WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Deepfake Image/Video Detection Using Deep Learning

J. Rahul[1], M. Rahul[2], G. Raj Kumar[3], N. Rajashekar[4], D. Rajeswari[5], Prof Thanish Kumar[6]

[1,2,3,4,5,6] [CSE]-AI&ML, MRUH

[1]2111cs020375@mallareddyuniversity.ac.in, [2] 2111cs020376@mallareddyuniversity.ac.in, [3]2111cs020377@mallareddyuniversity.ac.in, [4]2111cs020378@mallareddyuniversity.ac.in, [5]2111cs020380@mallareddyuniversity.ac.in

ABSTRACT

Deepfake (DF) technology presents a significant challenge in the digital era, allowing for the creation of highly realistic fake videos. While tools for generating Deepfakes are plentiful, effective detection mechanisms remain sparse. This paper introduces a new approach to combat the dissemination of Deepfake content online through a web-based platform. Users can upload videos to the platform, which utilizes advanced algorithms to classify and detect potential Deepfakes. Integration with widely-used applications like WhatsApp and Facebook enables real-time verification of content authenticity before distribution. Emphasizing security, usability, accuracy, and reliability, the platform aims to detect various forms of Deepfakes including substitution, impersonation, and interpersonal deception. Evaluation metrics will assess the platform's performance and efficacy in addressing the proliferation of Deepfake content on the internet. This research seeks to contribute to the ongoing efforts to safeguard digital media integrity and mitigate the societal impact of manipulated video content.

1. INTRODUCTION

The proliferation of Deepfake technology has created significant challenges in verifying the authenticity of digital images. Deepfakes, generated using advanced AI techniques, can convincingly alter visual content, posing risks for misinformation and malicious activities. Detecting these manipulated images is crucial due to their increasing sophistication, which makes them difficult to distinguish from genuine photographs. This paper introduces an approach to tackle this issue by employing machine learning and image analysis techniques. By developing robust detection methods, we aim to accurately differentiate between authentic and synthetic images. Our research seeks to enhance the trustworthiness of digital visual content, addressing both the technical challenges and the broader societal impacts of Deepfake proliferation.

2. LITERATURE REVIEW

The detection of deepfake images has become a significant focus within the research community due to the sophisticated techniques used to create these forgeries. Researchers have developed various methods to address this issue, with several promising approaches emerging. Traditional Convolutional Neural Network (CNN)-based methods are widely used, leveraging their ability to learn and detect intricate patterns within images. However, these methods often face challenges in generalizing across different datasets, which affects their performance when encountering new types of deepfakes (MDPI) (MDPI). To enhance detection capabilities, more advanced techniques have been developed, including the use of transformers. Transformer-based approaches are particularly effective in capturing long-range dependencies within image data, which are crucial for identifying subtle manipulations that may not be evident with CNNs alone (MDPI). Additionally, detection methods utilizing biological signals, such as eye blinking and pulse detection, offer another layer of robustness, as these signals are challenging to replicate accurately in deepfake videos (MDPI). Datasets play a critical role in the development and testing of deepfake detection algorithms. The FaceForensics++ dataset is a prominent example, providing a comprehensive collection of real and manipulated videos using various techniques. Similarly, the DeepFake Detection Challenge (DFDC) dataset, sponsored by major technology companies, offers a diverse set of deepfake videos created with different methods, facilitating the development of more generalized detection models (MDPI). The Celeb-DF dataset further contributes by offering high-quality and realistic deepfake videos, which help improve the robustness of detection models (MDPI).

3. PROPOSED SYSTEM

Numerous tools are available for generating Deepfake (DF) content, whereas only a limited number focus on detecting it. Our initiative aims to effectively prevent the dissemination of DF content across the World Wide Web. We propose developing a web-based platform that allows users to upload videos for classification as potentially altered. Leading applications such as WhatsApp and Facebook could seamlessly integrate this solution to verify the authenticity of content before it is shared with other users. Our primary goals include evaluating performance and ensuring the platform's security,

usability, accuracy, and reliability. Our approach is designed to detect various types of DF, including substitution, impersonation, and interpersonal deception.

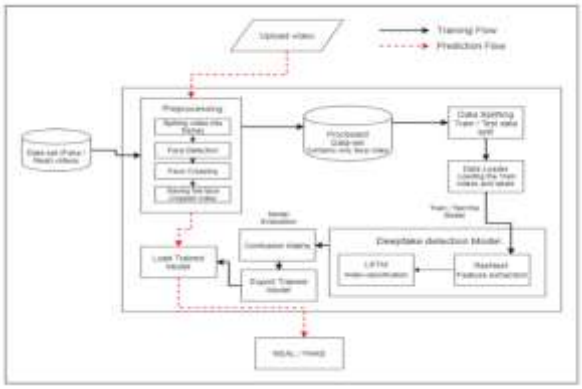

Fig. 1: System Architecture

4.METHODOLOGY

A. Dataset: We utilize a comprehensive dataset amalgamating videos from various sources, including FaceForensics++ [14], and the Deepfakes Detection Challenge dataset [13]. Our new dataset composition comprises 50% original photographs and 50% manipulated artworks. The dataset is partitioned into two segments: 70% allocated for training and 30% for testing.

B. Preprocessing: Dataset preprocessing encompasses several steps. Initially, videos are segmented into individual frames. Subsequently, face detection is performed, and the frames containing faces are cropped accordingly. To maintain consistency in the number of frames, we average the frame count across the video dataset and implement a new face cropping method that ensures equal frame averages. In the initial phase, frames without detected faces are discarded.

C. Model: Our model architecture includes a ResNeXt-50 32x4d network followed by a Long Short-Term Memory (LSTM) layer. The Data Loader is responsible for loading the preprocessed, face-cropped videos and dividing them into training and testing sets. The frames from these processed videos are then fed into the model in mini-batches for both training and testing purposes.

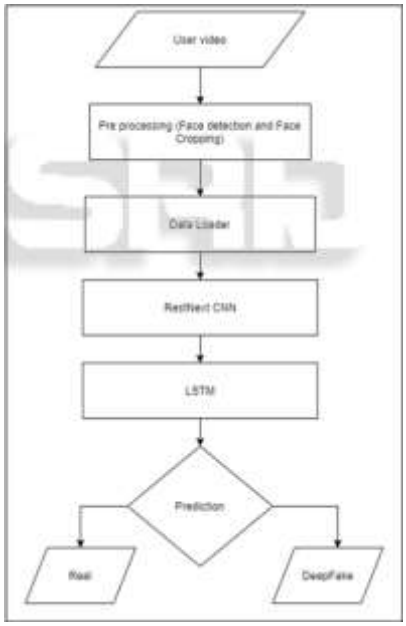

Figure 2 :Data flow diagram

D. ResNext CNN for feature extraction We recommend using the ResNext CNN classifier to extract features and provide frame-level features rather than rewriting the classifier. Next, we will finetune the network by adding the necessary layers and selecting the appropriate training value to accurately distinguish the slope of the model. The 2048dimensional feature vector after the last pooling layer will be used as the LSTM input array.

5.EXPERIMENTAL RESULTS

Figure 3:Application output

Figure 4:Resultant screen

6.CONCLUSION

We propose a neural networkbased approach to classify videos as deep or real by relying on the proposed model. The scheme is inspired by the way GANs create deepfakes with the help of autoencoders. Our method uses ResNext CNN for phase detection and RNN and LSTM for video classification. The plan can determine whether the video is deep or real based on what is not stated in the document. We believe that you will provide information very quickly.

7.FUTURE ENHANCEMENT

Autoencoder. Our method uses ResNext CNN for phase detection and RNN and LSTM for video classification. The plan can determine whether the video is deep or real based on what is not stated in the document. We believe that you will provide information very quickly.

8.REFERENCES

[1] "DeepFakes and beyond: Forgery and detection analysis", Yuezun Li, MingChing Chang, Siwei Lyu (2019). This paper provides an overview of the technology and findings. Thies, Matthias Nina (2019). This paper introduces a technique to detect deepfake movies by analyzing facial distortion artifacts. , Luisa Verdoliva, Christian Rees, Justus Thies, Matias Nina (2019). This paper introduces FaceForensics++ for forensics, data and benchmarks. This paper describes the measurement data for measuring depth measurement.