International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

DEEPFAKE DETECTION USING LSTM & RESNEXT-50

[1] MS. M. VARALAKSHMI, [2] SAKETH PAIDA, [3] L. RUTHWIK REDDY, [4] J. SAMEL

[1] Assistant Professor

Department of Information Technology,

Mahatma Gandhi Institution of Technology, Gandipet, Hyderabad.

mvaralakshmi_it@mgit.ac.in

[2] Department of Computer Science and Business Systems,

Mahatma Gandhi Institution of Technology, Gandipet, Hyderabad.

psaketh_csb203247@mgit.ac.in

[3] Department of Computer Science and Business Systems,

Mahatma Gandhi Institution of Technology, Gandipet, Hyderabad.

lruthwikreddy_csb203232@mgit.ac.in

[4] Department of Computer Science and Business Systems,

Mahatma Gandhi Institution of Technology, Gandipet, Hyderabad.

jsamel_csb203223@mgit.ac.in

Abstract:

As computing power increases, "deep fakes," or films that seem like they were created by a real person, are becoming feasible thanks to deep learning algorithms. Political instability, terrorism, revenge porn, or extortion can be the motivation for these lifelike face swapped deep fakes. This research presents a novel deep learning system that can recognize the difference between real and AI-generated films.We have developed a method that can identify deep fake replacements and reenactments automatically. We're using AI to combat AI. Our system's Res-Next Convolution neural network retrieves properties at the frame level. A recurrent neural network (RNN) trained using long short-term memory (LSTM) characteristics can distinguish between deep films

and regular movies. On big, balanced, and mixed datasets, our technique is tested by Face-Forensic++ [1], Deepfake Detection Challenge [2], and Celeb-DF [3]. The quality of real-time data is enhanced by this. We prove that our technology consistently outperforms the competitors.

Computer vision, Res-Next Convolution neural network, RNN, and LSTM are some of the index phrases.

1.INTRODUCTION

Deepfakes pose the greatest threat from AI due to the proliferation of social media platforms. Political instability, fabricated terrorist attacks, revenge porn, or blackmail may all benefit from these realistic face swapping deepfakes.As an example, there exist videos of Brad Pitt and Angelina Jolie in their underwear.

Find out what authentic videos are and what deepfake films are. Use AI to combat AI.In order to build deep connections, FaceApp [11] and Face Swap [12] use trained neural networks like GANs or autoencoders. Our approach integrates a pre-trained Res-Next CNN for data phase extraction with an LSTM-based neural network for snapshot selection over time. Frame-level characteristics are captured by the ResNext convolutional neural network. A long-term memory-based renewable energy source may learn to identify deepfake films by analyzing these characteristics. In order to train our system, we use a combination of data from Deepfake Detection Challenge [2], FaceForensic++ [1], and Celeb-DF [3]. The actual data is better organized as a result of this.

In an effort to streamline its use, we built a video-sharing front-end. The model will show how confident it is in its assessment of whether a video is a deepfake or not.Digital video creation and sharing has never been easier than with the rise of mobile cameras and social networking. Thanks to deep learning, technologies that were previously unimaginable are now within reach. Graphics, text, audio, and video may all be realistically generated using modern generative models. Medical image training data and text-to-speech have both benefited from these models. Incorporating deep generative models into video and audio samples has the ability to generate misleading material called "deep fakes," which brings up questions of ethics and presents new obstacles.There has been an increase in the number of synthetic media snippets with the introduction of several open-source deep fake generating methods and tools in late 2017. While some of these could be fun to watch, others might be harmful to people and the world at large. The prevalence of both fake and convincingly realistic films has grown as editing technology has become more accessible and experts are in high demand. Bullying and disinformation are common outcomes of the prevalence of deep fakes on social media. Envision our prime minister going to war with nearby countries or a famous person attacking their fans in a profound false. People will be scared and tricked by these horrible deep fakes. Avoiding the potential harm that deep fakes might inflict requires the capacity to detect them with precision. This article presents a novel deep learning approach that can distinguish between authentic films and deep fakes created by artificial intelligence. We need to develop systems that can identify deep fakes so they don't get circulated online.

2. LITERATURE SURVEY

The authors of Confront Distorting Artifacts [15] used a specific Convolutional Neural Network demonstration to detect artifacts by comparing the received face locations to their surroundings. In terms of front artifacts, there are two types of work.

The limitation of the deepfake algorithm to generate images of a certain size is the basis for their approach. These photographs will need to be edited since the source video faces need to be corrected. That method doesn't take frame timing into account.

One new way to spot deep fakes is eye movement detection [16], which uses the subject's eye movements to verify the

legitimacy of a video. In reduced frames, the Long-term Recurrent Convolutional Network (LRCN) monitored blinking eyes. Because deepfake algorithms are so effective, there are other ways to spot them than not blinking. Fake teeth, facial wrinkles, repositioned eyebrows, etc. are telltale signs of a deep-fake.

How to use capsule networks to detect manipulated media Repetitive attacks and computer-generated films are examples of images and videos that may be manipulated or falsified using a capsule network [17]. Worst of all, they practiced their strategy with erratic noise. The presentation worked great on their dataset, but it may not have worked on real-time data because of all the chaos in the preparation. We recommend practicing on real-time datasets that are devoid of noise. For deepfake localization, we used a pre-prepared Ima-geNet demonstration in conjunction with a Repetitive Neural Network (RNN). We used the HOHO [19] library, which has 600 movies.

Using real-time data can be a challenge for their creation of a few of comparable films. We will feed our model a mountain of data in real time so it can learn. Analyzing the face area of both real and fake portrait videos, Synthetic Portrait Videos using Biological Signals [20] creates a composite of the two. In order to train a support vector machine (SVM) and a convolutional neural network (CNN), as well as to record signal properties in feature vector and photoplethysmography (PPG) maps, transformations were used. The next step in determining whether the video is real or a deepfake is to look at the average authenticity odds. not real False recordings of any form, measurement, quality, or creation may be found by Catcher. The insights they made on natural flag upkeep were in vain as they lacked a discriminator. Creating a differentiable disaster that monitors flag handling steps is a challenging task.

3. METHODOLOGY

i) Proposed System:

The problem of deepfake can be resolved by using deep learning to differentiate between real and AI-made films. We get frame-level properties from a ResNext CNN, which are then passed on to LSTM-based RNNs. Automated detection of replication and alternative deepfakes is a feature of this system. Deepfakes may be exploited for political gain, fabricated terrorist events, revenge porn, or blackmail; our method leverages AI to fight AI to lessen these hazards. To make our method more useful, we put it through its paces on the Deepfake Detection Challenge, Celeb-DF, and Face-Forensic++ datasets. Competing with more complex deepfake production techniques, our approach is both easy and trustworthy, and it allows for fast real-time processing. and (ii) The Architecture of the System: ATo ensure that our PyTorch deepfake detection model was free of bias, we trained it on an equal number of real and fake videos. Before work began, we preprocessed a dataset to remove everything except the face-cut films. With the help of GANs and autoencoders, we can improve the quality of movie faces and learn how deepfakes are made. Deepfakes may seem realistic, but they hide all the finer nuances. Our goal is to find and collect all of these subtle flaws so that we can tell real movies from fake ones.

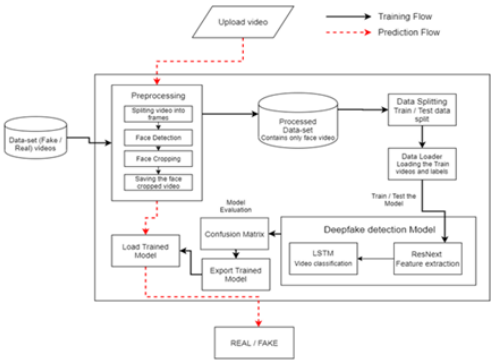

Fig 1 Proposed Architecture

iii) Dataset:

The integration of data from FaceForensic++ (FF), the Deepfake Detection Challenge (DFDC), and Celeb-DF enhanced real-time predictions. You can tell different types of movies apart with the assistance of our exclusive collection of 50% real and 50% fake films. In an effort to maintain coherence, DFDC omitted films with altered scores. Following the editing process, we selected 1500 genuine and 1500 inauthentic DFDC films, 1000 authentic and 1000 inauthentic FF movies, and 500 authentic and 1000 inauthentic Celeb-DF videos. Out of the total of 6,000 videos, 3,000 were authentic and 3,000 were fake. With this equitable allocation, the model becomes more stable and adaptable. As shown in Figure 2, our method is able to handle a wide variety of video sources and attributes because of the way our dataset is distributed.

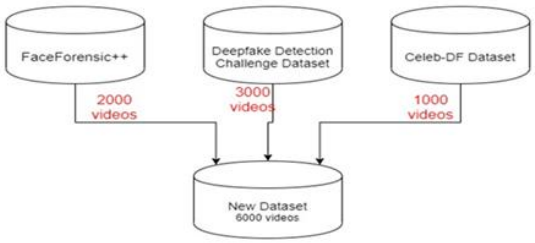

Fig 2 Dataset

iv) Pre-processing:

The process of editing involves cutting movies into individual frames. Each clip undergoes face recognition and editing to produce a library of films containing just faces. A threshold is established using the average number of frames every movie to maintain consistency and monitor system resources. The GPU's limitations led to the selection of a 150-frame threshold. Only the first 150 images are retained in order to maintain simplicity. A random arrangement of frames will not allow LSTM to function. All edited videos are stored at 30 frames per second with a resolution of 112 by 112 pixels. This approach ensures consistent findings that are compatible with other studies and speeds up processing. For the purpose of detecting deepfakes, this facilitates the application of deep learning algorithms.

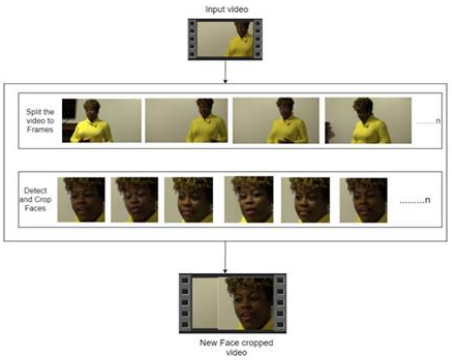

Fig 3 Pre-processing of video

v) Training & Testing:

The train dataset has 4,200 films, while the test dataset contains 1,800 movies. The 70/30 split is comprised of this. Half of the movies in the test group are fictitious and half are actual, making for an equal split between the two groups.

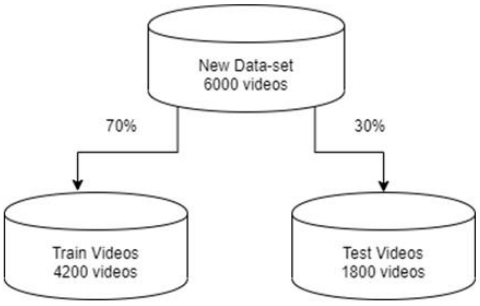

Fig 4 Split the dataset

vi) Model Architecture:

On our plan, you can find both CNN and RNN. We extracted the frame-level features using a ResNext CNN model that had already been trained. After that, a long short-term memory (LSTM) network was taught to distinguish between authentic and deepfake movies. To train the model, the Data Loader loads the labels from the training split of videos into the model.
ResNext: We used the ResNext show, which had been told to acquire features, instead of writing code. If you're looking for a Leftover CNN structure that's perfect for deep neural networks, ResNext might be it. Resnext50 32x4d was used for laboratory testing. We used a ResNext with fifty layers and a 32x4 grid.
By including layers and selecting an appropriate learning rate, we may optimize the arrangement for merging the model's gradient descent. Highlight vectors of 2048 dimensions are sent into the next LSTM after ResNext's final pooling layers.

LSTM for Analysis of Sequences: Include vectors of 2048 dimensions are fed into the LSTM. To prove our claim, all you need is a single LSTM layer with 2048 idle measures, 2048 hidden layers, and 0.4 dropout. Looking at the progression of the film from moment 't' to moment 't-n' reveals its development. Lstm is used in this case. There could be a variety of outlines before t. One of the show's components is an enactment of Defective Relu. The typical input-output association rate may be calculated using a direct layer with 2048 input and 2 yield highlights. The picture estimate to transmit is determined by H x W, and the show has a flexible normal surveying layer with a yield parameter of 1. In a sequential fashion, outlines are formed by a Successive Layer. For group preparation, four individuals are used. The forecast confidence is shown by the SoftMax layers.

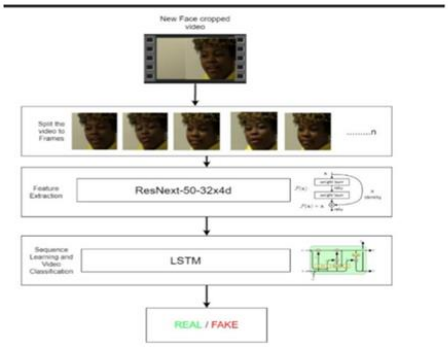

Fig 6 Overview of our model

vii) Hyper-parameter tuning:

After a lot of trial and error, we have settled on the optimal hyperparameters for our dataset. We improve gradient descent and facilitate flexible learning by using the Adam optimizer, which has a learning rate of 1e-5 and a weight loss of 1e-3. For tasks involving classification, such as ours, we use cross-entropy loss. To maximize the use of the available computers, batch training is conducted with a size of 4. In our office, this has been a huge success. Using the Django framework, which facilitates the addition of more features down the road, the user experience was created. One may upload videos to the model for prediction purposes by clicking the "video file" button on the index.html page. On top of the playing video in predict.html, you can see predictions regarding the video's authenticity and the confidence of the model. This enhances the user experience and makes it simpler for them to grasp.

4. EXPERIMENTAL RESULTS

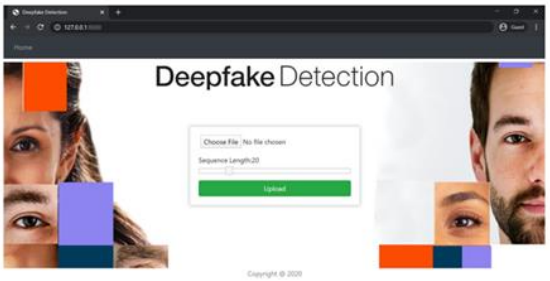

Fig 7 Home Page

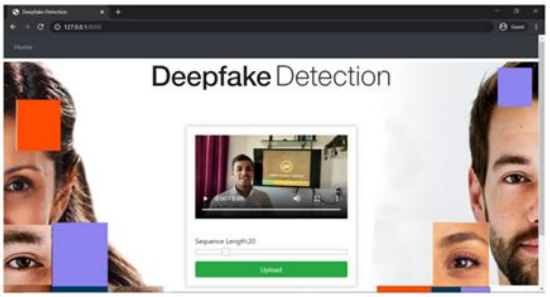

Fig 8 Uploading real video

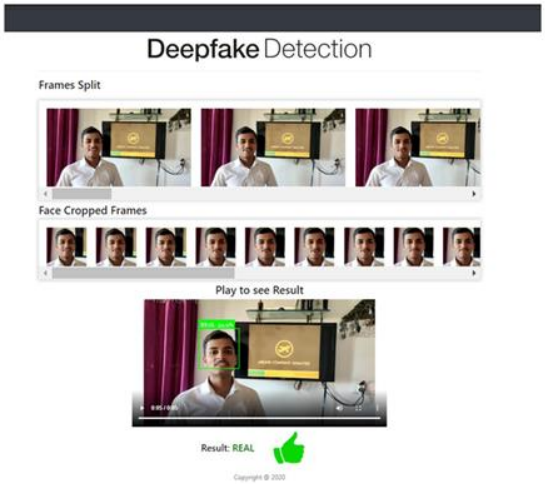

Fig 9 Real Video Output

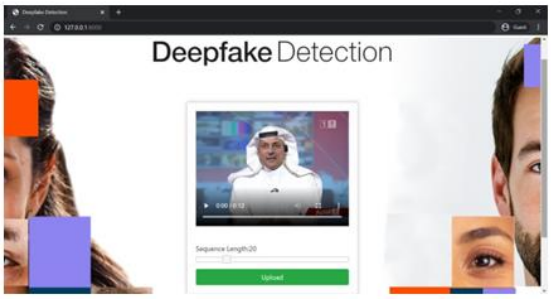

Fig 10 Uploading fake video

Fig 11 Fake video output

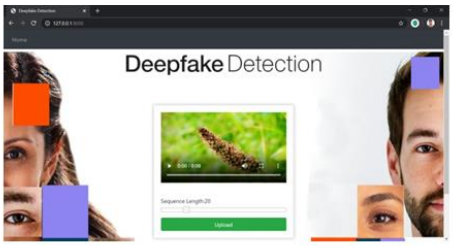

Fig 12 Uploading video with no faces

Fig 13 Output of uploaded video with no faces

5. CONCLUSION

To find out whether a video is a deep fake and how much faith we have in the model, we used neural networks. Once our method processes only one second of video (10 frames per second), it can accurately forecast the result. The model was built by combining a pre-trained ResNext CNN model for feature acquisition at the frame level with an LSTM for time sequence analysis, allowing us to find changes between frames t and t-1. The range of possible movie frame numbers is10,20,40,60,80, and 100.

VI. OUTLINE FOR THE REST OF THE WORK Thanks to its state-of-the-art construction, the developed system has room for future enhancements. To make it more user-friendly, it may be converted into an app for computers. The algorithm's utility and applicability are enhanced by the fact that it can be trained to detect both face and full-body deepfakes.

REFERENCES

[1] Andreas Rossler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies,Matthias Nießner, "FaceForensics++: Learning to Detect Manipulated Facial Images" in arXiv:1901.08971.

[2] Deepfake detection challenge dataset : https://www.kaggle.com/c/deepfake-detection-challenge/data Accessed on 26 March, 2020

[3] Yuezun Li , Xin Yang , Pu Sun , Honggang Qi and Siwei Lyu "Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics" in arXiv:1909.12962

[4] Deepfake Video of Mark Zuckerberg Goes Viral on Eve of House A.I. Hearing : https://fortune.com/2019/06/12/deepfake-mark-zuckerberg/ Accessed on 26 March, 2020

[5] 10 deepfake examples that terrified and amused the internet : https://www.creativebloq.com/features/deepfake-examples Accessed on 26 March, 2020

[6] TensorFlow: https://www.tensorflow.org/ (Accessed on 26 March, 2020)

[7] Keras: https://keras.io/ (Accessed on 26 March, 2020)

[8] PyTorch : https://pytorch.org/ (Accessed on 26 March, 2020)

[9] G. Antipov, M. Baccouche, and J.-L. Dugelay. Face aging with conditional gen- erative adversarial networks. arXiv:1702.01983, Feb. 2017

[10] J. Thies et al. Face2Face: Real-time face capture and reenactment of rgb videos. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2387–2395, June 2016. Las Vegas, NV.

[11] Face app: https://www.faceapp.com/ (Accessed on 26 March, 2020)

[12] Face Swap : https://faceswaponline.com/ (Accessed on 26 March, 2020)

[13] Deepfakes, Revenge Porn, And The Impact On Women : https://www.forbes.com/sites/chenxiwang/2019/11/01/deepfakes-revenge-porn-and- the-impact-on-women/

[14] The rise of the deepfake and the threat to democracy: https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of- the-deepfake-and-the-threat-to-democracy(Accessed on 26 March, 2020)

[15] Yuezun Li, Siwei Lyu, "ExposingDF Videos By Detecting Face Warping Arti- facts," in arXiv:1811.00656v3.

[16] Yuezun Li, Ming-Ching Chang and Siwei Lyu "Exposing AI Created Fake Videos by Detecting Eye Blinking" in arXiv:1806.02877v2.

[17] Huy H. Nguyen , Junichi Yamagishi, and Isao Echizen " Using capsule net- works to detect forged images and videos " in arXiv:1810.11215.

[18] D. Güera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," 2018 15th IEEE International Conference on Advanced Video and Sig- nal Based Surveillance (AVSS), Auckland, New Zealand, 2018, pp. 1-6.

[19] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld. Learning realistic hu- man actions from movies. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, June 2008. Anchorage, AK

[20] Umur Aybars Ciftci, ˙Ilke Demir, Lijun Yin "Detection of Synthetic Portrait Videos using Biological Signals" in arXiv:1901.02212v2

[21] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv:1412.6980, Dec. 2014.

[22] ResNext Model : https://pytorch.org/hub/pytorch_vision_resnext/ accessed on 06 April 2020

[23] https://www.geeksforgeeks.org/software-engineering-cocomo-model/ Accessed on 15 April 2020

[24] Deepfake Video Detection using Neural Networks
http://www.ijsrd.com/articles/IJSRDV8I10860.pdf

[25] International Journal for Scientific Research and Development http://ijsrd.com/