International Journal of Research in Engineering, Science and Management
Volume 7, Issue 4, April 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Deep Trace: Unravelling Deepfakes Through LSTM Analysis

D. Nandini[1], M. Abhishek[2*], M. Adithya[3], G. Anirudha[4], V. Charmyka Reddy[5]
[1]*Assistant Professor, Department of Information Science and Engineering, Jyothy Institute of Technology, Bengaluru, India*
[2,3,4,5]*UG Student, Department of Information Science and Engineering, Jyothy Institute of Technology, Bengaluru, India*

***Abstract*: Due to the intense rising of deepfake data which is causing huge risk in maintaining and protecting the consistency of different variety of data like video, audio, images and some interactive elements. These deepfakes are also harming the wellbeing and harmony of the society. Therefore, this paper brings in solution for detecting deepfakes by employing combination of different technology is called dual architecture. The spectaculated structure flawlessly, unified both Res-Net and LSTM with CNN. The Res-Net mainly emphasizes on categorizing and distinguishing between original or manipulated images. Additionally, Res-Net filter- outs patterns from images which includes facial contours and embedded text. LSTM with CNN unit pin points on temporal coherence and it also recognizes the realness of the image. Transfer Learning ways are used where understanding and intelligence is gained by training the dataset to magnify the functionality and performance of the spectaculated structure. In, Overview the entire deep fake detection system demonstrates unified collaboration of Res-Net and LSTM- CNN for delivering, a effective deep fake system which helps from attacks that is being created from AI tools. From the proposed model by increasing the accuracy and robustness a reliable system is generated to solve the attacks happening in digital world.**

***Keywords*: Datasets, Deepfakes, Feature extraction, LSTM-CNN.**

1. Introduction

In the modern age of revolving revolution in the field of technology and software artificial intelligence has been a boon and curse to mankind. Technological threats and unwanted usage of Artificial Intelligence in the field of crime has been increased over the decade and deepfake is one among them. There has been several cases of misuse in the field of technology which cannot be distinguished easily by any naked eye.

The perfection in the field of technology is making a huge threat in identifying or investigating the crime. There are several websites on the internet which indicates the threats of same kind. Face has been the most deepfake feature in recent times. Hence there is vigorous need of an algorithm which can identify the manipulation of expression in which the person is replaced with another person without interrupting the expression of the main person. Second type being the identity based one which can highly manipulate any persons identity in a particular video.

In the first type facial moments, actions, reactions and expression in the real time is replaced i.e., just the faces of two people can perfectly be swiped excluding the face. These types of manipulation can spread fake news and information which can highly mislead the crowd. There can be huge damage to any famous personalities which can rapidly harm their reputation in large scale.

This paper involves detection techniques of these type of deepfakes which is implemented using Convolution Neural Network (CNN). This can be successfully implemented by training the dataset in such a way that it can clearly differentiate between fake and the real. Any such algorithm typically involves some steps to be followed primarily conversion to frames followed by extracting the facial features and huge processed dataset which classifies them as real or fake.

2. Datasets

In the spectaculated structure of deep fake detection a data science platform called Kaggle is used to fetch the data set. Transfer Learning methods are used on these data sets for building effective model.

It is a huge data set extracted from Kaggle where data is split, training and testing is done for the labeled image to distinguish between real or fake image.

These data sets are used to solve the real-world problem that come across like deepfakes that is causing threat to society. Therefore, by implementing or training these datasets a deepfake model is built.

3. Proposed Methodology

This paper proposes two functions which are deep fake detection and deep fake creation.

Deep fake is used to identify the real output from the given input. The process is simple to execute, which collects the data sets and simplify, or perform LSTM with CNN.

Data collection from the data sets, in the deep fake selection the input added can be image or video dependent on user.

Face detection and alignment is done using LSTM with CNN to detect and align the face that covers the facial points or the main point from the image.

*Corresponding author: abhishekabhi58077@gmail.com

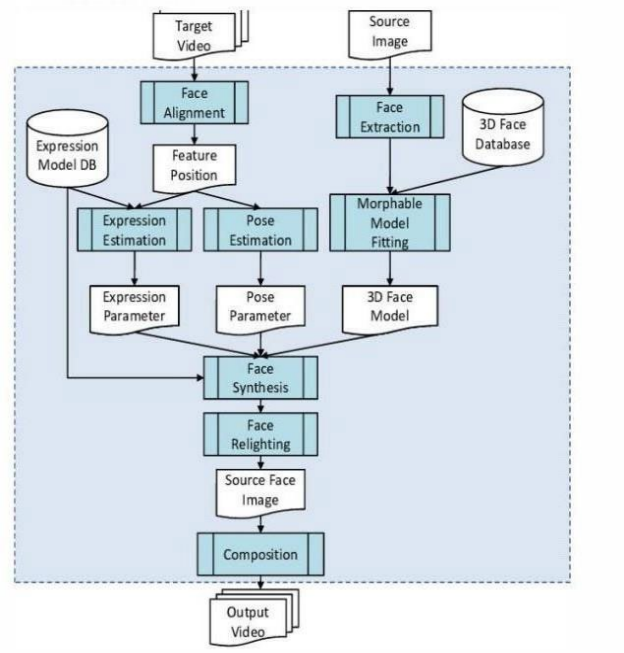

Fig. 1.

The extraction of the image is the crucial role for the deepfake process, and that is the first step of this project.

Feature extraction from the image or the video. From the targeted person it captures the face alignment and the extracted parameters are used to train the deep learning model.

Model training by the data sets user cancel select the images and the videos which will be compared to the data set and the data set will be trained.

4. Deepfake Creation

In deep fake creation that creates the input, which has a collection of the data sets. These data sets are obtained from a website called Kaggle. Which are highly trained data sets that can be used for the creation.

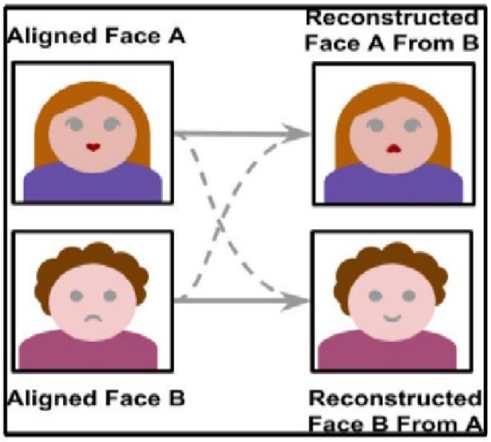

Fig. 2. Training of an image from Face A to Face B

The creation which undergoes in a various ways. it extracts, the targeted facial landmarks and the main highlighted point from the images.

Whereas, from the video, the procedure will be implemented same as the image extraction. This process will set the facial input for the task. The inputs from the data set will generate fake image. where the fake images will be taken as the input for the creation process.

The application is used to create a fake image from the data sets and that will be used for the input section. Considering the input section further implementation is done.

The extraction of the image is the crucial role for the deepfake process, and that is the first step of this project.

5. Fake Video Creation

In fake video detection the entire video is divided into frames from each frame pattern like facial contour is extracted from CNN feature extraction. Further by applying LSTM the realness of image is captured. Integrating both CNN and LSTM gives the clear representation of spatial and temporal data of image generated from video frames.

After integration data which is containing both real and fake images from video frames are trained to know the difference between real or fake image. Additionally, testing is done on certain set of data to check and evaluate accuracy, precision of the model. Later fine tuning is done on the dual architecture model that is designed for the proper output production by the given input. By using the above techniques fake videos are detected.

6. Deepfake Detection

The basic architecture to produce a deep fake is encoder and decoder architecture. Where in encoder acquires the features of the target and the source face. The decoder gets encoder feature of the target face and then generates the fake video or the fake image.

Using the high-level processing, the quality of video is been enhanced and left over or removed, but still few traces are left which are not visible by the naked human eye. These extracted features are used to train the recurrent neural network which is made to analysis the video that has been manipulated or not. The small portion of the video is being manipulated which is mean the deepfakes are shorter in time, therefore, the video has been split into small frames and these frames are given as input.

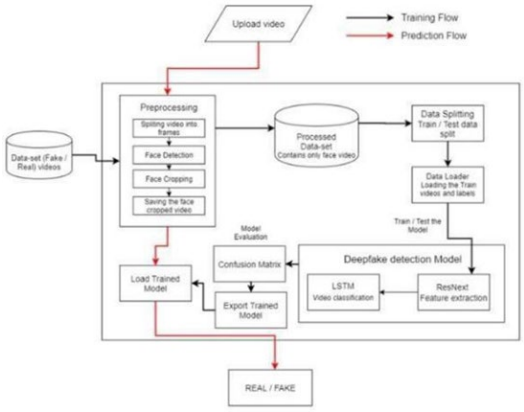

Fig. 3.

The study presented in the paper showcases a model that

achieves an accuracy of approximately 70% through image analysis-based feature learning. This model's performance is promising, suggesting that the features extracted can be leveraged for subsequent temporal analysis, facilitating the effective detection of deepfakes. During its training phase, the model was exposed to a dataset comprising approximately 4000 real frames and 5000 fake frames, with the test set accounting for 30% of the total data.

As training progresses, the loss diminishes, while accuracy ascends, indicating the model's improving performance. Enhancing the model's training by extending the number of epochs could potentially yield even better results. However, it is noteworthy that the dataset predominantly consists of low-resolution images, which pose challenges for effective learning. To further enhance the model's robustness, there is a pressing need to construct a high-resolution dataset tailored for deepfakes. Additionally, the issue of compression artifacts in images poses another challenge. While identifying signature styles of compression in deepfakes can aid in model training, it can also introduce learning errors. Thus, sophisticated temporal analysis techniques are essential to address these issues effectively.

7. Conclusion

The paper emphasizes the pressing need for robust deepfake detection techniques, given the potential socio-political ramifications of manipulated content. The proposed methodology employs transfer learning on the VGG-16 model, focusing specifically on facial manipulation detection. This approach streamlines the training process, necessitating minimal resources while yielding satisfactory accuracy. However, challenges persist, particularly with low-quality images and medium-quality videos, where the model's accuracy diminishes. To address these limitations, future research avenues could explore the integration of temporal and audio data to complement image-based feature extraction. Moreover, leveraging ensemble learning techniques and aggregating results across different frames and models could further enhance detection accuracy and account for dataset variance. Ultimately, the authors anticipate that their contributions will catalyze advancements in the realms of image and video forgery detection, fostering progress in digital media forensics. Enhanced datasets, innovative algorithmic approaches, and interdisciplinary collaborations are crucial for refining deepfake detection mechanisms and safeguarding the integrity of digital content.

References

[1] Oleg Alexander, Mike Rogers, William Lambeth, Jen-Yuan Chiang, Wan-Chun Ma, Chuan-Chang Wang, and Paul Debevec. The Digital Emily project: Achieving a photorealistic digital actor. IEEE Computer Graphics and Applications, 30(4):20-31, 2010.

[2] Antreas Antoniou, Amos J. Storkey, and Harrison Edwards. Augmenting image classifiers using data augmentation generative adversarial networks. In Artificial Neural Networks and Machine Learning - ICANN, pp. 594-603, 2018.

[3] Sercan Arik, Jitong Chen, Kainan Peng, Wei Ping, and Yanqi Zhou. Neural voice cloning with a few samples. In Proc. NIPS, pp. 10040-10050, 2018.

[4] Hadar Averbuch-Elor, Daniel Cohen-Or, Johannes Kopf, and Michael F Cohen. Bringing portraits to life. ACM Transactions on Graphics (TOG), 36(6):196, 2017.

[5] Facebook Wants to Stay 'Neutral' on Deepfakes. Congress Might Force it to Act.

[6] https://www.vox.com/future-perfect/2019/6/13/18677574/facebook-zuckerbergdeepfakes-congress-house-hearing

[7] A. K. Jain, A. Ross, and S. Pankanti, ''Biometrics: A tool for information security,'' IEEE Trans. Inf. Forensics Security, vol. 1, no. 2, pp. 125–143, Jun. 2006.