An Extended Local Binary Pattern for Gender Classification

Abbas Roayaei Ardakany, Mircea Nicolescu, Monica Nicolescu
Department of Computer Science and Engineering
University of Nevada, Reno
Reno, USA
Email: {roayaei, mircea, monica}@cse.unr.edu

***Abstract*— This paper addresses the problem of gender recognition by proposing a new feature descriptor to be used in classification. The contribution of this work is an extension to the local binary patterns traditionally used as descriptors. Local binary patterns include information about the relationship between a central pixel value and those of its neighboring pixels in a very compact manner. In the proposed method we incorporate into the descriptor more information from the neighborhood by using four predefined patterns, rather than just one, as in the classic model. We evaluate the performance of our method on the standard FERET database by comparing it to existing methods and show that we can extract more discriminative features and subsequently provide better gender recognition accuracy.**

Keywords-component; Local binary patterns; Gender recognition; Genetic algorithms.

I. INTRODUCTION

The human face is an important biometric feature, and being able to automatically recognize or classify faces is a challenging task in the object recognition research area. Successfully performing this task allows many applications in human computer interaction, psychology, and security [1]. Prior research has shown that it is possible to obtain information on ethnicity, identity, age, gender, and expression from face images [2]. This paper investigates a new approach that helps in gender classification from face images.

Gender plays a significant role in our interactions in society and with computer systems [3]. Gender classification is the binary classification problem of deciding whether a given face image contains a picture of a man or of a woman. Identifying gender from face images has received much attention recently due to its applications in improving search engine retrieval accuracy, demographic data collection, and human–computer interfaces (adjusting the software behavior with respect to the user's gender) [5]. Furthermore, gender classification can be used as a preprocessing step for face recognition since it could halve the number of face candidates, assuming an equal numbers of images from each gender before the recognition. Such preprocessing can sometimes double the speed of face recognition systems [1].

Similar to other image classification tasks, relevant features must be extracted first and then a classifier is applied. With regard to feature extraction, there are several types of methods. A very basic approach is to simply use gray-scale or color pixel vectors as features [6]. Second, many methods use approaches such as Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA) which transform images into a lower dimensional space [7]. The main drawback of this group of methods is that they are sensitive to variations irrelevant to gender such as face orientation and cannot tolerate large changes in this respect [8]. Third, texture information such as wrinkles and complexion has been used [9]. Alexandre *et al.* [5] combined local binary patterns (LBP) with intensity and shape features (histogram of edge directions) in a multi-scale fusion approach, while Ylioinas *et al.* [10] combined LBP with contrast information. Shan [11] used AdaBoost to learn discriminative LBP histogram bins. Finally, it is possible to extract explicit facial features for classification such as through the analysis of facial wrinkles and other shapes [8]. This is performed using a combination of facial feature detection with wavelet transforms [12, 13].

With regard to classifier learning, many different methods have been reported in the literature. [14, 15] used a two-layer neural network where the first layer is responsible for feature extraction and the second layer performs the classification. Radial basis function (RBF) networks and inductive decision trees were employed in [16]. Maximum likelihood classification is used in [17] for face detection, showing the superiority of nonlinear Support Vector Machines (SVM) over traditional linear pattern classifiers. [18] utilizes Gaussian process classifiers which are related to Bayesian kernel classifiers. This method emphasizes the problem of determining kernel parameters in SVMs. [19] tried to improve generalization for gender classification by using fuzzy SVMs. Another method which has been used for gender classification is AdaBoost. [20] applied a threshold Adaboost classifier for gender and ethnicity classification, and a look up table-based AdaBoost classifier is used in [21]. Pixel comparison operators with AdaBoost on low resolution grayscale face images are used in [22]. Other approaches [23, 1] combined face detection and gender classification while carrying out a comparison study for the state-of-the-art gender classification methods. It has also been shown that combining the outputs of different gender classifiers can increase the classification accuracy [1].

The contribution of this paper is focused on improving the discriminative power of feature descriptors used in gender classification, by proposing an extension to the local binary patterns (LBP). The following section provides an overview of descriptors traditionally used in gender classification, and also presents our extended feature descriptor. Section III describes our experimental evaluation and comparison to other methods. Section IV presents an additional stage for feature selection using genetic algorithms, that can further improve accuracy. The paper is concluded in Section V.

II. Feature Descriptors

A. Histogram of Oriented Gradients (HoG)

Local object appearance and shape can be represented well using the local distribution of intensity gradients (or edge directions). HoG features [24] are computed by building histograms of edge gradients in local regions. In the first step, the gradient image in which each pixel is represented by its gradient magnitude and orientation is generated by convolving the input image by an appropriate filter mask (e.g., Sobel). For the experimental comparison presented in this paper, HoG features are extracted from 16×16 local regions as shown in Fig.1. To preserve locality, each region is divided into a 4×4 grid of 4×4 cells. Histograms of gradient orientations are calculated in each of the local cells. Each gradient votes into the corresponding orientation bin by using its magnitude. Since the number of orientations is quantized to 8, each local histogram corresponding to a cell has 8 bins. As a result, the total number of HoG features (bins) for the local region becomes 8×(4×4) = 128. Therefore, an 128-bit feature vector is obtained by concatenating the histograms from each cell.

In order to incorporate the locality in the input image, first each image is divided into a 3×3 grid. A HoG histogram is computed separately in each of the blocks (using the same process as explained above) and subsequently, the concatenation of all those block-vectors yields the final feature vector that represents the face image in terms of HoG features.

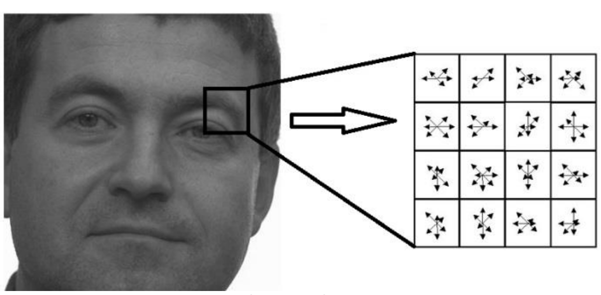

Figure 1. Extracted HoG features for a sample local region

B. Local Derivative Patterns (LDP)

The local derivative pattern (LDP) operator [25] is a high-order texture descriptor originally proposed for face recognition. While the local binary pattern operator (LBP, discussed next) encodes the binary result of the first order derivative among local neighbors, the LDP operator encodes the distribution of the derivative directions which is second order pattern information.

Z_1	Z_2	Z_3
Z_8	Z_0	Z_4
Z_7	Z_6	Z_5

Figure 2. A sample 3×3 neighborhood

Given an image $I(Z)$, the n^{th}-order derivative $I_\alpha^n(Z_i)$ along $\alpha = 0°, 45°, 90°$ and $135°$ is computed by subtracting neighboring pixel values in image I^{n-1} in the respective direction. Let Z_0 be a point in $I(Z)$, and Z_i, $i = 1, \dots, 8$ be the neighboring points around it (Fig. 2). The n^{th}-order directional local derivative pattern is defined as

$$LDP_\alpha^n(Z_0) = \left\{ f\left(I_\alpha^{n-1}(Z_0), I_\alpha^{n-1}(Z_1)\right), f\left(I_\alpha^{n-1}(Z_0), I_\alpha^{n-1}(Z_2)\right), \dots, f\left(I_\alpha^{n-1}(Z_0), I_\alpha^{n-1}(Z_8)\right) \right\} \quad (1)$$

where $f(\cdot,\cdot)$ defined in equation (2) encodes the $(n-1)^{th}$-order gradient transitions into binary patterns.

$$f(x_1, x_2) = \begin{cases} 0 & if\ x_1 \cdot x_2 > 0 \\ 1 & if\ x_1 \cdot x_2 \leq 0 \end{cases} \quad (2)$$

Subsequently, the $(n)^{th}$-order LDP is defined as

$$LDP^n(Z) = \{ LDP_\alpha^n(Z) | \alpha = 0°, 45°, 90°, 135° \} \quad (3)$$

LDP labels each pixel of the image with a 32-bit binary string encoding the local texture pattern around the pixel. In this paper, to preserve locality we divide each image into 3×3 blocks and compute the LDP histograms for each block separately. The final feature vector is obtained by concatenating these histograms in a fixed sequence. As in the case of LBP descriptors, the patterns extracted with the LDP operator are stored in histograms with their decimal values. Each bin of a histogram contains the number of occurrences of a given micro-pattern.

C. Local Binary Patterns (LBP)

Local binary patterns were first presented in [26]. These models can be used for describing textures or shapes in digital images and information in a specified neighborhood of a point can be encoded by using this method. A binary descriptor (LBP code) is obtained by comparing the intensity value of a point (the central point of a window) with its neighboring points. In an 8-bit neighborhood, if the intensity value of a pixel is greater than or equal to that of the central point, then "1" is assigned to the corresponding bit in an 8-bit binary code which is representative of the neighborhood texture with respect to the central point. Otherwise "0" is assigned to the corresponding bit (Fig. 3). This operator is applied on a sub-image, then a histogram is computed from the values produced by the operator. This histogram is used as a feature vector.

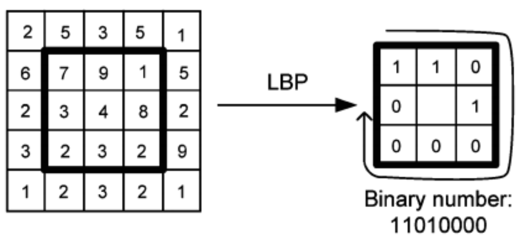

Figure 3. LBP code computation for a sample neighborhood

For the experimental comparison presented in this paper, we use uniform local binary pattern which is an extension to the original operator. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly.

D. The Proposed Descriptor

In traditional local binary patterns (LBP), the binary code that encodes a specific region is simply computed by comparing pixel values in a neighborhood (3×3) with the center pixel value and thresholding the result (Fig. 4).

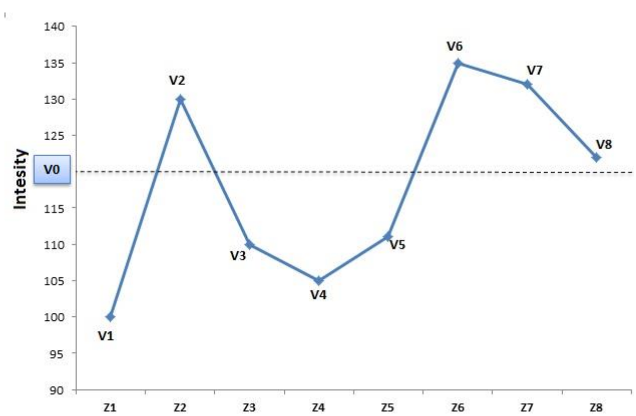

Figure 4. Thresholding neighboring pixel values in LBP

Although this encoding implicitly contains important information about the region and it has been proved to be quite discriminative in a wide variety of applications, it only covers a limited portion of the overall information available in that region. Ideally, for a 3×3 neighborhood of pixels a feature vector must be capable of specifying the relationship between each pair of pixels (in terms of their intensity values $v_i > v_j$). In our proposed extension we try to exploit more information in order to create better discriminating features, by incorporating four patterns of comparison between pixel values in a 3×3 neighborhood. The first comparison pattern is the same one used in the standard local binary pattern approach. Using this comparison pattern we generate a uniform LBP code for each pixel, as shown in Fig. 5.a. Each arrow shows a single comparison between two pixels. Numbers indicate the order of operations needed to create the binary code.

The second local binary code is generated by the second comparison pattern, in which each pixel is compared with its adjacent pixel in a clockwise direction, as shown in Figure 5.b. The third and fourth binary codes are generated using the comparison patterns shown in Fig. 5.c and Fig. 5.d respectively.

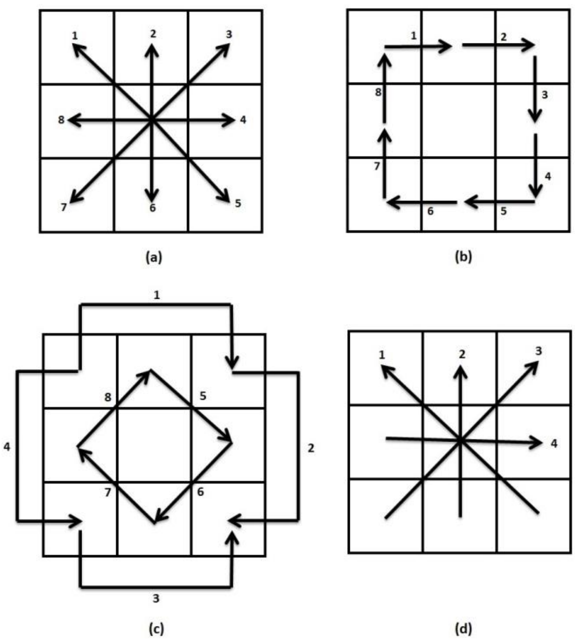

Figure 5. Four comparison patterns used for code generation

Mathematically these four comparison patterns can be expressed as follows:

$$code_1(i,j) = \sum_{k=1}^{8} f(\mathrm{v}_k > \mathrm{v}_0) \times (2^{k-1}) \quad (4)$$

$$code_2(i,j) = \sum_{k=1}^{7} \left(f(\mathrm{v}_{k+1} > \mathrm{v}_k) \times (2^{k-1}) \right) + f(\mathrm{v}_1 > \mathrm{v}_8) \times (2^7)$$

$$\begin{aligned} code_3(i,j) = {} & f(\mathrm{v}_3 > \mathrm{v}_1) \times (2^0) + f(\mathrm{v}_5 > \mathrm{v}_3) \times (2^1) + f(\mathrm{v}_7 > \mathrm{v}_5) \\ & \times (2^2) + f(\mathrm{v}_1 > \mathrm{v}_7) \times (2^3) + f(\mathrm{v}_4 > \mathrm{v}_2) \times (2^4) \\ & + f(\mathrm{v}_6 > \mathrm{v}_4) \times (2^5) + f(\mathrm{v}_8 > \mathrm{v}_6) \times (2^6) \\ & + f(\mathrm{v}_2 > \mathrm{v}_8) \times (2^7) \end{aligned}$$

$$\begin{aligned} code_4(i,j) = {} & f(\mathrm{v}_1 > \mathrm{v}_5) \times (2^0) + f(\mathrm{v}_2 > \mathrm{v}_6) \times (2^1) + f(\mathrm{v}_3 > \mathrm{v}_7) \\ & \times (2^2) + f(\mathrm{v}_4 > \mathrm{v}_8) \times (2^3) \end{aligned}$$

where v_i represents the intensity value of pixel Z_i and function $f(\cdot,\cdot)$ is defined as follows:

$$f(\mathrm{v}_1, \mathrm{v}_2) = \begin{cases} 1, if\ \mathrm{v}_1 \geq \mathrm{v}_2 \\ 0, if\ \mathrm{v}_1 < \mathrm{v}_2 \end{cases} \quad (5)$$

In order to extract features from a face image using the proposed method, each image is first divided into 3×3 grid of blocks with equal size. For each of these blocks, four histograms are built, each being computed on a different code image, as shown in equation (4) and Fig. 6. By concatenating these histograms, we create a feature vector for each block. The final feature vector for the input image is yielded by concatenating all block-vectors.

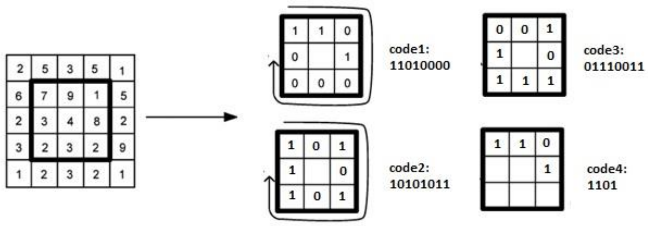

Figure 6. Sample code generation in the proposed method

It should be also mentioned that $code_1$ is converted to uniform LBP code after computation.

III. Experiments and Results

FERET [27] is a large-scale face database which currently contains 11,383 images of 994 individuals with 12 different poses. For the experimental comparison performed in this work we use the FERET subset containing frontal face images, which includes 1,364 images from 994 individuals. Table I provides details about the database used in our experiments and Fig. 7 shows a few sample face images.

Table I. Subset of the FERET database used in this study

Original image size	Total number of images	Number of Individuals	Female	Male
512×768	1364	994	504	860

Figure 7. Some samples from FERET database

In order to evaluate the performance of the proposed method, a set of experiments is conducted on the FERET database. In the first step we extract face regions from input images using Viola-Jones face detection algorithm [28]. Then we extract features from each image separately. To preserve locality, as discussed before, each image is divided into 3×3 grid of blocks and the feature extraction process is performed for each block separately and also once for the whole face region that contains the nine blocks altogether. The final feature vector for the input image is yielded by concatenating all extracted vectors.

In the first experiment we used our proposed descriptor in conjunction with four different classifiers: decision tree, k-nearest neighbors, AdaBoost and Support Vector Machine. The results show that the best performance is obtained for the Support Vector Machine classifier with 96.82% accuracy (Table II).

Table III. Gender classification accuracy for different classifiers trained on features extracted by the proposed descriptor

Classifier	Decision tree	k-NN	AdaBoost	SVM
Classification Accuracy	79.95%	86.55%	84.59%	96.82%

A second experiment was performed to show the comparative performance of our proposed descriptor. We compared it with three well-known, state-of-the-art local pattern descriptors: Histogram of Oriented Gradients (HoG), Local Derivative Patterns (LDP) and Local Binary Patterns (LBP), while using the same SVM classifier that performed best in the previous experiment. The results show that our approach offers better accuracy than the other methods in the context of the gender classification problem (Table IV). Table V also shows the average computation time of the feature extraction process for each image in the database. As shown, our approach performs faster than the HoG and LDP methods and slower than LBP. Indeed, the difference between the computation time of our method and LBP is a trade-off we incur for extracting more discriminative descriptors.

Table VI. Comparison of the proposed descriptor with state-of-the-art local pattern descriptors in terms of accuracy and computation time

Descriptor	Accuracy	Computation time
HOG	93.39%	0.2564 sec
LBP	93.64%	0.0845 sec
LDP	93.15%	0.4900 sec
Proposed	96.82%	0.1580 sec

IV. Feature Selection Using Genetic Algorithms

We have also investigated the use of genetic algorithms (GA) to select features that maximize classification accuracy for our SVM. For this purpose, the number of dimensions is reduced using Principal Component Analysis (PCA), and then a GA is employed to select the best combination of features that can achieve the highest classification accuracy. There are several reasons to justify applying feature selection on the output of PCA in a classification problem. First, there is always some noise in the resulting feature set that can be resolved using an appropriate feature selection algorithm. Second, reduced dimensionality can lead to faster classification. Third, non-linear feature interactions may result in lower principal components having higher discriminatory power in the particular classification task and for the given classification algorithm. We represent each individual chromosome as an array of ones and zeros. Each index position in the array corresponds to a feature. Having "1" at an index implies that the feature is selected, a "0" means that it is not selected. We evaluate the selected features by training a system using the SVM classifier.

In this experiment the algorithm is run for a population size of 100, and a probability $P_i = 0.5$ of being selected for each individual feature. The best result that we have obtained by following this approach is a 98.33% classification accuracy with 79 selected features (out of 166). Fig. 8 shows the classification accuracy plotted against the number of iterations in the genetic algorithm.

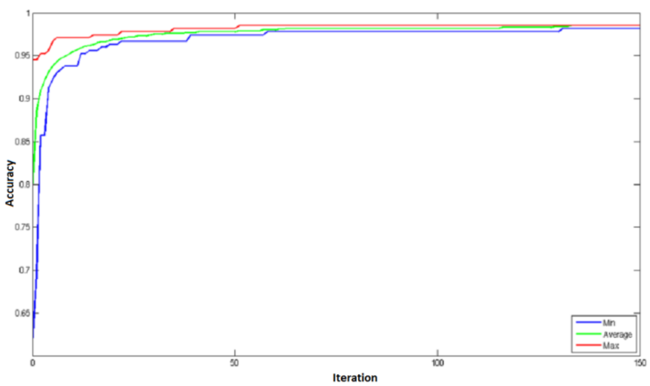

Figure 8. Accuracy after feature selection using genetic algorithms

V. CONCLUSION

In this paper we proposed a novel local binary descriptor capable of extracting more informative and discriminative local features for the purpose of gender classification. We have evaluated our method on the standard FERET database on a gender classification task and our experiments show that the proposed approach offers superior results compared to techniques using state-of-the-art descriptors such as LBP, LDP and HoG. We also show that by performing feature selection through a genetic algorithm, the classification accuracy could be significantly increased while also decreasing computation time during testing. Starting with 166 features, the genetic algorithm reduces the feature set size by approximately half, while increasing accuracy to a 98.33% rate.

REFERENCES

[1] E. Mäkinen and R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters,* vol. 29, pp. 1544-1556, 2008.

[2] B. Wu, *et al.*, "Real-time gender classification," in *3rd International Symposium on Multispectral Image Processing and Pattern Recognition*, 2003, pp. 498-503.

[3] J. Wu, *et al.*, "Semi-supervised feature selection for gender classification," in *Computer Vision – ACCV 2009*. vol. 5995, H. Zha, *et al.*, Eds., ed: Springer Berlin / Heidelberg, 2010, pp. 23-33.

[4] X. Ziyi, *et al.*, "A hybrid approach to gender classification from face images," in *19th International Conference on Pattern Recognition* Shanghai, China, 2008, pp. 1-4.

[5] L. A. Alexandre, "Gender recognition: a multiscale decision fusion approach," *Pattern Recognition Letters,* vol. 31, pp. 1422-1427, 2010.

[6] B. Moghaddam and M.-H. Yang, "Gender Classification with Support Vector Machines," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*. 2000.

[7] K. Balci and V. Atalay, "PCA for Gender Estimation: Which Eigenvectors Contribute?," in *Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02)*. 2002.

[8] H. C. Lian and B. L. Lu, "Multi-view gender classification using local binary patterns and support vector machines," in *Advances in Neural Networks*. vol. 3972, J. Wang, *et al.*, Eds., ed: Springer Berlin / Heidelberg, 2006, pp. 202-209.

[9] R. Iga, *et al.*, "A gender and age estimation system from face images," in *SICE 2003 Annual Conference*, 2003, pp. 756-761.

[10] J. Ylioinas, *et al.*, "Combining contrast information and local binary patterns for gender classification," in *Image Analysis*, ed: Springer, 2011, pp. 676-686.

[11] C. Shan, "Learning local binary patterns for gender classification on real-world face images," *Pattern Recognition Letters,* vol. 33, pp. 431-437, 2012.

[12] S. Hosoi, *et al.*, "Ethnicity estimation with facial images," in *Proceedings of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 195-200.

[13] H.-C. Lian, *et al.*, "Gender recognition using a min-max modular support vector machine," in *Advances in Natural Computation*, ed: Springer, 2005, pp. 438-441.

[14] M. K. Fleming and G. W. Cottrell, "Categorization of faces using unsupervised feature extraction," in *International Joint Conference on Neural Networks* 1990, pp. 65-70.

[15] B. GOLOMB, "SEXNET: A neural network identifies sex from human faces," *Advances in Neural Information Processing System,* vol. 3, pp. 572-577, 1991.

[16] S. Gutta, *et al.*, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Transactions on Neural Networks,* vol. 11, pp. 948-960, 2000.

[17] B. Moghaddam and M.-H. Yang, "Learning gender with support faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 707-711, 2002.

[18] A. M. BurtonU, *et al.*, "What's the difference between men and women? Evidence from facial measurement," *Perception,* vol. 22, pp. 153-176, 1993.

[19] X. Leng and Y. Wang, "Improving generalization for gender classification," in *Proceedings of 15th IEEE International Conference on Image Processing*, 2008, pp. 1656-1659.

[20] G. Shakhnarovich, *et al.*, "A unified learning framework for real time face detection and classification," in *Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 14-21.

[21] B. Wu, *et al.*, "LUT-based adaboost for gender classification," in *Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication*, Berlin, Heidelberg 2003, pp. 1062-1062.

[22] S. Baluja and H. Rowley, "Boosting sex identification performance," *International Journal of Computer Vision,* vol. 71, pp. 111-119, 2007.

[23] 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, pp. 541-547, 2008.

[24] Dalal, N.; Triggs, B., "Histograms of oriented gradients for human detection," *Computer Vision and Pattern Recognition, 2005. CVPR*

2005. IEEE Computer Society Conference on , vol.1, no., pp.886,893 vol. 1, 25-25 June 2005.

[25] B. Zhang, *et al*., "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor," *IEEE Transactions on Image Processing,* vol. 19, pp. 533-544, 2010.

[26] E. Mäkinen and R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters,* vol. 29, pp. 1544-1556, 2008.

[27] P. J. Phillips, *et al*., "The FERET database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing,* vol. 16, pp. 295-306, 1998.

[28] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2001, pp. 511-518.