

Stream Video based Human Identification using Ai

Karthikeyan¹

¹Department of ECE,
Anna University, India

Alagu Raja B²

²Department of ECE,
Anna University, India

Hariharan C³

³Department of ECE,
Anna University, India

Abinesh G⁴

⁴Department of ECE,
Anna University, India

Abstract:- face recognition becomes most widely used application from Authentication. It is used in our hand mobiles such as unlocking .it is used as another form of biometric lock, where we use iris, finger print, palm scanner. face recognition has received much attention during past few years and in future also it becomes most useful and secure one. But it can also be fooled by using spoofing attacks such as photos (shown using by smart phones, framed pictures, printouts, masks). There are many ways to achieve it but no method gives a fighting fit solution to the problem, no one gives it accurately. We propose an effective method and also robust face detection scheme with Caffe based detection scheme ,and we propose our project in ARTIFICIAL INTELLIGENT specifically in deep learning.

Keywords:- Face Recognition, Face Detection, Caffe Detection, AI, Deep learning.

I. INTRODUCTION

As face detection application is done by storing the reference images at the initial stage from the form of video to picture. By storing large number of reference face we can enhance the Application in various domains. In this work we explore the problem of face capture in the frame of photo. Spoofing attacks upon face recognition system involvement of unauthorized person picture also been captured for security purpose, using these specification we can improve the security .in order to distinguish real face features from fake faces, liveness detection is a commonly used countermeasure. Automatic face recognition has attracted more attention in various access control applications , especially for mobile authentication. With the function of face unlocking functionality in android os and ios system. it becomes another biometric authentication technique. Unlike finger print authentication, face recognition does not need any additional sensors since all mobiles equipped with a front camera, and easier to acquire a person's face image than it is to acquire other biometric traits such as finger print, palm printand iris. The delicacy of face recognition system has motivated a number of studies on face detection. Published studies are limited in their scope training and testing phases used were are not given a robust solution for this face detection technique.

II.LITERATURE SURVEY

Now a days, Face recognition is the Emerging field of research in biometrics.For last few years research work has been done and discussed in this field. Face Recognition gathers vicious interest of all researchers in the era of biometrics. Review of literature goes further search for data and involves the recognition and connection between literature and our research.

Recently, multiple CovNets or deep CovNets have shown good results for face verification. According to Yi Sun et.al [1], existing methods generally address the problem of FR in two steps: feature extraction (design or learn features from each individual face image separately to acquire a better representation) and recognition (calculate similarity score between two compared faces using feature representation of each face [1]. For face recognition (FR), many approaches have been implemented earlier, like the use of neural networks [3,4,6], geometrical features, Eigen faces, template matching, and graph matching. CovNets has shown many promising results for FR [2,4,5,7,8,9]. Automatic feature extraction method using ratios of distances, presented by Kanade [7] used geometrical features and reported a recognition rate between 45-75% with a database of 20 people.

The approaches like self-organizing maps (SOM) and Karhunen-Loeve (KL) transform both can be used for dimensionality reduction, from which SOM proved to be an efficient algorithm [3]. Principal Component Analysis (PCA) has also been successfully implemented for same purpose [3, 4]. Though CovNets have shown promising results for FR, it remains still ambiguous to design a good CovNet architecture for a specific classification task due to the lack of theoretical guidance [3]. According to [6], CovNet-Restricted Boltzmann Machine (RBM) has shown 97.08% accuracy for matching two images of same person in unconstrained environment.

Brunelli and Poggio [8] computed a set of geometrical features such as nose width and length, mouth position and chin shape. They reported a recognition rate of 90% on a database of 47 people. However, they showed that a simple template matching scheme shows 100% recognition for the same database. Cox, *et al* [9] have introduced a mixture-distance technique which achieved a recognition rate of 95% using a query database of 95 images, where each face was represented by 30 manually extracted features.

By Pentland *et al*. [5,12] good results are reported on a large database (95% recognition of 200 people out of 3000). It is difficult to draw broad conclusions as many images of the same people looked very similar [10]. In [10], Lades *et al*. presented a dynamic link architecture for distortion invariant object recognition which employs elastic graph matching to find the closest stored graph. Sparse graphs whose vertices are labeled with a multi-resolution description in terms of a local power spectrum, and whose edges are labeled with geometrical distances. They presented good results with a database of 87 people and test images composed of different expressions and faces turned 15°. The matching process is computationally expensive, taking roughly 25s to compare with 87 stored objects when a parallel machine with 23 transputers is used. Thus, Eigen faces is a fast, simple, and practical

algorithm. However, it may be limited because optimal performance requires a high degree of correlation between the pixel intensities of the training and test images [3]. Graph matching is another approach to face recognition.

Wikott *et al.* [12] used an updated version of the technique and compare 300 faces against 300 different faces of the same people taken from the Face Recognition Technology (FERET) database. They report a recognition rate of 97.3%.

In constrained environments, hand-crafted features such as Local Binary Patterns (LBP) [16] and Local Phase Quantization (LPQ) [15,17] have received respectable performance in FR. However, the performance degrades dramatically when applied on images taken in unconstrained environments such as varying facial alignment, expression and illumination.

High-level recognition is typically modeled with many stages of processing as in Marr paradigm of processing from images to surfaces to three-dimensional (3D) models to matched models [10]. However, Turk and Pentland [18] argue that there is also a recognition process based on two-dimensional (2D) image processing. They presented a face recognition scheme in which face images are projected onto the principal components of the original set of training images. The resulting Eigen faces are classified by comparison with known individuals [3].

Use of haar cascade for extracting facial features and feeding them instead of raw pixel values helps in decreasing the complexity of neural network based recognition framework as the number of redundant input features has been decreased. Also the use of DNN instead of CovNets makes the process lighter and faster. Also, the accuracy is not compromised.[19]

None of the previous methods have used the idea of feeding only the extracted features into deep neural networks to accomplish the task of FR. The paper proposes the use of haar cascade (frontal face) for pre- processing the images which are then fed as input to deep neural networks for face recognition rather than directly passing the pixel values to CovNets.

III.METHODOLOGY

A Neural Network is human brain inspired algorithm designed to recognize pattern in numerical datasets. The real world data for example image, text audio, video etc; needs to be transformed into numerical vectors to use neural nets. A neural network is composed of different layers and a layer is made up of multiple nodes. Based on the type of pattern the neural network is trying to learn each input data fed into a node is assigned some weight. These weights determine the importance of the input data in producing the end result. The weighted sum of input data is calculated and depending on some threshold biases the output for the node is determined. The mapping of input to output is performed by some activation function.

The goal of a neural network is to approximate some function 'f'. Task of a simple classifier function $y = f(x)$, is to map the input data x to a class y , while the neural network identifies the parameter β , that results in best approximation function, $y = f(x)$.

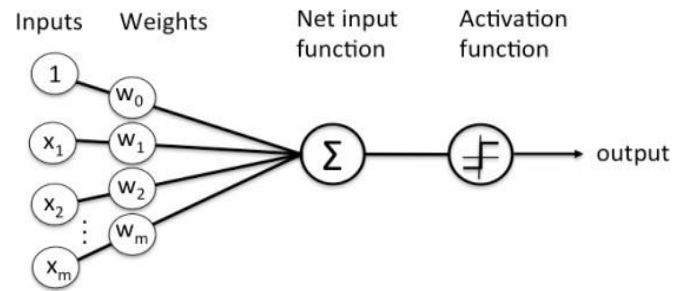


Fig.1. A neural node

A simple neural network is a network of such functions, that may be defined as $f(x) = f^3(f^2(f^1(x)))$. In the chain, f^1 is called the first layer, similarly f^2 is the second layer and so on. The length of this chain determines the depth of the neural network. Final layer is called the output layer. A schematic representation of a neural network is depicted in fig2. While training the desired output of each layer is not visible therefore the middle layers are called the hidden layer. A Deep Neural Network (DNN) is a feed forward Artificial Neural network (ANN), with multiple hidden layers and higher level of abstraction.

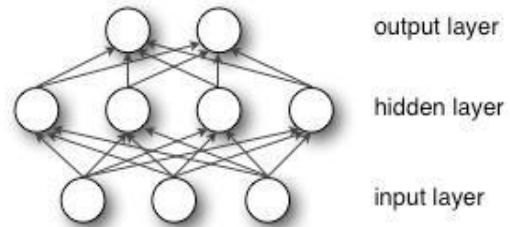


Fig.2. A Small Neural Network

The width of the DNN is determined from the dimensionality of the hidden layer. The hidden layer values are calculated through activation function. Learning in deep neural networks requires minimizing the cost function, like in case of classification cost function is the difference between actual label and the predicted label. Generally gradient descent is used for this purpose. In modern neural network, it is recommended to use Rectilinear Unit as activation function. A single hidden unit $h^{(i)}$ activation is given by

$$h^{(i)} = \sigma(w^{(i)T}x) \tag{1}$$

Where, σ is the tanh function $\sigma(x) = \frac{1}{1 + e^{-x}}$, with the weight vector for the i^{th} hidden unit, and x is the input. It gives a nonlinear transformation still it remains very close to the linearity making linear models to be easily optimized by Gradient Descent [13,20].

Generally limited data causes problem of over fitting in DNN. To avoid this dropout is used [6]. It randomly drops some nodes from the layers based on their probability. "dropping out" indicates temporarily removing units along with its incoming and outgoing edges. this is depicted in figure 3.

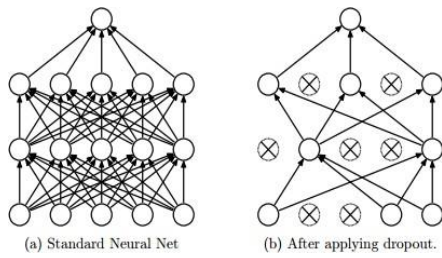


Fig.3.Solu.for Overfitting

IV.IMPLIMENTATION

In the working process of the face detection there some steps to predict the persons face in the live video. those are

1. Preprocessing stage
2. Learning stage
3. Recognition stage



Fig.4.Stage representation

IV.1. PREPROCESSING STAGE

preprocessing stage is the initial and fundamental step for recognition process through this stage reference images are stored , the input is given through in the form of video& picture. And the face of the person is captured as a picture, then it is stored in the database which is used to store the reference images this is done with the help of haar classifier. One of the main advantage of our project is we can create folder at the time of face capturing simply by given a specification in the command prompt.

IV.2. LEARNING STAGE

Learning stage is used to train the recognizer which is done with the help of the reference images taken through the stream video. At the time of face capture in the stream video. This is done with simply by giving opencv code in command prompt.

IV.3. RECOGNITION STAGE

Recognition stage is the final stage where we can predict the person in the stream video. If there is any unauthenticated person is in the video means then automatically it creates the database by capturing his image.

IV.4. Overall diagram

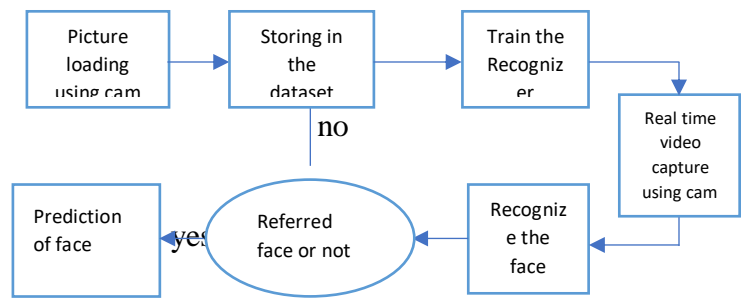


Fig.5.Stepbystep diagram

IV.5. RESULT & DISCUSSION

Proposed work is done in Haar cascade classifier which is used for the purpose of reducing the complexity in the neural network. We using the python 3.5.4 package and opencv 3 package is used for pre-processing using frontal face feature of haar cascade. Creation and training of neural network is done using keras, theano, and tensorflow . In command window we can the create the folder for reference (Fig.6) just by giving the name.

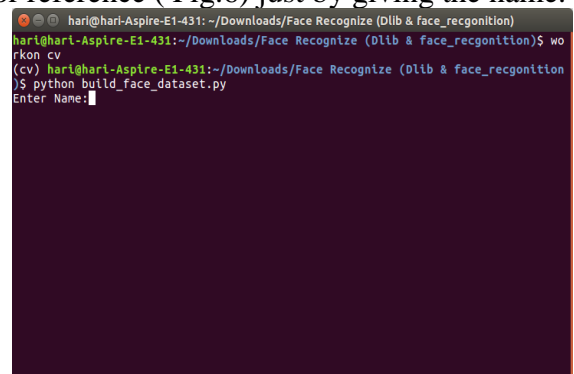


Fig.6.Folder Creation

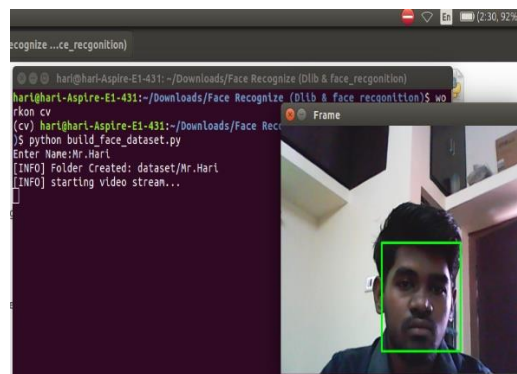


Fig.7.Live video stream

Then the capturing of images(Fig.8) for reference process is done cascade classifier, and the recognizer learning stage (Fig.9) is proposed to learn the recognizer which help the recognizer to predict the output in the live streaming video (Fig.10).

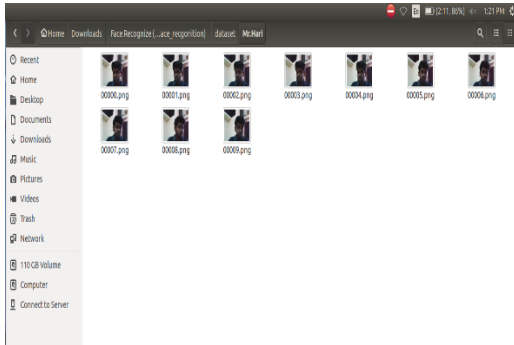


Fig.8.Dataset creation

```

1 # python extract_embeddings.py --dataset dataset \
2 --embeddings output/embeddings.pickle \
3 --detector face_detection_model \
4 --embedding-model openface_nn4_small12.v1.t7
5 [INFO] loading face detector...
6 [INFO] loading face recognizer...
7 [INFO] quantifying faces...
8 [INFO] processing image 1/18
9 [INFO] processing image 2/18
10 [INFO] processing image 3/18
11 [INFO] processing image 4/18
12 [INFO] processing image 5/18
13 [INFO] processing image 6/18
14 [INFO] processing image 7/18
15 [INFO] processing image 8/18
16 [INFO] processing image 9/18
17 [INFO] processing image 10/18
18 [INFO] processing image 11/18
19 [INFO] processing image 12/18
20 [INFO] processing image 13/18
21 [INFO] processing image 14/18
22 [INFO] processing image 15/18
23 [INFO] processing image 16/18
24 [INFO] processing image 17/18
25 [INFO] processing image 18/18
26 [INFO] serializing 18 encodings...
    
```

Fig.9.Learning

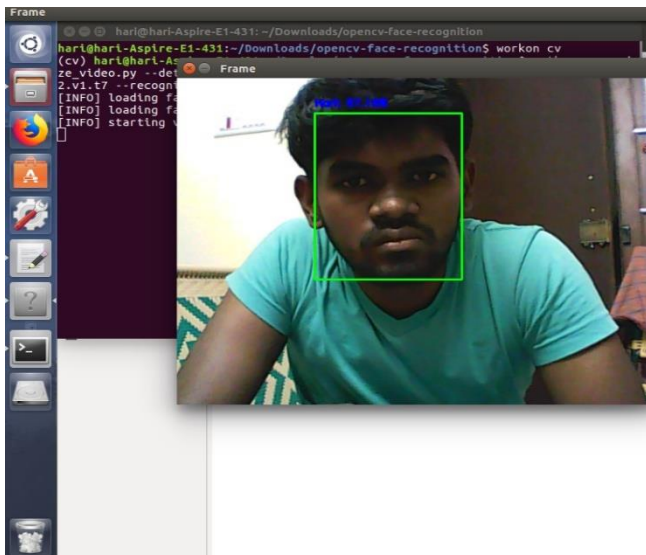


Fig.10.Face recognition

V.CONCLUSION

Use of haar cascade for extracting facial features and feeding them instead of raw pixel values helps in decreasing the complexity of neural network based recognition framework as the number of redundant input features has been decreased. Here we can capture the Unauthorized person face using our method. That must be very useful for surveillance purpose.

VI.FUTURE WORK

Though we increase no of reference faces we can enhance the project in various sector.

REFERENCES

- [1] Sun, Yi, Xiaogang Wang, and Xiaoou Tang. "Hybrid Deep Learning for Face Verification." IEEE Transactions on Pattern Analysis and Machine Intelligence 38.10 (2016): 1997-2009.
- [2] Hu, Guosheng, et al. "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition." Proceedings of the IEEE International Conference on Computer Vision Workshops. 2015.
- [3] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 2016.
- [4] Zhang, Tong, et al. "A deep neural network driven feature learning method for multi-view facial expression recognition." IEEE Trans. Multimed 99 (2016): 1.
- [5] Lawrence, Steve, et al. "Face recognition: A convolutional neural-network approach." IEEE transactions on neural networks 8.1 (1997): 98-113.
- [6] Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." Journal of Machine Learning Research 15.1 (2014): 1929-1958.
- [7] Kanade, Takeo. "Picture processing system by computer complex and recognition of human faces" Doctoral dissertation, Kyoto University 3952 (1973): 83-97.
- [8] Brunelli, Roberto, and Tomaso Poggio. "Face recognition: Features versus templates." IEEE transactions on pattern analysis and machine intelligence 15.10 (1993): 1042-1052.
- [9] Cox, Ingemar J., Joumana Ghosn, and Peter N. Yianilos. "Feature-based face recognition using mixture-distance." Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on. IEEE, 1996.
- [10] Ahonen, Timo, et al. "Recognition of blurred faces using local phase quantization." Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008.
- [11] Kohonen, Teuvo, and Self-Organizing Maps. "vol. 30." Berlin, Heidelberg, New York: Springer 1997 (1995): 2001.
- [12] Wiskott, Laurenz, et al. "Face recognition and gender determination." (1995): 92-97.
- [13] Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." Proc. ICML. Vol. 30. No. 1. 2013.
- [14] Chan, Chi Ho, et al. "Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors." IEEE Transactions on Pattern Analysis and Machine Intelligence 35.5 (2013): 1164-1177.
- [15] S. Moore and R. Bowden, Local binary patterns for multi-view facial expression recognition. Computer Vision and Image Understanding, 2011, pp. 541-558.
- [16] Ahonen, Timo, Abdenour Hadid, and Matti Pietikäinen. "Face recognition with local binary patterns." Computer vision-eccv 2004 (2004): 469-481.
- [17] M. Dahmane and J. Meunier, Emotion recognition using dynamic gridbased HoG features. Proc. IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, 2011, pp. 884- 888.

- [18] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." *Journal of cognitive neuroscience* 3.1 (1991): 71-86.
- [19] Dr. Priya Gupta, Nidhi Saxena, Meetika Sharma, Jagriti Tripathi. "Deep neural network for Human face recognition". *I.J.Engineering Manufacturing* 63-71; DOI:10.5815/ijem. 2018.01.06