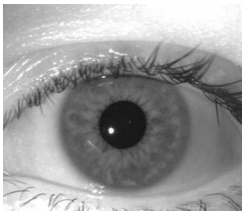
Iris or Periocular? Exploring Sex Prediction from Near Infrared Ocular Images

Denton Bobeldyk[1] and Arun Ross[2]

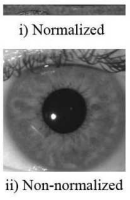
Abstract: Recent research has explored the possibility of automatically deducing the sex of an individual based on near infrared (NIR) images of the iris. Previous articles on this topic have extracted and used only the iris region, while most operational iris biometric systems typically acquire the extended ocular region for processing. Therefore, in this work, we investigate the sex-predictive accuracy associated with four different regions: (a) the extended ocular region; (b) the iris-excluded ocular region; (c) the iris-only region and (d) the normalized iris-only region. We employ the BSIF (Binarized Statistical Image Feature) texture operator to extract features from these regions, and train a Support Vector Machine (SVM) to classify the extracted feature set as Male or Female. Experiments on a dataset containing 3314 images suggest that the iris region only provides modest sex-specific cues compared to the surrounding periocular region. This research further underscores the importance of using the periocular region in iris recognition systems.

Keywords: iris, periocular, biometrics, soft biometrics, sex prediction, binarized statistical image feature (BSIF)

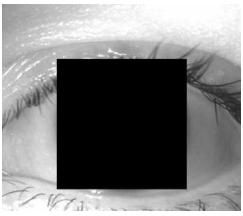
1 Introduction



(a) Ocular Image



(b) Iris-Only Image



(c) Iris-Excluded Ocular Image

Fig. 1: How much sex information is encoded in the iris compared to the surrounding periocular region? Original image taken from [DBF13].

A biometric system uses the physical or behavioral trait of an individual to automatically recognize the individual. The physical or behavioral trait that is used for recognition is referred to as a *biometric trait*. Examples of biometric traits include face, fingerprint, iris, voice, gait and hand geometry [JRN11]. The focus of this current work is on the iris trait. Iris has been observed to be a powerful biometric cue and is currently being used successfully in several large scale projects (e.g., United Arab of Emirates border crossing system and India's Aadhaar program).

[1] PhD Student, Michigan State University and Assistant Professor, Davenport University, denny@bobeldyk.org
[2] Professor, Michigan State University, rossarun@cse.msu.edu

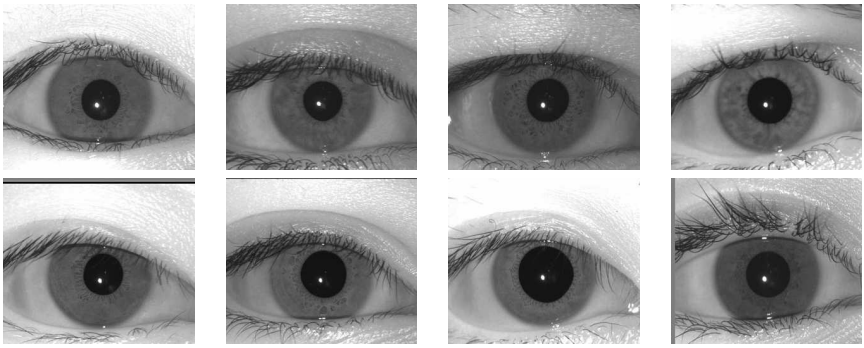


Fig. 2: The top row shows sample ocular images from male subjects, while the bottom row shows samples from female subjects. The images are from [DBF13].

Besides biometric traits, there are other attributes, that may not be unique to an individual, but that can offer useful information about the individual. These are known as soft biometrics. Gender, age, and ethnicity are all examples of soft biometrics [DER16].

There are several advantages to developing algorithms capable of automatically extracting soft biometric traits. While a single soft biometric trait may not offer sufficient information to uniquely recognize an individual, it can be combined along with primary biometric traits in a fusion framework to improve recognition accuracy [JDN04]. Besides improving the recognition accuracy, these attributes provide additional semantic information about an unknown subject that bridges the gap between human and machine descriptions of individuals (e.g., "middle-aged male Caucasian"). Soft biometric algorithms can also be used in scenarios where traditional biometric matchers may not be easily used. For example, it may be possible to extract soft biometric information from poor quality images or multispectral images (e.g., visible versus near-infrared) where traditional biometric matchers are likely to fail. Predicting soft biometric attributes accurately has applications in criminal investigations as well. Specifically, they can be used to *exclude* some suspects from further consideration. In addition to the above advantages, it may be possible to automatically glean aggregate demographic information from a central biometrics database (e.g., age distribution of subjects).

Sex[1]prediction from NIR iris is a relatively new problem.[2] Previous research in this area has explored predicting sex only from the iris region; however, most iris systems capture images of the *extended ocular* region, rather than just the iris (see Figure 1). In the biometrics literature, the region around the eyes is often referred to as the *periocular* region. Thus we raise the following questions:

- Does the extended ocular region offer more sex cues than the iris region?

[1] We use the term "sex" in this paper, rather than "gender", since the former has a genetic basis while the latter is assumed to be influenced by social, personal and psychological factors.

[2] It should be noted that there is a related publication that explores this problem in the RGB color spectrum by cropping out and using the periocular region from a face image [MJS10].

- Does the normalized iris image[3] result in better sex discrimination than the non-normalized iris image?

In this paper we will explore these questions systematically and by designing an experimental protocol that would allow us to draw conclusions for each of them.

2 Related Work

One of the earliest work on automatically deducing sex from iris was conducted by Thomas et al. [Th07]. The authors assembled a dataset of 57,137 ocular images. From each image the iris was automatically segmented and transformed into a 20×240 image using the rubber sheet transformation model [Da03]. Then a 1D Gabor filter was used to generate a feature vector. An information gain metric was used to perform feature selection. A decision tree algorithm was then used to classify the reduced feature vector as 'Male' or 'Female'. The paper does not indicate if a subject disjoint[4] training and test dataset was used; however, it states that only left iris images were used. Their best performance with bias reduction due to ethnicity was 'upwards of 80% with Bagging'.

Bansal et al. [BAS12] were able to achieve an 83.06% sex classification accuracy using statistical and wavelet features along with an SVM classifier. Occlusions from the iris region were removed (i.e., eyelids, eyelashes) using an unspecified masking algorithm. The size of their dataset, however, was quite small with only 150 subjects and 300 iris images. 100 of the subjects were male and 50 of the subjects were female. It is not clear if they used a subject-disjoint evaluation protocol.

An earlier 2011 paper by Lagree and Bowyer [LB11] used texture descriptors to predict gender from an NIR iris image, and explicitly stated that they had used a subject-disjoint training and test set. They collected a total of 600 images from 60 male and 60 female subjects. The iris images were normalized using Daugman's rubber sheet method [Da03] but using a different sampling frequency, resulting in a normalized iris image of 40x240 (as opposed to 20x240). The features were extracted using 6 spot/line detectors and 3 Laws' texture measures [La80]. They used the SMO support vector algorithm but were unable to achieve an accuracy greater than 62% on this dataset.

3 Feature Extraction

Existence of sex-specific attributes in the iris has been alluded to in the medical literature [SL09, LPS03, LP04]. From a computer vision standpoint, it is essential to use a texture descriptor that can potentially extract these attributes. While a number of texture descriptors have been discussed in the literature, the use of Binarized Statistical Image

[3] The normalized iris is a rectangular rendition of the annular iris and is obtained by sampling the segmented iris region in the radial and angular directions using a rubber sheet model [Da04].

[4] A *subject disjoint* experimental protocol stipulates that subjects in the training and test sets should be mutually exclusive.

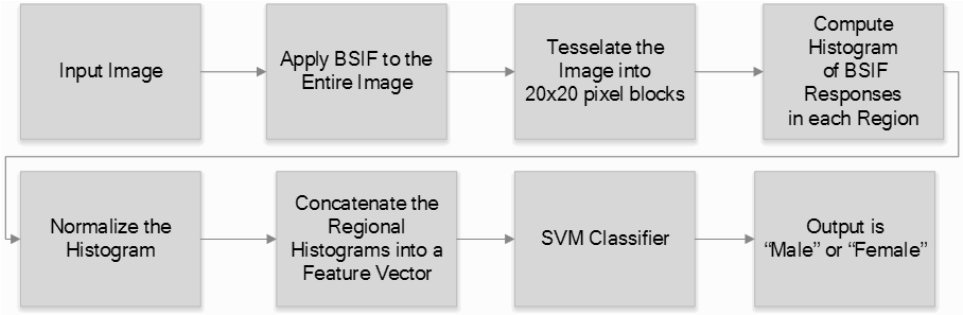
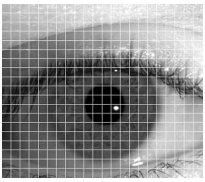
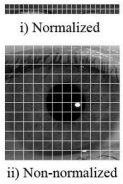


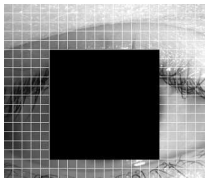
Fig. 3: Block diagram depicting the various stages of the sex prediction algorithm used in this work



(a) Ocular Image



(b) Iris-Only Image



(c) Iris-Excluded Ocular Image

Fig. 4: Tessellations applied to the four image regions. Original image taken from [DBF13].

Features (BSIF) is explored in the context of this work, primarily because our preliminary experiments suggest that this operator outperforms other descriptors on the task of sex classification.

Kannala and Rahtu [KR12] introduced the BSIF texture descriptor. In their paper, the authors showed that BSIF was able to outperform both Local Binary Patterns and Local Phase Quantization on the Outex and CURet texture datasets [KR12]. BSIF projects the image into a subspace generated by convolving the image with pre-learned filters. In order to learn the filters, the authors randomly sampled 50,000 patches of size $k \times k$ from the 13 natural images provided in [HHH09]. Principal Components Analysis was then applied to the sampled data keeping only the top $\mathbf{n}$ principal components. The $\mathbf{n}$ principal components were then subjected to a whitening transform. Independent Component Analysis was then applied to the dimension reduced data resulting in $\mathbf{n}$ filters each of size $k \times k$ (after reshaping each independent component vector into a matrix).

Each of the generated filters is convolved with the input image. The spatial filter responses due to each filter are then binarized by comparing them to a threshold. Thus, at each pixel there are $\mathbf{n}$ binary responses corresponding to the $\mathbf{n}$ filters, i.e., each pixel can now be encoded as a $\mathbf{n}$-bit string. This binary string is then converted to a decimal value.

For example, if the binarized responses for the first, second, third, fourth and fifth filter are 0, 1, 1, 0, and 1 respectively, the resulting decimal value would be 13, since $(01101)_2 = (13)_{10}$.

When the aforementioned procedure is applied to an image, the resulting response matrix will be the same size as the image. Each response will be in the integral range of $[0, 2^{\mathbf{n}-1}]$. For example, when using $\mathbf{n} = 6$ filters, the response values will fall in the range of [0, 63]. The aforementioned process was repeated by changing n in the interval [5,12] and using the following patch sizes: 3×3, 5×5, 7×7, 9×9, 11×11, 13×13, 15×15, 17×17. If the raw response values were used to create a feature vector, the length of the feature vector would be the same size as the image. In order to reduce the size of the feature vector and to obtain local statistical information, the image was tesselated into 20×20 square regions (see Figure 4). A histogram of the BSIF responses was calculated for each region. Each of the regional histograms was normalized and then concatenated to form a single feature vector. This vector was then input to a trained binary SVM classifier in order to predict the sex: male or female.

4 BioCOP Dataset

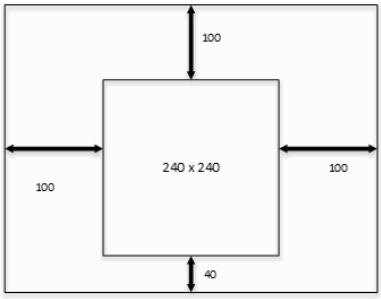


Fig. 5: Geometrically adjusted ocular image

(a) Before (b) After

Fig. 6: Example of a geometrically adjusted image. Original image taken from [DBF13].

The anonymized BioCOP iris dataset was used in this research. The images in the dataset were obtained using a near-infrared (NIR) sensor. Using a commercial off-the-shelf iris SDK, the center of the iris and its radius were automatically located. The iris was then centered horizontally, and the image was geometrically scaled such that the iris had a fixed radius of 120 pixels. The scaled image was then cropped around the repositioned iris region so as to have a 40-pixel border below the iris and 100-pixel borders on the top and sides. The size of the scaled and cropped image was 380×440. See Figure 5. A total of 181 images, corresponding to about 5% of the entire dataset, were discarded during this step (for example, some images did not include the whole iris or could not be centered

appropriately). The final dataset that was used consisted of 580 male subjects with 1720 images and 503 female subjects with 1594 images (please see Table 1 for a more complete breakdown). For each subject, images from both the left and right irides were included when available.

Tab. 1: The subset of the BioCOP iris dataset that was used in our experiments

Sex	#of Subjects	#Left Images	#Right Images	Total # Images
Male	580	889	831	1720
Female	503	822	772	1594

5 Experiments

Previous work [TPB14] has demonstrated success when the uniform LBP(8,2) texture operator was used for sex classification of the iris. However, our initial experiments showed that BSIF outperforms uLBP(8,2) in this problem domain.[5]

We conduct experiments to compare the sex prediction accuracy of the iris with that of the periocular region. From each geometrically adjusted ocular image in the dataset, three different sub-images were extracted: iris-excluded ocular image, iris-only image, and normalized iris-only image. This resulted in the following four regions.

Ocular Image: This is the entire scaled and cropped operational iris image (380×440). See Figure 1(a).

Iris-Only Image: The portion of the ocular image which encloses the entire iris region. The center of the image coincides with the center of the iris and the width of the image is twice the iris radius resulting in 240×240 images. No masking was performed to remove the eyelid or eyelash pixels. See Figure 1(b)(ii).

Normalized Iris-Only Image: The unwrapped iris-only image using Daugman's rubber sheet method [Da03]. The iris was sampled 20 times radially and 240 times angularly resulting in a 20×240 rectangular image. See Figure 1(b)(i).

Iris-Excluded Ocular Image: The ocular image with the iris-only region excluded. The portion of the image that was removed was zeroed out; essentially creating a black square in the middle of each of the images. See Figure 1(c).

In order to capture both local and global spatial information, each image was tessellated into 20×20 blocks. Due to the small size of the normalized iris-only image, it was tesselated into 10×10 blocks. The BSIF operator was applied to the entire image and a his-

[5] In [TPB14] the authors claim that their dataset has 1500 unique subjects, though our investigation has revealed a much smaller number of distinct subjects. The authors of [TPB14] have confirmed via email that there were significant errors in their subject labels. When the *incorrect* labels are used, the proposed BSIF approach resulted in a 96.1% accuracy on the left iris compared to the 91.33% accuracy obtained in [TPB14]. The high accuracy is due to *overlapping* subjects in the training and test sets that is a consequence of incorrect labeling.

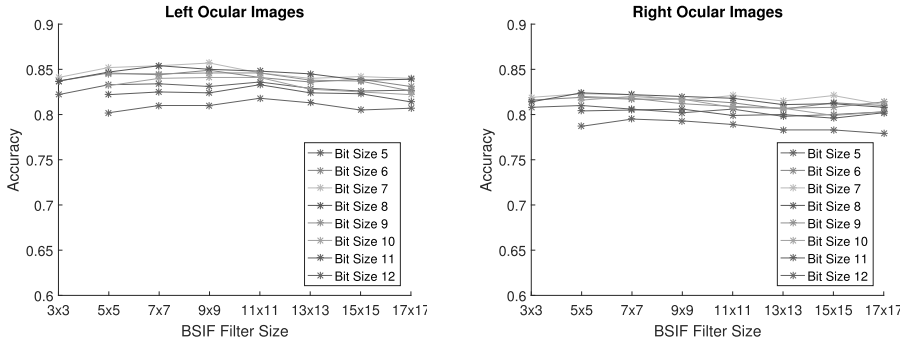


Fig. 7: **Ocular image:** Sex prediction average accuracies for various BSIF bit lengths and filter sizes

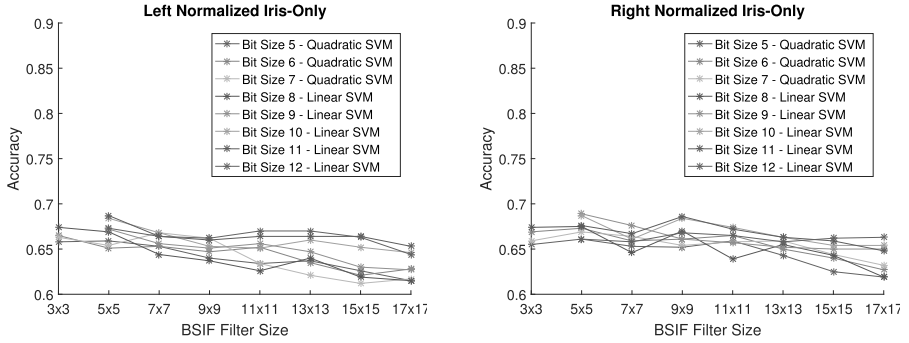


Fig. 8: **Normalized iris-only image:** Sex prediction average accuracies for various BSIF bit lengths and filter sizes

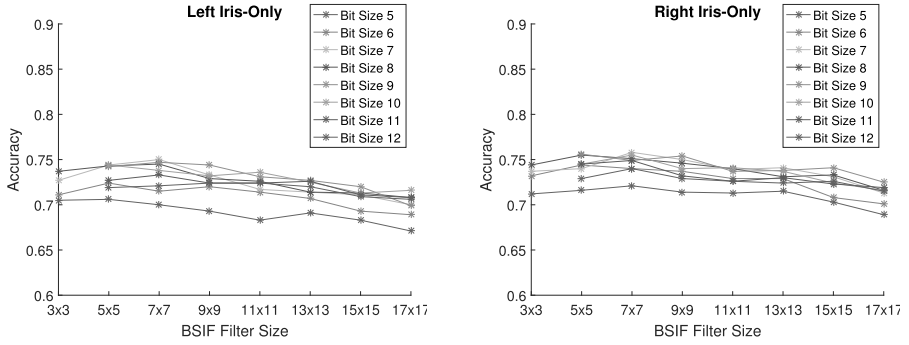


Fig. 9: **Iris-only image:** Sex prediction average accuracies for various BSIF bit lengths and filter sizes

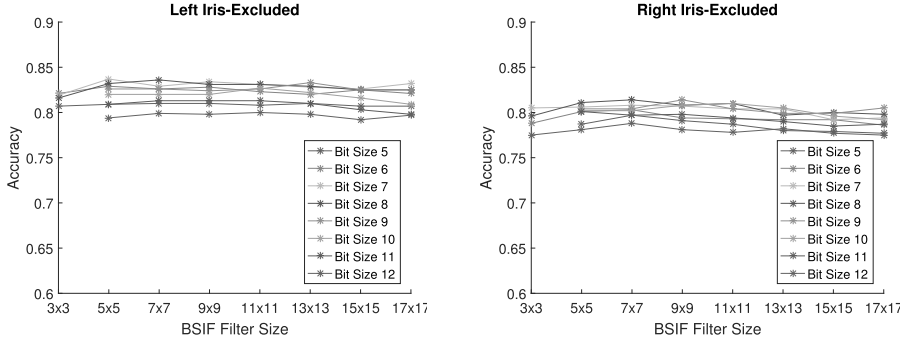


Fig. 10: **Iris-excluded ocular image:** Sex prediction average accuracies for various BSIF bit lengths and filter sizes

togram of the BSIF responses was computed for each block. Each histogram value was divided by the sum of the histogram values for that block, thereby normalizing it. The normalized histograms were concatenated together to form a feature vector that was input to a Support Vector Machine with a linear kernel.[6]

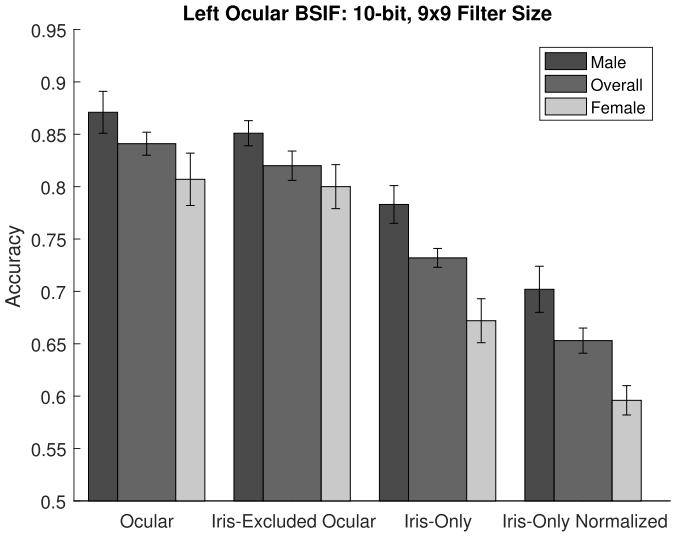


Fig. 11: Results of the sex prediction accuracy for a particular combination of k and $\mathbf{n}$ on each of the four regions considered in our experiments

5.1 Results

Each experiment was conducted using 60% of the subjects in the BioCOP dataset for training and 40% for testing. This subject-disjoint partitioning exercise was done 5 times. Further, the impact of the number of filters (the bit length, $\mathbf{n}$) and the size of each filter (k) on prediction accuracy was studied. Figures 7, 8, 9 and 10 report the accuracies corresponding to the four regions considered in this work. Results for the left and right eyes are shown separately in each figure. In each graph, the average classification accuracy (over the 5 different trials) for different combinations of $\mathbf{n}$ and k is reported. While change in filter size does not seem to have a drastic impact on sex prediction for all 4 regions, change in bit length does have a somewhat discernible impact. In general the smallest bit lengths (5-6) were outperformed by the slightly larger bit lengths (7-10). Increasing the bit length to 11 or 12, however, seemed to lower the accuracy for all 4 regions.

Figure 11 shows the male and female classification accuracies, along with the overall accuracy, for each of the four regions. The performance corresponds to the 10-bit BSIF operator

[6] For some combinations of k and $\mathbf{n}$, the quadratic kernel outperformed the linear kernel. This has been noted in the legend of the performance graphs.

Tab. 2: Percentage of test images correctly/incorrectly classified by the iris-excluded ocular region **and** correctly/incorrectly classified by the iris-only region (Left eye, BSIF parameters: bit length = 10, filter size = 9×9)

		Iris-Only	
		Correct	Incorrect
Iris-Excluded Ocular	Correct	$64.3 \pm 2\%$	$18.4 \pm 0.7\%$
	Incorrect	$8.7 \pm 1\%$	$8.6 \pm 0.4\%$

with 9×9 filters. The ocular region and the iris-excluded ocular region exhibit the best performance while the normalized iris-only image exhibits the worst performance, with almost a 20% difference in performance over the ocular region. Further, the male classification accuracies are observed to be higher than the female classification accuracies. This could be partly attributed to the larger number of male subjects than female subjects in the dataset.

Table 2 shows the prediction relationship between the left iris-excluded ocular region and the left iris-only region. The values in this table are based on the 10-bit BSIF operator with 9×9 filters. The table indicates that there is a potential for fusing the outputs of these two regions which could possibly result in a higher overall prediction accuracy.

6 Discussion

A number of different observations can be made based on the experiments conducted in this work.

- Both the iris and surrounding ocular region manifest sex-specific attributes that can be captured using a texture descriptor.
- The ocular region surrounding the iris results in better sex classification accuracy than the iris-only region (normalized or non-normalized). We speculate that the *morphology* and *structure* of the ocular region provide sex-specific cues in the iris-excluded and extended ocular regions. For the iris-only region, the sex-specific cues could be due to the stromal *texture* with its crypts.
- The ocular region - which could be viewed as the spatial addition of the iris-excluded region to the iris-only region - provides only a modest increase in accuracy over the iris-excluded ocular region. This suggests that (also see Table 2) there could be a better way to combine the extracted features from iris-only region with that of the iris-excluded region (as opposed to simply concatenating them into a single feature vector).
- The non-normalized NIR iris image results in better sex prediction performance than the normalized iris region, thereby suggesting that the normalization process may be filtering out some useful information.

- The proposed approach outperforms the previous best results reported by Lagree and Bowyer [LB11] that is known to have a subject-disjoint train-test protocol.[7] In [LB11], the best performance was 62%, while in this paper the best performance is 85.7%.

The size of the feature vectors generated from the ocular and iris-only images also vary substantially in size. For 5-bit BSIF, the ocular image results in a feature vector of length 13,376 while the iris-only image results in a feature vector that is approximately one third of that (4,608). This disparity in feature dimensions may also have impacted the sex classification accuracies of the two regions. Even though iris recognition algorithms use the normalized iris-only image, it appears that for sex prediction, the non-normalized iris region and the surrounding periocular region provide more cues.

7 Future Work

In the current work, the same feature descriptor (viz., BSIF) was used to encode both the iris and the periocular regions. However, it is necessary to investigate if these regions have to be encoded using different feature descriptors prior to combining them for the sex prediction task. It is also necessary to determine if race or age of the subject impacts the accuracy of sex prediction.

We would like to study the possibility of combining the results of multiple patch sizes (k) and multiple filters ($\mathbf{n}$) of the BSIF approach in order to further improve the sex classification accuracy. We would also like to combine the results of the left and right ocular regions for improving performance. However, it is likely that there is an upper bound on accuracy that is dictated by nature itself that would preempt the possibility of obtaining perfect classification.

The BSIF feature vector used in this work could also be applied to the problem of race and age classification from the ocular image. Our preliminary investigation in this regard is promising. We also plan on developing a more holistic approach that predicts sex and race (and possibly age group) simultaneously.

8 Acknowledgement

This work was partially funded by NSF-CITeR.

References

[BAS12] Bansal, Atul; Agarwal, Ravinder; Sharma, Rajendra: SVM Based Gender Classification Using Iris Images. In: Proc. of International Conference on Computational Intelligence and Communication Networks (CICN). pp. 425–429, Nov 2012.

[7] Bansal et al. [BAS12] use a 10-fold cross-validation. So the number of test images in each trial is only 30 images. Given this extremely limited number of test images, we do not include their results in our comparison.

[Da03] Daugman, John: The Importance of Being random: Statistical Principles of Iris Recognition. Pattern recognition, 36(2):279–291, 2003.

[Da04] Daugman, John: How Iris Recognition Works. IEEE Transactions on Circuits and Systems for Video Technology, 14(1):21–30, 2004.

[DBF13] Doyle, James S; Bowyer, Kevin W; Flynn, Patrick J: Variation in Accuracy of Textured Contact Lens Detection Based on Sensor and Lens Pattern. In: Proc. of IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS). pp. 1–7, 2013.

[DER16] Dantcheva, Antitza; Elia, Petros; Ross, Arun: What Else Does Your Biometric Data Reveal? A Survey on Soft Biometrics. In: IEEE Transactions on Information Forensics And Security (TIFS). volume 11, pp. 441–467, 2016.

[HHH09] Hyvärinen, Aapo; Hurri, Jarmo; Hoyer, Patrick O: Natural Image Statistics: A Probabilistic Approach to Early Computational Vision, volume 39. Springer Science & Business Media, 2009.

[JDN04] Jain, Anil K; Dass, Sarat C; Nandakumar, Karthik: Can Soft Biometric Traits Assist User Recognition? In: Defense and Security. International Society for Optics and Photonics, pp. 561–572, April 2004.

[JRN11] Jain, Anil K.; Ross, Arun; Nandakumar, Karthik: Introduction to Biometrics. Springer, New York, 2011.

[KR12] Kannala, Juho; Rahtu, Esa: BSIF: Binarized Statistical Image Features. In: Proc. of International Conference on Pattern Recognition (ICPR). pp. 1363–1366, 2012.

[La80] Laws, Kenneth: Textured Image Segmentation. PhD thesis, University of Southern California, January 1980.

[LB11] Lagree, Stephen; Bowyer, Kevin W: Predicting Ethnicity and Gender from Iris Texture. In: IEEE International Conference on Technologies for Homeland Security (HST). IEEE, pp. 440–445, 2011.

[LP04] Larsson, Mats; Pedersen, Nancy L: Genetic Correlations Among Texture Characteristics in the Human Iris. Molecular Vision, 10:821–831, 2004.

[LPS03] Larsson, Mats; Pedersen, Nancy L; Stattin, Håkan: Importance of Genetic Effects for Characteristics of the Human Iris. Twin Research, 6(03):192–200, 2003.

[MJS10] Merkow, Jameson; Jou, Brendan; Savvides, Marios: An Exploration of Gender Identification Using Only the Periocular Region. In: Proc. of IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS). pp. 1–5, 2010.

[SL09] Sturm, Richard A; Larsson, Mats: Genetics of Human Iris Colour and Patterns. Pigment Cell & Melanoma Research, 22(5):544–562, 2009.

[Th07] Thomas, Vince; Chawla, Nitesh V; Bowyer, Kevin W; Flynn, Patrick J: Learning to Predict Gender from Iris Images. In: Proc. of IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS). pp. 1–5, 2007.

[TPB14] Tapia, Juan E; Perez, Claudio A; Bowyer, Kevin W: Gender Classification from Iris Images Using Fusion of Uniform Local Binary Patterns. In: Proc. of ECCV Workshops. Springer, pp. 751–763, 2014.