ISSN No:-2456-2165

Hybrid Motion and Face Recognition with Detection for Criminal Identifications

Anumol K. B.[1], Anagha Ashokan[2], Safna Sainudeen[3], Shijina B.[4]
[1]B.Tech UG Students, [2-3]Assistant Professor, [4]Head of the Department,
[1]Department of Computer Science and Engineering, TOMS college of Engineering
APJ Abdul Kalam Technological University, Kerala, India

Abstract:- Identification and detection of a criminal takes time and effort. Today's criminals are becoming more intelligent, leaving no biological evidence or fingerprints at the scene of the crime, making motion detection software one of the most essential security features in recent years. It is employed to improve already-installed security tools, like the motion sensor lighting on indoor and outdoor security cameras. Through the use of CCTV cameras placed at various locations, an advanced facial identification system like this may be automated to find offenders. During emergencies and disasters, this method can also be utilized to find those who have gone missing. This method may be improved to recognize numerous faces at once and recognize faces in damaged or cropped photos.

Keywords:- *Python, machine learning, biological evidences, motion detection, identification of criminals, detection of criminals.*

I. INTRODUCTION

Motion detection tracks shifts in an object's position with relation to its surroundings, and vice versa. We can identify moving objects in front of the camera with the aid of this motion detection program. This software can be used to carry out the following activities, among others: 1) Take a screenshot while working at home; 2) Watch your youngster before screen time; 3) Get unauthorized entry to your yard; 4) Discover unwelcome people/animals moving around your room/house/tunnel. Finding a criminal has been a challenging procedure over the years. Formerly, the entire procedure was predicated on leads derived from evidence discovered at the crime site. Finding biological evidence is simple. But, criminals are more skilled than ever at hiding their traces and avoiding leaving behind any sort of forensic proof. Here, face detection and recognition are relevant. Every face is different because it can be recognized, which is important for establishing a person's identity. Facial recognition is a unique biometric method for criminal identification that has the advantages of high accuracy and minimal intrusion. It is a method that employs facial recognition to automatically identify and validate a person from video or image frames. The face recognition system described in this paper combines the best face detection, feature extraction, and classification methods currently in use. MTCNN for detection and Face Net for embeddings are two examples of deep learning techniques that have historically been shown to be sophisticated and modern. Our layer for streaming video is incredibly steady and doesn't switch between successive video frames. In order to make a significant change, we watch the backdrop if we can model it.If there is a significant shift, we can accommodate it; this change frequently resembles the motion of our video. It is now obvious from experience that this belief is susceptible to disbelief. Our background may appear quite differently from the numerous video frames due to factors including dignity, thinking, lighting, and other other environmental changes. In addition, our methods may be disregarded if the background appears different. Because of this, efficient retrieval and pre-discovery systems make use of mounted cameras on horses and artificial lights. By using automatic face recognition, the system extracts important facial characteristics from the subject, such as the size of the nose or jawline, the space between the eyes, the color of the eyes, etc. These characteristics are helpful for categorizing items and running database matches. This approach uses the two crucial stages of detection and identification. Two key procedures are triggered by face recognition: training and evaluation. As an algorithm is being trained, it is fed a sample set of photos from the training set. The freshly captured test image is compared to the existing database during the facial recognition evaluation step. The following is a breakdown of the structure of this work.

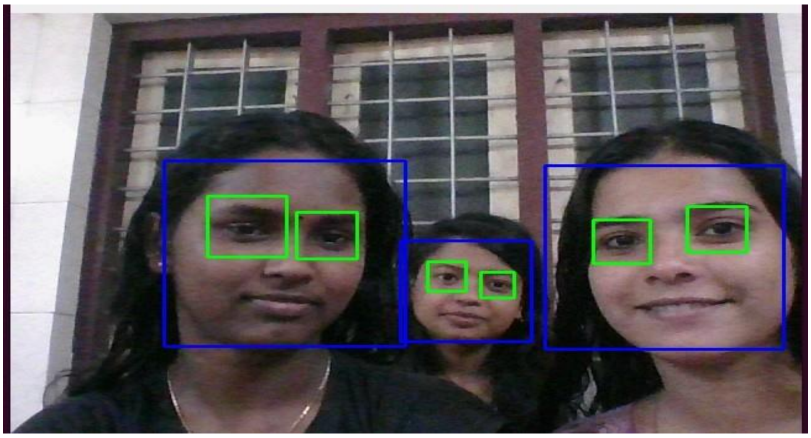

Fig. 1: Face detection example

The following section covers a review of the literature, followed by an overview of human detection in the third section, an explanation of the MTCNN and FaceNet modeling techniques in the fourth section, and the set of stages in the fifth. Results and discussion are covered in the sixth section, while conclusions and future work are covered in the seventh and eighth areas, separately.

II. LITERATURE REVIEW

Face recognition applications have been developing since the 1960s, as noted in [2], which also popularized the strategy of using a RAND tablet for coordinating facial features. A RAND device had been a gadget that allowed users to input vertical and horizontal coordinates on a grid using a pen that produced electromagnetic pulses. The entire system was used to manually record the coordinate locations of the eyes, hairline, mouth, and nose, among other facial features. By computing the phase difference between the original frame and the next frame, the movement will be determined. The new frames will be referred to as Delta frames, and they will allow you to use pixel power to enhance your delta frame. The Threshold framework will be the name of the revised structure. Larger objects are then caught utilizing additional sophisticated image processing techniques as Shadow Removal, Dilation, Contouring, etc. within the threshold framework [6]. The world was exposed to eigenfaces and the statistical method to face recognition by [4] and [5] in the late 1980s.To reduce dimensionality and project sample/training data into tiny feature faces, Eigenfaces uses Eigenvalues and Eigenvectors. This is the primary premise behind Principal Component Analysis (PCA). Referencing [3] uses 21 unique facial traits, such as the skin tone, chin, nose elevation, and hair color, to raise the bar for face identification.

III. RELATED WORKS

These section reviews a list of papers that are connected to this research. Search terms like "face detection," "face recognition," "feature extractions," and "computer vision" were used to find the paper;

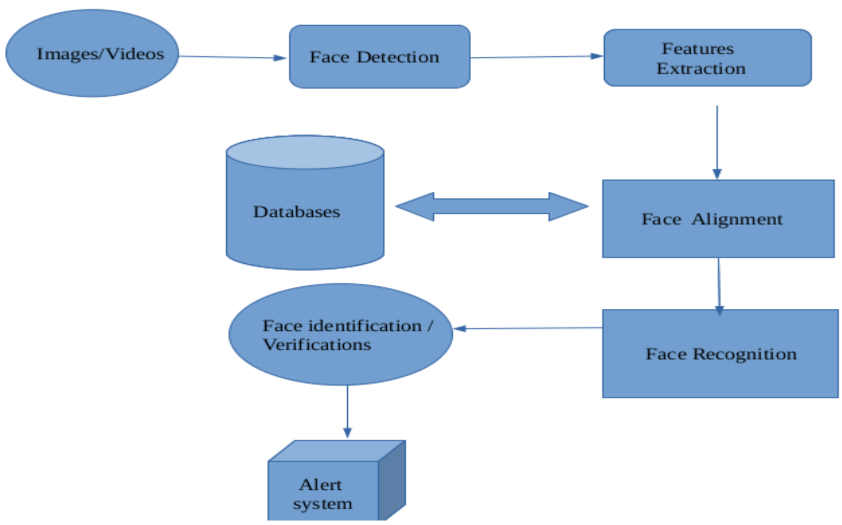

Fig. 2: Working Flow chart of criminal identification system

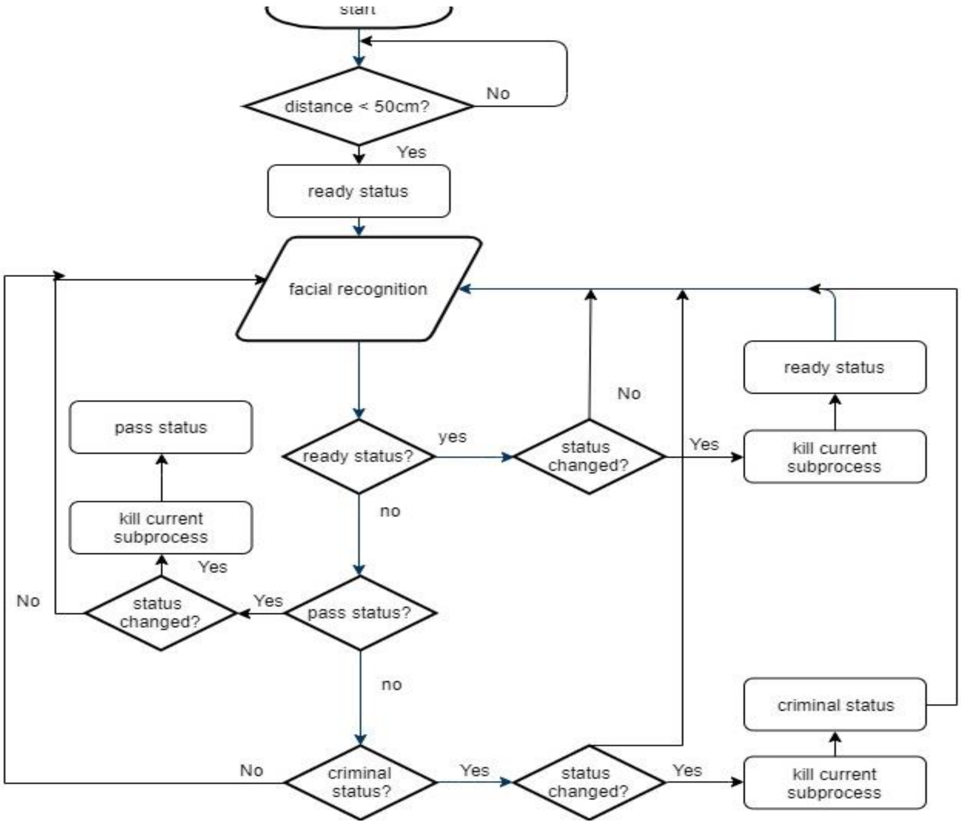

Fig. 3: Flow chart for motion detection and face recognitions

A. The MTCNN's Three Tasks

Three outputs are required from the network: localization of facial landmarks, bounding box regression, and face/non-face categorization.

- *Face classification is a cross-entropy loss-based binary classification problem:*
  Li ^ det = -(yi ^ det log (Pi) + (1-yi ^ det) (1-log (Pi)))

  Where,
  yi = Ground truth table, yi ^ det is an element of {0,1}
  Pi = Probability produced by the network

- *Regression using a bounding box:*
  A regression issue constitutes the learning aim. The offset between a candidate and the closest ground truth is calculated for each candidate window. For this challenge, Euclidean loss is used:

  Li ^ box = || yi ^ box – yi ^ box ||2 ^2

  here,
  yi ^ box = Target obtained from networks
  -yi ^ box ||2 ^ 2 = Ground truth coordinate

- *Localization of facial landmarks:*
  The localization of face landmarks is modeled as a regression issue with Euclidean distance as the loss function:

  Li ^ landmark = || yi ^ landmark - yi ^ landmark ||2 ^ 2

B. Gaussian Mixture Modelling (GMM)

Another kind of clustering algorithm is GMM. Each cluster is modeled using a unique Gaussian distribution, as suggested by its name. We have soft assignments into clusters as opposed to rigid ones like k-means because of this flexible and probabilistic approach to modeling the data. This implies that each data point might, with a given probability, have come from any of the distributions. In reality, every distribution shares some of the “responsibility" for producing a certain data point. How can we estimate a model of this kind? We may introduce a latent variable (gamma) for each data point, for example. This presupposes that some knowledge of the latent variable was used to generate each data point. the two components of the EM algorithm an M-step or maximized step, as well as an E-step or expectation step. Let's imagine we have our data points X and some latent variables, which are unseen and represented by the vector Z below. The vector represents our objective, which is to maximize the marginal likelihood of X given our parameters. Essentially, using the sum rule of probability, we can determine the marginal distribution as the product of X and Z plus the sum of all Zs.

$$ln\, p(X|\Theta) = ln\left\{\Sigma_z p(X, Z|\Theta)\right\}$$

IV. FINAL SUBMISSION

A. Training Data

- The three MTCNN steps can be explained more simply as follows:

✓ In the first stage the MTCNN creates multiple frames which scans through the entire image starting from the top left corner and eventually progressing towards the bottom right corner. The information retrieval process is called P-Net(Proposal Net) which is a shallow, fully connected CNN.

✓ The second stage involves feeding the R-Net(Refinement Network), a fully connected, complicated CNN that rejects the majority of frames that don't include faces, with all the data from the P-Net.

✓ The third and last stage uses a more advanced CNN

✓ called O-Net (Output Network), which, as its name implies, outputs the facial landmark position after identifying a face in the provided picture or video.

B. The efficiency of hard sample mining online

We train two O-Nets (one with and one without online hard sample mining) and compare their loss curves in order to assess the impact of the suggested online hard sample mining technique. We solely train the O-Nets for the face classification task in order to more directly compare the two systems. In these two O-Nets, all training parameters, including network initialization, are identical. Fix learning rate is used to compare them more easily.

C. The efficiency of joint alignment and detection

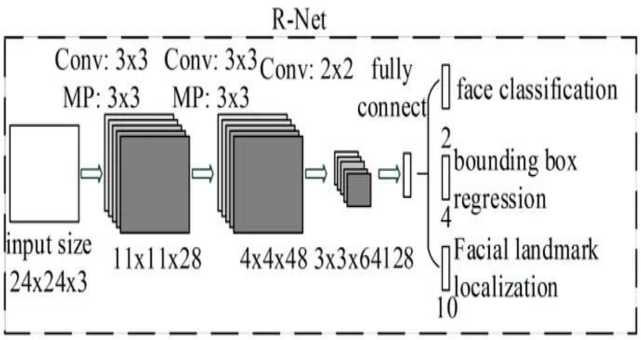

Fig. 4: R-Net configuration og mtcnn

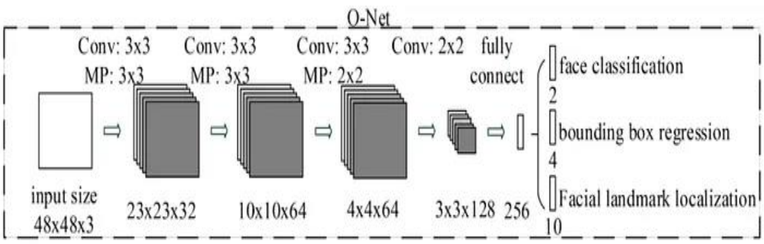

Fig. 5: O-Net configuration og mtcnn

We compare the performances of two different O-Nets (joint facial landmarks regression task and do not joint it) on FDDB (with the same P-Net and R-Net for fair comparison) in order to assess the contribution of joint detection and alignment. We contrast the effectiveness of bounding box regression in these two O-Nets as well.

D. Face Embeddings

In order to compare faces with various vectors, face embeddings must be made. The FaceNet model was applied in this stage to generate embeddings. The pixel values must be normalized after loading the compressed file of recognized faces because FaceNet requires this. The loaded Keras facenet model has already been trained. To determine each face's prediction and embedding from the train and test sets, each face is counted. A compressed NumPy array was used to store the embeddings.

ISSN No:-2456-2165

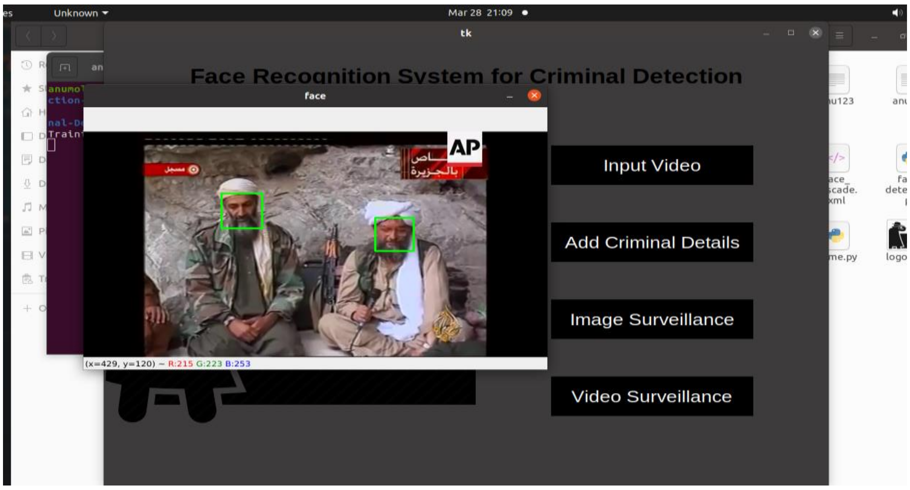

Fig. 6: Detected faces from sample video

E. Classifications of faces

Using models trained with machine learning, embeddings are categorized at this step of the procedure in order to be recognized as one of the criminals. Model vector normalization is used to scale values prior to categorization. For this, the scikit learn normalization library is employed. The criminals' names are then changed from string to integer format. Scikit Learn's Label Encoder is used for this. Linear Support Vector Machine was chosen as the classification model since it is efficient at distinguishing between the various face embeddings. On the training set of data, the linear SVM model is fitted.

Fig. 7: Samples of criminal data registed successfully

V. CONCLUSIONS

This essay outlines a novel method for facial recognition and shows how it might be used for the crucial task of criminal detection and identification. Applications for face recognition technology are numerous, and many issues in everyday life can be solved using comparable techniques. We think creating such a system is an eager step toward hastening and streamlining the process of apprehending criminals and law enforcement. With a dynamic dataset, this system can be further developed to identify criminals in real-time. Although computer vision applications can be tough, they make complex problems easier to solve.

ACKNOWLEDGEMENT

The task has been successfully completed, and I am very appreciative to God Almighty for his grace and blessings, without which this study would not have been possible.

REFERENCES

[1.] Richardson, M., & Wallace, S.(2012). Getting started with raspberry P I."O'Reilly Media, Inc.".
[2.] Bledsoe, "Manual measurements",1960
[3.] A.J. Goldstein, L.D. Harmon and A.B. Lesk, "Identification of human faces, "in proceedings of the IEEE, vol 59, pp. 748-760, May 1971.
[4.] L.Sirovich and M. Kirby, "Low dimensional procedure for the characterization of human faces," in Journal of the Optical Society of America A, vol 4, pp. 519-524, 1987.

[5.] M. Turk and A. Pentland, "Eigenfaces for Recognition," in Journal of cognitive neuroscience, vol 3, pp. 71-86, Jan 1991.

[6.] Roqueiro, Petrushin, counting people using video cameras, Department of Computer Science, University of Illinois at Chicago, Chicago, IL

[7.] 60607, USA, 2006

[8.] Security enhancement utilizing motion detection, Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), 2011 International Conference on, T huckalay, 2011, pp. 552–557. T homas, Ashraf, Lal, Mathew, Jayashree.

[9.] Automatic criminal identification by face recognition using open computer vision classifiers, in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 775–778. P. Apoorva, H. C. Impana, S. L. Siri, M. R. Varshitha, and B. Ramesh.

[10.] Choksi, V. M., Jadhav, S., and Potdar, M. B., together with Patel, P. B. (2016). smart raspberry pi motion detecting system. IJAIS, 10(5), 37–40. International Journal of Applied Information Systems.

[11.] Yue, Wei, and Li. A Video Sequence Segmentation Algorithm for Foreground-Background.China's Jiangnan University.2015

[12.] (2014) November. Abaya, W. F., Basa, J., Sy, M., Abad, A. C., and Dadios, E. P. With a Raspberry Pi and OpenCV, a smart security camera with night vision functionality is inexpensive.2014 saw the first-ever Humanoids, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) international conference (pp. 1-6).IEEE.

[13.] Joint face detection and alignment using multi-task cascaded convolutional networks, K. Zhang, Z. Zhang, Z. Li, and Y. Qiao.

[14.] FaceNet: A unified embedding for face identification and clustering, F. Schroff, D. Kalenichenko, and J. Philbin, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815-823.