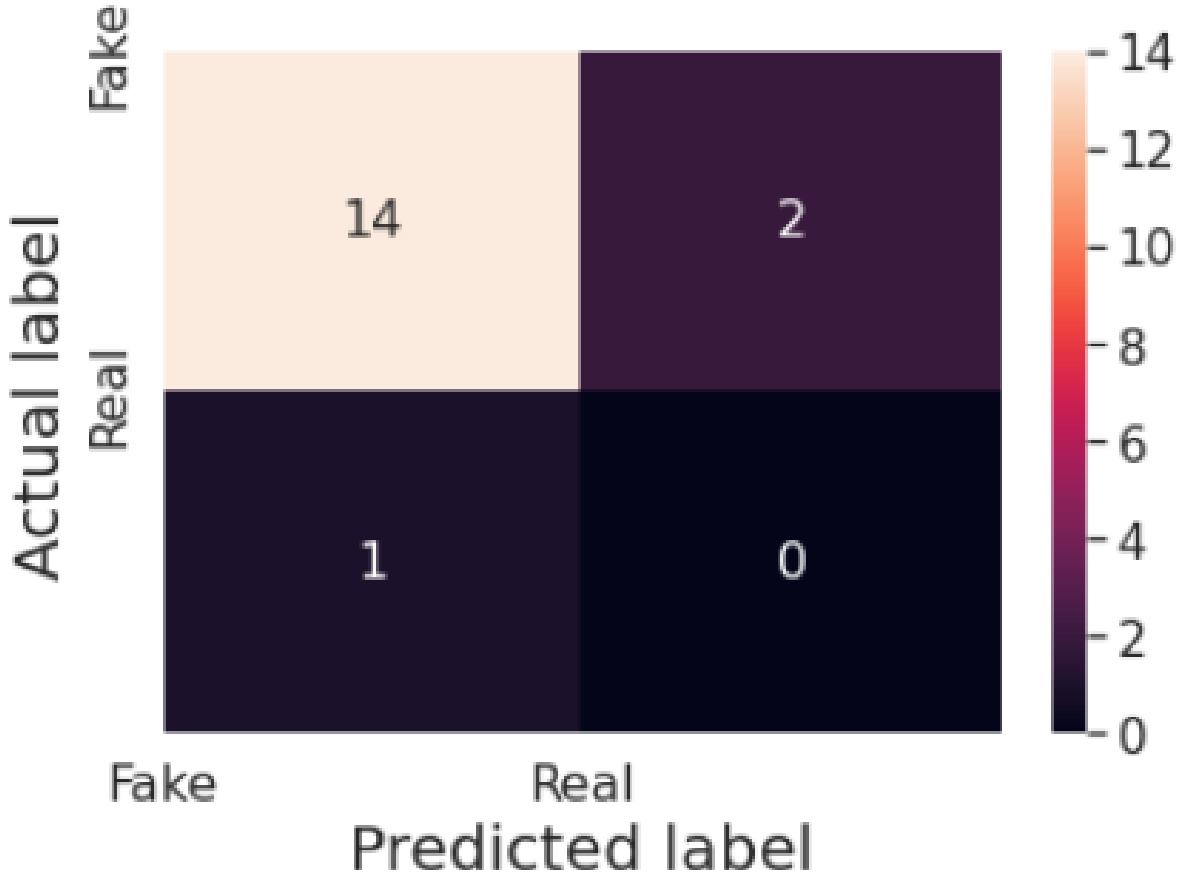
- **Face Recognition**

. Recognize and manipulate faces from Python or from the command line with the world’s simplest face recognition library

Result and analysis

- Plots


Training and Testing accuracy
Training accuracy
Testing accuracy
Accuracy
100
95
90
85
80
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epochs

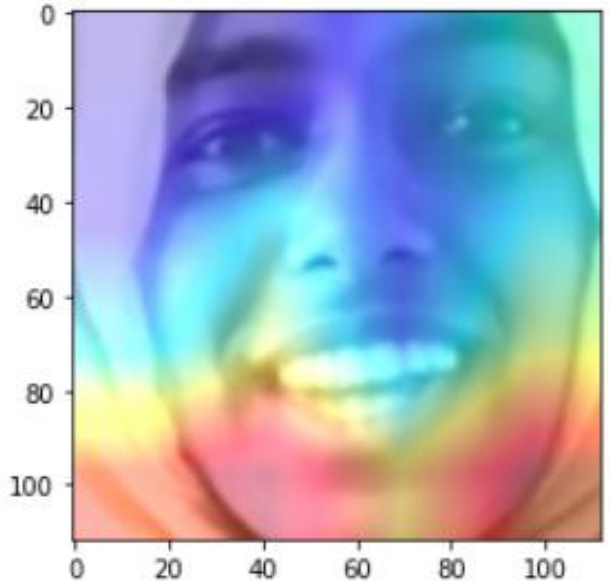


True positive = 14
False positive = 2
False negative = 1
True negative = 0

Calculated Accuracy 82.35294117647058

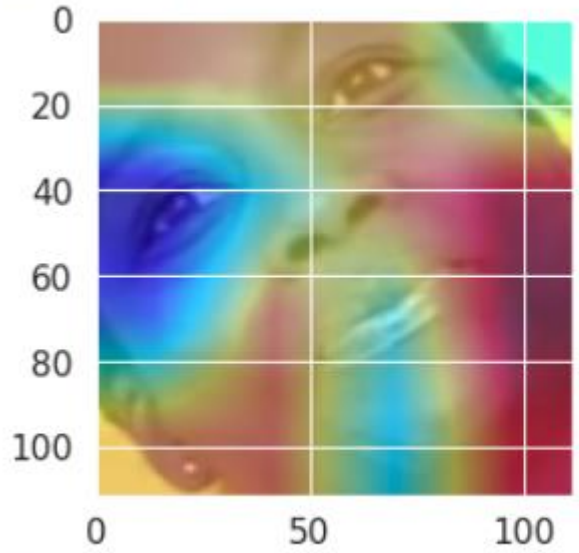
- Predictions

confidence of prediction: 51.155149936676025



FAKE

confidence of prediction: 82.90736675262451



REAL

Conclusion

- The implemented algorithms were able to predict real and fake videos correctly.
- Got an average accuracy of 82 %
- Day by day deepfakes are improving thus we need to train the model for new fakes and improve it's accuracy

References

- [1] Li, Y., Chang, M.-C. and Lyu, S. (2018) In Ictu Oculi: Exposing AI Generated FakeFace Videos by Detecting Eye Blinking. 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, 11-13 December 2018, 1-7. https://doi.org/10.1109/WIFS.2018.8630787
- [2] Ciftci, U.A., Demir, I. and Yin, L. (2020) How Do the Hearts of Deep Fakes Beat? Deep Fake Source Detection via Interpreting Residuals with Biological Signals. 2020 IEEE International Joint Conference on Biometrics (IJCB), Houston, 28 September- 1 October 2020, 1-10. https://doi.org/10.1109/IJCB48548.2020.9304909
- [3] Güera, D. and Delp, E.J. (2018) Deepfake Video Detection Using Recurrent Neural Networks. 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, 27-30 November 2018, 1-6. https://doi.org/10.1109/AVSS.2018.8639163
- [4] How to cite this paper: Almars, A.M. (2021) Deepfakes Detection Techniques Using Deep Learning: A Survey. Journal of Computer and Communications , 9, 20-35. https://doi.org/10.4236/jcc.2021.95003 College of Computer Science and Engineering, Taibah University, Yanbu, Saudi Arabia
- [5] https://www.kaggle.com/competitions/deepfake-detection-challenge/data

THAK YOU