Deep Neural Networks for Face Recognition and Feature Extraction from Multi-lateral Images

Venkata Srinivasu Veesam[1], Suban Ravichandran[2], and Gatram Rama Mohan Babu[3],

1 Research Scholar, Department of Information Technology, Annamalai Univerity, Annamalai Nagar, TN, India
2 Associate Professor, Department of Information Technology, Annamalai Univeristy, Annamalai Nagar, TN, India
3 Professor, Department of Information Technology, R.V.R. & J.C. College of Engineering, Guntur-Dist, AP, India

Summary
Face recognition has been a strong secure system which doesn't requires any keys or cards. To protect law enforcement, safekeeping of home, to expose phone, even to find missing persons and in many more applications face recognition has been a popular method. The practice of ascertaining operators that appeal right to use to a system, network, or device is called Authentication. Cybercriminals uninterruptedly advance their outbreaks. Security teams are facing plenty of authentication-related challenges. In addition, government interest in face recognition technologies has increased because of their high-security level and accessibility. This made the firms to focus on more intellectual instance response tactics. Facial acknowledgment technologies are quickly becoming integral parts of user security, allowing for a secondary level of user authentication. To maintain the skill to afford protected right to use without a fall in speed there is a requisite for authentication methods. We already had inspiring domino effects by boosting traditional username and password security with facial biometrics, on the other hand, swotting these techniques is obligatory to conclude how effective these methods are within different constraints. A2 stage human recognition structure embroils face identification in general. Strategy and development of a face recognition system through deep learning using Open-CV in python is proposed with this paper. Performance of face detection and recognition is affected and damaged because occlusion often leads to missed detection. The proposed work is of an optical consideration tool supervision model to shrink the recognition exactitude initiated by facial occlusion and augment the precision of face detection. The uniqueness and endorsement of the proposed work takes the method to be more strategic over other algorithms. We need speed and accuracy to identify.

Keywords:
facial recognition, face detection, feature extraction, convolutional neural networks, deep neural networks

I. INTRODUCTION

Face detection plays an important role in biometrics. For safekeeping assurances in communal dwellings we have shadowing cameras and for video capture as well. Face recognition is certainly a series of several associated glitches. The process of identifying and verifying people from real-time scene images against stored database faces is called face recognition [4]. Face recognition system has compulsorily stored human face images earlier in the database. Face recognition applications can be embedded in existing video surveillance systems by minimum updating and producing great security. Automation in face recognition system is divided into the following two major parts:

- **Face Detection:** It detects face(s) region from the whole image background. The image has background objects and other body parts which are not required in face recognition activity.
- **Face Identification:** This activity performs the verification stage of detected face region towards previously-stored database faces region as human recognition activity.

Face recognition is one of the major research interests in the area of pattern recognition and computer vision. Numerous practical applications of face recognition are developed and used in the area of business application, the security system of military organizations, access control, an Employee attendance system is used in industries or institutes, law enforcement, social networking sites to tag the user automatically by their profile picture, etc. [3].

II. Literature Survey

The superlative technique for spotting a person is face detection and face recognition which doesn't look for any human work to recognize faces. Since a lot of methods were invented for face recognition and face detection. The concept of face detection and recognition was instigated in 1872 by Darwin's article The Expression of Emotion in Man and Animals [1]. In 1991, Turk and Pentland used the method Eigen faces [5]. They used residual error to detect the face(s) from an image. They have enabled partially reliable real-time face recognition methods [4]. Francis Galton, a famous criminologist had used face images for criminal identification in 1999 [2]. He stated distinct cardinal points on the face region outline. They use i) Distance between basic cardinal points, ii) Angles between

Manuscript received April 5, 2022
Manuscript revised April 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.4.82

points iii) Areas between the points asset of face features. Successful human identification applications were developed in the year 2000 but it requires human interaction intentionally such as fingerprint identification, iris identification, voice recognition, signature recognition, and retina identification [8]. Since 2010 many researchers and scientists have been working to add intelligence for automatic face detection and recognition. But it is not up to the mark when recognizing front and side view face [6]. By 2012 only limited surveillance applications were working well in their control environment that require precise face orientation and supports only front face for recognition [7]. The high implementation cost accompanying the use of Deep Convolution Neural Networks (DCNN) is still a challenging task despite of the improvements in accurate face recognition system, and the need to balance accuracy requirements with time and resource constraints. Some concerns which disturbs the face recognition with blockage, brightness and variations in pose significantly brought down the conventional solutions for face recognition.

III. Database

A huge training dataset is required to work with Deep Convolution Neural Networks (DCNN) [9]. To train and test every facial identification structure necessitates the use of face datasets in particularly a large dataset. For our proposed work we are using Mega face dataset which contains millions of images over 4.7 Million Photos 672,057 Unique Identities which has endorsed the improvement of accurate DL object detection systems [10] [11] [12]. A subclass of which is shown in Figure 1, 2. The required mean for testing and ranking the recognition system with highly challenging images in unimpeded settings is provided by the labelled faces in the dataset [13].

Figure-1 Images of different celebrities from Mega face dataset.

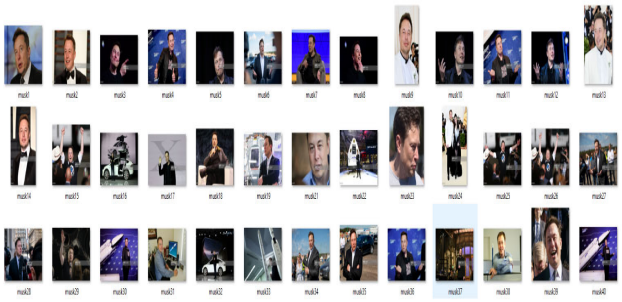

Figure-2 Set of images of a single person

The Mega face database was created to provide a large scale public and celebrity's dataset for training and testing the model, allowing fair comparison of algorithms without diminishing the accuracy.

IV. Proposed work

An impulsive face detection[14] system is developed using deep neural network and open CV using the landmarks i.e. facial key points we will extract the features which are very important for to develop machine learning model to detect a face[15], and some other requirements like identity, gender, emotion identification.
The frame work for the proposed work is given in the figure, in which we are having 3 main modules as Data Pre-processing (Embedding Features from faces), Training the model, and finally saving the model. In data pre-processing, we have three steps which are followed as collecting data, segregating the data and lastly extracting features using Face Descriptors, and Labelling. Here it is very important to collect at least 100 to 150 images of each person.

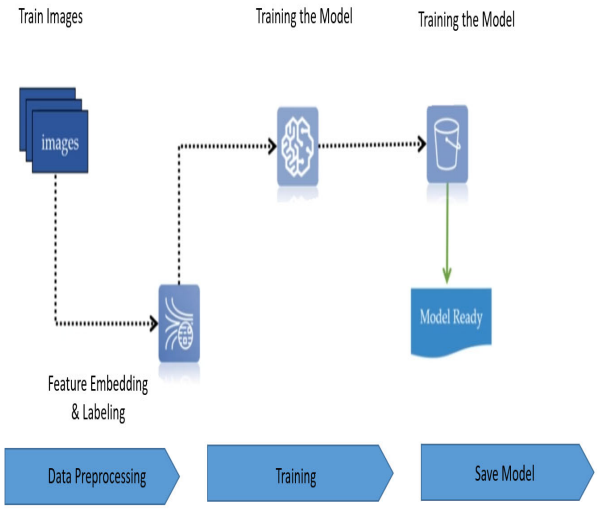

Figure-3 Frame work for the proposed DNN Face Detector

Face deduction using the Viola-Jones algorithm which is using harr features is having few disadvantages with the harr features. It may not work accurately for most of the faces. Now in the proposed work, for deep learning we are applying viola-Jones with a modified version. More specifically, the open-CV DNN face detector is based on a single-shot multi-box detector (SSD). Which is a framework using a ResNet-10 network. Open CV supports the DNN module, which uses common deep learning frameworks like Caffe, Tensor-Flow, Torch, and Darknet to construct a forward pass with pre-trained deep neural networks. Open-CV provides two models: Face detection FP 16 and Face detection UINT8. To build a model first the image is loaded followed by loading the model. DNN can able to detect the face accurately.

Features can be pull-out using Open-Cv face Extraction method and the flow to extract these features using the open CV method is shown below in figure 4. We are utilising the DNN to detect the face; once the face has been detected, the next step is to crop the face using Numpy slicing; finally, we will apply a shape predictor model to the cropped face, which will effectively identify the face's landmarks. The eyes, nose, and chin will be the focal points of the facial essential point. The deep neural network is used to pass the face shape predictor, and the result is a total of 128 dimensions. Clustering, similarity detection, and classification models can all benefit from this.

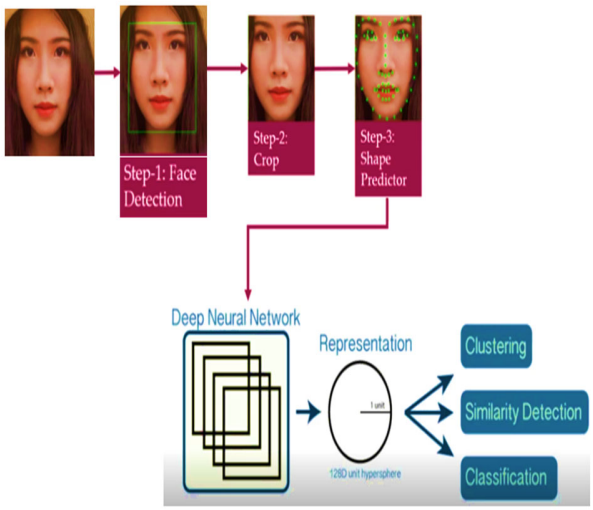

Figure-4 Flow used for extracting the features using DNN

```
[[-0.04844692  0.07973776 -0.09963042  0.05219208  0.02248172  0.20198324
   0.13852596 -0.06266069 -0.06219365  0.01123172  0.00230609  0.07717833
   0.01544594 -0.03280813 -0.0198783  -0.10874771 -0.09319499 -0.01268345
   0.10149115 -0.03454342  0.03998281 -0.14238802 -0.11665415 -0.01645241
   0.07435051  0.00277318 -0.17686342 -0.1490673   0.00644332  0.10620254
   0.08508859  0.08288751 -0.08299872 -0.03750012  0.04876548  0.05993147
  -0.07223055 -0.06693352  0.0433912  -0.07137258  0.14999217 -0.08112221
   0.04310593  0.04121759 -0.09240613 -0.03493492  0.11330134 -0.07407252
  -0.12562084 -0.14079288 -0.07809541 -0.01990954  0.07394529 -0.03580379
   0.0848316   0.13799967 -0.11584102  0.19711758 -0.0861319  -0.01615878
  -0.10057885  0.07192277  0.20695479 -0.216513    0.16872853  0.14802304
   0.04371198 -0.0987719  -0.14126581 -0.00213614  0.02870125  0.1082524
  -0.06683977  0.02168819  0.04144515  0.09735291 -0.01360501 -0.0795196
   0.00046234  0.01137843 -0.1640726  -0.00431227  0.03686861  0.03058025
  -0.09984109  0.01245098  0.08487714 -0.05692347  0.00904014 -0.03183766
   0.06037861 -0.09516985  0.03175969 -0.03873225 -0.01773375 -0.07007197
   0.07324421  0.02092896  0.02143688  0.08312549 -0.031659    0.05294375
   0.00321782 -0.0147114  -0.1374136   0.18237102  0.02833916  0.08900888
  -0.11359514  0.02033491  0.01212917 -0.07991205 -0.088426   -0.14238769
   0.07347389  0.01259384 -0.01749079 -0.0322057   0.14097632  0.10567141
   0.07477561 -0.177918   -0.01567982 -0.0292855  -0.01142321  0.15842962
   0.02104699  0.04293206]]
```

Figure-5 128 vectors obtained for a single image after feature extraction.

Once the vectors are obtained for a single image, the same can be applied to all the images in the folder to get complete vectors.

As of now the features of all images in a given folder are extracted. The further step in the proposed work is to compare the similarity of the given images. We want to know how similar the faces are, so we will compute the distance between the vector points i.e. the Euclidean distance. If the Euclidean distance is less than 0.6 the detected faces match with each other and if the distance is more than 0.6 it indicates that the faces are different.

Once the features extracted followed by the similarity check we need to train the model. Before training the model the data is loaded from a pickle file, in which the feature extracted are stored. Once the data is loaded, it is splited into dependent and independent followed by splitting the data into train and test set.

V Results and Discussion

The main Moto of face recognition method is to identify the face of a human present in the given image from database reference human face images. The archive reference face images are stored well in advance. But target scene image is passed at runtime. The co-ordination and mien of archive reference images may be very meticulous match to the given image faces. A noise from face region is removed and edge smoothing in performed in noise removal component.

The input images to the model can be an image with a single face, image with multiple faces, screen through webcam or a video clip. Face features boundary coordinates are extracted and stored well in the form of feature vectors. Feature extraction components identified like eyes, nose and mouth regions.

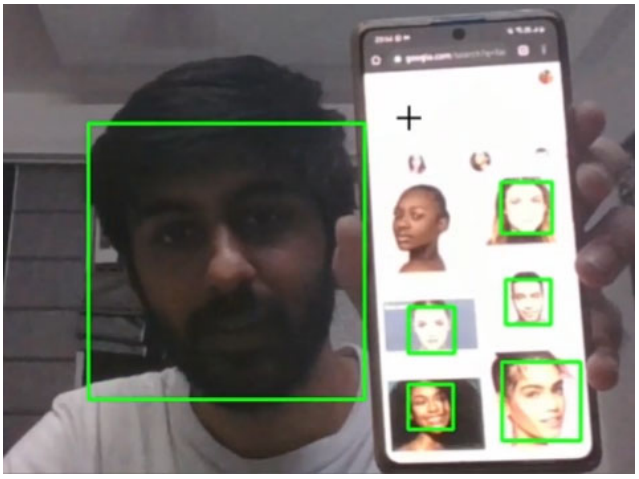

Figure-6 Face features boundary coordinates of screen capturing through webcam.

Figure-7 Feature extraction components identified like eyes, nose and mouth regions with multiple images**.**

Figure-8 Feature extraction components identified like eyes, nose and mouth regions with multiple images.

And finally based on the Euclidean distance measure, we can decide whether the given images are of the same person or are of different persons.

Figure-9 Comparison of images based on a Euclidean distance measure

In the proposed method analysis of the accuracy and F1 score of the open CV DNN face detector centered on a single-shot multi-box detector (SSD), Which is a framework using a ResNet-10 network are done. The below table shows the Comparative analysis of accuracy and F-1

Score for machine learning models with DNN face detector, which is based on single-shot multi-box detector (SSD).
F-1 Score is the Harmonic mean of Precision & Recall.

$$F1=\frac{2*Precision*Recall}{Precision+Recall}$$

Table 1. Comparative analysis of accuracy and F-1 score for machine learning models with DNN face detector

.

Model	Accuracy	F1-Score
Logistic Regression	0.74	0.74
Support Vector Classifier	0.84	0.84
Random Forest	1.00	1.00
DNN face detector using Single Shot Multi-Box Detector	0.94	0.94

The above output shows the object detection using the SSD model and Open-CV's DNNs. This shows how exactly SSD works with Open-CV's DNNs.

VI. CONCLUSION

The proposed face detection method is to detect person present in the given image from database reference persons face images. The database reference face images are stored well in advance. But target image is passed at runtime. The pre-processing component is to detect the face using deep neural networks which uses the models ResNet10 and Caffe for face detection and Torch for face descriptor instead of using Viola- Jones algorithm for detection. Face deduction using Viola Jones algorithm which is using harr features is having few disadvantages with the harr features. It may not work accurately for the most of the faces. After face detection cropping and shape predictor are applied. Feature extraction component identified and harmonized facial features like eyes, nose and mouth regions. Face feature's boundary coordinate are extracted and stored well in form of feature vector. Face similarity component checks face feature vector value of scene image with all database reference face images. The proposed work detects a faces using deep neural network algorithms using SSD which is framework of ResNet-10. The SSD algorithm along with multilayer convolution network to accomplish high accuracy in real time for detection of object or faces. The accuracy of the proposed model is around 94%. The presentation of our algorithm is good in still images and videos.

References

[1] C. Darwin "The Expression of the Emotions in Man and Animals" Oxford University Press, pp. 401-, ISBN 978-0-19-977197-4, Retrieved 4 August 2013.

[2] F Galton "Personal Identification and Description" Nature Journal, Paper Id 168, Vol. 38, pp. 173–177, 201-202, June 1888.

[3] A. Goldstein, L. Harmon, A. Lesk "Identification of Human Faces" in Proc. IEEE Conference on Computer Vision and Pattern Recognition, Vol. 59, No. 5, pp. 748–760, May 1971.

[4] W. Zhao, R. Chellappa, A. Rosenfeld, P. Phillips "Face recognition: a literature survey" ACM Computing Surveys, Vol. 35, No. 4, pp. 399–458, DOI 10.1145/954339.954342, December 2003.

[5] M. Turk, A. Pentland "Eigenfaces for Recognition" Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.

[6] S. Eickler, S. Muller "Recognition of JPEG compressed face images based on statistical methods" Image vision computing, Vol. 18, No. 4, pp. 279-287, https://doi.org/10.1016/S0262-8856(99)00055-4, 2007.

[7] Q. Chen, N. Georganas, E. Petriu "Real time vision based hand gesture recognition using Haar like features" Proceedings of IEEE Instrumentation and Measurement Technology Conference, Poland, INSPEC Number: 9717995, DOI: 10.1109/IMTC.2007.379068, ISSN: 1091-5281, 2007.

[8] R. Naik, K. Lad "Face Features Based Side View Face Recognition System" National Journal Of System And Information Technology, Vol. 9, No. 1, ISSN: 0974-3308, June 2016.

[9] Jalali, A., R. Mallipeddi, and M. Lee, Sensitive deep convolutional neural network for face recognition at large standoffs with small dataset. Expert Systems With Applications, 2017. 87: p. 304-315.

[10] Krizhevsky, A., I. Sutskever, and G. Hinton, ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017. 60(6): p. 84-90.

[11] Deng, J., et al. ImageNet: A large-scale hierarchical image database. in 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009.

[12] Nech, A. and I. Kemelmacher, Level Playing Field for Million Scale Face Recognition. 2017. 3406-3415.

[13] M. Parkhi, O., A. Vedaldi, and A. Zisserman, Deep Face Recognition. Vol. 1. 2015. 41.1- 41.12.

[14] Vidit Jain, E.L.-m., FDDB: A benchmark for face detection in unconstrained settings. 2010.

[15] X. Lu, A. Jain, D. Colbry "Matching 2.5d face scans to 3d models" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, No.1, pp.31-43, ISSN Information: PubMed ID: 16402617, INSPEC Accession Number: 8830239, DOI: 10.1109/TPAMI.2006.15 , 2006.