Sheffield Hallam University

Exploring gender prediction from iris biometrics

DA COSTA ABREU, Marjory <http://orcid.org/0000-0001-7461-7570>, FAIRHURST, M. and ERBILEK, M.

Available from Sheffield Hallam University Research Archive (SHURA) at:

http://shura.shu.ac.uk/27292/

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

DA COSTA ABREU, Marjory, FAIRHURST, M. and ERBILEK, M. (2015). Exploring gender prediction from iris biometrics. In: 2015 International Conference of the Biometrics Special Interest Group (BIOSIG). IEEE.

Copyright and re-use policy

See http://shura.shu.ac.uk/information.html

Sheffield Hallam University Research Archive
http://shura.shu.ac.uk

EXPLORING GENDER PREDICTION FROM IRIS BIOMETRICS

Michael Fairhurst[1], Meryem Erbilek[2], Marjory Da Costa-Abreu[3]

[1]School of Engineering and Digital Arts, University of Kent, Canterbury, Kent CT2 7NT, UK
[2]Computer Engineering Department at the Girne American University, Kyrenia, Cyprus
[3]DIMAp, UFRN, Natal, RN 59078-970, Brasil

Abstract: Prediction of gender characteristics from iris images has been investigated and some successful results have been reported in the literature, but without considering performance for different iris features and classifiers. This paper investigates for the first time an approach to gender prediction from iris images using different types of features (including a small number of very simple geometric features, texture features and a combination of geometric and texture features) and a more versatile and intelligent classifier structure. Our proposed approaches can achieve gender prediction accuracies of up to 90% in the BioSecure Database.

1 INTRODUCTION

The estimation of soft-biometric characteristics of individuals based on extractable features of conventional biometric data has become a very important research topic. Biometric-based estimation of characteristics such as gender, age, and ethnicity is performed by using physical and/or behavioural characteristics embedded in an individual's biometric data. This can be particularly useful in many practical scenarios (checking entitlement claims, for example) including, obviously, forensic investigations. In this paper, our focus is gender prediction from iris biometrics. The literature shows that face biometrics have received the greatest attention in relation to gender prediction [RB11, CCP$^+$11, WM09, FD12]. This is perhaps not surprising since it is particularly natural and easy to obtain face images for applications such as criminal investigations or profiling from CCTV cameras. However, considerable effort has also been invested in estimating gender from other biometric modalities such as voice [MAE$^+$07] and text [PDVV11] characteristics. On the other hand, if we consider the predictive properties of the iris in relation to gender characteristics of individuals, only two relevant reported studies [TCBF07, LB11] can be found. Indeed, this is a potentially very challenging task, since gender information is not evident from direct human visual inspection of iris images.

In [TCBF07], gender prediction is carried out using both geometric and texture features of iris images, and using bagging with the C4.5 decision tree classifier. This proposed gender prediction method was able to achieve 75% and 80% accuracy when tested respectively on the whole dataset and on a subset of this dataset corresponding only to Caucasian subjects. By contrast, in [LB11], gender prediction is carried out using only texture features of iris images, but adopting a different type, and a larger number of texture features than in [TCBF07] while using a support vector machine classifier. When tested on the whole dataset and on a subset corresponding only to single ethnicity subjects, this method was able to achieve an accuracy of around 62% in both cases. Possible reasons for this reduction in the attainable accuracy have been set out and explained in [LB11], summarised as follows:

- Differences in the dataset sizes: experiments in [TCBF07] used over 28,000 images whereas in [LB11] 600 images were used, with a factor of around 50 difference in the training set size.
- Differences in the feature vectors: the results in [TCBF07] are obtained with combined features computed on the log-Gabor filtered version of the iris image and geometric features, whereas in [LB11] features based on simple spot, line and Laws texture measures were used, without geometric features.
- Differences in the classification structure: the results in [TCBF07] were obtained using a multiclassifier configuration (bagging 100 C4.5 decision tree [Qui93]), whereas results in [LB11] were obtained with a single classifier (support vector machine).

A proposed technique for gender prediction from iris samples was presented in [TPB15]. In this paper, once again, the authors use only iris texture and they claim up to 91% accuracy using a variation of fusion of uniform local binary patterns.

An analysis of ageing issues in iris biometrics [FE11] shows that physical ageing effects in iris samples are primarily the result of the physiology of pupil dilation mechanisms, with pupil dilation responsiveness decreasing with age. Hence, pupil dilation is very likely to be related to the geometric appearance of the pupil and the iris, where these findings suggest that geometric features of the iris may also provide useful information for the gender-based biometric prediction task.

Therefore, in this paper, we will investigate and explore the gender prediction task with respect to three different approaches which respectively use (a) only geometric features, (b) only texture features and (c) both geometric and texture features extracted from iris images, and we will use more versatile and intelligence-rich classification structures. We will compare achievable error-rate performance and execution times for each approach.

As a result, this empirical comparative study will allow us to work towards developing an optimal approach to adopt for the gender prediction task by taking into account both feature types and classification structure choices. It will thus shed some new light on the issues noted above (effects of feature vectors and classifiers on achievable gender prediction accuracy). The effects of data size are not discussed because this does not apply in our work, since we do not experience such discrepancies in dataset size. The study will provide useful information to inform a choice of features in relation to particular application requirements. Specifically, this will define the flexibility available for trade-off between accuracy and classifier processing speed as a means of responding to application-specific constraints.

2 GENDER ESTIMATION USING IRIS IMAGES

The basic processing of biometric data in our iris-based gender prediction approach is illustrated in Figure 1. These systems typically adopt a process based on the following:

- An eye image is captured in the *Acquisition* step.
- The *Segmentation* step localises the iris region from the acquired eye image. This step involves detection of the sclera/iris and pupil/iris boundaries.
- The *Feature extraction* step extracts geometric, texture or both geometric and texture features of the iris according to the configuration required.
- The *Prediction* step uses the data generated at the output of the previous step and performs the gender classification task itself.

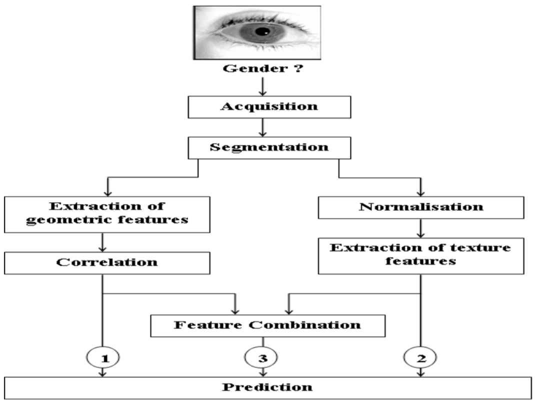

Figure 1: Gender estimation processing: path 1, 2 and 3 refers to the approach 1, approach 2 and approach 3 respectively

Each of these steps will be explained and discussed in more detail in the following subsections.

2.1 Acquisition

The Data Set 2 (DS2) of the BioSecure Multimodal Database (BMDB) [OGFAF$^+$10] is used in this study. The samples were collected as part of an extensive (and commercially available) multimodal database by 11 European institutions participating in the BioSecure Network of Excellence. The eye images were acquired in a standard "office" environment managed by a supervisor and using the LG Iris Access EOU3000 set-up. During the acquisition, spectacles were not allowed to be worn by subjects, although contact lenses were allowed. Four eye images (two left and two right) were acquired in two different sessions with a resolution of 640*480 pixels, for 210 subjects in total. However, the iris samples of 10 subjects were found to be incorrectly labelled in this database (some of the left eye samples labelled as right or vice versa), and were thus discarded. Hence, this decreased the available number of subjects to 200 (a total of 1600 images).

2.2 Iris segmentation

Using the defined iris dataset, each eye sample is first segmented using the automatic segmentation algorithm as described in [FE11, EF11]. In the event of segmentation failure (this occurred for only 1.87% of images), we segment the irises manually and make sure that all eye images are correctly segmented in order to guarantee the reliability of the further analysis. Subsequently, the obtained iris and pupil parameters from the segmentation process are stored for each eye, to be used in the further processing stages. The same features, also used in [FE11, EF11] can be described in Table 1.

Feature No.	**Feature Calculation**
TF1	Mean of the real components of the complex numbers in row X
TF2	Standard deviation of the real components of the complex numbers in row X
TF3	Variance of the real components of the complex numbers in row X
TF4	Mean of the real components of the complex numbers in col Y
TF5	Standard deviation of the real components of the complex numbers in col Y
TF6	Variance of the real components of the complex numbers in col Y

Table 1: Texture features

2.2.1 Approach 1: Geometric feature extraction and correlation

By using the iris and the pupil parameters saved during the segmentation stage, several features which are related to the geometric characteristics of the iris are extracted. Here, it is important to note that the extraction of these features is computationally simple and fast, since none of them requires the extraction of texture information relating to the iris patterning (see Figure 1 (path1)).

The parameters which were obtained at the segmentation stage are; p_x (which is the x-coordinate of the centre of the pupil), i_x (which is the x-coordinate of the centre of the iris), p_y (which is the y-coordinate of the centre of the pupil), i_y (which is the y-coordinate of the centre of the iris), i_r (which is the iris radius), and p_r (which is the pupil radius). By using the pupil and iris parameters defined above, 12 (GF1-GF12) geometric features are extracted for our experimental study. Features GF1-GF7 were similarly defined and adopted as in [TCBF07], while the remaining five features are specific to this study and adopted from [EFDCA13]. A brief description of these features (specified at the pixel level) is shown in Table 2.

Following the extraction of the 12 geometric features defined in Table 2, a correlation

Feature No.	Feature Calculation
GF1	$\|p_x - i_x\|$ (distance in x)
GF2	$\|p_y - i_y\|$ (distance in y)
GF3	$\|GF1 - GF2\|$ (distance from centres)
GF4	$\pi * i_r^2$ (area iris)
GF5	$\pi * p_r^2$ (area pupil)
GF6	$GF4 - GF5$ (true area iris)
GF7	$GF4/GF5$ (area ratio)
GF8	i_r/p_r (dilation ratio)
GF9	$p_i * 2 * i_r$ (iris circumference)
GF10	$p_i * 2 * p_r$ (pupil circumference)
GF11	$GF9/GF10$ (circumference ratio)
GF12	$GF9 - GF10$ (circumference diff)

Table 2: Geometric features

evaluation across the features is carried out as in [EFDCA13]. By removing the highly correlated features, efficiency is increased by adopting only the more distinguishing and non-redundant features. The inter-feature correlations were evaluated by using Spearman's rank correlation [Spe04] (a nonparametric-based estimate of correlation).

2.2.2 Approach 2: Normalisation and texture feature extraction

As illustrated in Figure 1 (path2), after the segmentation stage, this approach performs a normalisation step. This step transforms the iris region into a fixed rectangular block, so that the iris region extracted from the overall eye image is presented at the fixed size necessary for comparisons between samples. A technique [Mas03] based on Daugman's rubber sheet model is employed, which produces a 2D array with horizontal dimensions of angular resolution and vertical dimensions of radial resolution. This produces an unwrapped image of size 20*240 pixels.

Following the normalisation, 1D Log-Gabor wavelets are used to encode features [Mas03]. Each row of the 2D normalised iris pattern corresponds to a circular ring on the iris region. These rows are divided into a number of 1D signals and convolved with 1D Log-Gabor wavelets which outputs a template of size 20*480 with both real and imaginary components. As in [TCBF07], we only use the real components (which correspond to the array of complex numbers of size 20*240 of the template) to extract texture features, which are defined in Table 1. Features $TF1$, $TF2$, $TF6$ were similarly defined and adopted in [TCBF07], while the remaining three features are specific to this study and adopted from [EFDCA14].

2.2.3 Approach 3: Combining geometric and texture features

As shown in Figure 1 (path3), this approach simply adopts the combination of approach 1 and approach 2. Hence, geometric and texture features obtained from approach 1 and approach 2 respectively, are combined simply by concatenating them.

2.3 Prediction

The gender prediction task involves the specification of how to form the training and testing sets as well as the classification method to be applied, which may be described as follows:

2.3.1 Forming testing and training sets

In order reliably to evaluate the performance of the gender classification task, we divide the available samples into person-disjoint testing and training sets. Thus, samples from approximately 72% of the male and the female subjects are used as a training set and the remaining subjects' samples are used as a testing set. The available number of images in the testing and the training sets for each gender group is shown in Table 3.

Sets	**Male group**	**Female group**
All	808	792
Training	576	568
Testing	232	224

Table 3: Number of images

2.3.2 Classification techniques

One of the more difficult aspects of designing any classification task is making the best choice of classifier or, in the case of a multiclassifier approach, choosing the set of base classifiers for the fusion method. A guarantee of high diversity among the individual components is essential in the latter context. In order to achieve diversity, we have selected a pool of well known classifiers that have fundamentally different base structures for this experimental study which are listed below:

- Multi-Layer Perceptron (MLP) [Hay99]
- Support Vector Machine (SVM) [FAE08]
- Optimised IREP (Incremental Reduced Error Pruning) (JRip) [FW94]
- Decision Tree (DT) [Qui93]

- K-Nearest Neighbour (KNN) [Ary98]

In order to analyse the full potential of using geometrical, texture and both geometrical and texture features, we have also considered a range of traditional fusion techniques and more intelligent combination techniques which can be described as follows:

- Sum-based fusion (Sum) [KA03] is a linear fusion-based method that takes into account the confidence degree for each class of each classifier. Thus, when an input pattern is presented to the base classifiers, the degrees of confidence for each class output are added to the other related outputs giving an overall score for that class. The winner class, and hence the identity label computed by the system, is the class with the highest score.
- Majority Voting (Vote) [Kun04] is a non-linear fusion-based classifier combination method that takes into account only the top outputs of the component experts/classifiers. The outputs of the classifiers are represented in a winner-takes-all form (for each classifier, the output of the winner is 1 and the remaining outputs are 0) and the weights for all the component experts are equal to 1.
- Bagging [BB96] is a multiclassifier technique which attempts to neutralise the instability of learning methods by simulating the process of sampling a fresh, independent training dataset each time, the original training data is altered by deleting some instances and replicating others. Instances are randomly sampled, with replacement, from the original dataset to create a new one of the same size. In our approach, we have used the C4.5 decision tree [Qui93] algorithm as the base classifier.

We are especially interested in the use of intelligent agent-based architectures, which we have shown to be well suited to processing biometric data (see, for example, [DCAF11, AF09]). In this paper, we have chosen to analyse the performance of two different techniques that will be described below.

- The Sensitivity-based Negotiation Method (Sens) uses the idea of decreasing the confidence level of an individual agent based on a sensitivity analysis during the testing phase. This analysis can be achieved by excluding and/or varying the values of an input feature and analysing the variation in the performance of the classifier. The main aim of this analysis is to investigate the sensitivity of a classifier to a certain feature and to use this information in the negotiation process.
- The Game Theory-based Negotiation Method (GT) has been used as a cooperation tool in multi-agent systems. In game theory, the systematic description of the results can be carried out through the use of the concept of strategic games. A strategic game is a game in which a player chooses a plan of action only once and at the same time as his opponent. In order to help the players to make their decisions, a payoff matrix is used, in which each cell represents the payoff values which the players will have in a situation where these actions are chosen. The cell with the highest value is chosen.

3 EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we will present experimental results for the three approaches defined above, which are based on geometric (approach 1), texture (approach 2) and both geometric and texture (approach 3) features. We will analyse the proposed gender prediction approaches with respect to both the accuracy achieved and the execution time incurred at the classification stage, after the features were extracted and selected, using a Pentium IV computer with 2.40 GHz processor and 2048 MB RAM. The classifiers were implemented in Java.

For approach 1, all iris samples in the dataset are processed to form the biometric templates, passing through the steps of segmentation, geometric feature extraction and correlation as described in Section 2. Here, highly correlated features are designated as those with a correlation value greater than 0.4 ($-0.4 \leq \rho \leq 0.4$) as in [EFDCA13]. These features are discarded. The remaining uncorrelated 5 features ($GF1$ - $GF4$, and $GF8$) are used to form a feature vector for each iris sample in the dataset (with size of $1 * 5$).

For approach 2, all iris samples in the dataset are processed to form the biometric templates, passing through the steps of segmentation, normalisation and texture feature extraction as described in Section 2. Six texture-related features are used to form a feature vector for each iris sample in the dataset (with size of $1 * 780$).

For approach 3, the geometric and texture features from approach 1 and approach 2 are combined to form a feature vector for each iris sample in the dataset (with size of $1 * 785$).

An initial experiment is performed to test the accuracy achieved and the execution time incurred at the classification stage of the proposed prediction approaches by using the defined feature vectors. The results are shown in Table 4.

Approach	**Results**	**SVM**	**MLP**	**Jrip**	**KNN**	**DT**
1	ACC (%)	55.68	57.86	56.64	49.61	56.81
	ET (sec)	0.39	0.98	0.29	0.37	0.51
2	ACC (%)	65.68	67.86	56.03	59.61	66.81
	ET (sec)	1.47	0.49	0.41	0.74	1.47
3	ACC (%)	81.43	76.64	64.51	73.72	81.43
	ET (sec)	1.97	0.89	1.27	1.31	1.97

Table 4: Accuracy (ACC) and execution time (ET) of individual classifiers

The results obtained show that approach 2 (texture features) achieves a better prediction accuracy rate than approach 1 (geometric features) with all classifiers (except the Jrip classifier) while approach 1 completes the classification stage with lower execution time than approach 2 with all classifiers. This suggests that texture features provide more useful information for the gender prediction task. The results also show that approach 3 achieves the best error-rate performance with all classifiers, but with the highest execution time. Of course, this result is not surprising, since approach 3 is the combination of approach 1 and approach 2 (i.e. adopts both geometric and texture features).

Considering these results further from the classification perspective, it is unsurprising to note that different classifiers return the best performance for different approaches, since they perform solution space search in different ways. However, it is very encouraging to see that these initial results for the process of gender prediction from iris images show that our approaches can outperform the systems previously described in the literature, where peak accuracy currently reported is typically around 75-80% [TCBF07].

Hence, following these observations, and in order better to exploit the full potential of using the chosen geometrical and texture features, a second experiment is performed to investigate the attainable accuracy and execution time of the proposed gender prediction approaches when using the defined feature vectors with the combination-based classifiers presented in Section 2, with respect to the adopted dataset. The results obtained are shown in Table 5.

Approach	**Results**	**GT**	**Sens**	**Sum**	**Vote**	**Bagging**
1	ACC (%)	70.89	72.46	69.23	59.18	59.72
	ET (sec)	1.83	1.96	0.86	1.31	0.54
2	ACC (%)	72.46	75.96	70.86	70.30	68.00
	ET (sec)	2.05	2.37	1.42	1.47	1.09
3	ACC (%)	87.31	89.74	85.39	85.03	71.24
	ET (sec)	2.84	2.59	1.99	1.84	1.58

Table 5: Accuracy (ACC) and execution time (ET) of combined based classifiers

Thomas et al. [TCBF07], reported around 80% accuracy by using a multiclassifier bagging with the C4.5 approach. In the work presented here, the proposed iris based gender prediction approach 1 uses only five simple geometric features of iris images and can reach accuracies close to 73% within approximately 2 seconds for classification (with the multiagent system using negotiation). Also our approach 3, which adopts both geometric and texture features as in [TCBF07], is able to reach accuracies close to 90% within approximately 3 seconds using also the multiagent system.

4 CONCLUSION

In this paper we have investigated experimentally three approaches to gender prediction from iris images which use a combination of a small number of very simple (and therefore easily and efficiently computable) geometric features (ignoring texture-based information), or which uses texture features alone, or which uses both geometric and texture features. By also adopting an intelligent classification structure, which we have previously found to be especially well suited to more conventional identity prediction from biometric data, we have developed a particularly effective gender prediction approach. Thus, our study has investigated how performance is influenced by the choice of the types of features used, and we have shown how implementing a more flexible and "intelligent" classification technique can support more efficient prediction using smaller number of features.

The performance we have been able to achieve - assigning each tested subject to one of two gender groups (corresponding to male and female categories) in relation to prediction accuracy, even with a small and limited feature set, is seen to be comparable to that reported elsewhere for the prediction of a gender determination problem, but which used a much larger and more diverse feature set. This comparative study based on different feature sets (i.e. geometric, texture and both geometric and texture features) and different classification approaches, provides valuable information to inform and guide the choice of feature and classification approaches in relation to particular application requirements.

This is a very positive outcome in a task domain which has been relatively little investigated to date. Although further work can still be carried out to improve and enhance the levels of achievable performance, our reported results show real promise in relation to the suitability of our basic techniques for application to a number of practical scenarios of importance and considerable current interest.

References

[AF09] M.C.C. Abreu and M.C. Fairhurst. Analysing the Benefits of a Novel Multiagent Approach in a Multimodal Biometrics Identification Task. *IEEE Systems Journal*, 3(4):410–417, December 2009.

[Ary98] A. Arya. An optimal algorithm for approximate nearest neighbors searching fixed dimensions. *Journal of ACM*, 45(6):891–923, 1998.

[BB96] L. Breiman and L. Breiman. Bagging predictors. In *Machine Learning*, pages 123–140, 1996.

[CCP+11] D. Cao, C. Chen, M. Piccirilli, D. Adjeroh, T. Bourlai, and A. Ross. Can facial metrology predict gender? In *2011 International Joint Conference on Biometrics*, IJCB, pages 1–8, October 2011.

[DCAF11] M.C. Da Costa-Abreu and M.C. Fairhurst. Combining multiagent negotiation and an interacting verification process to enhance biometric-based identification. In *The COST 2101 European conference on Biometrics and ID management*, BioID 2011, pages 95–105, Berlin, Heidelberg, 2011. Springer-Verlag.

[EF11] M. Erbilek and M.C. Fairhurst. Evaluating iris segmentation for scenario optimisation. In *The 4th International Conference on Imaging for Crime Detection and Prevention*, ICDP 2011, pages 1–6, November 2011.

[EFDCA13] M. Erbilek, M. Fairhurst, and M. Da Costa-Abreu. Age Prediction from Iris Biometrics. *IET Conference Proceedings*, pages 1.07–1.07(1), January 2013.

[EFDCA14] M. Erbilek, M.C. Fairhurst, and M. Da Costa-Abreu. Analysis of physical ageing effects in iris biometrics. In *IEEE International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2014.

[FAE08] M.S. Fahmy, A.F. Atyia, and R.S. Elfouly. Biometric Fusion Using Enhanced SVM Classification. In *The International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, IIHMSP 2008, pages 1043–1048, August 2008.

[FD12] G. Farinella and J. Dugelay. Demographic classification: Do gender and ethnicity affect each other? In *International Conference on Informatics, Electronics Vision*, ICIEV 2012, pages 383–390, May 2012.

[FE11] M. Fairhurst and M. Erbilek. Analysis of Physical Ageing Effects in Iris Biometrics. *IET Computer Vision*, 5(6):358–366, November 2011. Special issue on Future Trends in Biometric Processing.

[FW94] J. Furnkranz and G. Widmer. Incremental Reduced Error Pruning. In *Proceedings the 11st International Conference on Machine Learning*, ICML 1994, pages 70–77, New Brunswick, NJ, 1994.

[Hay99] S. Haykin. *Neural networks: a comprehensive foundation*, volume 13. Cambridge University Press, New York, NY, USA, 1999.

[KA03] J. Kittler and F. M. Alkoot. Sum versus vote fusion in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):110–115, January 2003.

[Kun04] L.I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.

[LB11] S. Lagree and K.W. Bowyer. Predicting ethnicity and gender from iris texture. In *IEEE International Conference on Technologies for Homeland Security*, HST 2011, pages 440–445, November 2011.

[MAE+07] F. Metze, J. Ajmera, R. Englert, U. Bub, F. Burkhardt, J. Stegmann, C. Muller, R. Huber, B. Andrassy, J.G. Bauer, and B. Littel. Comparison of Four Approaches to Age and Gender Recognition for Telephone Applications. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 4 of *ICASSP 2007*, pages 1089–1092, Honolulu, Hawaii, USA, April 2007.

[Mas03] L. Masek. Recognition of Human Iris Patterns for Biometric Identification. Bachelor of engineering degree of the school of computer science and software engineering, The University of Western Australia, 2003.

[OGFAF+10] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M.R. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J.-L. Alba-Castro, E. Gonzalez-Agulla, E. Otero-Muras, S. Garcia-Salicetti, L. Allano, B. Ly-Van, B. Dorizzi, J. Kittler, T. Bourlai, N. Poh, F. Deravi, M.W.R. Ng, M.C. Fairhurst, J. Hennebert, A. Humm, M. Tistarelli, L. Brodo, J. Richiardi, A. Drygajlo, H. Ganster, F.M. Sukno, S.-K. Pavani, A. Frangi, L. Akarun, and A. Savran. The Multiscenario Multienvironment BioSecure Multimodal Database (BMDB). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:1097–1111, 2010.

[PDVV11] C. Peersman, W. Daelemans, and L. Van-Vaerenbergh. Predicting Age and Gender in Online Social Networks. In *The 3rd International Workshop on Search and Mining User-generated Contents*, SMUC 2011, pages 37–44, New York, NY, USA, 2011. ACM.

[Qui93] J.R. Quinlan. *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

[RB11] K. Ricanek and B. Barbour. What Are Soft Biometrics and How Can They Be Used? *Computer - IDENTITY SCIENCES*, 44(9):106–108, September 2011.

[Spe04] C.E. Spearman. *The Proof and Measurement of Association Between Two Things*, volume 15. Reprinted Int J Epidemiol, 1904.

[TCBF07] V. Thomas, N.V. Chawla, K.W. Bowyer, and P.J. Flynn. Learning to predict gender from iris images. In *The 1st IEEE International Conference on Biometrics: Theory, Applications, and Systems*, BTAS 2007, pages 1–5, September 2007.

[TPB15] J.E. Tapia, C.A. Perez, and K.W. Bowyer. Gender Classification from Iris Images Using Fusion of Uniform Local Binary Patterns. In L. Agapito, M.M. Bronstein, and C. Rother, editors, *Computer Vision - ECCV 2014 Workshops*, volume 8926 of *Lecture Notes in Computer Science*, pages 751–763. Springer International Publishing, 2015.

[WM09] Z.-H. Wang and Z.-C. Mu. Gender classification using selected independent-features based on Genetic Algorithm. In *International Conference on Machine Learning and Cybernetics*, volume 1, pages 394–398, July 2009.