Image Authentication Using PCA And BP Neural Network

Mohammed Hussein Miry*, Akel A. Alzaiez**
& Abbas Hussein Miry***

Received on: 2/12/2009
Accepted on:4/11/2010

Abstract

In this paper, a recognition system for image identification by using principal component analysis (PCA) and back propagation (BP) Neural Network is proposed. The system consists of three steps. At the very outset some pre-processing are applied on the input image. Secondly image features are extracted by using PCA, which will be taken as the input to the Back-propagation Neural Network (BPN) in the third step and classification. Principal Component Analysis (PCA) is one of the most popular appearance-based methods used mainly for dimensionality reduction in compression and recognition problems, this will reduce the size of training data which it entered to neural network. In our work, The proposed model is tested on a number of images with different value of learning rate. Experimental results demonstrate the proposed model is better, efficient and it reduces the ratio of the number of iteration training to half comparing with results of the Neural Network.

Keywords: Image Identification, Principal Component Analysis (PCA), Back Propagation (BP) Neural Network

تشخيص الصور باستعمال PCA و الشبكات العصبية BP

الخلاصة

في هذا البحث, اقترح نظام التمييز لتشخيص الصور بواسطة استخدام تحليل المكوّن الرئيسي (PCA) و الشبكات العصبية من نوع الانتشار الخلفي (BP) . هذا النظام يتكون من ثلاث خطوات. الخطوة الأولى: معالجة أولية على الصور المدخلة, الخطوة الثانية: استخلاص المعلومات من بيانات الصورة باستعمال (PCA), الخطوة الثالثة:تسليط المعلومات المستخلصة من الخطوة الثانية على الشبكات العصبية من نوع الانتشار الخلفي (BP) ومن ثم تصنيفها. تحليل المكوّن الرئيسي هي إحدى الطرق المشهورة لتقليل حجم البيانات في مسائل الضغط و التمييز, هذا يؤدي إلى تقليل حجم البيانات التي تدخل إلى الشبكة العصبية. في هذا العمل, اختبرت الطريقة المقترحة للعديد من الصور لقيم مختلفة لنسبة التعلم. النتائج التجريبية تبين إن النموذج المقترح أفضل و أكفأ و يقلل عدد المرات الأزمة للتعلم إلى النصف بالمقارنة مع الشبكة العصبية.

1. Introduction

In recent years, there is a great interest of many researchers on the recognition problem. Among these researchers are the engineers, studying this popular problem in different fields and in different points of view. There are several application areas of recognition in our real life such as identification of personnel using credit cards, passport checks, etc [1,2]. Neural networks have been shown to obtain successful results in system identification. But they use static mapping schemes, the weights' updating do not utilize information on

* Electrical & Electronic Engineering Department, University of Technology /Baghdad
** Engineering College, University of Baghdad/Baghdad
*** Engineering College ,AL-Mustansiriyah University / Baghdad

the local data structure, so the function approximation is sensitive to the training data [3].Machine learning offers one of the most cost effective and practical approaches to the design of pattern classifiers for a broad range of pattern recognition applications. The performances of the resulting classifier rely heavily on the availability of a representative set of training examples. In many practical applications, acquisition of a representative training data is expensive and time consuming[4]. Consequently, it is not uncommon for such data to become available in small batches over a period of time. In such settings, it is necessary to update an existing classifier in an incremental fashion to accommodate new data without compromising classification performance on old data [4].

In our work we have combined two approaches. We have applied back propagation neural network (BP) with principle component analysis (PCA) to extract a compressed representation of a set of images. Then testing was used to calculate similarity between training and test image.

The rest of this paper is organized as follows. Section 2 gives the brief introduction of the mathematical preliminaries of back propagation neural network. Section 3 introduction of the PCA. The proposed method is described in Section4. Experimental results and discussion are presented in Section 5.

2. Back Propagation Neural Network

The type of neural network used in this approach is the multilayer neural network with the sigmoid function as activation function. This type of network consists of an input layer, an output layer, and one or more hidden layers [5] as shown in figure 1. The choice of multi layers neural network is based on the fact that this type of network is a supervised neural network that learns through a process called back propagation, which is a form of gradient descent, which is suitable for our problem. Here, the network is supplied with a series of input and corresponding correct (desired) output. The network then tries to set its own parameter until it can approximate an unknown function that can associate input data with corresponding desired output [6]. BP learning consists of two passes through the different layers of the network[7]:

1-Forward pass.

2-Backward pass.

During the forward pass, the synaptic weights of the network are all fixed. During the backward pass, on the other hand, the synaptic weights are all adjusted in accordance with an error-correction rule [8].A BP consist of layers of interconnected perceptions denoted as the input layer, the hidden layer and the output layer. The number of the input units and the output units are fixed to a problem, but the choice of the number of hidden units is somehow flexible. Too many hidden units may cause over fitting, but if the number of hidden units is too small, the problem may not converge at all. So the number of nodes in the hidden layer can be varied based on the complexity of the problem and the size of the input information. The learning factor that significantly affects convergence speed as well as accomplish avoiding local minimization, is the learning rate. The learning rate (η) determines the portion of weight needed to be

adjusted. However, the optimum value of η depends on the problem. Even though as small learning rate guarantees a true gradient descent, it slows down the network convergence process. If the chosen value of η is too large for error surface, the search path will oscillate about the ideal path and converges more slowly than a direct descent., the learning is adjusted to bring the network out of its local minimal and to accelerate the convergence of the network[8] .

BP Program-Training Process:

Step 1: Design the structure of neural network and input parameters of the network.

Step 2: Get initial weights W and initial q (threshold values) from randomizing.

Step 3: Input training data matrix X and output matrix T.

Step 4: Compute the output vector of each neural units.

(a) Compute the output vector H of the hidden layer

$$net_K = \sum w_{ik} x_i - q_k \quad (1)$$

$$H_k = f(net_k) \quad (2)$$

(b) Compute the output vector Y of the output layer

$$net_j = \sum w_{kj} H_i - q_j \quad (3)$$

$$Y_j = f(net_j) \quad (4)$$

(c) Compute the root of mean square

$$RMS = \sqrt{\frac{\sum (y_i - T_j)^2}{n}} \quad (5)$$

Step 5: Compute the distance d

(a) Compute the distance d of the output layer

$$d_j = (T_j - Y_j) f'(net_j) \quad (6)$$

(b) Compute the distance d of the hidden layer

$$d_k = (\sum_j d_j w_{kj}) f'(net_j) \quad (7)$$

Step 6: Compute the modification of W and q (h is the learning rate, a is the momentum coefficient)

(a) Compute the modification of W and q of the output layer

$$\Delta w_{kj}(n) = h d_j H_k + a \Delta w_{kj}(n-1) \quad (8)$$

$$\Delta q_j(n) = -h d_j + a \Delta q_j(n-1) \quad (9)$$

(b) Compute the modification of W and q of the hidden layer

$$\Delta w_{ik}(n) = h d_k X_i + a \Delta w_{ik}(n-1) \quad (10)$$

$$\Delta q_k(n) = -h d_k + a \Delta q_k(n-1) \quad (11)$$

Step 7: Renew W and q

(a) Renew W and q of the output layer

$$w_{kj}(p) = w_{kj}(p-1) + \Delta w_{kj} \quad (12)$$

$$q_j(p) = q_j(p-1) + \Delta q_j \quad (13)$$

(b) Renew W and q of the hidden layer

$$w_{ik}(p) = w_{ik}(p-1) + \Delta w_{ik} \quad (14)$$

$$q_k(p) = q_k(p-1) + \Delta q_k \quad (15)$$

Step 8: Repeat step 3 to step 7 until converge.

BP Program-Testing Process:

Step 1: Input the parameters of the network

Step 2: Input the W and q

Step 3: Input the unknown data matrix X

Step 4: Compute the output vector

(a) Compute the output vector H of the hidden layer (according to equation (1) and equation (2))

$$net_K = \sum w_{ik} x_i - q_k$$

$$H_k = f(net_k)$$

(b) Compute the output vector Y of the output layer(according to equation (3) and equation (4))

$$net_j = \sum w_{kj} H_i - q_j$$

$$Y_j = f(net_j)$$

3. Principle Component Analysis

Principle component analysis (PCA) is a powerful methodology for a wide variety of applications. For example, it can be applied to data compression and feature extraction in pattern recognition. The principal component is the respective eigenvectors corresponding to the largest and smallest eigenvalues of the autocorrelation matrix of the input signals [9].Principal Component Analysis (PCA) is also known as Eigenspace Projection or Karhunen-Loeve Transformation [10]. It projects images into a subspace such that the first orthogonal dimension of this subspace captures the greatest amount of variance among the images and the last dimension of this subspace captures the least amount of variance among the images. The main goal of PCA is the dimensionality reduction, therefore the eigenvectors of the covariance matrix should be found in order to reach the solution. The eigenvectors correspond to the directions of the principal components of the original data; their statistical significance is given by their corresponding eigenvalues [11].

Basic Steps of PCA Algorithm

Determine PCA subspace from training data. *i*th image vector containing *N* pixels is in the form

$$x^i = [x_1^i, \quad x_2^i, \quad x_3^i, \quad ....., \quad x_n^i] \tag{16}$$

Store all *p* images in the image matrix

$$X = [x^1, \quad x^2, \quad x^3, \quad ....., \quad x^p] \tag{17}$$

Compute covariance matrix

$$\Omega = XX^T \tag{18}$$

Compute eigenvalues and eigenvectors

$$\Omega V = \Lambda V \tag{19}$$

where Λ is the vector of eigenvalues of the covariance matrix.

Order eigenvectors

$$V = [v_1, \quad v_2, \quad v_3, \quad ....., \quad v_p] \tag{19}$$

Order the eigenvectors in V according to their corresponding eigenvalues in descending order. Keep only the eigenvectors associated with non-zero eigenvalues. This matrix of eigenvectors forms the eigenspace V, where each column of V is the eigenvector. Visualized eigenvectors of the covariance matrix are called eigenfaces [12].

4. Proposed Method

In this section, we use the PCA with BP network in the recognition with different number of images (4, 6 and 8) shown in figure 2. The procedure of the recognition can be described as:

Step1: Obtain PCA coefficient for each input image. The percentage of coefficient is 50% of original image is taken which reduce the training of neural network.

Step 2: Input the PCA coefficient into the BP network, use the algorithm described in section 2 to train the network until the training error. When the training process is complete, some characteristics of these coefficient are stored in the weights and transfer functions of the network.

Step 3: Input the test images which are polluted by the noise and

disturbance, we can get the recognition result from the output of the network.

5 . Experimental Results

To evaluate the performance of the proposed method, we used face recognition databases(figure 3). All of the test images resize into 128*128. We used three-layer Artificial Neural Network (ANN) with Back-Propagation (BP) as the classifier. The number of nodes in the input layer is set to the number of features provide by the coefficients selected from PCA. The number of nodes in the output layer is equal to the number of testers to be recognized in the testing images. The number of nodes in the hidden layer is set to 5~10 depending on the number of training samples, the number of features used to train the ANN, and the number of ANN outputs. The results obtained based on the following test

1- Testing the NN with different images.

2- Various experiments were done with noisy faces images. We add different types of noise for the testing images (figure 4). All the algorithms are implemented in MATLAB 7.0.1 and executed o using Mat lab 7.0 on a Celeron M 1.73GHz platform with 2G memory. To evaluate our experiments, we defined a performance metrics to gauge the success of our schemes.

Identification rate (IR) [13] :

IR(%) = (True Identify / Number of tested Faces) * 100% ….(20)

Table 1, shows the identification rate for the proposed method and authentication based only NN. Tables (2 and 3) give the comparison of recognition result between the PCA-BP neural network and BP network, with different value of learning value (0.1 and 0.3) with training error are 0.01 and 0.001. From the tables we notice that the number of training reduced in the proposed method especially in the large number of image (8 images). Table 4, shows the time required for training for proposed method and authentication based only NN, from this table we notice that the propped method need small time for training compare with other method.

6. Conclusions

Image authentication is an important problem, which is extremely difficult in practice, since no contextual information is available. In this work, we proposed method a novel solution to the previous problem by using a PCA-BP based classifier. The proposed method is based on PCA and BP networks as a feature extractor and as the classifier respectively. It shown that PCA and BP can be applied to extract a relatively small number of features from the high-dimensional data. While PCA reduces feature dimensionality (reduction the amount of training data led to minimize the time and the number of training), the proposed method of authentication gave an accurate result on a given database. We can observe that this combination of principle component analysis for information representation and back propagation Artificial Neural Network for final classification gives good results. These performances are better, give high classification accuracy and is extremely fast than those given by the Neural Network only, In any case these results show how a good preprocessing technique, as the

principle component analysis, and a non-linear decision process, as neural networks, can extract very complicate rules even from a reduced training set.

References

[1] O. Toygar, A. Acan, "Face Recognition Using PCA, LDA AND ICA Approaches on Colored Images" Journal of Electrical & Electronics Engineering Vol.3, No.1, pp.735-743, 2003.

[2] J. L. Raheja, U. Kumar, ''Human Facial Expression Detection From Detected In Captured Image Using Back Propagation Neural Network'', International Journal of Computer Science and Information Technology (IJCSIT), Vol.2, No.1, February 2010.

[3] W. Yu and A. Ferreyra, "System Identification with State-Space Recurrent Fuzzy Neural Networks" IEEE Conference on Decision and Control December 14-17, 2004 pp.5106- 5111.

[4] R. Polikar, , L. Udpa,, S. S. Udpa,, V. Honavar, "Learn++: An Incremental Learning Algorithm for Supervised Neural Networks" IEEE Transactions on Systems , Man, and Cybernetics-Part c: Applications and Reviews , Vol . 31, No. 4, pp. 497-508, November 2001 .

[5[A. Choudhary, R. Rishi, S. Ahlawat, V. S. Dhaka ''Optimal Feed Forward MLP Architecture for Off-Line Cursive Numeral Recognition'' International Journal on Computer Science and Engineering (IJCSE) Vol. 02, No.01S, 2010, 1-7

[6] B.S. Murdianto1, R. Sutattyo, A. Haris, "Predicting Lateral Sonic Log Using Artificial Neural Network and Multiattribute Transform", Seminar Nasional MIPA 2005 ,FMIPA-Universitas Indonesia Depok J1E-04,24-26 November 2005

[7] M. C. Lee and C. To ''Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress'' International Journal of Artificial Intelligence & Applications (IJAIA), Vol.1, No.3, July 2010

[8]D. C. Reddy and K. Ghosh, "Identification and Interpretation of Manufacturing Process Patterns Through Neural Networks", Mathl.Comput. Modeling, Vol. 27, No. 5, pp. 15-30, 1998.

[9] Q. Zhang and Y.W. Leung ,"A Class of Learning Algorithms for Principal Component Analysis and Minor Component Analysis",IEEE Transactions on Neural Network ,Vol. 11, No. 2, March 2000.

[10]A. Pentland , "Face recognition using eigenfaces", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, 1991.

[11] R. Calvo .,M. Partridge , and M. Jabri , "A Comparative Study of Principal Components Analysis Techniques", In Proc. Ninth Australian Conf. on Neural Networks, Brisbane, QLD., pp. 276-281, 1998.

[12] J. Mazanec , M. Meli·sek ,M. Oravec ,J. Pavlovicov," Support Vector Machines, PCA And LDA In Face Recognition" Journal of Electrical Engineering, VOL. 59, NO. 4, 2008, 203-209

[13] G.Sugiarta 1, R. Bambang , Suhardi ," Feature Level Fusion of Speech and Face Image based Person Identification System'' 2010 Second International Conference on Computer Engineering and Applications. Vol.2 Year: 2010 , Page(s): 221 - 225.IEEE Conferences.

Table 1: Identification Rate for Different Methods

Method	No. of Test Images	IR(%)
Proposed	25	96%
NN only	25	83%

Table 2: Numbers of training with different number of images with η=0.1.

No.of test image	training error 0.01		training error 0.001	
	No. of train (NN)	No. of train (PCA-NN)	No. of train (NN)	No. of train (PCA-NN)
4	2961	2005	8225	6349
6	6193	4789	10001	8097
8	8121	6849	18529	15247

Table 3: Numbers of training with different number of images and learning rateη=0.3.

No . of test image	training error 0.01		training error 0.001	
	No. of train (NN)	No. of train (PCA-NN)	No. of train (NN)	No. of train (PCA-NN)
4	1165	877	2953	2297
6	2925	2293	6145	5575
8	4969	3057	8401	6369

Table 4: Times to train (Minute)

No. of test image	training error 0.01		training error 0.001	
	(NN)	(PCA-NN)	(NN)	(PCA-NN)
4	21	14	26	19
6	24	18	30	24
8	28	23	33	27

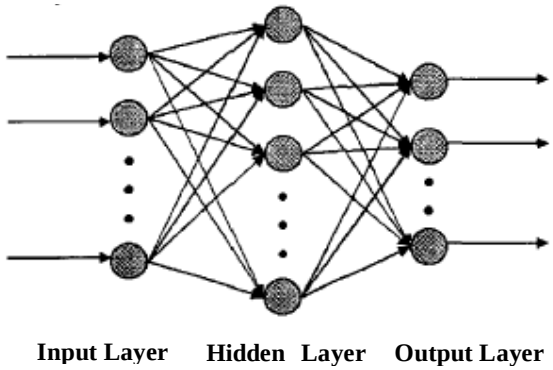

Figure 1: The BP Neural Network

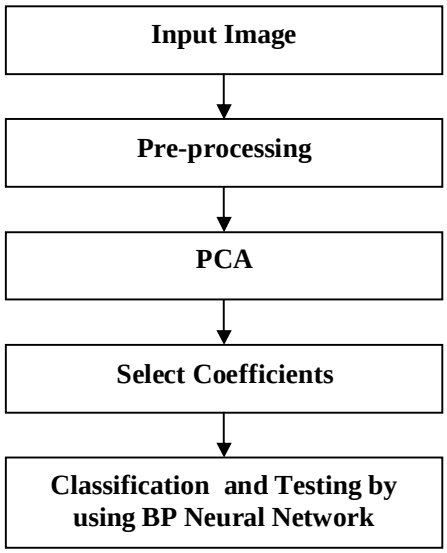

Figure 2: The Proposed Algorithm

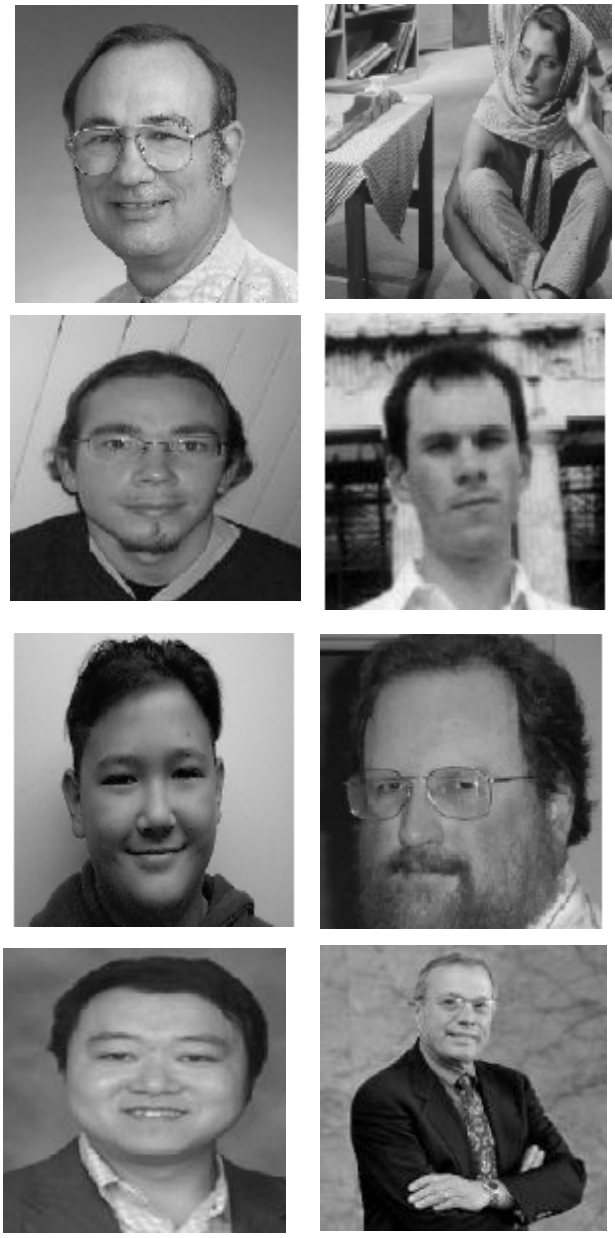

Figure 3: Test Images.

Figure 4: Test Images with Add Noise .