JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DETECTION OF THE AI GENERATED DEEPFAKE VIDEO BY USING MULTI-TASK CASCADED CONVOLUTION NEURAL NETWORK

[1]ABHISHEK N , [2]SHILPA H L
[1]Student, [2] Assistant Professor
Department of MCA,
PES College of Engineering, Mandya, Karnataka, India

Abstract : In recent years, deepfake has gained popularity via the creation of deep fake pictures and videos, as well as the detection technique for false videos. There are many methods available for creating a deepfake video, including the generative adversarial network (GAN), which is used to generate false pictures or movies by superimposing different images on top of one another. The discriminator in the GAN is used to identify the generated fake pictures. In this study, we had created a web platform for detecting artificially produced deep fake video using Multi-task Cascaded Convolution Neural Network (MTCNN) and calculating the noise from face landmarks. We achieved an accuracy of 78.8 percent in the experimental environment.

.

IndexTerms **– Generative Adversarial Networks, Multi-Task Cascaded Convolution Networks algorithm, Deepfake**

I. Introduction

Digital Media has been a goal for content stealers to chisel off the work from others' efforts and to get a name for them. The same is the motive for many unofficial and unknown cases where the creator's work has been stolen for multiple purposes and the scope for the original product has declined due to the overpowering performance of the duped product. This is not only a problem for the content creators but also for the consumers or the users who are being cheated in the name of brand and might be influenced over the recreated and fake content. Image and text modification have already developed and are booming all over the world with the introduction of many media editors such as Microsoft Office Word which works as a text editor and Adobe Photoshop CS6 which works as Image and Template editor. These have already taken their major form in the market within the early years of 19th Century. This is being applied slowly but steadily in a more advanced and sophisticated manner with the development in AI and ML techniques over the motion pictures too. Due to rise in usage and knowledge about the various techniques possible to do the manipulation of the content, many technological endeavors have concentrated on the possibilities of digital video modification leaving few or possibly no traces behind which indicate fraudulency in the 20th Century The main purpose of this study is to detect the deepfake video that is generated by some of the fake video generated tools that uses the Generative Adversarial Networks (GAN) Algorithm for creation of fake video. That by using this tool they can spread the fake news by using the deepfake video. To solve this kind of Real world problem we have build a system which could detect and classify the fake videos based on the provided video input

II. Literature Survey

Authors David Gu¨era and Edward J. Delp had published a Paper on the Detection of the deepfake videos. In the paper they had discussed about the detection of the deepfake videos by using the CNN and RNN and also discussed about the how the deepfake is emerging as the big threat for the further. In this work the authors had created a pipeline to detection the deepfake videos by detecting the temporal inconstancy between each frame based on that it will detect that the given video is real or fake. Hear the authors are used the 600 videos of the dataset for training the model from the different hosting websites which of 2 second of video with 40 frames for detection of the fake videos. In the proposed system they had got the experimental efficiency for the subsequence accuracy for 20 40, 80 frames of an accuracy of 96 , 97, 97.1 percent respectively for detection of fake video. [1]

Author Md. Shohel Rana and Andrew H. Sung had published the paper on the detection of the fake videos. In this paper they had discussed about the some of the deep ensemble learning techniques that are stacked into gather called deepstack for the detection of the deepfake videos. The main objective of this paper is creation of training data sets that can be adapted to the test dataset, creating the neural network for the detection of facial features ,performing the detection of the video weather the given Video is real or fake .In the paper they had used the face features technique for the detection of the deepfake by using the deep stack that they had created and they performed the comparative analysis on the different deep ensemble learning techniques for the detection of the weather the given video is real or fake. [2]

Author Worku Muluye Wubet had published a paper on the deep fake challenge and deepfake detection. In this paper they had discussed about the detection of the deepfake video by calculating the eye blinking rate in each frame sequence by using the Convolutional neural network and Long Short Term Memory (LSTM) algorithm for the feature extraction and the sequence processing of the videos. [3]

Author Yuezun Li, Ming-Ching Chang and Siwei Lyu. Had published the paper on Exposing AI Generated Fake Face Videos by Detecting Eye Blinking. In this paper the they had proposed the methodology of detection of the deepfake videos by using the detection the eye blinking rate for the specific time period by cropping the eye frames and analysing the each eye frame in the sequence by using the recurrent neural network (RNN) for the calculating the time sequence of each eye coordination of the each eye frame by comparing the degree of the openness of the eye in each video frame. [4]

Author Lakshmanan Natarajan and team had published a paper on the Detecting GAN generated Fake Images using Co-occurrence Matrices. In this paper they had proposed the deep fake images detection by using the Co-occurrence matrices for extract the three pixel channels and the model is trained by using the deep neural network for the co occurrence matrices. In this paper they had used the starGAN dataset and CycleGAN dataset for the detection of the fake images and they had performed the comparative analysis on the different state of art Algorithms where they had got the 94.99% and 93.42% on the ImageNet and XceptionNet respectively [5]

worku Muluye Wubet in 2020 In this paper they had conducted the deepfake manipulation tools that are available dataset and detection challenges by using the eye-blinking technique. In this study they had used the UADFV dataset for the deepfake video detection. And here they had used the CNN and the LSTM algorithm for the detection of the fake videos and the experimental results are for the real video for the 49 number of the video the average video length is 11.26 sec for 28fps the eye blinking rate is 14.42/min whereas for the fake video for the 49 video the average video length is 11.26 second for 28fps we got the 4.28/min as the eye blinking rate [6]

Md. Shohel Rana in 2020 had proposed and developed the deep ensembles learning techniques that are the stacking ensemble and randomized weight ensemble. In this study they had use the faceforensics++ dataset. In this study they had performed the comparative study on the different convolution neural network that we got the XCEPN-96.88 accuracy for the MOBNET 90.74%. RSNET101 94.95%. And DNS121 93.34%. DSNET69 94.13% [7]

Sebastien Marcel and Pavel Korshunov In 2018, had proposed a study on deepfake as a new threat for the recognition, assessment, and detection of deepfake video detection. This paper contribution is the first utilised the VIDTIMIT dataset, which is a publicly accessible collection of low-level and high-level vidTIMIT movies which is Using a GAN-based approach, with swapped faces.. The research utilises visual data to detect discrepancies between the lip movements and audio speech of different versions of the deepfake video. [8]

Luca Guarnera University of Catania - iCTLab Catania, Italy in 2020 In this document a novel deep detection technique focusing on pictures of human faces will be presented which in this research using the exception maximizing algorithm extracts a set of local features especially addressed to the model and the Convolution image trace cloud. The indigenous classifier is utilised to train the 5 most realistic CGDWCT datasets, STSRGAN, ATTGAN andSTARGAN. Precision is CELEBA against ATTGAN –92.67%-, CELEBA against GDWCT –88.40%-, CELEBA Versus STARGAN 93.17% [9].

Yuezanli et al. published a study in 2020 describing a novel deep learning-based technique for effectively distinguishing deepfake from genuine footage. Due to the limits of the deepfake algorithm, this process is base on the property of the deepfake video. Due to the limitations of the deepfake algorithm, this method can only generate fake pictures of fixed size that must go through affine matching to the configuration of the face sources. In this study, we are detecting deepfake videos using the warping technique. We used the UADFV dataset and performed analysis on the different methods, including the two-stream NN, which has an accuracy of 85.1 percent, the meso Inception, which has an accuracy of 82.1 percent, and the head pose technique, which has an accuracy of 89.0 percent. [10]

III. Proposed Model

In the Proposed methodology we have developed the system to detect the deep fake video. Our technique is based on the detection of the deepfake video for the specific size due to the resources limitation in experimental level. This proposed system our contribution is to detect the deepfake video that is spreading the fake news over the internet. We are developed a web platform for the detection of the deepfake videos that is used to spread the false information over the internet. This technique can be used the top companies like face book, whatsapp , and YouTube or other video sharing platforms where we can detect the fake video before sharing the video. Our method is used to detect the all kinds of deepfake video like enhancement deepfake or the interpersonal deepfake. The figure 1 shows the simple architecture of proposed methodology. system where firstly we upload the video into the system that video is converted into frames next after the extraction of the frames from the video next we perform the face detection and then extract the face from the each frame and then we find the facial landmarks from the face and next we calculated the noise as of the each face and we classify weather the given video is fake or real

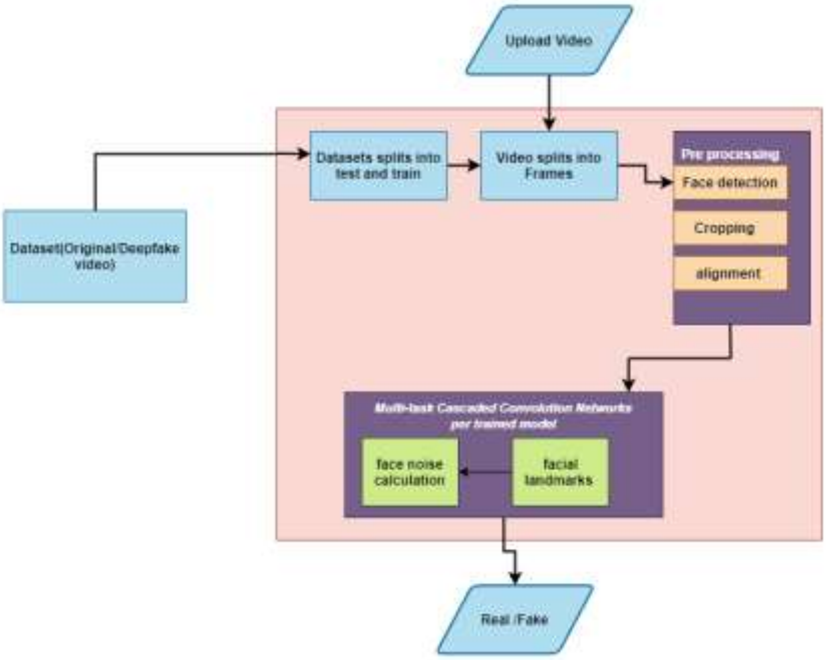


Figure 1 Simple Architecture of Proposed System

RESEARCH METHODOLOGY

Dataset used: In this proposed work we have used the deep fake detection challenge dataset which consists of the 800 videos that consists of the 10 sec video each, this dataset is created and organized by the kaggle for the spur researchers to build innovative and new technique for the detecting the deep fake and manipulated images or media . In this proposed system we have dividing the whole dataset into test and train we have used 70% for the training purpose and 20% for testing purpose.

Preprocessing: In the dataset preprocessing includes the splitting of the frames of the given video where we divide the frames from the video up to 30 frames that each framed will be capture 0.5 second time interval. After the extracting the frames from the video we have to detect the face that is present in the frame after the face detection we have to crop the face from the frame and we have to see the face alignment will be preformed. The video that does not have any faces that will be not detected in the pre processing stage.

Multi-task cascaded convolution neural network: In this Proposed, we utilized multi-task cascaded convolution networks to reorganize the face and identify facial landmarks. MTCNN is an acronym for Multi-task Cascaded Convolution Networks, which was developed toward address both face identification and face alignment. Three layers of convolution networks are used in the technique to identify faces and facial landmarks such as the eyes, nose, and mouth. In this proposed system after the face detection next we have to perform the face embeddings from the cropped face that where it consider the whole face and facial landmarks from the each face by using MTCNN framework next we have to calculate the noise from the centroid of the face to the face embeddings then it calculate the noise from the face and calculate the noise if the probability of the noise is less that 0.5 then it will classify weather the given video is fake or not.

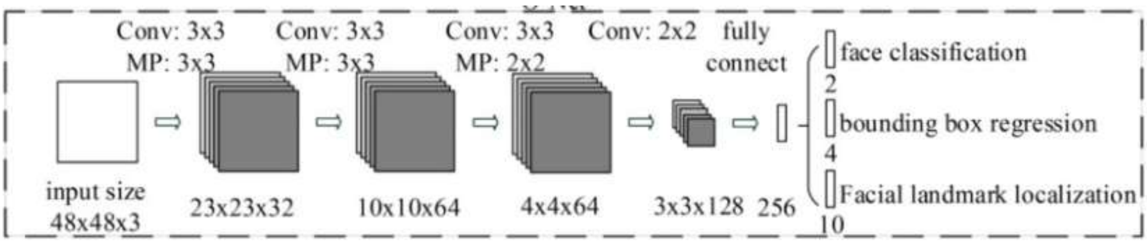


Figure 2 Simple architecture of MTCNN

The formula for the face Classification will be done by calculating the binary cross entropy loss

- L_i^{det}= -(y_i^{det} log(pi)+(1-y_i^{det})(1-log(p_i)))

Where L_i^{det} is the binary cross entropy loss and y_i^{det} is the ground truth label and p_i is the probability produced by the Network

To determine the offset between each candidate window and the closest ground truth, the bounding box regression formula is employed. This job is accomplished via the application of Euclidean loss.

- $L_i^{box} = ||yi^{box} - y_i^{box}||_2^2$

Where L_i^{box} is the bounding box regression and Y_i is the target obtained from the network and yi^{box} ground truth coordinate

The formula for the facial landmark is determined by formulating the localization of facial landmarks as a regression issue using Euclidean distance as the loss function.

- $L_i^{ladnmark} = ||y_i^{| \, landmarks} - y_i^{landmarks}||_2^2$

Where $L_i^{landmark}$ is the face landmark that is calculated from the mod value of the both the coordinates for detecting the facial landmarks

After performing the next we have to calculate the noise to calculate the noise we have to subtract the value of the centroid with the face embeddings that value will be used to classify weather the given video is deep fake or real.

FLOWCHART

The system flow of the proposed methodology starts with the uploading the video after he uploading the video that frames will be extracted from the video in the 0.5 second time interval after the extracting the frames from the video next we detect the face from the frame and then the face embeddings from the each face that is extracted and next we calculate the noise from the face embedding and classify weather the video is real or fake that is shown in figure 3

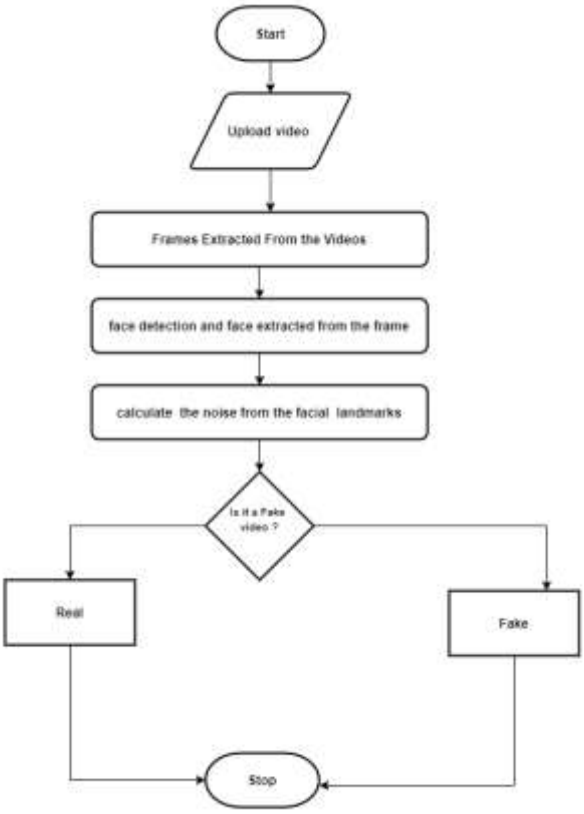
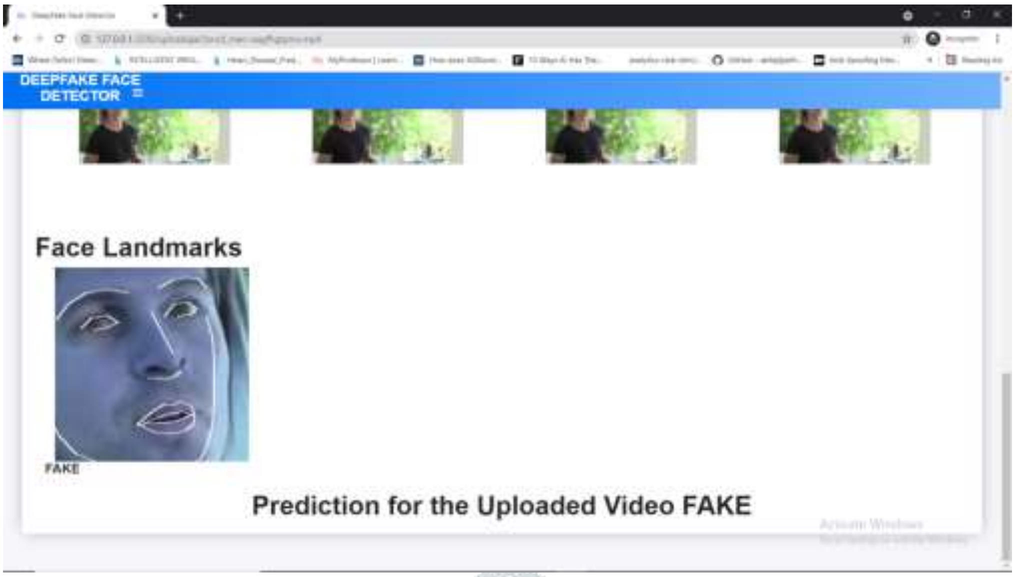

Start
Upload video
Frames Extracted From the Videos
face detection and face extracted from the frame
calculate the noise from the facial landmarks
Is it a Fake video ?
Real
Fake
Stop

Figure 3 Flow Chart of Proposed System

IV. RESULTS AND DISCUSSION



DEEPFAKE FACE DETECTOR

Face Landmarks

FAKE

Prediction for the Uploaded Video FAKE

The output of the proposed work classify the weather the given video is fake or real with the face embeddings after the calculating the noise it will classify weather the given video fake or real. where we upload the video into the system that video is converted into frames next after the extraction of the frames from the video next we perform the face detection and then extract the face from the each frame and then we find the facial landmarks from the face and next we calculated the noise as of the each face and we classify weather the given video is fake or real

V. Conclusion

We had proposed a methodology to classify weather the given video is deep fake or real by using MTCNN algorithm .The Proposed methodology is inspired by generation of the deep fake generation GAN Algorithm to solve this problem we have proposed a methodology for detecting of the deep fake our method is used to detect the deep fake by calculating the face features like facial landmarks of the whole face from the center of the face and calculate the noise based on that noise we are calculating weather the given video is deep fake or real. In future The first recommendation for the future enhancement is to detect the deepfake for the multiple faces that is present in the same video. The second recommendation for the future enhancement adding the block chain technology where we can use the smart contract technique for the video authority for each frame of the video.

VI. Acknowledgment

The authors wish to express sincere thanks to Assistant Professor Mrs.Shilpa H L, Department of MCA of PES College of Engineering Mandya for his continuous inspiration towards this work.

References

[1] Yuezun Li, Ming-Ching Chang and Siwei Lyu , In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking , https://arxiv.org/abs/1806.02877

[2] David Gu ̈era Edward J. Delp, Deepfake Video Detection Using Recurrent Neural Networks , 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)

[3] Divya Babu1, Uppala Santosh Kumar2, Konduri Ajith Kumar3, Yennana Jayanth Sai ,2020 Deepfake Video Detection using Image Processing and Hashing Tools,https://www.irjet.net/archives/V7/i3/IRJET-V7I374.pdf

[4] Thanh Thi Nguyen, Cuong M. Nguyen Thanh Thi Nguyen, Deep Learning for Deepfakes Creation and Detection: A Survey,https://arxiv.org/abs/1909.11573

[5] Lakshmanan Nataraj, Tajuddin Manhar Mohammed, Shivkumar Chandrasekaran, Arjuna Flenner, Jawadul H. Bappy, Amit K. Roy-Chowdhury and B. S. Manjunath , Detecting GAN generated Fake Images using Co-occurrence Matrices , https://arxiv.org/abs/1903.06836

[6] Worku Muluye Wubet , The Deepfake Challenges and Deepfake Video Detection, https://www.ijitee.org/wp-content/uploads/papers/v9i6/E2779039520.pdf

[7] Md. Shohel Rana, Andrew H. Sung , DeepfakeStack: A Deep Ensemble-based Learning Technique for Deepfake Detection, https://ieeexplore.ieee.org/abstract/document/9171002

[8] Pavel Korshunov and Sebastien Marcel , DeepFakes: a New Threat to Face Recognition? Assessment and Detection, https://arxiv.org/abs/1812.08685

[9] Luca Guarnera, Oliver Giudice, Sebastiano Battiato, DeepFake Detection by Analyzing ConvolutionalTraces,https://openaccess.thecvf.com/content_CVPRW_2020/html/w39/Guarnera_DeepFake_Detection_by_Analyzing_Convolutional_Traces_CVPRW_2020_paper.html

[10] Xin Yang , Yuezun Li , Siwei Lyu, exposing deep fakes using inconsistent head poses, https://arxiv.org/pdf/1811.00661.pdf.

[11] https://www.kaggle.com/c/deepfake-detection-challenge/overview

[12] R. Raghavendra, Kiran B. Raja, Sushma Venkatesh, and Christoph Busch, “Transferable deep-CNN features for detecting digital and print-scanned morphed face images,” in CVPRW. IEEE, 2017.

[13] Tiago de Freitas Pereira, Andr´e Anjos, Jos´e Mario De Martino, and S´ebastien Marcel, “Can face anti spoofing countermeasures work in a real world scenario?,”in ICB. IEEE, 2013.

[14]Nicolas Rahmouni, Vincent Nozick, Junichi Yamagishi, and Isao Echizen, “Distinguishing computer graphics from natural images using convolution neural networks,” in WIFS. IEEE, 2017.

[15] F. Song, X. Tan, X. Liu, and S. Chen, “Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients,” Pattern Recognition, vol. 47, no. 9, pp. 2825–2838, 2014. [21]D. E. King, “Dlib-ml: A machine learning toolkit,” JMLR, vol. 10, pp. 1755–1758, 2009.

[16]Long Short-Term Memory: From Zero to Hero with Pytorch: https://blog.floydhub.com/long-short-termmemory-from-zero-to-hero-with-pytorch/ [11]Sequence Models And LSTM Networks https://pytorch.org/tutorials/beginner/nlp/sequence_models_tutorial.html

[17]https://discuss.pytorch.org/t/confused-about-the-imagepreprocessing-in-classification/3965

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