

PSO-GA Based Optimized Feature Selection Using Facial and Clothing Information for Gender Classification

M. Nazir ^{*}¹, A. Majid-Mirza^{1, 2}, S. Ali-Khan³

¹ Department of Computer Science,
National University of Computer & Emerging sciences,
Islamabad, Pakistan.

^{*}Muhammad.nazir@nu.edu.pk

² College of Computer & Information Sciences,
King Saud University,
Riyadh, Kingdom of Saudi Arabia.

ammirza@ksu.edu.sa

³ Department of Computer Science,
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology,
Islamabad, Pakistan.

Sajidalibn@gmail.com

ABSTRACT

Gender classification is a fundamental face analysis task. In previous studies, the focus of most researchers has been on face images acquired under controlled conditions. Real-world face images contain different illumination effects and variations in facial expressions and poses, all together make gender classification a more challenging task. In this paper, we propose an efficient gender classification technique for real-world face images (Labeled faces in the Wild). In this work, we extracted facial local features using local binary pattern (LBP) and then, we fuse these features with clothing features, which enhance the classification accuracy rate remarkably. In the following step, particle swarm optimization (PSO) and genetic algorithms (GA) are combined to select the most important features' set which more clearly represent the gender and thus, the data size dimension is reduced. Optimized features are then passed to support vector machine (SVM) and thus, classification accuracy rate of 98.3% is obtained. Experiments are performed on real-world face image database.

Keywords: gender classification, real-world face images, particle swarm optimization, genetic algorithm, local binary pattern, features fusion.

1. Introduction

Gender classification is gaining popularity in the current research field. There is a wide range of applications that involves the use of this mechanism; intelligent user interfaces, demographics and customer-oriented publicizing are some of them.

With the advancement in human-computer interaction field, there is a growing demand for secure protection techniques such as face recognition [1]; finger prints recognition [2]; gesture recognition, and gender classification. Generally speaking, the features of an object are of two types; appearance based (global) and geometric based. The former extract the features from the full facial expression, which is considered an easy task. The latter method is used to get the characteristics of

every gesture of the face; giving the benefit of face rotation to have a look from various angles.

In the 1990s, Golomb et al. [3], trained two-layer neural network called SEX-NET and attained 91.9 % classification accuracy rate. They used a total of 90 frontal face images. In 1995, Brunelli and Poggio et al. [4], extracted geometric features and utilized them to train the networks. They claimed a 79% classification accuracy rate. Sun et al. (2002) [5], asserted that genetic algorithm (GA) works considerably well for vital feature's selection task. They used principal component analysis (PCA) to create features' vector and GA to select the vital features. They attained a 95.3% accuracy rate after training support vector machine classifier by using

those vital features. In 2004, A. Jain et al. [6], extracted facial features using independent component analysis (ICA), and then categorized gender using liner discriminant analysis (LDA). They provided results using normalized FERET database and attained a 99.3 % classification accuracy rate. In 2006, Sun et al., [7] used local binary pattern (LBP) to create features' vectors in the input of AdaBoost and attained a 95.75% classification accuracy rate. In 2007, Baluja and H Rowley [8] attained a 93% classification accuracy rate. They used pixel comparison operators with Adaboost classifier. In 2010, Nazir et al., [9] utilized discrete cosine change (DCT) to extract the vital facial features and utilized K-nearest neighbor classifier to categorize gender. In 2011, S. A. Khan et al., [10] used discrete wavelet change (DWT) to extract facial features. They claimed that classifier performance could be improved by assembling different classifiers using weight majority technique. They performed experiments at Stanford University Medical Students (SUMS) face database and received a 95.63% classification accuracy rate.

A major problem with all these proposed techniques is that these take into account only frontal face images (e.g., SUMS, FERET). The images in these databases consist of pictures that have a clean background, are occlusions free, furnish merely frontal faces, encompass limited facial expressions and have consistent lighting effects. But, real-time images are normally captured in unconstrained settings and conditions. A real-time picture normally contains momentous emergence variations like illumination change, poor picture quality, makeup or occlusions and disparate facial expressions. Gender recognition in an unconstrained environment is a very challenging task compared to images captured in constrained settings like FERET, and SUMS face databases. This problem has been highlighted by few researchers. Shakharovitch et al. (2002) [11] gathered 3500 face pictures from the web. They used Haar-like features and obtained a 79% accuracy rate with Adaboost and 75.5% with SVM. W. Gao and H. Ai (2009) [12], claimed a 95.5% classification accuracy rate by performing experiments on 10,100 real-time images. They also used Haar-like features but used probabilistic boosting tree. We supposed they used these results as benchmarks as their databases are not publicly available. N. Kumar et al. (2009) [13] performed experiments on real-time images. They received a

81.22% classification accuracy rate by training countless binary "attributes" classifiers. Their emphasis was put more on face verification and not on gender classification.

The major contributions of our work are as follows:

- Proposed technique is robust to illumination and expression changes due to local texture feature selection.
- Combination and optimization of facial and clothing features provides improved accuracy while using less number of features.
- Technique presented in this manuscript gives superior results on real-world face image dataset when compare to the existing methods.

2. Proposed methodology

The proposed method consists of the following steps as depicted in Figure 1.

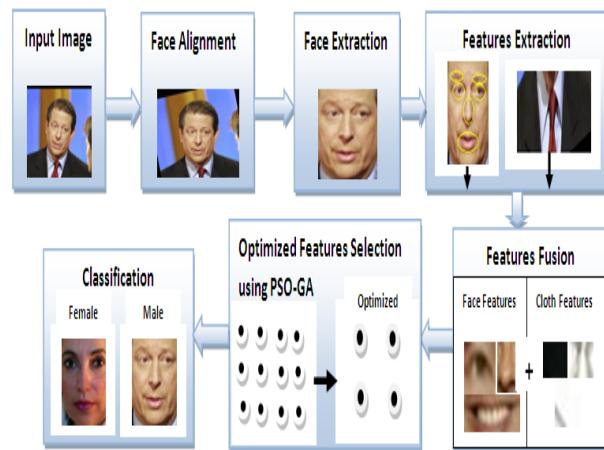


Figure 1. Proposed system architecture

Pre-processing: First, we align the images by using commercial software [14] and then perform histogram equalization to normalize the face.

Face extraction: Facial portion is extracted and unwanted area is removed using spatial co-the ordinate system.

Features extraction: Facial local features like nose, mouth and eyes are extracted using the local

binary pattern (LBP). A Color histogram is used to save the color information while LBP carries the textures' information of clothes.

Features fusion: In this step, facial and clothes features are fused by a concatenation method.

Optimized features selection: Only most important feature sets are selected using hybrid PSO-GA algorithm.

Classification: In this step, support vector machine (SVM) is used to classify gender.

2.1 Facial portion extraction

A solitary tiny picture normally contains thousands of pixels. This huge data set seriously affects the computational time and makes the system slow that is why pixel-based approaches are very expensive. Thus, to remove the unwanted area and to extract merely facial portion we use spatial co-ordinate system. By using hit and trial method we set the X and Y coordinates values. Figure 2 depicts a sample of extracted images.



Figure 2. Samples of extracted faces

2.2 Feature extraction

To get improved results, we extract both facial and clothing features of gender. The features' extraction mechanism is explained below.

Facial local features extraction: Face portion plays an important role in representing the gender. Features get from the whole face portion are known as global features. Global features are affected if a face contains some occlusion or illumination change. The sub-face individual features like nose, eyebrows and mouth, etc., are known as local features. Local features are robust to variations like illumination, face occlusions and alignment. In this work, we extract the features of nose, mouth, chin, eyes and forehead. First, facial components are located using active shape model (ASM) [15], and then facial components are

cropped from the image. In the next step, local binary pattern (LBP) is applied to extract the features from facial components.

Lian and Lu [16] used LBP for gender classification. The working algorithm of LBP is simple but efficient to carry texture information and also stable under illumination change. The operator of LBP labels the neighbor pixels values by thresholding with center value.

$$S(f_p - f_c) = \begin{cases} 1, & f_p - f_c \geq 0 \\ 0, & f_p - f_c < 0 \end{cases} \quad (1)$$

Here represent the center value, and represents the neighbors' pixels. Equation 2, represents the LBP operator at the center.

$$LBP(f_c) = \sum_{p=0}^7 S(f_p - f_c) 2^p \quad (2)$$

In LBP, the center pixel is selected, and the neighborhood pixels are either, modified to 0 if the gray levels are smaller than center, or to 1 if the gray levels are larger than a center. The center pixel is next substituted by the binary code of its neighborhood like 00011011, as it is shown in Figure 3.

1	4	2
8	5	7
3	3	6

0	0	0
1		1
0	0	1

Figure 3. Binary code conversion

Cloth features extraction: Hauling out various characteristics from clothes is very exigent due to diverse classification in garment's styles, consistency, outlines and insignia comparatively from facial appearance of human being.

Even matching garments represent different characteristics with different human bodies, although cloth attribute is generally superfluous because of its bulky deviation in today's research work. Therefore, we consider identical gender cloths with the same features regardless of dissimilarity as shown in Figure 4.

Based on these arguments, we apply color histogram and local binary pattern algorithm while extracting characteristics of the clothes. Both schemes are resistant to illumination and rotation transformation. Local binary pattern carried texture information and color histogram contains color information of clothes. The clothing texture images are shown in Figure 4.



Figure 4. Texture information of clothes

2.3 Features normalization and fusion

The feature vector of the face and clothes are represented by and Dimensions of both feature vectors are different, that is why we normalize both feature vectors between

Diversity is present between the facial and clothing feature vectors, which affect the fusion process of both feature vectors. For example, if the values of are in between and values of are in between {0 and 1}, then the distance is more sensitive to as compare to. In order to solve this problem, we normalized both feature's vectors.

Let n and n' represents the features' values before and after normalization. We compute the normalized features by using Min-Max technique.

$$n' = \frac{p - \min(f_n)}{\max(f_n) - \min(f_n)} \quad (3)$$

In Equation 3, the function generated at and the and represents the minimum and maximum values of all possible 'n' values.

The advantages of feature level fusion are obvious. By combining and fusing different features, we not only are able to preserve the discriminative information about the face but we also can eliminate redundant features. To take advantage of different features' extraction (techniques for better classification) and to make a proposed technique more robust to variations in expressions, illumination and pose changes, we have fused face and clothing features by the concatenation method.

Let the normalized features vectors of ' then; ' represent the concatenation features vector. Procedure for feature fusion is provided below.

2.4 Optimized features selection

In machine learning, the features' selection procedure is a process of selecting the optimal facial features and discarding redundant features that increase the classification accuracy rate. To select the optimal features, a combination of particle swarm optimization and genetic algorithm is used.

Genetic algorithm: In 1970, Holland provided the idea of a genetic algorithm. GAs is a stochastic search algorithms modeled on the procedure of natural selection; that underlines biological evolution. GAs has successfully been applied in many search, optimization, and machine learning problems [17]. GAs is used to simulate procedures in natural systems that are vital for evolution. GA searched for the best solution in search space intelligently and works in an iterative order, such as that new generation is created from the old one. The strings in GA's are represented as binary codes. In GA's, a fitness function computes the fitness of every string. The standard operators used in GA's are mutation, crossover and selection.

Particle swarm optimization: is a population-based stochastic optimization technique that was developed by Kennedy and Eberhart in 1995 [18]. A particle in search space can be considered as "an individual bird of a flock." To find the best solution PSO uses local and global information. The fittest solution is found by utilizing the fitness function and velocities of the particles. In PSO, the particle position gets updated by using local and global positions of every particle around its neighbor. The particles move across the problem space searching for the optimal particles (features). The procedure is then repeated for a fixed number of periods or until a minimum, error is attained [18].

K-nearest neighbor: The K-nearest acquaintance (K-NN) method was introduced by Fix and Hodges in 1951, and is one of the most accepted nonparametric methods [19]. KNN classifies the new data based on the knowledge of training data. Random procedures are used to solve the lazy problems. The working procedure of KNN is that it calculates the distance from the

query instance and the training data. In our work, we used leave-one-out cross validation strategy of 1-NN and the distance between neighbors is calculated using Euclidean distance. 1-NN strategy is easy to implement as it is independent of user-specified parameters. The particles are coded as binary strings, i.e., $S = F_1 F_2 \dots F_n$, $n = 1, 2, \dots, m$; where bit value '1' represents selected features, and '0' represent unselected features. The initial population is generated randomly, and leave-one-out cross validation strategy is used to evaluate the fitness value. To select the best solution from population, mutation and crossover operators are applied without modification. We used rand-based roulette-wheel scheme. In this work, the two cutting points are chosen using the 2-point crossover operator. First, a check is performed for mutation, and if mutation exists, then the offspring are mutated, and the code conversion is performed from 1 to 0 and similarly from 0 to 1. Next, if the mutated chromosomes resulted worthier than parents, then the replacement is performed with the worst parent and in another case, the replacement is performed with the lower quality one. The performance of GA is enhanced by applying the PSO for each new generation. In this work, the adaptive values are used to renew the particle. $pbest$ and $gbest$ are the best adaptive values in $pbest$ group. The particle position and speed are tracked after the values of $pbest$ and $gbest$ are obtained. The below equations are used to perform the updating process.

$$v_{pd}^{new} = w x_{pd}^{new} + c_1 x_{rand 1} (p_{best, pd}^{old} - x) + c_2 x_{rand 2} (p_{best, pd}^{old} - x) \quad (4)$$

$$s(V_{pd}^{new}) = \frac{1}{1 + e^{-v_{pd}^{new}}} \quad (5)$$

$$\text{if}(rand < S(V_{pd}^{new})) \text{then} x_{pd}^{new} = 1; \text{else} x_{pd}^{new} = 0 \quad (6)$$

Before updating the particle, the worth of $gbest$ is checked to prevent particle from getting trapped in local optimum. The $gbest$ value is updated if it is found that its value remains a number of times.

Equation 4 is used to calculate the features after the update. Here V represents the velocity.

The features are selected and represented by "1" if value of V is greater than random produced disorder {0.0~1.0}, and the features are unselected and represent by "0" if value of V is less than the random produced disorder {0.0~1.0}. The GA was configured to encompass 20 chromosomes and was run for 100 generations in each trial. The crossover rate and mutation rate was 1.0 and 0.1 respectively. The number of particles used is 20. The two factors $rand1$ and $rand2$ are random numbers between (0, 1), whereas $C1$ and $C2$ are acceleration (learning) factors, with $C1=C2=2$. The inertia weight 'w' was set as 0.9. The maximum number of iterations utilized in our PSO was 100.

2.5 Classification using support vector machine

The gender classification problem can be thought of as a two-class problem and the goal of this problem is to separate the two classes by mean of a function. SVM is a useful technique for data classification and is easier to use than Neural Networks. SVM takes data (features) as input and predicts based on training data set to which class these features belong. The goal of SVM is to find the optimal hyper plan such that the error rate is minimized for an unseen test sample. According to the structural risk minimization principals and VC dimension minimization principle [20] a linear SVM uses a systematic approach to find a linear function with the lowest capacity.

The SVM classifier correctly separates the training data of the labeled set of M training samples where ϵ and ω are the associated label, i.e., ($\epsilon \in \{-1, 1\}$). The hyper plan is defined as:

$$f(x) = \sum_{i=1}^M y_i \omega_i k(x, x_i) + b \quad (7)$$

3. Results and discussion

We used MATLAB 2009a environment for our experiments. Label faces in the wild, (LFW) [21], face database was used for experiments. This database contains 13,233 images collected from a web. To remove unwanted area, face portion was extracted using spatial co-ordinate system and then, as shown in Figure 5, histogram equalization was applied to normalize the face image. We

selected 400 face images, 200 male and 200 female for our experimental setting. As shown in Figure 6, all the faces were aligned using commercial aligned software (Wolf et al 2009) [14].



Figure 5. Image before (a) and after (b) normalization



Figure 6. Original (a) and align (b) image

We did avoid those images for which it was difficult to establish the ground truth as shown in Figure 7.



Figure 7. Images avoided to selected

The 5-fold cross validation is used in all experiments. First, the image is resized to 32x32 and then it is divided into blocks of size 8x8. For geometric-based features extraction, first the five facial points are located using active shape model and then local binary pattern is used to extract the local features from these points. In the next step, we extract the clothing features by applying the LBP, which carry texture information and color histogram, which carry the color information of cloths. After facial and cloths features extraction we

pass these features to support vector machine (SVM) to find the classification accuracy rate.

Figure 8, shows the classification accuracy of facial and clothing features; extracted using different features set (i.e., 30, 40, 50 and 60). Figure 8, also shows that in case of clothing features, FS-30 produce the higher accuracy rate (i.e., 75%) as compared to other features set and FS-50 produce 74% in case of clothing features.

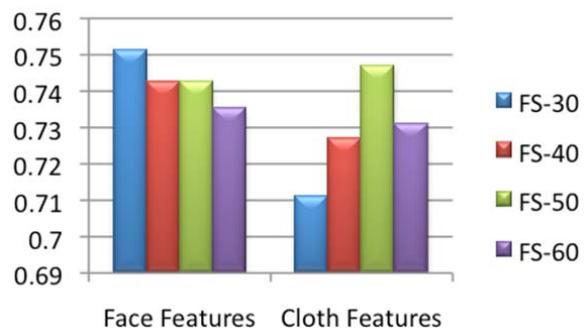


Figure 8. Facial and clothing extracted feature accuracy

After the extraction of facial and clothing features, fusion of features is performed using the concatenation method and features vectors of size 60, 80, 100 and 120 are generated. Figure 9, shows the classification accuracy rate after features fusion. We examine that classification accuracy rate of face and cloth features are low separately but when features of both types are fused then classification accuracy rate increases dramatically. The drawback is that the data dimension size also increases due to fusion process; hence, to reduce the data dimension size, optimized features are selected in the next step.

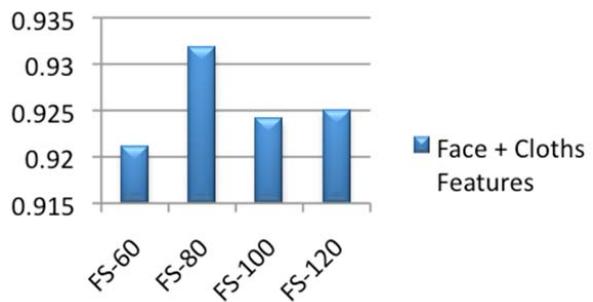


Figure 9. Classification accuracy of face and cloths fused features

In the next step GA and PSO were combined using leave-one-out cross validation strategy of 1-NN and distance between neighbors was calculated using Euclidean distance. The features set (i.e., 80), were passed to GA-PSO, which eliminates the redundant features and GA-PSO get back the result as optimized features (i.e., 40). After optimization step the support vector machine was trained and tested by optimized features. 1:3 and 3:1 testing to training ratio was used for SVM. Figure 10 shows the accuracy rate after features optimization. The results depict that the hybrid GA-PSO was very useful for features optimization. By using this technique, the enhancement in the classification accuracy rate (i.e., 98%) is noteworthy and it also reduced the data dimension size (i.e., from 80 features set to 40 features) that is shown in Figure 10.

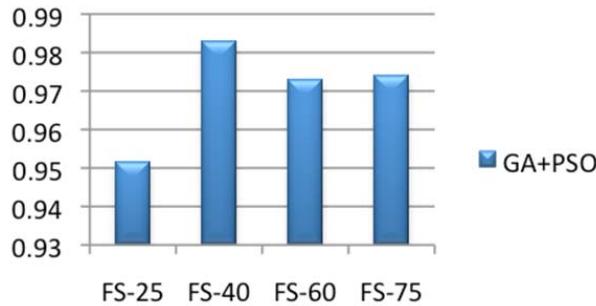


Figure 10. Optimized features classification accuracy rate

Table 1, presents a comparison of our proposed technique with other gender classification techniques. In our proposed technique we overcome the problems like high dimensions (i.e., by utilizing the minimum number of features), variation in pose or occlusions (i.e., by combining facial and clothing features). We used real life images that have large amount of variations in facial expression and poses. The database that we have used is also publically available which helps to benchmark for future. In Table 1, the proposed technique using real life face images is also publically available as compared to other techniques face databases. The high recognition rate of 98% is obtained.

Methods	Data Dimensions	Real Life	Publically Available	Recognition Rate
Proposed	30	Yes	Yes	98.3%
Sun et al [5]	150	No	No	95.3%
Jain and Hung [6]	200	No	Yes	96%
Baluja et al [8]	2409	No	Yes	94.3%
Nazir et al [9]	256	No	Yes	99.3%
Shakhnarovich et al [11]	35,00	Yes	No	79%
Gao et al [12]	10,100	Yes	No	95.5%
Sajid et al [10]	20	No	Yes	95.63%
Shan [22]	2891	Yes	Yes	95%

Table 1. Proposed method comparison with state-of-the-art gender classification techniques

4. Conclusion and future work

In previous studies most of the researchers worked on frontal face images. In this work both face and clothing-based features were extracted and fused which made the systems supportive of variations in expressions and poses. Furthermore, we also concentrated on reducing the data dimensions and struggled to produce optimal features' set that extra precisely embodies a gender face. In this paper, features were optimized by employing hybrid GA-PSO algorithm. Hybrid optimization results in improved 98.3% classification accuracy rate using label in wild face images (LFW) database. Furthermore, we aim to discover swarm-based optimization algorithms (SOAs) like Ant dominion optimization and make our arrangement more precise and stable.

References

[1] L. Ramírez-Valdez and R. Hasimoto-Beltran, "3D facial expression synthesis and its application to face recognition systems," *Journal of Applied Research and Technology*, vol. 7, no. 3, pp. 323-339, 2009.

[2] G. Aguilar-Torres et al., "Fingerprint recognition using local features and Hu moments," *Journal of Applied Research and Technology*, vol. 10, no. 5, pp. 745-754, 2012.

[3] B. Golomb and D. Lawrence, "Sexnet: A neural network identifies sex from human faces," *Proceeding of the Advance in neural information processing systems*, 2009, pp. 323-339.

[4] R. Brunelli and T. Poggio, "HyperBF networks for gender classification," *Proceeding of the DARPA image understanding workshop*, 1992, pp. 311-314.

[5] Z. Sun et al., "Neural network based gender classification using genetic algorithm for Eigen-feature selection," *Proceeding of the EEE International conference on neural networks*, 2002, pp. 2433-2438.

[6] A. Jain et al., "Gender classification using frontal facial images," *Proceeding of the IEEE international conference on pattern Recognition*, 2008, pp. 1-4.

[7] N. Sun et al., "Gender classification based on boosting local binary pattern," *Proceedings of the Third international conference on Advances in Neural Networks*, 2006, pp. 194-201.

[8] S. Baluja and H. Rowley, "Boosting sex identification performance," *International Journal of computer vision*, vol. 71, no. 1, pp. 572-577, 1990.

[9] M. Nazir et al., "Feature selection for efficient gender classification," *Proceeding of the 11th WSEAS International conference*, 2010, pp. 70-75.

[10] S.A Khan et al., "Efficient gender classification methodology using DWT and PCA," *Proceeding of the 14th IEEE international Multi-topic conference*, 2011, pp. 155-158.

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

[12] W. Gao and H. Ai, "Face gender classification on consumer images in a multiethnic environment," *Proceedings of the International Conference on Biometrics*, 2009, pp. 169-178.

[13] N. Kumar et al., "A search engine for large collections of images with faces," *Proceeding of the European Conf. on Computer Vision*, 2009, pp. 340-353.

[14] L. Wolf et al., "Similarity scores based on background samples," *Proceeding of the Asian Conf. on Computer Vision*, 2010, pp. 88-97.

[15] T. Cootes et al., "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.

[16] H. Lian and B. Lu, "Multi-view gender classification using multi-resolution local binary patterns and support vector machines," *Proceedings of the Third international conference on Advances in Neural Networks*, 2006, pp. 202-209.

[17] J. Holland, "Adaptation in Nature and Artificial Systems", In MIT Press, 1992.

[18] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," *Proceeding of the IEEE International conference on neural networks*, 1995, pp. 1942-1948.

[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," *Proc. IEEE Trans. Information Theory*, vol. 11, pp. 21-27, 1967.

[20] V. Vapnik, "The nature of statistical learning theory," *Proceeding of the Springer conference*, 1995.

[21] G. Huang, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments", *Tech. Rep. 07-49, University of Massachusetts, Amherst*, 2007.

[22] C. Shan, "Learning local binary pattern for gender classification on real-world face images," *Pattern Recognition Letter*, vol. 33, no. 4, pp. 431-437, 2011.