Analysis of Training Parameters for Classifiers Based on Haar-like Features to Detect Human Faces

Supratim Gupta
Department of Electrical Engineering
National Institute of Technology, Rourkela
Odisha - 769008
sgupta.iitkgp@gmail.com

Anirban Dasgupta, Aurobinda Routray
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur
West Bengal - 721302
anirban1828@gmail.com,aurobinda.routray@gmail.com

Abstract— This paper analyzes the performance of the Haar-like feature based classifier for detection of face with fewer features. The lower dimensional feature space representation of the image may reduce the computational burden compromising the accuracy in detection of faces with varying orientations. In this work we train the classifier with positive instances of different orientations under such feature constraint. The training parameters like maximum deviation and maximum angle are varied to form different classifiers. Experimental results show optimum values of the design parameters can produce good performance of the classifier to detect frontal as well as tilted human faces.

Keywords- Face Detection, Haar-like Feature, Classifier's Performance

I. INTRODUCTION

Face detection has become essential in many computer vision applications. Haar-like features have been widely used for robust and fast detection of faces from an image or a video or from a camera in real-time. It classifies the whole image into face and non-face categories using a supervised learning method. The performance of this method is found to be reasonably well for detection of frontal faces using Gentle AdaBoost algorithm [1]–[4]. However it has limited accuracy in the detection of faces with in-plane and off-plane rotation. Several other approaches [5]–[7] were made for detection of such faces but these methods used complex algorithms whose computations were complex. Lienhart et al. have introduced the concept of a tilted (45°) Haar-like features in [6] to improve the accuracy in detection of objects in the images. This method was found to be successful in a number of cases. Messom et al. had suggested to include the off-plane Haar-like features [7]. Increase of these features raises the computational burden in terms of both memory and time of execution.

The objective of this paper is to analyze the detection performance of the algorithm considering a few Haar-like features and including varieties of human faces with different orientations. Towards this goal, the issues related to the selection of design parameter values to improve the detection rate are critically reviewed and tested. Experiment is carried out to record faces with different orientations under varying illumination condition in laboratory. Five classifiers are developed by varying two design parameters (a) maximum angles and (b) maximum deviation during the training process. These classifiers are tested on 8 test image sets generated from the recorded images different from the training set. Receiver Operating Characteristics (ROC) curves are generated for each of the classifiers and area under the curve (AUC) of these ROC's are computed to analyze their performances.

This paper is organized as follows. Section II discusses the training process of the Haar-like features. Section III describes the experimentation for the classifiers with different design parameters and presents a comparative judgment of their performances by computing the ROC curves and the respective AUC. Section IV explains the results obtained in the previous section. Finally, the work is concluded in Section V.

II. TRAINING OF HAAR-LIKE FEATURES FOR FACE DETECTION

A. *Haar-like Features*

Haar-like features are generally used to detect and recognize objects [1]. These are named as Haar-like because of their similarity with Haar wavelets. A Haar-like feature considers the adjacent rectangular regions at a specific location within a detection window. Then the pixel intensities in these regions are summed. Finally it calculates the difference between these regions. The difference is used to place different subsections of an image into different categories.

B. *Training Steps*

Training of Haar-like feature based classifiers includes several steps as described in [8]. It is required to collect positive images that contain only objects of interest (faces in the present context) as well as the negative images which is devoid of the object of interest to build the classifier. The following steps are taken to obtain the final classifier in XML format using Intel's Open Computer Vision (OpenCV) version 1.0 library.

1) Collection of training images containing faces in different orientation under varying illumination.

2) Graphical User Interfaces (GUI) in Matlab has been developed to speed up such operations like cropping images to obtain the training data.
3) Positive and negative images were generated along with corresponding information texts files like ‘positive.txt’ and ‘negative.txt’ in the Matlab GUI. The text files contain the coordinates of the positive and negative training samples along with their names in a specified order. The name of the image file contains the following format.

   < File Name > _< x > _ < y >_ < width _> <height > .jpg

   where x, y, width, and height define the object bounding rectangle.
4) A vector (.vec) file is obtained from the text files generated in the previous step. The vector file contains compact information of positive instances of objects and the negative image such as background.
5) The classifiers are trained using the vector file.
6) Once the Haar training was complete, the output was stored in the specified directory in the form of text files. The final step of the process was to convert these files into a single xml file. This xml file was the final classifier.

III. EXPERIMENT AND RESULTS

A. *Experiment Design*

The objective of the experiment is to analyze the effect of variation in training parameters on the detection of faces in images. Facial images of 13 subjects were recorded under laboratory condition with varying orientations of faces for the purpose. The videos were saved in Audio Video Interleave (AVI) format at 30 fps. The training and testing data sets were obtained from these videos.

1) Training Data Set: The frames for training of the classifier were extracted from the recorded videos using an application software “FreeVideoToJPGConverter” version 1.8. The training database is formed from these extracted frames chosen randomly. A few example images are shown in Fig. 1.

Fig. 1. Some training images

2) Testing Data Set: The test data set has also been formed from the earlier extracted frames excluding the images contained in the training database. Some non-face images have been added with this test data set. All the images in this database are kept at same resolutions. Subsequently, this data set is divided into 8 groups corresponding to 8 subjects. Each of these groups contains 115 images of both faces and non-faces. Each of these images within a group is manually identified and noted for computation of the ROC curve for each classifier described in Section III.B.

3) Training Parameters: Table I shows the different training commands used to train a Haar-like feature. It also indicates the respective values used in the present work.

B. *Classifier Design and Generation*

The classifiers are formed using the conventional Haar training process. Five classifiers are developed by varying the maximum deviations and the maximum x angle, maximum y angle and maximum z angle. The classifiers formed are named according to the format ‘face maxdeviation maximumangle.xml’. An Intel i5 CORE processor 2.53 GHz, 4 GB RAM has been used for the experiment. Open CV 1.0 has been used for the entire training process. The classifiers are defined in Table II.

C. *ROC Curve and AUC Calculation*

An ROC curve provides a mean of visualizing the performance of a machine learning algorithm like the classifiers considered in this work [9]. This is a mapping between true positive rate (TPR) and false positive rate (FPR). The TPR and FPR are obtained by constructing a confusion matrix as shown in Table III. The points on the ROC curve computed from the test database are interpolated to construct a smooth average version of it for each of the classifier.

The TPR (hit rate) and FPR (false alarm rate) can be related to the elements in the confusion matrix by the following equations 1 and 2 respectively.

$$tpr = \frac{tp}{tp+fn} \tag{1}$$

where tpr = true positive rate, tp = number of true positives, fn = number of false negatives.

$$fpr = \frac{fp}{fp+tn} \tag{2}$$

Where fpr = false positive rate, fp = number of false positives, tn = number of true negatives

A MATLAB GUI is developed to detect the face using Haar-like features and the FPR and TPR for each classifier from the test database. MATLAB 2009b is used for testing the performance of each classifier. The ROC curves are plotted in MATLAB using the ’cftool’ (Curve Fitting Tool). Shape Preserving Interpolant curve fitting was used to plot the curve. The integral of the curve produce the Area Under the Curve (AUC) which indicates the accuracy of the classifier.

TABLE I

PARAMETERS AND THEIR DESCRIPTION

Commands	Descriptions
mem.	The memory assigned for the training process to run in MB. We have used 1024 in our case.
vec.	Location of the input .vec file.
bg.	This is the location of the txt file containing the list of background images.
h., w.	It is the height and width that are used to create the .vec file. We have used 24 for both the h and w.
nstages.	It is the number of stages (number of strong classifiers). We have used 20 in our case.
nosym.	This is only included if the object is not symmetric. We have used a non-symmetric case.
minhitrate.	This is the minimum hit rate for a stage/strong classifier. If the minhitrate = 1, you impose no false negatives in the training data. We have used 0.999 for training
maxangle	This is the maximum angle of tilt from vertical in a specified direction up to which the classifier can detect the object. We have trained classifiers for angles of 30, 60 and 90 degrees.
maxdeviation	This is the maximum deviation of intensity from the training set which the classifier is trained to detect. We have trained classifiers for maximum deviations of 80, 100 and 120 units.
maxfalsealarm.	Maximum false positive rate for each stage classifier. In each stage, features are added and the false positive rate is decreased till the max false rate is satisfied. We have used 0.5 for this parameter.
npos.	Number of faces in the .vec file. Our case had 588.
nneg.	Number of negative samples in the set. In our case it was just the background images and hence it was just the same number, i.e. 588 as the positive samples.

TABLE II

CLASSIFIERS WITH DIFFERENT TRAINING PARAMETERS

Classifier Name	Max Angle	Max Deviation
Classifier 1	30	100
Classifier 2	45	80
Classifier 3	45	100
Classifier 4	45	120
Classifier 5	60	100

Fig. 2. Some examples of true positives

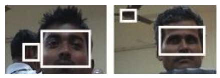

Fig. 3. Some examples of false positives along with true positive

TABLE III

CONFUSION MATRIX

		True Class: p	True Class: n
Hypothesized Class	Y	True Positive	False Positive
	N	False Negative	True Negative

TABLE IV

AREA UNDER THE ROC CURVE FOR DIFFERENT CLASSIFIER

Classifier	AUC
Classifier 1	0.8362
Classifier 2	0.8167
Classifier 3	0.8447
Classifier 4	0.834
Classifier 5	0.8249

D. Results

The program for face detection is run and the number of true positives, false positives, false negatives and true negatives are computed. Fig. 2 and Fig. 3 show some detection results using Classifier 1.

1) ROC curves for different classifiers: Fig. 4 shows the comparative visualization of different ROC curves shown together.

2) AUC for each ROC: The AUC for each ROC were obtained by integrating the curves from limits 0 to 1. Table IV shows the different AUC values for different classifiers.

IV. DISCUSSION

The AUC of each classifier is above 80% which indicates that the classifiers have a high hit rate with good accuracy and low false alarming. Classifier 3 gives the highest AUC which is an indication that it is the best among these 5 classifiers. The mean of the AUC found to be 0.8313 with a standard deviation of ± 0.0108. The mean value shows that all the classifiers perform nicely on an average basis. It can be observed from Table II that in Classifiers 1, 3, & 5, only maximum angle (MA) is varied keeping the maximum

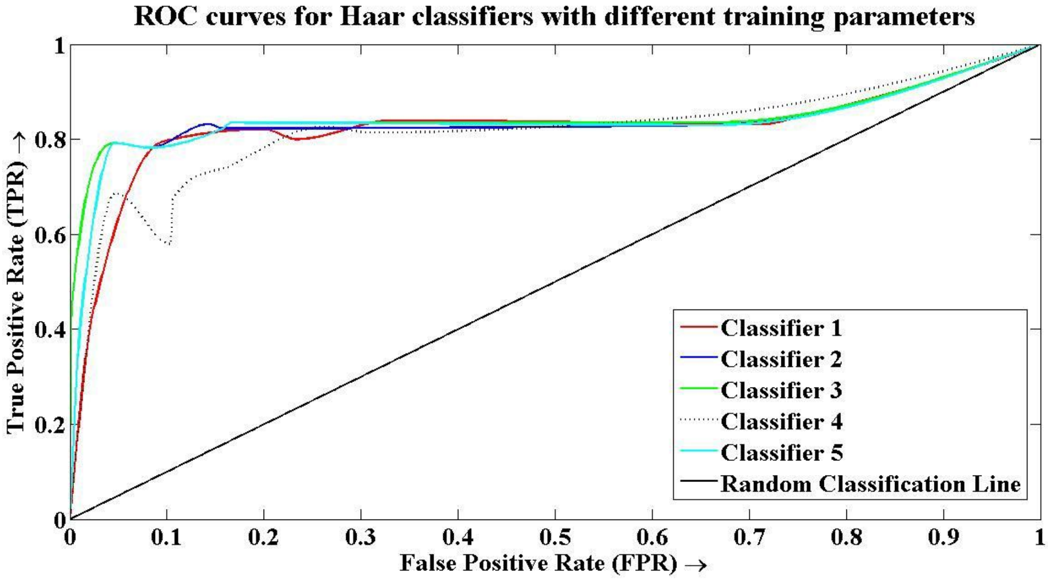

Fig. 4. Comparative study of ROC curves for Haar classifiers with different training parameters

deviation (MD) constant while MD is varied with fixed MA for Classifiers 2, 3, & 4. In both the cases, it is observed that the performance of the classifiers reaches maximum with a moderate value of either MA or MD as displayed in Table IV. This reveals that we cannot choose the MA and MD arbitrarily high or low. Albeit, a moderate values of them may produce classifier with better performance.

V. CONCLUSION

The face detector based on the Haar-like feature-based classification method has been analyzed. The classifier detects the frontal faces with high accuracy. The researchers have introduced different extra features to improve detection of faces with other orientations (like in-plane and off-plane rotated faces). Such increment in features introduces higher computational burden. In this paper we have analyzed the effect of variation of two training parameters on classifier's performance. The ROC curves for each of the classifiers indicates that their performances are more than 80%. It has been observed that the best performance among the classifiers is achieved at moderate values of the maximum deviation and maximum angle. This indicates that optimum values of the design parameters can train a classifier to provide good performance for detection frontal as well as tilted human faces.

REFERENCES

[1] P. Viola, and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", *Conference on Computer Vision And Pattern Recognition*, 2001.

[2] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *in Computational Learning Theory: Eurocolt 95*, pp. 23-37, Springer-Verlag, 1995.

[3] P. Viola, and M. Jones, "Robust Real-time Object Detection", 2nd Int. Workshop on Statistical and Computational Theories of Vision Modeling, Learning, Computing, and Sampling.

[4] M. Jones, and P. Viola, "Fast Multi-view Face Detection",*Mitsubiushi Electric Research Laboratories*, TR2003-96, Aug. 2003

[5] H. A. Rowley, "Neural Network-Based Face Detection", *CMU-CS-99-117*, May 1999

[6] R. Lienhart, and J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection", *ICIP02*, pp. I:900-903, 2002

[7] C. H. Messom, and A. L. C. Barczak, "Fast and Efficient Rotated Haar-like Features Using Rotated Integral Images", *Australian Conference on Robotics and Automation (ACRA2006)*, pp. 1-6, 2006

[8] N. Seo, "Tutorial: OpenCV Haar training (Rapid Object Detection With A Cascade of Boosted Classifiers Based on Haar-like Features)" *available on* **http://note.sonots.com/scisoftware/haartraining.html**

[9] T. Fawett, "ROC Graphs: Notes and Practical Considerations for Researchers", Mar. 2004, *available on* **http://home.comcast.net /~tomfawcett/public_html/papers/ROC101.pdf**