IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12615

Deep fake video detection through deep learning

Manish Kumar[1], Pravin Chavhan[2], Prathamesh Shingare[3],Surekha Suryawanshi[4] ,

Prof. Chetana P.Shravage[5]

Student, Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune, India[1]

Student, Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune, India[2]

Student, Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune, India[3]

Student, Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune, India[4]

Professor, Department Computer Engineering, Dr. D. Y. Patil Institute of Technology Pimpri, Pune, India [5]

Abstract*:* Over the last years, the increase in smartphones and social networks has made digital images and videos common digital objects. According to reports, almost two billion pictures are uploaded every day on the internet. This tremendous use of digital images has been followed by a rise of techniques to alter image contents, using editing software like Photoshop for instance. Fake videos and images created by deepFake techniques have become a great public issue recently. Nowadays several techniques for facial manipulation in videos have been successfully developed like Face Swap, deep Fake, etc. On one side, this technological advancement increase scope to new areas (e.g., movie making, visual effect, visual arts, etc.). On the other side, contradicting, it also increases the ease in the generation of video forgeries by malicious users. Therefore by using deep learning techniques we can detect the video is fake or not. In order to detect these malicious images, we are going to develop a system that can automatically detect and assess the integrity of digital visual media is therefore indispensable. Deepfake is a technique for human image synthesis based on artificial intelligence, i.e., to superimpose the existing (source) images or videos onto destination images or videos using neural networks (NNs). Deepfake enthusiasts have been using NNs to produce convincing face swaps. Deep fakes are a type of video or image forgery developed to spread misinformation, invade privacy, and mask the truth using advanced technologies such as trained algorithms, deep learning applications, and artificial intelligence. They have become a nuisance to social media users by publishing fake videos created by fusing a celebrity's face over an explicit video. The impact of deepFakes is alarming, with politicians, senior corporate officers, and world leaders being targeted by nefarious actors. An approach to detect deepFake videos of politicians using temporal sequential frames is proposed. The proposed approach uses the forged video to extract the frames at the first level followed by a deep depth-based convolutional long short-term memory model to identify the fake frames at the second level. Also, the proposed model is evaluated on our newly collected ground truth dataset of forged videos using source and destination video frames of famous politicians. Experimental results demonstrate the effectiveness of ourmethod.

Keywords: Deepfake, Deep Learning, Deep fake Technology, Deep fake Detection, Forensic Verification, Fake Images, Fake Image Detection,Etc.

I. INTRODUCTION

Photos and videos are frequently used as evidence in police investigations to resolve legal cases since they are considered to be reliable sources. However, sophisticated technology increases the development of fake videos, and photos that have potentially made these pieces of evidence unreliable. Fake videos and images created by deep fake techniques have been become a great public issue recently. So, predicting them becomes an important subject. A prediction that can be accurate and relied on is the need for resolve all forensic cases.It gets us ready for all the worst possible scenarios and hence we focus on understanding deep learning algorithms, the necessity apparatus as well as theory required to do so.

II. EXISTING SYSTEM

This paper introduces the Face Forensics++ dataset, which includes manipulated facial images. The authors propose a deep learning approach that learns to identify visual artifacts and inconsistencies present in deepfake images.[1]

This paper presents the Celeb-DF dataset, which contains a large-scale collection of real and deepfake videos featuring celebrities. The authors highlight the challenges associated with deepfake detection and discuss various deep learning

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12615

techniques for analyzing and detecting deep fake videos.[2]

This review article provides an overview of deep fake video detection methods based on deep learning. It covers various techniques, including CNNs, recurrent neural networks (RNNs), and generative models, along with their advantages and limitations.[3]

This paper focuses on the specific problem of deep fake videos targeting world leaders. The authors propose a deep learning framework that detects facial manipulations in videos to safeguard the integrity of public figures.[4]

This paper presents MesoNet, a lightweight deep learning network designed for detecting facial video forgeries, including deep fake videos. The network utilizes spatiotemporal information from facial regions to classify manipulated videos.[5]

This paper present Investigates the use of capsule networks, a type of deep neural network, for detecting deepfake videos based on learned hierarchical features.[6]

III. PROPOSED SYSTEM

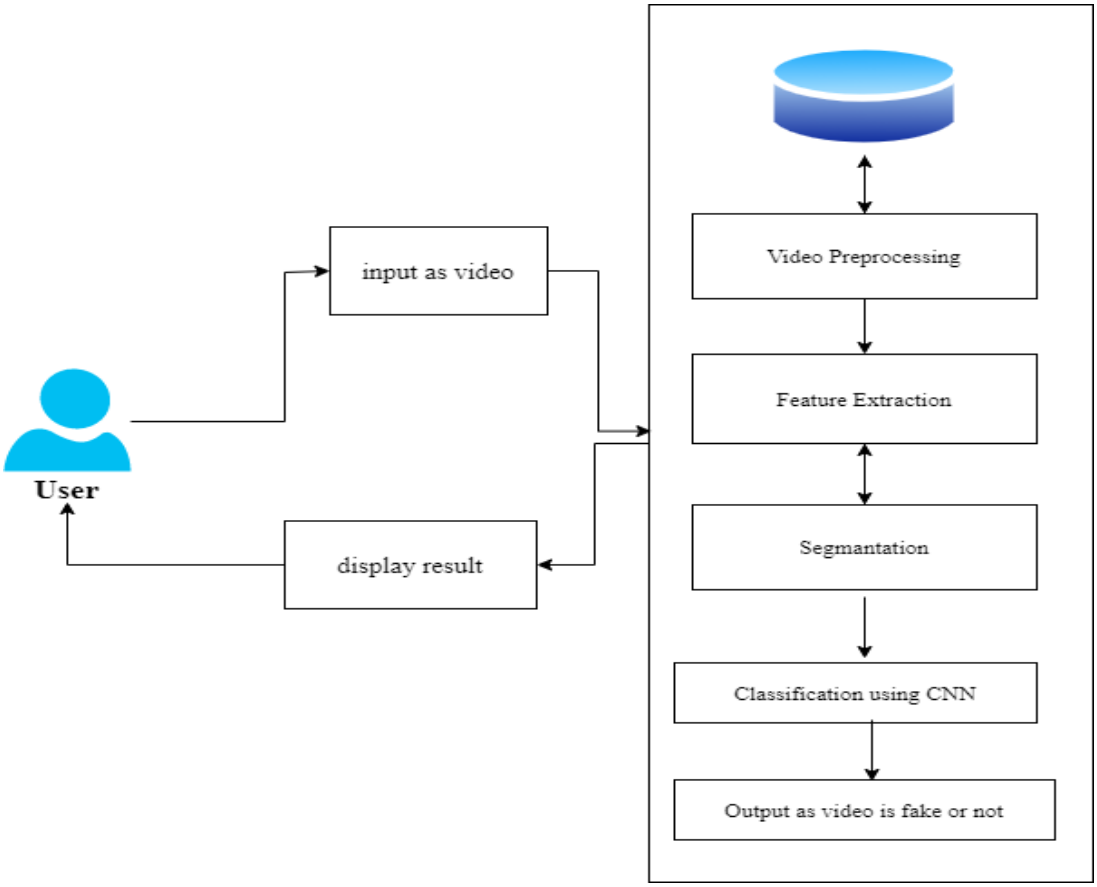


Fig 1: System Architecture

As fig.1 shows the flow of system and its divided into multiple modules. These module are defined below:
[1].Video Preprocessing: The video is preprocessed to extract individual frames and audio segments. These frames and audio segments are then passed through the deep learning model for analysis.
[2].Feature Extraction: Extract features from the pre-trained deep learning model. These features can include visual cues such as facial landmarks, eye movement, and inconsistencies in lighting, as well as audio cues like speech patterns and voice characteristics.
[3].Segmantation: Choose a deep learning model suitable for segmentation tasks. Commonly used models include U-Net, Mask R-CNN, or DeepLab. These models are designed to identify and segment regions of interest within an image.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

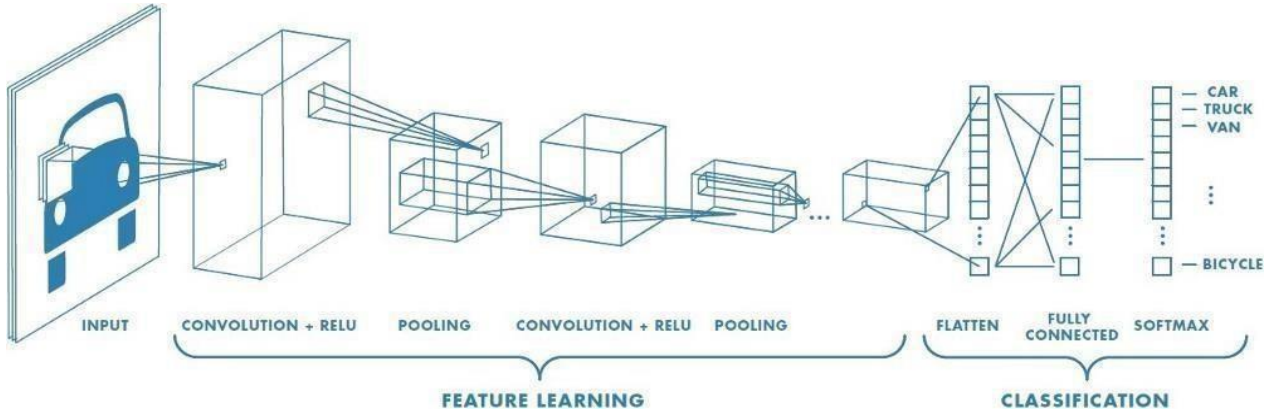
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12615

[4].Classification using CNN: Deepfake video detection using deep learning classification involves training a CNN model on a dataset of deepfake and real videos. The model learns to extract visual features from video frames to distinguish between manipulated and authentic content. By analyzing these features, the model can classify new videos as either fake or real, providing an effective solution for identifying deepfake videos.
[5].Output as video is fake or not: The system generates an output indicating the likelihood of the input video being a deepfake. It can be a binary classification (fake/real) or a probability score indicating the confidence level.

III. IMPLEMENTATION

Convolutional Neural Network(CNN):
A Convolutional Neural Network is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other It is a class of artificial neural networks that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including radiology. CNN is designed to automatically learn spatial hierarchies of features through back propagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. The first two layers convolution and pooling perform feature extraction whereas the third layer a fully connected map the extracted features into final output, such as classification.



Convolution layer :
A convolution layer is a fundamental component of the CNN architecture that performs feature extraction.It is consists of a combination of linear and nonlinear operations, i.e., convolution operation and activation function

Pooling layer:
A pooling layer provides a down sampling operation which reduces the in-plane dimensionality of the feature maps in order to introduce translation invariance to small shifts and distortions, and decrease the number of subsequent learnable parameters. It is of note that there is no learnable parameter in any of the pooling layers, whereas filter size, stride, and padding are hyper parameters in pooling operations, similar to convolution operations.

Fully connected layer :
The output feature maps of the final convolution or pooling layer is to transformed into a one- dimensional array of numbers , and connected to one or more fully connected layers, also known as dense layers, in which every input is connected to every output by a learnable weight. Once the features extracted by the convolution layers and down sampled by the pooling layers are created, they are mapped by a subset of fully connected layers to the final outputs of the network,such as the probabilities for each class in classification tasks. The final fully connected layer typically has the same number of output nodes as the number of classes.

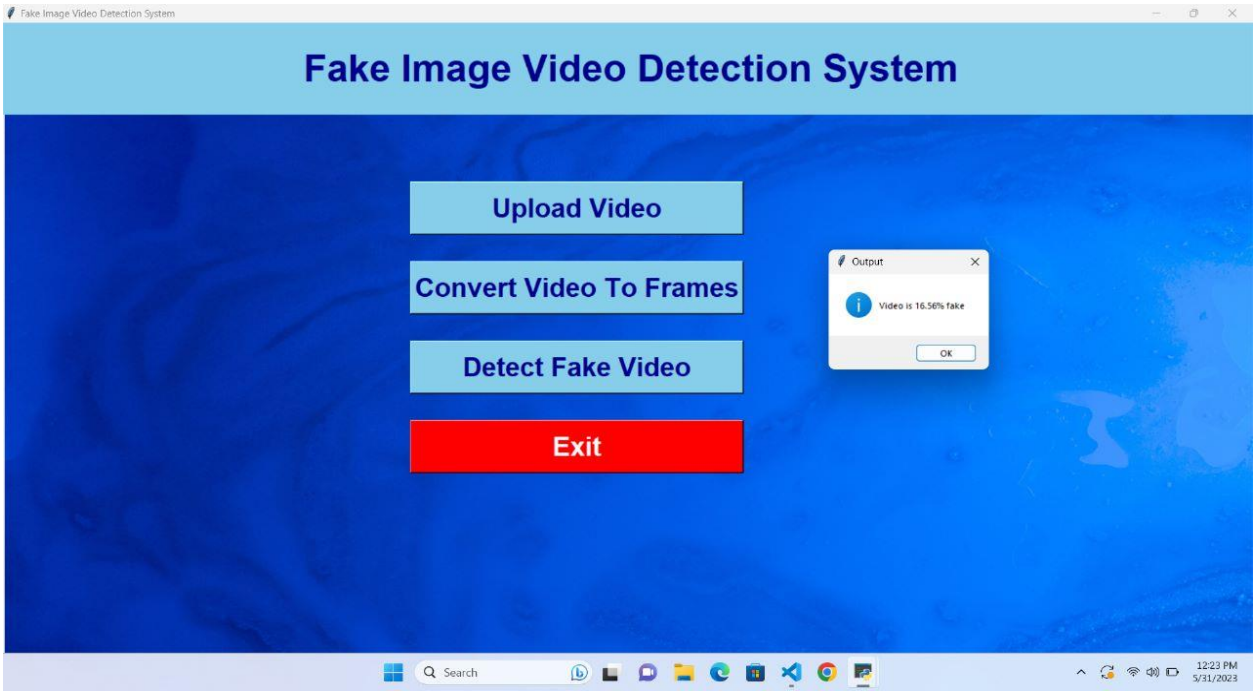
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12615

RESULT



The system generates an output indicating the likelihood of the input video being a deepfake. It can be a binary classification (fake/real) or a probability score indicating the confidence level is 16.56 % .

IV. CONCLUSION

Deep fake video detection through deep learning is an effective approach to identify manipulated videos. By training deep learning models, such as CNNs, on datasets of deep fake and real videos, we can teach the models to recognize visual patterns and features that distinguish between fake and authentic content. These models can then be deployed to classify new videos, helping to combat the spread of deep fake videos and ensuring the integrity of multimedia content.

REFERENCES

[1]. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). "FaceForensics++: Learning to Detect Manipulated Facial Images."
[2]. Li, Y., Li, X., & Lyu, S. (2020). "Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics."
[3]. Sabir, A., Fauzi, M. F., Ahmad, W., & Saba, T. (2020). "DeepFake Video Detection Using Deep Learning
[4]. Agarwal, A., & Namboodiri, A. M. (2020). "Protecting World Leaders against Deep Fakes."
[5]. Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). "MesoNet: A Compact Facial Video Forgery Detection Network."
[6]. Nguyen et al. (2020): "Capsule Networks for Deepfake Detection."
[7]. Lyu, S. (2018, August 29). Detecting deepfake videos in the blink of an eye. Retrieved from http://theconversation.com/detecting-deepfake-videos-in-the-blink-of-an-eye-101072
[8]. Yang, W., Hui, C., Chen, Z., Xue, J. H., and Liao, Q. (2019). FV-GAN: Finger vein representation using generative adversarial networks. IEEE Transactions on Information Forensics and Security, 14(9), 2512-2524.
[9]. Zakharov, E., Shysheya, A., Burkov, E., and Lempitsky, V. (2019). Few-shot adversarial learning ofrealistic neural talking head models. arXiv preprintarXiv:1905.08233.
Korshunov, P., and Marcel, S. (2018, September). Speaker inconsistency detection in tamperedvideo. In 2018 26th European Signal Processing Conference (EUSIPCO) (pp. 2375-2379).IEEE.