Learning to predict gender from iris images

Vince Thomas, Nitesh V. Chawla, Kevin W. Bowyer, and Patrick J. Flynn

***Abstract*— This paper employs machine learning techniques to develop models that predict gender based on the iris texture features. While there is a large body of research that explores biometrics as a means of verifying identity, there has been very little work done to determine if biometric measures can be used to determine specific human attributes. If it is possible to discover such attributes, they would be useful in situations where a biometric system fails to identify an individual that has not been enrolled, yet still needs to be identified. The iris was selected as the biometric to analyze for two major reasons: (1) quality methods have already been developed to segment and encode an iris image, (2) current iris encoding methods are conducive to selecting and extracting attributes from an iris texture and creating a meaningful feature vector.**

Index Terms— **gender classification, iris, biometric**

I. Introduction

Traditionally, every subject that would need to be identified in a biometric system must be *enrolled*. This means that the subject's biometric data is gathered and stored in the system. Whenever a user attempts to be identified by the system, it matches the user's biometric data with all of the enrolled data and will identify the user if a successful match is made. However, this method does not account for subjects who are not enrolled in the system. Supposing a system attempts to identify a subject that is not enrolled, the only information the system can provide is that there is no matching enrolled subject. However, a system capable of analyzing biometric data could provide enough information about an unknown subject that alternative identity methods could be pursued. It can also enable impostor identification, that is disguised genders.

The focus of this paper is to predict gender based on iris image data. Unlike properties such as ethnicity, the distribution of gender does not vary based on external constraints. Thus, we can be assured that any gender-related patterns discovered in the iris will not be subject to the various biases that may afflict other properties. Moreover, gender allows us to eliminate various biases of ethnicities, procurement environment, etc. We can run a controlled group within each ethnicity, demographic, and procurement environment to weed out the variations in gender predictions and presence of biases, if any. To the best of our knowledge, we are not aware of any other study that attempts to predict gender from irises.

The paper is organized as follows. In section 2 we discuss the related work. In section 3 we present the procedure of extracting feature vector from an iris image. In section 3, we discuss the machine learning algorithm used for training the model from the derived feature vector. Section 4 presents our core results and observations. We conclude with Section 5.

V. Thomas, N. V. Chawla (contact author), K. W. Bowyer and P. J. Flynn are with the Department of Computer Science and Engineering at the University of Notre Dame {vthomas, nchawla, kwb, flynn}@nd.edu

II. Related Work

Currently, few studies of this nature have been made, but none, to the best of our knowledge, have used the properties of an iris image to predict gender. Qiu, Sun, and Tan developed a model to predict ethnicity using iris data ([1]) using the AdaBoost algorithm ([2]) and achieved an accuracy of 85.95%. Their model attempts to predict if a subject is either Asian or non-Asian. While the accuracy is quite high, one concern that must be addressed is potential bias associated with the iris databases they use for each class. All of the Asian images come from the CASIA dataset, while non-Asian images come from the UPOL and UBIRIS datasets. Naturally, this forms a very strong sample selection bias in the paper given different databases of ethnicities. In this paper, it is never addressed whether there is any concern over the prediction models learning on subtle differences in the iris databases. However, simply looking at sample images from each dataset suggests that there may be potential for bias since images from one dataset appear significantly darker than images from the others.

Fig. 1. LG EOU 2200 Station.

Some biometric work has been done on *gender* classification, but all existing work involves the analysis of face image data. Moghaddam and Yang [3] used Support Vector Machines to achieve an accuracy of 96.6% using low resolution ”thumbnails” of face images. Gutta et al. [4] employ an ensemble of RBF networks and inductive decision trees to obtain an accuracy of 96%. Sun et al. also use face images for gender identification [5]. Jain et al [6] approach gender classification from the frontal facial images using a combination of ICA and Support Vector Machines.

III. Data and Feature Creation

In this section we discuss the core step and contribution of our work: *derivation of a feature vector from iris images to predict gender.* We believe the most critical component of a learning task is providing a cogent representation of data.

A. Data Acquisition

We used Iridian LG EOU200 system [7] to acquire all our data. The system has one infra-red LED on the top, one at the bottom left, and one at the bottom right. Figure 1 shows the Iridian data acquisition station. The subjects' images were acquired on a weekly basis, with no subject repeating during a week. The distance between the eyes and the system was around 6 inches. Once a subject is appropriately positioned, the images are acquired in three stages with each of the LED lights getting activated one after the other. The resulting iris images carry a resolution of 640×480 with the intensity value scaled to use an 8-bit range. We ran a rigorous quality check to discard images with poor quality such as motion blur or out-of-focus. Note that both left and right images were acquired during each session.

B. Segmentation and Encoding

We used the “ND_IRIS” software [9] to perform segmentation on the iris images. ND_IRIS deploys canny edge detector [10] circular Hough transformation to perform the segmentation. First, the limbic boundary is detected and then the pupillary boundary. We refer the reader to Xiaomei Liu's dissertation [9] for complete details on the ND_Iris software. Figure 2 shows an acquired iris image before and after segmentation.

The segmentation procedure is followed by texture encoding. The original image is first translated from the cartesian (x, y) image coordinates to polar coordinates (r, θ) using Daugman's rubber sheet model [11]. This results in an unwrapped image of size $20(r) \times 240(\theta)$. Figure 3 shows an example. Each row represents a “ring” through the iris. The first row is closest to the iris and the last row is closest to the sclera. Once the images are unwrapped in this fashion, fast fourier transform (FFT) is applied to each row (treated as a signal). If x is a row of the image, then the FFT results in a vector Y.

$$Y_k = \sum_{1}^{N} x_n e^{\frac{-j2\pi(k-1)(n-1)}{N}} \tag{1}$$

The resulting vector Y is passed through a 1-D Gabor Filter (F), resulting in Z_k. Finally, the inverse FFT is applied to get a new convolved image as shown in vector I.

$$Z_k = Y_k \times e^{\frac{-log(f/f_0)^2}{2log(\sigma/f_0)^2}} \tag{2}$$

$$I_n = \frac{1}{N} \sum_{k=1}^{N} Z_k e^{\frac{-j2\pi(k-1)(n-1)}{N}} \tag{3}$$

The gabor filter out comprises of both the real and imaginary components. Since the phase angle is now quantized into two bits, we arrive at a template of size 20×480. Figure 3 shows a normalized and unwrapped image.

C. Feature extraction

As each iris image is processed, features of the iris are extracted during various phases of the processing. During iris segmentation, variables the radius and center of both the iris and pupil are available and are stored as “geometric” features. In the iris encoding phase, the iris image is represented as an array of complex numbers with 20 rows and 480 columns. We extract the real components at this stage, which will be called as the “texture” features. We considered the imaginary components as well, but our initial investigation implied no immediate advantage to adding the imaginary component so we just persisted with the real components. Both the geometric and texture features are then used for deriving additional features, as will be discussed below.

D. Geometric Features

In this context, geometric features refer to those that pertain to dimensions of the iris. These features are acquired during the segmentation phase of iris image processing and thus contain no information about the actual texture. The measurements for all of these features are in pixels (or, in the case of area-related features, pixels squared). The geometric features are:

- **distX:** The scalar distance between the X-coordinates of the center of the iris and the center of the pupil.
- **distY:** The scalar distance between the Y-coordinates of the center of the iris and the center of the pupil.
- **distCenter:** The scalar distance between the center of the iris and the center of the pupil. Calculated using disX and disY.
- **irisArea:** The total area of the iris and the pupil. Calculated using the iris radius.
- **pupilArea:** The area of the pupil. Calculated using the pupil radius.
- **diffArea:** The difference between irisArea and pupilArea. This gives the true ”area” of the iris alone.
- **areaRatio:** The ratio of irisArea to pupilArea.

E. Texture Features

Texture features are those that were extracted from the iris image during the encoding phase of the image processing. At one point in the encoding process, the iris is represented

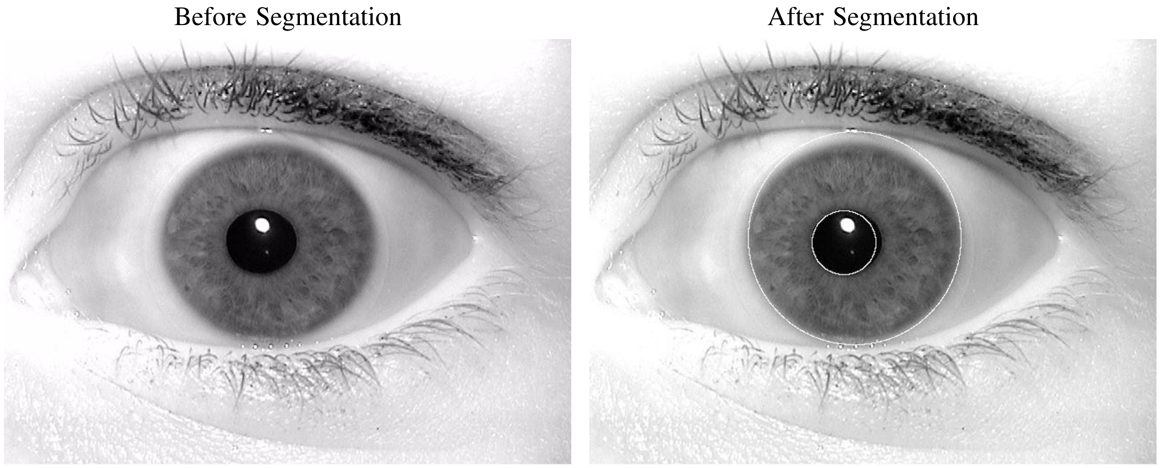

Fig. 2. Sample images of a subject in the training data.

Fig. 3. Unwrapped Iris.

by an array of complex numbers consisting of 20 rows and 240 columns (since we are only using the real component, there are 240 possible values for each row). The rows of the array correspond to concentric rings in the iris. Each column corresponds to a different angle from the center of the iris. Thus, traversing through a row would be analagous to tracing a circle around the iris that is concentric in relation to the pupil and iris boundaries and lies between the two. Traversing down a column is analagous to maintaining a fixed angle from the center of the pupil and travelling from the pupil-iris boundary outwards to the outer edge of the iris. All of the texture features involve operations on this encoded array:

- **rowXRealMean:** The mean value of the real components of the complex numbers in row X.
- **rowXRealStd:** The standard deviation of the real components of the complex numbers in row X.
- **rowXWindowVar:** A measure of variation of the real components of the complex numbers in a row within a window of size 4. This variance is calculated as follows:

$$V(n) = \frac{\sum_{k=2}^{N-2} V(i,j) - \frac{1}{4} * T(i,j)}{N-4} \quad (4)$$

  where

$$T(i,j) = V(i-1,j)+V(i-2,j)+V(i+1,j)+V(i+2,j) \quad (5)$$

- **colVar:** An overall measure of variation of the real components of the complex numbers in the entire array. Instead of a one-dimensional sweep of each row, this feature considers the variance along each column.

Once the feature vectors were constructed, each was assigned a class label of male or female based on the subject's gender.

IV. Machine Learning Framework

The machine learning framework consisted of feature selection, decision tree learning and ensemble methods.

A. Feature Selection

Feature selection is a common process in machine learning, in which a subset of the features available from the data are selected for application of a learning algorithm, such as for classification. Feature selection based on information gain is the most popular method. In this a feature is selected based on the maximum possible information gain from a partition based on the feature.

Specifically, there be d features denoted by $\mathbf{X} = \{X_1, X_2, \ldots, X_d\}$ and c class labels denoted by the set $\mathbf{Y}$. Let S_n denote the set of n instances (image feature vectors). In particular, $S_n = (\mathbf{x}_{(1)}, y_{(1)}), (\mathbf{x}_{(2)}, y_{(2)}), \ldots, (\mathbf{x}_{(n)}, y_{(n)})$ is the set of n ordered pairs, each consisting of a feature vector $\mathbf{x}_{(i)}$ and a class label $y_{(i)}$. Suppose we wanted to determine the information gain resulting from partitioning the examples in S_n at $X_i = v_i$. Let S_n^L be the subset of examples such that $X_i \leq v_i$ and S_n^R be the remaining examples in S_n. Additionally, let $S_{n,y}$ denote the examples in S_n with the class label y. The information entropy in S_n is defined as

$$I(S_n) = \sum_{y \in \mathbf{Y}} \left(-\frac{|S_{n,y}|}{|S_n|} \lg \frac{|S_{n,y}|}{|S_n|} \right). \quad (6)$$

The information gain by partitioning at $X_i = v_i$ is defined as the decrease in entropy due to the partitioning, i.e.,

$$\text{gain}(S_n, X_i, v_i) = I(S_n) - \left[\frac{|S_n^L|}{|S_n|} I(S_n^L) + \frac{|S_n^R|}{|S_n|} I(S_n^R) \right] \quad (7)$$

B. Classifier Learning

We used the C4.5 decision tree algorithm [12] as the core learning technique. Our initial experiments included SVM and neural networks, but they resulted in no obvious gain over just using a simple and fast decision tree technique. So for clarity of experiments and expression in the paper, we restricted to the usage of C4.5 decision tree. Moreover, we wanted to use an unstable classifier to get gains from ensemble methods.

Using C4.5 as the base classifier, we learned *ensembles* using bagging [13] and random subspaces [14], [15] to further improve the accuracy of prediction. Bagging involves making many "bootstrap" aggregates of a training set. A bootstrap replicate is typically of the same size as the original dataset. The random subspace method constructs an ensemble of classifiers on independently selected feature subsets, and combines them using a heuristic such as majority voting, sum rule, etc. A decision tree is learned on each of the subspaces or bootstraps generating an ensemble. We learned ensembles with size up to 100. With random subspaces we tried feature subsets of sizes 5, 10, 15, and 20.

V. Experiments

We performed a comprehensive array of experiments to evaluate the accuracy of our approach for gender prediction. We just included the left irises for all the subjects. We considered the following independent sets.

1) **All**: This included all the images without any partitioning to eliminate the bias from LED illumination or ethnicity or subject. This gave us a total of 57,137 images. We maintained the balance of distribution of males and females.
2) **LED**: We considered one LED at time to eliminate the effect of any specular highlight introduced during acquisition. Each time an iris is enrolled, three images from slightly different angles are captured simutaneously. In order to ensure that the prediction model is not, for example, learning that two of the three images are male and, thus, since the third image has very similar data, must also be male, a subset of the data is created that includes only one image from each acquisition. 10-fold CV was employed in this experiment as well. Each of the three LED sets carried 16,469 images, again balanced. Note that they don't exactly sum up to all images, as we had to drop some images due to imbalance in class distribution between genders.
3) **Ethnicity**: We considered different ethnicities such that each ethnicity was considered separately. Finally, it must be established that predictions are not being made on ethnicity. For example, if it turns out that the majority of the males in the data are asian, it could be that the prediction model is built on features that differentiate ethnicity instead of gender. To ensure this, the data is divided based on ethnicity. White and asian data account for the majority of instances, thus a prediction model is built for each of these ethnicities. We considered the two largest ethnicities in our set — Caucasian and Asians. The Caucasian set comprised of 36,850 images and the Asian set comprised of 16,385 images.

We used 10-fold cross-validation on each of the sets to validate the approach. 10-fold CV involves randomly splitting the instances into ten equally sized partitions. Each partition is used for testing exactly once, while the other nine partitions are used to train a classifier. Thus, 10-fold CV involves ten iterations of training a classifier and testing the model. The mean of the accuracies from each iteration represents an overall accuracy of the model.

A. Results

1) All: No Bias Reduction: Figure 4 shows the results obtained from using the unbiased set. Interestingly, we were able to approach a performance of 75% in accuracy, which is significantly better than random. However, we have not eliminated the effect from any of the biases in the dataset. Another noteworthy observation that will persist throughout each experiment is that bagging produces the highest accuracy, random subspaces with 5 features selected produces the lowest accuracy, and as the number of features selected increases, the accuracy approaches that of bagging.

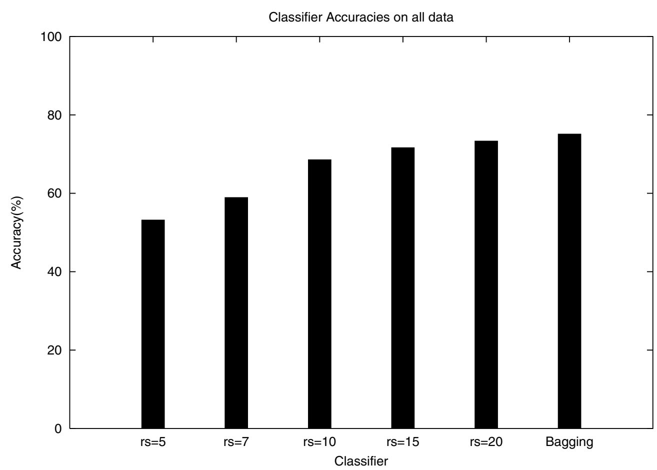

Fig. 4. All subjects.

2) LED Bias Reduction: Since, different LED illuminations introduce different images of a subject, potentially captured on the same day, we wanted to carefully eliminate the artifact of predicting on the same subject, albeit captured under the different light. Figure 5 shows the results. The bagging accuracy for LED bias reduction is very close to the bagging accuracy for no bias reduction, which seems to suggest that using all three images from each acquisition does not substantially improve the results. It is interesting to note that using random subspaces with relatively few features out-performs the corresponding experiments with no bias reduction. This could be due to the fact that the size of the LED data set is one third of the size of the full set. Thus, it may have been able to "overfit" the data with a small number of features.

3) Ethnic Bias Reduction: Figure 6 shows the results from predicting gender for caucasian ethnicity. Interestingly, there is a significant jump in the performance for predicting gender for caucasians. We are able to approach accuracies in the upwards of 80% with Bagging. It is compelling that there

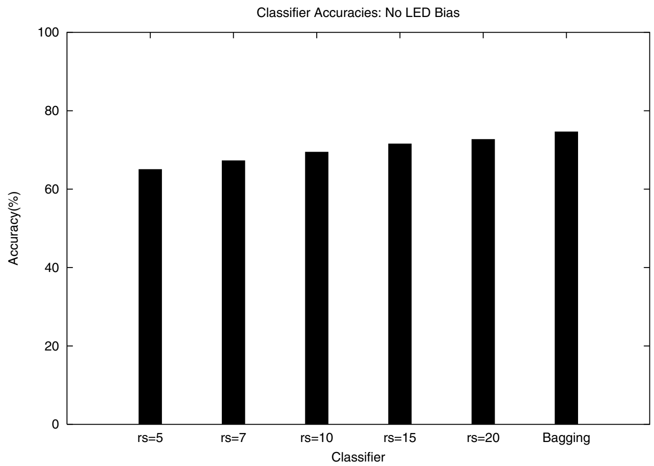

Fig. 5. LED Bias.

is a significant improvement in prediction accuracies over using all the ethnicities together.

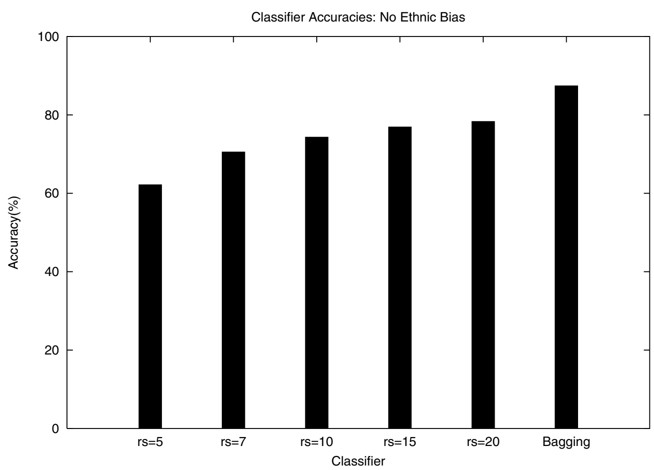

Fig. 6. Ethnic Bias Reduction

VI. Conclusion

Employing decision tree learning made it possible to develop gender classification models that can reach accuracies close to 80%. This is a notable result as it has previously been an unchartered territory. Our ballpark was: *can we do better than guessing at random*. We obviously established that by achieving statistically significantly improved performances. Moreover, we also showed that by eliminating certain biases the performance persists.

One logical extension of this work would be to devise a more complicated feature vector and determine if there are more features, either global or local, that are useful in classifying gender. From here, it may be possible to discover if there are specific singular features that determine gender. In other words, it may be interesting and useful to search for features of particularly high *quality*, while this experiment relies on a large *quantity* of features and possibly interactions between them.

Using similar techniques as those found in this experiment, it may be possible to classify an individual based on attributes besides gender using iris image data.

VII. Acknowledgements

Biometrics research at the University of Notre Dame is supported by the National Science Foundation under grant CNS01-30839, by the Central Intelligence Agency, by the US Department of Justice/ National Institute for Justice under grants 2005-DD-CX-K078 and 2006-IJ-CX-K041, by the National Geo-spatial Intelligence Agency, and by UNISYS Corp.

References

[1] X. Qiu, Z. Sun, and T. Tan, "Global texture analysis of iris images for ethnic classification.," in *Springer LNCS 3832: Int. Conf. on Biometrics*, pp. 411–418, 2006.

[2] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," in *Thirteenth International Conference on Machine Learning*, 1996.

[3] B. Moghaddam and M.-H. Yang, "Gender classification with support vector machines," in *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 306 – 311, 2000.

[4] S. Gutta, H. Wechsler, and P. J. Phillips, "Gender and ethnic of face images," in *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 194 – 199, 1998.

[5] Z. Sun, G. Bebis, X. Yuan, and S. Louis, "Genetic feature subset selection for gender classification: A comparison study," in *Sixth IEEE Workshop on Applications of Computer Vision*.

[6] A. Jain, J. Huang, and S. Fang, "Gender identification using frontal facial images," in *IEEE International Conference on Multimedia and Expo*, 2005.

[7] I. System, "http://iris.nist.gov/ice."

[8] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, University of Western Australia, http://www.csse.uwa.edu.au/pk/studentprojects/libor/, 2003.

[9] X. Liu, *Optimizations in Iris Recognition*. PhD thesis, University of Notre Dame, 2006.

[10] J. Canny, "A computational approach to edge detection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679 – 698, 1986.

[11] J. Daugman, "How iris recognition works," tech. rep., University of Cambridge, http://www.CL.cam.ac.uk/users/jgd1000.

[12] J. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann, 1992.

[13] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[14] T. Ho, "Random subspace method for constructing decision forests," *IEEE Transactions on PAMI*, vol. 20, no. 8, pp. 832 – 844, 1998.

[15] N. V. Chawla and K. W. Bowyer, "Random subspaces and subsampling for 2-d face recognition," in *IEEE CVPR*, pp. 582–589, 2005.