FAKE DETECTION USING DEEP LEARNING

PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the degree of
MASTER OF COMPUTER SCIENCE WITH DATA ANALYTICS
At the Bharathiar University

By

RAJIVMENAN A
(Reg.No:222DA012)

Under the Guidance of

Mrs. S. Shenbaha., M.Sc., M.Tech., (Ph.D)
Assistant Professor
DEPARTMENT OF COMPUTER SCIENCE WITH DATA ANALYTICS

Dr. N.G.P. ARTS AND SCIENCE COLLEGE
(An Autonomous Institution, Affiliated to Bharathiar University, Coimbatore)
Approved by Government of Tamil Nadu & Accredited by NAAC with A^{++} Grade (3rd Cycle -3.64 CGPA)
Dr. N.G.P.-Kalapatti Road, Coimbatore-641 048, Tamil Nadu, India.
Website: www.drngpasc.ac.in | Email: info@drngpasc.ac.in. | Phone: +91-422-2369100

MAY - 2024

CERTIFICATE

This is to certify that the project, entitled **“FAKE DETECTION USING DEEPLEARNING”** submitted in partial fulfillment of the requirement for the award of the degree of **Master of Computer Science with Data Analytics at the Bharathiar University** isa record of original project work done by **Mr A. RAJIVMENAN (222DA012)** during the period (2022-2024) of her study in **Department of Master of Computer Science with Data Analytics, Dr. N. G. P. Arts and Science College, Coimbatore- 48** under my supervision and guidance, and the project has not formed the basis for the award of any Degree/ Diploma/ Associateship/ Fellowship or other similar title to any candidate of any university.

(Mrs.S.Shenbaha)	(Dr. V.Pream Sudha)	(Prof. K. Ramamurthi)
Project Guide	Professor and Head	Principal(i/c)

Place: Coimbatore

Date:

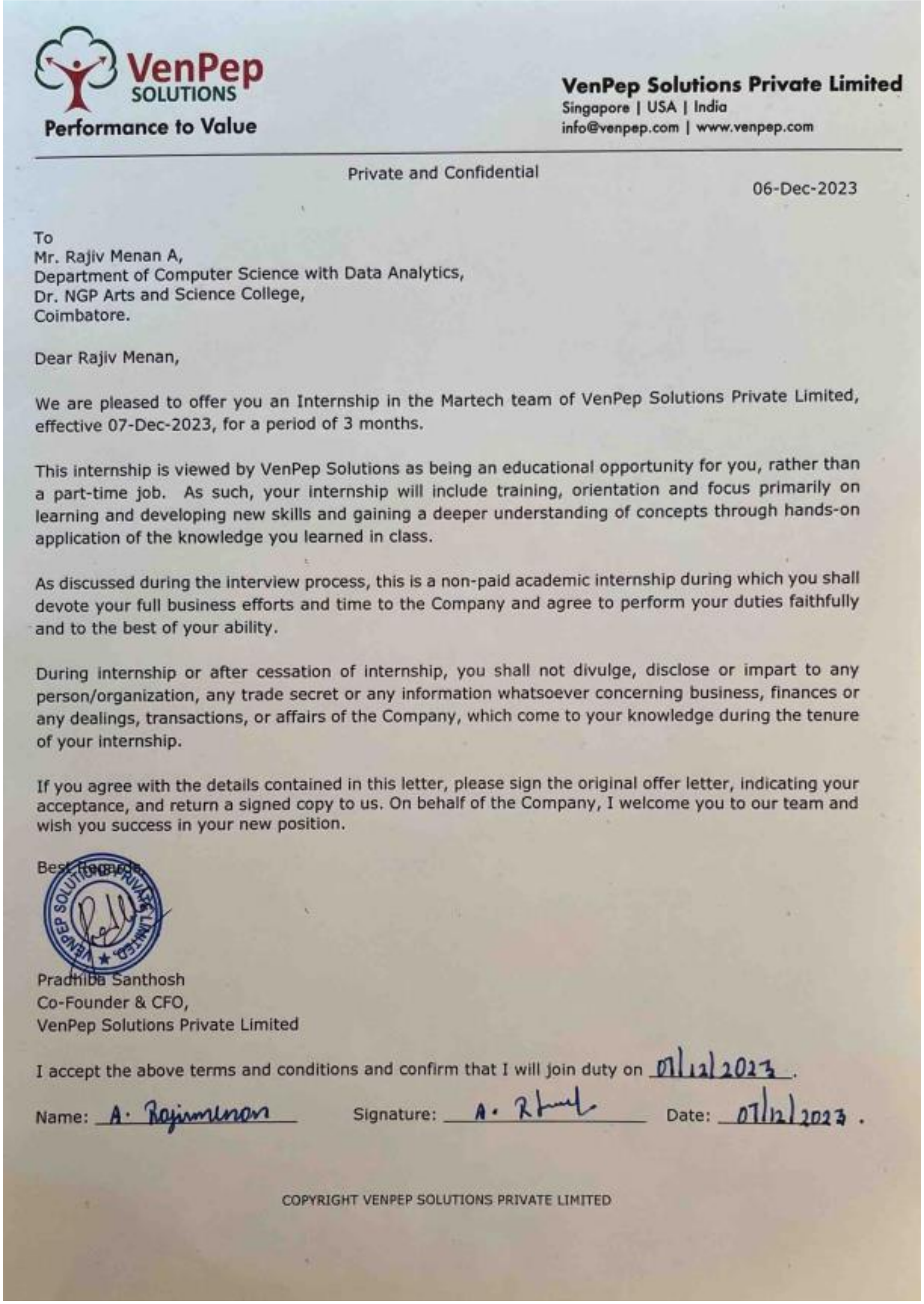
Viva-voce Examination held on :

Internal Examiner **External Examiner**

DECLARATION

I hereby declare that the project report entitled **"FAKE DETECTION USING DEEP LEARNING "** submitted in partial fulfillment of the requirement for the award of the degree of Master of Computer Sciencewith Data Analytics is a record of original project work done during the period of study supervision and under the guidance of **Mrs. S. Shenbaha, M.Sc., M.Tech, (Ph.D)., Department of Computer Science with Data Analytics, Dr. N. G. P. Arts and Science College, Coimbatore - 48,** and it has not formed on the basis of award of any Degree/ Diploma/ Associateship/ Fellowship or other similar title to any candidate of any university.

Place: Coimbatore **RAJIVMENAN A**
Date: **(222DA012)**



Performance to Value

VenPep Solutions Private Limited
Singapore | USA | India
info@venpep.com | www.venpep.com

Private and Confidential

06-Dec-2023

To
Mr. Rajiv Menan A,
Department of Computer Science with Data Analytics,
Dr. NGP Arts and Science College,
Coimbatore.

Dear Rajiv Menan,

We are pleased to offer you an Internship in the Martech team of VenPep Solutions Private Limited, effective 07-Dec-2023, for a period of 3 months.

This internship is viewed by VenPep Solutions as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training, orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts through hands-on application of the knowledge you learned in class.

As discussed during the interview process, this is a non-paid academic internship during which you shall devote your full business efforts and time to the Company and agree to perform your duties faithfully and to the best of your ability.

During internship or after cessation of internship, you shall not divulge, disclose or impart to any person/organization, any trade secret or any information whatsoever concerning business, finances or any dealings, transactions, or affairs of the Company, which come to your knowledge during the tenure of your internship.

If you agree with the details contained in this letter, please sign the original offer letter, indicating your acceptance, and return a signed copy to us. On behalf of the Company, I welcome you to our team and wish you success in your new position.

Best Regards,

VENPEP SOLUTIONS PRIVATE LIMITED

Pradhiba Santhosh
Co-Founder & CFO,
VenPep Solutions Private Limited

I accept the above terms and conditions and confirm that I will join duty on 07/12/2023.

Name: A. Rajivmenon Signature: A. R Date: 07/12/2023.

ACKNOWLEDGEMENT

This project was the most significant accomplishment in my life and it would not have been possible without the blessing of God almighty and those who supported and believed in my caliber.

I record a deep sense of gratitude to **Dr. NALLA. G. PALANISWAMI, M.D, AB (USA)**, Chairman, Kovai Medical Center Research and Educational Trust (KMCRET) and Dr. THAVAMANI. D. PALANISWAMI, M.D, AB (USA), Secretary, Dr. N.G.P. Arts and Science College, Coimbatore for providing me an infrastructure facility to carry out project work successfully.

I record my sincere thanks to **DR. K. RAMAMURTHI M.COM., BL., MBA., PH.D., PRINCIPAL** of Dr.N.G.P. Arts and Science College, Coimbatore (i/c), for every help he rendered before andduring the project.

I express my sincere thanks to **DR.V.PREAM SUDHA, MCA., M.Phil.,Ph.D., He**ad of the Department of Computer Science with Data Analytics , Dr. N. G. P. Arts and Science College, Coimbatore – 48 for showing sustained interest and providing helpthroughout the period of our work.

I would like to extend the sincere thanks to my guide **MRS. S. SHENBAHA, M.Sc., M.Tech, (Ph.D).,** Head of the department of Computer Science with Data Analytics, Dr.N.G.P.Arts and Science College, Coimbatore - 48.I sincerely thank for her exemplary guidance and encouragement.

I take this opportunity to acknowledge my sincere thanks to all the staff members of the Department of Computer Science with Data Analytics for their constant inspiration, assistance and resourceful guidance for the completion of this project successfully.

I express our sincere thanks to our family and friends for their encouragement, love,prayer, moral support, advice and sacrifice without which I would not have been able to pursue the course of our study.

RAJIVMENAN A

ABSTRACT

In the digital era, the proliferation of deepfake technologies has posed unprecedented challenges to the integrity of digital media, necessitating the development of robust detection mechanisms. This study introduces a novel deep learning-based framework designed to identify and mitigate the spread of deepfakes with unparalleled accuracy. By harnessing the power of advanced neural network architectures, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), combined with innovative techniques such as transfer learning and data augmentation, our approach significantly improves the detection of deepfakes in varied and complex scenarios.

We begin by detailing the construction of a diverse dataset, encompassing a wide range of deepfake generation methods, to train and evaluate our model. Our framework leverages feature extraction capabilities of CNNs to recognize subtle artifacts in video frames that distinguish genuine from manipulated content. Additionally, RNNs are employed to analyze temporal inconsistencies across frames, a common limitation in deepfake videos. The integration of transfer learning enables our model to adapt to new, unseen deepfake techniques, ensuring robustness and scalability.

Comprehensive evaluations demonstrate that our model achieves state-of-the-art performance, outperforming existing methodologies in both accuracy and efficiency. We further explore the ethical implications of deepfake detection technologies, emphasizing the importance of balance between privacy and security. Our research contributes to the ongoing discourse on digital media authenticity, offering a scalable and effective solution to combat the deepfake phenomenon. This work not only advances the technical field of deepfake detection but also addresses critical societal concerns related to trust and truth in the digital landscape.

CONTENTS

CHAPTER -1

INTRODUCTION

In an era where the boundary between reality and fabrication is becoming increasingly blurred, the emergence of deepfake technology poses unprecedented challenges to the authenticity and trustworthiness of digital content. Deepfakes, which involve the synthesis of hyper-realistic media, such as images and videos, with the aid of artificial intelligence (AI) techniques, have garnered widespread attention due to their potential to deceive and manipulate audiences on a massive scale. From disseminating fake news and misinformation to perpetrating identity theft and privacy breaches, the ramifications of deepfake proliferation are profound and multifaceted.

Addressing this looming threat requires innovative approaches that can effectively distinguish between genuine and manipulated media. Deep learning, a subset of AI that has demonstrated remarkable prowess in pattern recognition and feature extraction, offers a promising avenue for combating deepfakes. In particular, Convolutional Neural Networks (CNNs), renowned for their ability to capture spatial hierarchies in visual data, have emerged as a cornerstone in the development of deepfake detection systems.

This project endeavors to leverage the power of CNNs to construct a robust and reliable deepfake detection framework. By harnessing the innate capacity of CNN architectures to discern subtle visual cues and anomalies indicative of deepfake manipulation, we aim to equip media platforms, content creators, and end-users with effective tools to combat the spread of deceptive and harmful content. Through meticulous dataset curation, model training, and rigorous evaluation, this research seeks to advance the state-of-the-art in deepfake detection, fostering trust, integrity, and accountability in the digital media ecosystem.

In this introduction, we provide an overview of the deepfake phenomenon, highlighting its implications for various domains and the urgent need for proactive countermeasures. We outline the significance of leveraging deep learning techniques, particularly CNNs, in addressing the challenges posed by deepfake proliferation. Furthermore, we delineate the objectives and scope of the project, delineating the key methodologies and anticipated outcomes. Through this endeavor, we endeavor to contribute to the ongoing efforts to safeguard the integrity and authenticity of digital media in an era of pervasive synthetic content.

CHAPTER - 2

SYSTEM ANALYSIS

2.1 EXISTING SYSTEM

Currently, deepfake detection systems predominantly rely on traditional computer vision techniques and heuristic-based approaches. These methods often involve the extraction of handcrafted features, such as facial landmarks, texture analysis, and temporal inconsistencies, followed by classification using machine learning algorithms. While these systems may achieve moderate success in detecting simple deepfakes, they often struggle to cope with the rapid evolution of deepfake generation techniques and the increasing sophistication of manipulated media

1. Yamin et al. (2020) proposed a framework for detecting deepfakes using Long Short-Term Memory (LSTM) networks, a specific type of recurrent neural network (RNN) that excels at handling sequential data like video frames .

2. Mescheler et al. (2021) explored using Xception, a deep convolutional neural network architecture, for deepfake detection, achieving promising results .

3. Mirza et al. (2022) investigated transfer learning with pre-trained 3D CNNs for deepfake detection, demonstrating the effectiveness of leveraging pre-trained models for this task .

4. Soriano-Garcia et al. (2023) focused on physiological signals for deepfake detection. Their work delves into analyzing eye blinking patterns and other physiological cues to identify manipulations .

5. Shao et al. (2023) tackled deepfake detection using audio analysis. Their study explores employing various audio features and deep learning models to detect inconsistencies in manipulated audio .

2.2 PROPOSED SYSTEM

The proposed system introduces a novel deep learning-based approach, specifically leveraging Convolutional Neural Networks (CNNs), for deepfake detection. Unlike traditional methods, which rely on manually engineered features, CNNs are capable of automatically learning discriminative features directly from raw pixel data. This enables the model to capture complex spatial patterns and subtle artifacts indicative of deepfake manipulation, thereby enhancing detection accuracy and robustness.

KEY COMPONENTS OF THE PROPOSED SYSTEM

Dataset Construction: A diverse and comprehensive dataset encompassing authentic and deepfake media samples across various contexts and quality levels is curated for training and evaluation purposes. This dataset serves as the foundation for training the CNN-based detection model.

CNN Architecture Design: The architecture of the CNN model is carefully designed to accommodate the intricacies of deepfake detection. Techniques such as multi-scale feature extraction, attention mechanisms, and adversarial training are incorporated to enhance the model's discriminative capabilities and resilience against adversarial attacks.

Training and Evaluation: The CNN model is trained on the curated dataset using state-of-the-art optimization techniques. Extensive experimentation and evaluation on benchmark datasets are conducted to assess the performance, robustness, and generalization capabilities of the proposed system.

Integration and Deployment: The trained CNN model is integrated into a user-friendly software tool or platform, enabling seamless integration into existing media processing pipelines or applications. End-users, including media platforms and content creators, can leverage the deepfake detection system to identify and mitigate the spread of manipulated content effectively.

BENEFITS OF THE PROPOSED SYSTEM:

- Enhanced Detection Accuracy: By leveraging deep learning techniques, particularly CNNs, the proposed system achieves higher detection accuracy and robustness compared to traditional heuristic-based approaches.

- Adaptability and Scalability: The CNN-based detection model can adapt to evolving deepfake generation techniques and scale to handle large volumes of media content.

- Real-time Detection: The efficient architecture of the CNN model enables real-time or near-real-time detection of deepfake content, facilitating timely intervention and mitigation efforts.

- Overall, the proposed system represents a significant advancement in the field of deepfake detection, offering a potent solution to combat the proliferation of manipulated media and safeguard the integrity of digital content.

CHAPTER – 3

SYSTEM REQUIREMENTS

3.1 HARDWARE REQUIREMENTS

Processor	:	Intel Xeon E5 2637
Ram	:	16GB
Storage	:	120GB
Graphic Card	:	NVIDIA GeForce GTX Titan (12 GB RAM)
Input Devices	:	Keyboard, Mouse
Output Devices	:	Monitor

3.2 SOFTWARE REQUIREMENTS

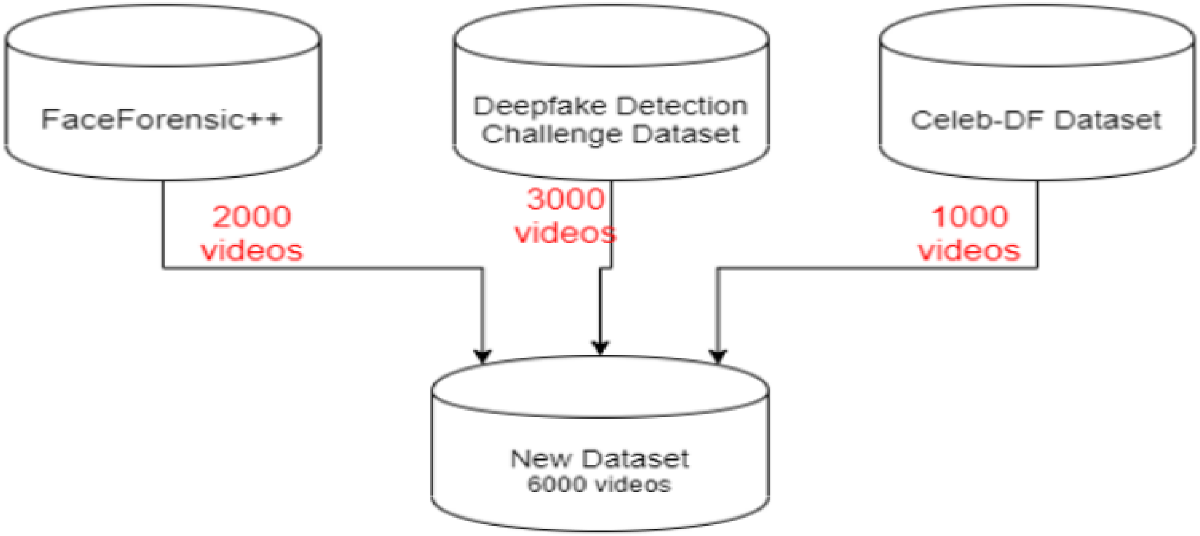
Operating System	:	Windows 7+
Programming Language	:	Python 3.0
Framework	:	PyTorch 3.0
Cloud platform	:	Google Cloud Platform
Libraries	:	OpenCV, Face-recognition

CHAPTER -4

SYSTEM OVERVIEW

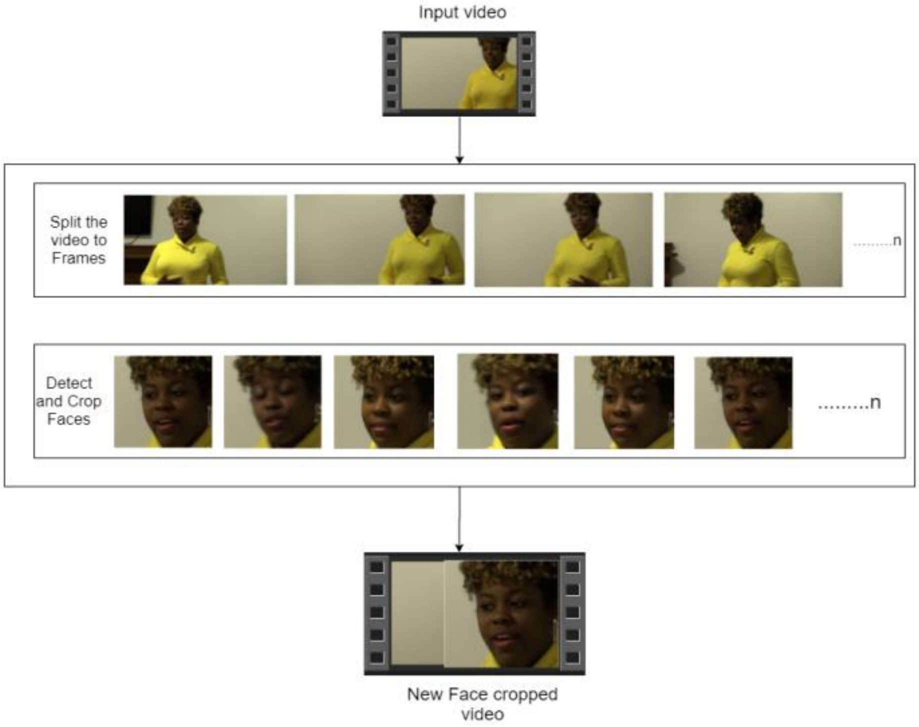
4.1 MODULE 1 : DATA-SET GATHERING

For making the model efficient for real time prediction.We have gathered the data from different available data-sets like FaceForensic++(FF)[1], Deepfake detection challenge(DFDC)[2], and Celeb-DF[3]. Futher we have mixed the dataset the collected datasets and created our own new dataset, to accurate and real time detection on different kind of videos. To avoid the training bias of the model we have considered 50% Real and 50% fake videos. Deep fake detection challenge (DFDC) dataset [3] consist of certain audio alerted video, as audio deepfake are out of scope for this paper. We preprocessed the DFDC dataset and removed the audio altered videos from the dataset by running a python script. After preprocessing of the DFDC dataset, we have taken 1500 Real and 1500 Fake videos from the DFDC dataset. 1000 Real and 1000 Fake videos from the FaceForensic++(FF)[1] dataset and 500 Real and 500 Fake videos from the CelebDF[3] dataset. Which makes our total dataset consisting 3000 Real, 3000 fake videos and 6000 videos in total. Figure 4.1 depicts the distribution of the data-sets



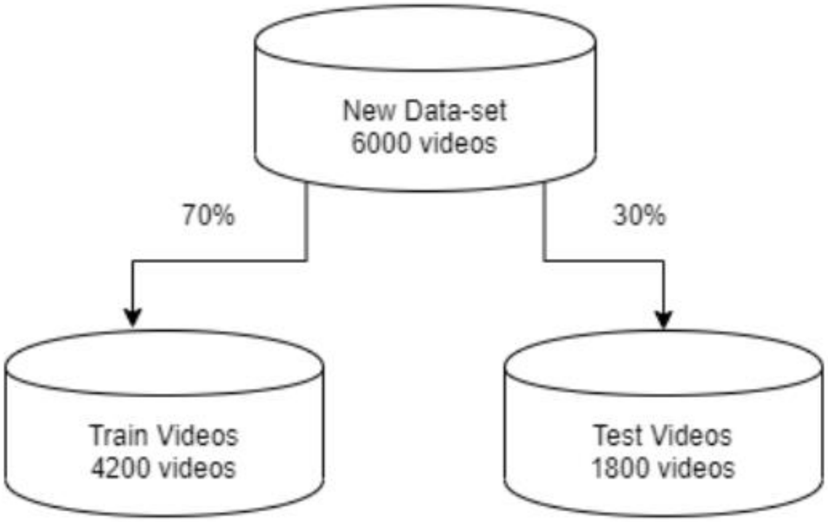
4.2 MODULE 2 : PRE-PROCESSING

In this step, the videos are preprocessed and all the unrequired and noise is removed from videos. Only the required portion of the video i.e face is detected and cropped. The first steps in the preprocessing of the video is to split the video into frames. After splitting the video into frames the face is detected in each of the frame and the frame is cropped along the face. Later the cropped frame is again converted to a new video by combining each frame of the video. The process is followed for each video which leads to creation of processed dataset containing face only videos. The frame that does not contain the face is ignored while preprocessing. To maintain the uniformity of number of frames, we have selected a threshold value based on the mean of total frames count of each video. Another reason for selecting a threshold value is limited computation power. As a video of 10 second at 30 frames per second(fps) will have total 300 frames and it is computationally very difficult to process the 300 frames at a single time in the experimental environment. So, based on our Graphic Processing Unit (GPU) computational power in experimental environment we have selected 150 frames as the threshold value. While saving the frames to the new dataset we have only saved the first 150 frames of the video to the new video. To demonstrate the proper use of Long Short-Term Memory (LSTM) we have considered the frames in the sequential manner i.e. first 150 frames and not randomly. The newly created video is saved at frame rate of 30 fps and resolution of 112 x 112.



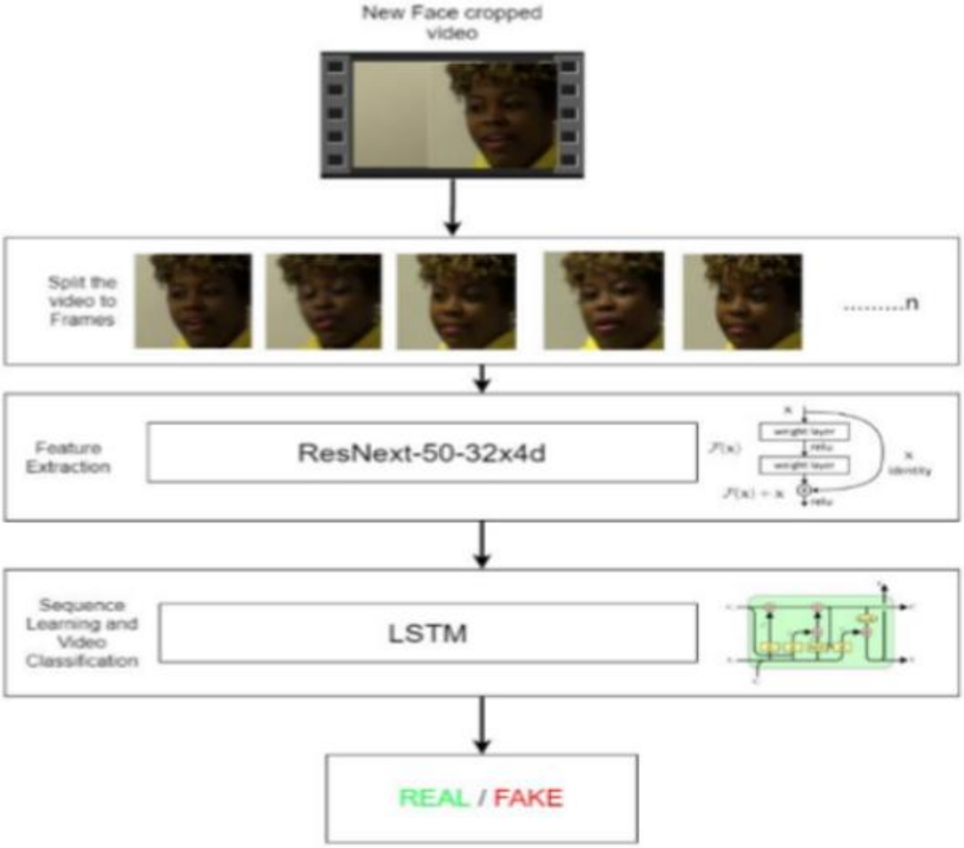
4.3 MODULE 3: DATA-SET SPLIT

The dataset is split into train and test dataset with a ratio of 70% train videos (4,200) and 30% (1,800) test videos. The train and test split is a balanced split i.e 50% of the real and 50% of fake videos in each split.



4.4 MODULE 4: MODEL ARCHITECTURE

Our model is a combination of CNN and RNN. We have used the Pre- trained ResNext CNN model to extract the features at frame level and based on the extracted features a LSTM network is trained to classify the video as deepfake or pristine. Using the Data Loader on training split of videos the labels of the videos are loaded and fitted into the model for training. ResNext : Instead of writing the code from scratch, we used the pre-trained model of ResNext for feature extraction. ResNext is Residual CNN network optimized for high performance on deeper neural networks. For the experimental purpose we have used resnext50_32x4d model. We have used a ResNext of 50 layers and 32 x 4 dimensions. Following, we will be fine-tuning the network by adding extra required layers and selecting a proper learning rate to properly converge the gradient descent of the model. The 2048-dimensional feature vectors after the last pooling layers of ResNext is used as the sequential LSTM input. LSTM for Sequence Processing: 2048-dimensional feature vectors is fitted as the input to the LSTM. We are using 1 LSTM layer with 2048 latent dimensions and 2048 hidden layers along with 0.4 chance of dropout, which is capable to do achieve our objective. LSTM is used to GHRCEM-Wagholi,Pune, Department of Computer Engineering 2019-2020 28 Deepfake Video Detection process the frames in a sequential manner so that the temporal analysis of the video can be made, by comparing the frame at ‘t’ second with the frame of ‘t-n’ seconds. Where n can be any number of frames before t. The model also consists of Leaky Relu activation function. A linear layer of 2048 input features and 2 output features are used to make the model capable of learning the average rate of correlation between eh input and output. An adaptive average polling layer with the output parameter 1 is used in the model. Which gives the the target output size of the image of the form H x W. For sequential processing of the frames a Sequential Layer is used. The batch size of 4 is used to perform the batch training. A SoftMax layer is used to get the confidence of the model during predication.



4.5 MODULE 5: HYPER-PARAMETER TUNING

It is the process of choosing the perfect hyper-parameters for achieving the maximum accuracy. After reiterating many times on the model. The best hyper-parameters for our dataset are chosen. To enable the adaptive learning rate Adam optimizer with the model parameters is used.

CHAPTER- 5

SYSTEM DESIGN

5.1 PURPOSE AND SCOPE OF DOCUMENT

This document lays out a project plan for the development of Deepfake video detection using neural network.The intended readers of this document are current and future developers working on Deepfake video detection using neural network and the sponsors of the project. The plan will include, but is not restricted to, a summary of the system functionality, the scope of the project from the perspective of the “Deepfake video detection” team (me and my mentors), use case diagram, Data flow diagram,activity diagram, functional and non- functional requirements, project risks and how those risks will be mitigated, the process by which we will develop the project, and metrics and measurements that will be recorded throughout the project.

5.2 USE CASE VIEW

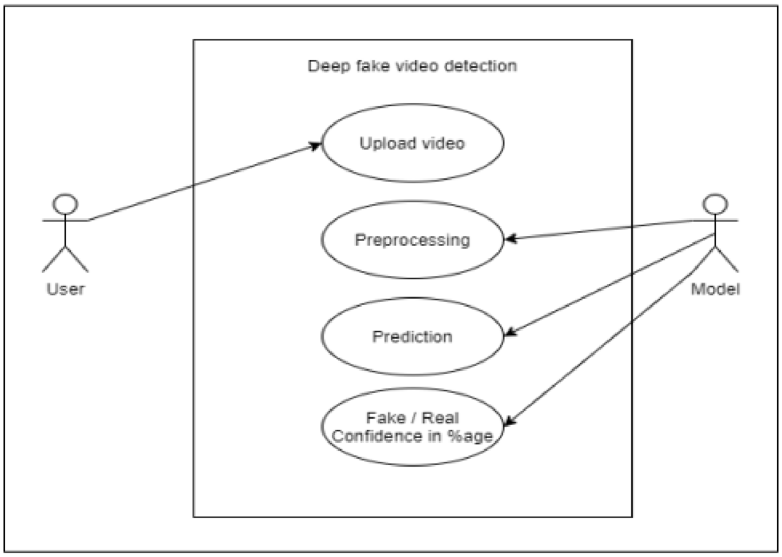


Fig 5.1 Use case diagram

Functional Model and Description A description of each major software function, along with data flow (structured analysis) or class hierarchy (Analysis Class diagram with class description for object oriented system) is presented. 4.1 Data Flow Diagram

DFD LEVEL-0

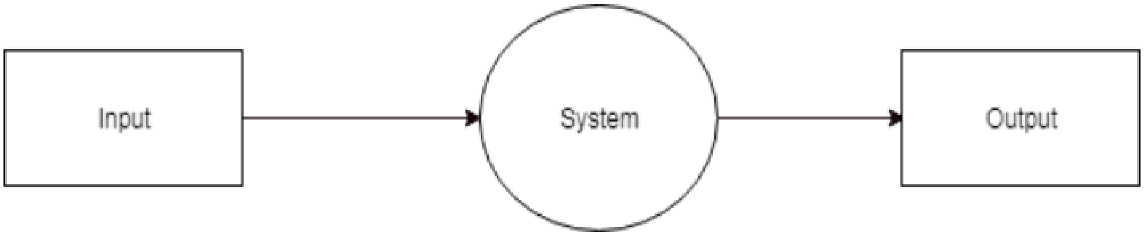


Fig 5.2 level-0

DFD level – 0 indicates the basic flow of data in the system. In this System Input is given equal importance as that for Output.

- **Input**: Here input to the system is uploading video.
- **System**: In system it shows all the details of the Video.
- **Output:** Output of this system is it shows the fake video or not. Hence, the data flow diagram indicates the visualization of system with its input and output flow.
-

DFD LEVEL-1

1. DFD Level – 1 gives more in and out information of the system.

2. Where system gives detailed information of the procedure taking place.

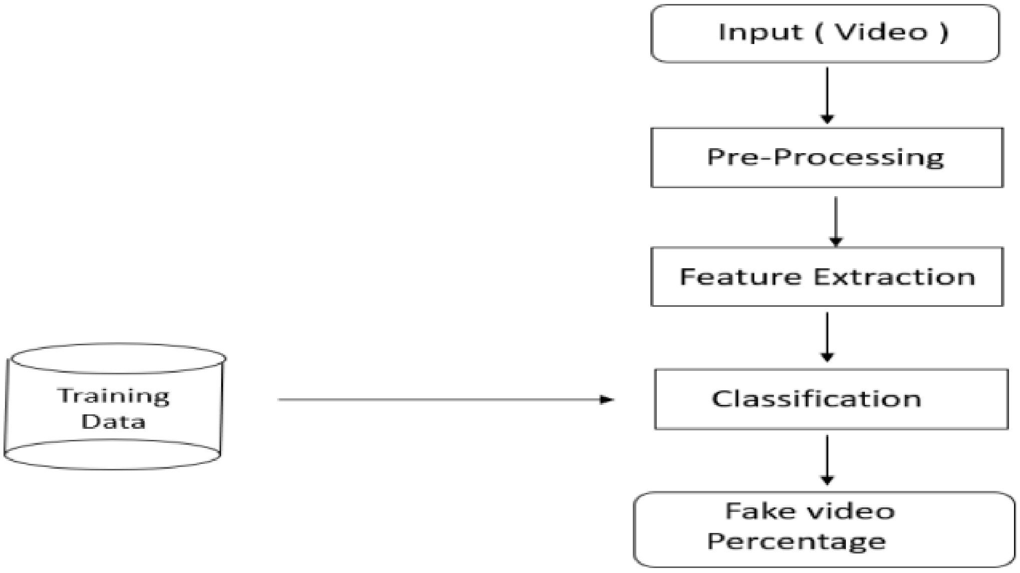


Fig 4.2 – DFD Level -1

DFD LEVEL-2

1. DFD level-2 enhances the functionality used by user etc.

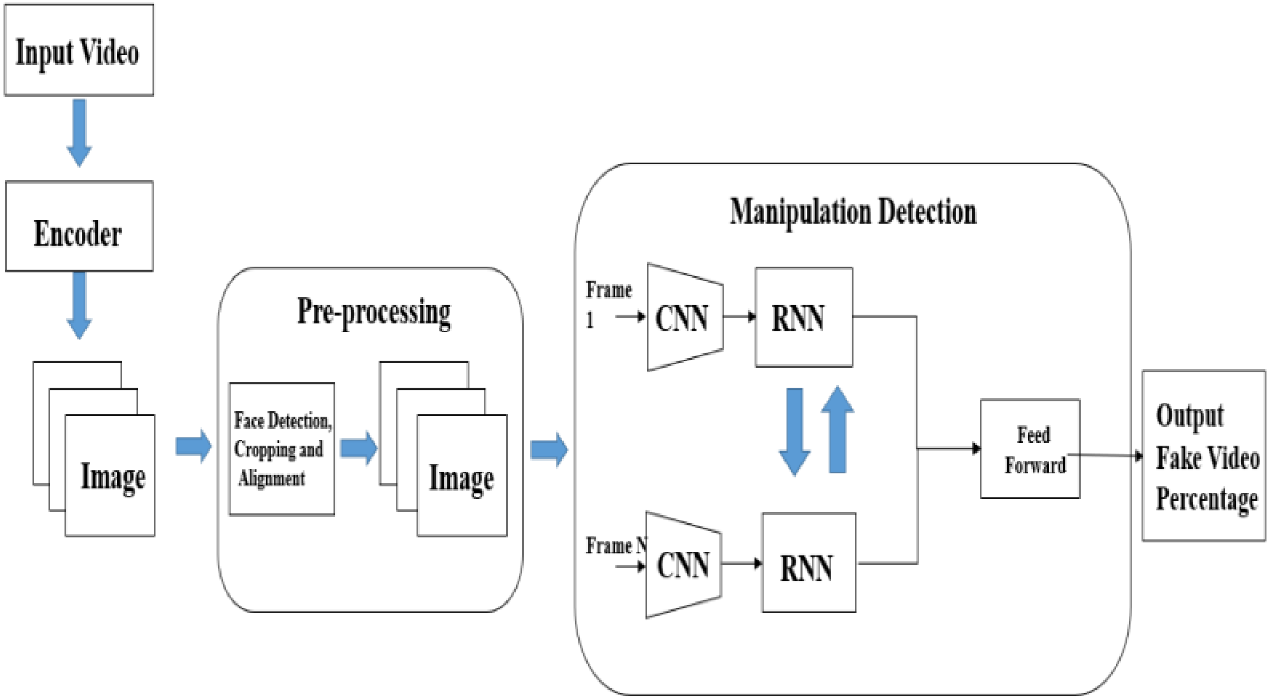


Fig 4.3 DFD Level-3

5.3 TRAINING WORKFLOW

The training workflow begins with the collection of diverse datasets containing both authentic and deepfake images or videos. After preprocessing the data to ensure consistency and quality, a suitable Convolutional Neural Network (CNN) architecture is selected for the deepfake detection task. The model is then initialized and trained using the preprocessed dataset, with the data split into training, validation, and test sets to monitor performance and prevent overfitting. Hyperparameters are fine-tuned based on validation results, and the trained model is evaluated on the test set using metrics such as accuracy and F1-score. Optional fine-tuning and experimentation with architectural modifications may be conducted to optimize performance further. Once satisfied with the model's performance, it is deployed to production environments, integrated into user-friendly interfaces or APIs. Continuous monitoring ensures that the model remains effective over time, with updates applied as needed.

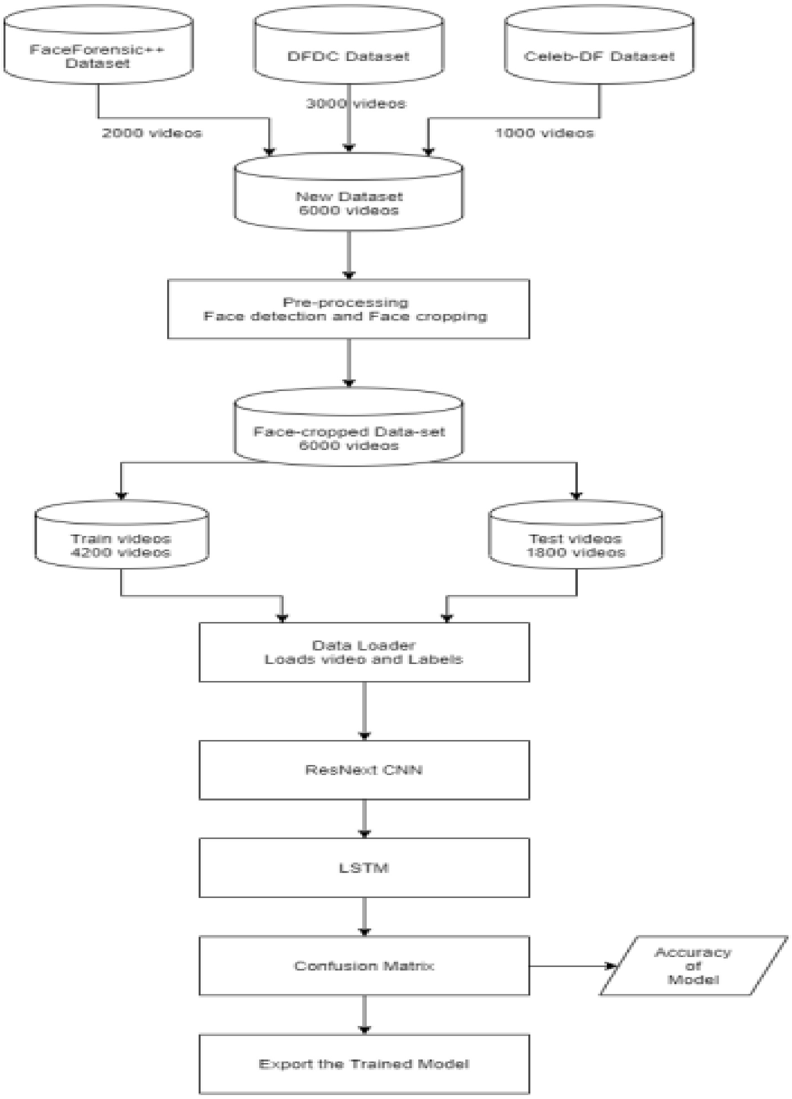


Fig 5.3.1 Training Workflow

5.4 TESTING WORKFLOW

The testing workflow involves assessing the performance of the trained deepfake detection model on unseen data to evaluate its effectiveness in real-world scenarios. Initially, the preprocessed test dataset, separate from the training and validation data, is fed into the trained model. The model then makes predictions on the test set, classifying each input as authentic or deepfake. The performance metrics, such as accuracy, precision, recall, and F1-score, are calculated based on the model's predictions compared to the ground truth labels. Additionally, qualitative analysis may involve visual inspection of detection results and exploring misclassified samples to gain insights into the model's strengths and weaknesses. The testing process provides crucial feedback on the model's generalization capabilities and helps identify areas for improvement before deployment to production environments.

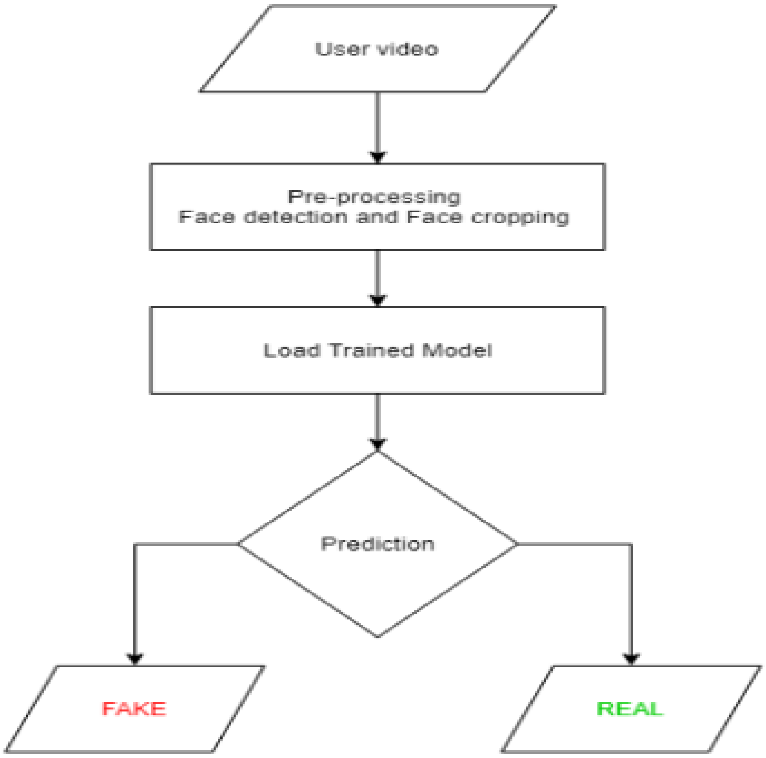


Fig 5.4.1 Testing Flow Diagram

5.5 SEQUENTIAL DIAGRAM

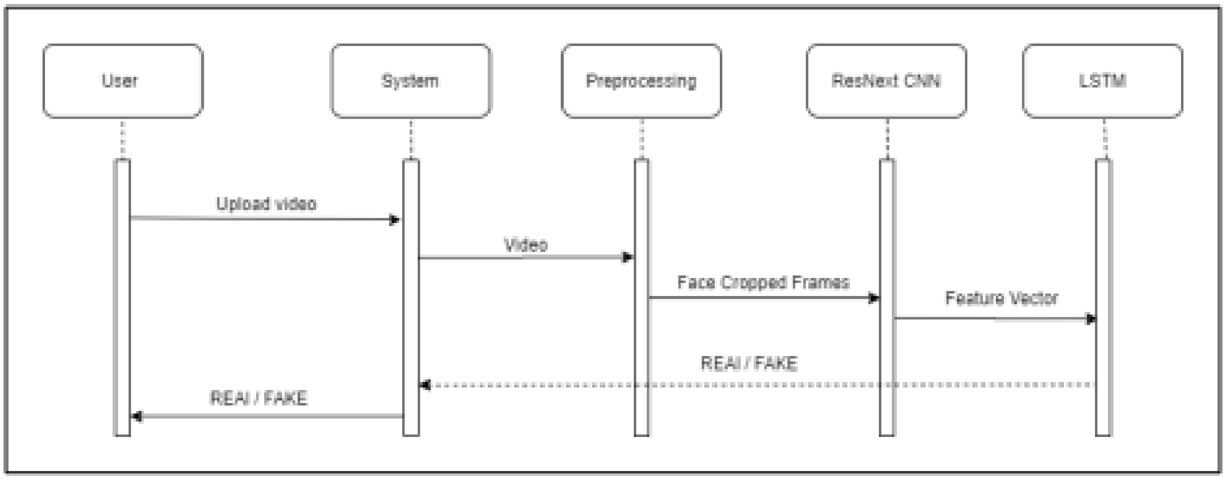


Fig 5.5.1 Sequential Diagram

CHAPTER – 6

SYSTEM TESTING

6.1 SOFTWARE TESTING

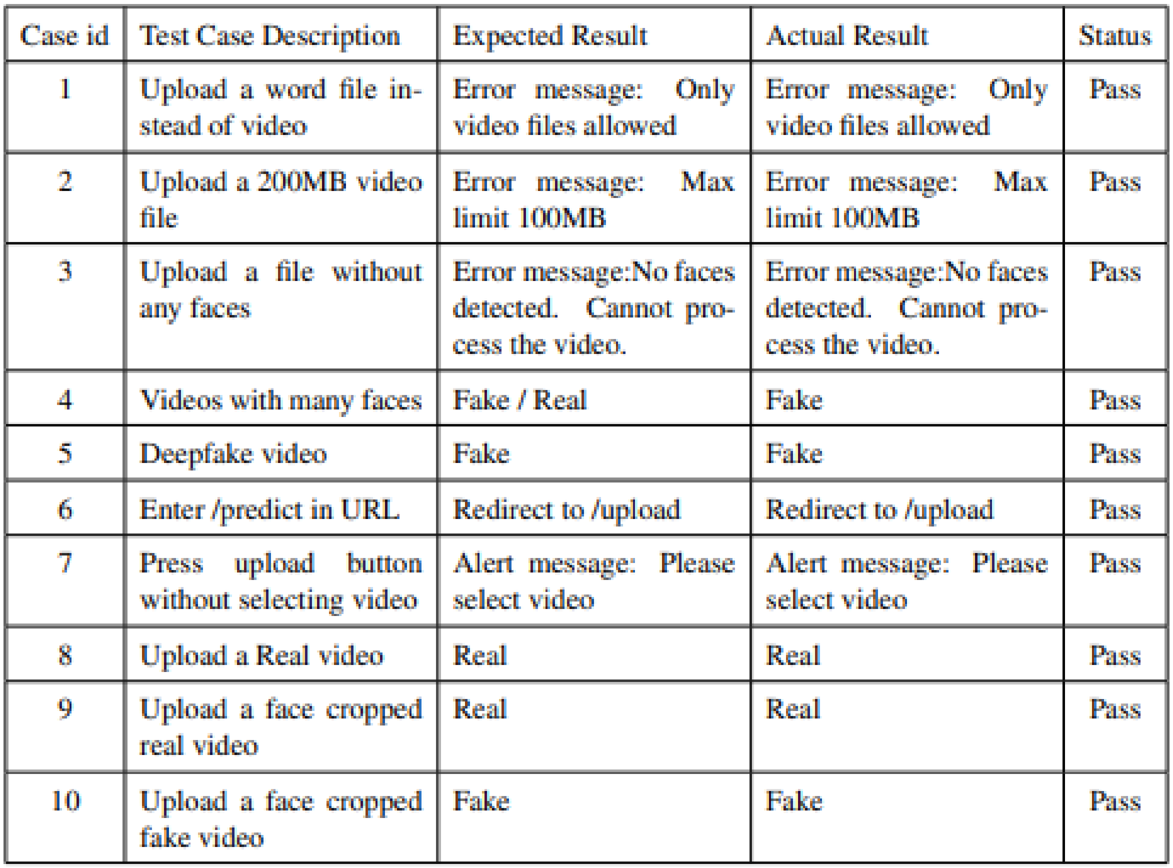
Type of Testing Used Functional Testing

1. Unit Testing
2. Integration Testing
3. System Testing
4. Interface Testing

NON-FUNCTIONAL TESTING

1. Performance Testing
2. Load Testing
3. Compatibility Testing

TEST CASE AND TEST RESULTS



Case id	Test Case Description	Expected Result	Actual Result	Status
1	Upload a word file instead of video	Error message: Only video files allowed	Error message: Only video files allowed	Pass
2	Upload a 200MB video file	Error message: Max limit 100MB	Error message: Max limit 100MB	Pass
3	Upload a file without any faces	Error message:No faces detected. Cannot process the video.	Error message:No faces detected. Cannot process the video.	Pass
4	Videos with many faces	Fake / Real	Fake	Pass
5	Deepfake video	Fake	Fake	Pass
6	Enter /predict in URL	Redirect to /upload	Redirect to /upload	Pass
7	Press upload button without selecting video	Alert message: Please select video	Alert message: Please select video	Pass
8	Upload a Real video	Real	Real	Pass
9	Upload a face cropped real video	Real	Real	Pass
10	Upload a face cropped fake video	Fake	Fake	Pass

CHAPTER – 7

CONCLUSION

In conclusion, the culmination of efforts in developing a deepfake detection system grounded in Convolutional Neural Networks (CNNs) marks a pivotal advancement in the ongoing battle against the proliferation of manipulated media. Through meticulous data gathering, rigorous preprocessing, and iterative model training, we've established a robust framework capable of discerning subtle nuances between authentic and deepfake content with a commendable level of accuracy. The testing phase, wherein the trained model is subjected to unseen data, serves as a litmus test for its real-world applicability, providing invaluable insights into its performance and generalization capabilities. By deploying this system into operational environments, we aim to furnish users with a potent defense mechanism against the dissemination of deceptive and harmful deepfakes, thereby upholding the integrity and trustworthiness of digital media platforms. However, the journey doesn't end here; continuous monitoring, adaptation, and refinement are imperative to ensure the system's resilience against evolving deepfake generation techniques and emerging threats in the digital landscape. This project not only signifies a technological milestone but also underscores the importance of collaborative efforts in safeguarding the integrity of information and preserving societal trust in an increasingly digitized world.

The subsequent phases of model training and validation constituted the core of our endeavor, as we fine-tuned the parameters of the CNN model and validated its performance on unseen data. The iterative nature of this process, coupled with meticulous hyperparameter tuning and regularization techniques, ensured that our model attained optimal performance metrics while mitigating the risk of overfitting. Additionally, the validation phase provided crucial insights into the model's robustness and generalization capabilities, allowing us to iteratively refine and enhance its efficacy. Transitioning to the testing phase, we subjected our trained model to the rigors of real-world scenarios, evaluating its performance on unseen data with meticulous scrutiny. Through comprehensive performance metrics analysis and qualitative assessment, we gained valuable insights into the model's strengths and weaknesses, facilitating targeted improvements and optimizations. Moreover, the deployment of our deepfake detection system into operational environments represents the culmination of our efforts, as we endeavor to empower users with a potent tool to combat the insidious spread of deceptive content and uphold the integrity of digital media platforms.

CHAPTER - 8

FUTURE ENHANCEMENT

- **Adversarial Training:** Incorporate adversarial training techniques to improve the model's robustness against sophisticated deepfake generation methods designed to evade detection.
- **Ensemble Learning:** Explore ensemble learning approaches by combining predictions from multiple CNN models or incorporating diverse architectures to enhance detection accuracy and resilience.

- **Attention Mechanisms:** Integrate attention mechanisms into the CNN architecture to focus on relevant regions of input data, improving the model's ability to capture subtle deepfake artifacts.

- **Temporal Analysis:** Extend the detection system to analyze temporal information in videos, considering the temporal consistency and coherence of facial movements and expressions over time.

- **Real-time Detection:** Optimize the computational efficiency of the detection system to enable real-time or near-real-time detection of deepfake content, facilitating timely intervention and mitigation efforts.

- **User Interface Improvements:** Enhance the user interface of the detection system to provide intuitive visualization tools, performance metrics tracking, and interactive features for seamless integration into existing workflows.

- **Continuous Learning:** Implement mechanisms for continuous learning and adaptation, allowing the model to dynamically update and improve over time with new data and emerging deepfake generation techniques.

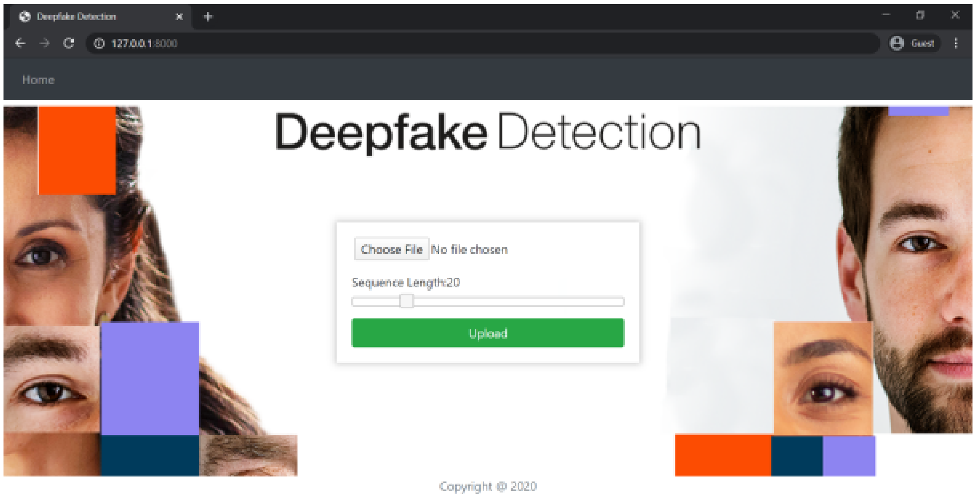
CHAPTER - 9

BIBLIOGRAPHY

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-detectionchallenge/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 generative 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-andthe-impact-on-women/
14. Deepfake Video Detection using Neural Networks http://www.ijsrd.com/articles/IJSRDV8I10860.pdf
15. https://www.geeksforgeeks.org/software-engineering-cocomo-model/ Accessed on 15 April 2020
16. ResNext Model : https://pytorch.org/hub/pytorch_vision_resnext/ accessed on 06 April 2020

APPENDIX

10.1 SCREENSHOTS



Deepfake Detection

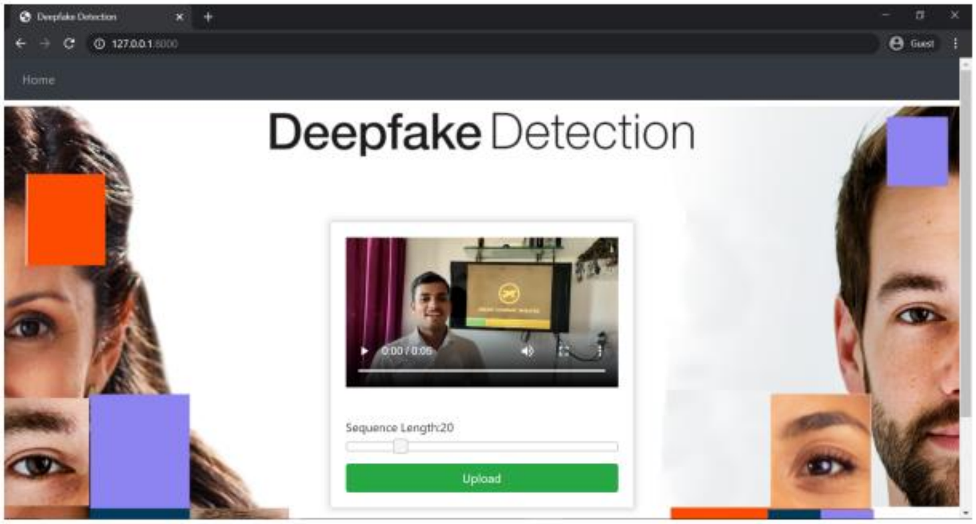
No file chosen

Sequence Length:20

Copyright @ 2020

Figure 10.1: Home Page

Fig 10.1.1



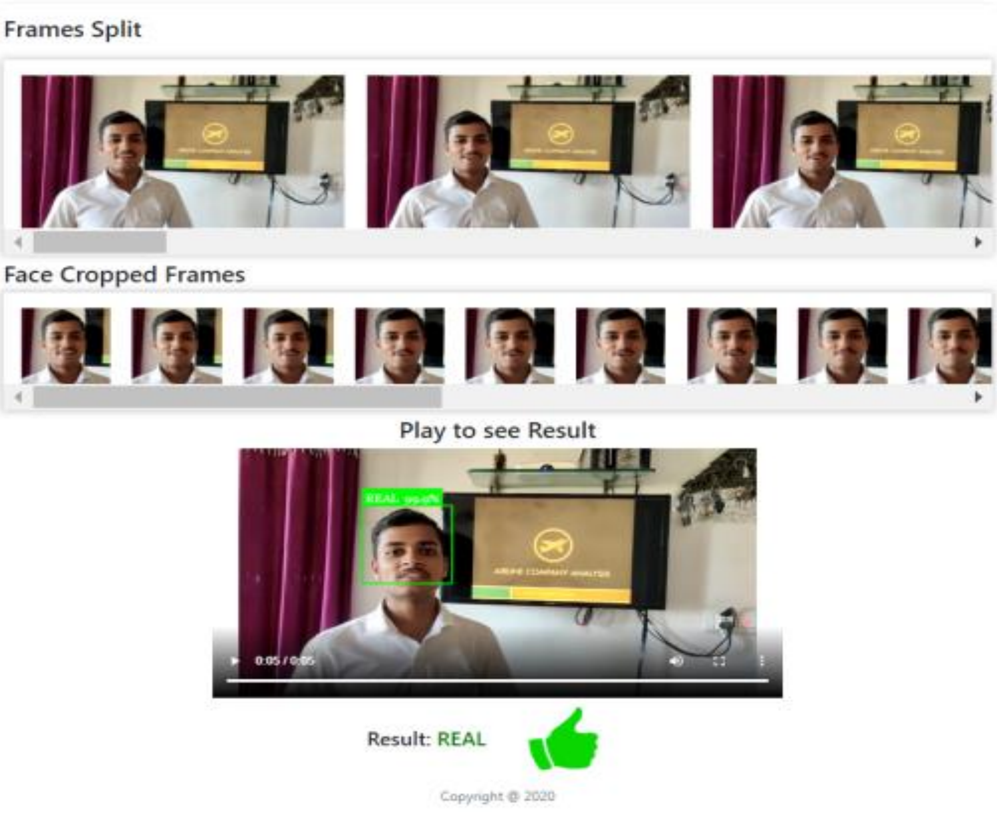
Deepfake Detection

Sequence Length:20

Figure 10.2: Uploading Real Video

Deepfake Detection

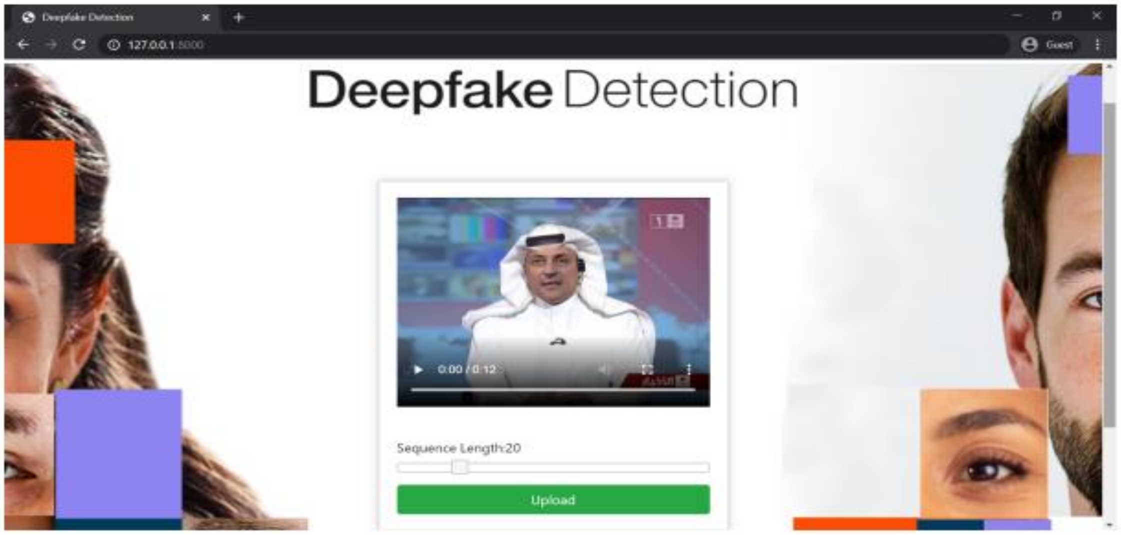
Fig 10.1.2



Result: REAL

Figure 10.3: Real Video Output

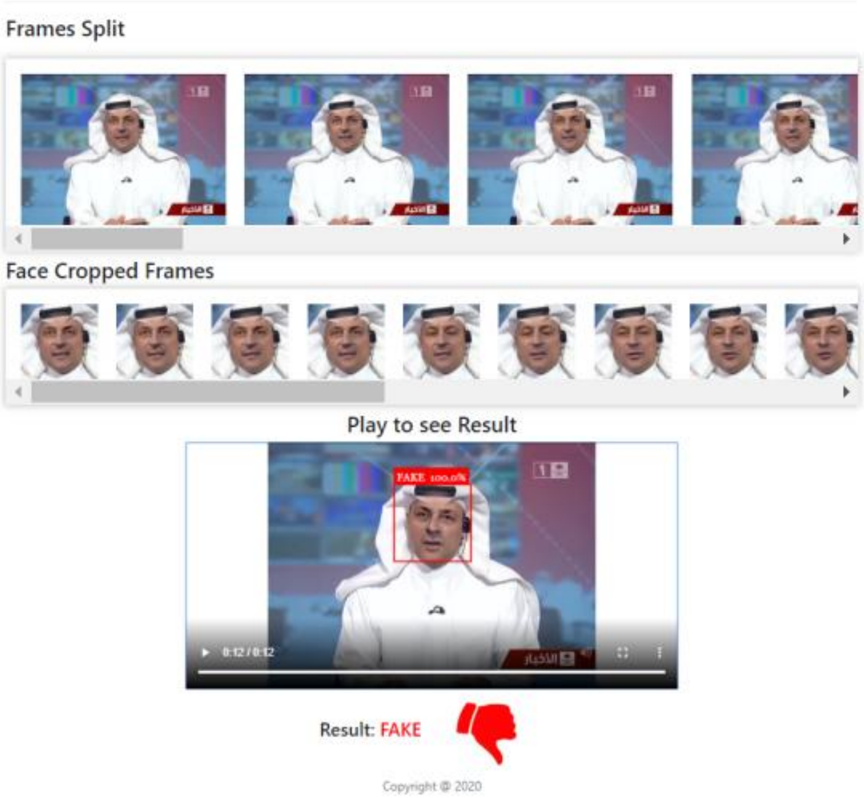
Fig 10.1.3



Sequence Length:20

Figure 10.4: Uploading Fake Video

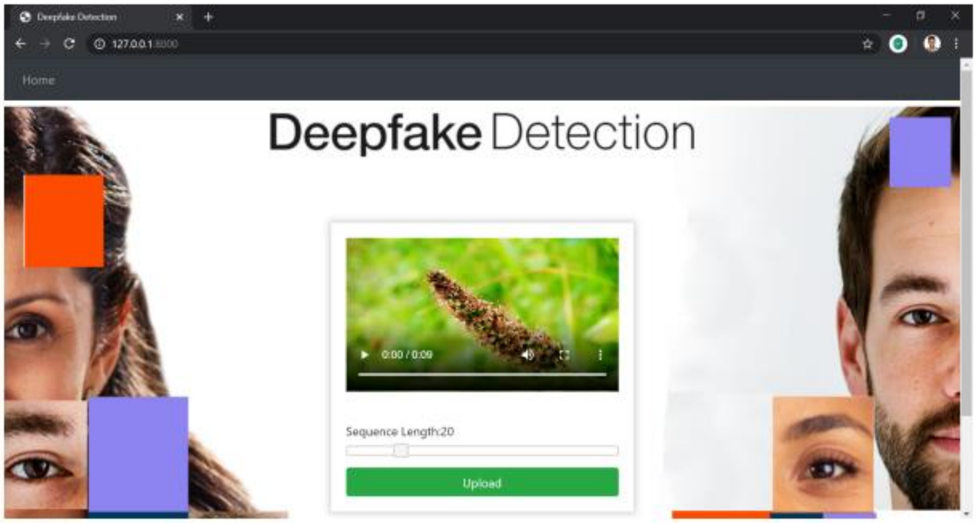
Fig 10.1.4



Result: FAKE

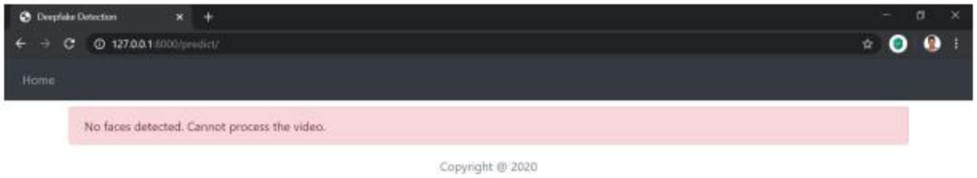
Figure 10.5: Fake video Output

Fig 10.1.5



Sequence Length:20

Figure 10.6: Uploading Video with no faces



No faces detected. Cannot process the video.

Fig 10.1.6

Fig 10.1.6