

Deep Fake Video Detection using Neural Networks

Ancy Y Das¹, Sinchana G², Sushmitha C S³, Venkatesh L⁴ and Ysaswvne⁵

¹Assistant Professor, Department of Computer Science and Engineering at BGS Institute of Technology, Adichunchanagiri University, BG Nagara, Karnataka
Email: ancydas@bgsit.ac.in

²⁻⁵Department of Computer Science and Engineering at BGS Institute of Technology, Adichunchanagiri University, BG Nagara, Karnataka
Email: gowdasinchana324@gmail.com, cssushmitha38@gmail.com, venkatesh.l8402@gmail.com, yashaswigowda902@gmail.com

Abstract—As of late, the multiplication of free profound learning programming apparatuses has fundamentally worked with the advancement of practical face trades in recordings, bringing about the development of "DeepFake" (DF) recordings with negligible hints of control. These AI-generated media, commonly known as DF, can now be generated with relative ease using artificially intelligent tools. However, despite the simplicity of generating DF using AI tools, detecting such manipulated content remains a significant challenge. Training algorithms to identify DF is inherently complex. We expect to achieve promising results against a comprehensive dataset of fake videos sourced from standard repositories.

I. INTRODUCTION

The rising openness of top notch web associations overall has broadened the extending impact of virtual entertainment. Also, media-sharing stages have improved on the creation and conveyance of computerized videos. more simple than at any other time. Progressions in computational power have moved profound figuring out how to remarkable degrees of capacity, outperforming what was considered unthinkable only a couple of years prior. As with any ground breaking technology, these advancements have inevitably ushered in a new set of challenges.

"The dissemination of false information (DF) on social media platforms has become increasingly prevalent, resulting in spamming and the circulation of inaccurate information. This phenomenon contributes to the spread of misinformation across these platforms." The existence of Deepfake (DF) technology poses significant risks, potentially leading to threats and the dissemination of misleading information among the general populace.[2] To address this critical issue, the detection of Deepfakes (DF) becomes paramount. Therefore, we present an innovative system capable of precisely distinguishing AI-generated fake videos (DF Videos) from genuine ones. The imperative lies in the development of technology equipped to detect fraudulent content, thereby thwarting the dissemination of Deepfakes (DF) across the internet. To effectively detect Deepfakes (DF), it's crucial to comprehend the underlying process through which Generative Adversarial Networks (GANs) generate such content. The Generative Adversarial Network (GAN) operates by taking a video input along with an image of a particular individual (referred to as the 'target'), They then generate a modified video in which the faces of the target are replaced with those of another individual (the 'source'). Deep adversarial neural networks serve as the foundation of Deepfake (DF) technology. They are prepared on datasets including face pictures and target recordings, empowering them to plan the facial highlights and demeanors of the source onto the objective

consequently. Following post processing refinement, subsequent recordings can attain a high degree of authenticity.

II. LITERATURE SURVEY

One method utilized for deepfake detection involves model to detect artifacts. This study addresses two distinct types of facial artifacts. It capitalizes on the observation that existing deepfake algorithms often produce images with limited resolutions, necessitating additional transformations to align the replaced faces with those in the source video. Assessment over eye-squinting discovery benchmarks shows promising execution in distinguishing recordings produced with Profound Brain Organization based programming for deepfake creation. While the essential dependence of this strategy on the shortfall of squinting for discovery is noted, it is underlined that different boundaries, like teeth upgrade and facial kinks, ought to be considered for powerful deepfake location. Our proposed strategy coordinates thought of every one of these parameters. The utilization of case organizations to identify fashioned pictures and recordings is investigated in [3]. While randomly adding noise during training may not be the most optimal choice, the model still shows promising results on the given dataset. Fake Catcher achieves high accuracy in detecting fake content, regardless of the generator used or the video's resolution and quality. However, the absence of a discriminator in their approach compromises fidelity in preserving biological signals.

III. PRELIMINARIES

There are many tools available to create Deep Fakes, but not enough resources to detect them. Our approach aims to help stop the spread of DeepFakes online. Our method is designed to detect all types of DeepFakes, including substitution, conservation, and relational ones.

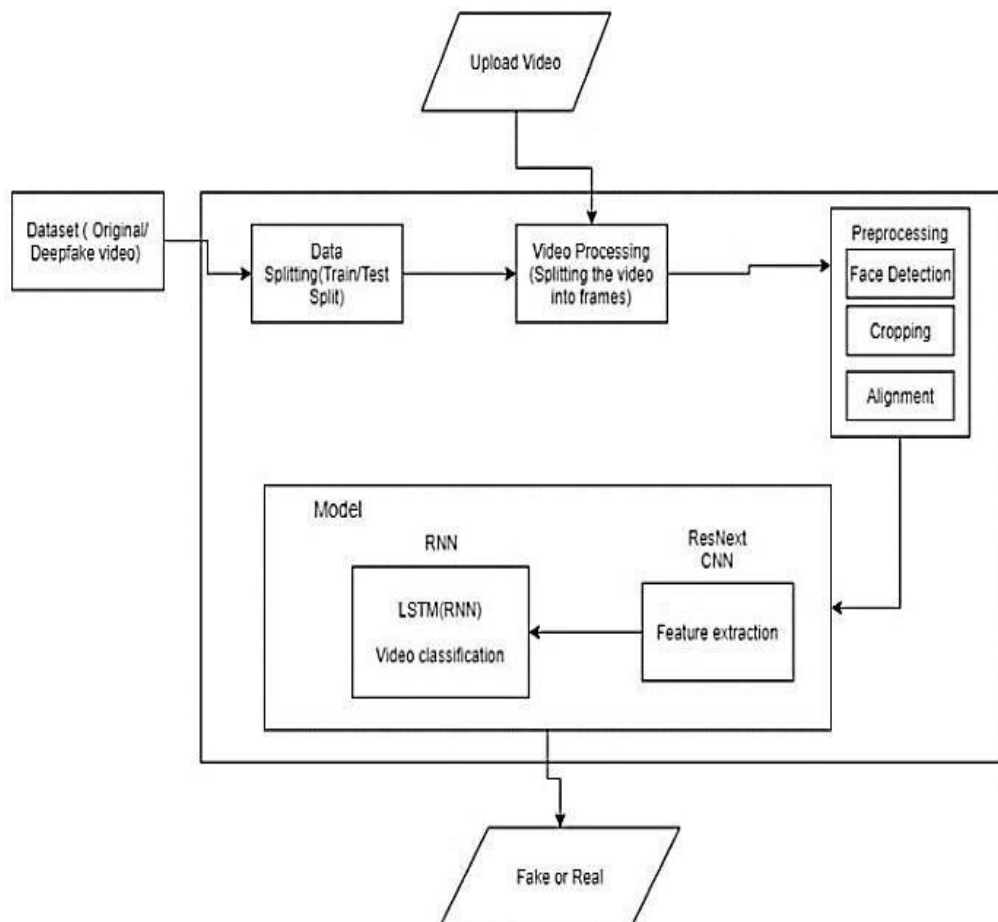


Figure 1: System Architecture

A. Data Collection

The dataset utilized is a blend of videos sourced equally from various datasets such as youtube Face Forensics++ and the Deepfake Detection Challenge dataset.

B. Pre-processing

At first, the dataset goes through preprocessing, including the division of recordings into individual casings. Following this, facial detection algorithms are applied to identify and extract the regions containing human faces within each frame. Frames without detectable faces are excluded from further processing. Given the computational demands of processing videos at a standard rate of 30 frames per second, a pragmatic approach is adopted, restricting the training data to the initial 100 frames per video for experimental purposes.

C. Model

The proposed model architecture comprises a resnext50_32x4d backbone augmented with a single Long Short Term Memory layer. The Data Loader module is responsible for ingesting preprocessed videos containing cropped facial regions, which are subsequently partitioned into distinct training and testing sets. Smaller than usual groups of casings got from the preprocessed recordings are then taken care of into the model for both preparation and testing stages.

D. ResNext Convolutional Neural Network

As an alternative to building the classifier from scratch, we suggest using the ResNeXt Convolutional Neural Network model. This pre-trained model can effectively extract features and accurately identify components. To fine tune the model, we'll add crucial additional layers and determine an appropriate learning rate.

E. Long Short-Term Memory

which serves as the foundation for temporal management. We use a sequence of ResNext Convolutional Neural Network feature vectors from input frames as input data, along with a two-node neural network for sequence classification probabilities.

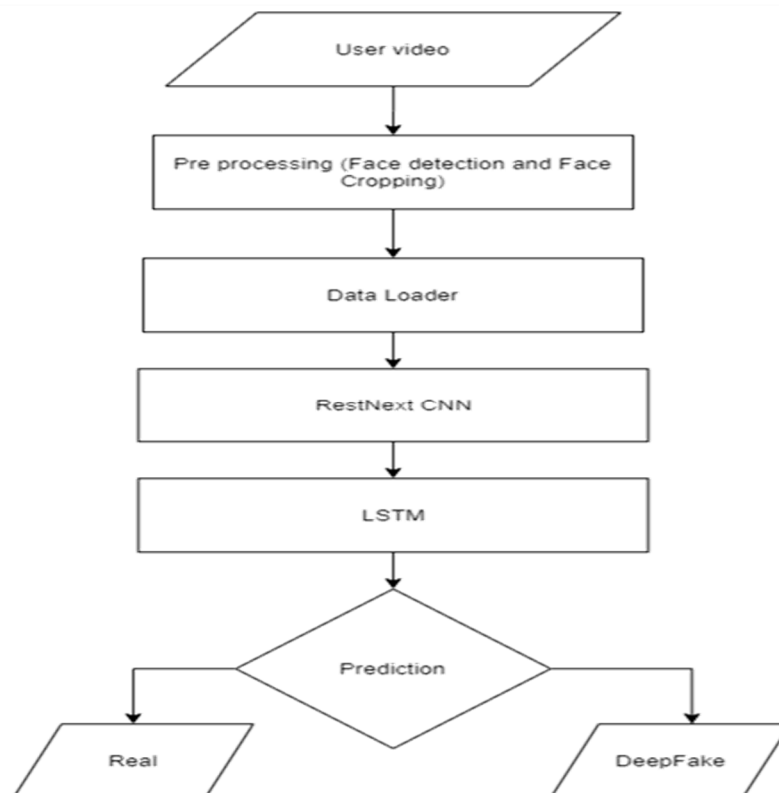


Figure 2: Preparing Stream



Figure 3: Expected Results

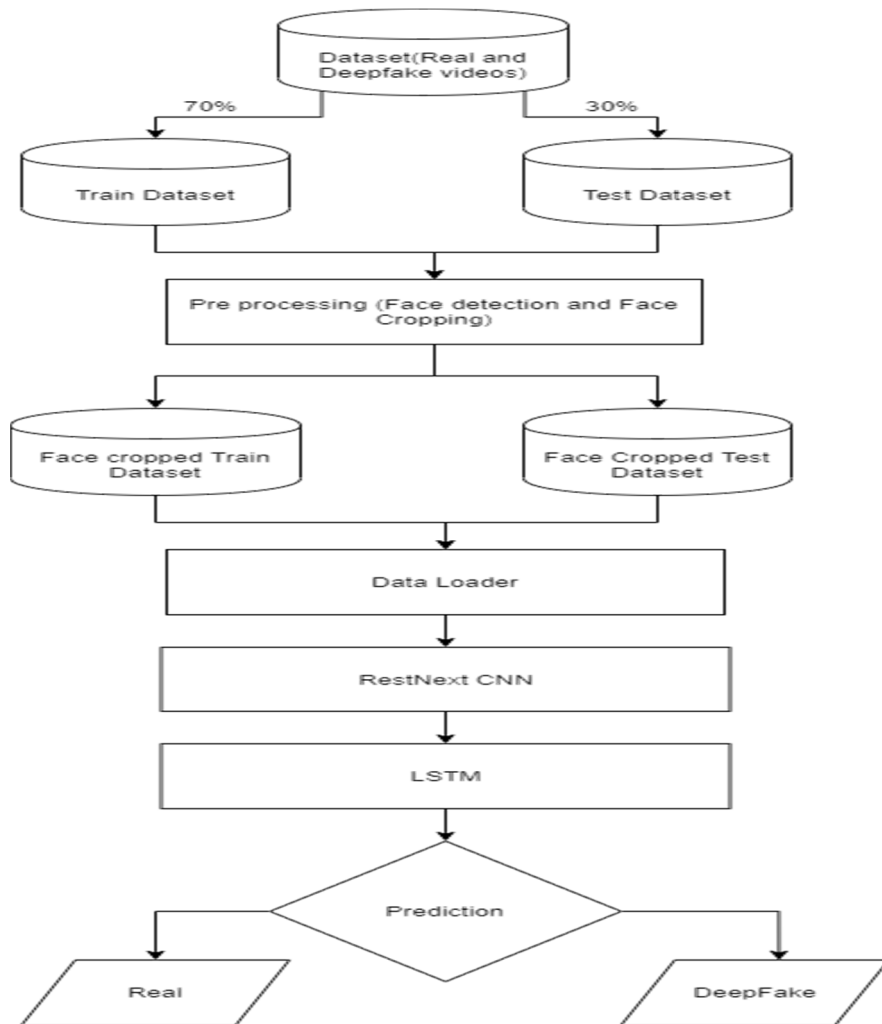


Figure 4: Prediction flow

IV. CONCLUSION

Our study introduces a neural network-driven approach for differentiating between authentic and deep fake videos, while also offering insights into the model's confidence levels. Inspired by the techniques employed in generating deep fakes using Generative Adversarial Networks (GANs) in tandem with Auto encoders, By considering various parameters outlined in our research, our proposed approach demonstrates the capability to accurately identify deep fake content in real-time scenarios. We are confident in the effectiveness of our method and anticipate achieving high accuracy rates when applied to real-world data.

REFERENCES

- [1] Yuezun Li, Siwei Lyu, "ExposingDF Videos By Detecting Face Warping Artifacts".
- [2] Yuezun Li, Ming-Ching Chang and Siwei Lyu "Exposing AI Created Fake Videos by DetectingEye Blinking".
- [3] Huy H. Nguyen , Junichi Yamagishi, and Isao Echizen " Using capsule networks to detect forgedimages and videos ".
- [4] Hyeonwoo Kim, Pablo Garrido, Ayush Tewari and Weipeng Xu "Deep Video Portraits".
- [5] Darius Afchar,Vincent Nozick,Junichi Yamagishi,Isao Echizen "Compact Facial Vedio ForgeryDetection Network".
- [6] Ian Good fellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, SherjilOzair,Aaron Courville, and Yoshua Bengio. Generative adversarial nets.
- [7] David G`uera and Edward J Delp. Deepfake video detection using recurrent neural networks.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning orimagerecognition.
- [9] Nurzhan Zhumabekuly Aitzhan; Davor Svetinovic "Exposing Deep Fakes Using InconsistentHeadPoses".
- [10] Long Short-Term Memory: From Zero pytorch: <https://blog.floydhub.com/long-short-term-memory-from-zero-to-hero-with-pytorch/>