Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Criminal face identification system using deep learning algorithm multi-task cascade neural network (MTCNN)

K. Kranthi Kumar [a,*], Y. Kasiviswanadham [a], D.V.S.N.V. Indira [a], Pushpa Priyanka palesetti [b], Ch.V. Bhargavi [c]

[a] IT Department, Gudlavalleru Engineering College, Gudlavalleru, India
[b] IT Department, VNR Vignan Jyothi Institute of Engineering and Technology, Hyderabad, India
[c] IT Department, Vignan Institute of Information Technology, Visakhapatnam, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Crime
Detection
Neural network
Deep learning
Algorithm
Multi-task

ABSTRACT

Nowadays criminal activities are growing at an exponential rate. Crime prevention by effective identification of criminals is the main issue before the police and on the other hand, the availability of police officers is inadequate. There are various technological solutions for detecting criminals but they are not up to the mark. In this project, a face detection and recognition system for criminal identification is developed using the multi-task cascade neural network. This system will be able to detect faces and recognize faces of criminals automatically in real-time. This system would also just require a single image of the criminal to recognize him, also known as one-shot learning. The purpose is to identify the criminal face, retrieve 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 he was under the surveillance of the camera.

Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

1. Introduction

Crime is one of the greatest and overwhelming issue in our general public and it's anything but a significant errand. Every day there are tremendous quantities of wrongdoings perpetrated regularly. This require monitoring every one of the wrongdoings and keeping an information base for same which might be utilized for future reference. Various kinds of wrongdoings and the full thought of the insurance and security of residents in any general public are huge segments that assume an imperative part straightforwardly in the nature of the existences of inhabitants [1]. Specific sorts of criminal episodes like burglary, data fraud, or even pickstashing can cause unsettling influence and stress in a person's life and influence his psychological harmony [2]. In response to increasing anxieties about crime and its threat to security and safety, the utilization of substantial numbers of closed-circuit television system (CCTV) in both public and private spaces have been considered a necessity [3]. The use of these significant video footages is essential to incident investigations. But as the number of these systems rises, so as the need for human operator monitoring tasks [4]. 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 [5].

Numerous facial applications, for example, face acknowledgment and look examination, include face discovery and arrangement [6]. Be that as it may, in true applications, enormous visual changes of countenances, like impediments, considerable posture vacillations, and serious lightings, offer critical obstacles for these assignments [7]. 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 [8]. 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. In addition to the cascade structure, Mathias provided deformable component models for face detection, which performed admirably [9]. They are, however, computationally demanding and may necessitate costly annotation during the

* Corresponding author at: Assistant Professor, IT Department, Gudlavalleru Engineering College.
E-mail addresses: kk97976@gmail.com (K.K. Kumar), ykvnath@gmail.com (Y. Kasiviswanadham), indiragamini@gmail.com (D.V.S.N.V. Indira), pushpapriyanka_p@vnrvjiet.in (P. Priyanka palesetti), bhargavicherukumalli@gmail.com (C.V. Bhargavi).

https://doi.org/10.1016/j.matpr.2021.06.373
2214-7853/
Selection and peer-review under responsibility of the scientific committee of the International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology.

training step [10]. Convolutional neural networks (CNNs), have as of late made huge advances in a scope of PC vision applications, including picture order and face acknowledgment [11]. Deep CNNs are prepared for facial property distinguishing proof to get high reaction in face locales, which prompts up-and-comer face windows [12]. This method, in any case, is burning-through by and by because of its refined CNN structure [13]. The utilization of fell CNNs for face recognition requires additional handling cost for jumping box adjustment from face identification and disregards the characteristic connection between facial milestones confinement and bouncing box relapse [14]. At present, various technical means have been established in the research on the practical application of video facial recognition technology like criminal Identification [15]. However, existing technologies are cannot tackle relevant difficulties such as light intensity, substantial pose fluctuations, severe lightings monitoring, and facial expression, which limits the application of video face recognition technology. **"A Convolutional Neural Network Cascade for Face Detection".** In the current framework the course contains six CNNs, including three for paired order of faces versus non-countenances and three for bouncing box alignment, which is expressed as multiclass characterization of discretized removal designs. In these CNNs, without explicit clarification they followed AlexNet to apply ReLU nonlinearity work after the pooling layer and completely associated layer. The six CNNs in existing framework are in particular 12-net, 12-adjustment net, 24-net, 24-alignment net, 48-net, 48-adjustment net. The underlying CNN in the test pipeline is alluded as 12-net. To rapidly examine the testing picture, 12-net is an amazingly shallow twofold characterization CNN. As a general rule, if the allowed least face size is F, the test picture is first framed into a picture pyramid to cover faces at different scales, and afterward each level of the picture pyramid is extended by 12/F as the 12-net info picture. The CNN after the 12-net for jumping box adjustment is alluded to as 12-alignment net. The locale is cut out and scaled to 12 × 12 as the information picture for the 12-alignment net given an identification window.24-net is a twofold order CNN that lessens the quantity of recognition windows considerably more. The 12-alignment net location windows are edited and contracted into 24 × 24 pictures, which are then assessed by the 24-net. The 24-alignment net is an adjustment net with N adjustment designs, like the 12-adjustment net. The last paired grouping CNN is 48-net. It is feasible to utilize an all the more remarkable yet more slow CNN now in the course. We utilize a multi-goal plan in the 48-net, with an extra information duplicate in 24 × 24 and a substructure like the 24-net. The course's last stage is 48-alignment net. To get more exact adjustment, we just utilize one pooling layer in this CNN (Figs. 1–3).

2. Limitations of existing system

The difficulty in detecting faces are primarily due to two factors: 1) the wide range of visual alterations of human faces in cluttered backdrops; 2) the enormous search space of possible face positions and face sizes (Figs. 4–6).

2.1. Face detection in uncontrolled environments

The existing system cannot detect faces in environment where there is low light intensity, substantial pose fluctuations, severe lighting.

2.2. Computationally expensive

The existing System requires bounding box calibration from face detection with extra computational expense and ignores the inherent correlation between facial landmarks localization and bounding box regression (Figs. 7–9).

2.3. Accuracy

The proposed algorithm achieved more accuracy than the existing algorithm.

3. Related 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) 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.

b) 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.

c) 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.

d) 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 (Fig. 10).

(a) Send notification (If person is criminal):

Twilio library in python allows the system to send SMS in an enhanced Manner.

In this paper the following algorithms are used:

i) MTCNN (for face detection)
ii) Siamese Neural Network (for face identification)

3.1. MTCNN (Multi-task cascade neural networks)

MTCNN (Multi-task Cascaded Neural Network) distinguishes countenances and facial milestones on pictures/recordings. The entire idea of MTCNN can be clarified in three phases out of which, in the third stage, facial recognition and facial milestones are performed at the same time. These stages comprise of different CNN's with changing intricacies.

A less difficult clarification of the three phases of MTCNN can be as per the following

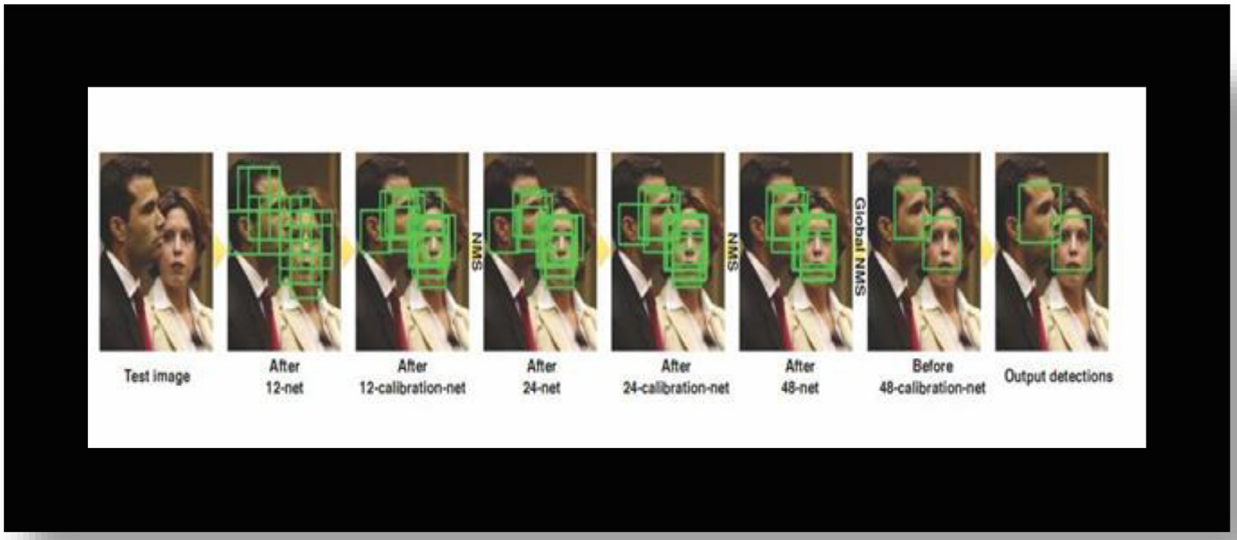

Fig. 1. Facial expressions.

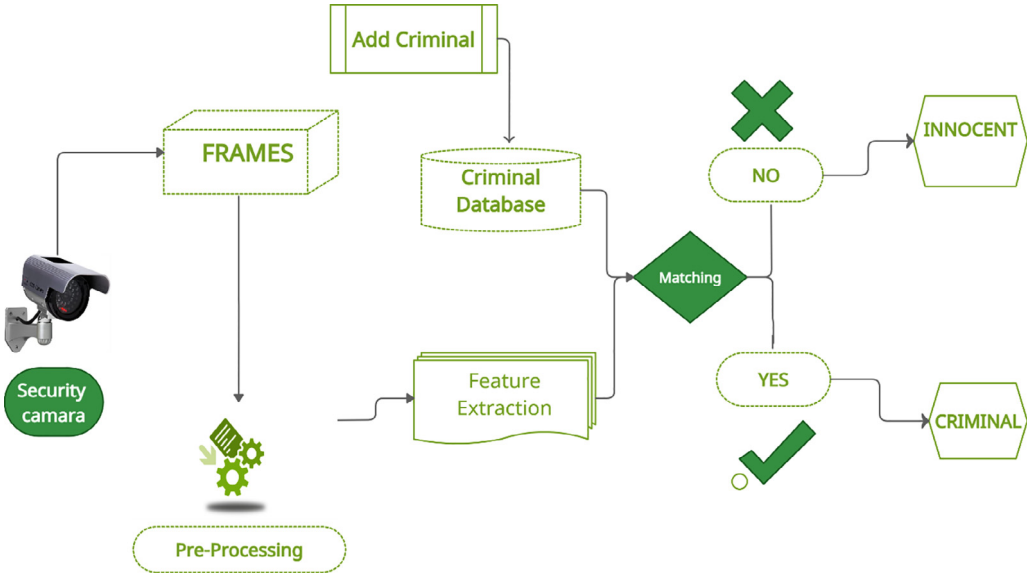

Fig. 2. Face recognition process.

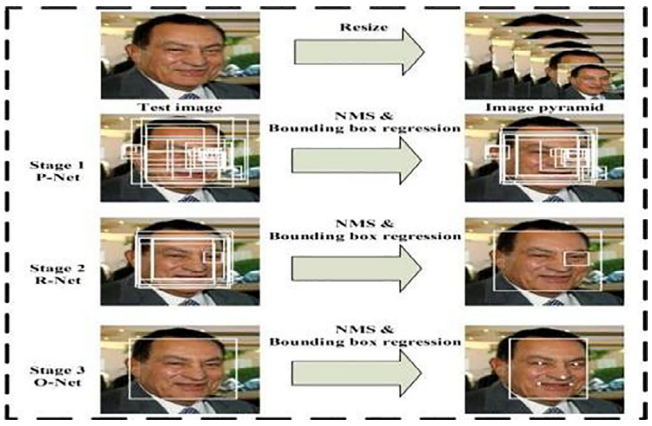

Fig. 3. Face recognition process.

- In the main stage the MTCNN makes numerous casings which look over the whole picture beginning from the upper left corner and in the long run advancing towards the base right corner. The data recovery measure is called P-Net (Proposition Net) which is a shallow, completely associated CNN.
- In the second stage all the data from P-Net is utilized as a contribution for the following layer of CNN called as R-Net (Refinement Organization), a completely associated, complex CNN which dismisses a larger part of the edges which don't contain faces.
- In the third and last stage, an all the more impressive and complex CNN, known as O-Net (Yield Organization), which as the name recommends, yields the facial milestone position identifying a face from the given picture/video.

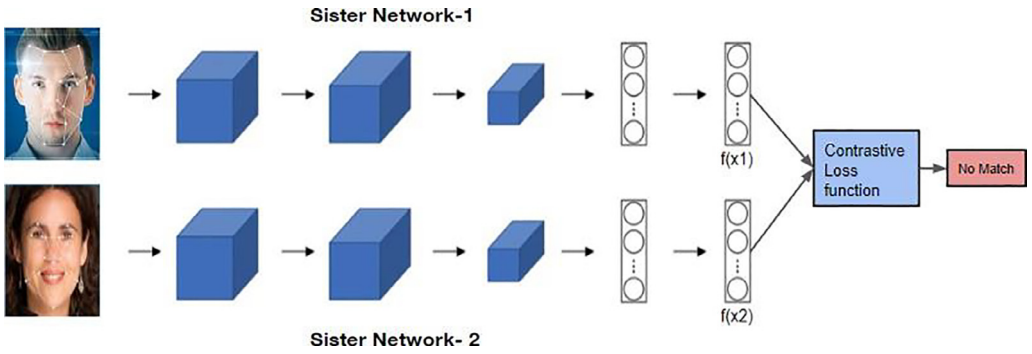

Fig. 4. Architecture of Siamese networks.

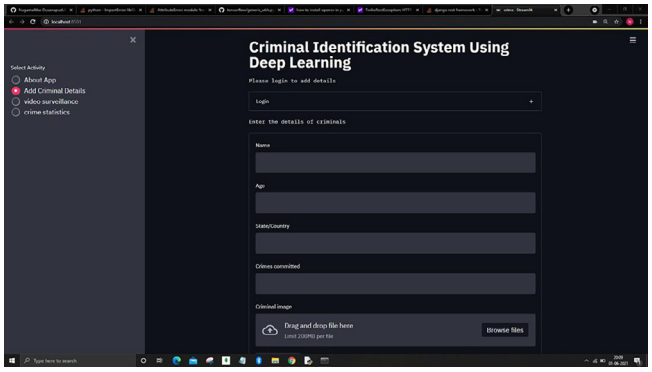

Fig. 5. Home page.

Fig. 6. Add criminal details page.

Fig. 7. Video Surveillance for real-time criminal identification page.

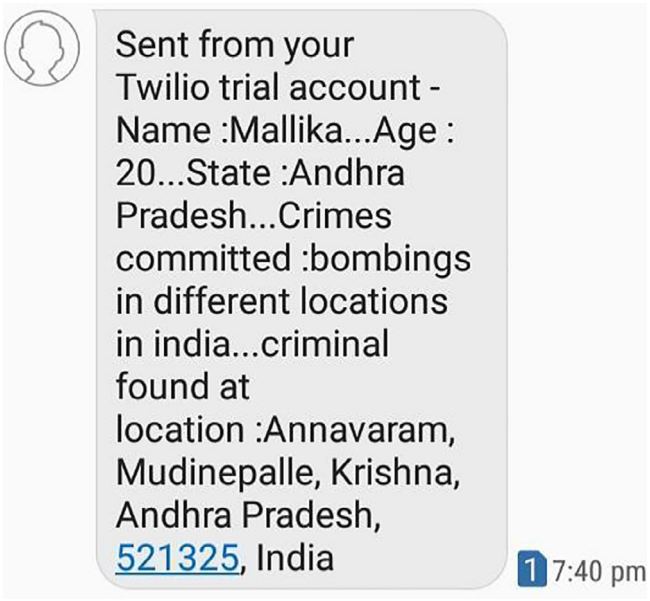

Fig. 8. Instance when person is identified.

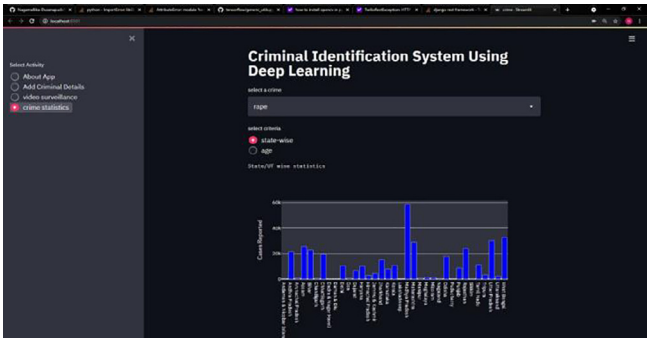

Fig. 9. Example of notification send to police when criminal is identified.

- Face Acknowledgment Utilizing Siamese Neural Organization: Face Net is a model that, given an image of a face, extracts high-quality features and predicts a 128-element vector representation of these features, known as a face embedding.
- Siamese neural network is a state-of-the-art algorithm that is quick and efficient in identifying dissimilarities in images. Algorithms nature to learn key features quickly and effectively differentiate between these features opens up new opportunities for its implementation especially in the field of video surveillance, facial recognition, self-learning tasks, etc.

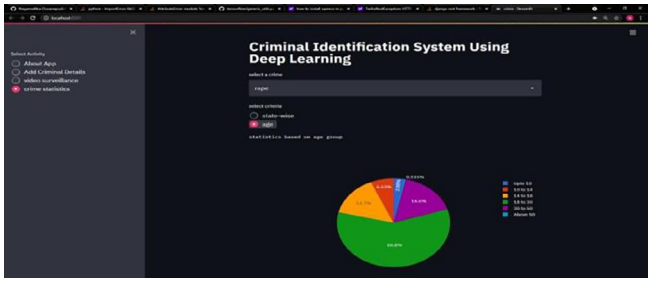

Fig. 10. Criminal statistics based on selected criteria.

- One-shot learning was practically achieved by Siamese networks that are a special type of neural networks, where instead of a model learning to classify its inputs, the network learns the similarity between two points and then differentiates them. Imagine you compliment your friends," You look far better than him/her". The Siamese neural network takes into account how far (distance) better looking you are.

4. Results

The criminal database with 50 records is collected and trained. The accuracy achieved is 86%. The system considers threshold (which tells how accurate faces should be compared) parameter which can be adjusted according to our requirement.

5. Conclusion and future scope

In the current world, practically all individuals know about the significance of CCTV film, yet much of the time, this recording is being utilized for examination purposes after a wrongdoing/occurrence has occurred. The proposed model has the advantage of recognizing and getting hoodlums before they perpetrate another wrongdoing. The constant CCTV film is being followed and investigated. The consequence of the examination is an order to the individual position to make any move if in the event that the framework distinguishes the crook. Henceforth this can be halted. Despite the fact that the proposed framework is restricted to the scholarly region, this can likewise be utilized to distinguish hoodlums in broad daylight or private spots. The model can be utilized in any situation where wrongdoings are bound to occur. Rather than looking through the whole data set to analyze the faces, model execution can be improved by considering different qualities like the age and sex of an individual.

CRediT authorship contribution statement

K. Kranthi Kumar: Conceptualization, Methodology. **Y. Kasiviswanadham:** Data curation, Project administration. **D.V.S. N.V. Indira:** Writing - review & editing. **Pushpa Priyanka palesetti:** Visualization, Investigation, Validation. **Ch.V. Bhargavi:** Writing - original draft, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Chauhan, B.K. Varghese, L.A. Rahman, V. Mohapatra, T. Badal, in: WIDER Face Challenge using Multi-Task Cascading Neural Network, IEEE, 2019 December, pp. 188–192.
[2] X. Lv, M. Su, Z. Wang, Application of face recognition method under deep learning algorithm in embedded systems, Microprocess. Microsyst. (2021) 104034, https://doi.org/10.1016/j.micpro.2021.104034.
[3] R. Jayaswal, M. Dixit, in: Comparative Analysis of Human Face Recognition by Traditional Methods and Deep Learning in Real-Time Environment, IEEE, 2020 April, pp. 66–71.
[4] S. Meivel, K.I. Devi, T.M. Selvam, S.U. Maheswari, Real time analysis of unmask face detection in human skin using tensor flow package and IoT algorithm, Mater. Today:. Proc. (2021).
[5] C. Peng, W. Bu, J. Xiao, K.C. Wong, M. Yang, An improved neural network cascade for face detection in large scene surveillance, Appl. Sci. 8 (11) (2018) 2222.
[6] Q. Guo, Z. Wang, C. Wang, D. Cui, in: Multi-face detection algorithm suitable for video surveillance, IEEE, 2020 July, pp. 27–33.
[7] C. Yu, H. Pei, Face recognition framework based on effective computing and adversarial neural network and its implementation in machine vision for social robots, Comput. Electr. Eng. 92 (2021) 107128, https://doi.org/10.1016/j.compeleceng.2021.107128.
[8] K. Li, H. Chen, 2019, December. Implementation of Automatic Face Detection System Based on ARM, in: 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), pp. 1–4. IEEE.
[9] C. Wu, Y. Zhang, MTCNN and FACENET based access control system for face detection and recognition, Autom. Control Comput. Sci. 55 (1) (2021) 102–112.
[10] Z.X. Ooi, Face recognition using deep learning (Doctoral dissertation UTAR), 2019.
[11] J.R.T. Garcia, Face detection from video streaming, 2018.
[12] Q. Ye, 2018 March, Masked Face Detection Via a Novel Framework, in: 2018 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018). Atlantis Press.
[13] S. Bakkali, M.M. Luqman, Z. Ming, J.C. Burie, Face Detection in Camera Captured Images of Identity Documents under Challenging Conditions 4 (2019 September) 55–60.
[14] S.T. Ratnaparkhi, A. Tandasi, S. Saraswat, in: Face Detection and Recognition for Criminal Identification System, IEEE, 2021 January, pp. 773–777.
[15] Seelam, V., kumar Penugonda, A., Kalyan, B. P., Priya, M. B., Prakash, M. D., Smart attendance using deep learning and computer vision, Mater. Today: Proc. (2021).