Gender Classification with Decision Trees

Muhammad Naeem Ahmed Khan, Sheraz Ahmed Qureshi and Naveed Riaz

Department of Computing,
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST),
Street # 09, Plot # 67, Sector H-8/4, Islamabad, Pakistan
mnak2010@gmail.com, shrzqureshi@yahoo.com, n.r.ansari@szabist-isb.edu.pk

Abstract

With the evolution of HCI (Human-Computer Interaction), the computer vision systems are playing an important role in our lives. Some of the prime areas of computer vision applications include gender detection, face recognition, body tracking and ethnicity identification etc. Automated data analyses techniques help discover regularities and hidden associations in larger volumes of datasets. Classification being a data mining technique is largely used to group categorical data as well as a blend of continuous numeric values and categorical data. A number of classification techniques like decision trees, support vector machine (SVM), nearest neighbors and neural networks etc. have gained popularity in numerous areas of data mining practices. Among these classification techniques, decision trees offer an added advantage of producing easily interpretable rules and logic statements along with generating the classification tree for the given dataset. This study offers a distinct method for gender classification of facial images. We have used a variant of the decision tree algorithm for gender classification of frontal images due to its distinctive features. Our technique demonstrates robustness and relative scale invariance for gender classification. Details of the experimental design and the results are reported herein.

Keywords: *Decision Trees, Gender Classification, Data Mining, Machine Learning, Feature Extraction, Human Computer Interaction*

1. Introduction

Gender classification has become an area of extensive research due to its increasing application in the existing HCI system *e.g.*, passive surveillance (restricting access to premises based on gender) and collecting demographics *etc.* The traditional pattern classifiers such as linear and quadratic discriminant, nearest-neighbor as well as modern techniques like neural networks and SVM have been extensively used for gender classification in the contemporary literature. Classification, much like clustering and regression, is one of the data mining techniques that serves as an effective tool for identifying hidden affinities and distinctive patterns in large datasets. The detection of regularities and affinities in different parameters of a dataset serves as a useful tool for decision making and drawing predictions. Decision tree is a classification technique which is mostly used in procedural research as a predictive model to classify the large volumes of datasets [27]. Facial image processing has always been of prime concern for commercial applications as well as specialized applications used by the security and law enforcement agencies. Ensuing to this fact, gender classification has been an active area of research for decades due to mercantile and security challenges worldwide.

One of the main characteristics of face recognition is to identify gender of the person from a still image or an animated video. The visual information extracted from human faces serves as an important source of information for gender classification [11]. A large number of psychophysical studies have investigated gender classification from facial perception in the humans [22]. To date, there is no single technique that offers a robust solution for those applications which involve face recognition [6] as part of their functionality. In view of the above, this paper focuses on the issue of gender classification. We have used SUMS dataset of facial images to realize an efficient classification mechanism. The classification results are supplemented with easy to understand rule-sets that evolve during the course of the classification process.

Gender classification refers to designate an image of a person into one of the categories of male or female. Precise image-based gender classification could have central value in *Human-Computer Interaction (HCI)* [28]. Computer vision systems not only boost existing *HCI* systems but can also assist passive surveillance and control [21] of precincts and estates (*e.g.*, restricting entry to certain premises based on gender), performing valuable analysis (*e.g.*, comparing the consumption of specific items of menswear and womenswear in a super store), voice recognition purposes *(e.g.*, identifying gender of the speaker in the audio/video files) and reducing uncertainties in audio-visual aids (*e.g.*, relating only female voice to a female appearance).

The characteristic extraction of features and appropriate selection of classifier are of prime significance for classification of facial images. Facial feature extraction techniques for static or even video images are generally categorized into geometric and template matching methods [6]. Facial features extraction methodology pertains to the measurement of conspicuous facial characters like eyes, mouth, nose, chin and other key characteristic facial elements. Brunelli & Poggio (1993) [3] used geometrical features, such as nose width and length, mouth position and chin shape for recognition of human faces. A variety of methods for gender classification have been proposed for extraction of facial features in the earlier studies like *2-D Gabor Wavelet Transform* [7], *SVM* [34, 11], *Adaboost* [28], *hyper-plane* [8] and *Discrete Cosine Transformation* [17]. However, it has been observed that a blend of these techniques could be considered more effective and appropriate [31]. Each of the facial features extraction technique entails the following characteristics:

- *Geometry-based features* are generally extracted using geometric characteristics such as relative positions and size of the facial parts. This approach, however, results in a large number of features to be used by the classifier.
- *Template-based approach* matches facial elements to previously designed templates. This approach could be too complex due to extensive computation involved and is only effective when query and model images have the same scale, orientation and illumination properties.
- *Colour segmentation-based technique* uses colour of the skin to isolate the face from rest of the body. In this approach, the quality of image coupled with illuminations and hue plays an important role for image recognition as well as on the rate of classification.
- *Appearance-based technique* refers to any of the extracted characteristics of the image, known as a feature. The technique, however, requires good-quality images to extract the features correctly.

Irrespective of the feature set obtained through employing any of the above-mentioned approaches, a variety of classification techniques can be used for categorization purposes such as Decision Trees, Neural Networks, Nearest Neighbors, Support Vector Machines (SVM), Artificial Immune System (AIS), Bayesian Networks, Fuzzy Logic, Rough Set

Theory and Genetic Algorithms *etc.* However, decision trees-based classification technique has proven to be a more practical approach in developing predictive models in many domains [9, 4].

Keeping in view its usefulness as a classifier, we decided to use decision trees methodology for gender classification in our study. This paper proposes a distinctive method for gender classification of facial images using decision tree classification. We have used frontal face images obtained from SUMS face database for identifying gender of the person based on the geometrical facial features extracted from the images. Empirical results show that our decision tree based gender classification technique is comparatively more efficient and produces promising accuracy.

This paper is organized into seven sections. This section being the introduction to gender classification is followed by a summary of the related work in the field of gender classification. An insight into the decision trees classification technique and the possible use for gender classification is recapitulated in Section III. The fourth section outlines our proposed methodology for classification to identify gender of a person from the facial image. The experimentation design and scenarios' details are explicated in this section. An elaborated discussion of the experimental results is provided in the successive section. The prospective future dimensions to this work are described in Section VI and, finally, we conclude in the last section.

2. Related Work and Motivation

Gender is a critical cue in social activities [32] and its automatic classification in the digital world has gained much popularity. A number of different techniques and methods have evolved in the past for gender classification. A critical analysis of various gender classification methods is enunciated in (Mäkinen & Raisamo, 2008) [18]. Gender identification from images that differ in pose, illumination and facial expression have been a challenging task as processing such images is much trickier than the snapshot images. To perform gender classification on consumer images in a multiethnic environment, a *Probabilistic Boosting Tree (PBT)* approach [10] has been used since it caters for many variances in facial expression. However, experimental results of both snapshots and consumer images show that *PBT* method is better than the *Support Vector Machines (SVM)* [2] and *Real AdaBoost* [28] algorithm.

Strategy to mine statistical information of images based on the *appearance features* carries the advantage of being swift and simple but may have imperfection when local appearance deviation occurs [31]. In geometric feature-based approach, some prior (self-governing) knowledge is applied [14] and facial geometry features such as eyebrows' thickness and nose width *etc.* are extracted first. It may bear advantage of translation and rotation, but it may discard a large number of supportive information [30]. Hybrid gender classification uses *Active Appearance Model (AAM)* [31] to extract geometry features that are normalized using min-max method and are subsequently classified by applying *AdaBoost* algorithm.

Critical analysis of the performance of different automatic gender classification methods like SVM, neural networks and threshold AdaBoost with various facial alignment methods has been reported by (Makinen & Raisamo, 2008) [19]. Such studies lead to development of a standard dataset for gender classification that consists of a set of facial images detected using face detector used in (Viola & Jones, 2004) [29] when applied to a subset of images from the FERET face recognition database [23].

For optimizing face recognition scheme, *Discrete Cosine Transform (DCT)* - that uses features extraction method - leads to the development of a hybrid *Flexible Neural Tree (FNT)* classification model for facial recognition [6]. The *FNT* model employed *Olivetti Research Laboratory (ORL)* database of facial images and it focuses on reducing the input features for refining the face recognition functionality. Gender classification based on iris texture has also been studied extensively. Lagree & Bowyer (2011) [16] attempted to predict ethnicity and gender from features of the iris texture and report that gender prediction is harder than ethnicity prediction and, over and above, ethnicity prediction is more difficult for females than for males. Thomas, *et al.*, (2007) [26] emphasize that selection of iris as the biometric is more appropriate as a number of methods are available to segment and encode an iris image which in turn as quite helpful in extracting attributes from an iris texture and creating a meaningful feature vector.

Sparse classification method using sparse Support Vector Machines (SVM) has also been used by (Costen, *et al.*, 2004) [8] for selecting the most essential features and to maximize the classification margin. Performance comparison among *Linear Discriminant Analysis (LDA)*, *SVMs* and *Exploratory Basis Pursuit Classification (EBPC)* methodologies shows that *LDA* produces suboptimal results but *EBPC* and *SVM* produce promising results [8]. Sparse classification method is tried for gender classification in (Costen, *et al.*, 2004) [8] which not only selects the most important features but also maximizes the classification margin, which makes such a technique much closer to SVM. Likewise, nonlinear *SVM* for appearance-based gender classification with 1800 hairless low-resolution "thumbnail" faces obtained from FERET facial images database has been used in (Moghaddam & Yang, 2002) [21]. Gender classification based on gait is also studied by (Yu, *et al.*, 2009) [32]; however, such techniques are feasible only in the controlled environments as it suffers from some caveats like view variation, shoes changes, clothing or image with additional objects. Manual alignment (*i.e.*, when location of facial features are detected by human) produces slightly better results (Makinen & Raisamo, 2008) [19], but such methodologies lack practical implementation due to inadequacy of automation requirements. Being motivated of the high accuracy of results and simple to comprehend rule set generated by the decision trees algorithm, we employed it to classify the gender of a person from his/her frontal image.

3. Decision Tree Classification Approach

The C4.5 (Quinlan, 1993) [24] is an improved version of CLS (Concept Learning System) and ID3 algorithms. It was designed by Quinlan Ross in 1993. Since C4.5 is a descendant of CLS and ID3, it not only generates more intelligible classification rule sets that are expressed in the form of a decision tree much like CLS and ID3 algorithms, but is also capable to account for missing data. C4.5 uses a *divide and conquers* methodology to construct a suitable tree from a training dataset. The C5.0 and J48 are the enhanced versions of C4.5 algorithms and use the same classification approach as of C4.5 which is based on the following procedure [15]:

- If dataset is small or all the instances in a dataset belong to the same class, then the generated decision tree would consist of leaf nodes labeled with the most frequent class in the dataset; or
- otherwise, it classifies dataset on single attribute with two or more outcomes and applies the same procedure recursively to each subset.

The C4.5, C5.0 and J48 algorithms are computationally powerful and easy to use since the rule set derived from these techniques are relatively straightforward [27, 9]. Decision tree based classification has been previously used for several purposes, such as improvement of personnel selection [5], prediction of electricity energy consumption [27], prediction of breast cancer survivability [9] and analysis of accident frequency [4] to name a few. Some of the well-known decision tree algorithms are ID3 (Iterative Dichotomiser 3), C4.5/C5.0/J48 and CART (Classification and Regression Trees) [1]. Among these algorithms, C4.5 has gained much popularity. A decision tree is a hierarchical data structure where each branch node represents a choice between two alternatives and each leaf node represents a classification or decision. For instance, a leaf node in a credit rating system may indicate a red or green classification flag indicating the decision to reject or accept a customer's loan application respectively. Decision trees are usually implemented as a set of nested if-then-else statements. Classification through decision trees endures the following characteristics [5, 27]:

- Decision tree-based classification can be performed on dataset comprising both continuous and categorical values.
- It produces a model that contains easy to explicable and interpretable rule set which helps in understanding the logical outcome of the classification results.
- It is also capable to deal with noisy data as it focuses on the more prominent features.
- The rule sets generated by the decision tree classifier can be readily expressed in English language; whereas, such type of characteristics is not supported by other counterpart classification techniques *e.g.*, neural network whose classification criteria always remain enigmatic and inexplicable.
- Being a straightforward method for construction of a tree from the available data and rule sets, decision tree offers easily explainable classification paths to the users.
- It can also discover the substantial high-order interactions among different data objects rapidly.
- Decision trees can be constructed relatively quickly as compared to the other methods.

WEKA (*Waikato Environment for Knowledge Analysis*) [30] is a popular data mining suite developed in Java language at the University of Waikato, New Zealand. WEKA is an open source freeware tool and is available under the GNU General Public License. WEKA offers a uniform interface to implement numerous classification and prediction algorithms along with the methods for pre-processing and post-processing of data. It is equally useful for evaluating the results of classifiers for a numerical, categorical and enumerated datasets. We have used WEKA version 3.6.4 for experimentation purposes in our research. WEKA toolkit contains a customized J48 classifier which is the optimized implementation of C4.5.

4. Proposed Methodology and Experimentation

Our proposed methodology for gender classification comprises the following three phases:

- The first phase *feature extraction* relates to data collection process which involves extraction of several facial features. These features are obtained through geometrically-based features extraction technique.
- The second phase *classification* necessities learning process of the training dataset followed by generation of classification rule set. This phase is accomplished by using WEKA J48 classifier.
- The third phase *validation* verifies the accuracy classification rule set using the test dataset.

The attributes selected by us for both training and test datasets are based on the feature-set selection criteria proposed by (Xu, *et al.*, 2008) [31]. The list of key attributes that we have used for experimentation is described in Table 1. All the numeric attributes are obtained through geometry-based feature extraction technique; however, the classification attribute named as *"Classification Flag"* is represented with values "M" or "F" for denoting the male or female image respectively.

Table 1. List of Key Facial Attributes used in the Experiments

Attribute/Factor	**Description**
LEW	Left Eye Width (From one corner to the other)
REW	Right Eye Width (From one corner to the other)
NW	Nose Width
LEML	Left Eye center to Mouth Left corner
REMR	Right Eye center to Mouth Right corner
D1	Left Eye center to Mouth Right corner
D2	Right Eye center to Mouth Left corner
MLC	Mouth Left corner to middle of Chin
MRC	Mouth Right corner to middle of Chin
Classification Flag	Classification Label - either M (Male) or F (Female)

For training our classifier, we used a dataset of 300 (150 male and 150 female) grayscale images of unique subjects obtained from Stanford University Medical Student (SUMS) frontal facial images database. The selected images had fixed frame size of 200 x 200 pixels and all the images were frontal facial images with some variations in pose and expression. A sample of few images is provided in Figure 1.

Figure 1. A Sample of SUMS Images Face Database

Features extraction of the facial images was achieved by using the distance tool available in the *Image Processing Toolbox* of MATLAB [12]. We have used MATLAB version 7.5.0 for this purpose. The distance tool displays the Euclidean distance between the two endpoints of the line which is displayed in a label tip superimposed over the line. The tool specifies the distance in data units determined by the X-coordinate and Y-coordinate of the pixel positions corresponding to a specific selected feature. A selected image along with the X, Y-coordinates of different facial features, as mentioned in Table 1, were then viewed in the *Image Tool* utility in the MATLAB workspace by using the following set of commands.

```
img = imread('face1.jpg');
imtool(img);

or

imtool('face1.jpg');
```

Subsequently, we trained 300 selected images through WEKA J48 classifier. The training dataset was organized in a comma-separated format using “.arff” file extension type. It is worth mentioning that the training process of the J48 classifier (one of the implementations of C4.5 algorithm) does not require discretization of numeric attributes (in contrast to the ID3 algorithm from which C4.5 has evolved). Afterward, we choose 200 images, having the same frame size and feature set as those of the images used in the training phase in order to verify the accuracy of the classifier. It is worth mentioning that in addition to the 100 images reserved for test dataset, we added 100 additional images into the test dataset that were already used to train the classifier. In this way, we got 200 images in the test dataset comprising a blend of known and unknown images by the classifier. The results obtained for the test dataset were in conformity with the desired functionality as required for a perfect classifier.

5. Results & Discussion

As described earlier, we used 400 (200 male and 200 female) grayscale frontal facial images to train and test the classifier. We used 75% and 25% of the images for training and testing the classifier respectively. All the images for training and test datasets were obtained from SUMS database and were trained on WEKA J48 classifier by using nine distinctive geometrical features for each image. As stated earlier, a key significance of C4.5 classifier (the one used in WEKA) is that it produces human readable and easy to understandable classification rule set which does not require any domain expert knowledge. We used geometrically-based features extraction method and performance evaluation summary of the classification process on the training dataset is shown in Table 2.

Table 2. Classification Accuracy of the Training Dataset

Classification Statistics	Results	%age
Correctly Classified Instances	*292*	*97.33 %*
Incorrectly Classified Instances	*8*	*2.67 %*
Kappa statistic	*0.9449*	
Mean absolute error	*0.0487*	
Root mean squared error	*0.156*	
Relative absolute error	*9.8192 %*	
Root relative squared error	*31.3379 %*	
Total Number of Instances	*300*	

The WEKA J48 classifier also generated a tree exhibiting decision taken at each branch based on selecting the more appropriate and fitting parameters which helped to achieve the desired accuracy. Figure 2 exhibits tree view of rule set generated by the classifier. The size of the decision tree generated by the classifier (Figure 2) consists of 25 nodes. The leaf nodes (13 in total) represent a classification or *decision* and the remaining 12 branch nodes articulate a choice between the two alternatives; which can also be called branching factor.

The C4.5 classifier can also help to uncover the interesting and the most pertinent attributes from the training dataset through the number of hits for each of the attributes from the automatically generated classification rule set. A list of key attributes along with their frequency of hits generated by the classifier is shown in Table 3.

Table 3. List of Key Attributes Identified by the WEKA Classifier

Attribute	Description	Hits
LEW	Left Eye Width	4
NW	Nose Width	3
LEML	Left Eye center to Mouth Left corner	1
REMR	Right Eye center to Mouth Right corner	1
D1	Left Eye center to Mouth Right corner	1
MRC	Mouth Right corner to middle of Chin	2

As evident by Table 3, out of nine selected attributes, excluding C*lassification Flag*, six of the attributes were chosen as a hit by the aforementioned classification rule set. The results shown in Table III also reveal that the *Left Eye Width* attribute has the highest number of hits followed by *Nose Width* whose frequency of hits stands at three. Key to analysis of the experimental results is to comprehend the relevance and importance of the generated rule set. The rule set that extracted from the classification tree shown in Figure 2 are presented in Table 4 to highlight the criterion used for formulating classification rule set by WEKA classifier. A thorough review of the classification rule set also gives the impression which it is, in fact, a set of *nested if* statements that can readily be expressed in English language and are straightforwardly understandable. Though the findings obtained from our experimentation appear to be quite interesting; however, further experiments are required to establish the relevance and aptness of these rule set.

Table 4. Summary of Classification Rule Set

Rule #	Classification Rule Set	Classification
1	*IF (MRC <= 38.28) AND IF (LEW <= 15.34)*	M
2	*IF (MRC <= 38.28) AND IF (LEW > 15.34) AND IF (REMR <= 34.02)*	F
3	*IF (MRC <= 38.28) AND IF (LEW > 15.34) AND IF (REMR > 34.02) AND IF (LEW <= 16.28) AND IF (MRC <= 32.04)*	F
4	*IF (MRC <= 38.28) AND IF (LEW > 15.34) AND IF (REMR > 34.02) AND IF (LEW <= 16.28) AND IF (MRC > 32.04)*	M

Makinen and Raisamo (2008) [19] reported that the average gender classification rate for all the dataset stood at 82.1% when alignment was used, 84.6% without alignment and 87.1% when manual alignment was used. The classification accuracy achieved in our study stands at 94% which is much more than the accuracy reported by (Makinen and Raisamo, 2008) [19]. Chen, *et al.*, (2005) [6] reported that all the assessments conducted realized high grounding and recognition speed, as well as high recognition rate; and their results are also in agreement with results of our study. Moghaddam & Yang (2002) [21] reported that a Gaussian kernel *SVM* produced 3.4% more error rate as compared to typical classifiers (like *linear*, *quadratic*, *FLD and Nearest Neighbor - NN*) but their study did not select appropriate robust dimensions to classify the face which is not the case in our study.

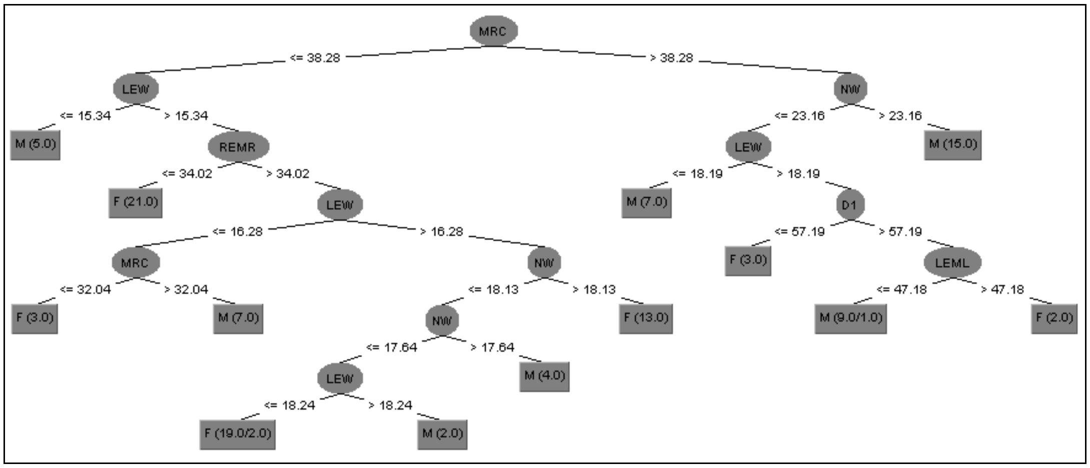

Figure 2. Tree View Sample of Rule Set using J48 Classifier

In this study, we empirically observed the remarkable potential of C4.5 as a classification method for gender identification. Our experiments using training and test datasets generated promising results and equally ascertained the suitability of C4.5 classifier. Furthermore, we intend to carry out experiments with other decision tree classifiers (such NBTree, SimpleCart, REPTree and BFTree) to further augment our findings as a prospective future work. Table 5 illustrates accuracy of the WEKA J48 classifier when used with test dataset of 200 facial images in this study.

Table 5. Classifier Accuracy

Status	**Number of Data Items**	**Accuracy**
Correctly Classified	*197*	*98.5 %*
Incorrectly Classified	*3*	*1.5 %*
Total:	***200***	

Sun, *et al.*, (2002) [25] employed fusing infrared with visible images on the Equinox face dataset to exploit lower sensitivity of visible imagery to occlusions caused by eyeglasses. For this purpose, genetic algorithms are employed to find an optimum strategy to perform image-based fusion in the wavelet domain and feature-based fusion in the eigenspace domain. Whereas, Jain, *et al.*, (2005) [13] employed independent component analysis (ICA) to represent images (of FERET facial database) as a feature vector in a low dimensional subspace; and that approach helped achieve higher classification accuracy by using SVM as a classifier. Zahedi & Yousefi (2011) [33] applied scale invariant feature transform (SIFT) method to eliminate the pre-processing steps for aligning facial images so that facial landmarks like eyes, nose, lips and chin are placed in uniform locations. Likewise, Brunelli & Poggio (1993) [3] used geometrical features, such as nose width and length, mouth position and chin shape for recognition of human faces. We also used the similar set of facial features, as used Zahedi & Yousefi (2011) [33] and Brunelli & Poggio (1993) [3], for gender classification; and this makes our approach in line with the generally accepted criteria for selection of feature set.

Moghaddam & Yang (2000) [20] investigated gender classification by employing SVM as a classifier on the low-resolution thumbnail images from the FERET face database and reported 3.4% error rate. The average classification error rate reported in studies stands at 32% for thumbnail images and 6.7% for higher resolution images [20]. Nevertheless, the accuracy that we achieved using the decision trees makes our approach a promising one. Furthermore, in real-world applications, a gender classification system is required to handle a large amount of facial variations such as posture, illumination intensities, skin color, personal age *etc.* However, our proposed technique is merely affected by the pose and is not affected by the other factors due to the obvious reasons of the features/parameters selected from the facial image for gender classification.

6. Future Work

This paper described the importance and utility of decision trees for gender classification. However, other data classification techniques for data mining such as SVM, Fuzzy Logic, Nearest Neighbors, Neural Networks, and Bayesian techniques etc. may be tested by using the same dataset to demonstrate a comparison of these techniques and this could be potential future work. Moreover, in the proposed methodology the geometrically-based feature extraction of facial images was in semi-automated fashion using the *Image Processing Toolbox* of MATLAB [12]. However, the facial features as mentioned in Table I need to be obtained automatically in an integrated environment together with the classification tool. Therefore, as a prospective work, we intend to develop an integrated package which can perform the same set of activities to streamlining and sequencing of the whole phases starting from automatic extraction of features dataset up to the classification phase. Further, rule set extracted from the classification need to be embedded into the decision support system. Another prospective and challenging future dimension to this research could be to contemplate upon to design an algorithm that can also identify the age group of the person by using the facial features obtained from an image.

7. Conclusion

In this study, we conducted experiment using decision tree — a machine learning technique — on the frontal facial images to classify them on the basis of their gender. There was some variation in images with regard to pose and expression as well as with certain accessories like eyeglasses. It emerged from the classification rule set generated by the classifier that at least six facial features were required for training the classifier. Using the WEKA J48 classifier, we also achieved dimensionality reduction as initially, we fed nine attributes to the classifier, which were reduced to six by the WEKA J48 as the discriminant attributes. Form a series of experimentation conducted as part of this research, we observed that WEKA J48 classifier proved to be of much use for gender classification as the classification results obtained using WEKA classifier was very promising. The classification accuracy was relatively high when output of our classifier was arithmetically compared with other classification techniques. The experiments provide a new insight into gender classification methodology that not only serves as a classifier but also outlines the clear and easy to understand discriminant rules. This study may be useful for extracting the rule set for a decision support system which involves a large volume of dataset having several decision parameters.

References

[1] M. N. Anyanwu and S. G. Shiva, “Comparative Analysis of Serial Decision Tree Classification Algorithms”, International Journal of Computer Science and Security, vol. 3, no. 3, **(2009)**, pp. 230-240.
[2] W. Ashour, F. Hussin and K. M. Mahar, “Supervised Texture Classification Using Several Features Extraction Techniques Based on ANN and SVM”, In: Proceedings of the IEEE/ACS International Conference on Computer Systems and Applications, Alexandria, Egypt, **(2008)** April 4, pp. 567-574.
[3] R. Brunelli and T. Poggio, “Face recognition: features versus templates”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Istituto per la Ricerca Sci. e Tecnologica, Trento, vol. 15, Issue 10, **(1993)** October, pp. 1042-1052.
[4] L. Y. Chang and W. C. Chen, “Data mining of tree-based models to analyze freeway accident frequency”, Journal of Safety Research, vol. 36, **(2005)**, pp. 365-375.
[5] C. F. Chien and L. F. Chen, “Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry”, Expert Systems and Applications, vol. 34, **(2008)**, pp. 380-290.
[6] Y. Chen, S. Jiang and A. Abraham, “Face Recognition Using DCT and Hybrid Flexible Neural Tree”, In: Proceedings of the IEEE International Conference on Neural Networks and Brain (ICNN&B), Beijing, China, vol. 3, **(2005)** October 13-15, pp. 1459-1463.
[7] W. Chuan-xu, L. Yun and L. Zuo-yong, “Algorithm Research of Face Image Gender Classification Based on 2-D Gabor Wavelet Transform and SVM. In: Proceedings of the IEEE International Symposium Computer Science and Computational Technology (ISCSCT), Shanghai, China, **(2008)** December 20-22, pp. 312-315.
[8] N. P. Costen, M. Brown and S. Akamatsu, “Sparse Models for Gender Classification”, Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG'04), Seoul, Korea, **(2004)** May 17-19, pp. 201-206.
[9] D. Delen, G. Walker and A. Kadam, “Predicting breast cancer survivability: A comparison of three data mining methods”, Artificial Intelligent in Medicine, vol. 34, **(2005)**, pp. 113-127.
[10] W. Gao and H. Ai, “Face Gender Classification on Consumer Images in a Multiethnic Environment”, In: Proceedings of 3rd ACM International Conference on Advances in Biometrics, Alghero, Italy, **(2009)**, pp. 169-178.
[11] A. B. A. Graf and F. A. Wichmann, “Gender classification of human faces”, In Biologically Motivated Computer Vision 2002, LNCS 2525, **(2002)**, pp. 491–501.
[12] R. F. Gonzalez, R. E. Woods and S. L. Eddins, “Digital Image Processing Using MATLAB”, second ed. Gatesmark, ISBN: 9780982085400, **(2009)**.
[13] A. Jain, J. Huang and S. Fang, “Gender identification using frontal facial images”, In: Proceedings of IEEE Conference on Multimedia and Expo, **(2005)**, pp. 1082–1085.
[14] T. Ji and W. Wang, “A Fast Face Detection Method Based on Skin Color”, Journal of Convergence Information Technology, vol. 6, no. 9, **(2011)**, pp. 50-58.
[15] R. Kohavi and R. Quinlan, “Decision Tree Discovery”, Data Mining, vol. 6, **(2000)**, pp. 10-18.
[16] S. Lagree and K. W. Bowyer, “Predicting Ethnicity and Gender from Iris Texture”, 2011 IEEE International Conference on Technologies for Homeland Security (HST), University of Notre Dame, Notre Dame, IN, USA, **(2011)** November 15-17, pp. 440-445.
[17] A. Majid, K. Asifullah and M. Anwar, “Gender Classification using Discrete Cosine Transformation: a Comparison of Different Classifiers” In: Proceedings of the IEEE 7th International Multi Topic Conference (INMIC), Islamabad, Pakistan, **(2003)** December 9, pp. 59-64.
[18] E. Mäkinen and R. Raisamo, “An experimental comparison of gender classification methods”, Pattern Recognition Letters, vol. 29, **(2008)**, pp. 1544–1556.
[19] E. Makinen and R. Raisamo, “Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 3, **(2008)**, pp. 541-547.
[20] B. Moghaddam and Yang, “Gender classification with support vector machines”, In Proceedings of the 4th IEEE Conference on Automatic Face and Gesture Recognition, **(2000)**, pp. 306-311.
[21] B. Moghaddam and M. -H. Yang, “Learning Gender with Support Faces”, In: Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, **(2002)**, pp. 707-711.
[22] A. J. O’Toole, K. A. Deffenbacher, D. Valentin, K. McKee, D. Huff and H. Abdi, “The Perception of Face Gender: the Role of Stimulus Structure in Recognition and Classification”, Memory & Cognition, vol. 26, no. 1, **(1998)**.
[23] P. J. Phillips, M. Hyeonjoon, P. Rauss and S. A. Rizvi, “The FERET Evaluation Methodology for Face-Recognition Algorithms”, In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), San Juan, Puerto Rico, **(1997)** June 17-19, pp. 137-143.
[24] J. R. Quinlan, “C4.5: Programs for Machine Learning”, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, **(1993)**.

[25] Z. Sun, G. Bebis, X. Yuan and S. Louis, "Genetic feature subset selection for gender classification: A comparison study", In proceedings of the 6th IEEE Workshop on Applications of Computer Vision, **(2002)**.

[26] V. Thomas, N. Chawla, K. W. Bowyer and P. J. Flynn, "Learning to predict gender from iris images", In: Proceedings of the IEEE International Conference on Biometrics: Theory, Applications, and Systems (BATS), **(2007)** September, pp. 1–5.

[27] G. K. F. Tso and K. K. W. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks", Energy, vol. 32, **(2007)**, pp. 1761-1768.

[28] R. Verschae, J. Ruiz-del-Solar and M. Correa, "Gender Classification of Faces Using Adaboost", Progress in Pattern Recognition, Image Analysis and Applications, vol. 4226, **(2006)**, pp. 68-78.

[29] P. Viola and M. J. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, vol. 57, no. 2, **(2004)**, pp. 137–154.

[30] I. H. Witten and E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques", second ed. Elsevier, ISBN: 0-12-088407-0, **(2005)**.

[31] Z. Xu, L. Lu and P. Shi, "A Hybrid Approach to Gender Classification from Face Images. In: Proceedings of the IEEE 19th International Conference on Pattern Recognition (ICPR), Tampa, Florida, **(2008)** December 11, pp. 1-4.

[32] S. Yu, T. Tan, K. Huang, K. Jia and X. Wu, "A Study on Gait-Based Gender Classification", IEEE Transactions on Image Processing, vol. 18, no. 8, **(2009)**, pp. 1905-1910.

[33] M. Zahedi and S. Yousefi, "Gender recognition based on SIFT features", International Journal of Artificial Intelligence & Applications (IJAIA), vol. 2, no. 3, **(2011)**.

[34] Y. Zhiguang, L. Ming and A. Haizhou, "An Experimental Study on Automatic Face Gender Classification", In: Proceedings of 18th International Conference on Pattern Recognition (ICPR), Hong Kong, vol. 3, **(2006)** August 24, pp. 1099-1102.

Authors

Muhammad Naeem Ahmed Khan obtained D.Phil. degree in Computer System Engineering from the University of Sussex, Brighton, England. His research interests are in the fields of software engineering, cyber administration, digital forensic analysis and machine learning techniques.

Sheraz Ahmad Qureshi obtained his Master of Sciences (MS) degree from the Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad in 2011. His research interest is mainly on the application of data mining techniques.

Naveed Riaz received a Ph.D. from Graz University of Technology (2008) in Software Engineering, and an M.S. in Software Engineering (2005) from National University of Sciences and Technology (NUST), Pakistan. His research interests include Model-based and Qualitative Reasoning, Theoretical Computer Sciences, Distributed Computing and Digital Image Processing.