Narcodes

Team Id : 1443

CMR Technical Campus

Team Lead : Sahasransu Kar

Team Members : Kundam Venkata Nikhil

Manpreet Singh

S. Vishnu Teja

H. K. Varun Rao

Mahigna Vemula

Team Mentor : Dr. T. S. Mastan Rao

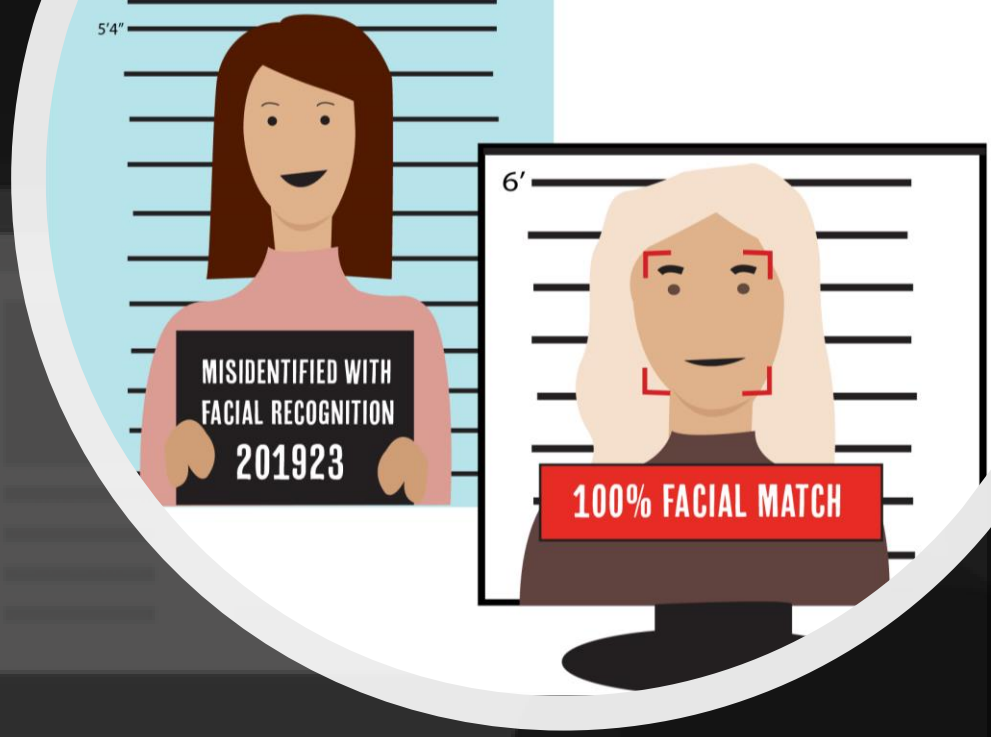
Company/Ministry : Govt. of Madhya Pradesh

PS No. MS331
Face, expression and gesture recognition and compilation in database software

Description of criminal activities

Objectives

- To recognize criminal activities and mis-happenings through the captured live CCTV footage.
- To provide immediate information to the responsible authorities about such activities.
- Recognize the criminal associated with those activities through face recognition.
- Determination of the emotion, age and gender of the person associated with those activities.

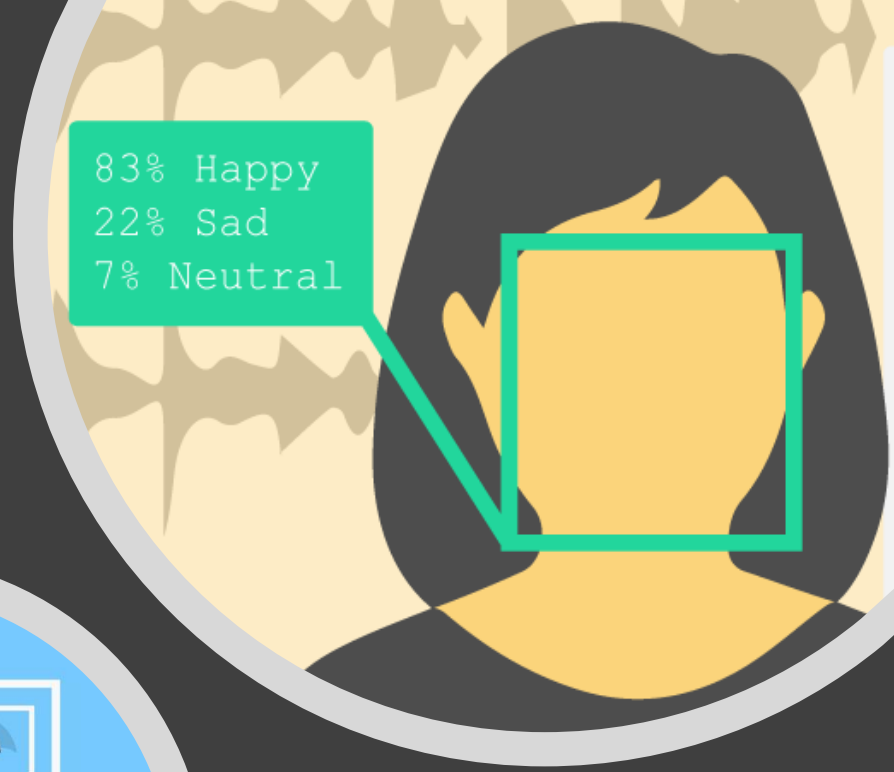


Merits of the Proposed System

- To stop the criminal activities immediately and reduce the delay in alerting the authorities.
- To reduce the loss of property, money and to minimize fatality and crime rate.
- To get hold of the culprit within a short period of time.
- To provide the culprit’s details such as gender and age to the concerned authorities.
- To reduce the tedious job of continuous monitoring of the live CCTV footage including remote places where there is no frequent monitoring.

Approach

- Using activity, face and emotion recognition to predict the activity and suspect details as soon as possible and alert authorities about the same.



Activity Recognition

- We have trained various CCTV captured video footage obtained from online sources and marked various events under its respective attributes.
- Some of the attributes on which we have trained our module are as follows:

1. Abuse : This activity is identified from the CCTV footage if any bad, cruel or violent behaviour against children, old people and animals is recognized.

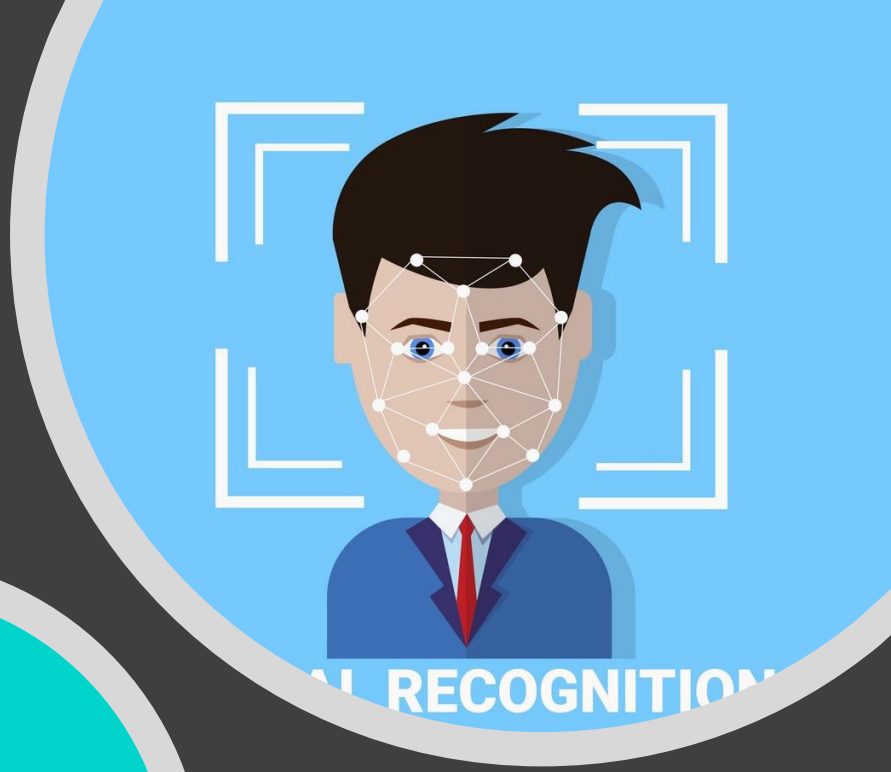
2. Burglary : This event is recognized when people(thieves) are entering into a building or a house with the intention to commit theft.

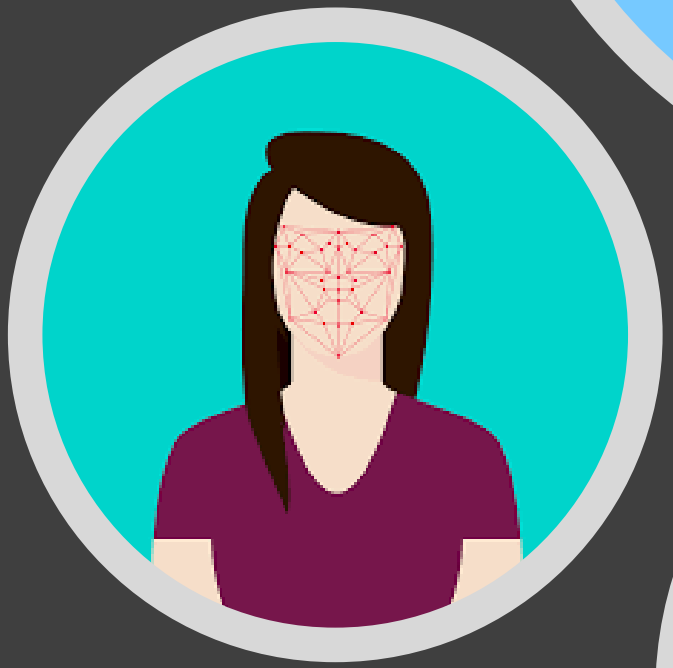
3. Assault : This event recognizes sudden physical or violent attack on someone.

4. It also includes other activities such as Arrest, Explosion, Arson and Fighting.

Face Recognition

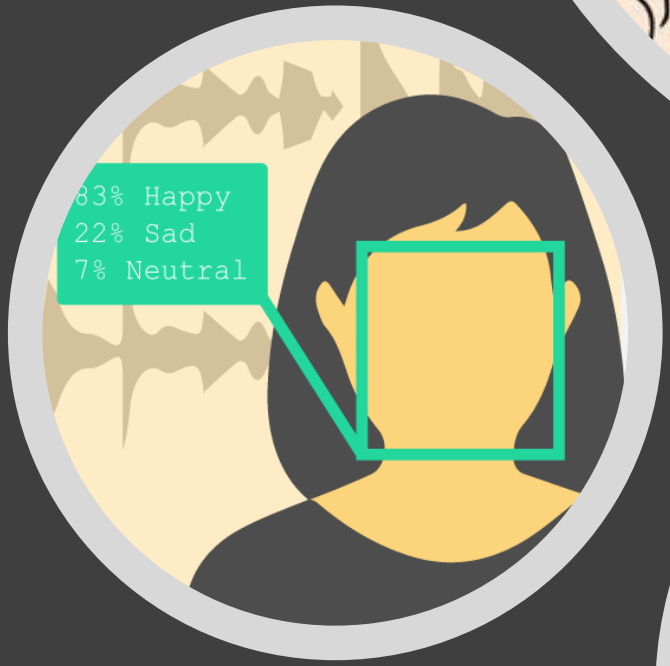
- By using the obtained face data of a person from criminal records, we'll train the face recognition module.
- For now, we have used few of our teammate's face dataset and trained our module in such a way that it can recognize them on any live video footage.
- For the face recognition module to effectively identify the person, he needs to be at a minimum distance of 5 mts. in the live video footage.

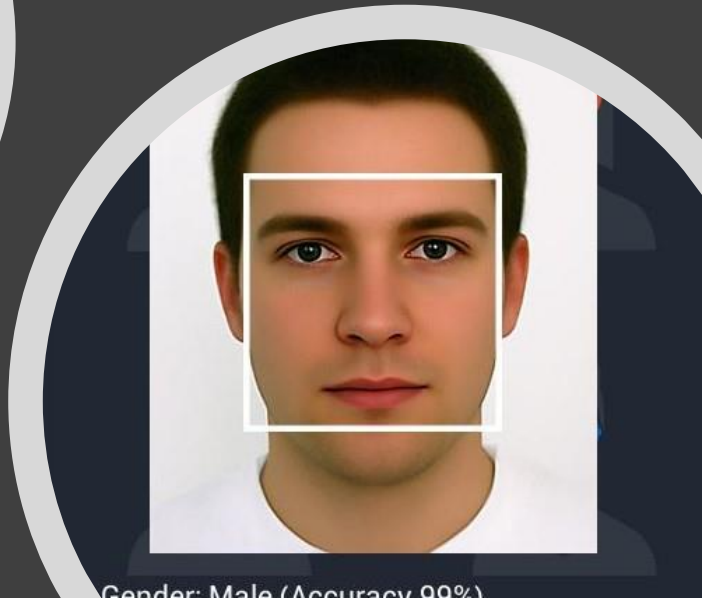


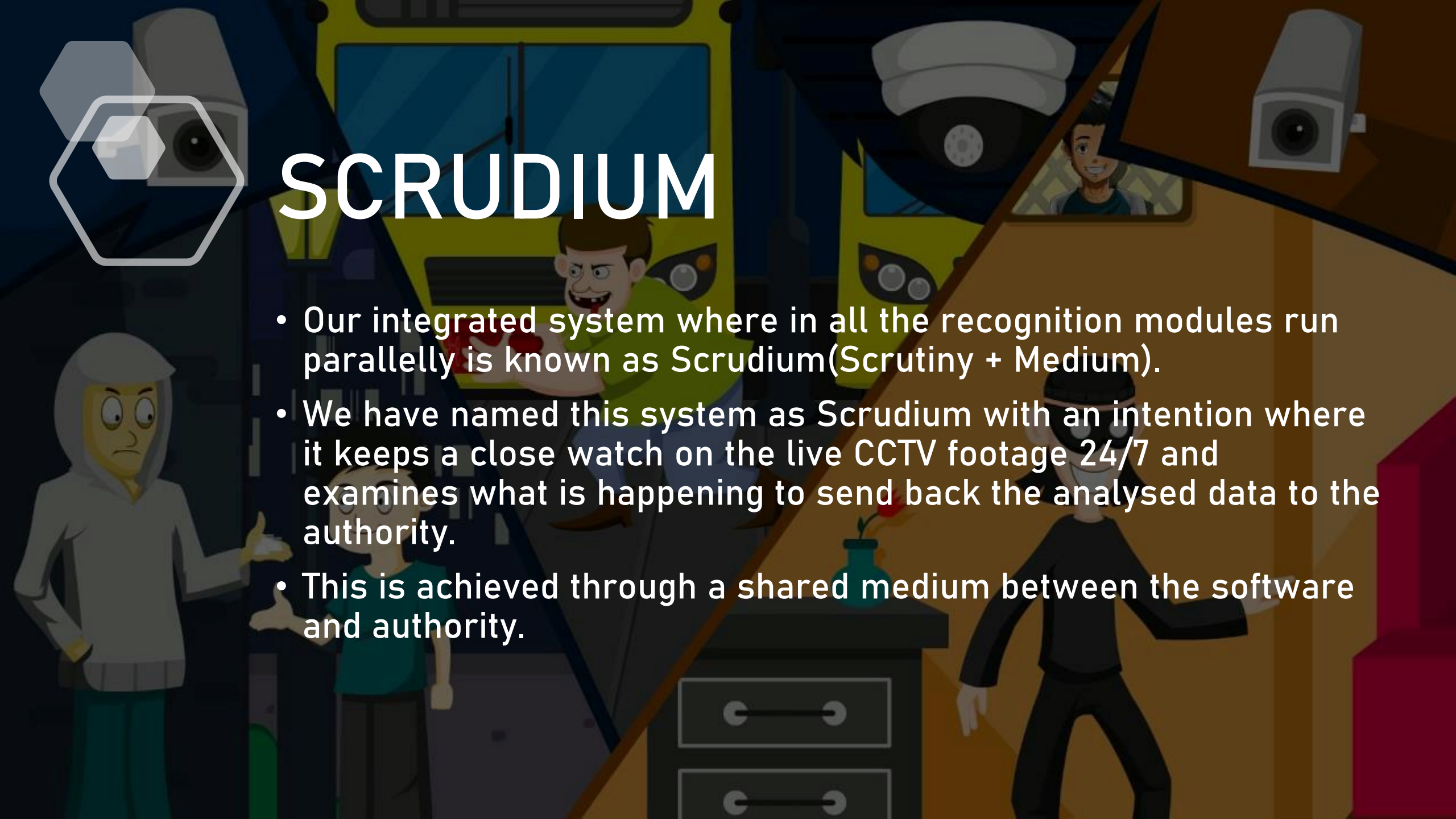


Emotion, Age and Gender Recognition

- When a person comes in the range of the CCTV camera, emotion recognition module will try to capture his emotion and define it under the predefined set of emotions such as happy, sad, angry, neutral etc.
- We are also trying to determine the age and gender of the person in the footage which can be used for further investigation if required.







SCRUDIUM

- Our integrated system where in all the recognition modules run parallelly is known as Scrudium(Scrutiny + Medium).
- We have named this system as Scrudium with an intention where it keeps a close watch on the live CCTV footage 24/7 and examines what is happening to send back the analysed data to the authority.
- This is achieved through a shared medium between the software and authority.

How Scrudium works ?

- Scrudium software is designed to take the input from the live CCTV camera stream and run all the integrated recognition modules parallelly at the same time.
- It identifies each activity under any of the following trained activity recognition attributes namely “Abuse”,” Burglary” etc.
- At the same time if the associated person’s face comes into the system defined detection range, it’ll try to match that person’s face with the available dataset or mark him as unknown.
- Parallelly, it determines the person’s emotion, age and gender.

- All these recognised attributes can be viewed and examined on Scrudium platform.

How are the authorities alerted ?

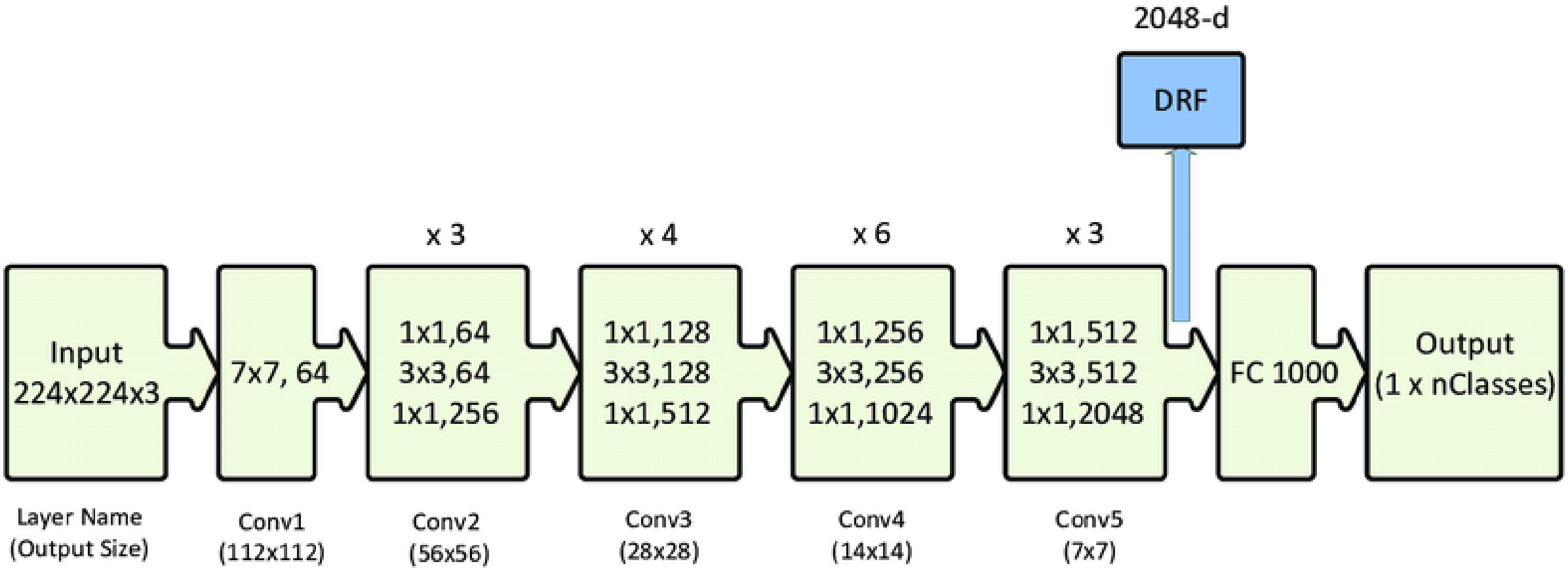
- If any of these activities or a suspect's details are detected by the system, those details will be stored in our database.
- These details can be acquired from the database in real-time according to our requirement.

Firebase as Database

- Firebase provides a real-time database and backend as a service.
- The service provides application developers an API that allows application data to be synchronized across platforms and store on Firebase's cloud.
- We are using this service to store all the recognized activity and criminal details which we would be using as per our requirements.

Criminal Activity Recognition

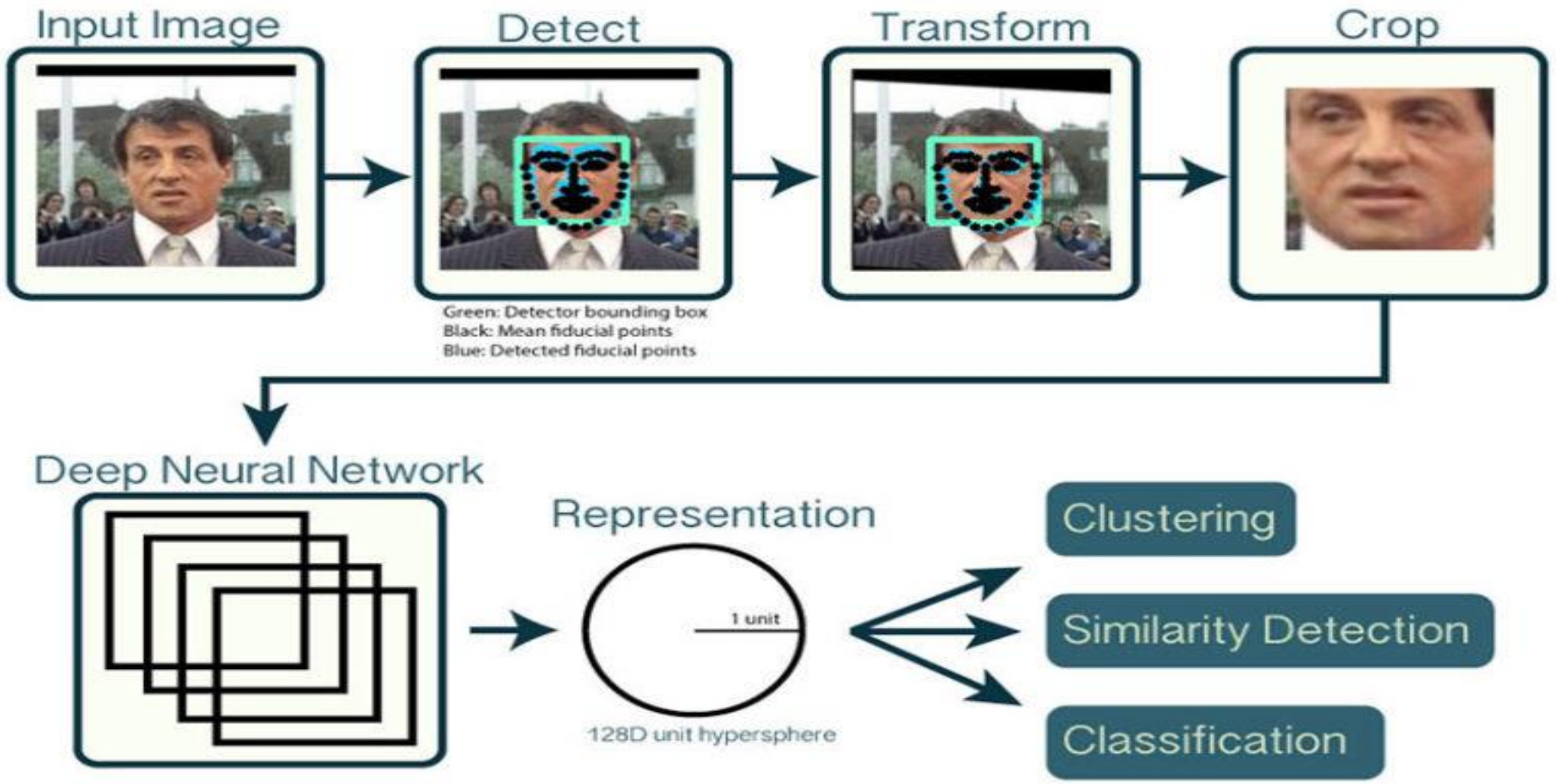
- The architecture used for this module is the ResNet50 alternative of the Convolutional Neural Network model that has 5 main convolutional layers and a fully connected layer to output a predicted class on processing a set of frames of dimensions 224x224x3 .
- Each convolutional layer filters out prominent features by convoluting the input frames to half their size.
- A Keras training script grabs the dataset class images, loads the ResNet50 CNN, and applies transfer learning of obtained weights to train our model. The training script generates/outputs as:-
- A fine-tuned classifier based on ResNet50 for recognizing criminal activities(Arson, Assault etc.) named activity.model.
- A serialized label binarizer containing our unique class labels named lb.pickle.



ResNet50 CNN Architecture

Face Recognition

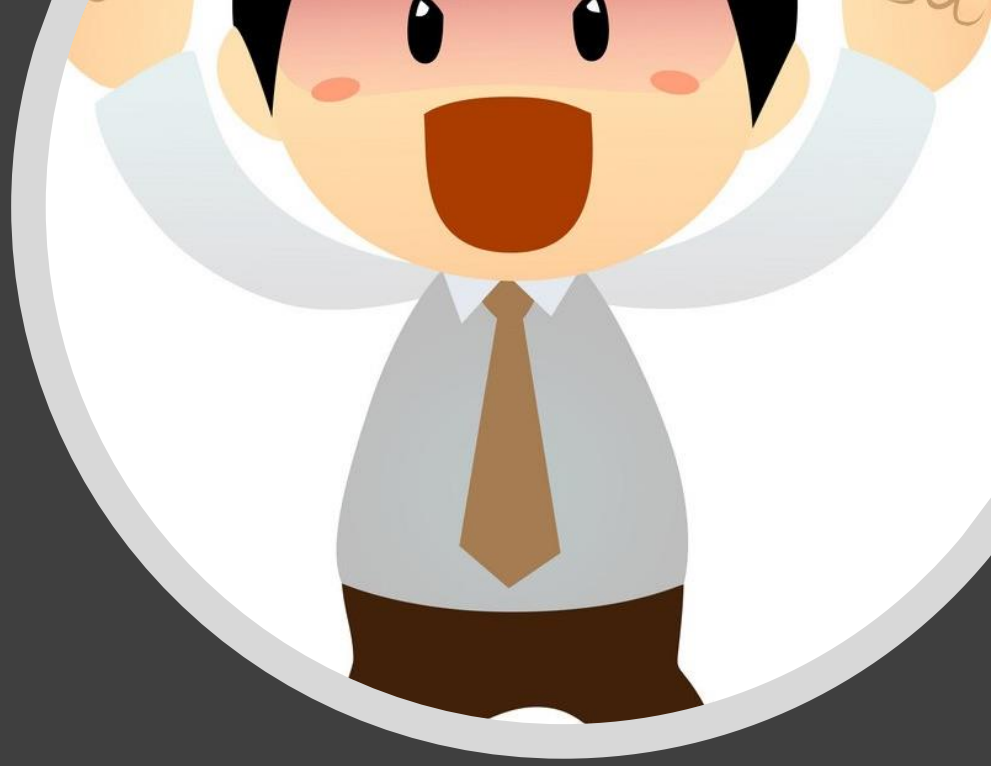
- The architecture used Support Vector Machine for this module is the Res10 model that divides data into ‘n’ classes based on an anchor, positive image and a negative image. This model detects and localizes faces in frames.
- Embeddings have been computed for every face in the dataset and are stored in a file named embeddings.pickle.
- le.pickle contains the name labels for the people that our model can recognize.
- We’ll detect faces, extract embeddings using the torch deep learning model openface_nn4.small2.v1.t7, and fit our SVM model to the embeddings data which are in 128 d format.
- In the video or live stream we get the predicted class of the face recognized in it by our model.

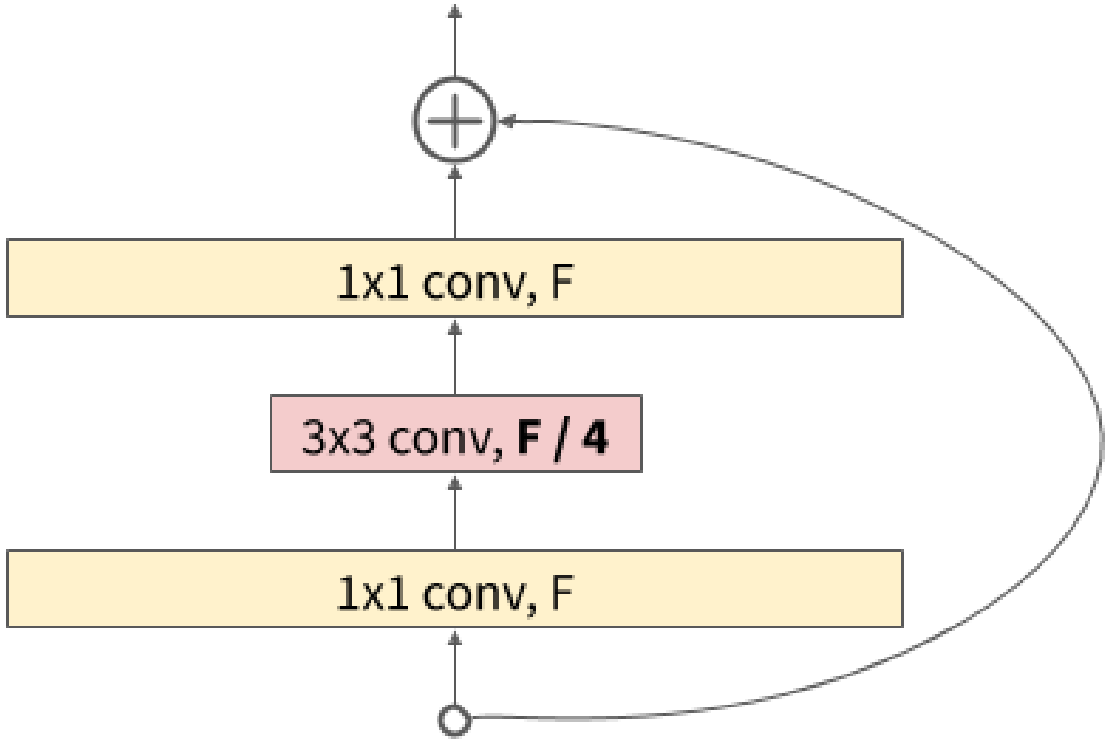


Face Recognition Module Architecture

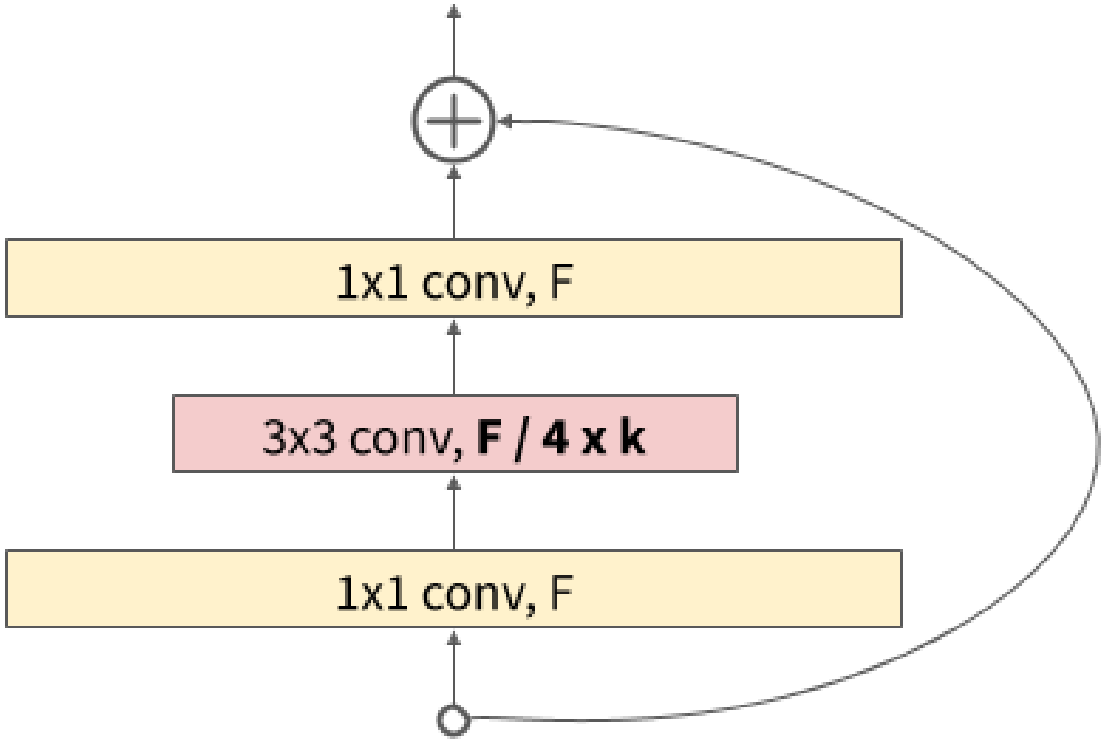
Emotion, Age and Gender Recognition

- This module uses a haarcascade frontal face emotion classifier and a trained WideResNet CNN model to predict the age and gender using argmax() function.
- haarcascade_frontalface_default.xml and emotion_little_vgg_3.h5 are used to classify the emotions that are projected by the suspects in the live stream or video stream.
- The WideResNet CNN model is used to predict the age and gender of the suspect in the stream.





ResNet bottleneck



Wide ResNet bottleneck

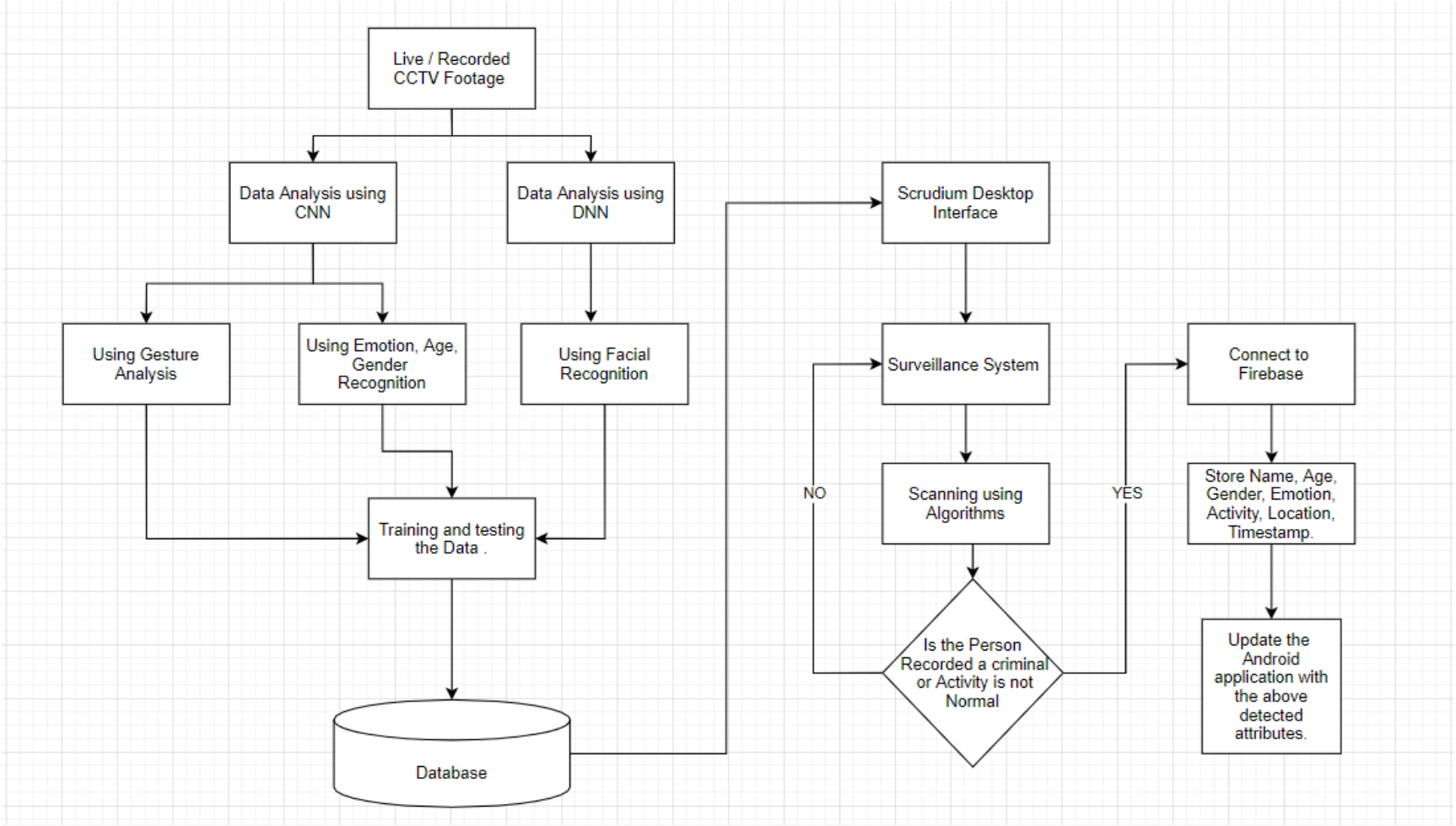
CNN vs WideResNet

Scrudium Android Application

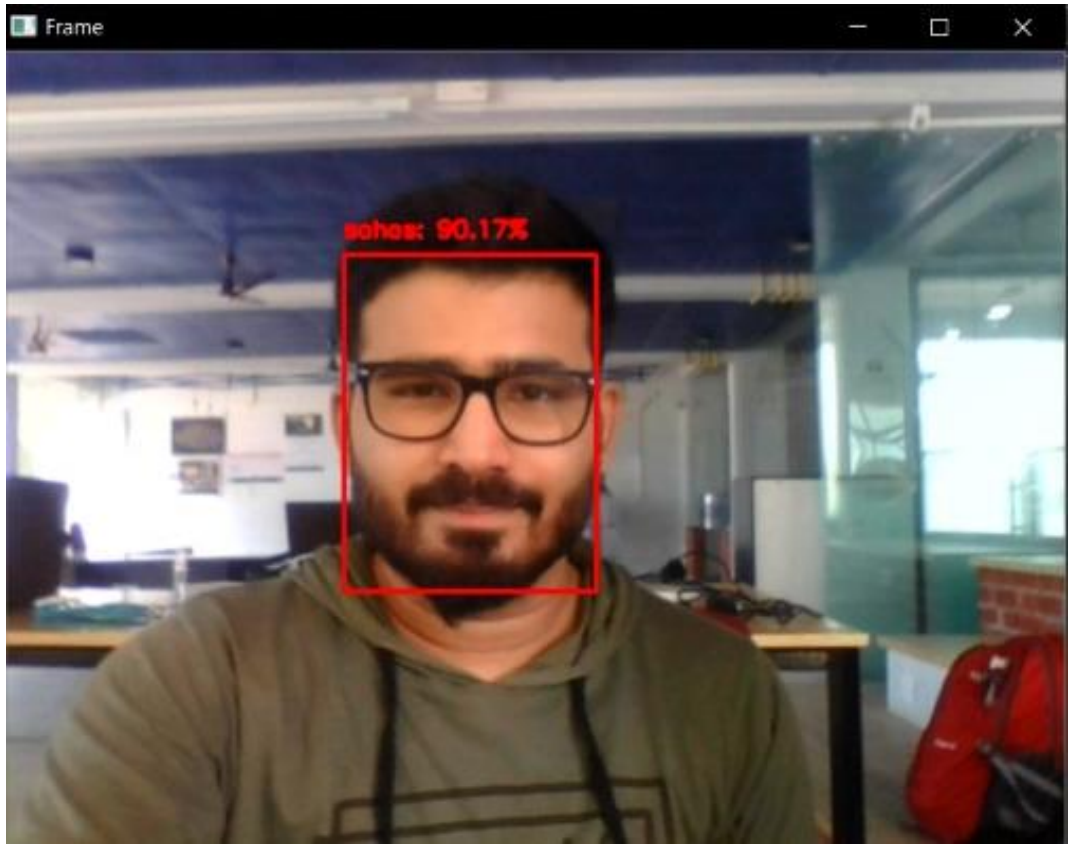
- The Scrudium android application is being developed using Android Studio and will be integrated with the database.
- We will be providing an interface in the application which would contain the following details for each incident occurred :

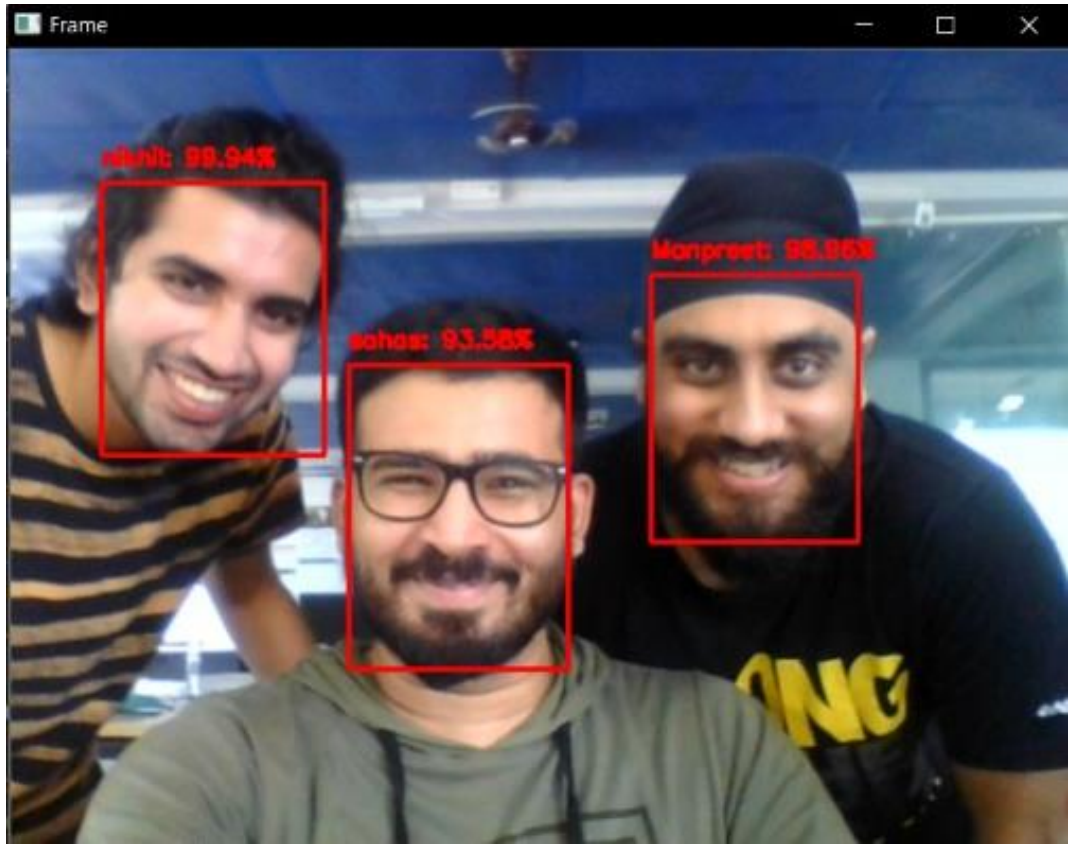
1. Name
2. Age
3. Gender
4. Activity
5. Emotion
6. Accuracy
7. Timestamp
8. Location

Architecture Diagram

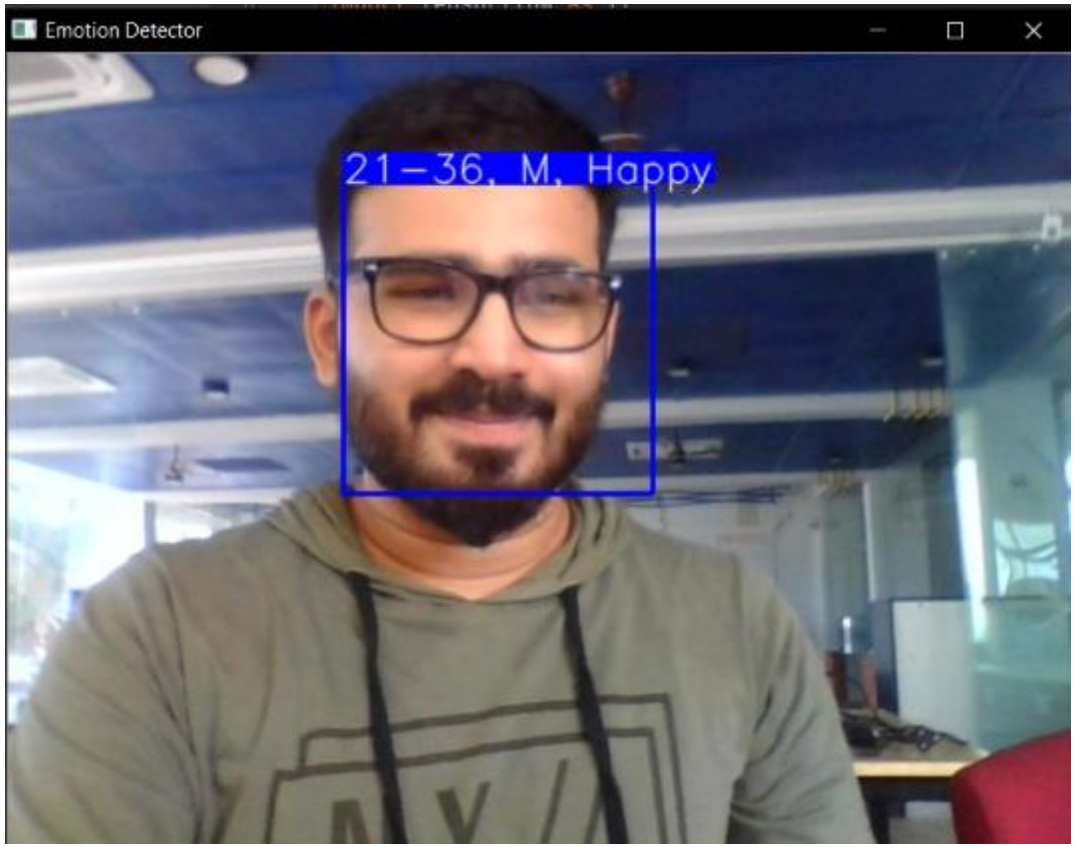


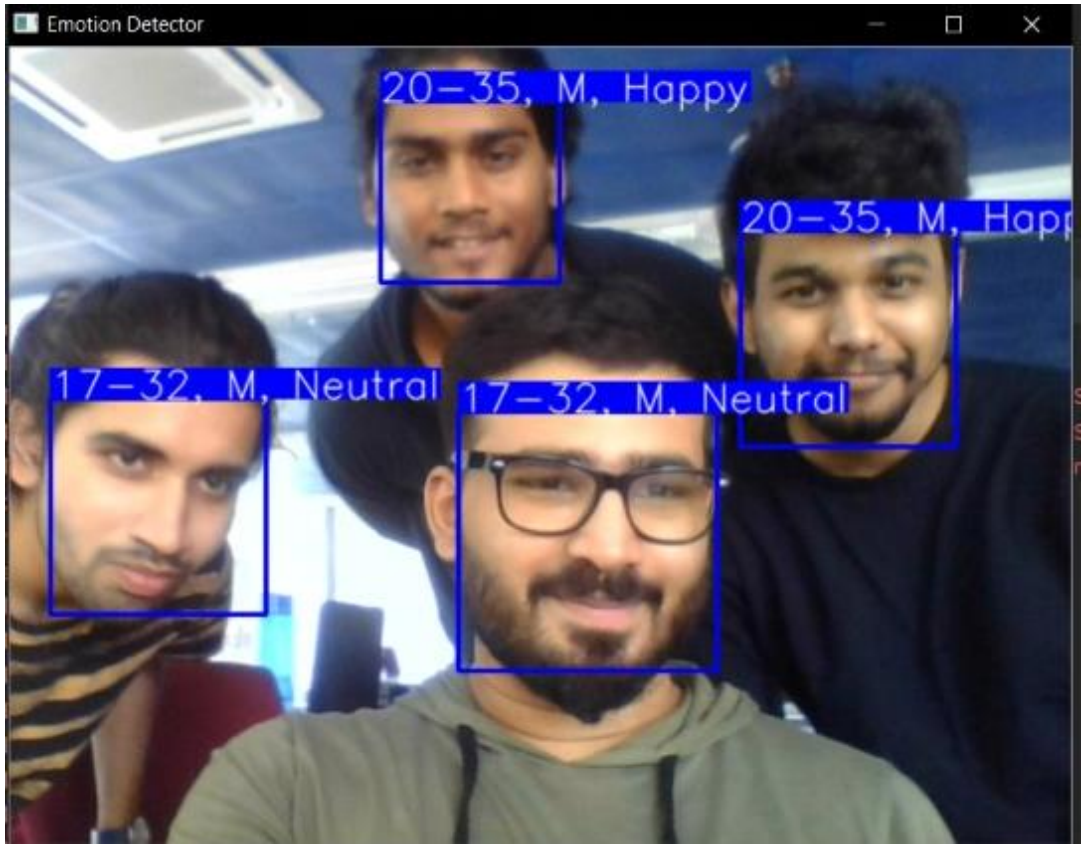
Activity Detection Model



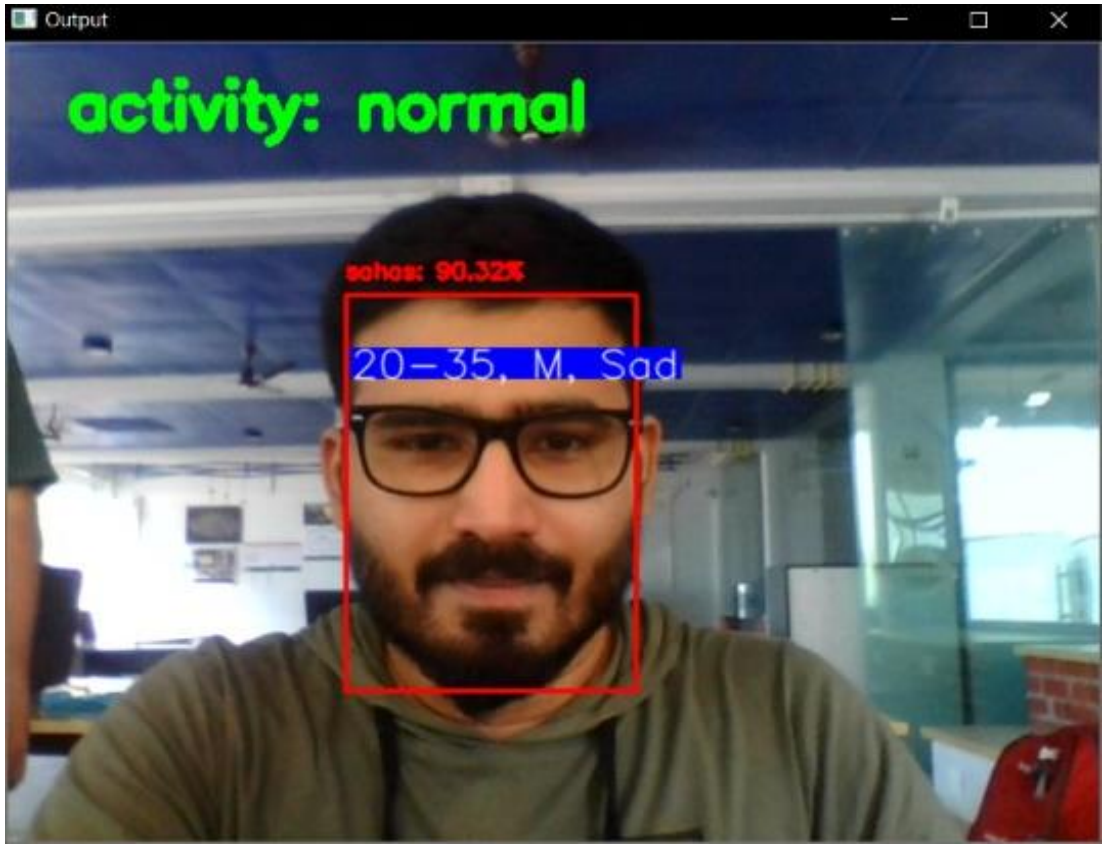


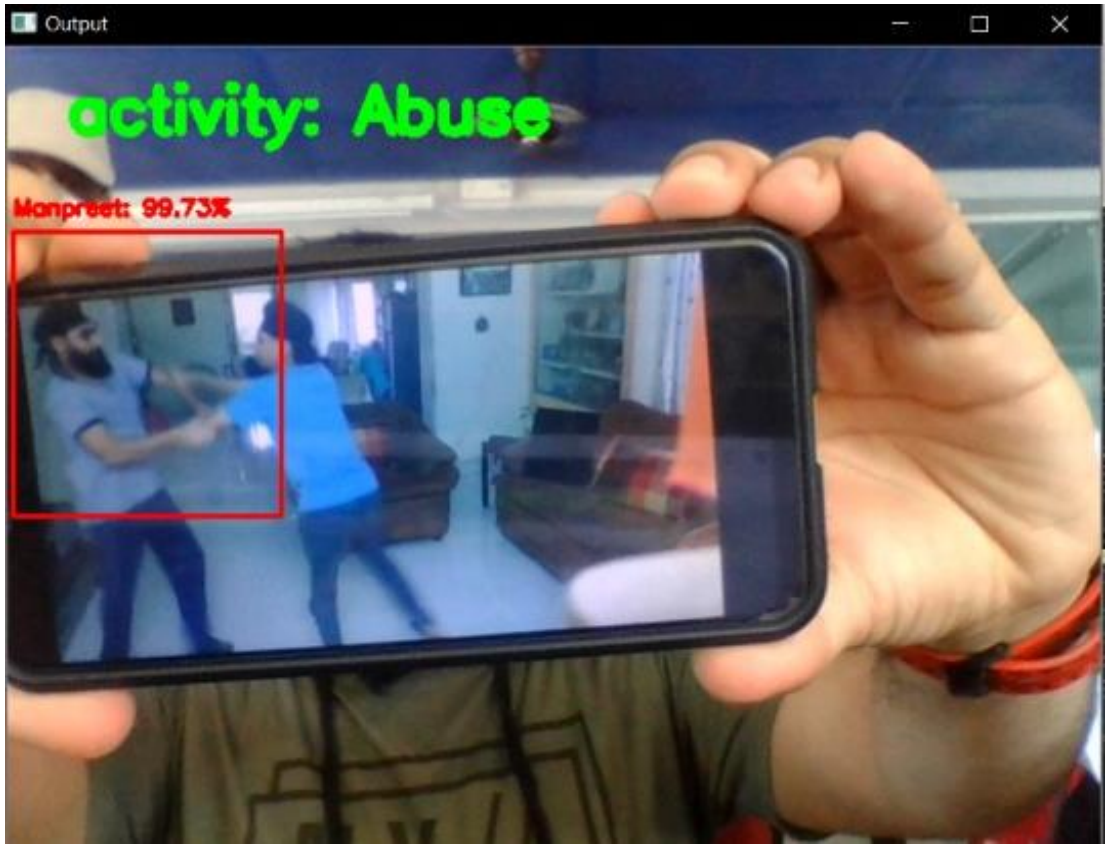
Face Detection Model





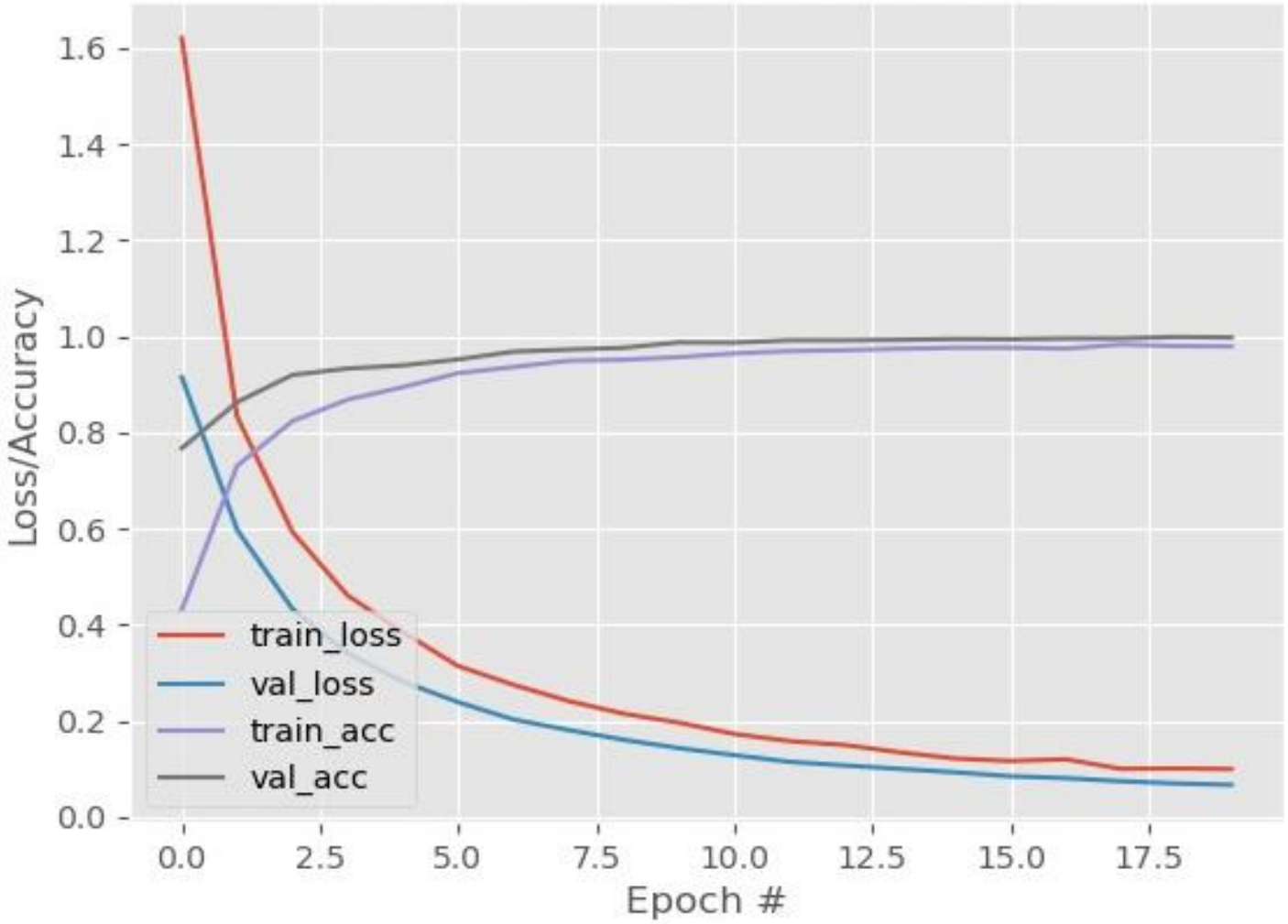
Emotion, Age and Gender Detection Model





Integrated Model Of Face, Gesture and Emotion, Age and Gender Detection

Integrated Model Of Face, Gesture and Emotion, Age and Gender Detection

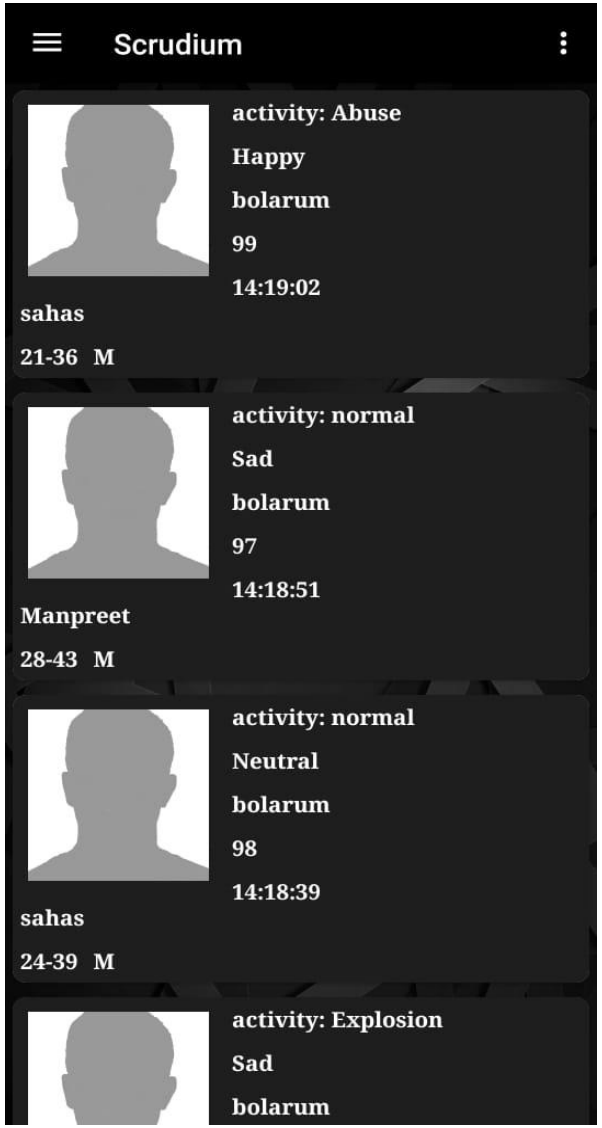


TRAINING LOSS AND ACCURACY ON DATASET

Firebase Database

Graphical User Interface

Web User Interface



Mobile Interface

THANK YOU