GENDER PREDICTION FROM FACIAL IMAGES USING LOCAL BINARY PATTERNS AND HISTOGRAMS OF ORIENTED GRADIENTS TRANSFORMATIONS

Tariq KHALIFA (ORCID: 0000-0003-4246-9846)¹
Gökhan ŞENGÜL (ORCID: 0000-0003-2273-4411)²*

¹Computer Science Department, Engineering Faculty, Al Jabal Al Gharbi University, Gharian, Libya
²Computer Engineering Department, Faculty of Engineering, Atılım University, Ankara, Turkey

ABSTRACT

Gender prediction from facial images can be used in a large number of applications including human-computer interaction, customer information measurement, access control, etc. Furthermore, it can substantially effect on many fields, such as security systems, biometric authentication, medical imaging systems, demographic studies, content based searching, and surveillance system. In this study, we proposed to use Local Binary Patterns (LBP) and Histograms of Oriented Gradients (HOG) as the feature extractor and k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) as the classifier in order to predict the gender of the people from facial images. We tested the proposed method in FERET and UTD databases. We used leave-one-out approach as the cross validation technique. The results are promising.

Keywords: Facial images, gender prediction, Local Binary Patterns, Histograms of Oriented Gradients

1. INTRODUCTION

Human beings can detect and recognize the gender of any person easily by only seeing to his/her face. On the other hand, it is a challenging task for computers, which modern world going to depend on it in everything. Predicting gender from facial images is an attractive research topic and important task for computer but still
there is a deficiency between system requirements and current performances. This deficiency extends due to the difference in illumination conditions, expressions, poses etc. Gender prediction from facial images in turn can boost the performance of a large number of applications including human-computer interaction, data collection related to the customers in shopping malls and stores (such as gender of the costumers visiting the shopping malls and their distribution according to time periods and days of the week), and access control. Furthermore, it can substantially effect on many fields, such as security systems, biometric authentication, medical imaging systems, demographic studies, content based searching, and surveillance systems [1].

Predicting the gender information from the face image has gained a lot of attention in last decades. Sun et al. [2] investigated the challenge of gender recognition from frontal facial images using Genetic Algorithm (GA). They used Principal Component Analysis (PCA) to extract features, and then appointed GA to select a subset of these features. Moreover, they used four different classifiers that are Bayesian, Neural Network, Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). The comparison results show that the superiority of PCA + SVM with 95.3% gender recognition accuracy. Mohgaddam and Yang [3] performed SVM classifier with Radial Basis Function (RBF) kernel to increase the gender classification accuracy. However, by applying the methods on 1755 facial images from FERET database, they got accuracy of 96.62% in SVM + Gaussian RBF kernel, and 95.12% in SVM + Cubic Polynomial kernel. Jian and Huang [4] tried to classify the gender by employing Independent Component Analysis (ICA) as a feature extraction method, and LDA as a classifier. They evaluated these classifiers on 500 frontal facial images taken from FERET database [5]. Using ICA with LDA leads to obtain an accuracy of 99.3%. Costen et al. [6] developed a sparse classifier for gender recognition. They used PCA to extract the facial features of 300 images taken from Japanese face images database. Moreover, they classified these data by using three classifiers: LDA, EBPC, and SVM. 10-folds cross-validation basis executed on the classification process, and then accuracy of 79.75%, 94.25%, 94.42% in PCA+LDA, PCA+EBPC, PCA+SVM, has achieved respectively. Sun et al. [7] used Local Binary Patterns (LBP) to extract the features of images from FERET database for gender recognition. In addition, they employed both Self Organizing Map (SOM) and threshold Adaboost classifiers in their experiments and they claimed 95.75% recognition rate. Similarly, Lian and Lu [8] applied same LBP features with SVM classifier and reported 96.75% gender recognition rate. Makinen and Raisamo [9] had done several experiments by combining gender prediction methods with real time face detection. In details, they compared the results of two classifiers that are Multi-Layer Neural Network and SVM on FERET and WWW databases. In addition, they used four feature extractors that are Look-Up Tables (LUT) Adaboost, Threshold Adaboost, Mean Adaboost, and LBP. As a result, the best classification accuracy they achieved on FERET database was 91.11% by applying LUT Adaboost and Neural Network, and 78.25% by applying LBP with SVM on WWW database. Yuchun and Zhan [10] have employed PCA to decrease the high density of LBP feature to classify gender from facial images. They worked on FERET database and got success ratio of 92.16%. Scalzo et al. [11] made experiments to recognize the gender by comparing the results of using three feature extraction methods that are PCA, Gabor, and Feature Fusion Hierarchies (FFH), and three classifiers that are LDA, Nearest Neighbours (NN), and SVM. They evaluated the performance of these classifiers on 400 images from FERET database [5] and they got best accuracy of 96.2% at FFH+SVM. Zafeiriou et al. [12] applied SVM and Kernel Fisher Discriminant Analysis (KFDA) with an RBF kernel on 2360 images (1256 male and 1104 female) taken from XM2VTS database [13]. They claimed 97.20% gender classification rate. Lu and Shi [14] divided the face into three regions before extracting there features. They used two feature extractions that are two-dimensional PCA (2DPCA) and PCA. In addition, SVM is used for classification. By applying this method on 800 grey scale facial images taken from FERET database, 2DPCA+SVM gave best result (95.33% gender recognition rate).

Singh et al. [15] were able to reach 95.56% success gender prediction by using "HOG+SVM", and accuracy of 89.43% by using "LBP+SVM" by applying given methods on 300 images selected from Indian face Database. Similarly, Liu et al. [16] have reached to accuracy of 94.38% with "HOG+SVM" and 91.43% with "LBP+SVM" by applying these methods on LFW face database, and then they combined "HOG & LBP" and used it with SVM to reach to gender prediction accuracy of 94.88%. Lin and Zhao [17] studied a gender prediction scheme depending on color information. They developed an eye detection algorithm by combining SVM with some color features. By combine SVM classifier with these features, they achieved an accuracy of gender prediction scheme with 80.7%. Shan [18] investigated gender recognition using LBP feature extractor with both Adaboost and SVM classifiers. The result of this method was 94.81% success recognition rate on Labelled Faces in the Wild (LFW) database. Ihsan Ullah et al. [19] investigated Weber’s Local Descriptor (WLD) for gender recognition. Their technique depends on dividing an image into a number of blocks, compute WLD histogram for each block and concatenate them to form a Spatial WLD descriptor (SWLD). They got overall accuracy of (99.08%) with (Chi-Square) technique. Chen Wang et al. [20] proposed a method called Local Circular Patterns (LCP) for gender classification, which is an improvement of traditional LBP. Their experiments carried out on the FERET database. Then both LBP and LCP features extracted for different scales. SVM with a linear kernel
used for the classification. As a result, the gender classification accuracy was up to 95.36%, clearly highlighting the effectiveness of LCP method. Min Li et al. [21] have combined three methods that are Histograms of Oriented Gradients (HOG), LBP, and Gabor descriptors together to represent the head-shoulder part of human body for gender recognition. Consequently, an accuracy of 88.55% achieved. Emon et al. [22] have described the gender detection from facial images in a novel approach. The system can detect facial area of image by Region of Interest (ROI), and then detect gender by using Discrete Cosine Transformation (DCT). Moreover, they used LBP to improve ROI matching accuracy. At the beginning, they got success ratio of 70%, but after applying Histogram Equalization, the ratio increased to 78.91%.

Iga et al. [23] have used graph matching with Gabor Wavelet Transformation (GWT) [24] to extract images information such as, skin color, moustache, hair, etc. All this information classified by applying SVM on 300 facial images (150 males and 150 females) taken from Softtopia Japan HOIP database. Consequently, they achieved to 97.3% gender classification rate. Yang and Ai [25] employed LBP feature extractor to know the Chi square measure between a reference histogram and the extracted LBPH. Moreover, real Adaboost algorithm is used as a strong classifier, which learns a sequence of best local features. By using 3540 facial images from FERET database, accuracy of 93.30% gender prediction achieved. Guo et al. [26] studied the differences between "no crossings" and "crossing", when they found that crossing race and gender could increase the estimation error rate. Notably, they implement their system on large database called MORPH-II [27], which includes more than 55,000 facial images. For experiments, they used the Biologically-Inspired Features (BIF) classifier with four feature extraction techniques that are PCA, Marginal Fisher Analysis (MFA) [28], Orthogonal Locality Preserving Projections (OLPP) [29], and Locality Sensitive Discriminant Analysis (LSDA) [30]. Consequently, the accuracies of 100% gender classification reported.

Shirkey and Gupta [31] developed a steady gender prediction algorithm depending on rectangle features method that are used to describe sub-regions of faces. In results, they achieved a ratio of 90% for gender classification. Fazl-Ersi et al. [32] have compared between three methods that are LBP, Color Histogram (CH) and Scale-Invariant Feature Transform (SIFT) to recognize the gender. Moreover, they used SVM classifier to classify the features extracted from Gallagher facial images database [33]. As a result, the best accuracy they obtained was by combining LBP, CH, and SIFT features, which is reported as 91.59%.

In this study, we used LBP and HOG features in order to predict the gender of the people from facial images.

2. MATERIAL AND METHODS

The design of a pattern recognition system for gender prediction from facial images by using LBP and HOG transformation involves three main parts: pre-processing, features extraction, and classification. In pre-processing part, we isolate unnecessary samples, select training and test samples, and segment patterns from one another and from the background. Features extraction part is related to find features, define pattern classes and representation, and reduce the data by measuring these features. Finally, classification part is very critical section, where a classifier divides the features space into regions and assigns a pattern to a category [34]. These parts are briefly discussed in following sub-sections.

2.1. Databases and Pre-processing Step

In this step, we used three main image databases that are Grey Scale FERET database (Figure 1), Color FERET database (Figure 2), and UTD database (Figure 3). Grey Scale FERET database includes 531 female images and 731 male images, whereas color FERET database contains 400 female images and 543 male images [5]. Similarly, University of Texas at Dallas (UTD) database [44] contains 352 female images and 228 male images of people from 18 to 99 years old (see Table 1). On the other hand, to make the classifier obtains high accuracy performance; our improvements for combining the feature extractors, dimensions alignment, and illumination normalization of the facial images have applied on two separate types of gender prediction experiments. The first type of these experiments applied on FERET grey scale, and FERET color images, while the second type have applied by using FERET database as a Training Set and UTD database as a Test Set.

In the pre-processing phase, we resized the images into "128×192" dimensions by using bicubic interpolation method. Dimensions alignment (image resizing) has advantages of helping the feature extractor to extract same number of features from all images (in our study it helps us to extract 128×192 features from each image using LBP and HOG approaches), and it decreases the computation power, hence increasing the system performance. Whereas, normalizing the illumination by applying Histogram Equalization technique helps to unify luminosity of images, which reflected positively on the accuracy and performance of our system.
GENDER PREDICTION FROM FACIAL IMAGES USING LOCAL BINARY PATTERNS AND HISTOGRAMS OF ORIENTED GRADIENTS TRANSFORMATIONS

Table 1. Number of Images taken from FERET database

<table>
<thead>
<tr>
<th>Data Sets</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grey Scale FERET images</td>
<td>731</td>
<td>531</td>
<td>1,262</td>
</tr>
<tr>
<td>Color FERET images</td>
<td>543</td>
<td>400</td>
<td>943</td>
</tr>
<tr>
<td>UTD</td>
<td>228</td>
<td>352</td>
<td>580</td>
</tr>
<tr>
<td>Total</td>
<td>1,502</td>
<td>1,283</td>
<td>2,785</td>
</tr>
</tbody>
</table>

Figure 1. Examples from Grey Scale Images of FERET database

Figure 2. Examples from Color FERET database

Figure 3. Examples from UTD database

2.2. Feature Extraction Phase

Extracting the image features is the most important part of the gender prediction approach. In this study, two different feature extraction approaches are followed. The first one is HOG approach and the second one is the LBP approach. The details of these approaches are given below.

2.2.1. HOG Features

Histograms of Oriented Gradients (HOG) is one of the local descriptors that largely been used as an appearance based feature extraction method. It provides high performance in many of computer vision problems related to object detection and recognition. The main idea of HOG is that objects in images are described by the apportionment of the edge directions or gradients intensity. The image is divided into small cells, and then the
histogram of gradient directions is computed. For better accuracy, the local histograms can be calculated as a value of the intensity across large spatially connected cells in the image called blocks. Then use this value to normalize all cells within the block from the upper-left corner down to the right one, where the final HOG features are created by concatenating all the normalized cell histograms from each block into a single vector. This normalization gives better stability and invariability in shadowing and illumination [35].

2.2.2. LBP Features

Local Binary Patterns (LBP) is an efficient texture descriptor that is used as a feature representation. The main LBP mechanism is that the input image divided into (N×N) local regions, each local region is threshold a 3×3 neighborhood of each pixel by the central pixel’s value. Then, each pixel assigned with a label by a type of binary pattern, where the distribution of these binary patterns in local region represents the results with an 8-bits integer, where the calculation of these patterns used as a feature representation [36].

2.3. Classification

In this step, different classifiers are used on the facial features that are already extracted. The classification follows a standard steps that a classifier takes one image randomly from a Training set as a test image and use the rest as training images, this called Leave-One-Out technique. In our study, we used two different classifiers that are k-Nearest Neighbor (k-NN) and SVM. The details of the classifiers are given below.

2.3.1. k-NN

K-Nearest Neighbor (k-NN) classifier is one of the simplest classifiers used in machine learning, which depends on closest training examples in the feature range. The main idea of the k-NN classifier is finding a collection of k objects in the training set that are closest to the test object by calculating the distances of all training objects to test object, then gathering k closest objects and calculating the average of them [37]. However, determining k value not easy, because it is affected by some parameters like number of samples that we have in Training set, and the type of feature extractor algorithms used. In this paper, experiments were carried out in order to determine the best optimal k value that can give high performance. Therefore, we have done the experiments (as explained in Experiment 1, Experiment 2, and Experiment 3 parts), using HOG with k-NN, LBP with k-NN, and combined (HOG and LBP) with k-NN respectively.

2.3.2. SVM

Support Vector Machine (SVM) classifier was developed by Cortes and Vapnik in 1995 [38]. It has extensively used as a powerful supervised learning tool for general pattern recognition approaches, and applied to classify some tricky issues with excellent performance on a space of pattern recognition and many other fields. The main task of SVM is based on the ability to minimize both empirical and structural risk, which lead to greater generalization for data classification even when the number of test set is high with limited training set [39].

3. RESULTS AND DISCUSSION

We used three image databases: FERET grey scale images database, which contains 531 female images and 731 male images, FERET color images database, which contains 400 female images and 543 male images, and UTD database, which contains 352 female images and 228 male images. In the pre-processing stage, Dimensions alignment (image resizing) with dimensions of "128×192" pixels, and Histogram Equalization (HE) technique applied on all the facial images. Then, both HOG and LBP features extracted from each facial image and used for gender recognition experiments. In order to measure the performance of the proposed method, we conducted 3 case studies, given as follows:

Case 1: SVM Based Classification:
In this case, we used LBP, HOG and combined feature set (LBP+HOG) as the features and SVM as the classifier. We used FERET database, and we used leave-one-out cross validation technique in order to calculate the performance of the proposed methods. The performance of the suggested methods is given in Table 2.
Table 2. The performance results of SVM classifier when different feature extractors are used

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOG + SVM</td>
<td>FERET Grey Scale</td>
<td>99.68%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>100.00%</td>
</tr>
<tr>
<td>LBP + SVM</td>
<td>FERET Grey Scale</td>
<td>99.84%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>100.00%</td>
</tr>
<tr>
<td>Combined (HOG&amp;LBP) + SVM</td>
<td>FERET Grey Scale</td>
<td>100.00%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Case 2: k-NN Based Classification:
In this case, we used LBP, HOG and combined feature set (LBP+HOG) as the features and k-NN as the classifier. We used FERET database, and we used leave-one-out cross validation technique in order to calculate the performance of the proposed methods (see Table 3). The performance of the suggested methods with respect to k is given in Figure 4.

Table 3. The performance results of the k-NN classifier when different feature extractors are used

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOG + k-NN</td>
<td>FERET Grey Scale</td>
<td>99.60%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>98.30%</td>
</tr>
<tr>
<td>LBP + k-NN</td>
<td>FERET Grey Scale</td>
<td>99.76%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>80.49%</td>
</tr>
<tr>
<td>Combined (HOG&amp;LBP) + k-NN</td>
<td>FERET Grey Scale</td>
<td>100.00%</td>
</tr>
<tr>
<td></td>
<td>FERET Color</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Figure 4. The performance of k-NN with respect to ‘k’ value according to Case 2 experiment.

Case 3: Training with FERET Database and Testing with UTD Database:
All previous experiments trained and tested only on FERET database. Therefore, we supported our complementary experimental investigations on gender prediction by doing another approach of experiments. In these experiments, we used FERET database, which includes 1,262 facial images, as the Training Set, and UTD database, which includes 580 facial images, as the Test Set. The results are presented in Table 4. In addition, the performance of the suggested methods with respect to k is given in Figure 5.
Table 4. Accuracies when using FERET database as the Training Set and UTD database as a Test Set

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOG + SVM</td>
<td>87.76%</td>
</tr>
<tr>
<td>LBP + SVM</td>
<td>94.74%</td>
</tr>
<tr>
<td>Combined HOG &amp; LBP + SVM</td>
<td>98.79%</td>
</tr>
<tr>
<td>HOG + k-NN ( (k = 35) )</td>
<td>91.55%</td>
</tr>
<tr>
<td>LBP + k-NN ( (k = 18) )</td>
<td>83.70%</td>
</tr>
<tr>
<td>Combined HOG &amp; LBP + k-NN ( (k = 24) )</td>
<td>98.08%</td>
</tr>
</tbody>
</table>

As can be seen from the Table 4, the best results are obtained when the combined feature set (LBP and HOG) is used. For this feature set, SVM classifier gives 98.79% performance and \( k \)-NN classifier gives 98.08% performance. The performance ratios are very close to each other and because of that we propose to use any of the classifiers.

Figure 5. The performance of \( k \)-NN with respect to \( k \) value according to Case 3 experiment

4. CONCLUSIONS

In this study, we proposed to use LBP and HOG feature extractors and SVM and \( k \)-NN classifiers in order to predict the gender of the people from facial images. We resized the image in order to reduce the computation cost, and we used histogram equalization in order to minimize the illumination effects in different images. We tested the proposed methods in FERET and UTD databases and we obtained promising results.

This paper shows how the combining of HOG and LBP features can result in much improvement in the recognition accuracy of gender prediction problem. In addition, dimensions alignment and illumination normalization of the facial images are significant factors to improve the performance of our system. Furthermore, our investigation confirms that using correct \( k \) value when using \( k \)-NN classifier in addition to extracting and using the correct image features can strongly affect the prediction accuracy even if we use different number of data. Comparing to other similar studies, our work attains honorable and significant results at the area of identifying the gender information from facial images. As a case study, we used the images from the FERET database as the training set, and images from the UTD database as the test set. In this case, we obtained 98.79% accuracy when the SVM classifier is used and 98.08% when the \( k \)-NN classifier is used. The results show that the proposed method is a robust one, which can work on the images from different databases.

One of disadvantages of the proposed method is that we used directly face images as the input to the predictor. Detecting the facial parts in the images is a challenging problem and may cause significant error in real test conditions. There are many approaches for detecting faces in the images such as Viola and Jones [40], Rowley-Baluja-Kanad [41], Schneiderman-Kanade [42], Roth-Yang-Ahuja [43] etc. Nevertheless, detecting the faces in the images is out of the scope of this work and we directly used facial images to test the performance of the proposed methods.
REFERENCES


T. KHALIFA, G. ŞENGÜL