



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12    **Issue:** X    **Month of publication:** October 2024

**DOI:** <https://doi.org/10.22214/ijraset.2024.64654>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# Real-Time Criminal Detection System Using Deep Learning Technique

Dhananjay Shendge A<sup>1</sup>, Shubham Singh A<sup>2</sup>, Aditya Barde A<sup>3</sup>, Prof. Snehal Khartad A<sup>4</sup>

Department of Electronics and Telecommunication Engineering, Dhole Patil College of engineering, Pune

**Abstract:** *In this paper, we may developed a system for detecting criminal faces, for this, we have used deep learning algorithms. Since deep learning is now the most famous technology, it is used in different applications. One such application is crime detection and prevention. This system identifies the criminal face, retrieves the information stored in the database for the identified criminal and a notification is sent to the police personnel with all the details and the location at which the criminal was under the surveillance of the camera.*

## I. INTRODUCTION

Crime is one of our societies most serious and pervasive issues, and preventing it is a critical duty. Different types of crimes and the full consideration of the protection and safety of citizens in any society are significant components that play a vital role directly in the quality of the lives of residents. Certain types of criminal incidents such as larceny, identity theft, or even pickpocketing can cause disturbance and stress in an individual's life and affect his mental peace. The use of large numbers of closedcircuit television systems (CCTV) in both public and private settings has been considered a necessary in response to rising concerns about crime and its danger to security and safety. A deep learning-based approach is employed as it provides a better performance and faster results as compared to the existing techniques, thereby providing real-time data for police forces to function more efficiently.

Deep Learning pre-trained models are built models that assist users in learning about algorithms or experimenting with existing frameworks for better outcomes without explicitly developing. In addition, deep learning neural networks have five layers, which include input and output layers with Convolution, Max-Pooling, and Fully connected layers. Because of limited time, memory, and resources such as CPUs and Processors, many people prefer Deep Learning pre-trained concepts. And, when compared to machine learning, which requires us to construct explicitly, these pre-trained models will provide the best and most accurate outcomes. There is a significant amount of human interaction required to detect firearms in surveillance videos, which is prone to human error. It is required to design an automatic surveillance system that detects firearms with less computing time in order to limit crime incidence.[3]

## II. RELATED WORK

Many facial applications, such as face recognition and facial expression analysis, involve face detection and alignment. However, in real-world applications, large visual fluctuations of faces, such as occlusions, substantial pose fluctuations, and severe lightings, offer significant hurdles for these tasks.

- 1) Viola and Jones offer a cascade face detector that uses Haar-Like features and AdaBoost to train cascaded classifiers, resulting in good performance and real-time efficiency. However, several studies show that this type of detector, even with more advanced features and classifiers, may suffer dramatically in real-world applications with larger visual variances of human faces.
- 2) In addition to the cascade structure, Mathias provided deformable component models for face detection, which performed admirably. They are, however, computationally demanding and may necessitate costly annotation during the training step.
- 3) Convolutional neural networks (CNNs) have recently made significant advances in a range of computer vision applications, including image classification and face recognition.
- 4) Deep CNNs are trained for facial attribute identification to get high response in face regions, which leads to candidate face windows. This technique, however, is very consuming in practise due to its sophisticated CNN structure.
- 5) The use of cascaded CNNs for face detection necessitates extra processing cost for bounding box calibration from face detection and ignores the intrinsic relationship between facial landmarks localization and bounding box regression.[1]

### III. METHODOLOGY

#### A. MTCNN (Multi Task Cascade Neural Network)

MTCNN (Multi-task Cascaded Neural Network) detects faces and facial landmarks on images/videos. The MTCNN concept may be broken down into three stages, the third of which involves performing facial detection and facial landmarks at the same time. These stages consists of various CNN's with varying complexities.

- The MTCNN creates numerous frames in the first step, which scan the entire image starting from the top left corner and progressing to the bottom right corner. The information retrieval process is called P-Net (Proposal Net) which is a shallow, fully connected CNN.
- In the second stage all the information from P-Net is used as an input for the next layer of CNN called as R-Net (Refinement Network), a fully connected, complex CNN which rejects a majority of the frames which do not contain faces.
- In the third and final stage, a more powerful and complex CNN, known as O-Net (Output Network), which as the name suggests, outputs the facial landmark position detecting a face from the given image/video.[1]

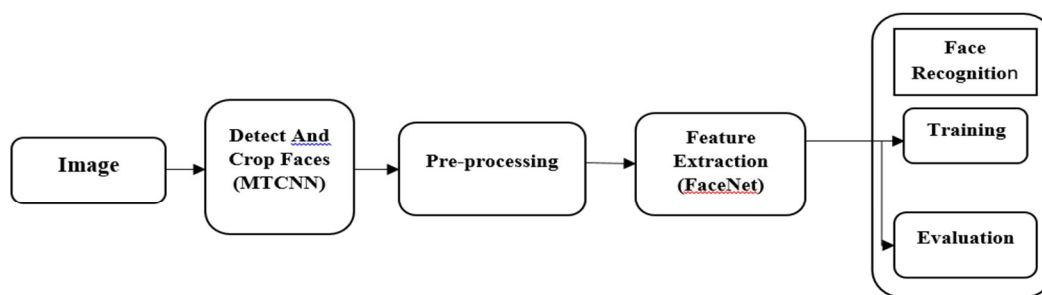
#### B. FaceNet

Face-embedding were created using the FaceNet algorithm. The facial features of a person's face are represented by the embedding vectors. As a result, the embedding vectors of two separate photos of the same person will be closer together, whereas those of a different person will be more apart. The distance between face encodings created by the Encoder network (InceptionResNet-v1) is used as a metric to compare two faces' similarity. The Triplet Loss is used to train the Encoder network, which necessitates effective Triplet Mining.[2]

#### C. OpenCV

OpenCV is a video and image processing library that is used for image and video analysis such as facial detection, licence plate reading, photo editing, advanced robotic vision, and many more applications. OpenCV is used to do real-time face detection from a live stream via our webcam. The collection contains over 2500 optimised algorithms, which include a complete mix of both classic and cutting-edge computer vision and machine learning techniques. These algorithms can be used to detect and recognise faces, identify objects, classify human actions in videos, track camera movements, track moving objects, produce 3D point clouds from stereo cameras, extract 3D models of objects, stitch images together to create a high-resolution image of an entire scene, and find similar images from an image database.[2]

#### D. Proposed Diagram



In a block diagram, each component would be represented as a distinct block, with arrows indicating the flow of data from one block to the next.

This structure outlines a comprehensive approach to face detection and recognition, enhancing the accuracy and efficiency of identifying individuals in various applications.

**E. Block Diagram Components**

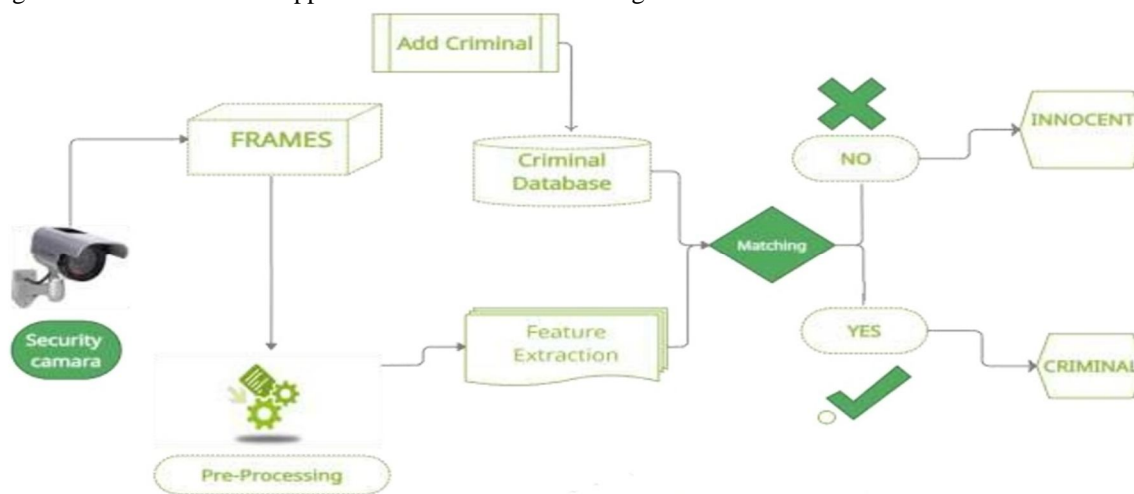
- 1) Image Input: Captures images from cameras, databases, or user uploads.
- 2) Detect And Crop Faces: Identifies and locates faces within the input image. And extracts the detected face region from the image.
- 3) Pre-Processing: Prepares cropped faces for analysis including the standardizes face dimensions.
- 4) Feature Extraction: Extracts relevant features from pre-processed face images and it is based on deep learning.
- 5) Face Recognition: Matches extracted features against a database of known faces.
- 6) Training: Involves training the recognition model using labeled datasets and Machine learning or deep learning algorithms are applied to learn from features.
- 7) Evaluation: Assesses the model's performance using metrics like accuracy, precision, recall.

**IV. FLOW OF WORK**

The criminal face identification is implemented by extracting the face from video or image, identify the face. The face is searched in the database to look for the details about the criminal.

**A. Pre-processing**

- 1) Clean and pre-process the data to remove noise and inconsistencies. This step might involve data augmentation, normalization, and other techniques to make the data suitable for deep learning models.
- 2) Data Augmentation can also be implemented.
- 3) The input will be multiplied by a factor known as "rescale" before any further processing is performed. The RGB values of our original images range from 0 to 255, but at a typical learning rate, these values would be too high for our models to comprehend. To achieve values between 0 and 1, we scale our original photographs by a factor of 1/255.
- 4) Shearing transformations can be applied at random with shear range.



- 5) *Registering New Criminal*: This is the first step of implementing face detection as the criminal face with id, name, age, state and crime committed is registered to the database.
- 6) *Pre-processing Images*: Processing the features that are to be extracted, for improving the rate of recognizing the face. The facial image is cropped and is resized at lesser pixel value. Ascertain images contain disturbances it will be hard to train the model, results in the inaccurate histogram.
- 7) *Feature Extraction*: The performance of the entire system depends on this step. Different facial features are extracted using different mtcnn classifier. Grayscale images from this step used for identification of the criminal and train the model.
- 8) *Matching*: Compare the resultant image with the existing images in database. If match is found then return the data related to that image from the database otherwise the recognized person is not criminal.
- 9) *Send Notification (If person is criminal)*: The system has been developed in an open source platform using python. Sending of SMS is done by creating an account in Twilio and installed the Twilio library in python. Twilio allows programmatically make and receive phone calls, send and receive text messages.

## V. CONCLUSION

In conclusion, the face detection and recognition system employs a systematic approach to identify and verify individuals through a series of structured processes. Starting from image input, it effectively detects and crops faces, followed by pre-processing to enhance image quality. Feature extraction and recognition algorithms then match these features against a known database. The model is continuously trained and evaluated to improve its accuracy and reliability. This comprehensive methodology ensures effective implementation in various applications, enhancing security and user authentication capabilities.

## VI. FUTURE ENHANCEMENTS

Even though the proposed system is limited to the academic area, this can also be used to identify criminals in public or private places. The model can be used in any scenario where crimes are more likely to happen. Instead of searching the entire database to compare the faces, model performance can be improved by considering other traits like the age and gender of a person.

## REFERENCES

- [1] MTCNN method which is used in this project was proposed by Kaipeng Zhang et al. in their paper 'Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks', arXiv:1604.02878. DOI: 10.1109/LSP.2016.2603342.
- [2] Florian Schroff, Dmitry Kalenichenko: "Face Net: A Unified Embedding for Face Recognition and Clustering", 2015; arXiv: 1503.03832. DOI: 10.1109/CVPR.2015.7298682.
- [3] Sanika Tanmay, Aamani Tandasi, Shipra Saraswat at 2021 11th international conference on cloud computing, Data Science & Engineering : "Face Detection and Recognition for criminal Identification system", DOI: 10.1109/Confluence51648.2021.9377205.
- [4] Facenet : A Unified Embedding for Face Recognition and Clustering by Florian Schroff, Dmitry Kalenichenko, James Philbin (arXiv:1503.03832v3 [cs.CV] 17 Jun 2015).
- [5] Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep face recognition: A survey," in Proc. 31st SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI), Oct. 2018, pp. 471–478, doi: 10.1109/SIBGRAPI.2018.00067.
- [6] D. King. (2017). High Quality Face Recognition With Deep Metric Learning. [Online]. Available: <http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.htm>.
- [7] Y. Zheng, J. Chang, Z. Zheng, and Z. Wang, "3D face reconstruction from stereo: A model based approach," in Proc. IEEE Int. Conf. Image Process., 2007.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)