Classification of Human Faces and Non Faces Using Machine Learning Techniques

Lal Hussain
University of Azad Jammu and Kashmir, Quality Enhancement Cell/ DCS &IT, Muzaffarabad, Pakistan
Email: lall_hussain2008@live.com

Wajid Aziz, Zaki H. Kazmi, and Imtiaz A. Awan
University of Azad Jammu and Kashmir, DCS & IT, Muzaffarabad, Pakistan
Email: kh_wajid@yahoo.com, zaki.mzd@gmail.com, thisisawan@gmail.com

Abstract—**Face detection technique is used for face authentication and verification and face detection is a front part of face recognition. It is used in many fields such as authentication security, video surveillance and human interaction system. In this paper we have collected data of 400 faces from school students in Muzaffarabad, Azad Kashmir. Besides, 50 non-faces are also collected. Both faces and non-faces are preprocessed using Background Elimination, Noise Reduction, Width Normalization and Thinning. After the preprocessing, we have extracted features from 400 faces and 50 non-faces including Geometric Features such as Image Cropping, Vertical/Horizontal Projection, Global Features such as Aspect Ratio, Normalized Area of Faces and Non-faces, Center of Gravity, Slope of Line joining the center of Gravity and texture features. Finally, we have applied Machine Learning Methods such as Bayes, Function, Lazy, Meta, Misc, Rules and Tree to classify the faces and non-faces using 10 fold cross validation. HyperPipes gives an overall higher accuracy of 99.8%, while ADTree, LWL and LogiBoost gives accuracy of more than 99%. The average AUC of ROC value was calculated as 96.08%.**

***Index Terms*—classification, receiver operating curve, feature extraction, preprocessing, cross validation**

I. Introduction

Face recognition is a technique that is used to identify a person from his /her digital image .It is helpful in daily life such as for security access, control systems, content based indexing and bank teller machines. In face recognition, feature based approaches are used. [1]

Various approaches have been used in classifying and recognizing faces including principles component analysis, local feature extraction, neural networks comparative analysis and radial basis function. Face detection is front end of face recognition. It locates and segments face regions from cluttered images, either obtained from video or still image. [2]

The Principal Component Analysis (PCA) is one of the mathematical techniques that have been used in image recognition and compression. The jobs which PCA can do are prediction, redundancy removal, feature extraction, data compression etc. Because PCA is a classical technique which can do something in the linear domain applications having linear models are suitable, such as signal processing, image processing system and control theory, communications, etc. [3]

From many years lots of work on Face Detection and Recognition has been carried out as it does not need human cooperation. We have dataset of Face images after Detection framed faces are formed from which removed background then extracted faces are obtained. Preprocessing is also performed then we will trained the dataset for which we use training classifiers and then we recognize the face [4]

Facial images are essential for intelligent based human computer interaction and it does not need the human cooperation. Many techniques are used for face detection from a single image. When a face region is extracted in preprocessing then localization is done. In preprocessing of image illumination to determine specific features and image size then localized image is matched with database by using matching algorithms. [5]

Over the last few decade lots of work is been done in face detection and recognition. Since lots of methods are introduced for detection and recognition which considered as a milestone. [6]

Face recognition has acquired considerable attention from both your own computer vision and also value processing. The interest can be motivated from applications ranging from static matching of controlled photographs just as in mug shot matching in addition to verification in order to surveillance video images. Your first step throughout automated face recognition is face detection in which Metropolis along width size of each face will be determined. The reliability has an major influence to the performance and usability of the whole face id system. [7]

To produce fully automated systems, robust and efficient face identification algorithms usually are required. Your own face can be detected immediately after a person’s face comes into a good view right after a face will be detected, the face region will be cropped from the visual to provide As “Probe” into your current knowledge to check on for possible matches. Ones face

Manuscript received November 1, 2013; revised February 11, 2014.

doi: 10.12720/ijeee.2.2.116-123

visible is usually pre-processed regarding items like aesthetic size as well as illumination and for you to detect Particular possesses. The details by the graphic are then matched against your knowledge. Your own matching algorithm will probably produce a similarity measure to its match of the probe face into your own knowledge [8].

Training the neural network for its face i.e. Employment is difficult from the difficulty in characterizing prototypical “nonface” images. Unlike face recognition, which the classes in order to be discriminated are various other faces, you’re a couple of classes in order to end up being discriminated throughout face recognition usually are “images containing faces” in addition to “images not containing faces”. The idea is simple to obtain a representative sample of images that contain faces, but much harder for getting a great representative sample of the person which do not. [9].

II. Data Acquisition

In the present work we have taken faces from 400 students of a school from both male and female aging between 12 to 20 years. 10 images are taken using digital camera of 12 mega pixel from each student. Besides, we have taken 50 nonface images from environment.

III. Proposed Methodologies

The process of recognition is the identification of something already known or acknowledgement of something as valid the state or quality of being recognized or acknowledged. It is broadly divided into two phases, identification of object and verification of the object. In proposed system the object is signature which we will get from the scanned image.

A. Preprocessing

The preprocessing step is applied on scanned gray faces. The purpose in this phase is to make faces standard and ready for feature extraction. These stages include the four steps: Background elimination, noise reduction, width normalization and skeletonization [10].

Background Elimination

Data area cropping must be done for extracting features. P-tile thresholding is applied to capture faces from the background. After the thresholding the pixels of the faces would be “1” and the other pixels which belong to the back-ground would be “0”.

Noise Reduction

A noise reduction filter is applied to the binary image for eliminating single black pixels on white background. 8-neighbors of a chosen pixel are examined. In this case the back pixel greater than the number of white pixel, the chosen pixel will be black otherwise it will be white.

Width Normalization

Face dimensions may have intrapersonal and interpersonal differences. So the image width is adjusted to a default value and the height will change without any change on height-to-width ratio. The width normalization is adjusted to 100 after normalizing the width.

Thinning

The purpose of thinning is to eliminate the thickness differences of pen. The image is made one pixel think using Hilditch’s Algorithm.

B. Feature Extraction Methods

Feature extraction is essential classifying the face detection. Before classification, we extracted the features of faces and non-faces scanned images such as normalized area of faces and non-faces, accept ration, center of gravity, slop of the line joining the centers of gravity, cropping maximum horizontal projection, maximum vertical projection, edge detection and texture features using Matlab. The features are extracted against each face and non-face. We have used of 40x50,80x50, 100x50,120x50,160x50,200x50,240x50,280x50,320x50,360x50 and 400x50 (faces vs non-faces) datasets with 27 features extracted using Matlab and prepared data in arff format for processing for classification using Weka classifier.

C. Classification

For structural activity relationship analysis, we have used Weka software for classification. The above data prepared in ARFF format was then processed for classification. We have applied seven classification methods on faces and non faces datasets such as Bayes, Function, Lazy, Meta, Misc, Rules and Tree. The classification performance tested for:

Bayes methods includes Bayesian Logistic Regression (BLR), Bayes Net (BN), Complement Naïve Bayes (CNB), DMNB Text, Naïve Bayes(NB), Naïve Bayes Multinomial (NBMN), Naïve Bayes Multinomial Updateable (NBMNU), Naïve Bayes Simple (NBS), Naïve Bayes Updateable (NBU).

Function Method includes LibLinear (LL), LibSVM(LSVM), Logistic, Multilayer Perceptron (MP), Radial Base Function Network(RBFN), Simple Logistic (SL), SPegasos(SP), SMO, Voted Perceptron (VP).

Lazy method includes IBI, IBK, Kstar, LWL

All of the above classification method have been tested for performance analysis from given each method using 10 fold cross validation. However, we only depicted those classifiers in below tables in the discussion section which have accuracy of more than 95%. Few of the classifiers with higher performance measures are narrated below:

Naive Bayes

It is a probabilistic classifier based on Bayes theorem. Naive Bayes is independent of features i.e. the presence or absence of feature is unrelated to the presence or absence of another feature of given class variable. For example if a thing is white and has oval shape then it is egg [11].

SMO (Sequential minimal optimization)

John Platt invented sequential minimal optimization algorithm (SMO) in 1998. It is a function’s algorithm and is widely used for solving optimization problem in the training of support vector machine (SVM) [12].

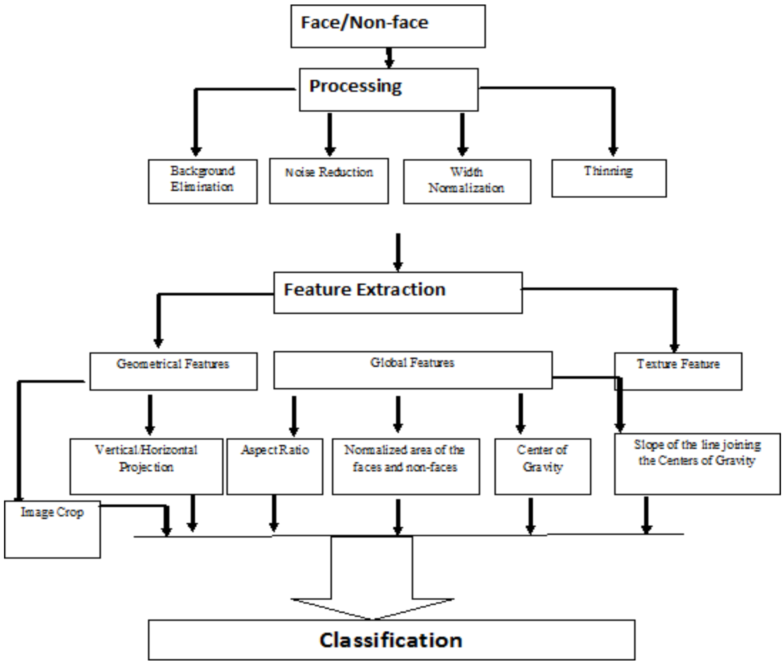


Figure 1. Block diagram of preprocessing and feature extraction

Meta method includes AdaBoostMI(ABMI), Bagging, Classification via Clustering(CC), Classification via Regression (CR), Dagging, Decorate, Filtered Classifiers, Grading, LogiBoost(LB), MultiBoost AB(MBAB), MultiClass Classifier(MCC).

Misc Method includes Hyper Pipes(HP), Serialized Classifier (SC), VFI.

Rule method includes Conjunctive Rule (CR), Decision

Table (DT), JRip, NNge,PART, Ridor, Zero.

Tree method includes ADTree, BFTree, Decision Stump (DS), J48, LADTree, Random Forest (RF), Random Tree (RT), SEP Tree.

DECORATE (Diverse Ensemble Creation by Oppositional Relabeling of Artificial Training Examples) is presented that uses a learner to build diverse committee. This is accomplished by adding different randomly constructed examples to the training set when building new committee members. These Artificially constructed examples are given category labels that discourage with the current decision of the committee [13].

Bayes

Naïve Bayes is an algorithm of Baye's rule. It is statistical algorithm and gives the simplified result of given inputs of an example. Naïve Bayes says that each feature of a given class variable is independent and cannot be related to other features of that class for example, a thing that is round and it's colour is black is considered as a ball, Naïve Bayes will consider these features to participate independently to probability [11].

$$\mathrm{p(C|F1,\ldots\ldots,Fn)} = \frac{1}{\mathrm{Z}}\mathrm{p(C)}\prod_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{p(Fi|C)}$$

where Z is scale dependent on F1……Fn (constant if values are known).

Accuracy

The accuracy (AC) is the proportion of the total number of predictions that were correct. It is determined using the equation [14], [15].

$$AC = \frac{a+b}{a+b+c+d}$$

where

- a is the number of correct predictions that an instance is negative,
- b is the number of incorrect predictions that an instance is positive,
- c is the number of incorrect of predictions that an instance negative, and
- d is the number of correct predictions that an instance is positive.

True positive

The recall or true positive rate (TP) is the proportion of positive cases that were correctly identified, as calculated using the equation:

$$TP = \frac{d}{c+d}$$

False positive

The false positive rate (FP) is the proportion of negatives cases that were incorrectly classified as positive, as calculated using the equation:

$$FP = \frac{b}{a+b}$$

True negative

The true negative rate (TN) is defined as the proportion of negatives cases that were classified correctly, as calculated using the equation:

$$TN = \frac{a}{a+b}$$

False negative

The false negative rate (FN) is the proportion of positives cases that were incorrectly classified as negative, as calculated using the equation:

$$FN = \frac{c}{c+d}$$

Precision

Precision (P) is the proportion of the predicted positive cases that were correct, as calculated using the equation:

$$P = \frac{d}{b+d}$$

IV. Result and Discussion

By using Weka classifiers we get different values for different classifiers. Comparing the submenus of each classifier we can get the best one which gives us more accurate value. In classifier Bayes and Meta the accurate value is higher than other classifiers. In Bayes Naive Bayes and in Meta Decorator gives the accuracy of 98%.In these classifiers two classifiers have more accurate values and those two classifiers are Naive Bayes and Decorate accuracy is 98% and 98% respectively. Now we will discuss about these two tables in details. We will calculate their values as follow:

TABLE I. Classification Using Meta Method

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
ABMI	98.7%	1.7%	98.7%	98.7%	98.7%	100%
Log.B	**99.3%**	0.3%	99.3%	99.3%	99.3%	100%
MBAB	**99.3%**	0.3%	99.3%	99.3%	99.3%	99.9%

TABLE II. Classification Using Function Method

Classifier	TP Rate	FP Rate	Precision	Recall	F measure	Roc Area
RBFN	98%	1%	98.1%	98%	98%	98.5%
Log.	97.3%	4.3%	97.3%	97.3%	97.3%	98.7%
Slog.	97.3%	5.3%	97.4%	97.3%	97.3%	99.9%

The Tables I, II, III, IV, and V show the classification measures such as True Positive (TP) rate, False Positive (FP) rate, Precision, Recall, F-measure and ROC calculated using Meta, Function, Lazy and Bayes Methods. In each of the Tables above we depicted the classifiers which show the accuracy percentage more than 96% to classify the 100 faces and 50 non-faces. Among the above classification methods, LogiBoost and MultiBoostAB of Meta method gives the classification performance of 99.3%. Likewise, RBFNetwork of function method give the accuracy of 98% higher than other classifiers. Similarly, LWL classifier of Lazy method, NNge& PART of Rule method and NB & NBU of Bayes method gives higher performance than other classifier in that method of 97.3%, 98% and 98% respectively. From the Tables I-V above, it is seen that LogiBoost and MultiBoostAB of Meta method gives classification accuracy of 99.3% at 100 faces and 50 non-faces higher than the other methods and classifiers as depicted in the Tables. In each of the above case we have first computed 27 features from 100 faces and 50 non-faces.

Confusion Matrix

a	b	Classified as
99	1	**a=f**
0	50	**b=nf**

$$AC = \frac{a+b}{a+b+c+d}$$

$$AC = \frac{149}{150} x100 = 99.3\%$$

From the above Confusion Matrix, it is seen that out of 100 faces 99 were correctly classified as faces, whereas out of 50 non-faces, 50 are classified as non-faces with an accuracy of 99.3% using LogiBoost classifier. The accuracy for all other classifiers is also illustrated in the Tables I-V against each classifier.

TABLE III. Classification Using Lazy Method

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
IBK	96%	5%	96%	96%	96%	95.5%
IB1	96%	5%	96%	96%	96%	95.5%
LWL	97.3%	1.3	97.5	97.3%	97.4%	98.9%

TABLE IV. Classification Using Rules Method

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
NNge	98%	1%	98.1%	98%	98%	98.5%
PART	98%	1%	98.1%	98%	98%	98.5%
Ridor	97.3%	2.3%	97.4%	97.3%	97.3%	97.5%

TABLE V. Classification Using Bayes Method

Classifier	TP Rate	Fp Rate	Precision	Recall	F measure	Roc Area
BN	97.3%	1.3%	97.5%	97.3%	97.4%	99.1%
NB	98%	1%	98.1%	98%	98%	99.4%
NBU	98%	1%	98.1%	98%	98%	99.4%

TABLE VI. Classification Using 40 Faces and 50 Non-Faces

Classifier	TP Rate	FP Rate	Precision	Recall`	F measure	Roc Area
NB	97..8%	2.8%	97.9%	97.8%	97.8%	98.5%
RBFNetwork	97.8%	2.8%	97.9%	97.8%	97.8%	98.1%
LWL	97.8%	2.8%	97.9%	97.8%	97.8%	98.6%
LogiBoost	**98.9%**	1.4%	98.9%	98.9%	98.9%	99.9%
HyperPipes	96.7%	4.2%	96.9%	96.7%	96.7%	99.4%
PART	97.8%	2.8%	97.9	97.8%	97.8%	99.5%
ADTree	97.8%	2.8%	97.9%	97.8%	97.8%	99.6%

TABLE VII. Classification Using 100 Faces

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
RBFNet.	98%	1%	98.1%	98%	98%	98.5%
NB	98%	15	98.1%	98%	98%	99.4%
PART	98%	1%	98.1%	98%	98%	98.5%
LB	**99.3%**	0.3%	99.3%	99.3%	99.3%	1%
ADTree	98.7%	0.7%	98.7%	98.7%	98.7%	100%
LWL	97.3%	1.3%	97.5%	97.3%	97.4%	98.9%
HP	98%	1%	98.1%	98%	98%	99.6%

TABLE VIII. CLASSIFICATION USING 200 FACES

Classifier	TP Rate	FP Rate	Precision	Recall	F measure	Roc Area
NB	97.6%	0.6%	97.9 %	97.6%	97.6%	99.6%
RBFNet.	98.4%	0.4%	98.5%	98.4%	98.4%	99%
LWL	98.8%	3.3%	98.8%	98.8%	98.8%	97.4%
LB	**98.8%**	4.8%	98.8%	98.8%	98.8%	100%
HP	99.6%	0.1%	99.6%	99.6%	99.6%	100%
PART	98%	6.5%	98%	98%	98%	96.7%
ADTree	98.8%	3.3%	98.8%	98.8%	98.8%	100%

TABLE IX. CLASSIFICATION USING 320 FACES

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
NB	98.4%	0.3%	98.6%	98.4%	98.4%	99.8%
RBFNt.	98.4%	0.3%	98.6%	98.4%	98.4%	99.1
LWL	98.4%	10.4%	98.4%	98.4%	98.3%	97.1%
HP	**99.7%**	0%	99.7%	99.7%	99.7%	99.7%
LB	**99.2%**	5.2%	99.2%	99.2%	99.2%	100%
PART	97.6%	13.9%	97.6%	97.6%	97.5%	96.2%
ADTree	99.2%	5.2%	99.2%	99.2%	99.2%	100%

TABLE X. CLASSIFICATION USING 400 FACES

Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
NB	97.1%	0.4%	97.7%	97.1%	97.3%	99.9%
RBFNet.	98.7%	0.2%	98.8%	98.7%	98.7%	99.4%
LWL	98.7%	7.2%	98.7%	98.7%	98.7%	99.9%
LB	**99.1%**	3.6%	99.1%	99.1%	99.1%	99.9%
HP	**99.1%**	0.1%	99.2%	99.1%	99.1%	99.7%
PART	97.8%	10.8%	97.8%	97.8%	97.8%	95.5%
ADTree	99.1%	3.6%	99.1%	99.1%	99.1%	99.9%

Tables VI-X show the classification measures using various classifiers of greater than 96% of accuracy out of all the classifiers tested among all the machine learning methods using Weka software. In each of the Tables VI-X we have computed the classification measures using classifiers such as Naïve Bayes, RBFNetwork, LWL, LogiBoost, HyperPipes, PART and ADTree. For each of the classifiers we have computed the classification measures at 40, 80, 100, 120, 160, 200, 240, 280, 320, 360, 400 as faces and 50 as non-faces for each case in order to judge the classification performance at different number of subjects for each classifier. In this case, we have depicted few of the subjects such as 40, 100, 200, 320 and 400 as faces and 50 as non-faces as shown in the Tables VI-X. In the Table VI, the LogiBoost classifier gives the accuracy of 98.9% at 40 faces and 50 non-faces. However, this accuracy increased to 99.3% when the faces increased to 100 with 50 as non-faces. While, the performance measure slightly decreased when the number of faces increased such as 98.8%, 99.2% and 99.1% at 200, 320 and 400 faces.

Another classifier such as HyperPipes of Misc method shows an accuracy of 96.7 at 40 faces and 50 non-faces as depicted in Table VI. However, by increasing the number of faces such as 100, 200, 320 its accuracy also increased to 98%, 99.6% and 99.7% respectively. However, it slightly decreased to 99.1 % at 400 faces and 50 non-faces as shown in Table X. The performance measure for all other classifiers is depicted in the Tables VI-X.

K-fold Cross Validation for Performance measure

In each of the cases we used k-fold cross validation taking k=1, 2,3,4,5 and 10. In k-fold cross-validation the data is first partitioned into k equally (or nearly equally) sized segments or folds. Subsequently k iterations of training and validation are performed such that within each iteration a different fold of the data is held-out for validation while the remaining k _ 1 folds are used for learning. Data is commonly stratified prior to being split into k folds. However, when k=10, we have get better classification performance than other k-folds. The results so far depicted here computed using 10-fold cross validation.

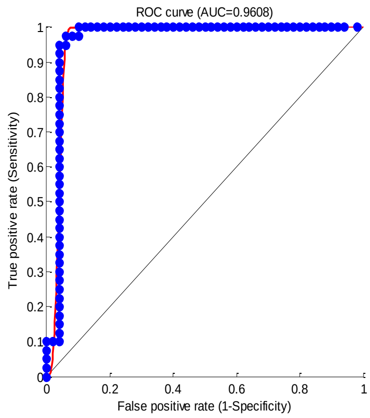


Figure 2. Receiver Operating Curve (ROC) at average number of features of faces and non-faces

A receiver operating characteristics (ROC) graph is a Technique used to visualize, organize and select classifier based on their performance. It is used since long time to detect the signals and shows a tradeoff between hit rate and false alarm rate of classifiers (Egan, 1975; Swets *et al.*, 2000). ROC analysis is also used to visualize and analyze the behavior of diagnostic systems (Swets, 1988). Besides, the medical decision making community has an extensive literature on the use of ROC graphs for diagnostic testing (Zou, 2002). Swets *et al.* (2000) brought ROC curves to the attention of the wider public with their Scientific American article [15].

ROC graph is a two-dimensional graph. The True Positive (TP) rate i.e. sensitivity is plotted on Y-axis while, False Positive (FP) rate i.e. Specificity is plotted on X-axis as shown in Fig. 2. In order to measure the classifiers ROC performance is reduced to a single value known as Area under the ROC curve, abbreviated as AUC (Bradley, 1997; Hanley and McNeil, 1982). In this case AUC=0.9608. AUC is a portion of the area of the unit square, so the value of AUC will always be between 0 and 1. And every realistic classifiers performance should never be less than 0.5. AUC has one of the most important statistical property that classifier will rank a randomly chosen positive instance higher than randomly chosen negative instance as claimed by Wilcoxon test of ranks (Hanley and MCneil, 1982).

This ROC curve shows that how the classifiers separates the faces from non-faces. If the area under the ROC is 100% it means perfect test, however, if the ROC value is 90% to 100 %, it is an excellent test i.e. the classifier excellently separate the positive examples from negative examples in this case faces from non-face. So, area under the ROC curve is a spread which shows higher the spread better the separation among positive and negative case. In our case we have tested 100 faces and 50 non-faces using LogiBoost classifier and the value AUC=0.9608 shows the excellent separation performance of faces from that of non-faces.

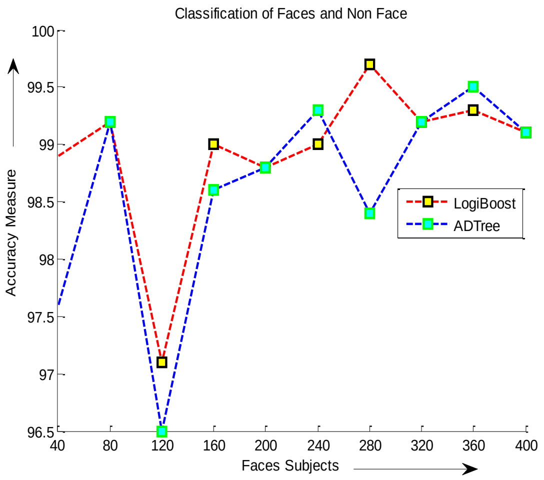


Figure 3. Accuracy measure using LogiBoost and ADTree classifiers at different number of faces subjects

In the Fig. 3 above, we have measured the performance accuracy using LogiBoost and ADTree of ML classifiers at different number of faces and non-faces. The x-axis shows the number of faces subjects with non-faces as 50 against each faces subject, i.e. 40 faces plus 50 non faces equivalent to 90 (40 vs 50 = faces vs non-faces),………400 x 50. For simplicity, we have shown here only the number of faces subjects on x-axis, while non-faces are fixed in each case which is 50. In each case, the accuracy measure is more than 96.5 %. Both LogiBoost and ADTree classifiers shows of less than 97% when the number of face subjects are 120, however, in all other cases the accuracy measure as shown in Fig. 3 is more than 98%. The ADTree shows higher accuracy measure when the number of faces subjects is 80, 240 and 360 and non-faces in each case was 50. While LogiBoost gives higher accuracy when the number of face subjects are 80 and 280. The above discussions give the directions to classify the faces and non-faces, the number of subjects and classifiers with higher performance.

The Fig. 4 below shows the accuracy and ROC measure values using 360 faces and 50 non-faces. Here, we would like to check the performance measure using different Machine Learning Methods such as Tree, Lazy, Meta, Function, Rules, Bayes and Misc. The classification methods such as Tree, Lazy, Meta and Misc give an accuracy of more than 99.2%. The highest accuracy is obtained from HyperPipes of Misc method, i.e. 99.8% higher than all other classifiers.

Likewise, in Fig. 5, we have shown that which classifier gives higher performance using 360 faces. From the Fig. 5, it is seen that LWL, ADTree and HyperPipes gives measuring accuracy of 99.6%, 99.5% and 99.8% than other classifiers as depicted in the figure. The ROC values from both Fig. 4 and Fig. 5 are also depicted against each classification method The ROC values are depicted in each case.

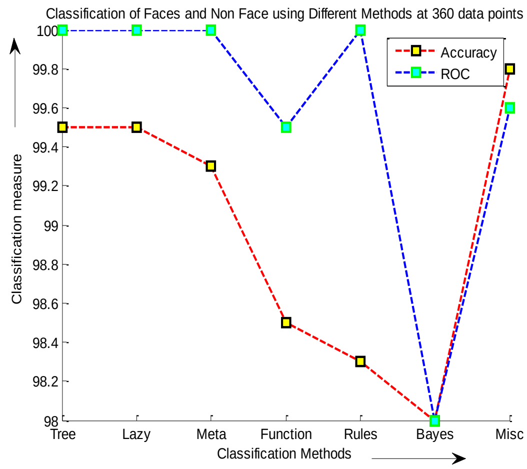


Figure 4. Classification of faces and non-faces using different methods and 360 faces subjects

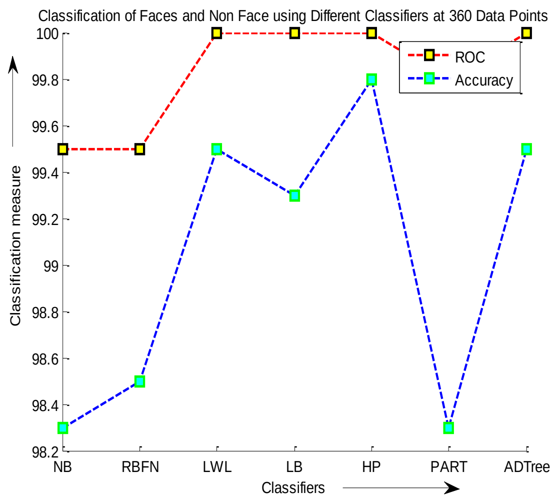


Figure 5. Classification of faces and non-faces using different classifiers and 360 faces subjects

The Table XI shows the summary of ROC Analysis with specificity, sensitivity and efficiency values shown against each cut-off. The maximum sensitivity, specificity, cost effective and efficiency cut-off point values are shown in the Table and Figure at right.

The Table XI also shows the summary of ROC Curve data with Accuracy as AUC=0.9608, standard error (S.E) value of 0.02232 less than 0.05 for 95% Confidence Interval (C.I) and value of ROC greater than 0.5 for C.I. The overall test performance is excellent i.e. excellent separation of faces from non-faces.

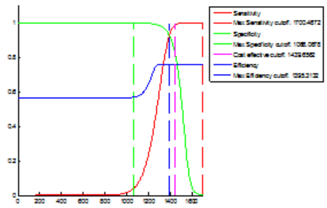
TABLE XI. SUMMARY OF PERFORMANCE MEASURE FOR ROC ANALYSIS WITH SENSITIVITY, SPECIFICITY AND EFFICIENCY

AUC	S.E.	95% C.I.		Comment
0.96083	0.02232	0.91708	1.00000	Excellent test

Standardized AUC	1-tail p-value	
20.6434	0.0000e+000	The area is statistically greater than 0.5

ROC CURVE DATA

Cut-off	Sensitivity	Specificity	Efficiency
1700.47	1.0000	0.0200	0.4556
1699.47	1.0000	0.0600	0.4778
1698.68	1.0000	0.0800	0.4889
1603.19	1.0000	0.1000	0.5000
1580.15	1.0000	0.1200	0.5111
1573.92	1.0000	0.1400	0.5222
1563.98	1.0000	0.1600	0.5333
1558.68	1.0000	0.1800	0.5444
1553.17	1.0000	0.2000	0.5556
1550.91	1.0000	0.2200	0.5667
1549.52	1.0000	0.2400	0.5778
1540.35	1.0000	0.2600	0.5889
1538.08	1.0000	0.2800	0.6000
1536.46	1.0000	0.3000	0.6111
1531.23	1.0000	0.3200	0.6222
1531.04	1.0000	0.3400	0.6333
1530.27	1.0000	0.3600	0.6444
1403.00	0.9750	0.9000	0.9333
1397.78	0.9750	0.9200	0.9444
1395.31	0.9750	0.9400	0.9556
1390.46	0.9500	0.9400	0.9444
1388.14	0.9500	0.9600	0.9556
1374.06	0.9250	0.9600	0.9444
1369.19	0.9000	0.9600	0.9333
1362.57	0.8750	0.9600	0.9222
1358.30	0.8500	0.9600	0.9111
1348.41	0.8250	0.9600	0.9000
1345.38	0.8000	0.9600	0.8889
1338.22	0.7750	0.9600	0.8778
1336.48	0.7500	0.9600	0.8667
1332.54	0.7250	0.9600	0.8556
1325.78	0.7000	0.9600	0.8444
1319.20	0.6750	0.9600	0.8333
1317.94	0.6500	0.9600	0.8222
1313.79	0.6250	0.9600	0.8111
1309.70	0.6000	0.9600	0.8000
1303.32	0.5750	0.9600	0.7889



1) Max Sensitivity Cut-off point= 1700.47
2) Max Specificity Cut-off point= 1066.07
3) Cost effective Cut-off point= 1439.66
4) Max Efficiency Cut-off point= 1395.31

V. CONCLUSION

In this paper we have classify faces and non-faces using Machine Learning Classifiers. We have developed our primary data of faces and non-faces using Digital Camera of 12 Mega Pixel from male and female children of class 5th to 8th in School at Muzaffarabad, Azad Jammu and Kashmir. After collecting the data, we have preprocessed it for proper feature extraction and better classification performance. We have developed program in Matlab for preprocessing and features extraction as shown in Fig. 1 above. We have applied all Machine learning classifiers of Weka Software using 10-fold cross validation. The accuracy is checked for varying number of faces and non-faces subject using different classifiers. The classification methods such as Tree, Lazy, Meta and Misc give higher performance of 98% than other methods. While the classifiers LWL, ADTree and HyperPipes gives performance accuracy of more than 99% than all other classifiers. The average ROC analysis value of 96.08% was obtained to show the separation of faces from non-faces to correctly classified positive examples as positive and negative examples as negative.

REFERENCES

[1] A. Hamed, M. Malekzadeh, and S. Sanei, "A new neural network approach for face recognition based on conjugate gradient algorithms and principal component analysis," *Journal of Mathematics and Computer Science*, vol. 6, pp. 166-175, 2013.

[2] H. S. Mittal and H. Kaur, "Face recognition using PCA & neural network," *IJESE*, vol. 1, pp. 570-578, April 2013.

[3] K. Kyungnam, "Face recognition using principle component analysis," *International Conference on Computer Vision and Pattern Recognition*, 1996, pp. 586-591.

[4] S. Z. Li and A. K. Jain, eds. *Handbook of Face Recognition*, Springer, ch. 8, 2011.

[5] P. Latha, L. Ganesan, and S. Annadurai, "Face recognition using neural networks," *Signal Processing: An International Journal, CSC Journals, Kuala Lumpur, Malaysia*, vol. 3.5, 2009, pp. 153-160.

[6] A. Faizan, A. Najam, and Z. Ahmed, "Image-based face detection and recognition: State of the art," *NASA Astrophysics Data System*, pp. 1302.6379, 2013.

[7] S. Z. Li and J. W. Lu, "Face detection by neural learning," 1999.

[8] S. Anupam, R. Tiwari, and R. Kala, *Real Life Applications of Soft Computing*, Boca Raton: CRC Press, 2010.

[9] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 23-38, 1998.

[10] O. Emre, T. Senturk, and M. ElifKarsligil, "Off-line signature verification and recognition by support vector machine," in *Proc. European Signal Processing Conference*, 2005.

[11] B. Ahmad, L. Khan, and M. Awad, "Bayesian networks," *Encyclopedia of Data Warehousing and Mining: IZ*, pp. 89-92, 2006.

[12] P. John. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.43.4376

[13] M. Prem and R. J. Mooney, "Constructing diverse classifier ensembles using artificial training examples," *IJCAI*, vol. 3, 2003.

[14] S. Heiko, *et al.*, "Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma," *Journal of Clinical Oncology*, vol. 23, pp. 4643-4651, 2005.

[15] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861-874, 2006.

Lal Hussain received his MCS degree in Computer Science from University of Azad Jammu and Kashmir, Pakistan in 2005 and MS in Communication and Networks from Iqra University, Islamabad, Pakistan in 2012 with Gold medal. Currently, he is pursuing Ph. D at Department of CS & IT, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan. He is working as Assistant Director Quality Enhancement Cell, University of Azad Jammu and Kashmir since 2006. His responsibilities includes to arrange trainings for faculty members, assist program teams and assessment team in preparing self-assessment reports, preparing institutional performance evaluation data, program assessment, research activities etc. His research interest includes Biomedical Signal Processing with concentration on complexity analysis of brain and physiological signals, Neural Networks and Machine Learning classification problems etc.