Face Tracker: A Multi Camera Oriented Surveillance Approach

Shivank Sharma[1], Pritam khot[2], Omkar Patil[3]
[123](Students, Computer Science, Pdea College of Engineering MANJARI (BK), Pune, Maharashtra.)

Prof. Raghunath M Kawale[4]
[4](Professor, Computer Science, Pdea College of Engineering MANJARI (BK), Pune, Maharashtra.)

Abstract: ***Nowadays the crime rates in society are increasing, it is hard for the police department to give complete surveillance due to lack of manpower.***

This paper presents the need for facial recognition technology, as part of video surveillance systems. Our system performs face recognition by using the HAAR classifier, extracts the features using the CNN algorithm. The face detection done by the system is more robust, as we are using multiple cameras that capture all possible direction of the human's face. This system gives effective assistance to the police department by using a technique like a face recognition and artificial intelligence in reducing the crime rate of the city and assists in surveillance in real-time.

Keywords: Video Surveillance, Face Detection, HAAR Classifier, Convolution Neural Network (CNN), Preprocessing, Feature Extraction, Tracking of Criminals, Identification of the Missing Person.

I. INTRODUCTION

In a country like India where the population is increasing day-by-day, where, controlling the crimes has become a tedious job. Such existing of crime, violence, and fear in cities challenges to law enforcement authorities. Public safety is emerged as an important factor, due to the growing crime rates, increasing terrorism, missing children/runaway cases. Therefore there has to be a direct application in any urban safety strategy, aimed at reducing and preventing problems of crime and insecurity. There are many terrorists, criminals, thieves who have been staying in the country from years fearlessly due to the poor surveillance system.

The video Surveillance system is a widely used system to increase security in public places like governments' agencies, businesses, schools, Hospitals, Airports, accident detection, traffic monitoring and controlling [2][3]. Lots of investigation has been done before for automatic surveillance systems to support monitoring indoor and outdoor environments.
But we have to think before believing that the CCTV can be efficiently used to prevent terrorist attacks. Committing acts of terrorism are usually unknown before the act takes place, and it is very difficult to conduct surveillance of someone whose identity is unknown.

The effectiveness of CCTV surveillance can be improved by using some biometric characters. Face recognition in surveillance environments is essential to recognize possible terrorists and criminals on a watch list. Even though the performance of face recognition has improved significantly in the last decade. But the existing most of the face detection system is based on single view object detection. Such a single view object detection methods may affect the information loss that occurs during image formation. The obstruction of facial features, such as skin color, pose variation, illumination, expression, aging raises fundamental challenges in designing robust face recognition algorithms. A recent report by the Metropolitan Police in London revealed that the city's one million-plus cameras have helped to solve only a handful of crimes [1].
So there is a need for a multi-camera surveillance system that can detect the person face in all possible directions to increase the face detection efficiency. The coverage of multi-cameras surveillance is defined by the cameras' field of view (FOV) which is either overlap or non-overlap depending on the direction of the camera is installed [6].

The proposed system gives effective assistance to the police department by using a technique like a face recognition and artificial intelligence in reducing the crime rate of the city and assists in surveillance in real-time. We are implementing a multi-viewpoint human tracking system based on face detection using Haar-like features and convolution neural network (CNN).

II. LITERATURE REVIEW

Michael Charles Davis et al. [1] proposed a system that can perform detection and recognition rapidly in real-time. The system detects faces using the Viola-Jones face detector and then extracts local features to build a shape-based feature vector. The feature vector is constructed from ratios of lengths and differences in tangents of angles, to be robust to changes in scale and rotations in-plane and out-of-plane. Matching is performed using a Support Vector Machine (SVM).

Le An et al.[4] proposes a proba-bilistic approach for face recognition suitable for a multi-camera video surveillance network. A Dynamic Bayesian Network (DBN) is used to integrate the data getting from different cameras as well as the temporal clues from successive frames. This method is tested on a public surveillance video dataset. The paper concludes that by dynamically modeling the face the recognition performance in a multi-camera network can be improved.

The author of [5] proposes a robust feature for multi-view recognition that is insensitive to pose variations. A robust feature by exploring the fact that the subspaces panned by Spherical Harmonics are irreducible representations for the SO (3) group is presented. A multi-view video tracking algorithm is presented to mechanize the feature acquisition in a camera network setting.

An extensive survey of an active-vision ace Recognition at a Distance (FRAD) system we call the Biometric Surveillance System. is given in [7]. Bellotto et al. [8] describe an architecture for face capture with an active multi-camera surveillance system. Person trackers are associated with each Wide Field Of View (WFOV) camera, and high-level reasoning algorithms share data via an SQL database. Person detections from WFOV trackers are used to automatically assign and target Narrow Field Of View (NFOV) cameras to particular subjects. Using the NFOV cameras, a face is tracked and follows the face with a velocity control system.

Josh Harguess et al.[9] propose a method to track the subject using a cylinder head model (CHM) in multiple cameras. The CHM is a cylinder that has is a close approximation of a 3D head, allowing for correct tracking of the face. Knowing the pose from the CHM allows us to produce a frontal view of the face even though some information from the face might be missing due to self-occlusion. Eigenfaces are used for face recognition along with the average-half-face to reduce the effect of transformation errors.

Wei Niu et al.[10] presents a real-time multi-person tracking system. By using adaptive background subtraction, foreground regions are first recognized and segmented. A clustering algorithm is then used to collect the foreground pixels in an unsupervised manner to evaluate the image location of individual persons. A Kalman filter is used to keep track of each person and a unique label is assigned to each tracked individual. By using this method we can find the people enter and leave the scene at random.

Mahdi, Fahad Parvez, et al.[11] proposed a real-time system for surveillance using cameras. Initially, the system tracks and selects the faces of the detected persons. Face detection is done by using an efficient recognition algorithm with a known database. The Viola-Jones algorithm is used for face detection, the Kanade-Lucas-Tomasi algorithm is used as a feature tracker and face recognition is done with Principal Component Analysis (PCA). This system can be implemented in different restricted areas, such as at the office or house to detect the doubtful person at the entrance of a sensitive installation. The system works practically under reasonable lighting conditions and image depths.

III. PROPOSED SYSTEM

This system gives an effective assistance to police department by using technique like face recognition and artificial intelligence in reducing crime rate of city and gives assistance in surveillance in real time. The proposed system can detect the face by the multi-camera to cover all the possible variations of the people's move to capture the face. With this variation, we can detect the face with higher accuracy. With the use of this system, we can track criminal activities, helps in the identification of a missing person, etc. Figure 1 shows the architecture of the proposed system followed by system description.

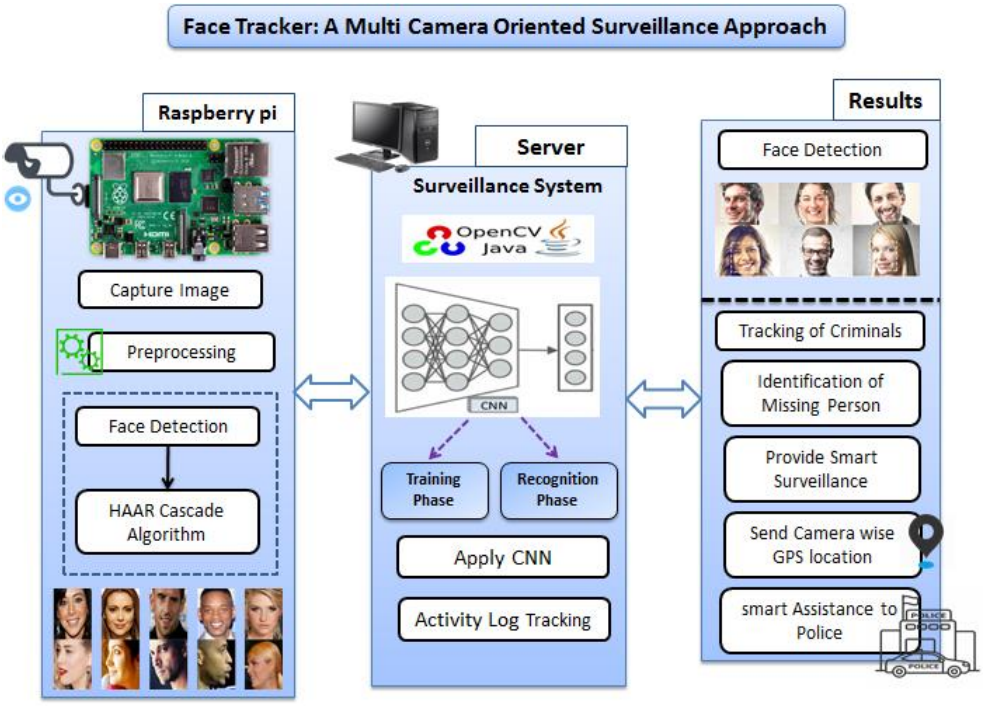

Figure: - System Architecture

1. Smart surveillance:

In simple words the smart surveillance is nothing but it just detects the changes based on the human motion in the video and save it.

2. Capture Images using Raspberry Pi camera:

The original high-definition color images, taken with the Raspberry Pi camera, were converted to downscale to a resolution of 640×480 pixels.

3. Pre-Processing on images:

Once we get the faces, apply the preprocessing on images like noise removal, grayscale conversion, etc.

a. **RGB to Gray Scale Image:**

Convert the image into Gray scale by taking the average of the each pixel RGB.

b. Apply HAAR Classifier

A Haar Cascade is a classifier that is used to detect the face images. The Haar Cascade is trained by superimposing the positive image over a set of negative images. The training is generally done on a server and various stages.

c. Haar Feature Classifier

A Haar features a stage threshold. Each stage does not have a set number of Haar features. Depending on the parameters of the training data individual stages can have a varying number of Haar features.

d. Haar Features:

Haar features are composed of either two or three rectangles. Face candidates are scanned and searched for Haar features of the current stage. Each Haar feature has a value that is calculated by taking the area of each rectangle, multiplying each by their respective weights, and then summing the results.

4. CNN Training dataset:

A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of Convolutional layers, pooling layers, fully connected layers, and normalization layers. CNN will be used to train the Images for recognizing phase.

Training Input: - Face Images Dataset

Training Output: - CNN Model

5. Apply CNN:

CNN Input: - Captured Images using Raspberry Pi camera

CNN Output: - Persons Name.

The Convolutional Neural Networks (CNN) is applied on the preprocessed image. The output of the polling is forwarded to the fully connected layer and then we get the output in the form of face detection.

6. Face Recognition:

Using Convolutional Neural Networks (CNN) algorithm analysis on the preprocessed images, Detect The Face of the user for smart surveillance.

IV. CONCLUSION

Face recognition for police investigative purposes solve more crimes faster with diminishing budgets. Police agencies want to use more automated tools, including facial recognition for crime detection. The multi-cameras environment covers more physical space than a single camera view which provides a more comprehensive view about the crime scene.

Here an attempt is made to implement a system that based on face detection using Haar-like features and convolution neural network (CNN). This system gives effective assistance to the police department by using a technique like a face recognition and artificial intelligence in reducing the crime rate of the city and assists in surveillance in real-time.

REFERENCES

[1] Michael Davis,Stefan Popa,CristinaSurlea "Real-TimeFaceRecognitionfromSurveillance Video", : https://www.researchgate.net/publication/225322442 ,2011.

[2] Prof. Archana D. Gaikawad, Prof. Paresh D. Sonawane,"An Efficient Video Surveillance System Using Video Based Face Recognition on Real World Data",International Journal of Science, Engineering and Technology Research (IJSETR)2016.

[3] Ľuboš OVSENÍK, Anna KAŽIMÍROVÁ KOLESÁROVÁ, Ján TURÁN, "VIDEO SURVEILLANCE SYSTEMS ",Acta Electrotechnica et Informatica, Vol. 10, No. 4, 2010, 46–53.

[4] Le An, Mehran Kafai, Bir Bhanu ,"Face Recognition in Multi-Camera Surveillance Videos using Dynamic Bayesian Network ".

[5] Ming Du, Student Member, IEEE, Aswin C. Sankaranarayanan, Member, IEEE, and Rama Chellappa, Fellow, IEEE, "Robust Face Recognition from Multi-View Videos", IEEE TRANSACTIONS ON IMAGE PROCESSING.

[6] Nor Nadirah Abdul Aziz*, Yasir Mohd Mustafah†, Amelia Wong Azman‡, Amir Akramin Shafie§, Muhammad Izad Yusoff¶, Nor Afiqah Zainuddin|| and Mohammad Ariff Rashidan,"Features-Based Moving Objects Tracking for Smart Video Surveillances: A Review",International Journal on Artificial Intelligence Tools Vol. 27, No. 2 (2018).

[7] Frederick W. Wheeler, Richard L. Weiss and Peter H. Tu,"Face Recognition at a Distance System for Surveillance Applications".

[8] Bellotto, E. Sommerlade, B. Benfold, C. Bibby, I. Reid, D. Roth,C. Fern´andez, L. V. Gool, and J. Gonz`alez, "A distributed camerasystem for multi-resolution surveillance," inProc. of the ACM/IEEEIntl. Conf. on Distributed Smart Cameras (ICDSC), 2009.

[9] Josh Harguess, Changbo Hu, J. K. Aggarwal,"Fusing Face Recognition from Multiple Cameras",2007..

[10] Wei Niu, Long Jiao, Dan Han, and Yuan-Fang Wang ,"Real-Time Multi-person Tracking in Video Surveillance ".

[11] Mahdi, Fahad Parveza; * | Habib, Md. Mahmudulb | Ahad, Md. Atiqur Rahmanb | Mckeever, Susanc | Moslehuddin, A.S.M.b | Vasant, Pandian,"Face recognition-based real-time system for surveillance ",Intelligent Decision Technologies, vol. 11, no. 1, pp. 79-92, 2017