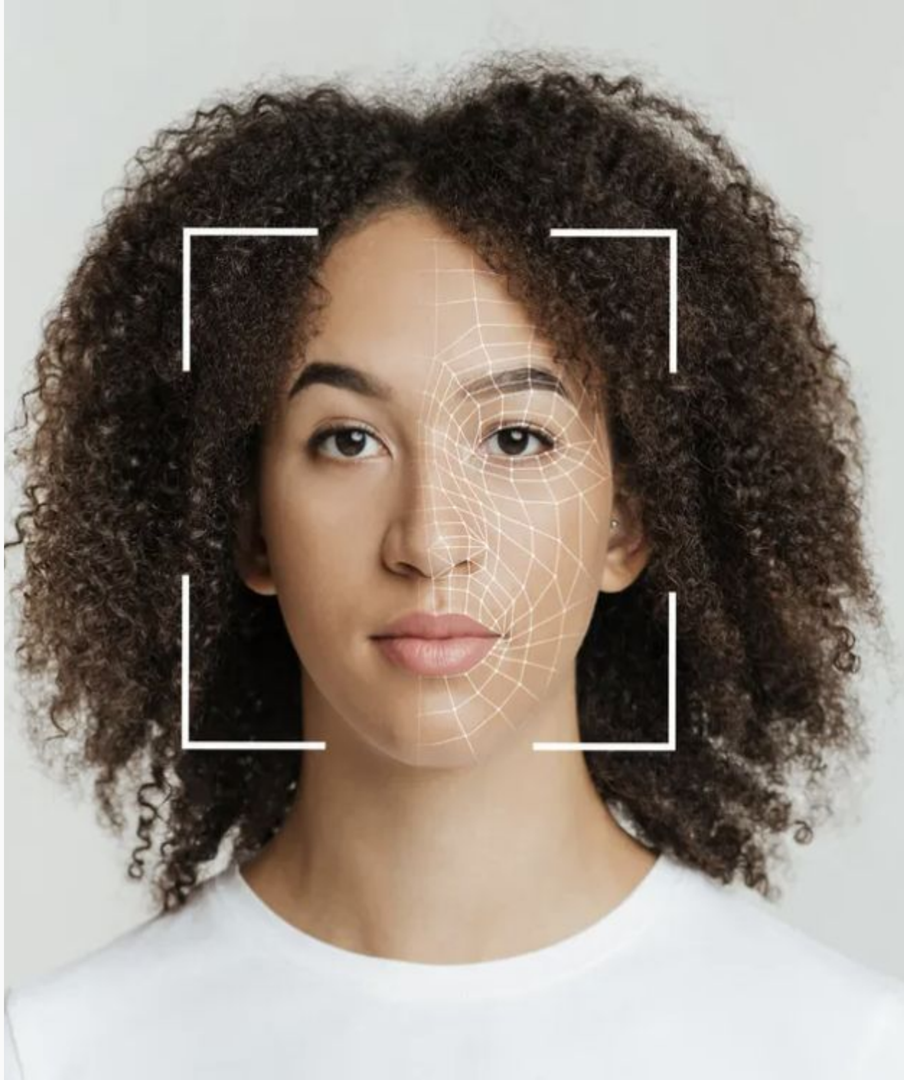
CPG-161

BEYOND THE PIXELS

Uncovering Synthetic Video Fabrications used in Cybercrimes

Team Members:
102003389 Abhishek Aggarwal
102003448 Arushi
102053051 Jatin Narula
102003468 Priyal Kaler
102003374 Yashaswini Khanna

Project Mentors:
Dr Husanbir Singh Pannu
Dr Ashima Singh



PROJECT INTRODUCTION

PROBLEM DEFINITION

Deepfakes and misinformation erode trust in media, harm reputations, spread bias, and invade privacy, posing cybersecurity challenges and societal division.

PROJECT SCOPE

Developing accessible deepfake detection software for platforms, fact-checkers, and users; verifying authenticity, countering manipulation's impact, ensuring privacy.

OBJECTIVES

EXISTING RESEARCH

To study and analyse the already proposed pre-existing state-of-the-art deepfake models

MODEL CREATION

To propose, design and implement an identifier that can recognise synthetically AI-generated videos

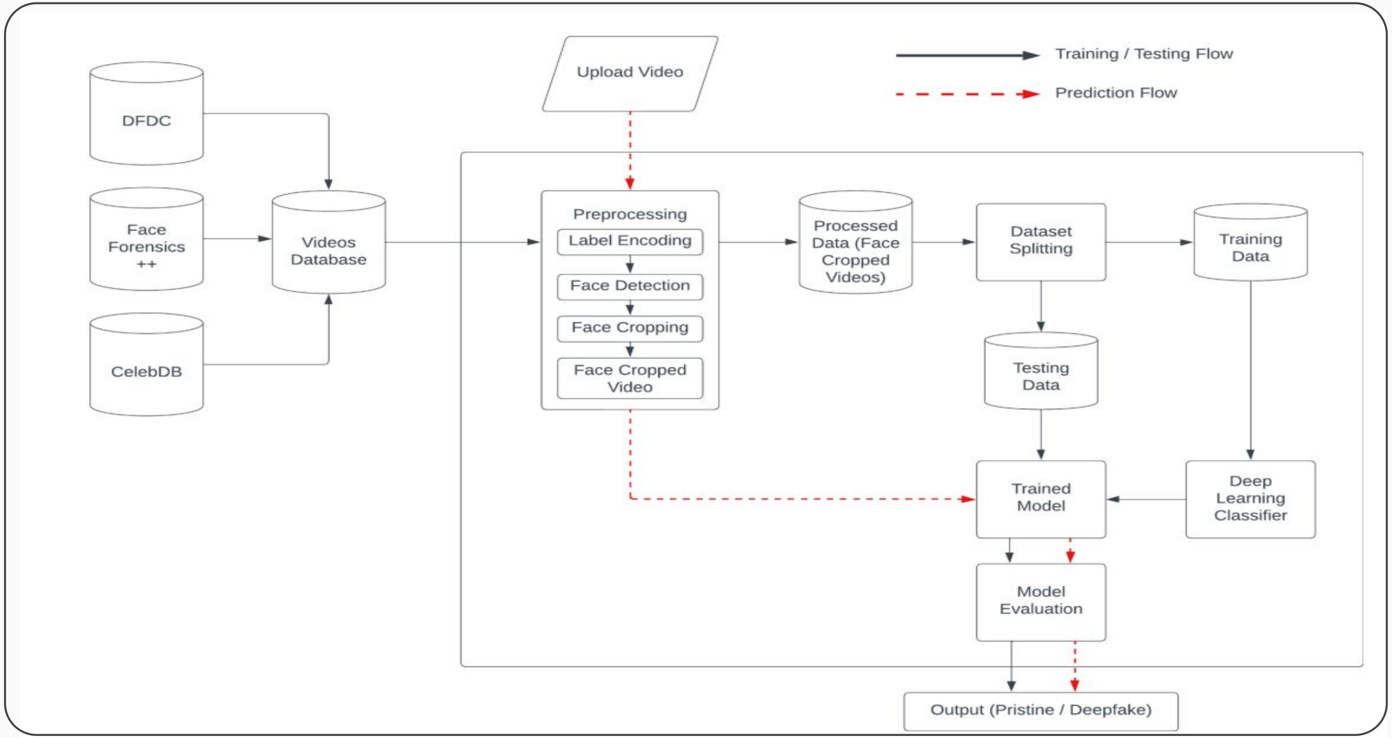
MODEL TESTING

To test and validate the proposed model

PROTOTYPE DEVELOPMENT

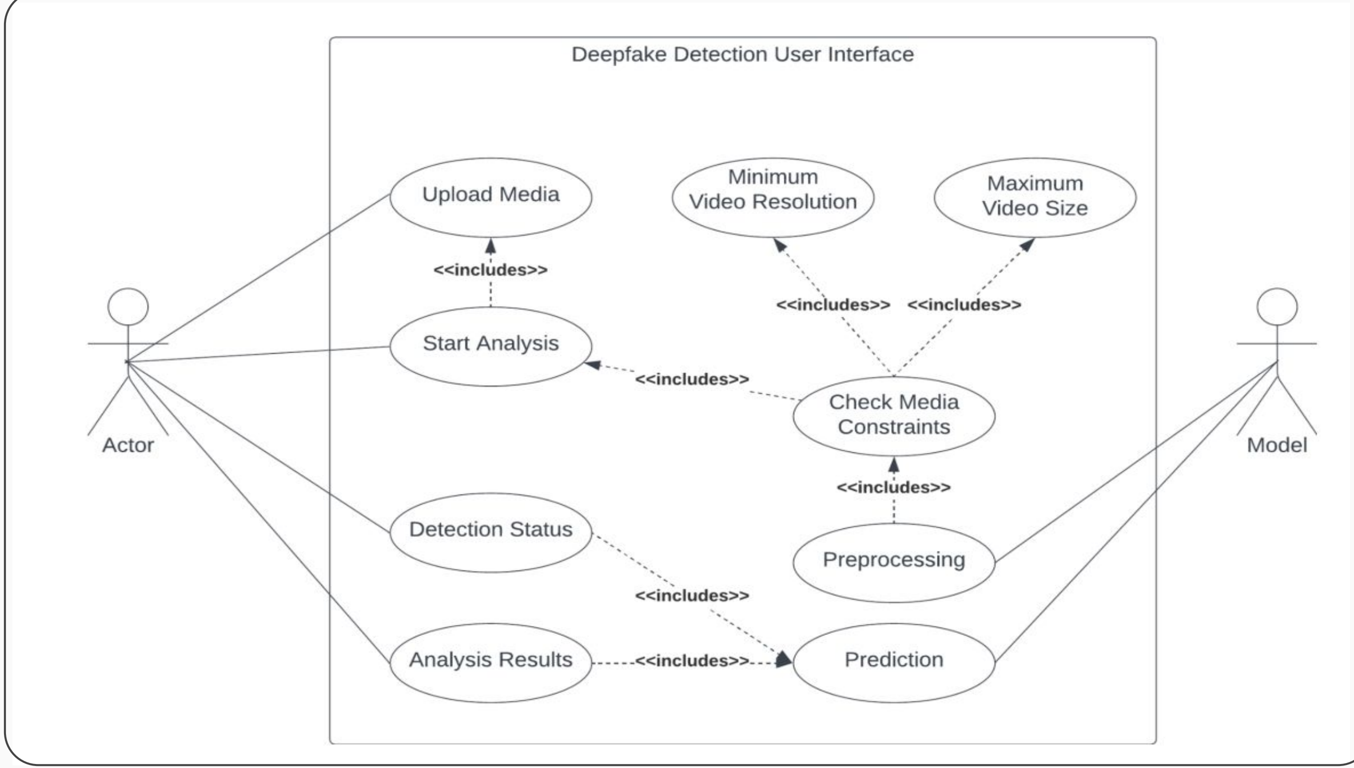
Development of a clean and intuitive user interface for the software

PROJECT ANALYSIS AND DESIGN



BLOCK DIAGRAM OF DEEPFAKE DETECTION MODEL

USER INTERFACE DESIGN



TOOLS AND TECHNOLOGY

PYTHON

PYTORCH

OPENCV

NVIDIA GPU

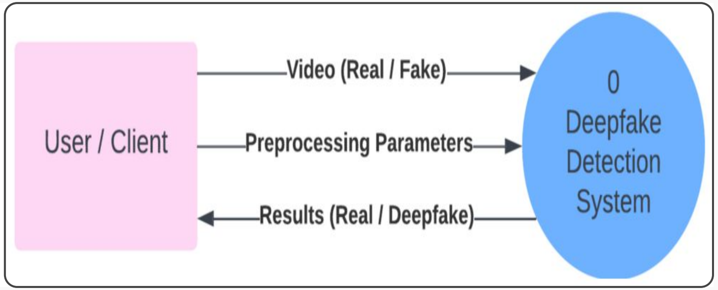
DJANGO

AWS

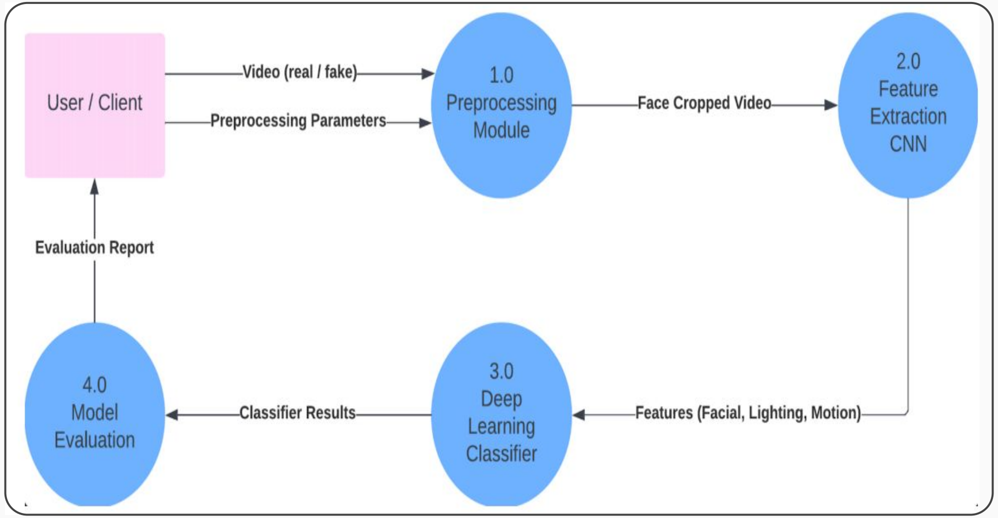
GIT

PRE-TRAINED CNN MODELS

DATA DESIGN

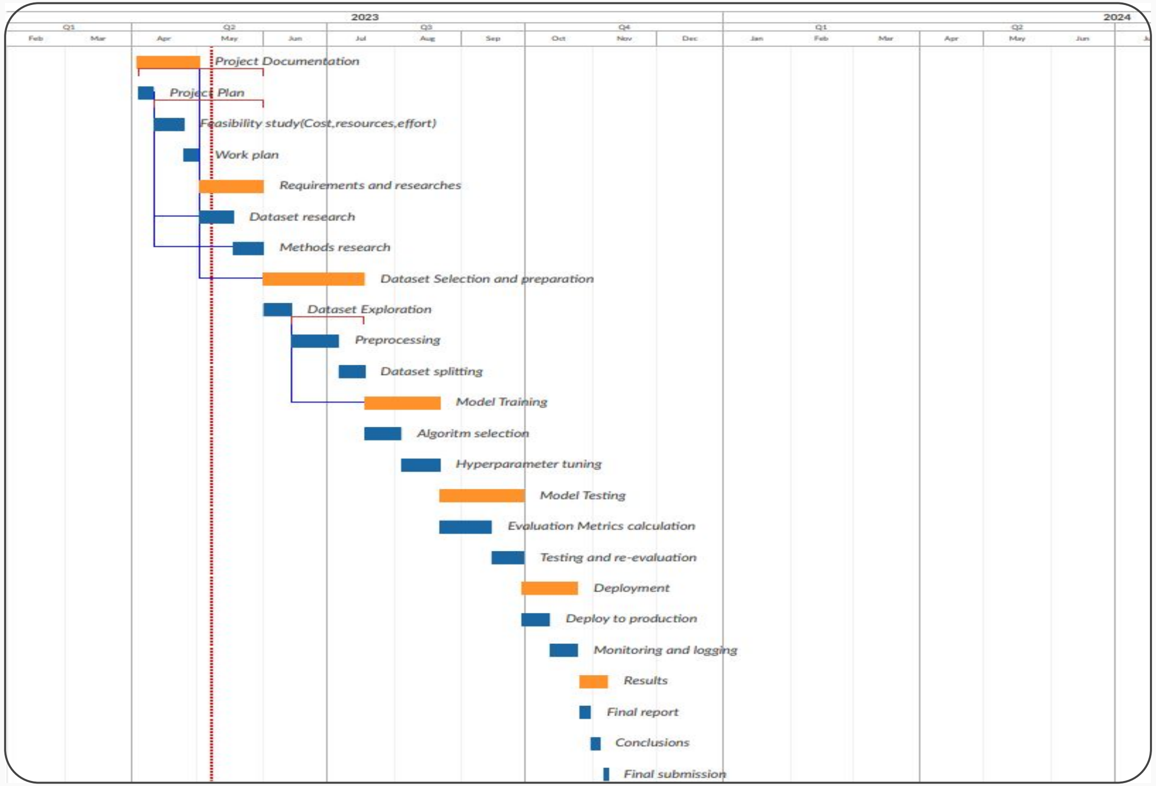


LEVEL-0



LEVEL-1

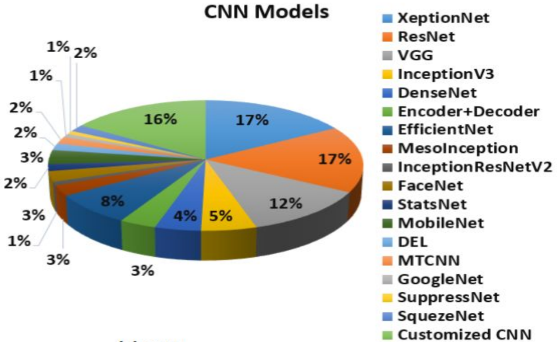
WORK PLAN


2023
2024
Q1
Q2
Q3
Q4
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Project Documentation
Project Plan
Feasibility study(Cost,resources,effort)
Work plan
Requirements and researches
Dataset research
Methods research
Dataset Selection and preparation
Dataset Exploration
Preprocessing
Dataset splitting
Model Training
Algoritm selection
Hyperparameter tuning
Model Testing
Evaluation Metrics calculation
Testing and re-evaluation
Deployment
Deploy to production
Monitoring and logging
Results
Final report
Conclusions
Final submission

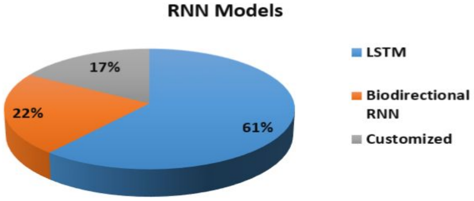
COST ANALYSIS

S. No.	Tool	Use	Approx Costs
1.	PyTorch	Deep Learning Framework	Open Souce (Free to use)
2.	AWS EC2 and S3	Cloud Computing Services	Approximate costs for storage and deployment: 7,012.10 INR
3.	NVIDIA GPUs	GPU Acceleration	500 Compute Units from Google Colab: 4,852.50 INR
4.	PyCharm	Development IDE	Free on Educational Licence
5.	GitHub	Version Control	Open Souce (Free to use)
		Total Costs	11,864.60 INR

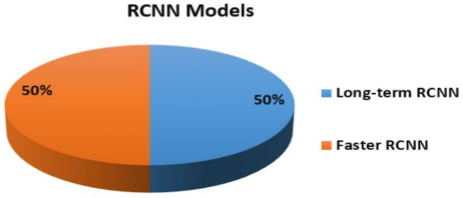
LITERATURE SURVEY AND RELATED WORK



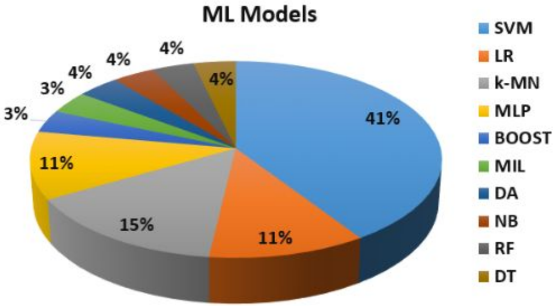
(a) CNN



(b) RNN



(c) RCNN



Category	Model	#Studies
Deep Learning	CNN	71
	RNN	12
	RCNN	2
Machine Learning	SVM	11
	k-MN	4
	LR	3
	MLP	3
	BOOST	2
	RF	1
	DT	1
	DA	1
	NB	1
	MIL	1

Rana, Md Shohel, et al. "Deepfake detection: A systematic literature review." *IEEE Access* (2022).

DATASET RESEARCH

Out of very popularly available deepfake datasets, we decided to use a combination of videos from 3 prominent dataset:
FaceForensics++, Celeb-DF and DFDC.

	Real		Fake				
Dataset	**Video**	**Frame**	**Video**	**Frame**	**Generation Method**	**Release Date**	**Generation Group**
UADFV	49	17.3K	49	17.3K	FakeAPP	11/2018	1st
DF-TIMIT	320	34K	320	34K	Faceswap-GAN	12/2018	1st
*Real & Fake Face Detection	1081	405.2K	960	399.8K	Expert-generated high-quality photoshopped	01/2019	2st
FaceForensics++	1000	509.9k	1000	509.9K	DeepFakes, Face2Face, FaceSwap, NeuralTextures	01/2019	2nd
DeepFakeDetection	363	315.4K	3068	2242.7K	Similar to FaceForensics++	09/2019	2nd
DFDC	1131	488.4K	4113	1783.3K	Deepfake, GAN-based, and non-learned methods	10/2019	2nd
Celeb-DF	590	225.4K	5639	2116.8K	Improved DeepFake synthesis algorithm	11/2019	2nd
*140K Real & Fake Faces	70K	15.8M	70K	15.8M	StyleGAN	12/2019	2nd

https://www.researchgate.net/figure/Comparison-of-publicly-available-deepfake-datasets_tbl2_359067919

FACE FORENSICS ++

FaceForensics++ consisting of **1000 original video sequences** that have been manipulated with four automated face manipulation methods: **Deepfakes, Face2Face, FaceSwap and NeuralTextures.**

Rossler, Andreas, et al. "Faceforensics++: Learning to detect manipulated facial images." *Proceedings of the IEEE/CVF international conference on computer vision*. 2019.

CELEB-DF

Celeb-DF includes **590 original videos** collected from YouTube with subjects of different ages, ethnic groups and genders, and **5639** corresponding DeepFake videos.

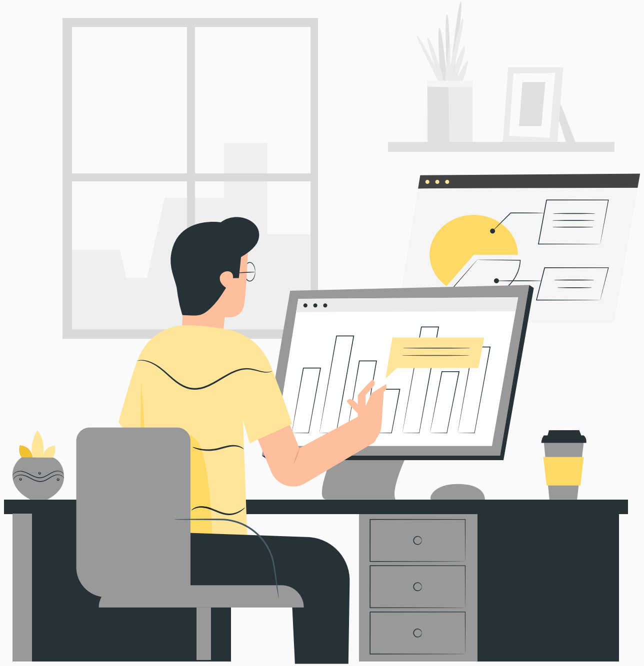
Li, Yuezun, et al. "Celeb-df: A large-scale challenging dataset for deepfake forensics." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.*

DFDC

DFDC is by far the largest currently and publicly available face swap video dataset, with over **100,000 total clips** sourced from 3,426 paid actors, produced with **several Deepfake, GAN-based, and non-learned methods.**

Dolhansky, Brian, et al. "The deepfake detection challenge (dfdc) dataset." *arXiv preprint arXiv:2006.07397* (2020).

PREPROCESSING RESEARCH



01

Extracting Frames from Video

02

Cropping Faces from Frames

03

Creating new video from cropped faces

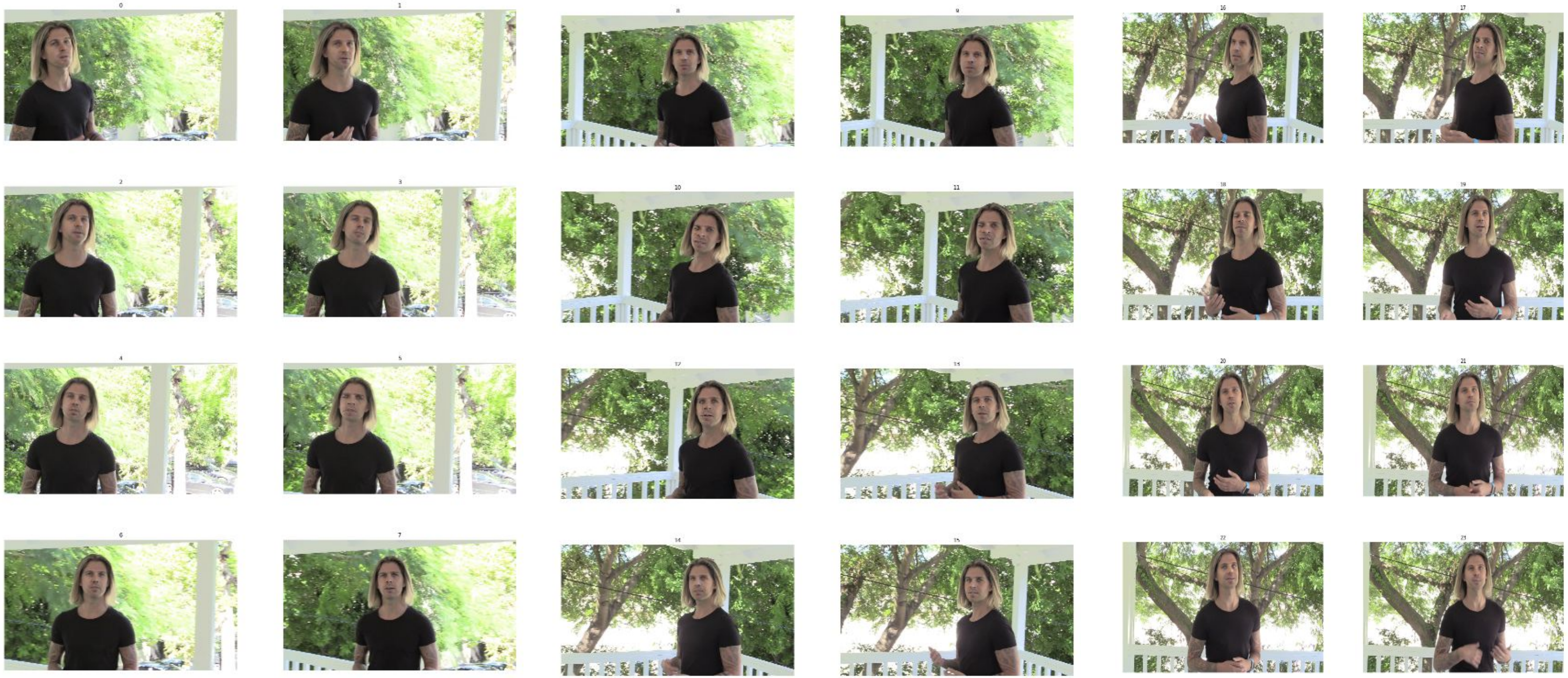
04

Extracting useful features from videos

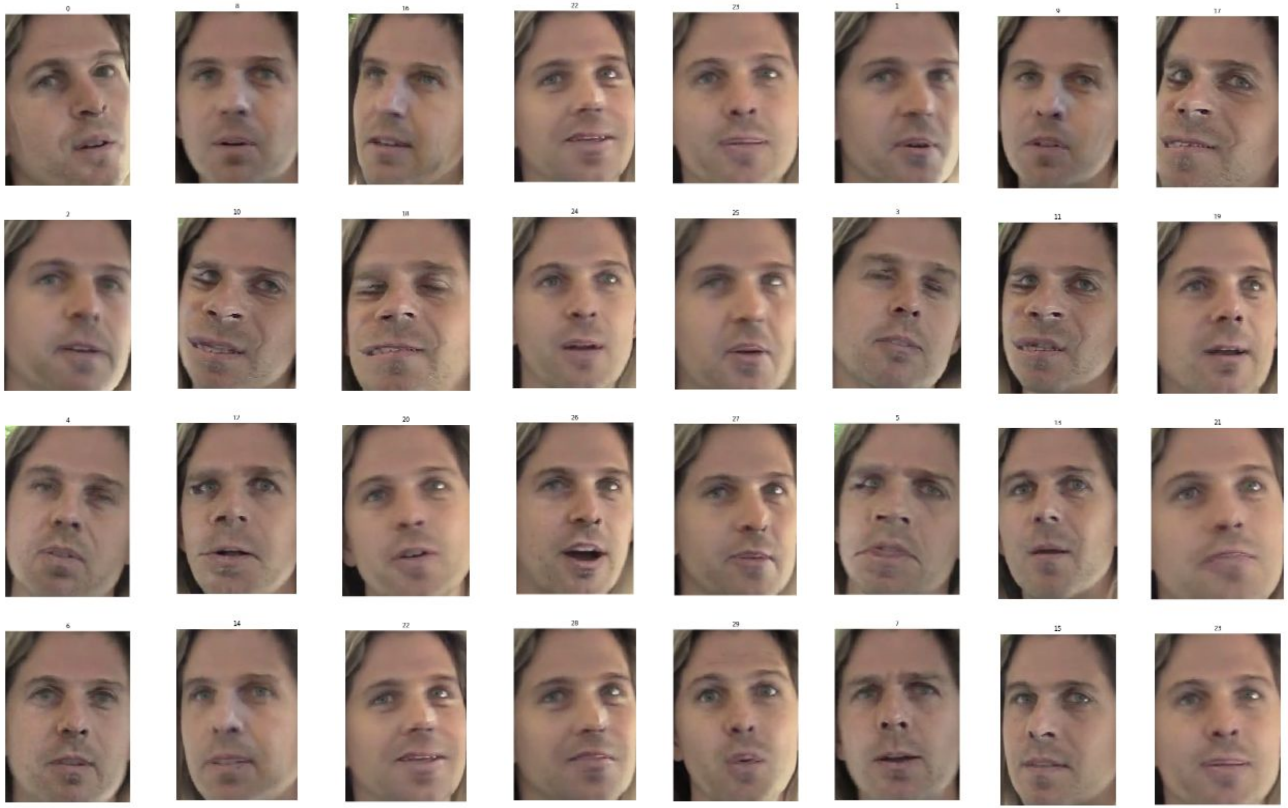
DEEP DIVE INTO PREPROCESSING

Video File: aagfhgtpmw.mp4

EXTRACTING FRAMES FROM VIDEO



CROPPING FACES FROM FRAMES



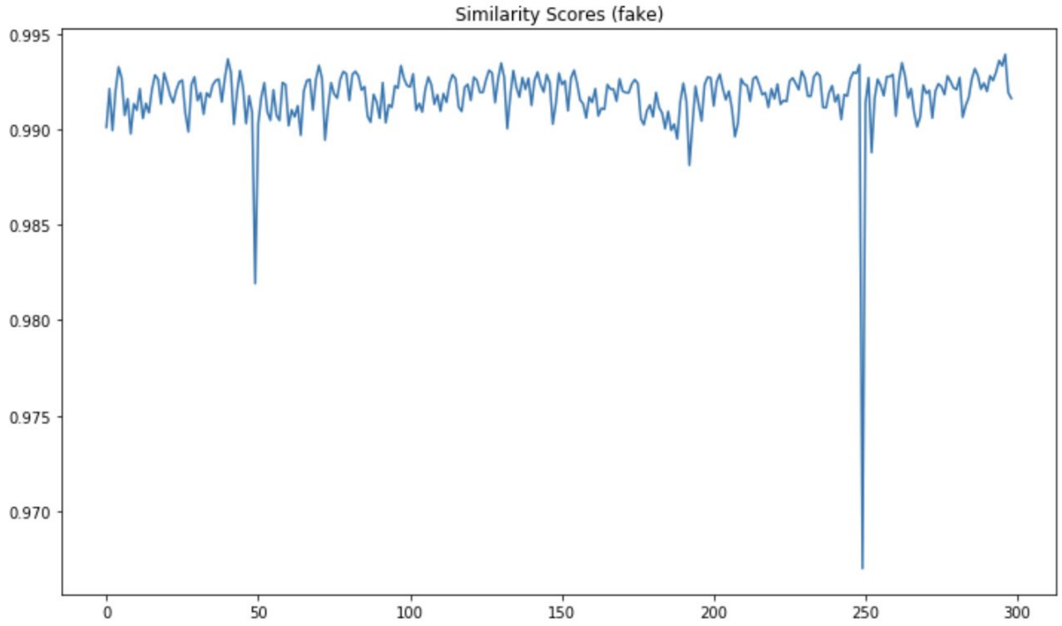
CREATING NEW VIDEO FROM CROPPED FACES

SIMILARITY BETWEEN FRAMES

Plotting the structural **similarity scores** of 2 consecutive frames for all the frames of the video. Here, We can observe some similarity drops between the frames.

300 frames

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$



SOME FEATURES OF DEEPFAKE VIDEOS

FACE DISCOLORATIONS: Slight discolorations of the face relative to the original light

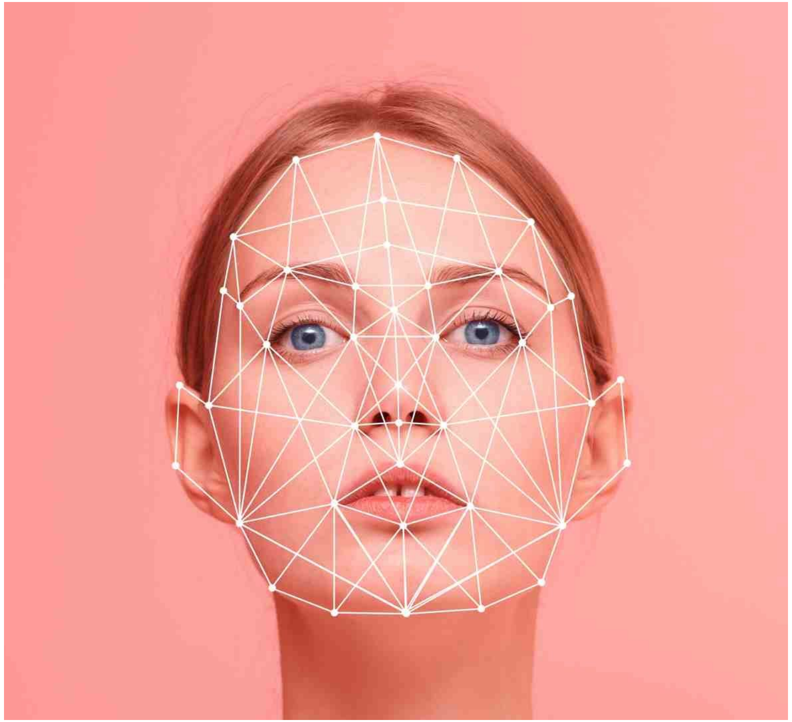
REDUCED BLINKING WITH THE EYES: Less natural eye blinking patterns than real videos

LIGHTING THAT ISN'T QUITE RIGHT: Inconsistent lighting effects on the face, such as shadows or reflections.

UNNATURAL POSITIONING OF FACIAL FEATURES: Someone's face and nose are pointed in different directions.

BLURRINESS: Blurry edges where the face is merged with the neck and hair

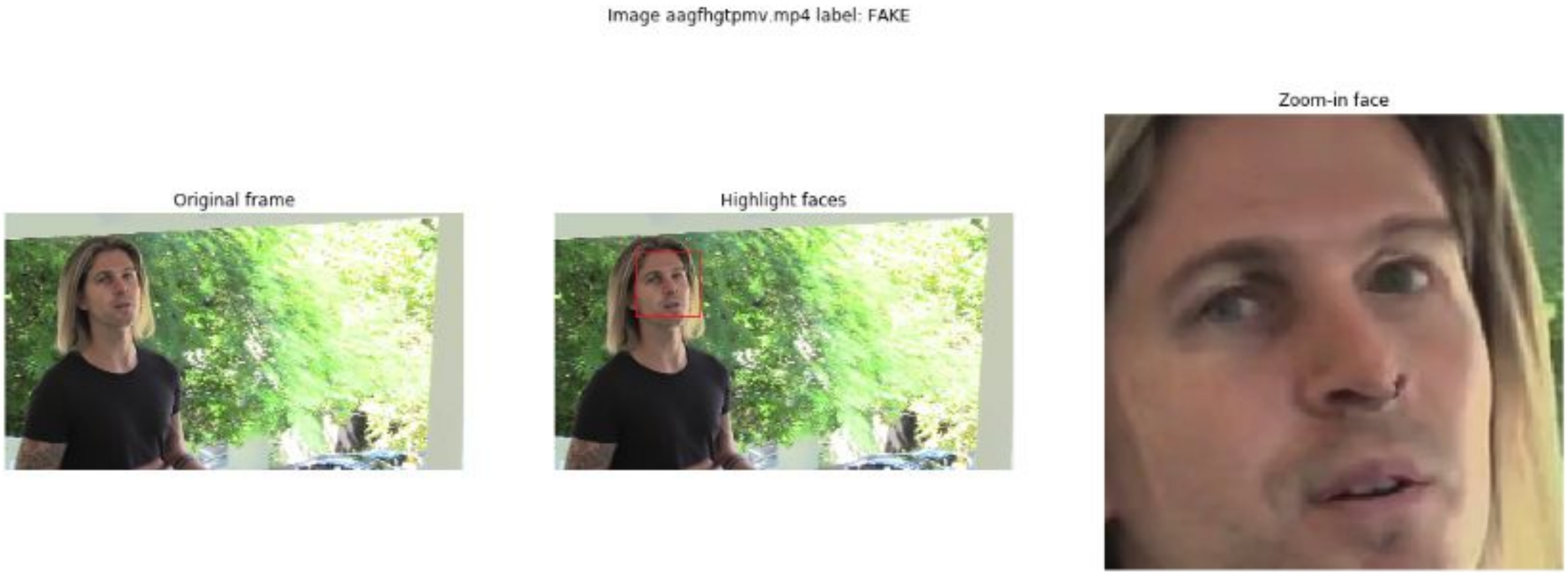
TEXTURE FEATURES: Different texture features than real video



https://www.deepfakenow.com/best-deepfake-detection-methods/

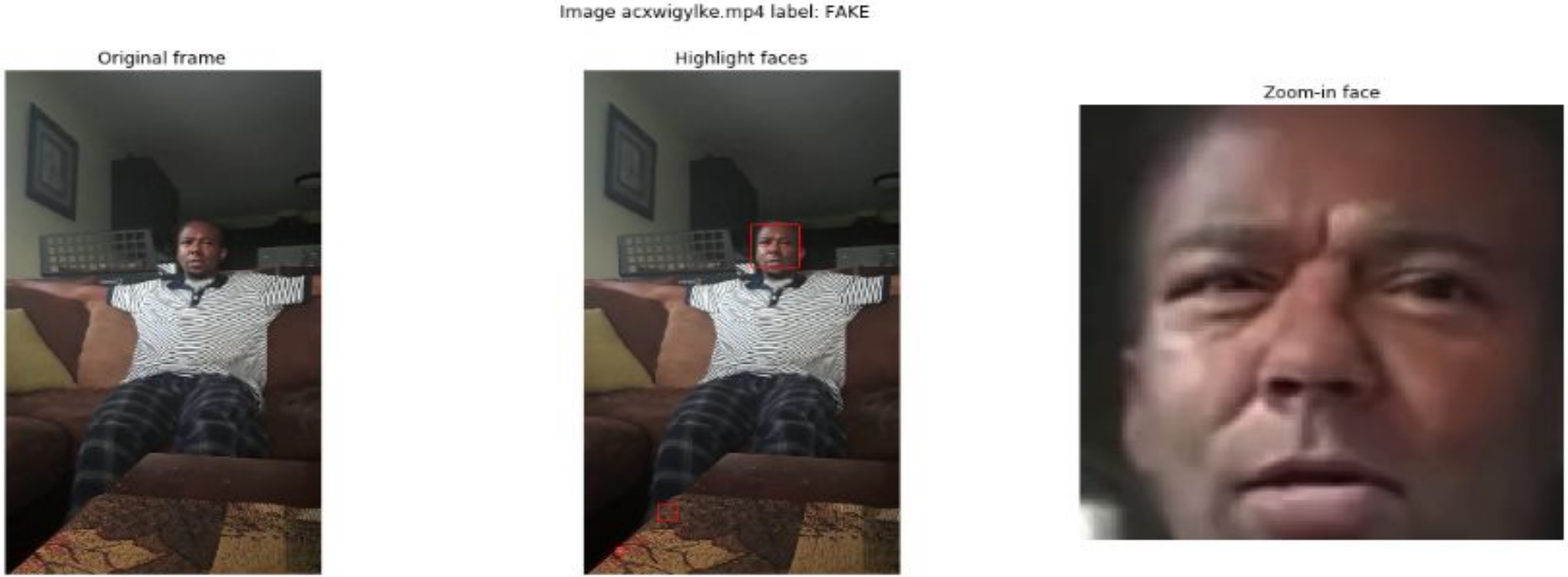
UNNATURAL POSITIONING OF FACIAL FEATURES

In this sample, We can observe that the nose of the person is strange.



BLURRINESS

In this sample, We can observe that the face of this person is so blurry.



INCONSISTENT TEXTURE & LIGHTING

In this sample, We can observe that the glasses of this don't look very realistic. There is also a strangely rounded shape around the right eye of the lady and a strange white spot to the right of the mouth.

Image abqwwspghj.mp4 label: FAKE

PROTOTYPE DEVELOPMENT

MODEL TRAINING DETAILS:

Test Train Split: 70% train videos and 30% testing videos

No of Epochs: 20

Batch Size: 4

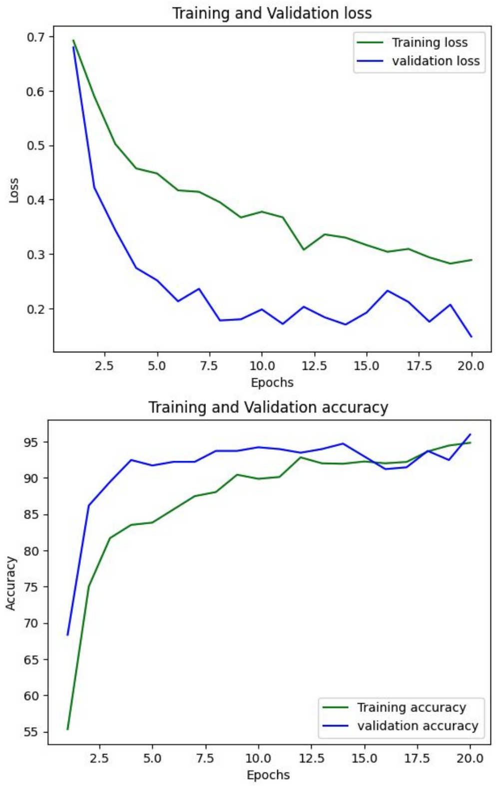
Learning Rate: 1e-5, i.e. 0.00001

Feature Extraction: ResNet-50

Sequential Learning: LTSM

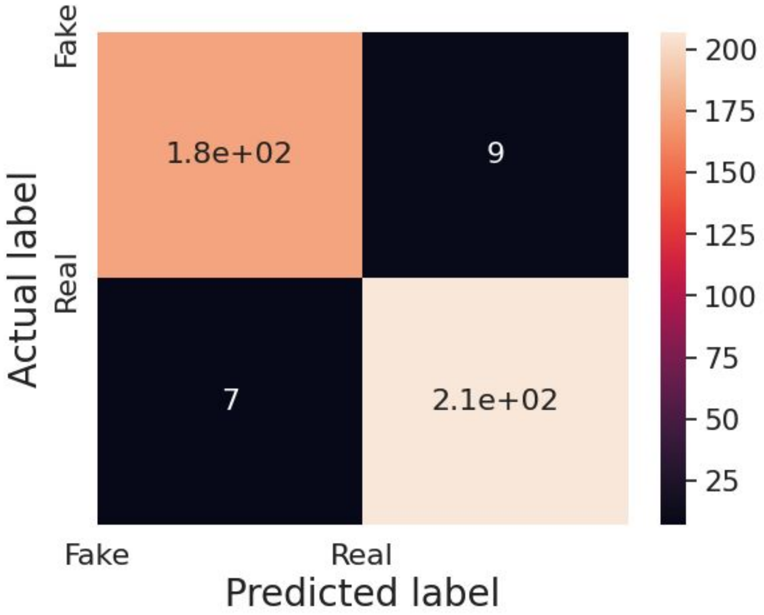
Optimiser: Adam Optimiser, to enable adaptive learning rate

Softmax layer with two output nodes, i.e., REAL or FAKE, also provides the confidence(probability) of prediction.



EVALUATION METRICS

```
True positive =  175
False positive =  9
False negative =  7
True negative =  207
```




```
Calculated Accuracy 95.97989949748744
```

OUTCOMES

01

Preventing the Spread of Fake Media

02

Criminal Prosecution

03

Increased Public confidence

04

Technological Developments

CONTRIBUTIONS OF INDIVIDUAL TEAM MEMBERS

ABHISHEK AND PRIYAL

Researching deep learning methods to develop algorithms for detection of synthetically generated fake videos

YASHASWINI AND JATIN

Skilled in developing data science will be responsible for experimenting on datasets of fake AI-generated videos, data analytics, pre-processing, visualization, etc.

ARUSHI

Experienced in full stack software development, responsible for developing the user interface for the system for detecting deepfakes.

The entire team will work on testing the platform, conducting surveys, and reaching out to experts in the field and researchers who are prepared to supply us with useful information.

CURRENT WORK PROGRESS

ANALYZING DATASETS & METHODS FOR DETECTION

The team analyzed multiple deepfake datasets & used diverse techniques. Reviewed key deepfake detection challenges in models.

REVIEW OF PRE-EXISTING DETECTION TECHNIQUES

The team analysed deep learning models, surveyed deepfake detection progress, pinpointing critical challenges.

DESIGNING SOFTWARE TO DETECT DEEPFAKES

The team preprocesses videos,employs deep learning to classify content as deepfake or genuine.

FUTURE WORK PLAN

MODEL TUNING

Performing appropriate hyperparameter tuning and data preprocessing on different parameters like learning rate, no of epochs, optimizers and even trying alternate CNN Architectures

SYSTEM TESTING

Conducting extensive model validation and evaluation using appropriate evaluation metrics to ensure the models' effectiveness

USER INTERFACE

Simultaneously building the FrontEnd User Interface for the system as a web application so that it can be accessed easily by many internet users

Thank you