# Deepfake Detection

Ahmed Hesham, Gehad Mohamad, Mostafa Ebrahim, Mostafa Abdelaziz, Robert Ayman, Samaa Hamdan

Dr. Ezz Eldin Badawy

Faculty of Computers & Artificial Intelligence

Benha University

## ABSTRACT

In recent months, free deep learning-based software tools have facilitated the creation of credible face exchanges in videos that leave few traces of manipulation, in what they are known as "Deepfake“ videos.

Creating the DF using artificially intelligent tools is a simple task. But, when it comes to detection of these DF, it is a major challenge, because training the algorithm to spot the DF is not simple.

We have taken a step forward in detecting the DF using EfficientNetB0. The system uses EfficientNetB0 to extract features at the frame level. These features are used to classify if a video has been subject to manipulation or not and can detect the visual artifacts within frames introduced by the DF creation tools

## OBJECTIVES

The project aims to minor imperfections found in doctored videos and exposed. The creation of a doctored video needs information to be examined in order to reveal the dishonest, particularly facial expression variables.

The minor imperfections could be boundary points, background incoherence, double eyebrows, or irregular twitch of the eye.

## MATERIAL & METHODS

**Neural network type:**

Use EfficientNetB0 convolutional neural network architecture and scaling method to classify the videos into real and fake videos.

**Procedures**:

### Database

First, the database, we are using a mixed dataset called FaceForensics++ videos and 140k images.

Videos are 10 seconds long on average with 30 fpm frame rate. The average frame dimension is 720*1280. Our newly prepared dataset contains 50% of the original video and 50% of the manipulated deepfake videos. The dataset is split into 80% train and 20% test set.

Image dataset consist of 140k Real and Fake Faces, all the images are resized into 256px. The dataset is split into 80% train and 20% test set

### Preprocessing

We have used “cv2” to split videos in the dataset to improve the diversity and richness of the image dataset. One frame image is extracted every 30 frames when the video is divided into frames.

Face recognition and cropping work are performed on the video frame images. Then, the new cropped face frames are created and saved.

### Model

we present our end-to-end trainable re-current deepfake video and image detection system. The proposed system is composed of an ExceptionNetB0 network followed by a 2D average pooling layer.

The network is similar to the traditional convolutional neural network.

### Predict

A new video/image is passed to the trained model for prediction. A new video is also preprocessed to bring in the format of the trained model. The video is split into frames followed by face recognition, then frame cropping, and instead of storing the video into local storage, the cropped frames are directly passed to the trained model for detection.

The prediction is identified by using the mode of all frame's prediction. In case of image detection, it is treated as a single frame video.

## RESULTS

At first we used the ResNext deep learning model but it has a lower accuracy and consumes a lot of resources, we then use the EfficientNet deep learning model on two datasets FaceForensics++ and 140k real and fake faces. We made a combination of the two datasets and used them to train the model.

| Deep learning Model | Dataset | Accuracy |
|---|---|---|
| RESNext | Faceforensics++(500 video) | 48.0% |
| EfficientNet | Combination of 140k real and fake faces and Faceforensics++ | 96.79% |
| | Faceforensics++ | 91.72% |
| | 140k real and fake faces | 92.87% |

## CONCLUSION

We presented a deep neural network-based approach to classify the video/image as deep fake or real, along with the confidence of the proposed model. The proposed method is inspired by the way the deep fakes are created by the GANs with the help of Autoencoders.

Our method does frame-level detection using EfficientNet0. The proposed method is capable of detecting the video/image as a deep fake or real. We believe that it will provide a very high accuracy on real-time data.

## REFERENCES

[1] Yuezun Li, Siwei Lyu, “ExposingDF Videos By Detecting Face Warping Artifacts,”

[2] Nguyen, Thanh Thi, et al. "Deep learning for deepfakes creation and detection."

[3] https://www.kaggle.com/c/deepfake-detection-challenge/data

[4] https://github.com/ondyari/FaceForensics

[5] R. Raghavendra, Kiran B. Raja, Sushma Venkatesh, and Christoph Busch, “Transferable deep-CNN features for detecting digital and print-scanned morphed face images,”

[6] Yu, Peipeng, et al. "A survey on deepfake video detection."

[7]https://antispoofing.org/EfficientNet_for_Deepfake_Detection:_Architecture,_Purpose_and_Practical_Application

## AKNOWLEDGEMENT

Foremost, we would like to express our sincere gratitude to our advisor Prof. Ezz Eldin Badawy for the continuous support of our project and research, and for his patience and motivation.

His guidance helped us all the time in working on the project.

Also, we would like to thank our college BFCAI for the support and motivation for us