Effects of Training Set Dimension on Recognition of Dysmorphic Faces with Statistical Classifiers

Şafak Saraydemir\textsuperscript{1}, Necmi Taşpınar\textsuperscript{2}, Osman Eroğul\textsuperscript{3} and Hülya Kayserili\textsuperscript{4}

\textsuperscript{1}Department of Electronics Engineering, Turkish Military Academy, Turkey
\textsuperscript{2}Department of Electrical and Electronics Engineering, Erciyes University, Turkey
\textsuperscript{3}Department of Biomedical Engineering Centre, Gülhane Military Medicine Academy, Turkey
\textsuperscript{4}Department of Medical Genetics, İstanbul University Medicine Faculty, Turkey

Abstract: In this paper, an evaluation using various training data sets for discrimination of dysmorphic facial features with distinctive information will be presented. We utilize Gabor Wavelet Transform (GWT) as feature extractor, K-Nearest Neighbor (K-NN) and Support Vector Machines (SVM) as statistical classifiers. We analyzed the classification accuracy according to increasing dimension of training data set, selecting kernel function for SVM and distance metric for K-NN. At the end of the overall classification task, GWT-SVM approach with Radial Basis Function (RBF) kernel type achieved the best classification accuracy rate as 97.5\% with 400 images in training data set.

Keywords: Dysmorphology, GWT, Principal component analysis, face recognition, SVM, K-NN.

Received February 26, 2013; accepted May 6, 2013; published online August 17, 2014

1. Introduction

Many syndromes with dysmorphic features display congenital malformations, minor or major anomalies [5]. Approximately 3500 different dysmorphic syndromes have already been defined and new ones are added to this number each day. Down syndrome is one of the most common and well defined dysmorphic syndromes. It occurs in approximately 1 in 800 live births. Down syndrome cases have noticeable characteristic facial features including upward slanted palpebral fissures, epicanthic folds, depressed nasal root, midfacial hypoplasia, short/ small nose, small mouth with protruding tongue and small ears as dysmorphism. Clinical geneticists use review articles from the literature and databases (LMD and POSSUM) [20] or defined syndrome descriptions when they do not have enough experience on the syndrome. Several investigations, cytogenetic, molecular cytogenetic, molecular and biochemical tests are performed to reach the definite diagnosis. To perform these tests are frequently very costly and time consuming. Clinical geneticists to finalize all these tests are not sufficient in numbers even in developed western countries so alternative ways to achieve diagnosis is underway.

Computer-aided high performance syndrome classifier techniques have been using to achieve a diagnosis at a short time span and at a cost effective way. In this context, as mentioned before [22] there have been successful cross-dysmorphic syndrome classification studies with face recognition approaches [4, 13, 28]. There are also syndrome recognition [2, 9, 15] and emotion perception [10, 14, 27] studies related to Down syndrome in literature. We aimed to discriminate Down syndrome affecteds from healthy individuals by their facial images in this study. Eroğul et al. [9] achieved classification accuracy of 68.7\% with Elastic Bunch Graph Matching (EBGM) method. Misclassification rate was obtained as 4.7\% with Local Binary Pattern (LBP) technique in the study of Kurt and Nabiye [15]. We introduced a Down syndrome discrimination method based on GWT and statistical classification techniques on Two Dimensional (2D) face images in previous study [22]. In 3-fold cross validation, we obtained classification accuracy of 96.0\% with K-Nearest Neighbor (K-NN) which uses Euclidean distance as distance metric and 97.4\% with Support Vector Machines (SVM) which use linear kernel type. In all these studies, only one frontal face image of an individual was used.

In contrary to these studies related with Down syndrome diagnosis, instead of one image per individual, we used 6 different poses for each 80 individuals (40 Down syndrome affecteds and 40 healthy individuals) in this study. Totally we utilized from a expanded face database including 480 images. We decided to utilize from 6 different poses for individuals, since it is not possible to obtain perfectly frontal images of individuals in daily life. We aimed to examine the effects of increasing dimension of training data set to the classification accuracy by using 6 different still images obtained from videos. Image pre-processing procedures were completed before the feature extraction step. We utilized from GWs for feature extraction process. For the classification of analyzed data after the statistical analysis, we considered K-NN and SVM statistical classifiers. The classification accuracy was evaluated according to different training data set dimensions, kernel functions
for SVM classifier and distance metrics for K-NN classifier.

2. Materials and Method

2.1. Acquisition and Selection of Images

We acquired videos and photographs of Down syndrome affecteds at Friendly Life Down syndrome foundation. The photographs and videos of healthy individuals were obtained at a day-care nursery. At the monthly meetings, the individuals and their parents were informed about the study design. Written informed consent was obtained from the parents. We utilized from the standard equipment and illumination for taking photos and videos. A soft illumination was obtained with two different lighting lamp. A homogeneous background was used for every imaging. We acquired both images and video sequences of the individuals with a digital camera (Nikon Coolpix S2500). Videos allowed extraction of optimal still images. Final image selection was based on pose and facial expression. Selection of sample poses from individuals can be viewed in Figure 1.

![Figure 1. Sample poses of healthy and Down’s syndrome individuals.](image)

Individual’s ages ranged from 1 to 14 years. There were 22 girls and 18 boys in Down syndrome group and 19 girls and 21 boys in healthy group. At least two independent clinical geneticists had established the diagnoses in Down syndrome group.

2.2. Image Preparation and Analysis

The overall block diagram of the proposed syndrome discrimination system can be seen from Figure 2. At the beginning of the proposed system, images were converted from color to the gray scale before the pre-processing procedures. The cropping, histogram equalization and scaling processes were applied to all face images in the database, respectively. The cropping realized to delete redundant parts of images and the histogram equalization fixed the contrast differences in images. At the end of the pre-processing step, the resolutions of all images were adjusted to 80×60 pixels with scaling. This resolution will be denoted as M×N in the following sections when it is necessary. We utilized from the MATLAB program for the implementation of this study.

GWs have been successfully applied on face recognition studies since, their robustness against variance of expression, illumination and pose [24]. Human visual system can be simulated by them. Optimized resolution both in the spatial and frequency domains can be obtained with these wavelets [7, 11]. GWs allow us to extract optimal local features for pattern recognition applications. We can define the GWs as follows [6, 16, 17, 19]:

\[
f_{\mu,v}(z) = \left[ \frac{1}{\sigma^2} \exp \left( \frac{-|z|^2}{2\sigma^2} \right) \right] \exp \left( \frac{i k_{\mu} \cdot z}{\sigma^2} \right)
\]

(1)

Where, \( z=(x, y) \), \( \mu \) and \( \nu \) define the location, scale and orientation of the GWs, respectively. \( \int \) denotes the absolute value and the wave vector \( k_{\mu, \nu} \) is defined to represent the central frequency components in the frequency domain in order to find the similarities between different GWs as follows:

\[
k_{\mu, \nu} = k_{\text{max}} e^{i \phi_{\mu}}
\]

(2)

Where:

\[
k_{\text{max}} = \frac{k_{\text{max}}}{f^\nu} \quad \text{and} \quad \phi_{\mu} = \frac{\mu \pi}{8}.
\]

\( k_{\text{max}} \) is the maximum frequency and \( f \) is the spacing factor between wavelets in the frequency domain [16]. As it seen in Equation 1, the GWs can be generated from one wavelet by scaling and rotation of the wave vector. As a result of this case, the GWs are all similar.

The feature vectors are created by the convolution of the face image points with 2D GWs. We utilized from the five different scales and eight different orientations of (totally 40 different) Gabor filters for each image point. The Gabor filter resolutions are selected as 6×6 pixels in order to reduce the overlap. The magnitudes of complex outputs of Gabor convolutions are used as feature descriptors. The convolution results give us a feature vector of size 40 at each image point. In this way, an individual can be represented as 192000×1 feature column vector by multiplying image resolution M×N and number of Gabor filters \( \nu \times \mu \). At the end of the feature extraction process, we obtain a 192000×480 image feature matrix.

![Figure 2. The overall block diagram of the proposed syndrome discrimination system.](image)
2.3. Training and Test Data Set Preparation

In this study, we utilized from a custom face database which includes 80×60 sized 480 face images. Five different training data sets were created. The created data sets include varying pose counts (from 1 to 5) for each individual and remaining poses are selected as the test data set. Our training data sets include 80, 160, 240, 320, 400 images according to the selected pose count.

2.4. Statistical Analysis and Classification

We utilized from PCA for statistical analysis process. PCA is a holistic method [26], which derives information from the whole face image. There have been many implementations about PCA in pattern recognition studies [1, 3, 23, 25, 30].

The PCA method is implemented as follows [8, 12]: At the end of the feature extraction process we obtained an image feature matrix \( X \) of dimension 192000×480. Considering an individual is represented with One Dimensional (1D) column vector \( X_i \), we can indicate image feature matrix \( X \) as follows, where \( 1 \leq i \leq 480 \).

\[
X = [X_1, X_2, \ldots, X_{480}] \tag{3}
\]

Initially \( \psi \), the mean of \( X \), is calculated by:

\[
\psi = \frac{1}{480} \sum_{i=1}^{480} X_i \tag{4}
\]

The corresponding mean-centered set \( D_1, D_2, \ldots, D_{480} \) is calculated by subtracting each image point value from the mean value over all input images.

\[
D_j = X_j - \psi, \quad i=1, \ldots, 480 \tag{5}
\]

Therefore, we found a 192000×480 difference matrix \( D = [D_1, D_2, \ldots, D_{480}] \). The covariance matrix \( C \) is 480×480, obtained by using this difference matrix as follows:

\[
C_x = D^T D \tag{6}
\]

We intend to achieve a transformation of a point \( D_i \) in data space into a corresponding point \( Y_i \) in the feature space. The points located in the feature space are uncorrelated with others. When we obtain to verify the following expression, we find the eigenvectors \( U \) and eigenvalues \( \Lambda \) of the covariance matrix \( C_x \):

\[
C_x U = U \Lambda \tag{7}
\]

\( \Lambda \) includes the eigenvalues \( \lambda \) of covariance matrix \( C_x \) and \( U \) is the related eigenvectors of \( \Lambda \) as follows:

\[
\Lambda = \begin{bmatrix}
\lambda_1 & 0 & L & 0 \\
0 & \lambda_2 & O & M \\
M & O & O & 0 \\
0 & L & 0 & \lambda_{480}
\end{bmatrix}
\]

\( U = [u_1, u_2, \ldots, u_{480}] \)  

(8)

By utilizing eigenvectors \( u_j \), the face image can be transformed into new point \( y_j \). This point represents PCs of face image. This transformation can be done with the following operation:

\[
y_j = u_j^T D_j, \quad j=1, \ldots, 480 \tag{9}
\]

Let \( U \) be the projection matrix and \( Y \) be the matrix comprising the vectors \( y_j \):

\[
Y = U^T D \tag{10}
\]

After these operations we can reduce the dimension by selecting first \( k < 480 \) eigenvectors and discarding the remaining ones. We investigated the effect of different numbers of eigenvectors from 20 to 100. Since, the satisfaction of the classification accuracy rate results, the number of eigenvectors was selected as 50.

Eventually, when we have performed the PCA, we can represent the image feature matrix as reduced through the PCs. After the PCA method, the image feature matrix size reduced from 192000×480 to 50×480. In this way, feature vector size of an individual decreased from 192000×1 to 50×1.

Several statistical classification techniques were employed to determine the classification accuracy rate for the proposed system by utilizing from 50×1 sized feature vectors. These statistical classification techniques are SVM and K-NN. These techniques differ in complexity and robustness. There are many pattern recognition applications using these classification techniques [18, 21, 29].

In pattern recognition applications, SVM have been using in classification and regression analysis problems. They have been working as supervised learning and help us for analyzing data sets and recognizing patterns. When an input data is given to the SVM, two possible classes which form the output are predicted by SVM. The separation of two data sets with maximum distance is aimed in SVM method.

K-NN classifies patterns according to the closest training examples in the feature space. Taking into account majority vote of neighbor of a pattern, we can assign it to the most common class between its K-NN.

In addition to euclidean distance, cityblock distance metric was used in K-NN method differently from the previous study [22]. The \( k \) value was selected as 1, 3, 5, 7 respectively. We found similar results due to the low variance between classes. We chose to give results according to the 3 for K-NN considering this similarity. The equations of distance metrics used in K-NN method are as described below.

Consider that \( X \) and \( Y \) patterns with dimension \( N \) will be compared. \( x_i \) and \( y_i \) denotes the \( i^{th} \) elements of \( X \) and patterns, respectively.

\[
\text{Euclidean distance}=D_e(X, Y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2} \tag{11}
\]

\[
	ext{Cityblock distance} = D_c(X, Y) = \sum_{i=1}^{N} |x_i - y_i| \tag{12}
\]
Linear, polynomial, quadratic and RBF kernel types were used for SVM. Classification performances of all kernels were examined. The results are given in tables in details.

The classification accuracy rates were used to compare statistical classifiers. We utilized from true and false matching results to compute classification accuracy rates of the classifier types. Proportion of true matching results to all results gives us the classification accuracy. Here, the true matching implies the individuals correctly identified as Down’s syndrome or healthy and is denoted with TP and TN, respectively. The meaning of false matching is incorrectly identification of status of Down’s syndrome and healthy individuals. These status are represented by FN and FP in tables, respectively.

3. Results

We applied GWT method with both SVM and K-NN classifiers on our custom face image database. We extracted 192000×1 sized feature vectors with an application program coded in MATLAB. The statistical analysis and classification procedures were implemented by statistics and bioinformatics toolboxes of MATLAB. Tests were performed on a computer with Intel Core i7 Q740 1.73Ghz CPU and 4 GB RAM.

In this study, color images were converted into gray scale image format with scaling their size to 80×60 resolution. For GWT method, a total of five different training and test data sets were created. While the training data set increases from 80 to 400, test data set decreases from 400 to 80.

During SVM classification, the best results were attained with order of 3 in polynomial kernel and scaling factor of 1 in RBF. We used the city block and Euclidean distances as a distance metric in K-NN classifier. The results were given in tables according to the 3NN.

The classification accuracy results obtained in case of one pose per individual in training data set are presented in order of success in Table 1 as an example. True and false matching results are seen in this table.

<table>
<thead>
<tr>
<th>Classifier Type</th>
<th>Pose Count Per Individual in Training/ (Number of Test Images)</th>
<th>Weighted Average (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-rbf</td>
<td>1/ (400) (%) 77,75, 2/ (320) (%) 83,13, 3/ (240) (%) 90,83, 4/ (160) (%) 93,75, 5/ (80) (%) 97,5</td>
<td>88,59</td>
</tr>
<tr>
<td>SVM-polynomial</td>
<td>1/ (400) (%) 75,75, 2/ (320) (%) 80,94, 3/ (240) (%) 90,42, 4/ (160) (%) 93,13, 5/ (80) (%) 96,25</td>
<td>86,45</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>1/ (400) (%) 68,75, 2/ (320) (%) 78,44, 3/ (240) (%) 89,58, 4/ (160) (%) 91,25, 5/ (80) (%) 95</td>
<td>84,2</td>
</tr>
<tr>
<td>SVM-city</td>
<td>1/ (400) (%) 58,75, 2/ (320) (%) 76,88, 3/ (240) (%) 88,75, 4/ (160) (%) 91,88, 5/ (80) (%) 93,75</td>
<td>82</td>
</tr>
<tr>
<td>SVM-euc</td>
<td>1/ (400) (%) 51,75, 2/ (320) (%) 68,44, 3/ (240) (%) 82,08, 4/ (160) (%) 85,63, 5/ (80) (%) 91,25</td>
<td>75,83</td>
</tr>
</tbody>
</table>

For SVM classification, according to weighted averages, we obtained highest average classification accuracy rates with RBF kernel and lowest average classification accuracy rates with linear kernel.

As one can notice from Table 2, as the test data set increases classification accuracy rates decreases. Between all classifiers, SVM-rbf classifier type has shown better results than the others. Considering the weighted average of classification accuracy rates for five training data sets, SVM-rbf classifier type has the highest average classification accuracy rate 88,59% and K-NN-euc classifier type has the lowest average classification accuracy rate 75,83%. We obtained the best classification accuracy rate 97,5% with SVM-rbf classifier type in case of five poses per individual in training data set (totally 400 training images). In case of one pose per individual in training data set (totally 80 training images) the lowest level (51,75%) was obtained with K-NN-euc classifier type.

We here presented the comparison between the studies [9, 15, 22] according to the classification accuracy results in Table 3. We used the results of in case of five poses per individual in training data set in this table. It can be observed from the table, the database (480 images) used in this study is larger than others. The best classification accuracy rates 97,5% and 96,25% were achieved with SVM-rbf and SVM-polynomial classifier types, respectively. The result of SVM-rbf classifier type was far better than the classifier performances in other studies. The SVM-polynomial classifier result was far better than the result of the K-NN-euc classifier performance in our previous study. The classification accuracy rates 93,75% and 91,25% obtained with K-NN-city and K-NN-euc classifier types, respectively.

Table 1. Matching results in case of one pose per individual in training data set.

<table>
<thead>
<tr>
<th>Matching Results/Classifier Type</th>
<th>True Matches</th>
<th>False Matches</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP 155</td>
<td>156</td>
<td>44</td>
<td>45</td>
</tr>
<tr>
<td>SVM-rbf</td>
<td>145</td>
<td>141</td>
<td>59</td>
</tr>
<tr>
<td>SVM-polynomial</td>
<td>135</td>
<td>132</td>
<td>68</td>
</tr>
<tr>
<td>SVM-quad</td>
<td>119</td>
<td>116</td>
<td>84</td>
</tr>
<tr>
<td>SVM-euc</td>
<td>111</td>
<td>107</td>
<td>93</td>
</tr>
<tr>
<td>K-NN-city</td>
<td>104</td>
<td>103</td>
<td>97</td>
</tr>
</tbody>
</table>
Table 3. The comparison table of accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Database Information (Down/ Healthy)</th>
<th>Pose Count Per Individual in Training</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBG [9]</td>
<td>36 images (18/ 18)</td>
<td>1</td>
<td>68,7</td>
</tr>
<tr>
<td>LBP [15]</td>
<td>107 images (53/ 56)</td>
<td>1</td>
<td>95,3</td>
</tr>
</tbody>
</table>

4. Conclusion

Novice clinical geneticists need quick and reliable information about the presence of dysmorphic syndrome. In this way they can inform the family about recurrence risk of syndrome. Computer aided syndrome discrimination systems can recognize facial features to provide preliminary diagnosis for geneticists. In this study, we utilized from GWs to extract important features on face, since their robustness against several factors and human visual system simulation ability.

For classification task by using the extracted features, both SVM (with four types of kernels) and K-NN (with two distance metrics) classification approaches were implemented. At the end of the classification process, the images belonging to individuals with Down syndrome discriminated from the images belonging to healthy individuals.

We have created five different training data sets according to the pose count per individual in training. In this study, we aimed to compare classification accuracy rates and analyze the effects of the dimension of the training data set to the classification process. In case of five poses per individual in training data set, the best classification accuracy rates were obtained as 97.5% and 96.25% with SVM-rbf and SVM-poly approaches, respectively. The lowest level (51.75%) was obtained with K-NN-euc method in case one pose per individual in training data set. Considering the weighted average of classification accuracy rates, SVM based classification approaches give better results than K-NN based approaches.

Acknowledgements

The videos and photographs of Down syndrome affected were obtained at Friendly Life Down Syndrome Foundation. The authors wish to thank foundation staff for their assistance in contacting the families and are also grateful to the parents of Down syndrome individuals and healthy individuals for their cooperation and participation in the study.

References


Şafak Saraydemir received BS and MS degrees from the Electrical and Electronics Engineering Department, Istanbul University, Turkey in 2002 and in 2006, and the PhD degree in Electrical and Electronics Engineering Department, Erciyes University, Turkey in 2014. His main research area consists of image processing and computer vision.
Necmi Taşpınar received the BS degree in electronic engineering from the Erciyes University, Turkey in 1983, the MS degree in electronic and telecommunications engineering from the Technical University of Istanbul, Turkey in 1988, and the PhD degree in electronic engineering from the Erciyes University, Turkey in 1991. From 1984 to 1992 he was a Research Assistant, during 1992-1996 he was an Assistant Professor, during 1996-2002 he was an Associate Professor and since 2002 he has been Full Professor in the Department of Electrical and Electronic Engineering, Engineering Faculty, Erciyes University, Turkey. His research interests include theory and applications of error-control coding, multi-user detection, channel estimation, PAPR reduction methods, CDMA systems, MC-CDMA systems, OFDM systems, MIMO systems and applications of neural networks, fuzzy logic, adaptive neuro-fuzzy inference system and heuristic approaches to communications problems.

Osman Eroğul received the BS degree in electrical and electronics engineering from the Military War Academy, Turkey in 1981, the MS degree in electrical and electronics engineering from the Middle East Technical University, Turkey in 1985, and the PhD degree in electronic engineering from the Ankara University, Turkey in 1997. He worked at Biomedical Engineering Centre of Gulhane Military Medical Academy. His research interests include biomedical signal processing, medical image processing, speech processing, sleep laboratory studies, biomedical signal classification and medical technology management.

Hülya Kayserili studied medicine at İstanbul Medical Faculty of İstanbul University. She received her PhD on medical genetics in 1998. She worked as a specialist and a lecturer at Medical Genetics Division of Pediatrics Department of İstanbul Medical Faculty and Prenatal Diagnosis and Research Center of İstanbul University between 1991-2004. She received her Associate Professorship degree on Medical Genetics in 2000 and full time professorship in 2009. She has been a faculty member of Medical Genetics Department of İstanbul Medical Faculty since 2004, and she is also the coordinator for Medical Genetics in the educational board of the Faculty.