A Project report on

IDENTIFYING CRIMINALS THROUGH SURVEILLANCE CAMERA

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

Submitted by

K RAJA GOPAL
(18B91A05B3)

M SRIRAM VARMA
(18B91A05C7)

G SAI SRUTHI
(18B91A0572)

G DATTU NAIK
(18B91A0577)

Under the Guidance of

Mr. L.V. Srinivas
Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SRKR ENGINEERING COLLEGE (A)

Chinna Amiram, Bhimavaram, West Godavari Dist., A.P.

[2021 – 2022]

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SRKR ENGINEERING COLLEGE (A)

ChinnaAmiram, Bhimavaram, West Godavari Dist., A.P.

BONAFIDE CERTIFICATE

This is to certify that the project work entitled **“IDENTIFYING CRIMINALS THROUGH SURVEILLANCE CAMERA”** is the bonafide work of **K. RAJA GOPAL, M. SRIRAM VARMA, G. SAI SRUTHI, G.DATTU NAIK with Regd.Nos 18B91A05B3, 18B91A05C7, 18B91A0572, 18B91A0577** who carried out the project work under my supervision in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering.

SUPERVISOR	**HEAD OF THE DEPARTMENT**
Mr.L.V.Srinivas	Dr. V. Chandra Sekhar
Assistant Professor	Professor

SELF DECLARATION

We hereby declare that the project work entitled "**Identifying Criminals through Surveillance Camera**" is a genuine work carried out by us in B.Tech., (Computer Science and Engineering) at SRKR Engineering College(A), Bhimavaram and has not been submitted either in part or full for the award of any other degree or diploma in any other institute or University.

Date:

1 K. RAJA GOPAL
(18B91A0B3)

2 M. SRIRAM VARMA
(18B91A05C7)

3 G. SAI SRUTHI
(18B91A0572)

4 G. DATTU NAIK
(18B91A0577)

TABLE OF CONTENTS

ABSTRACT

Criminal activities keep on increasing at a rate of 1.6% annually in India even when highly advanced surveillance systems are placed in most parts of the country Identifying criminals in crowded places such as train stations, malls, and airports is of utmost priority. In the current world, surveillance for criminal detection is a time-consuming task that necessitates numerous computations and man-hours, and training hence it is increasingly becoming a burden financially. Many organizations are searching for a better alternative that helps in increasing security with minimal resources. This method involves using Live Face Detection, Face Expression analysis and Face Recognition in finding the criminals, using a haar cascade frontal face classifier for face detection, CNN for face expression analysis, and a face_recognition package from python for face recognition. It takes the help of existing criminal records and searches for them in case of unusual cases.

LIST OF FIGURES

1.INTRODUCTION

In the modern era where highly Technological and Sophisticated systems are in place for every minuscule part of society, securing the world from criminal activities and avoiding them altogether is not focused enough. [1] This is why criminal activities keep on increasing at a rate of 1.6% annually in India even when highly advanced surveillance systems are in place in most parts of the country. According to the aforesaid reference, 51.5 lakh (5.15 million) cognizable crimes were filed nationally in 2019, including 19.4 lakh Special and Local Laws offences and 32.2 lakh IPC offences. The crime rate per 100,000 people grew from 383.5 in 2018 to 385.5 in 2019, indicating a 1.6 percent annual increase in case registration (50.7 lakh cases). Murder, kidnapping, assault, and death by negligence made up more than a fifth of all reported crimes (10.5 lakh). According to the United Nations, the homicide rate in 2020 was 2.95 per 100,000, down from a high of 5.46 per 100,000 in 1992 and basically stable since 2017, higher than most Asian and European countries and lower than most American and African countries, despite being numerically one of the highest due to the large population.[2] Fig 1.1 explains the increase in crimes in India from 1953-2007.

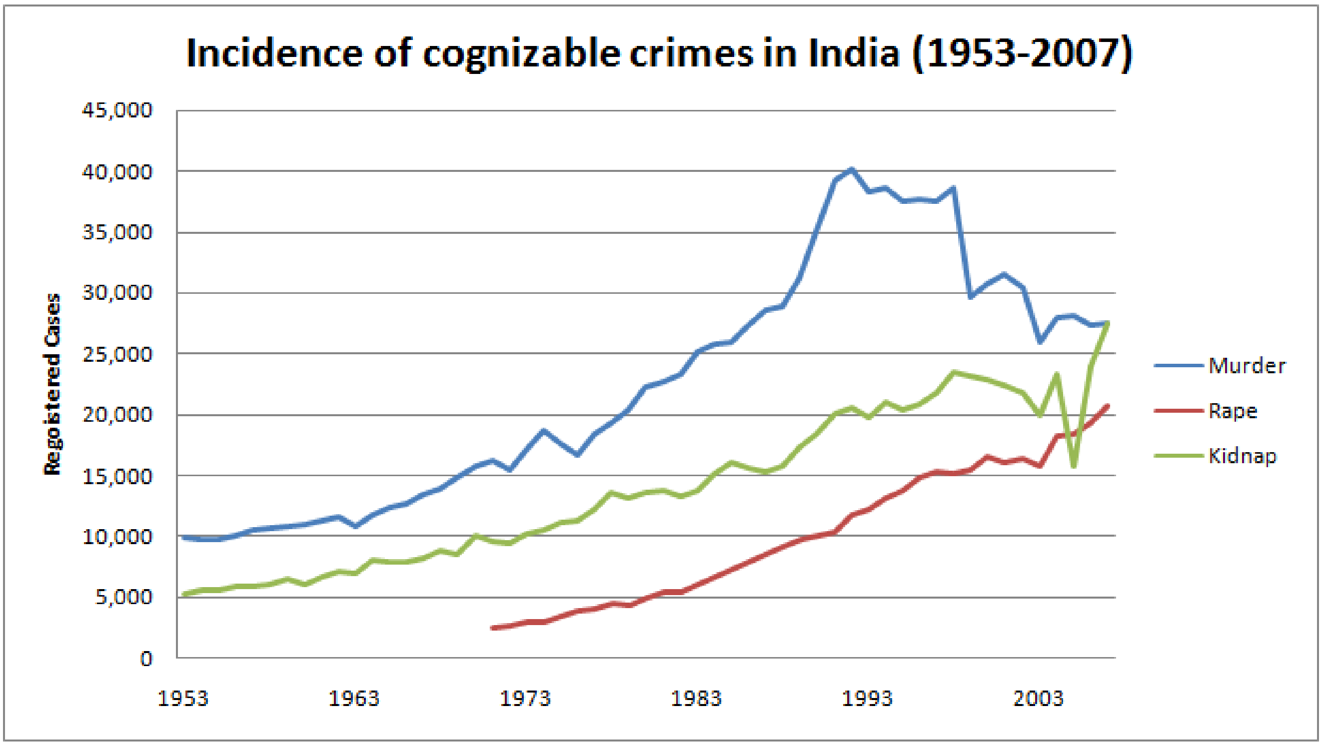


Fig 1.1 Incidence of Cognizable crimes in India

There are many factors that are influencing this gradual increase in criminal activities worldwide. Most important of them all is the high reliance of manual surveillance which not only is ineffective and inefficient but also infeasible on a global level. Manual surveillance requires highly trained professionals and with high expense of having large scale surveillance rooms which is quite infeasible because of the sheer number of surveillance cameras around the world and the sheer amount of people present around the world, as the below image suggests it is increasingly becoming difficult for restoring the security around the world using manual surveillance. [3] This might explain why, despite having 1.54 million cameras in the top 15 cities in India, crime rates are still growing at a massive rate. [4] Fig 1.2 shows the total crime statistics in the top 10 cities in India.

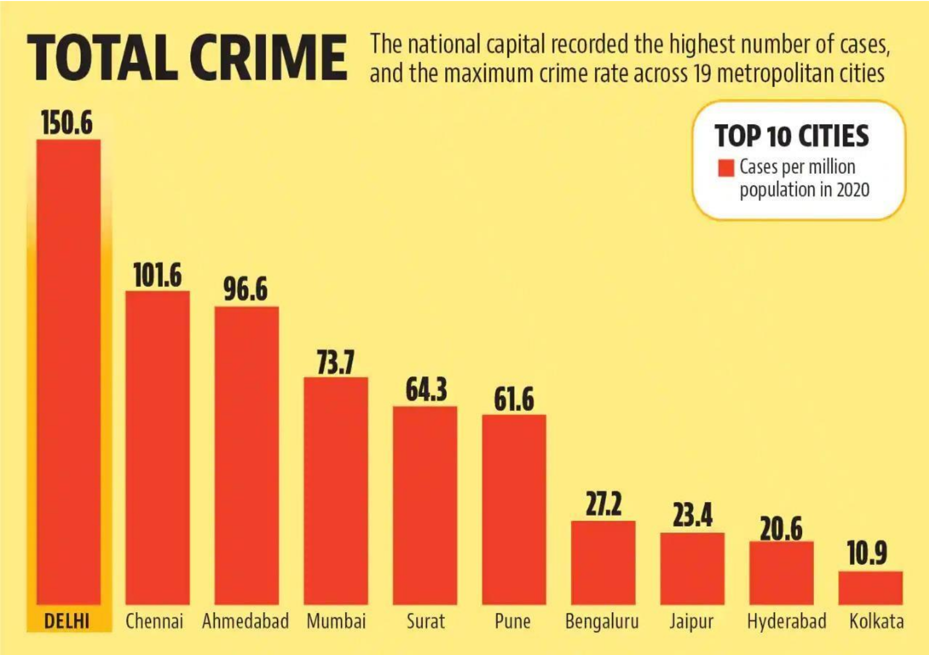


Fig 1.2 Crime statistics in top 10 cities.

[5] With a crime score of 44.63 and a safety value of 55.37, India is ranked 76th in the world. Ineffective criminal detection affects the society on multiple fronts including economic, social, and most importantly loss of human lives. This proves the immediate need for a change in system to a much more robust, efficient and most importantly effective solution in automated surveillance. [6] Fig 1.3 explains how manual surveillance cameras are installed in major cities.

Fig 1.3 Security Surveillance Camera

Automated surveillance decreases the amount of manual work required exponentially and is highly cost-effective compared to its preceding technologies. It is also highly robust as a large number of people can be tracked with modern algorithms and processing capabilities as shown in the below image. This system employs facial tracking, facial expression detection, and face recognition technology, and also a graphical alertsystem to decrease the manual work involved in this system by a large factor. [7] Fig 1.4 shows how automated surveillance using deep learning can help in recognizing criminals easily.

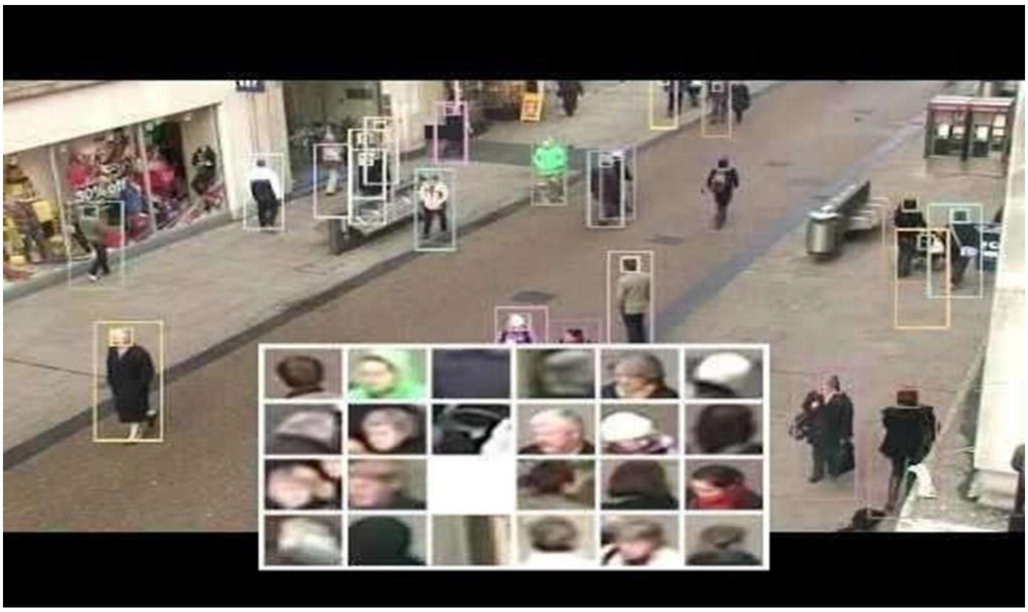


Fig 1.4 City Surveillance camera

Criminology theories state that the facial expression of criminals will be changed while they are committing a crime as the below image suggests. Facial expressions like surprise, fear or anger can be used to detect criminals. This system helps in detecting the people in highly crowded areas and is responsible to track the sudden change in facial expressions to surprise, fear or anger to take those facial images forward to further search them in the existing criminal record to effectively track down criminals even in large crowded areas with minimal effort. This system does not act as an investigative tool which tracks down criminals after an unfortunate incident, rather this system prevents the incident altogether by alerting the officials in real time. Thereby not only detecting criminals but doing it even before they commit a crime. This modification can not only save the economy Trillions of dollars but also most importantly save the lives of humanity. [8] Fig 1.5 shows what criminals' expressions look like.

Fig 1.5 Criminal expressions

2. LITERATURE SURVEY

Traditional criminal detection methods use thumb recognition, however detecting criminals using this method is becoming more difficult. The CCTV footage, which can be used to identify suspects on the spot, can help with this. Using the well-known PCA technique, Nurul Azma Abdullah et al [9] suggested an automatic facial recognition system. The results demonstrate that approximately 80% of the input images can be matched to the template data, implying that this strategy increases criminal detection for effective law enforcement.

Zhang et al. [10] suggested efficient and reliable real-time face detection and recognition techniques in complex backdrops. Cascade classifier, Ada Boost, Haar-like feature, Local Binary Pattern(LBP), Principal Component Analysis (PCA), and face image pre-processing are among the signal processing methods used in the approaches. To train the face and eye detectors with great detection accuracy, the Ada Boost method is employed in a cascade classifier. Face detection uses the LBP descriptor to extract facial characteristics. The eye detection method reduces the number of false faces detected. The identified facial images are then processed to improve the contrast and adjust the alignment of the face, resulting in good facial recognition accuracy. Ultimately, the PCA approach is used to recognise faces quickly.. Face detection and facial recognition algorithms are trained and validated using databases including 2492 images with faces and 4548 images without faces. Face detection algorithms produced a true-positive rate of 98.8%, and proper facial recognition algorithms achieved a rate of 99.2%.

In criminal identification scenarios, eyewitnesses' visual perception is crucial. It aids law enforcement officials in locating specific criminals based on their prior criminal records. It has been stated that manually searching a criminal record takes a long time to produce an accurate result. Avinash Kumar Singh et al. [11] proposed a visual perception-based criminal detection strategy that reduces search space while minimising computational cost. This method is implemented using a database of 240 photos. They employed a rapid query-based searching method based on a data structure called dynamic decision tree that allows for four levels of deconstruction to get relevant criminal information. The system has a 97 percent accuracy rate. Law enforcement can readily detect criminals using this method.

To increase the face identification rate, Ju-Yuan Hsiao et al [12] suggested a hybrid face recognition system based on various facial traits. In a grayscale image, the Adaboost algorithm is used to recognise the entire face as well as some specific facial parts such as the eyes, nose, and mouth. They successfully merged the features gathered from the whole-face and specific facial parts using principal component analysis. Finally, they used a voting technique including the support vector machine and the Euclidean distance to get a recognition rate of up to 99.16 percent, outperforming earlier methods greatly.

The process of identifying and recognizing a criminal is slow and difficult task. SANIKA TANMAY RATNAPARKHI et.al [13] proposed the recognition of criminals by using Face Detection. They used video footage from the camera and used to identify suspects, criminals, missing persons etc. They developed a criminal identification system using Machine Learning and deep neural networks. Face detection system is the combination of face detection, feature extraction and classification. In deep learning they used methods like MTCNN for detection and Face Net. Outputs confirm that the proposed system achieves 90% accuracy over 128 images. This system can be further implemented to detect criminals in real-time using a dynamic dataset and it is also challenging in real-time applications

Face detection techniques are widely used in various real-time applications. MD KHALED HASAN et.al [14] proposed comprehensive review of Human face detection techniques. They explored various face detection algorithms in different ways. They divided those algorithms based on feature and image-based approaches. They published NNs for high performance and feature based algorithms for real-time applications. They came to the conclusion that the best way to choose an algorithm is to understand the problem at hand and the algorithm that will best solve it.

Many real-time applications rely heavily on facial expression recognition. SYED INTHIYAZ et al. [15] proposed a method using KERAS to recognise facial expressions. Convolutional neural networks were used to create this application. OpenCV is also used to recognise faces and create boxes around them automatically. The dataset is divided into two parts: training and testing. The suggested model uses a trained CNN network to provide the proper expression, and it is validated using real-time photos with various expressions.

Surveillance systems are commonly employed in the security and safety industries. GUANGYONG ZHENG and YUMING XU [16] suggested using deep learning to recognize and track faces in video sequences. This model employs the SENResNet face detection model and the Regression network-based face tracking model, as well as ResNet to tackle the gradient disappearance problem in deep network training. The accuracy and performance of this model on public face and video datasets are superior. They attained an average accuracy of 80%.

A. S. JALAL et.al [17] proposed suspect detection by the Facial mole detection approach. Face moles are very useful when recognizing a criminal. In this system, they used Normalized Cross correlation to identify spots on the face and determined the type of spot by using the Modified ResNeXt-50 CNN model. Experimental evaluation and comparison with the state-of- the-art method show that the proposed model outperforms the previous models.

Mahyar Najibi et.al [18] proposed FA-RPN: Floating Anchor Region Proposal Network for Face Detection. For producing region ideas, they used a pooling-based technique. To reduce the number of anchor-boxes, they adopted an effective anchor placement approach. In this research, they show that for face identification, suggestions generated by their network (FA-RPN) are superior than RPN generated region proposals. On the WIDER dataset, which contains 158,989 training pictures and 39,496 testing images, their face detector based on FARPN achieves 89.4% mAP (Mean Average Precision).

S. Shirsat et.al [19] Proposed System for Criminal Detection and Recognition on CCTV Data Using Cloud and Machine Learning. The model was implemented using Microsoft Azure Cognitive Services and the Cloud system. They used the Haar classifier for detecting the faces and VGG Face recognition algorithm for recognizing the faces. They used the Dataset consisting of 2.6M images and achieved 99.2% training accuracy.

K. M. M. Thu et.al [20] proposed a Cloud Computing based edge crime assistance. Face identification utilising Haar face detection and face identifier creation using Local Binary Pattern from collected photos will be performed on edge devices. Face identifier matching is processed on the cloud. They used a Georgia tech face dataset which consists of 400 face images for training and testing. They achieved 95% accuracy**.**

Kar et al. [21] suggested a Face expression recognition system based on least square SVM and type II ripplet transform. Their method has three steps. The characteristics from facial photographs are extracted in the first step using the ripplet transform type II (ripplet- II). The following step combines a mix of PCA and LDA techniques to generate a more compact and discriminative feature set. In the last step of classification, a least squares form of a support vector machine with a rbf kernel is utilised. They used the dataset, which includes 210 adult faces. The accuracy of their system is 98.97%. The results of their experiments show that their system outperforms others.

Isha Talegaonkar et.al [22] proposed a Real Time Facial Expression Recognition sytem using Deep Learning. The face detection process was implemented in OpenCV using the Haar cascade classifiers. They trained a CNN model to detect real time face expression. They used the FER2013 dataset which consists of 3589 testing and 28709 training images. Their model achieved a test accuracy of 60.12% and a validation accuracy of 89.78%.

M. Georgescu et al. [23] suggested Facial Expression Recognition using Local Learning with Deep and Handcrafted Features. They experimented with several CNN architectures, pre-trained models, and training strategies to get automated features. To predict the class label for each test image, they used a local learning framework. Three phases make up the local learning framework. To choose the closest training samples for an input test picture, a k-nearest neighbours model was used first. Second, the chosen training samples were used to train a SVM classifier. Finally, just the test picture for which the SVM classifier was trained was utilised to predict the class label. On all data sets, they outscored cutting-edge techniques by more than 1%.

Surveillance systems are widely used for their security and safety. Traditional approaches are insufficient to identify suspect behaviour. MOSSAD BEN AYED et.al [24] proposed the recognition of fear using a camera. Using three techniques: Bandpass Filter, Eulerian Transformer, and Lagrangian Transformer, on a facial video, heart rate is estimated. A combination of the proposed method and facial expression recognition can be used for better results, but it is also challenging in the real-time applications.

For effective facial expression identification, Zhengning Wang et al [25] presented a directed attention ensemble. Their network model has two branches: a maintenance branch with many convolutional blocks to exploit high-level semantic characteristics, and an attention branch with a UNet-like design to extract local highlight information. They started by entering the facial image into the maintenance branch. They determined the correlation coefficient between a face and its sub-regions for the attention branch. After that, they create a weighted mask by combining the face landmarks and correlation coefficients. The attention branch was then supplied a weighted mask. Ultimately, the two branches were connected to produce categorization results. Experiments are conducted on four of the most popular face expression datasets, and acheived 50 percent accuracy.

Using text and facial recognition techniques, Rachna Jain et al. [26] created a machine learning-based practical crime detection system. They presented a system with two elements, one for text recognition and the other for face recognition. Where Text recognition includes extracting characters from Indian car number plates and comparing the expected output to the registered vehicle collection. Facial recognition is a feature that identifies criminal faces based on certain face areas and then maps those coordinates to the criminal database. Over 50 photos of individual faces and targets were used to increase facial recognition accuracy, with over 85 percent of successful recognitions in regular working situations.

For detecting criminals, Muhammad Sajjad et al. [27] suggested a Raspberry Pi and cloud-assisted facial recognition system. Because the Raspberry Pi has limited resources, they had to store and train their suggested classifier in the cloud. The system is trained using the world's biggest labelled data set, which contains four million faces from four thousand people. Experiments show that the suggested technique is capable of accurate detection in excess of 70% of cases. This method assists law enforcement in recognising a suspect's face and comparing it to picture databases of known offenders.

Kiran Amjad et al. [28] suggested a strategy and architectural concept for criminal detection based on Lombroso Theory that uses deep learning to anticipate criminals or criminal prospects. The ResNet50 model, which is built on CNN and SVM Classifiers for feature extraction from the dataset, was used for training. The model was trained on 2500 photos and tested on 800 images using two separate labelled datasets. With almost 99 percent accuracy, the suggested technology might assist investigating cops in narrowing down the suspect pool.

Surveillance security cameras are rapidly implemented for monitoring reasons in practically every site. After an event, the images captured by these cameras are utilized. Salim Afra and Reda Alhajj [29] developed an early warning system that recognizes humans in surveillance camera environments and then uses data from several sources to identify and classifies these people. The goal is to use convolutional neural networks to train a feature extraction model for facial identification using the Chokepoint dataset, which was acquired using surveillance cameras. The goal is to collect and assess information about unknown individuals by matching camera photos to a database of social media accounts. This concept proposes that establishing a comprehensive early warning system is effective in combating and identifying criminals so that authorities may respond quickly.

Face recognition is crucial in a variety of professions. Parekh Payal and Mahesh M. Goyani [30] offered a complete research on face recognition that included numerous face identification approaches and obstacles. Face recognition has a number of obstacles, including lighting, position variation, facial emotions, occlusions, and ageing. The basic phases in every face recognition system are pre-processing, face detection, feature extraction, optimal feature selection, and classification. They gave a thorough examination of each in this paper. Appearance-based approaches and geometry-based methods are two types of feature extraction techniques. Face recognition works well in most situations, but it fails in loud conditions. A variety of dimension reduction and region selection strategies can be used to optimize execution. Machine learning and deep learning can also help with performance.

Facial Detection and Recognition Using OpenCV on Raspberry Pi Zero with a Raspberry Pi camera module, capacitive touch sensor, and OLED display was discussed by G. S. Nagpal et al. [31]. For face identification in a picture, they employed the Haar Cascade classifier, followed by the Local Binary Pattern Histogram for facial recognition (LBPH). They used OpenCV to develop the LBPH method. They trained the model with up to 400 photos and obtained accuracy of up to 60% in normal settings and 75% in optimal conditions. The system was created using cutting-edge hardware, allowing it to function quickly and effectively.

A. A. Shah et al. [32] implemented a facial monitoring system by combining MATLAB's face detection and face tracking algorithms with the GPIO pins of the Raspberry Pi B via the RasPi command, so that the array of LEDS tracks the facial movement by detecting the face using the Haar classifier, tracking its position in the range assigned using the eigen features of the face, which are detected using MATLAB's eigenvectors, and face tracking. They created a unique kind of facial tracking on a live broadcast using the Viola Jones algorithm and an infrared camera. However, this application is only capable of monitoring a single face.

Criminal identification is a top priority for police, military, and security organizations. Kavushica Rasanayagam et al. [33] suggested a deep learning-based CNN algorithm for automated criminal identification based on study of faces, emotions, ages, and genders. The whole training is based on the IMDB dataset, which comprises around 20 million photos. They tested utilizing real-time photos and were able to identify offenders with an average accuracy of 80%.

Face detection analysis utilizing CCTV surveillance was proposed by Worawut Yimyam et al. [34] to detect criminals. Face detection using the Viola-Jones technique, face recognition utilizing eigen faces, and a notification system for prompt action are all included in this system. For training and testing, they utilized a database of 15 photos. The results found an average of 35- 55% accuracy over that database. When tilted beyond 90 degrees, however, the detection is unsuccessful.

For human-centered smart surveillance, X.Chen et al. [35] developed a technique for eyes-to-face synthesis and personal identification. The face information is generated via an end-to-end network based on conditional generative adversarial networks using just the data from the eyes area. A synthesis loss function based on feature loss, GAN loss, and total variation loss is utilized to steer the training process in order to acquire photorealistic faces and identity-preserving information. They employed 6730 eyes-face samples and got a 91 percent accuracy on average. Experiments show that the suggested technique not only produces compelling findings but also has a high level of resilience, paving the way for the creation of human-centered smart surveillance systems.

Yingruo Fan et al. [36] introduced a Multi-Region Ensemble CNN framework for facial emotion detection that increases CNN models' learning capability by collecting both global and local information from several sub regions of the human face. They tested the model with the RAF-DB dataset, which comprises 12271 training photos and 3068 testing images, and found it to be 74.78 percent accurate. In comparison to other state-of-the-art approaches, the MRE- CNN framework has obtained a highly attractive performance over RAF-DB dataset.

Teddy Mantoro et al. [37] developed a Multi-Faces Recognition Process based on Haar Cascades and Eigenfaces. They used an image with 60 faces and was able to detect upto 55 faces accurately. Haar Cascades algorithm was used for face detection and Eigenfaces method was used for face recognition. It was able perform with an accuracy 0f 91.67%. It was still able to detect faces until about 15 degrees tilt.

Brendan F. Klare et al. [38] suggested a Descriptive Facial Attributes-based Suspect Identification method. They created a collection of 46 face features to allow this method. The proposed qualities are manually remarked on a huge corpus of face photos. They utilized 1194 photos in their dataset. To increase the extraction accuracy, localized face components were used to extract attributes. They got an accuracy of at least 85.0 percent. The suggested technique aids sketch-based face identification by allowing authorities to explore face repositories without the time delay that sketch production causes.

3. PROBLEM STATEMENT

Criminal activities keep on increasing around the world, despite strict measures. Identifying criminals using manual surveillance is tedious and, in most cases, highly ineffective. There are many factors that are contributing to the failure of manual surveillance. Manual surveillance requires highly trained professionals and with the high expense of having large-scale surveillance rooms which is quite infeasible because of the sheer number of surveillance cameras around the world. Ineffective criminal detection affects society across multiple fronts including economic, social, and most importantly loss of human lives. This proves the immediate need for a change in the system to a more robust, efficient, and most importantly effective solution in automated surveillance.

Objectives:

- Face Detection
- Face expression recognition
- Face Recognition
- Trigger Alert

4.1 Dataset:

4.1.1 Face Expression Recognition dataset:

We are using a dataset from Kaggle [39] which consists of over 36,000 various kinds of expressions. The dataset includes 48x48-pixel images of diverse emotions which were shown in the below tables. The tables table 4.1.a consists of training dataset and 4.1.b consists of validation dataset

Name Of the Expression	No. of Images in dataset
Angry	3549
Disgust	403
Fear	3719
Surprise	3060
Happy	6228
Neutral	4152
Sad	4429

Table 4.1.a Training dataset

Name Of the Expression	No. of Images in dataset
Angry	960
Fear	1018
Disgust	111
Surprise	797
Happy	1825
Neutral	1216
Sad	1139

Table 4.1.b Validation dataset

4.1.2 Face Recognition dataset:

We are using 15 criminal images for getting the facial encodings of the criminals

4.2 System Architecture:

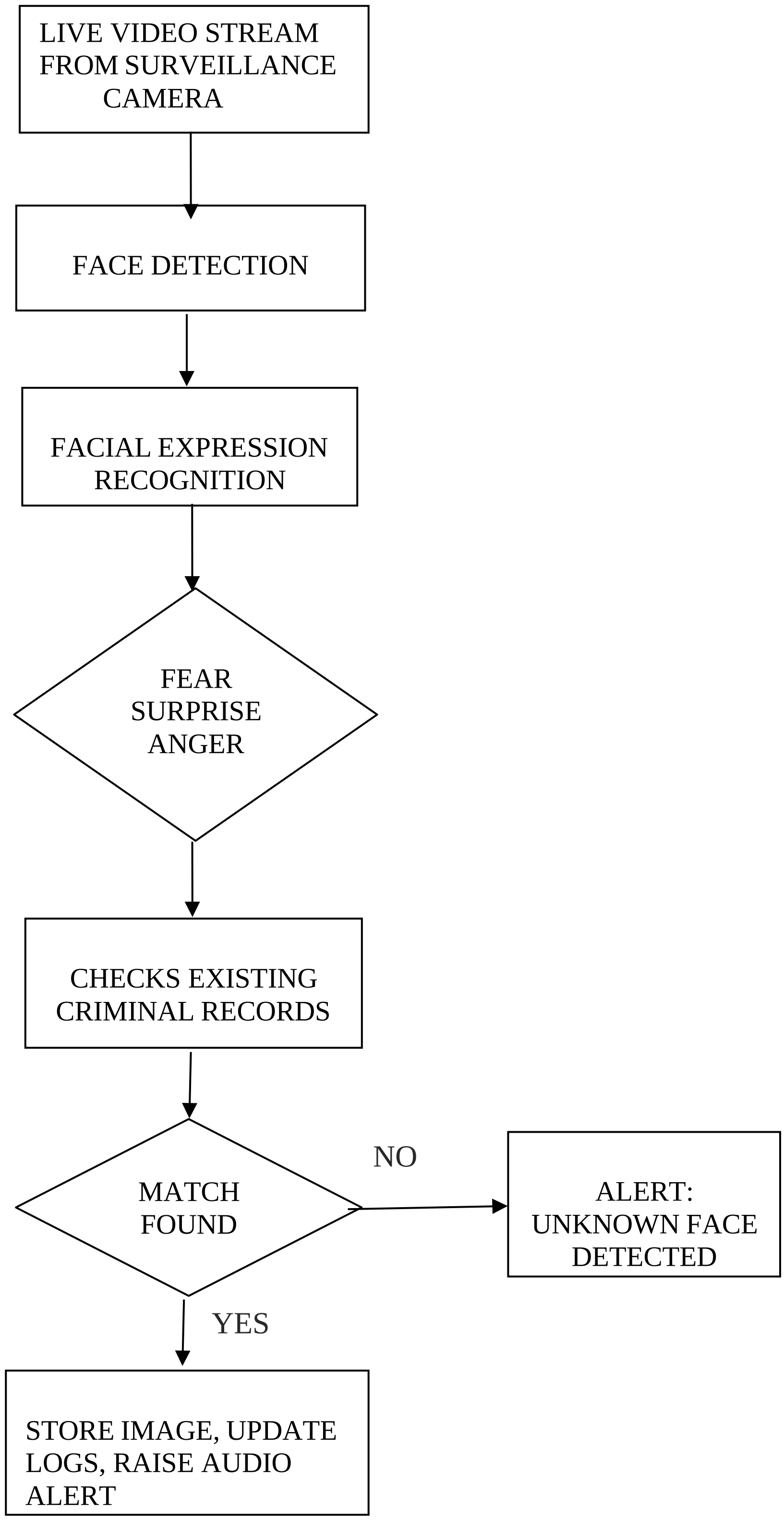


Fig 4.2 Flowchart for automated surveillance system

As Fig 4.2 explains that this project starts from taking live video from surveillance camera and detecting the faces available in the frame, identifying the facial expressions from the recognized face and if the expression is anger or surprise or fear then we try to match that face with the faces of known criminals if the match is found then we are storing the image, updating the logs, raising audio alert. If the match is not found then we are giving alert as unknown face detected

4.3 Models Used

4.3.1 Face Detection:

Face detection has a great amount of importance across various applications. We use the Haar Cascade Classifier for face detection along with OpenCV. It is less computationally expensive, has a quick algorithm, and provides great accuracy. Many pre-trained classifiers are included with OpenCV. The cascadeClassifier method of the cv2 module can load those XML files. For detecting faces, we're using haarcascade frontalface default.xml.

The following are the steps involved in face detection

4.3.1.1 Haar Features Calculation

Gathering the Haar properties is the initial stage. A Haar feature is a collection of computations conducted at a certain place on neighbouring rectangular pieces of a detection window. The procedure involves adding and subtracting the pixel intensities in each section. Haar properties include edge, line, and four-rectangle characteristics.

4.3.1.2 Integral Image Creation

Integral pictures, without going into too much detail about the mathematics, helps to compute haar characteristics quickly and constructs sub rectangles, as well as array references for each produced sub rectangle, instead of calculating at each and every pixel. Then Haar features are calculated utilizing them.

4.3.1.3 Adaboost Training

It just picks the most useful features and trains the classifiers to use them. A robust classifier is made by combining all the weak learners. Weak learners are produced by sliding window over the input image and calculating Haar characteristics for every portion of image. After that, the difference is compared to a learnt threshold for differentiating non-objects from objects. To create a robust classifier, a high number of Haar features are required because these are weak classifiers. The adaboost training is shown in Figure4.3.1.

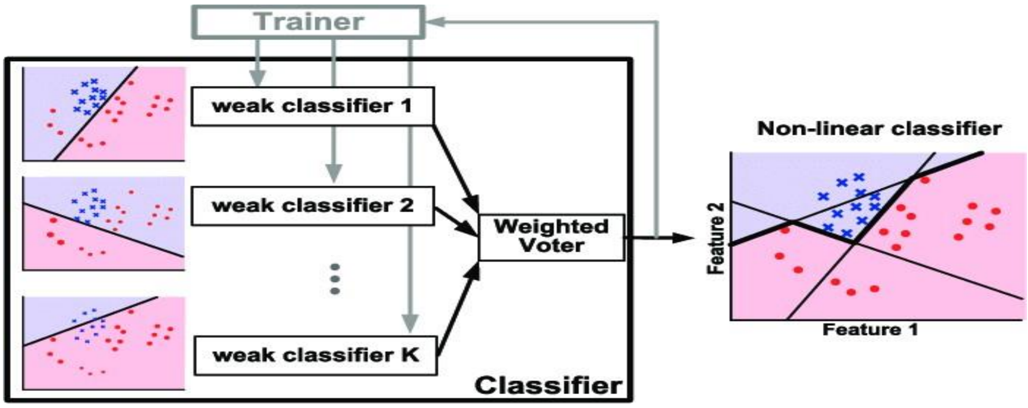


Fig 4.3.1 Adaboost training

4.3.1.4 Cascade Classifier Implementation

Weak learners are integrated in every stage of cascade classifier. They are trained with the help of a boosting method and produces a high accurate claasifier by taking the average of weak classifiers. Then created classifier is used to know whether it has identified an object or not. If no object was found then it simply moves to the next region**.**

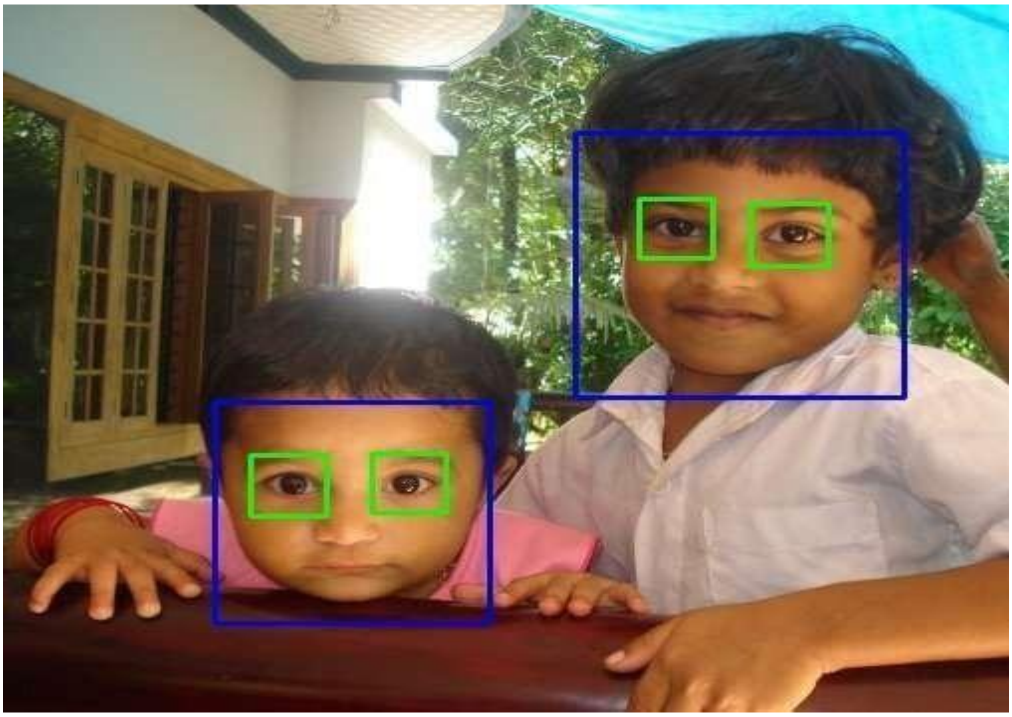


Fig 4.3.2 Face detection example

As we can see in the fig 4.3.2 picture the final results of face detection would be like this

4.3 .2 Facial Expression Identification:

After detecting the faces, we need to identify the facial expression. In order to identify facial expression.

We're going to employ a deep learning-based Convolutional Neural Network (CNN). A Convolutional Neural Network (CNN) is a deep learning neural network that is used to analyse organised arrays of data like photographs. CNNs are widely utilised in Computer Vision and have become the state-of-the-art for a variety of visual applications such as Image Classification, as well as in Natural Language Processing for Text Classification. Here in our project the architecture we are using is 2-D CNN layer along with RELU (Rectified Linear Unit) activation function. The input frame to this layer is of dimensions 48x48. The fig 4.3.3 demonstrates the basic CNN architecture

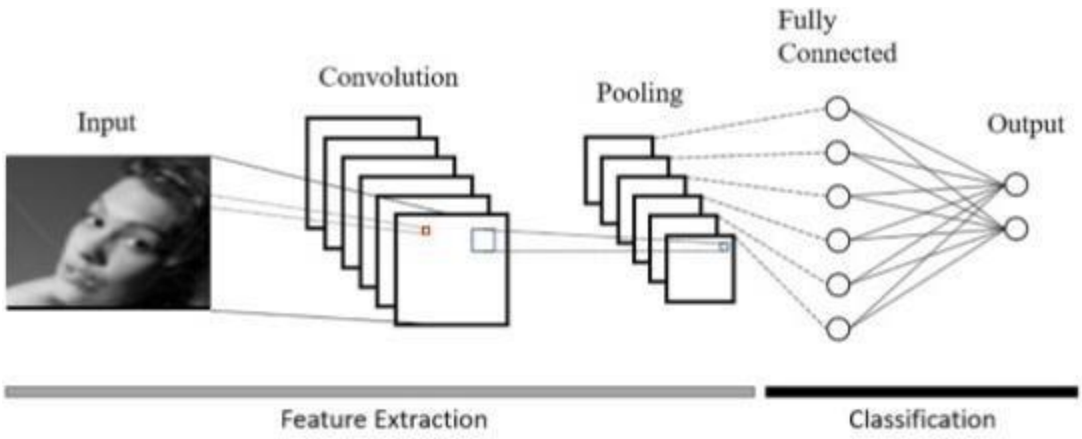


Fig 4.3.3 CNN Architecture

Pooling is also used in the above-mentioned architecture to minimize the size of feature maps which in turn reduces the number of learning parameters.

ReLU Activation function:

It is an activation function that outputs the input in the range (0, infinity). The mathematical function used by this activation function is max (0, x) where x is the input.

Fig 4.3.4demonstrates the ReLU activation function

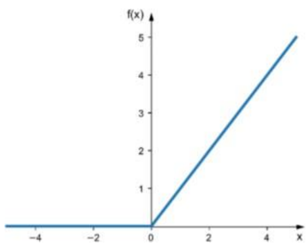


Fig 4.3.4 ReLU activation function

We are also using max pooling to minimize the feature map’s size.

Max Pooling – This identifies only important features of the previous feature map

The fig 4.3.5 demonstrates the max pooling

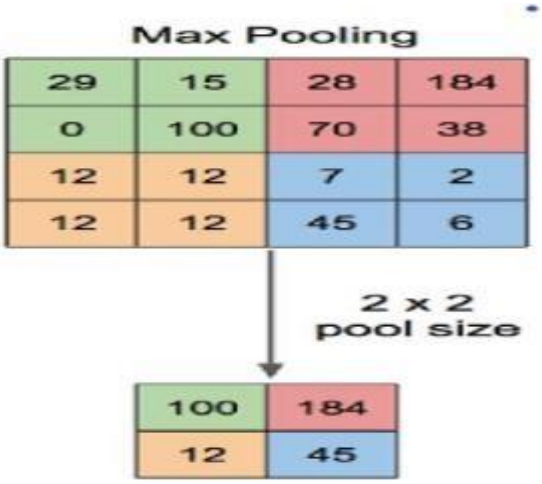


Fig 4.3.5 Max pooling

At the output layer we are going to use the softmax activation function.

Softmax activation function:

The Softmax activation function is a mathematical activation function that converts a number vector into a probability vector. The fig 4.3.6 demonstrates the softmax activation function.

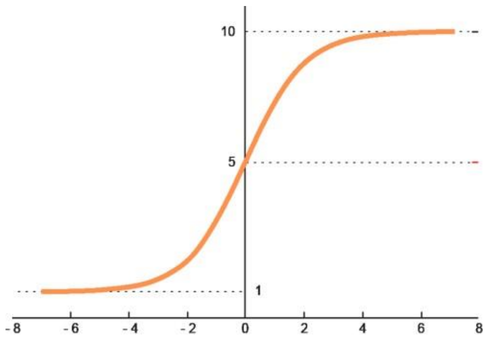


Fig 4.3.6 Softmax activation function

Convolution: Convolution is a process of identifying certain features in a feature map. Convolution helps the model or machine learn some important qualities of a picture through edge detection, noise reduction, blurring, sharpening and more.

Flattening: A 2-D feature matrix is flattened to form a feature vector, which may then be given as input into the dense layer

Various kinds of facial expressions are shown in the below fig 4.3.7

Fig 4.3.7 Various kinds of facial expressions

4.3.4 Face Recognition:

A facial recognition system compares a human face to a database of recognised faces. After identifying facial expressions, we decide to go for face recognition or not based on face expression. If the facial expression is fear or anger or surprise then we perform facial recognition. We will make use of the Python face recognition package. We collect facial encodings of criminal images using the face_encodings()method available in the face_recognition library. The syntax of face_encodings() method is

face_encodings(face_image, known_face_locations=None, num_jitters=1, model='small')

The above method takes a known face image as input and returns 128-dimensional face encodings.

If we find facial expressions as fear or anger or surprise then we get that face encoding of that particular face and we try to compare that encoding with known encodings of criminals. Comparison of encodings can be done using compare_faces() method available in face_recognition library.

The syntax of compare_faces() is

compare_faces(known_face_encodings, face_encoding_to_check, tolerance=0.6)

The above method takes already extracted facial encodings and face encodings and try to compare, then returns a list of True or Falsevalues.

4.3.5 Alert:

After applying the facial recognition method on the obtained faces with suspicious expressions, if the face is recognized as a criminal, then we store the cropped image, add data to log including name of the criminal, time of his detection and also name of the file stored. We also generate alert audio to notify the concerned authorities, we used os package available in python to achieve this.

If the face is not matched with any existing criminals, then we simply display that an unknown face is detected. No alert is raised in this case.

We are taking live video from surveillance camera as input and we are identifying faces in the input video using haar cascade classifier, then we are identifying the facial expression using CNN, later if any suspicious expressions are observed the image is compared with existing criminal records. If any match occurs, an alert is generated to notify the authorities.

This project consists of the following modules:

5.1 Train and label all criminal Images: We gathered images of various people for training the system as criminals. We took images from various resources including google for celebrity images and included some images related to students in our college as shown in fig 5.1.

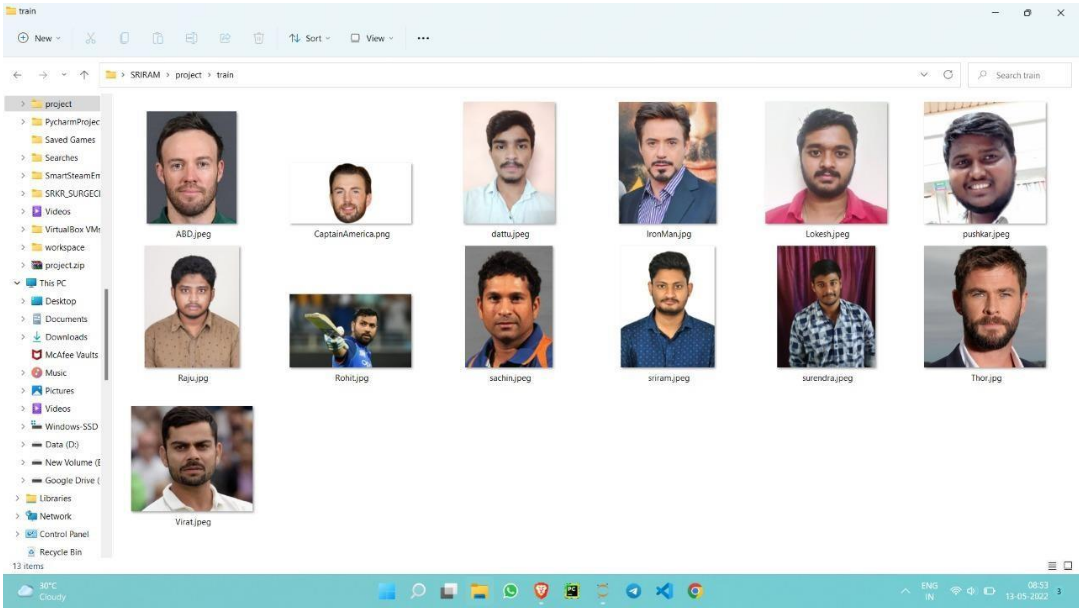


Fig. 5.1 The Training Images for Criminal Record

5.2 Face Detection: This method involves various steps, to load the model for detecting frontal faces, give a live surveillance video footage, and detect all the faces along with processing them accordingly to send it to Face Expression Recognition step

5.2.1 Load the Frontal Face Model: We used haarcascade_frontalface model to detect the faces in the live surveillance video footage.

5.2.2 Image Processing: The live surveillance video footage is taken and each of its frame is then converted to grayscale image and all the faces in the images are cropped and sent to expression detection. If no faces are detected a message is shown accordingly as shown in figure 5.2

Fig. 5.2 The Response Message if no faces are detected

5.3 Face Expression Detection: This step involves detecting the expression of every cropped frame sent from step 5.2 and if the expression is suspicious, then it is further sent to next step. As shown in fig 5.3 the face expression is displayed for reference as a surprise, since surprise is a suspicious expression, we send it to face recognition step

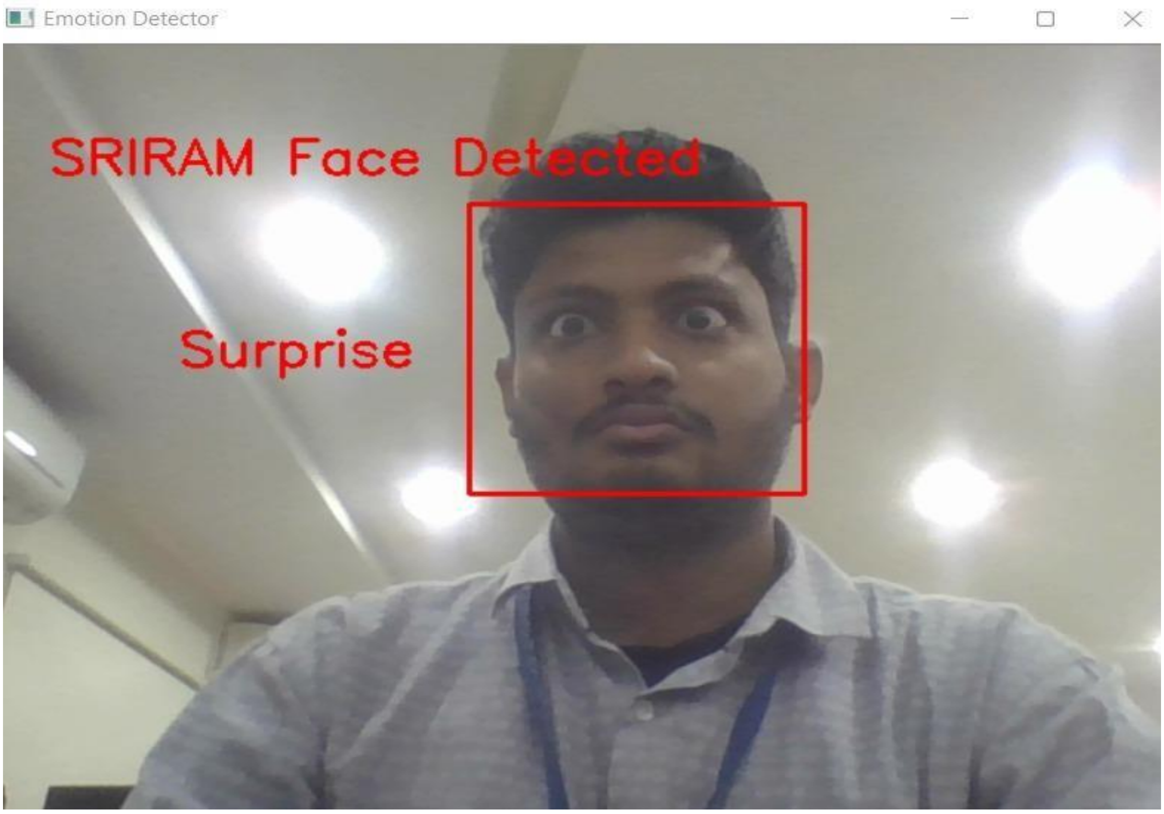


Fig. 5.3 The Detected Face of the criminal along with the expression

5.4 Face Recognition: This step involves the recognition of criminals by comparing them to the existing criminal record. The training data used as criminal images is taken from various sources as shown in fig 5.1, the processed frame of the footage is compared with these images to find a match. As shown in fig. 5.3 the face is recognized of a criminal named Sriram. If unknown face is shown, then a message is displayed accordingly.

5.5 Raise Alert to notify Authorities: If a proper match is found with an existing criminal the image is stored in the suspectedImages folder as shown in fig. 5.4. the images are saved with a pattern of (criminal name)(date of identification).jpg. A log is also maintained as shown in fig. 5.5. Finally, an audio alert is triggered to notify the concerned authorities.

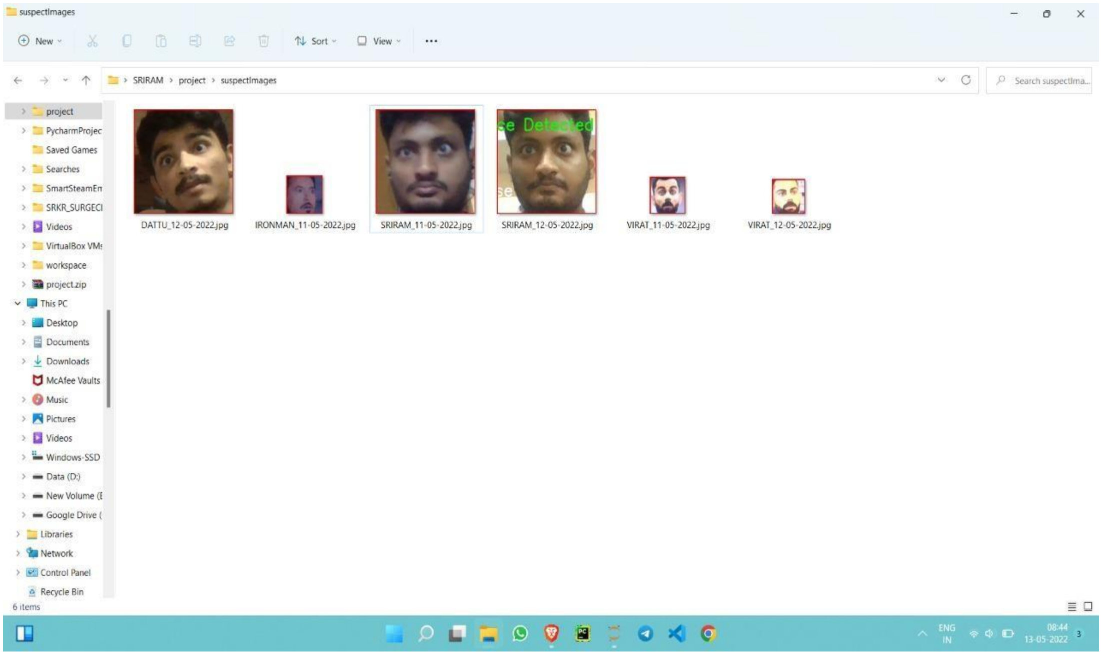
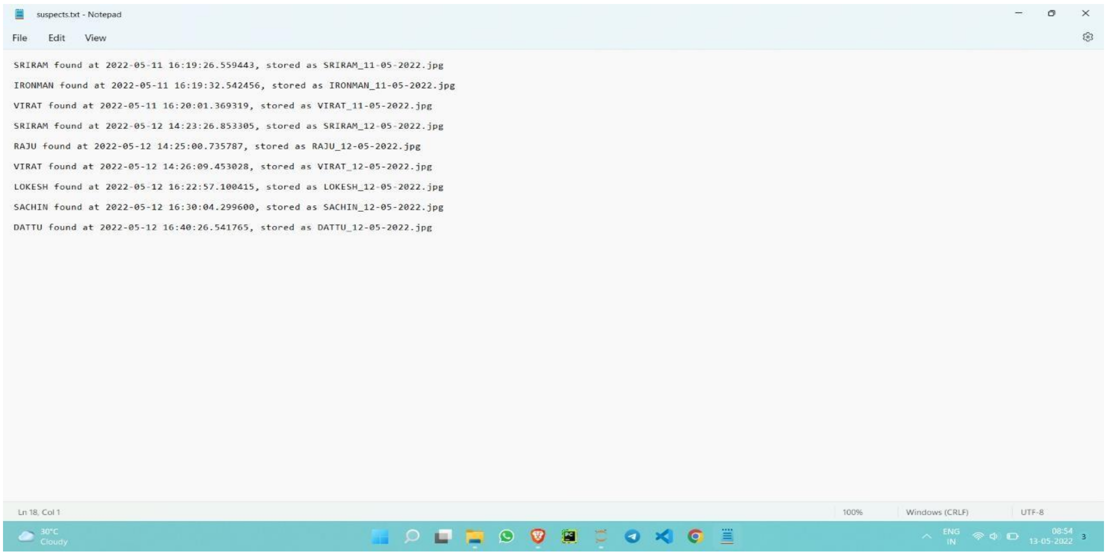


Fig. 5.4 The images from footage which matched with existing criminal record



suspects.txt - Notepad

```
SRIRAM found at 2022-05-11 16:19:26.559443, stored as SRIRAM_11-05-2022.jpg
IRONMAN found at 2022-05-11 16:19:32.542456, stored as IRONMAN_11-05-2022.jpg
VIRAT found at 2022-05-11 16:20:01.369319, stored as VIRAT_11-05-2022.jpg
SRIRAM found at 2022-05-12 14:23:26.853305, stored as SRIRAM_12-05-2022.jpg
RAJU found at 2022-05-12 14:25:00.735787, stored as RAJU_12-05-2022.jpg
VIRAT found at 2022-05-12 14:26:09.453028, stored as VIRAT_12-05-2022.jpg
LOKESH found at 2022-05-12 16:22:57.100415, stored as LOKESH_12-05-2022.jpg
SACHIN found at 2022-05-12 16:30:04.299600, stored as SACHIN_12-05-2022.jpg
DATTU found at 2022-05-12 16:40:26.541765, stored as DATTU_12-05-2022.jpg
```

Fig. 5.5 The Log maintained to find the time, name, and file name of a criminal detected

We have used 13 images as existing criminal record. The test is done through the live surveillance using various images for various test cases.

6.1 TEST CASES:

6.1.1 Image of an existing criminal:

In this test case, the images of people present in existing criminal record is shown in the live surveillance video. Since the image matches with existing criminal record, the image is stored along with other suspectedImages, a log is also updated with information related the criminal detected and time of recording, and an alert is raised successfully to notify the authorities. Fig. 6.1 show how the system shows result if a criminal is detected

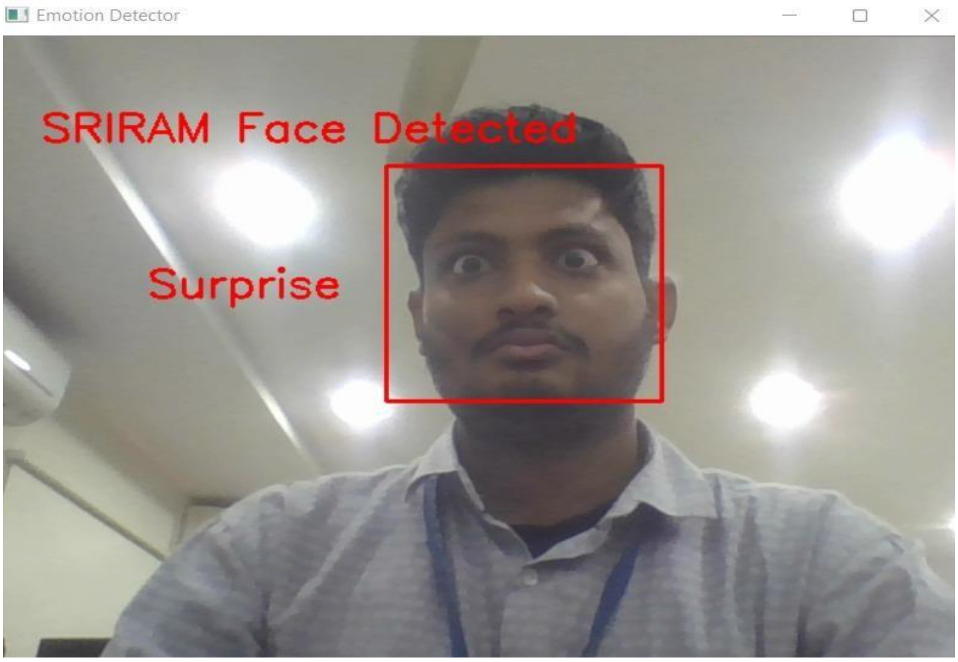


Fig. 6.1 Image of a criminal detected

6.1.2 Image of Unknown Person:

In this test case, the images of people that does not present in existing criminal record is shown in the live surveillance video with a suspicious expression. Since the image doesn't match with existing criminal record, no further alert is raised, but a message is displayed that an unknown person is detected as shown in the fig. 6.3.

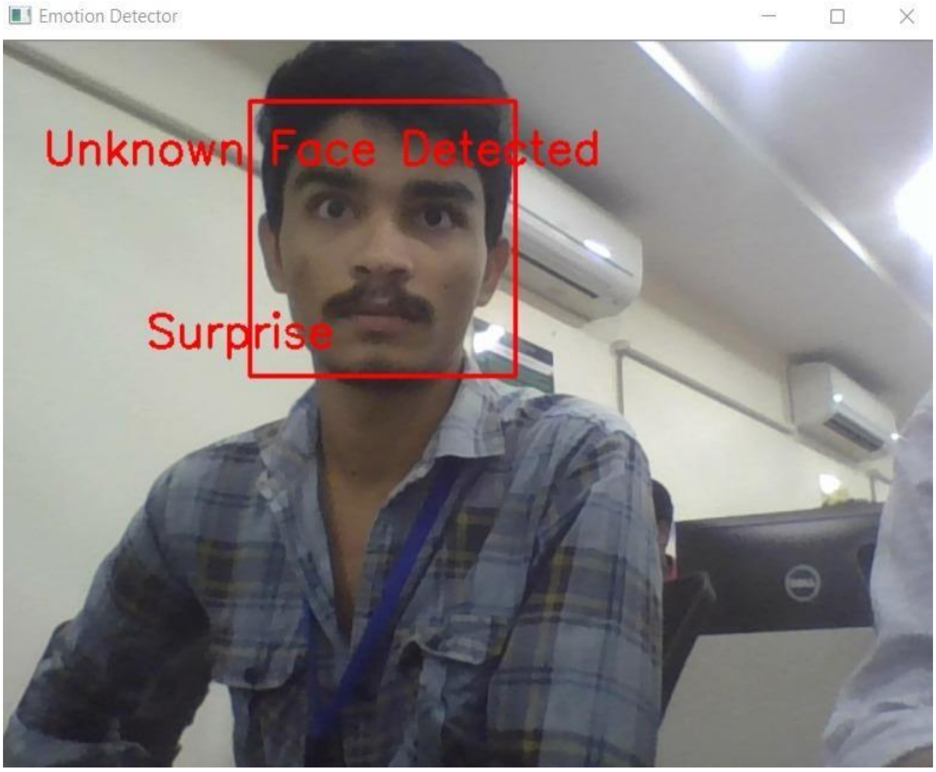


Fig. 6.2 An image of a non-criminal with a suspicious expression

6.1.3 Image with no faces:

In this test case, the images containing no people is given, since no face is detected, an message is displayed accordingly that no face is detected as shown in example figure 6.4

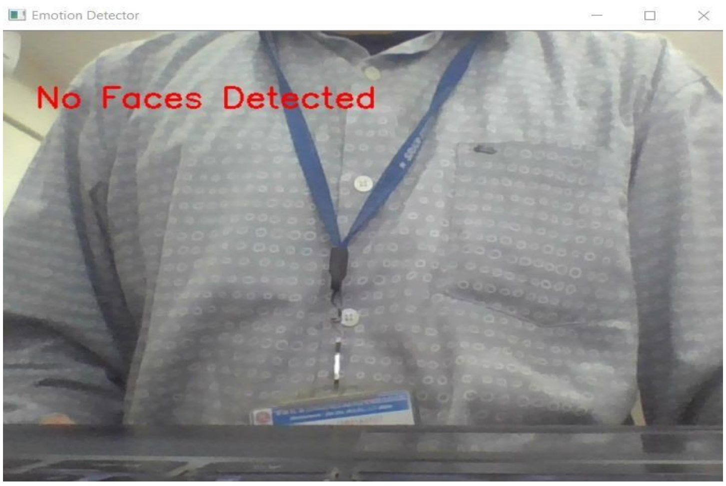


Fig. 6.3 shows the result in a test case where no face is detected

We have tested with images across the three test cases with 20 images of those 17 being predicted correctly under ideal lighting conditions, with an accuracy of 85%.

6.2 Face Expression analysis

The training and validation accuracy for facial expression recognition using CNN have been observed, as shown in fig 6.2.1.

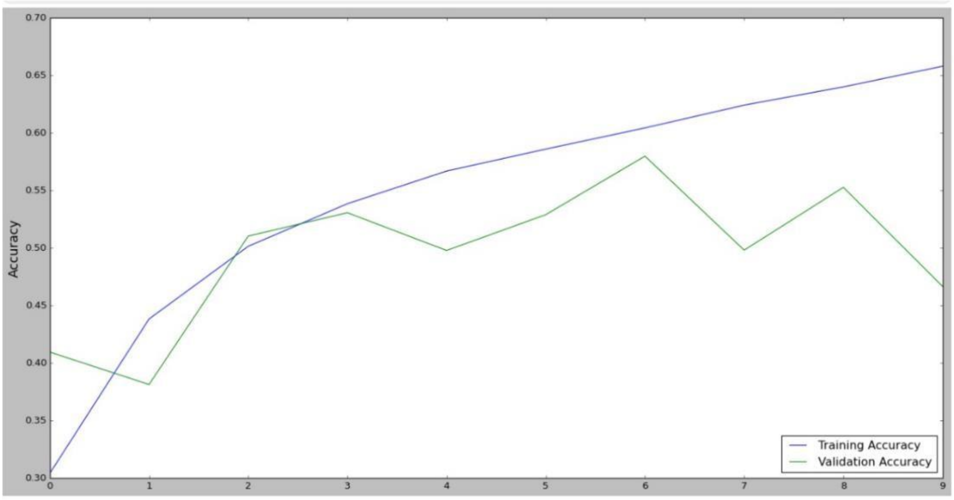


Fig 6.2.1 Training and validation accuracy

For facial expression recognition using CNNAccuracy for CNN up to 58% have been achieved.

7. CONCLUSION

This method involves using Live Face Detection, Face Expression analysis and Face Recognition in finding the criminals, using haarCaascade frontal face classifier for face detection, CNN for face expression analysis, and face_recognition package from python for face recognition. It takes the help of existing criminal records and searches for them in case of unusual cases. We have tested with images across the three test cases with 20 images of those 17 being predicted correctly under ideal lighting conditions, with an accuracy of 85%. This method works effieciently in certain light conditions and it has to be improved to achieve better results.

LIST OF REFERENCES

[1] https://en.wikipedia.org/wiki/Crimein_India#:~:text=Showing%20a%201.6%25%20annual%20increase, 2018%20to%20385.5%20in%202019.

[2] https://www.wikiwand.com/en/Crime_in_India

[3] https://thewire.in/rights/cctv-surveillance-is-rising-in-india-world-but-crime-rates-remain-unaffected#:~:text=Figure%201%20shows%20that%20around,surveillance%20cameras%20in%20the%20countr y.

[4] https://www.hindustantimes.com/cities/others/among-19-metro-cities-crime-rate-in-delhi-highest-in-2020-101631728619971.html

[5] https://www.numbeo.com/crime/rankings_by_country.jsp

[6] https://securitysurveillanceweb.wordpress.com/2017/10/03/benefits-of-city-surveillance-cameras/

[7] https://medium.com/nanonets/how-to-automate-surveillance-easily-with-deep-learning- 4eb4fa0cd68d

[8] https://www.pinterest.ca/pin/504543964481842145/

[9] Nurul Azma Adbullah, Md. Jamri Saidi, Nurul Hidayah Ab Rahman, Chuah Chai Wn and Isredza Rahmi A.Hamid, Face recognition for criminal identification: An implementation of principal component analysis for face recognition AIP Conference Proceedings 1891, 020002 (2017).

[10] Zhang, X. Gonnot, T. and Saniie, J. (2017) Real-Time Face Detection and Recognition in Complex Background. Journal of Signal and Information Processing, 8, 99-112.

[11] Avinash Kumar Singh, Neha Baranwal, G. C. Nandi. (2019) A rough set based reasoning approach for criminal identification. International Journal of Machine Learning and Cybernetics 10:3, pages 413-431.

[12] Ju-Yuan Hsiao, Shu-Ju Chuang & Po-Yueh Chen (2016) A hybrid face recognition system based on multiple facial features, International Journal of Computers and Applications, 38:1, 1-8

[13] S. T. Ratnaparkhi, A. Tandasi and S. Saraswat, "Face Detection and Recognition for Criminal Identification System," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 773-777

[14] Hasan, M.k.;Ashan, M.S.; Abdullah-Al-Mamun: Newaz, S.H.s.; Lee, G.M. Human Face Detection Techniques: A comprehensive Review and Future Research Directions. Electronics 2021,10,2354.

[15] Syed Inthiyaz and M Muzammil Parvez and M Siva Kumar and J Sri sai Srija and M Tarun Sai and V Amruth Vardhan, Facial Expression Recognition Using KERAS, 2021,02.J.Phys.: Conf. Ser: 1804 012202 .

[16] Guangyong Zheng, Yuming Xu, Efficient face detection and tracking in video sequences based on deep learning, Information Sciences, Volume 568, 2021, Pages 265-285, ISSN 0020-0255.

[17] A. S. Jalal, D. K. Sharma and B. Sikander, "Facial Mole Detection Approach for Suspect Face Identification using ResNeXt-50," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-6.

[18] Najibi, Mahyar and Singh, Bharat and Davis, Larry S. “FA-RPN: Floating Region Proposals for Face

Detection", arXiv, 2018.

[19] S. Shirsat, A. Naik, D. Tamse, J. Yadav, P. Shetgaonkar and S. Aswale, "Proposed System for Criminal Detection and Recognition on CCTV Data Using Cloud and Machine Learning," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 2019, pp. 1-6.

[20] K. M. M. Thu and N. N. M. Thein, "Edge based Crime Assistance System with Cloud Computing," 2019 International Conference on Advanced Information Technologies (ICAIT), 2019, pp. 43-48.

[21] Kar, N.B., Babu, K.S., Sangaiah, A.K. et al. Face expression recognition system based on ripplet transform type II and least square SVM. Multimed Tools Appl 78, 4789–4812 (2019).

[22] Talegaonkar, Isha and Joshi, Kalyani and Valunj, Shreya and Kohok, Rucha and Kulkarni, Anagha, Real Time Facial Expression Recognition using Deep Learning (May 18, 2019). Proceedings of International Conference on Communication and Information Processing (ICCIP) 2019.

[23] M. Georgescu, R. T. Ionescu and M. Popescu, "Local Learning With Deep and Handcrafted Features for Facial Expression Recognition," in IEEE Access, vol. 7, pp. 64827-64836, 2019.

[24] Ayed MB, Elkosantini S, Alshaya SA, Abid M. Suspicious behavior recognition based on face features. IEEE Access. 2019 Oct 14;7:149952-8.

[25] Zhengning Wang, Fanwei Zeng, Shuaicheng Liu, Bing Zeng, OAENet: Oriented attention ensemble for accurate facial expression recognition, Pattern Recognition, Volume 112, 2021.

[26] Jain, R., Nayyar, A., Bachhety, S. (2020). Factex: A Practical Approach to Crime Detection. In: Sharma, N., Chakrabarti, A., Balas, V. (eds) Data Management, Analytics, and Innovation. Advances in Intelligent Systems and Computing, vol 1042. Springer, Singapore.

[27] Muhammad Sajjad, Mansoor Nasir, Khan Muhammad, Siraj Khan, Zahoor Jan, Arun Kumar Sangaiah, Mohamed Elhoseny, Sung Wook Baik, Raspberry Pi assisted face recognition framework for enhanced law- enforcement services in smart cities, Future Generation Computer Systems, Volume 108, 2020.

[28] Kiran Amjad, Prof Dr. Aftab Ahmad Malik, Dr. Sumet Mehta, A technique and architectural design for criminal detection based on Lombroso Theory using deep learning, Vol.4 No. 3 (2020).

[29] Salim Afra, Reda Alhajj, Early warning system: From face recognition by surveillance cameras to social media analysis to detecting suspicious people, Physica A: Statistical Mechanics and its Applications, Volume 540, 2020.

[30] Parekh Payal & Mahesh M. Goyani (2020) A comprehensive study on face recognition: methods and challenges, The Imaging Science Journal, 68:2, 114-127.

[31] G. S. Nagpal, G. Singh, J. Singh and N. Yadav, "Facial Detection and Recognition using OpenCV on Raspberry Pi Zero," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2018.

[32] A. A. Shah, Z. A. Zaidi, B. S. Chowdhry and J. Daudpoto, "Real time face detection/monitor using raspberry pi and MATLAB," 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT), 2016, pp. 1-4

[33] K. Rasanayagam, S. D. D. C. Kumarasiri, W. A. D. D. Tharuka, N. T. Samaranayake, P. Samarasinghe and S. E. R. Siriwardana, "CIS: An Automated Criminal Identification System," 2018 IEEE International Conference on Information and Automation for Sustainability (ICIAfS), 2018, pp.1-6

[34] W. Yimyam, T. Pinthong, N. Chumuang and M. Ketcham, "Face Detection Criminals through CCTV Cameras," 2018 14th International Conference on Signal-Image Technology & Internet- Based Systems

(SITIS), 2018, pp. 351-357

[35] X. Chen, L. Qing, X. He, J. Su and Y. Peng, "From Eyes to Face Synthesis: a New Approach for Human- Centered Smart Surveillance," in IEEE Access, vol. 6, pp. 14567-14575, 2018

[36] Fan, Y., Lam, J.C.K., Li, V.O.K. (2018). Multi-region Ensemble Convolutional Neural Network for Facial Expression Recognition. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds) Artificial Neural Networks and Machine Learning – ICANN 2018. ICANN 2018. Lecture Notes in Computer Science(), vol 11139. Springer, Cham.

[37] T. Mantoro, M. A. Ayu and Suhendi, "Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods," 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), 2018, pp. 1-5

[38] B. F. Klare, S. Klum, J. C. Klontz, E. Taborsky, T. Akgul and A. K. Jain, "Suspect identification based on descriptive facial attributes," IEEE International Joint Conference on Biometrics, 2014, pp. 1-8

[39] https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset

APPENDIX 1: SAMPLE CODE

```
import face_recognition as fr
import cv2
import os
path="train/"
known_encodings=[]
names=[]
for i in os.listdir(path):
    img=fr.load_image_file(path+i)
    img=cv2.cvtColor(img,cv2.COLOR_BGR2RGB)
    enc=fr.face_encodings(img)[0]
    known_encodings.append(enc)
    names.append(i.split('.')[0])
criminalMap = {}
from datetime import datetime
import os


def issuspect(img,known_encodings,names):
    cv2.imshow('x',img)
    imgToStore = img
    img=cv2.cvtColor(img,cv2.COLOR_BGR2RGB)
    fl=fr.face_locations(img)
    if len(fl)!=0:
        encodings=fr.face_encodings(img,fl)
#        print('1')
        for e,f in zip(encodings,fl):
            matches=fr.compare_faces(known_encodings,e)
            facedis=fr.face_distance(known_encodings,e)
            matchindex=numpy.argmin(facedis)
#            print('2')
            if matches[matchindex]:
#                print('3')
                name=names[matchindex].upper()
                global criminalMap
                x = datetime.now()
                today = x.strftime("%d-%m-%Y")
                if name:
                    cv2.putText(video_frame,f'{name} Face
Detected',(30,80),cv2.FONT_HERSHEY_SIMPLEX,1,(0,0,255),2)

                    print(name)
                    print(criminalMap)
                    newImgName = f'suspectImages/{name}_{today}.jpg'
                    if today in criminalMap:
                        if name not in criminalMap[today]:
                            #img create
                            cv2.imwrite(newImgName, imgToStore)
                            # for playing note.mp3 file
                            os.system("danger.mp3")
```

```
                    print("danger")
                    with open('suspects.txt','a') as f:
                        msg = f'{name} found at {str(datetime.now())}, stored as {name}_{today}.jpg \n\n'
                        f.write(msg)
                        f.close()
                criminalMap[today].add(name)
            else:
                #img create
                cv2.imwrite(newImgName, imgToStore)
                print("danger")
                os.system("danger.mp3")

                criminalMap[today] = {name}
                with open('suspects.txt','a') as f:
                    msg = f'{name} found at {str(datetime.now())}, stored as {name}_{today}.jpg \n\n'
                    f.write(msg)
                    f.close()
    else:
        cv2.putText(video_frame,'Unknown Face
Detected',(30,80),cv2.FONT_HERSHEY_SIMPLEX,1,(0,0,255),2)
#           print('-',end= ' ')
```

```
from keras.models import load_model
from keras.preprocessing.image import img_to_array
from keras.preprocessing import image
import cv2 as cv
import numpy as np

face_classifier = cv.CascadeClassifier(r'haarcascade_frontalface_default.xml')
classifier =load_model(r'model.h5')

emotion_labels = ['Angry','Disgust','Fear','Happy','Neutral', 'Sad', 'Surprise']

video_capture = cv.VideoCapture(0)
while True:
    _, video_frame = video_capture.read()

    grey = cv.cvtColor(video_frame,cv.COLOR_BGR2GRAY)
    faces = face_classifier.detectMultiScale(grey)
    if len(faces)==0:
        cv.putText(video_frame,'Identified NO Faces',(30,80),cv.FONT_HERSHEY_SIMPLEX,1,(0,0,255),2)
    for (x,y,w,h) in faces:
        region_grey = grey[y:y+h,x:x+w]
        dt=video_frame[y:y+h,x:x+w]
        region_grey = cv.resize(region_grey,(48,48),interpolation=cv.INTER_AREA)
        if np.sum([region_grey])!=0:
            reg = region_grey.astype('float')/255.0
            reg = img_to_array(reg)
            reg = np.expand_dims(reg,axis=0)

            expression_predicted = classifier.predict(reg)[0]
```

```
            predicted_label_emotion=emotion_labels[expression_predicted.argmax()]
            if predicted_label_emotion=='Surprise' or predicted_label_emotion== 'Fear':

cv.putText(video_frame,f'{predicted_label_emotion}',(100,200),cv.FONT_HERSHEY_SIMPLEX,1,(0,0,255),2)
                cv.rectangle(video_frame,(x,y),(x+w,y+h),(0,0,255),2)
                issuspect(dt,known_encodings,names)
        else:
            cv.putText(video_frame,'Identified NO Faces',(30,80),cv.FONT_HERSHEY_SIMPLEX,1,(0,255,0),2)
    cv.imshow('Criminal Detection System',video_frame)
    if cv.waitKey(1) & 0xFF == ord('e'):
        break

video_capture.release()
cv.destroyAllWindows()
```