VIVA-Tech International Journal for Research and Innovation *Volume 1, Issue 4 (2021)*
ISSN (Online): 2581-7280 *Article No. X*
PP XX-XX

VIVA Institute of Technology
9thNational Conference on Role of Engineers in Nation Building – 2021 (NCRENB-2021)

VIVA-TECH IJRI
ANNUAL RESEARCH JOURNAL

VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION
ANNUAL RESEARCH JOURNAL
ISSN(ONLINE): 2581-7280

CRIMINAL IDENTIFICATION FOR LOW RESOLUTION SURVEILLANCE

Saniya Prashant Patil[1], Grishma Sunil Yadav[2], Shrutika Devdas Kudalkar[3], Sunita Naik[4]

[1](Department of Computer Engineering, Mumbai University, India)
Email: patilsaniya11@gmail.com
[2](Department of Computer Engineering, Mumbai University, India)
Email: grish30799@gmail.com
[3](Department of Computer Engineering, Mumbai University, India)
Email: 17305076shrutika@viva-technology.org
[4](Department of Computer Engineering, Mumbai University, India)
Email: sunitanaik@viva-technology.org

Abstract: *A Criminal Identification System allows the user to identify a certain criminal based on their biometrics. With advancements in security technology, CCTV cameras have been installed in many public and private areas to provide surveillance activities. The CCTV footage becomes crucial for understanding of the criminal activities that take place and to detect suspects. Additionallywhen a criminal is found it is difficult to locate and track him with just his image if he is on the run. Currently this procedure consists of finding such people in CCTV surveillance footage manually which is time consuming. It is also a tedious process as the resolution for such CCTV cameras is quite low. As a solution to these issues, the proposed system is developed to go through real time surveillance footage, detect and recognize the criminals based on reference datasets of criminals. The use of facial recognition for identifying criminals proves to bebeneficial. Once the best match is found the real time cropped image of the recognized criminal is saved which can be accessed by authorized officials for locating and tracking criminals or for further investigative use.*

Keywords -*Criminal Identification System, Detection, Face Recognition System, Facial Recognition*

I. INTRODUCTION

For security reasons CCTV cameras are installed in public places. This CCTV footage becomes an important piece of evidence for criminal cases. Thus it is used to identify suspects or criminals on a crime scene. But it is difficult to track them using just their image.

According to cyber cell members of Goa branch, they make multiple members of the department sit with laptops and computers to search through the CCTV footage to find and trace the culprit due to the lack of automated system for doing this task with them. This process is both time and labor intensive. These resources and time can be utilized more efficiently if an automated system is used. Moreover, the law enforcement agencies won't have to depend on other people to spot the criminal at a particular location and then inform them about it. Also, the resolution of surveillance footage is quite low which makes it difficult to identify a person.

The criminal detection system will ease this process. It will use facial recognition to identify criminals in low resolution frames. When the person will be spotted in the surveillance footage, the details of the criminal and his or her location details will be recorded in a folder. This folder can be accessed to locate and track any criminal when required. This will save time and manpower required to manually go through the surveillance footage. It can be used to get real-time location of wanted criminals on the run, missing persons etc.

The system will be developed using the Tiny Face Detector deep learning model. The output of the system will be saved in a folder and accessed on the web portal developed on Django framework.

II. RELATED WORKS

Apoorva.P, et. al. [1] have proposed a system for real-time face recognition using automated surveillance cameras. The system comprises of three datasets including data of citizens, data of criminals who are citizens of the country and, data of criminals who are not the citizens of the country. The system first detects a face in a particular frame. Then the face undergoes feature extraction which is done using Haar Cascade Classifier and OpenCV. This processed image is compared with the datasets to identify the person. The system proposes two approaches. The first is using multiple classifiers at a time and training and recognizing the faces with various rotation angles. The second approach is to use single classifier to recognize faces with multiple rotation angles but this process takes longer time. If a match is found from the criminal datasets then the time from the surveillance is noted.

Vishakha K, et. al. [2] have proposed a system for detection of human beings in a scene, and tracking those people as long as they stay in the scene by identifying individual persons. Functions of Haar classifier were used for human detection. Both positive as well as negative samples were used to train the haar classifier. It detects humans from the input frames using as many samples as possible to ensure better probability of identification. After detection the features are extracted for every distinct human being in a frame. Tracking of human beings is done using sampling and resampling algorithms. The next position where that human can be found is predicted with accuracy. Positions are represented visually using unique colored rectangles enclosing the human being.

Han Xia, et. al. [3] proposed system on Face Recognition and Application of Film and Television Actors Based on Dlib. In this paper, a face recognition model applied to film and television is proposed. Firstly, for a specific scene, they used the image information of all faces in the scene as input to calculate the face feature vector and as training data, then train the KNN classification model in the scene, and then can recognize the face in the scene. In the system, instead of building a face recognition network like Dlib from zero, they used the face recognition network model that Dlib has trained. The system detects the actor's face, extracts feature points, calculates face feature vectors, and then obtains the actor's identity information. Based on Dlib, this paper designed and implemented a film actor face recognition system, and designed experiments to verify the effectiveness of the face recognition system. The system has a high recognition rate for relatively clear faces, but the face recognition effect with occlusion needs to be improved.The model can recognize most faces accurately, but when the face is covered by sunglasses, hair or other occlusions, or when the face tilts to a certain angle, the recognition rate will decrease. In addition, there will also be recognition errors.

SatyaSathvik Kadambari, et. al. [4] have proposed an automated attendance system using facial recognition. The system consists of two phases, face detection and face recognition and records the attendance of recognized faces in excel. Face detection was implemented using Viola-Jones algorithm. Haar features are used along with Adaboost to describe a face with 6 sufficient accuracy and in less time. The face recognition phase was implemented using Deep Neural Networks (DNN) with OpenCV. The dlib library consisting CNN was used to detect a face from all the suitable angles. The face_recognition library was used for face recognition. Since Viola Jones algorithm was inadequate for face recognition, the dlib library was used. The accuracy of the system was 85% to 95%. For images with poor lighting conditions the accuracy was 70% to 80%.

Lirie Koraqi, et. al. [5] proposed a framework that detects fast moving object in an image or a video. The objective of the system was to detect persons in camera frame, construct their bounding box, localize their position in camera frame, tracking the person's movements and visualizing all these steps in GUI. The GUI of the system was built using Qt5 library and Qt creator IDE. The proposed system is an autonomous tracking system. The hardware requirements were a processor (i.e. a laptop with linux OS in this case), a microcontroller developed on Arduino platform and a physically powered camera. The communication between the processor and microcontroller for camera controls was done using MOBUS communication protocol. The frame data and support part of the necessary mathematical operator was manipulated using the OpenCV library. Person detection was done using TensorFlow as main attribute. The Haar Cascade algorithm is used to detect object specific features, face detection in this case. The essence of this paper is that it does not attempt to develop a system that will be able to classify all possible objects, but develop a framework that enables the user to classify as accurately as possible a rapid set of objects in the case of proposed system of detecting people and tracking them. Tensorflow was used instead of haar cascade algorithm because the web camera did not have good resolution while haar was used only for face detection. The decision to move the camera is made only by the person detection algorithm.

Ratna Yustiawati, et. al. [6] proposed a research on analysing of different features using haar cascade classifier. The research is on security of a room. Here the room can only be opened by specific face which already exists in the database. The technology used is Haar cascade classifier. The center of focus is to detect the

people who using hijab and spectacles is been applied in this research. Haar-like features image processing is used, where it uses 10 x10 pixel dimensional image. Implementation and integration for this research was done in measurement laboratory of PoliteknikNegeriSriwijaya. The resulting average confidence value between the object with hijab and without hijab here is 80-100. The smaller value of confidence the better and more accurate the feature would be.

B. Maga, et. al. [7] have proposed an efficient approach for object detection and tracking. Here a new proposed approach is provided for efficient object tracking using Kernel and feature based tracking methods. It is approach where vehicle classification performance can be done in surveillance videos. Here this approach requires shape and appearance of the object. Object basically contains various features and any of them is used to track object as kernel. Object tracking can be done easily if compute the motion of the kernel of the between more than two frames. Here the process is divided into two processes training and testing of objects in videos. First process is a trained image or frame in videos and trained object value based on shape and moving position with vehicle positive and negative results. It's store one database for testing video surveillance object values. Second process is extracted image in video after capture object value then tested in database object value, if object values are matched because result is positive then object tracked in given surveillance videos. Object matching processing use to template matching technique. Object detecting and tracking has a wide variety of applications in computer vision video surveillance, its gives more accuracy than other existing method like CNN, template matching. But it fails in multiple object tracking, learning and detection in term of more computing time required.

Igor Lashkov, et. al. [8] proposed system on driver dangerous state detection based on OpenCV & Dlib libraries using mobile video processing. The research is on real-time driving behaviour monitoring system in intelligent transportation systems. This system based on Vision-based approach including video cameras for dangerous situation detection is undoubtedly one of the most perspective and commonly used in sensing driver environment. In this case, the images of a driver, captured with video cameras, can describe its facial features, like head movements, eye state, mouth state, and, afterwards, identify a current level of fatigue state. In this paper, author leverage the built-in front-facing camera of the smart phone to continuously track driving facial features and early recognize driver's drowsiness and distraction dangerous states. Dlib presents one of the most popular and frequently used libraries for facial landmarks extraction using the trained model, the library allows to obtain the 68 facial landmarks, including the information about eyes, nose, lips and mouth, including its size and location. According to the received results of the study the developed framework for smart phones demonstrates 93% drowsiness state classification.

III. METHODOLOGY

The proposed model uses Tiny Face Detector model for face recognition. It is a prominent real-time face detector. It is a mobile and web friendly model. The quantized model i.e. the tiny_face_detector_model has the size of only 190 KB. It detects faces and facial landmarks on images or frames. It takes a person's face as input and gives a vector called embedding of 128 numbers as output. The vector represents the most important features of the face. Then the images are classified by comparing with the feature from the dataset of criminals. The web portal is developed using Django framework which allows the admin to view the results.

First the frames are fed to the convolution neural network where the faces or facial landmarks are detected in them. After the face is detected, the features are extracted. It takes a person's face as input and gives a vector called embedding of 128 numbers as output. The embeddings represent the most important features of the face and are similar for similar faces. Essentially, embeddings are low-dimensional representations of the high-dimensional data viz. images or frames. It identifies key features of a person's face which separate it from different faces. Then the features are compared with the features extracted from the images in the dataset to find a match and thus the face is recognized.

The fig.1 depicts the system flow diagram of the proposed system. The dataset comprises of photos of various criminals which are used to train the model of the proposed system. After training, the user can test the system by giving input of live surveillance video feed to identify any criminal. After the video dataset is fed to the system, it is divided into multiple frames and they are fed to the network for face detection. It is done using Tiny Face Detector Model. The network returns the bounding boxes of each face from the face landmark positions. The extracted and aligned face images are fed into the network Later the frames undergo feature extraction of the detected faces. It gives a vector called as embedding. These embeddings comprise of the distinctive features of the face. The said embeddings are compared to the features of the images from the dataset i.e., the reference data to find a match. It is judged using a threshold value, 0.6 in this case. The recognized

images are saved in PNG format in a folder which can be accessed through the portal. The image file will specify the name, date and time and the location for identification.

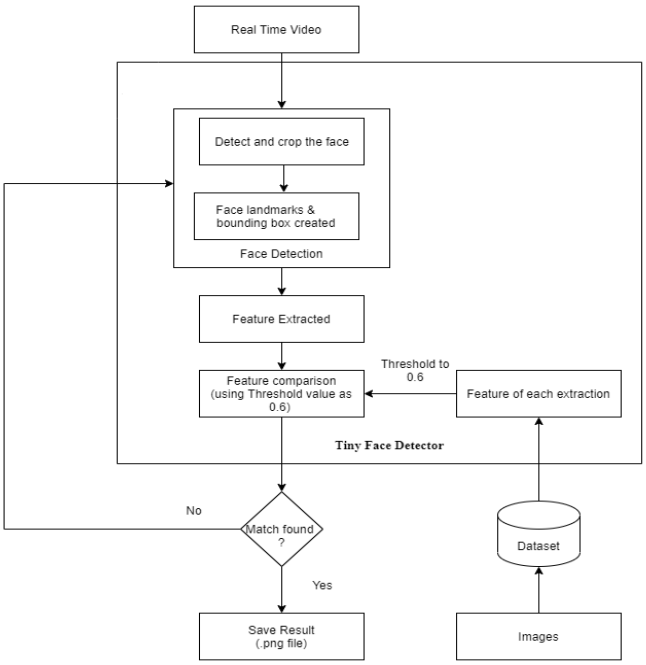

Figure 1: System Flow Diagram

IV. RESULT

The system is trained using 200 images of 3 people each. The web portal can be used to view the details of the criminals in the database. The system consists of tiny_face_detector_model. It returns the bounding boxes of each face from the face landmark positions. The features are extracted as a vector called embedding of 128 numbers. Then the features are compared with the features extracted from the images in the dataset to find a match and thus the face is recognized. The Euclidian distance is used to compare the features of the two faces with a threshold value of 0.6. This value is ideal for images with resolution 150x150 pixels. The model detects faces for images with lowest resolution of 200x200 pixels.

The following figures 2 and 3 are screenshots of live video feed in the portal. The model used is tiny_face_detector that correctly recognizes the face. Here age and expression of face is also recognized which is done by using age_expression model.

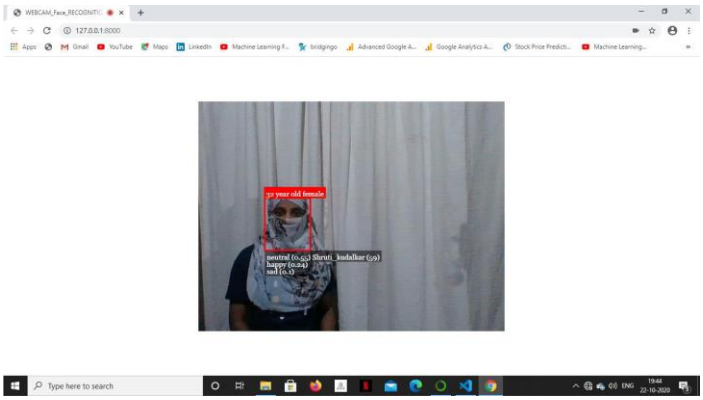

Figure 2: Result of face recognition of face wearing a hijab

Figure 2 is a screenshot of live video feed in the portal. The model used correctly recognizes the face wearing a hijab as it is trained with the dataset including labels with masked and unmasked faces. The model is able to predict the label of the face wearing a hijab with Euclidean distance of 59 i.e. 0.59.

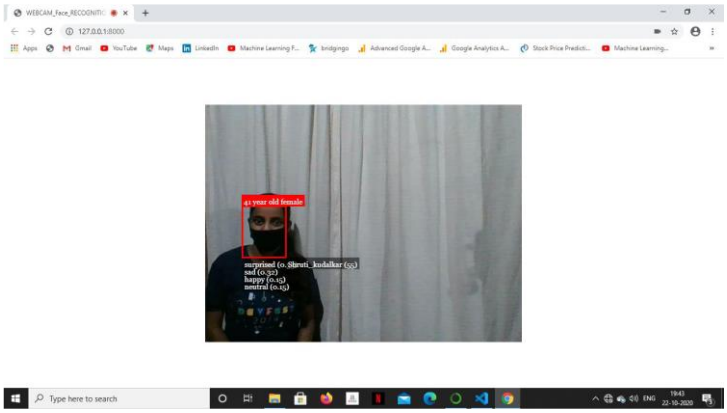

Figure 3: Result of Face Recognition of Face Wearing Mask

Figure 3 is a screenshot of live video feed in the portal. The tiny_face_detector model correctly recognizes the face and based on the dataset provided the model is trained to recognize the faces their label. The dataset also includes labels with masked and unmasked faces which is why it is able to predict face wearing a mask with euclidean distance of 55 i.e. 0.55.

Table 1: Result analysis table

Parameters	**Uncovered Face**	**Face Wearing Mask**	**Face Wearing Scarf**	**Face Wearing Spectacles**	**Left Sided**	**Right Sided**
Close In Light	31	48	59	42	42	49
Close In Dark	43	45	60	45	50	54
Far In Light	50	57	Only detected	49	Only detected	Only detected
Far In Dark	50	58	Not accurate	57	Not accurate	Not accurate

*ACCURACY COMPARED ON FOLLOWING CONDITIONS

Very Accurate	1-20
Accurate	21-40
Less Accurate	41-60
Unknown	61-100

Table 1 is a result analysis of the proposed model where above mentioned conditions are considered where it gives a accurate results for the covered and uncovered faces in day as well as in dark light but in certain scenario where face is covered and person is standing at far in light place it is good in only detecting faces but not in recognizing but when scenario changes in far in dark result drop and does not give good accuracy for detecting.

V. CONCLUSION

The proposed system is a deep learning approach to detect criminals in public places from live surveillance footage. The system does facial recognition for uncovered faces as well as faces wearing hijabs, masks and scarves. The need of developing such a system is the increasing crime rate at public places and lack of security. The system will ease the investigation process of the crimes. The output of the proposed system can be used to detect if any criminals were present at any particular location such as a crime scene. This system will be better than the current system that use OpenCV as they lag while real-time streaming. It can also be used to find missing persons by facial recognition by adding those particular images to the dataset.

Acknowledgements

We would like to express a deep sense of gratitude towards our guide Prof.Sunita Naik, Computer Engineering Department for her constant encouragement and valuable suggestions. The work that we are able to present is possible because of her timely guidance. We would like to pay gratitude to the panel of examiners Prof. Sunita Naik, Prof. Dnyaneshwar Bhabad, & Prof. Vinit Raut for their time, effort they put to evaluate our work and their valuable suggestions time to time. We would like to thank Project Head of the Computer Engineering Department, Prof. Janhavi Sangoi for her support and co-ordination. We would like to thank the Head of the Computer Engineering Department, Prof. Ashwini Save for her support and co-ordination. We are also grateful to the teaching and the non-teaching staff of Computer Engineering Department who lend their helping hands in providing continuous support.

REFERENCES

[1] Apoorva.P, Impana.H.C, Siri.S.L, Varshitha.M.R and Ramesh.B, "Automated Criminal Identification by Face Recognition Using Open Computer Vision Classifiers", *Third International Conference on Computing Methodologies and Communication (ICCMC 2019), IEEE, 2019, CFP19K25-ART; ISBN: 978-1-5386-7808-4*

[2] Visakha K and Sidharth S Prakash, "Detection and Tracking of Human Beings in a Video using Haar Classifier", *International Conference on Inventive Research in Computing Applications (ICIRCA 2018), IEEE, 2018, CFP18N67-ART; ISBN:978-1-5386-2456-2*

[3] Han Xia and Chunfang Li, "Face Recognition and Application of Film and Television Actors Based on Dlib", *12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, 2019, 978-1-7281-4852-6*

[4] SatyaSathvik Kadambari, Deep Mistry, Gauraang Prabhu and Prof.Monica Khanore, "Automation of Attendance System Using Facial Recognition", *International Conference on Advances in Computing, Communication and Control (ICAC3), IEEE, 2019*

[5] Lirie Koraqi and Florim Idrizi, "Detection, identification and tracking of objects during the motion", *IEEE, 2019, 978-1-7281-3789-6/19*

[6] Ratna Yustiawati, Nyayu Latifah Husni, Evelina, Sabilal Rasyad, Iskandar Lutfi, Ade Silvia, Niksen Alfarizal, Adella Rialita, "Analyzing Of Different Features Using Haar Cascade Classifier", *International Conference on Electrical Engineering and Computer Science (ICEECS), IEEE, 2018, 978-1-5386-5721-8/18*

[7] B.Maga and Mr. K.Jayasakthi Velmurgan, "An Efficient Approach for Object Detection and Tracking", *Third International Conference on Science Technology Engineering & Management (ICONSTEM), IEEE, 2017, 978-1-5090-4855-7/17*

[8] igor Lashkov, Alexey Kashevnik, Vladimir Parfenov and Anton Shabaev, "Driver Dangerous State Detection Based on OpenCV & Dlib Libraries Using Mobile Video Processing", *International Conference on Computational Science and Engineering (CSE) and International Conference on Embedde and Ubiquitous Computing (EUC), IEEE, 2019, 978-1-7281-1664-8*