

Face Detection and Recognition using Machine Learning Techniques

S. Vijayalakshmi¹, J. Uma Maheswari², K. Jananiye³

^{1,3}Department of CSE, PSG College of Technology, Coimbatore, India

²School of CSE, Vellore Institute of Technology, Chennai, India

E-mail: ¹svl.cse@psgtech.ac.in, ²umamaheswari.j@vit.ac.in, ³janani1180@gmail.com

Abstract

Face recognition of persons has received so much attention in the recent years due to its many applications in different fields such as security applications, video surveillance, biometric systems, identifying the criminals etc. This paper develops a system that can recognize the human face in the input image after it has been detected. The system is trained with set of faces and non faces, and when the input picture is given, the face is detected using Viola Jones Algorithm. In face recognition, the features are extracted from the training dataset using Principal Component Analysis (PCA) and then the system is trained to recognize the face using Support Vector Machine (SVM) classification. When the input image is given for face recognition, features are extracted from the input picture using PCA and multiclass classification is done by SVM.

Keywords: Multiclass classification, support vector machine, principal component analysis, viola jones algorithm

1. Introduction

Face detection and recognition has received a lot of attention in the recent years due to its numerous applications. Face detection searches for human face in the whole image and the output may be one or more than one face. The output will be a rectangle or rectangles on the faces in the image. Face recognition involves comparison of the input image. The input image is compared with all classes in face recognition system, and then the decision is made such that the output is a class name or an unknown face which is not present in the dataset. The dataset contains face images with different poses, different emotions etc., and also some non faces.

The proposed system consists of three steps, the first step is the training phase of face detection which involves training the model to identify whether the given image is face or not. Second step is the testing phase of face detection and if the input image is face then features are extracted using Principal Component Analysis (PCA) and the system will be trained with all the face images for recognition. Third step involves the testing phase of face recognition in which a test image is given and then face detection is done to check if the test image given is face or not using Viola Jones Algorithm (VJA) and if the test image is face then the Support Vector Machine (SVM) classifier says the classification label of the test image if that person's image is already trained.

Viola Jones algorithm involves four stages: 1. Haar feature extraction, 2. Creating an integral image, 3. AdaBoost training 4. Cascading classifier. In general, human faces share similar properties that can be matched by Haar feature extraction. Integral image creation is done for the efficient computation of the sum of the pixels in the rectangular region. AdaBoost is used for removing the irrelevant features. To eliminate the false candidate quickly, cascading of classifiers is done.

Liton Chandra Paul et al. [1] proposed that PCA is used for dimensionality reduction, that is feature extraction, so that only the significant features of the face are taken for face recognition. PCA uses eigen faces which are weighted eigen vectors. Eigen vectors are linear combination of every image in the training set. These weights obtained from PCA can be used for SVM classification. SVM is a binary classifier which has the classification labels +1 or -1 whereas face recognition is N class problem where N is the number of classes or persons. Therefore, a multiclass SVM classifier must be created from a regular 2 class SVM classifier. It is important to apply a bottom-up binary tree SVM classifier in which all the pertinent classification labels with features similar to those of the test face serve as the leaf nodes, and the nodes are then paired, and binary classification is carried out on each pair of them. Winning class labels will be considered in the next level. This process repeats until root node where one class is left which gives the classification label.

2. Literature Survey

Venkata Narayana et al. [2] proposed an SVM based face recognition technique. The work compared using PCA against using both PCA and SVM for face recognition and concluded that the combination yields better results. SVM-based face recognition with

occlusions was introduced by Hongjun et al. [3]. A standard that reduces the likelihood of overlap was established. How the global minimum of the error term is guaranteed under conditions and how the ensuing optimization issue is amenable to efficient solution were highlighted. Jonathon Phillips et al. [4] presented an SVM-based facial recognition technique. The formulation of the face recognition problem was as a difference space problem, which models the differences between two facial images. Differences between faces of the same individual and differences between the faces of different people are the classes.

A more advanced face recognition system employing PCA and artificial neural networks was described by Jindal Vikas Kumar in [5]. The non-human face images and the unidentified human faces that were absent from the input dataset of face images were discarded using a threshold value. In the research, face recognition tasks were carried out even for noisy faces with better accuracy and success rates. The PCA-based approach and the SVM-based techniques were compared on a challenging set of images. Monali Chaudhari et al. [6] conducted a review on face detection and study of Viola Jones method and concluded that Viola Jones method has faster face detection rate and increased computational efficiency.

Vijayalakshmi et al. [7] proposed biometric feature such as fingerprint with fuzzy commitment for user identification and verification. Face detection based on VJA was carried out by Mehul K. Dabhi et al.[8] A method of object detection that requires the least amount of computation time while still offering better accuracy of detection was offered. The method was used to develop a face detection that is around 15 times faster than any existing method. Vinoth Kumar B et al. [9] discussed various techniques that are available in the field of image segmentation and their pros and cons as well as Vinoth Kumar B et al. [10] used You Only Look Once (YOLO) with pre-trained weights for the detection of object in an image.

3. Proposed Work

The two phases of the proposed system are face detection and recognition. It goes through the following stages.

3.1 Face Detection using VJA

It is a very fast and perfect method of object detection. This algorithm is employed to find faces. This algorithm uses four steps to find faces, including the extraction of Haar features Integral image creation, AdaBoost (machine learning method), Cascade Classifiers.

3.1.1 Haar Features

To determine the variation in the image's dark and light intensity, Haar like features are used. The image uses a 24 x 24 sub-window. The Haar features consist of two rectangle features, three rectangle features, and four rectangle features. The difference between the sums of the pixels in the two rectangle sections represents the value of the two rectangle features. In a triangle with three rectangles, the value is the centre rectangle multiplied by the sum of the two rectangles on either side. The difference between the pixel sums in the diagonal pair of rectangles as illustrated in Figure 1 is the value of the four rectangle feature.

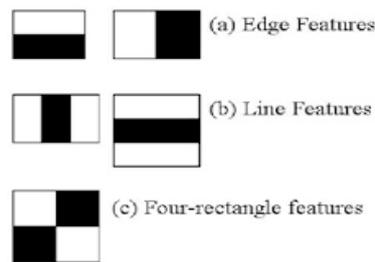


Figure 1. Haar Features

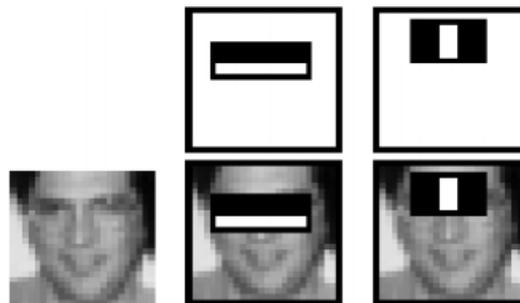


Figure 2. Applying Haar Features

In Figure 2, the first image is the original image, and the second image displays the Haar feature extraction using the contrast between the intensity of the upper cheeks and the eye region since the upper cheek region is lighter than the eye region. The last image compares the intensity in the eye region and across the bridge of the nose. The following formula can be used to compute the Haar like features:

$$\text{Value} = \text{sum (Pixels in white area)} - \text{sum (Pixels in black area)}$$

3.1.2 Integral image

The value at any location (x, y) of the integral image is the summation of the pixels in the original image above and to the left of location (x, y) . The calculation for adding

the pixels in the rectangle is simplified by integral picture construction, making it possible to do calculations more quickly. 1, 2, 3, and 4 are assumed to be the values of the integral image at the corner of a rectangle in Figure 3. The sum of the original image values inside the rectangle can be calculated using the equation below and any size of rectangle only requires three additions. The summation of original image values within the rectangle can be computed as, $\text{Sum} = 4 - 2 - 3 + 1$.

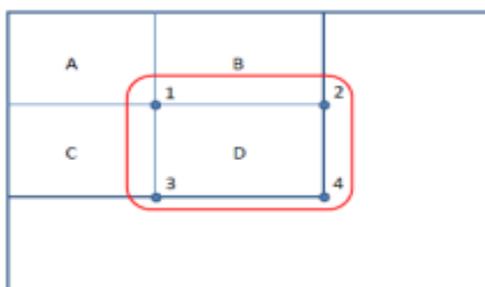


Figure 3. Integral Image

3.1.3 Adaboost

There are nearly 1,60,000 Haar features but not all these features are relevant to be applied on the face. AdaBoost is used to eliminate the irrelevant features and the selected features can be combined to form an effective classifier. Although it takes more time to train, this technique provides better accuracy and detection speed with only 1% false detection. Each feature selected by AdaBoost will be a weak classifier. Initially, given the set of weak classifiers, it iteratively combines those classifiers to form the linear combination of weak classifiers.

3.1.4 Cascade classifiers

To quickly eliminate the false candidate, a cascade classifier is used. If a candidate fails the first step, which requires that the candidate be a non-face image, the cascade eliminates the candidate. If it succeeds, it proceeds to a stage that is more difficult than the previous one. If a candidate advances past each stage, a face is discovered, as shown in Figure 4.

The complex classifiers should have low false positive rate and high detection. The overall False Positive rate (F) for the cascade is given below:

$$F = \prod_{i=1}^K f_i \quad (1)$$

Where, K is the stages of cascading, and f_i is the false positive rate in each stage.

Similarly, the overall Detection rate (D) is given as,

$$D = \prod_{i=1}^K d_i \quad (2)$$

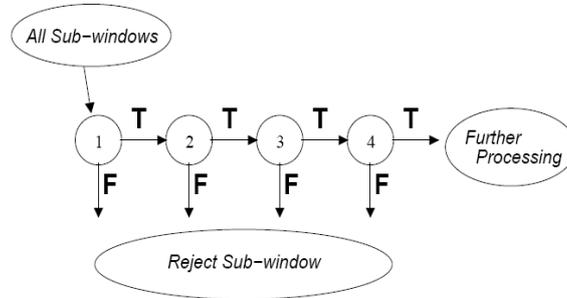


Figure 4. Stages of Cascade Classifier

For each stage, the objective is to have a very high detection rate ($\sim 100\%$), but a moderate false positive rate ($\sim 30\%$) is fine, in order to maintain a very low false positive rate F and a very high detection rate D .

3.2 Feature extraction using PCA

Feature extraction is relevant to dimension reduction. When the input data to an algorithm is too large to be processed or it is redundant, the number of dimensions in the input data can be reduced by selecting the more relevant features from the input data. PCA is a technique for reducing dimensionality. It is an easy and effective strategy. The advantages of PCA are its simplicity, speed, and potential for learning. The stages of the PCA are shown in Figure 5.

1. In the training phase, the training face dataset is obtained and the data is prepared by converting 2D face images to 1D face vectors.

2. The mean face image is found by considering the common facial features of all the faces in the training dataset.

3. The mean face image is subtracted from each of the face vector in order to obtain the normalized face vector.

4. The covariance matrix is found by $C=AA^T$, where A is the matrix consisting of M 1D face vectors of size N^2 , M is the number of images in the training dataset and $N \times N$ is the

size of each image in training dataset. Dimensionality reduction is used for the selection of K eigen vectors from the covariance matrix, which is tedious. The covariance matrix is calculated with reduced dimensionality, $C=AA^T$ of size $M \times M$.

5. Now, the eigen values and eigen vectors are determined from this matrix.

6. Only K eigen vectors are selected.

7. The lower dimensional K eigen vectors are converted to original face dimensionality which is given by, $u_i=Av_i$. Where u_i is the eigen vector in original face space and v_i is the eigen vector in lower dimensionality space.

8. Only K eigen vectors that are called as the eigen faces are selected. The remaining will have noise in it.

9. Similarly in testing phase, initially the feature extraction of test image is performed using the steps of PCA mentioned above and then the face is classified using SVM which gives the classification label of the test image if that face is recognized.

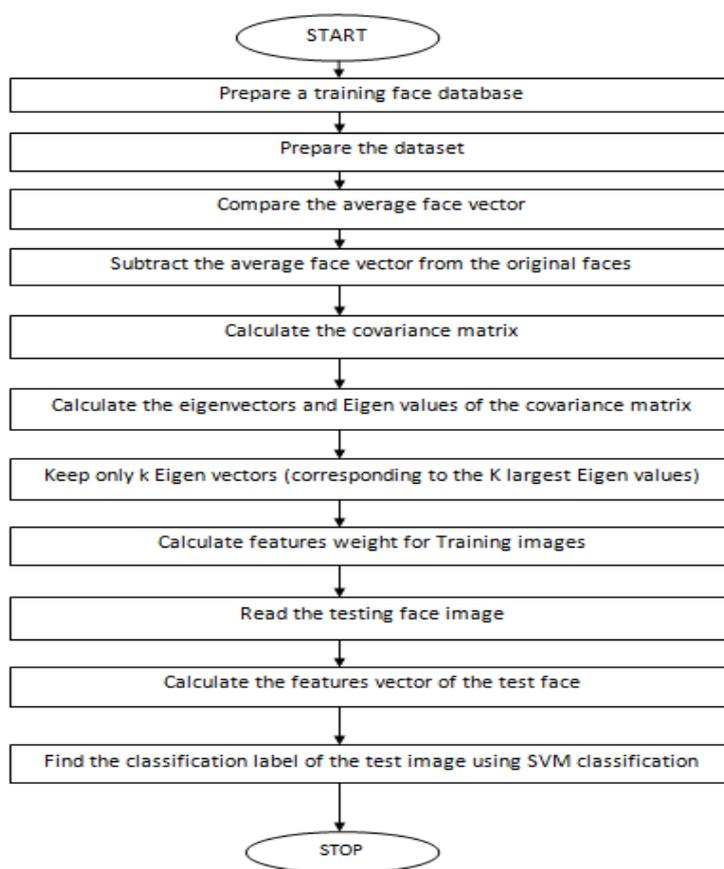


Figure 5. Steps of PCA

3.3 Face recognition using SVM classification

SVM is the supervised learning model which is used for classification. The goal of the SVM is to assign the new objects a specific classification label. This can be achieved by linear partition of the features using hyperplane. Based on the features in the new objects, it places an object either above or below the hyperplane. This hyper-plane helps to classify the objects based on the features. Several hyperplanes may be present, but the hyperplane that maximizes the margin must be chosen as shown in Figure 6. The reason for finding the optimal hyperplane with higher margin is to avoid misclassification due to the low margin. SVM provides two classification labels +1 and -1 above and below the hyperplane. Based on the features in the new objects, it places the object either above or below the hyperplane and returns the classification label +1 or -1 based on the classification label to which the object belongs to.

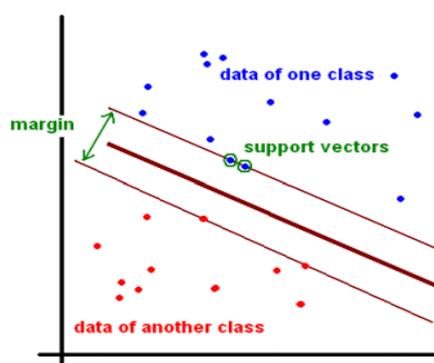


Figure 6. Optimal Hyper-plane

SVM is a binary classifier. Face recognition is multiclass problem. So the two class SVM classifier has to be mapped to multiclass SVM classifier. The relevant list and the irrelevant list for the given test face will be maintained based on the similarity in the features between the test face and the training set. The relevant list contains the faces of the persons who have similar features to that of the test face and the irrelevant list contains the faces of persons who don't have any similarity to the given test face. The relevant list of people's categorization labels will be placed in all of the leaf nodes of a bottom-up binary tree, and each pair will then undergo SVM binary classification. The level of the winning class labels will be increased. This cycle repeats until there is only one class left, which provides the classification label for the test face at the root node. The multiclass SVM classifier is shown in Figure 7. Relevant list and irrelevant list are maintained because if all the classification labels are considered in the leaf level, then there will be more number of leaf nodes in binary

tree and SVM classification will be done even for the faces that has no relevance to the test face which consumes time and cost. Therefore, the relevant list is obtained and then the binary tree is constructed.

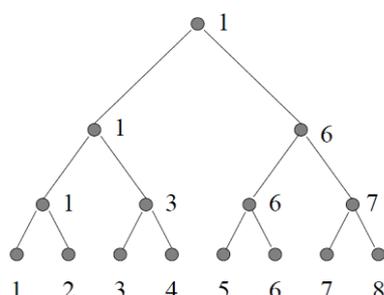


Figure 7. Bottom up Binary Tree SVM Classifier

It is considered that there are eight persons who have similar face to that of the test face for classification. The multiclass SVM classifier classifies the given person1 from eight persons and performs binary classification by pairing the two persons. The person1 and person2 are paired and the person1 who is mostly relevant to the given input is classified and that person is taken to the next level. Similarly, the other persons are also classified based on some relevant features and given to the next level. This process is repeated until the classification reaches the root node which classifies the person1 and then returns the result as in which classification label it belongs to.

4. Results and Discussion

Some animal faces having similar face to that of the human faces are used in non face dataset and also some other non faces to check for face detection performance. AT&T face dataset which contains 40 persons having 10 different facial expressions for each person is used to check the face recognition performance. Training set contains 7 faces for each person and testing set contains 3 faces for each person. Totally AT&T contains 400 faces and in that 280 faces are taken for training and 120 faces for testing. Figure 8 and Figure 9 show the Face Detection phase and Face Recognition phase respectively.



Figure 8. Face Detection

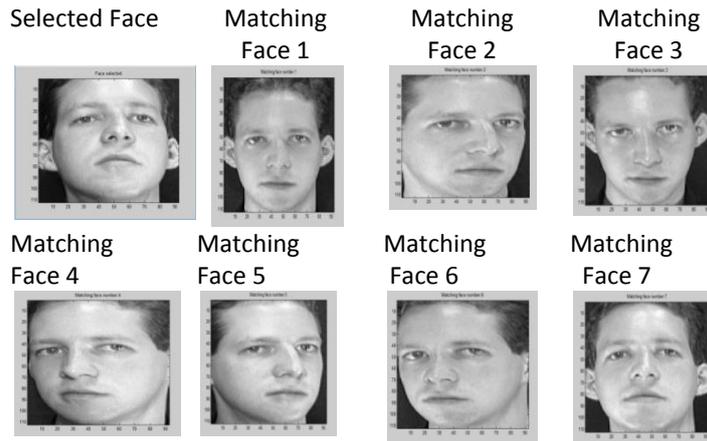


Figure 9. Face Recognition

In Table 1, Figure 10 and Figure 11, it is demonstrated the proposed system obtained 90% true positive rate (sensitivity) and 10% true negative rate (specificity) for face detection followed by face recognition which has 85% true positive rate and 15% true negative rate.

Table 1. Face detection and Face recognition rate (%)

Training Samples 40	Accuracy (%)
Face detection	90%
Face recognition	85%

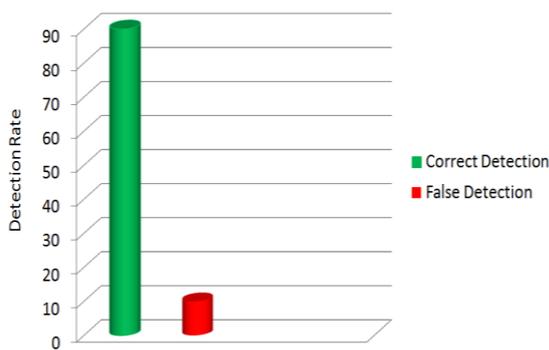


Figure 10. Face Detection Rate

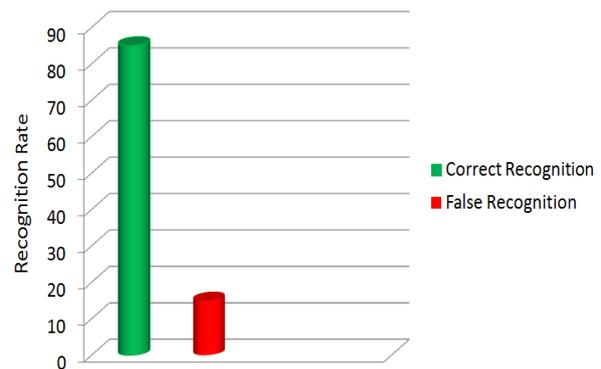


Figure 11. Face Recognition Rate

5. Conclusion

The proposed system incorporates the Viola Jones Algorithm for face detection because of its higher detection rate. In this research, the Principal Component Analysis technique has been used because of its unsupervised learning nature, which produces better

results over time, and SVM has been used due to its greater recognition rate. The system has been implemented using AT&T face dataset. The system can also take dynamic face data by capturing the different expressions of the user through web camera. Using the above mentioned techniques, the face detection and recognition system is successfully implemented and the system achieves good face detection and recognition rates such as 90% and 85% respectively.

References

- [1] Liton Chandra Paul, Abdulla Al Suman. "Face Recognition using Principal Component Analysis Method," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 1, no.09 (2012), 135-139.
- [2] Venkata Narayana, K., Manoj, V.V.R., Swathi, K. "Enhanced Face Recognition based on PCA and SVM." International Journal of Computer Applications, 7, no.02 (2015): 40-42.
- [3] Hongjun Jia & Aleix, M, Martinez. "Support Vector Machines in Face Recognition with Occlusions, Available": <http://www2.ece.ohio-state.edu/~aleix/CVPR09.pdf>. [Accessed: 07- Oct- 2022].
- [4] Jonathon Phillips P. "Support Vector Machines Applied to Face Recognition," National Institute of Standards and Technology, <https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir6241.pdf>. [Accessed: 17- Aug- 2022].
- [5] Jindal Vikas Kumar, N.(2013). Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks. International Journal of Advanced Research in Computer Science and Software Engineering, 3. no.06(2013)
- [6] Monali Chaudhari, Shanta sondur, Gauresh Vanjare."A review on Face Detection and study of Viola Jones method." International Journal of Computer Trends and Technology (IJCTT), 25. No.01((2015):54-61.
- [7] Vijayalakshmi, S., Karpagam, GR."Authentication as a service in cloud from a fuzzy perspective." International Journal of Enterprise Network Management," 9. no.03(2018):352-362,
- [8] Mehul, K., Dabhi, Bhavna, K., Pancholi."Face Detection System Based on Viola - Jones Algorithm." International Journal of Science and Research (IJSR), 5. no.04(2016):62-64.

- [9] Vinoth Kumar, B., Sabareeswaran, S., Madumitha, G. “A Decennary Survey on Artificial Intelligence Methods for Image Segmentation.” *Advanced Engineering Optimization Through Intelligent Techniques*, (2020). 291-311.
- [10] Vinoth Kumar, B., Abirami, S., Bharathi Lakshmi, R. J., Lohitha, R., Udhaya, R. B. “Detection and Content Retrieval of Object in an Image using YOLO.” *International Conference on Frontiers in Materials and Smart System Technologies*. (2019).