Detecting Real Time Deep Fake Video Using Neural Network

Mithun P M[1] and Sindhu Daniel[2]
Student, Department of Computer Applications[1]
Assistant Professor, Department of Computer Applications[2]
Musaliar College of Engineering & Technology, Pathanamthitta, Kerala

Abstract*: Because of realistic deepfake production technologies are always being developed, it is extremely difficult to detect these videos. These deep fakes are getting better over time to the point that it is difficult to tell whether they are real or fake, making them harder to catch. Deep-fake technology will have some advantages, but it will also do a lot of harm. Nothing is more dangerous than people accepting these movies at face value in a time when truth is rapidly eroding. These deepfakes can be used for a variety of evil intentions, including defaming public figures, fostering political bias, sabotaging personal relationships, inciting fear and exploitation, and spreading misleading information. This issue is addressed in the research by offering a model that evaluates video frames using a deep learning approach to find discrepancies generated during video creation, such as differences in compression rate and facial feature consistency. The model, which can detect these embedded faults in the deepfakes, is trained using a convolutional neural network and transfer learning. These differences created during deepfake construction surrounding the face are used to train the neural network.*

Keywords: CNN, MesoNet, deepfake, conv 2D

I. INTRODUCTION

There is a general consensus that we can and should believe that the person on the other end of our videoconferencing conversation is who they claim to be, despite the fact that many people have grown wary of videos they find online. However, it is becoming more challenging to believe even live video calls as more sophisticated deep fakes are being created in real time. Compared to off-line forensic analysis, detecting deep fakes in real time poses unique difficulties. We outline a method for quickly identifying deep fake videos sent through a live video conferencing platform. This method takes advantage of the fact that a video conversation often sets the user in front of a light source that can be controlled to cause a controlled alteration in the user's face look. This light source is the computer display . Real-time measurements of deviations from the anticipated change in appearance over time can be used to confirm a video conference participant's identity

II. LITERATURE SURVEY

Detecting Real-Time Deep-Fake Videos Using Active Illumination: Candice R. Gerstner National Security Agency Fort George G. Meade MD, Hany Farid University of California, Berkeley Berkeley CA, Despite our scepticism of internet viral photographs and videos, there is a general consensus that we can and should believe that the person on the other end of our videoconferencing discussion is who it appears to be. However, as increasingly complex deep fakes are being constructed in real time, it is getting harder to believe even live video calls. Real-time deep fake detection is more challenging than off-line forensic analysis. In this article, we present a technique for swiftly identifying deep fake videos transmitted via a live video conferencing platform. This technique makes use of the fact that a video discussion frequently places the user in front of a controlled light source to bring about a controlled change in the user's face look. This light source is the computer display.

Detection of fake 3D video using CNN: In Authors Shuvendu Rana, Sibaji Gaj, Arijit Sur and Prabin Kumar Bora in put forth a paper based on Neural Network. Convolution Neural Network (CNN), Dual Tree Complex Wavelet Transform (DT DCT), Depth Image-Based Rendering (DIBR), Multiview Video Plus Depth (MVD), and 3D Highly

DOI: 10.48175/568

Efficient Video Coding (3D-HEVC) are some of the topics covered in the study. Using CNN, the author attempted to identify a technique for identifying false 3D video from real 3D video. The dual tree complex wavelet transform is used in the author's pre-filtration process to reveal the edge and vertical and horizontal parallax properties of actual and false 3D films.. Over the training and testing datasets, the effectiveness of the fake 3D video is evaluated. Each video sequence contained in the training dataset is utilised to train the CNN. The author argued that in order to obtain the specified precision, enormous computational resources and a time-consuming process are needed. For instruction, high-definition video clips are employed. For the suggested scheme, the author implemented CNN architecture. The author can strive to develop a powerful, more effective system for distinguishing between authentic and false videos.

III. PROPOSED SYSTEM

The suggested deep fake detection system attempts to take advantage of deep learning breakthroughs, notably Convolutional Neural Networks (CNN), to create a more reliable and accurate deep fake detection system. The suggested method can successfully learn the visual patterns and features particular to deep fake material by training a CNN model on huge datasets of both real and altered videos, enabling accurate detection.

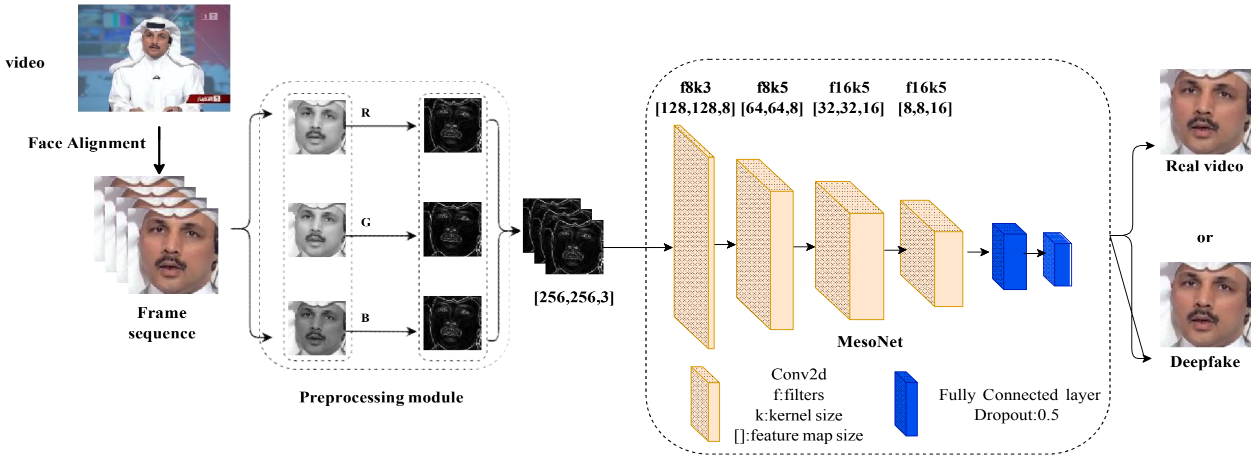

Figure 1. The pipeline of our proposed Deepfake detection method.

IV. METHODOLOGY

In this paper, we presented a method for identifying deep fake films that uses the neural network technique. Here, the system is operated by a convolutional neural network. Neurons, the computational units, are arranged in networks that make up neural networks. A number (the initial input or the output of the preceding layer) plus an activation function make up each neuron. The non-linear functions used to determine the output are called activation functions. Neurons in the convolutional neural network are partially connected to the layer above. The connections between the neurons depend on the type of filter that is being used; for example, if a 33 filter is being used, nine neurons in the nth layer will determine the output of one neuron in the (n+1)th layer. Convolution is applied to the image's normalised pixels. Features are extracted by the convolution procedure, which multiplies filter values and pixel values before adding the values. When the fxf filter is used on a nxn image, the output has the following dimensions:The dataset for our suggested system consists of deepfake videos, from which frames were selected. Frame level feature extraction was then carried out using a combination of dense and convolutional neural networks to identify pixels. This convolution technique extracts the features of the corresponding 9 pixels, which means that it does so for portions of images. Then, using the output of the convolution operator's computation, these properties of individual image components are used for detection. The largest computed value among the values picked in the n-dimensional array is subsequently chosen by the pooling layers. Less computations are required when using pooling layers. The final layer, which has two neurons and is entirely connected to the preceding layer, is then connected to these layers.

Three Layers of the Convolutional Neural Networks:

- **Convolution Layer:** The foundation of CNN is this layer. The majority of computations take place at this stratum. Three inputs make up this layer. Data input is first. A filter is the second, and a feature map is the third. Utilising filters on the image as input data, the convolution process takes place. A 33 matrix, which is iterated over the input data, which is likewise a matrix, is all that a filter is. An activation function that calculates the filter result and the area of the image is applied as the filter is fed over the input matrix's area. The feature matrix is then used to store the output.
- **Pooling Layer:** Using this layer, we can downscale the number of parameters in our input image. It functions similarly to how the convolution layer does. The sole distinction is that the pooling layer either takes the maximum value in the area of the input matrix or the average value in the area of the matrix, whereas the filters in convolution layers carry certain weights on which computation is conducted by activation function. Two different methods of pooling exist.:
- **Max Pooling**– The largest value across all the values in the region of the matrix is taken in this pooling.
- **Average Pooling**– In this pooling, we take the average of all the values in the patch of the matrix.
- **Fully Connected Layer**– In this phase, all the calculations we made in the earlier steps will finally pay off. The image is finally classified in this layer. Our input matrix is flattened and transmitted via the neural network's hidden layers in this layer.

DOI: 10.48175/568

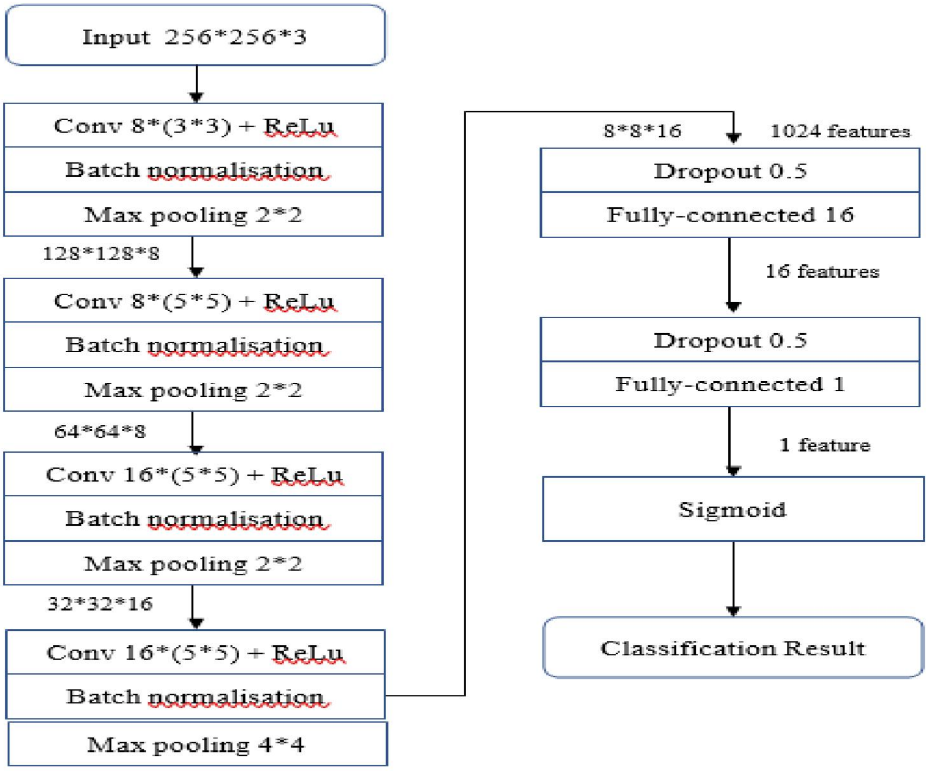

Figure 2. The network architecture of MesoNet

MesoNet:

The effectiveness of deep fake detection algorithms based on low-level noise features is greatly reduced when the video or image content is compressed. This is due to the fact that when the image is compressed, some feature information will be lost, and the noise information that the detector depends on will also be diminished. Similar to this, the network struggles to distinguish between real face images and Deepfake generated images at the semantic level of images, especially when the image is describing a face. Archfar et al. therefore suggest MesoNet based on the Inception module MesoNet is a neural network with a limited number of layers that is primarily concerned with picture mesoscopic features, which are situated in the middle between high-level and low-level features. After doing tests on a more complicated CNN network topology, Archfar et al. gradually simplified the network structure. It should be emphasised that the simplified network structure's detection ability is on par with that of the complex network structure. Last but not least, the proposed MesoNet consists mostly of two fully connected layers and four consecutive convolutional layers. A thick layer acting as a hidden layer is added to the network after four consecutive convolutions and pooling. Each convolutional layer has a nonlinear ReLu activation function to enhance generalisation. To avoid vanishing gradient effects, regularisation is done via batch normalisation. Additionally, each completely linked layer is regulated by the Dropout layer to increase the network's sturdiness. There are 27,977 trainable parameters in the network as a whole. The particular network

Comparison with Previous Methods

The outputs of those algorithm models can be analysed and compared, and various combinations of hyperparameters with regard to neural networks can be used for the study of deep fakes. By addressing deep fakes in the most effective manner possible, one of the major threats to the veracity of videos is mitigated. To maintain the originality of videos, immutable storage can be used with contemporary technologies like blockchain. DeepFakes employ specialised techniques that typically alter fixed facial features that serve as a base for superimposition. The algorithm uses a similar technique to produce several deepfakes, leaving minor inconsistencies during editing. suggests a technique for training the classifier using video frame input. The frames are sent to the classifier for training after being passed via face extraction and alignment fragment. Before training the model, the dataset is pre-processed. Face extraction and alignment are included in this. The suggested model focuses on flaws brought about by deepfake construction near the face shape. The area that needs to be processed will therefore be extracted during face extraction. In order to accommodate for potential variations in the target person's head position in the deepfake video, face alignment is used

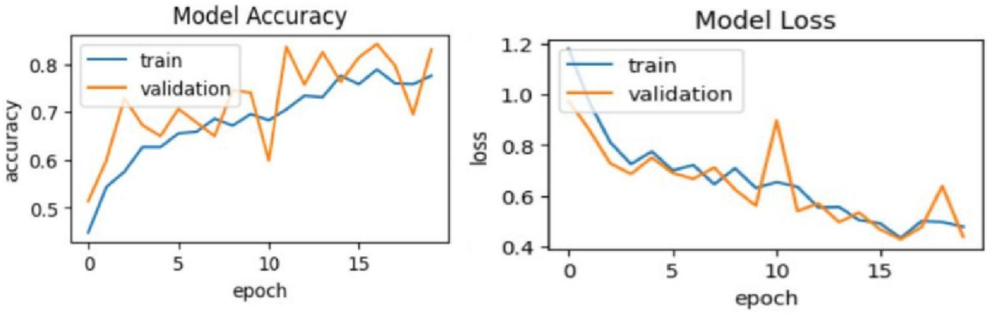

Figure 4 Model Accuracy

V. CONCLUSION AND FUTURE SCOPE

Using neural network methods, the deepfake video detection project has shown promising results for the detection and avoidance of modified and fake faces in videos. To accurately extract, preprocess, classify, and label data, the project used cutting-edge approaches. Based on the attributes discovered through image analysis, machine learning algorithms have been successfully used to detect the presence of deepfake films with an accuracy level of about 90%. Any video can be used with the proposed methods to find modified or false faces. Deepfake films have major potential effects on the veracity of multimedia material, including the dissemination of erroneous information and propaganda as well as the defamation of individuals. Therefore, it is essential to create a deepfake detection system that is accurate and dependable. The project's attributes can be further used to temporal analysis to more accurately identify deepfakes. As a result, by combining more sophisticated deep learning algorithms and investigating more data sources, the suggested methodology has the potential to be improved and refined. This project's contribution to the creation of a deepfake detection system, which can aid in maintaining the accuracy of multimedia information and halt the spread of misinformation and propaganda, is what makes it significant. The deepfake detection project's scope for future work covers a number of potential areas for extension and enhancement. These areas can improve the system's functionality, address new problems, and boost deepfake detection technology.

REFERENCES

[1] Sheng-Yu Wang, Oliver Wang, Richard Zhang, Andrew Owens, and Alexei A Efros. CNN-generated images are surprisingly easy to spot...for now. In IEEE Conference on Computer Vision and Pattern Recognition, 2020.

[2] Xin Yang, Yuezun Li, and Siwei Lyu. Exposing deep fakes using inconsistent head poses. In IEEE International Conference on Acoustics, Speech and Signal Processing, pages 8261–8265, 2019.

[3] Ning Yu, Larry Davis, and Mario Fritz. Attributing fake images to GANs: Learning and analyzing GAN fingerprints. In IEEE International Conference on Computer Vision, 2018.

[4] Xu Zhang, Svebor Karaman, and Shih-Fu Chang. Detecting and simulating artifacts in GAN fake images. arXiv: 1907.06515, 2019.

[5] Peng Zhou, Xintong Han, Vlad I. Morariu, and Larry S. Davis. Two-stream neural networks for tampered face detection. In IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017. 2 D. Alahakoon, X. Yu, "Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey," in IEEE Transactions on Industrial Informatics, vol. 12, no. 1, pp. 425-436, Feb. 2016, doi: 10.1109/TII.2015.2414355.

[6]Vineet Mehta, Parul Gupta, Ramanathan Subramanian, and Abhinav Dhall. FakeBuster: A deepfakes detection tool for video conferencing scenarios. arXiv: 2101.03321, 2021. 2.