Face Recognition System Based on Kernel Discriminant Analysis, K-Nearest Neighbor and Support Vector Machine


Mustafa Zuhaer Nayef Al-Dabagh
Knowledge University


Mustafa Alhabib
Cihan University


Firas Husham Almukhtar
Knowledge University, Erbil , Iraq

International Journal of Research and Engineering
ISSN: 2348-7860 (O) | 2348-7852 (P) | Vol. 5 No.3 | March 2018 | PP. 335-338
Digital Object Identifier DOI® http://dx.doi.org/10.21276/ijre.2018.5.3.3

Face Recognition System Based on Kernel Discriminant Analysis, K-Nearest Neighbor and Support Vector Machine

Author(s): [1]*Mustafa Zuhaer Nayef Al-Dabagh, [2]Mustafa H. Mohammed Alhabib, [1]Dr. Firas H. AL-Mukhtar

Affiliation(s): [1]Department of Computer Science, Knowledge University, Kurdistan Region, Iraq
[2]Department of Communications and Computer Engineering, Cihan University-Erbil, Kurdistan Region, Iraq
*Corresponding author: mustafa.nayef@knowledge.edu.krd

ORIGINAL ARTICLE

Abstract: Although many methods have been implemented in the past, face recognition is still an active field of research especially after the current increased interest in security. In this paper, a face recognition system using Kernel Discriminant Analysis (KDA) and Support Vector Machine (SVM) with K-nearest neighbor (KNN) methods is presented. The kernel discriminates analysis is applied for extracting features from input images. Furthermore, SVM and KNN are employed to classify the face image based on the extracted features. This procedure is applied on each of Yale and ORL databases to evaluate the performance of the suggested system. The experimental results show that the system has a high recognition rate with accuracy up to 95.25% on the Yale database and 96% on the ORL, which are considered very good results comparing with other reported face recognition systems.

Keywords: *Face recognition, kernel discriminate analysis, support vector machine, K-nearest neighbor.*

I. INTRODUCTION

Automated Authentication and identification is a need for the development of a new technology. These requirements are increased due to the fact that threats are also increased. There are many technologies that can satisfy the authentication and identification needs, some of these technologies are related to face or voice recognition. Great interest has been made by different academics and researchers on face recognition methods. These methods are gradually developed to include the means of extracting the features of each face image before classifying and recognizing it [1].

The feature extraction process is a very critical step in expressing facial images that could greatly affect the rate of recognition. This process also is very important in making computation and storage cheaper. Eigenface and Fisherface are popular and famous examples of feature extraction methods. With Eigenface, a method of reducing the input dimensionality such as Principle Component Analysis (PCA) is used. PCA method converts images into a low-dimension space and performs a linear matrix transformation that finds the data variance in the projection subspace [9]. Fisherface method focuses on the variation of light direction, and facial expressions, which can be implemented by Linear Discriminant Analysis (LDA) technique. LDA is a supervised scheme that aims at minimizing the within-class variances as well as maximizing the in-between-class distances in the projection subspace. However, in face classification and recognition tasks, we often have problems with small sample size or high dimensional data. Therefore, the traditional LDA is not suitable due to the fact that the within-class scatter matrix is always singular [2]. On the other hand, there is a method called Kernel method that can give an accurate and a robust classification. The main idea behind kernel-based approaches is to map input data to a novel feature space through nonlinear mapping, which produces a set of projective feature vectors by maximizing the between-class covariance and minimizing the within-class covariance [3].

Another important aspect of face recognition is the Classifier based algorithms. One of these algorithms is the Nearest Neighbor (NN). In NN classifier, a single training sample is used such that the samples can be classified by minimizing the distance between the samples, after that these samples can be represented in different classes. In NN method, it classifies the testing samples according to prediction. It predicts the best linear representation for each class [4].

Support Vector Machines (SVM) [5] and AdaBoost [6] methods are two statistical methods used as a framework for the theory of statistical learning. These methods are always suffering from the generalizability problem due to the dramatical different between the testing and training samples [7-8].

Section 2 of this paper explains the design of the face recognition system, involving kernel discriminated analysis (KDA) method and the principle of SVM as a classification method. Section 3 presents the experimental results that evaluate the adopted techniques performance. Conclusions are reviewed in Section 4.

II. DESIGN OF THE SYSTEM

The proposed system consists of three steps: Pre-processing, features extraction and classification. In the Pre-processing step, median filter and Histogram equalization are applied on the face image just to improve the illumination level of the image. In the second steps, non-linear feature extraction methods called kernel discriminated analysis is used to take unique features from the image. The third step is classification, in this step; both of support vectors machine and K-nearest neighbor are applied to classify the features of the face image. The steps of the proposed system are described in the FIGURE 1 below:

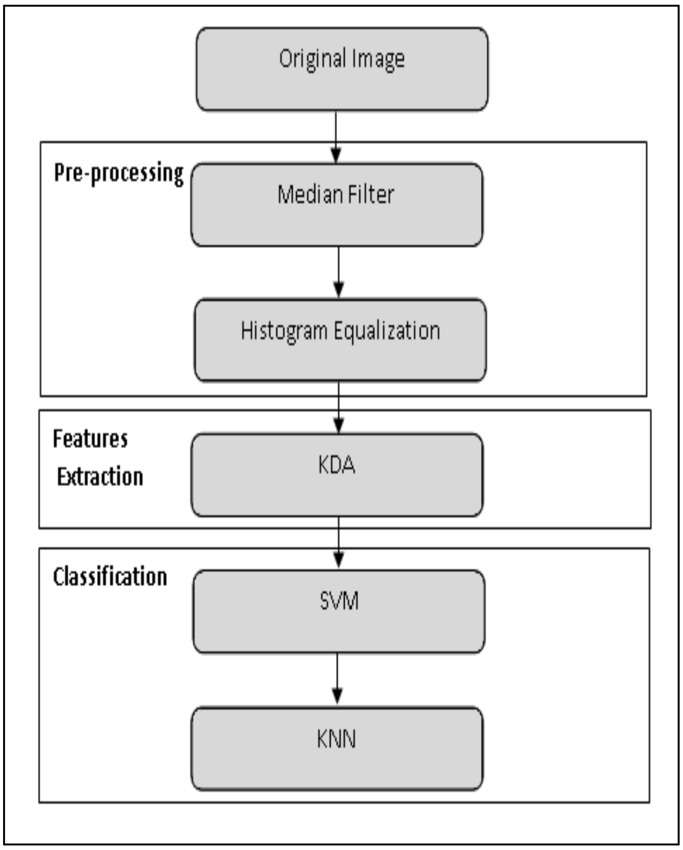

FIGURE 1: The steps of proposed system

A. Pre-processing

Pre-processing is considered the initial step of the system and one of the important steps in any face recognition system because this step affects the results of the next steps. Two common techniques are used: a median filter and histogram equalization are applied in this step. In the beginning, a median filter is used to filter the image from any noise, after that histogram equalization is applied for contrast enhancement.

B. Kernel Discriminant Analysis

KDA is a non-linear model of LDA. The goal of KDA is to obtain the direction that maximizes the separation among classes. To understand it, Let us consider that xi ∈ Rd , i = 1, …., n be training vectors described as an m×m kernel matrix K such that K(xi, xj) = (Ψ(xi), Ψ(xj))). Where, each of Ψ(xi) Ψ(xj) are the embedding of data items xi and xj respectively, which is described based on an inner product through a positive semi-definite kernel function for the feature space. The objective function of KDA is described in equation below:

$$\max_{\varphi} D(\varphi) = \frac{\varphi^T C_b \varphi}{\varphi^T C_t \varphi} \quad \textbf{(1)}$$

In the equation above, the φ denotes a projective function into the kernel feature space; Cb represents the in-between-class and is described in the equation below:

$$C_b = \sum_{k=1}^{c} m_k(\mu_{\Psi}^{(k)} - \mu_{\Psi})(\mu_{\Psi}^{(k)} - \mu_{\Psi})^T \quad \textbf{(2)}$$

C_t in equation (1) represents the total scatter matrices in the feature space and can be described in the equation below:

$$C_t = \sum_{i=1}^{m} (\Psi(x_i) - \mu_{\Psi})(\Psi(x_i) - \mu_{\Psi})^T \quad \textbf{(3)}$$

In the two equations above, μ_{ψ} represents the centroid of the global centroid and $\mu_{\psi}^{(k)}$ represents the k-th class centroids in the feature space [9].

C. Support Vector Machines

Let us suppose that we have two different classes and there are N points belongs to these two classes, such that:

$$\{(x_i, y_i)\}_{i=1}^{N} \; and \; y_i = \{+1, -1\} \quad \textbf{(4)}$$

Where xi represent n dimensional vector and yi represent the class label the vector belongs to. The method SVM uses a hyper plane to separate the two classes' points:

$$W^T x + b = 0 \quad (5)$$

Such that x represents the vector, while w is an adaptive weight vector, and b is the bias, then the functional margin of this hyper plane can be represented as:

$$\begin{cases} w_0^T x_i + b_0 \geq +1 \rightarrow & y_i = +1 \\ w_0^T x_i + b_0 \leq -1 \rightarrow & y_i = -1 \end{cases} \quad (6)$$

If we consider one point W_0 and b_0, the geometrical distance of a point x from the optimal hyper plane is:

$$d(w_0, b_0, x) = \frac{|w_0 x + b_0|}{||w_0||} \quad (7)$$

SVM tries to maximize the distance between the hyper plane and the closest point of these two classes. Therefore the optimal separation of the data is the one that minimized:

$$\emptyset(w) = \frac{1}{2}||w||_2^2 = \frac{1}{2}(w.w') \quad (8)$$

Maximization to the following dual Lagrange formula leads to the optimization of the problem:

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i . x_j) \quad (9)$$

Lagrange multiplier must subject to the constraints:

$$\alpha_i \geq 0, \qquad \sum_{i=1}^{N} \alpha_i y_i = 0 \quad (10)$$

Only a small number of α_i is none zero, each of which corresponds to one training data points. These data points called Support Vectors since they lie on the margin border [10-11].

D. Recognition using K-Nearest Neighbor

This method is used to classify the objects on the training examples in the features space.. It calculates the distance between an instance and the training examples. Euclidian rule is used to determine the distance between two vectors, X=(x1,x2,..,xn) and Y=(y1,y2,..yn) as the following:

$$D(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (11)$$

According to this equation, the object is assigned to the class of the nearest neighbor [12].

III. EXPERIMENTS AND RESULTS

The proposed system was implemented using MATLAB version 9.2. Yale [13] and Olivetti-Oracle Research Lab database (ORL) [14] databases are used to evaluate the performance of this system, which are freely available in the Internet. The Yale database includes 15 distinct classes with eleven various images of each one. These images were taken under different lighting conditions with different facial details and expressions, and were captured using a white homogeneous background. The size of the image was (243 x 320) pixels with 256 available gray levels per pixel. The ORL database includes 40 cases and each case contains 10 images. The resolutions of the images were 112 x 92 with 256 available grey levels per pixel. Samples of Yale and ORL databases are shown in FIGURE 2 and 3 and the recognition accuracy rate is shown in TABLE 1:

FIGURE 2: A sample of images from Yale database.

FIGURE 3: A sample of images from ORL database.

Database	No. of classes	No. of training samples	No. of training samples	Accuracy (%)
Yale	15	60	105	95.25%
ORL	40	160	240	96%

TABLE 1: The recognition accuracy rate

IV. CONCLUSIONS

In this paper, a system of face recognition based on Kernel Discriminant Analysis (KDA) that uses feature extraction and Support Vector Machine (SVM) with K-nearest neighbor (K-NN) classifier is proposed. To evaluate the system performance, Extensive experiments conducted on two common face databases. The experimental results have proven that the proposed system can work in a different condition such as different lighting exposure, facial details, and facial expressions. Also, the results show that the system has a high recognition rate and the accuracy is achieved up to 95.25% on the Yale and to 96% on the ORL databases. In the future work, other types of methods will be applied to enhance the classification such as convolution neural networks (CNNs).

V. DECLARATION

All authors had disclosed no conflicts of interest.

REFERENCE

[1] Xingang Liu, Lingyun Lu, Zhixin Shen, Kaixuan Lu, “A novel face recognition algorithm via weighted kernel sparse representation”, Future Generation Computer Systems, Vol.80, pp. 653-663, 2018,

[2] Hwang-Ki Min, Yuxi Hou, Sangwoo Park, Iickho Song, ”A computationally efficient scheme for feature extraction with kernel discriminant analysis”, Pattern Recognition, Vol. 50, pp. 45-55, 2016.

[3] G. Baudat and F. Anouar, "Generalized Discriminant Analysis Using a Kernel Approach," in Neural Computation, vol. 12, no. 10, pp. 2385-2404, Oct. 1 2000.

[4] T.M. Cover, and P.E. Hart, Nearest neighbor pattern classification, IEEE Transactions on Information Theory,vol. 13, pp.21-27, 1967.

[5] V.N. Vapnik, Statistical learning theory, John Wiley & Sons, New York, 1998.

[6] P. Yang, S. Shan, W. Gao, S. Li, and D. Zhang, Face recognition using Ada-Boosted Gabor features, In FGR ,pp.356-361,2004.

[7] W. Zhang, S. Shan, W. Gao, and X. Chen, Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition, in: Proceeding of International Conference on Computer Vision (ICCV), pp. 786-791,2005.

[8] Mustafa Zuhaer Nayef Al-Dabagh,“Face Recognition Using LBP, FLD and SVM with Single Training Sample Per Person”, in International Journal of Scientific & Engineering Research, Volume 5, Issue 5, May-2014.

[9] M. A. Tahir, E. Khan and A. A. Salem, "Medical text categorization using SEBLA and Kernel Discriminant Analysis," 2015 2nd World Symposium on Web Applications and Networking (WSWAN), Sousse, pp. 1-6,2015.

[10] Zhifeng Li and Xiaoou Tang, "Bayesian face recognition using support vector machine and face clustering," Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. II-374-II-380 Vol.2..2004.

[11] Mustafa Zuhaer AL-Dabagh, Dr. Firas H. AL-Mukhtar."Breast Cancer Diagnostic System Based on MR images Using KPCA-Wavelet Transform and Support Vector Machine", International Journal of Advanced Engineering Research and Science(ISSN : 2349-6495(P) | 2456-1908(O)),vol.4,no. 3, pp.258-263,2017.

[12] K. N. N. Hlaing and A. K. Gopalakrishnan, "Myanmar paper currency recognition using GLCM and k-NN," 2016 Second Asian Conference on Defence Technology (ACDT), Chiang Mai,pp. 67-72,2016.

[13] P. N. Bellhumer, J. Hespanha, and D. Kriegman. (1997, July). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. Retrieved from cvc.yale.edu/projects/yalefaces/yalefaces.html.

[14] Samaria F., Harter A. (1994, December). Parameterisation of a Stochastic Model for Human Face Identification. Retrieved from www.cl.cam.ac.uklresearch/dtglattarchive/facedatabase.html