

Face Recognition Model using Error Correcting Output Codes

Hamdollah ghamgin^{1*}, Mohammad Reza Rzaq Poor²

¹Instructor, Department of Engineering, Payame Noor University, po Box 19395-3697, IRAN

²Collegian, Department of Engineering, Payame Noor University, po Box 19395-3697, IRAN

*Corresponding Author Email: hamghamgin@yahoo.com

Abstract

Classifiers Ensemble One of the common methods to achieve high accuracy in classification problems. Error-correcting output code of the algorithms for combining multiple classifications in machine learning for classification problems have been proposed in recent years has become of great interest. In this method, first the class was given a two-class problem is converted to the number below. Then each of the issues separate binary classifiers are trained and tested and finally answer some basic classes of binary classifier combines the results obtained. In this paper, performance of different algorithms ECOC on the problem of faces recognition, a multi-class problem, and the review and then provides a useful model for face recognition process. Experimental results on ORL data including face images of 40 different faces is the case, the performance of the proposed algorithm compared with existing methods show.

Keywords: Combining classifiers, Face detection, Error correcting output codes neural networks.

Introduction

The ultimate goal of a pattern recognition system to achieve the highest possible classification for the problem is desired. Since there is no classification algorithm is the only fully are suitable for all problems, the combinations of classifiers as a way to increase the performance of classifiers have been proposed. The combination of classifiers is proven to increase recognition rates can be more independent. So we try to increase the diversity among the base classifiers. There are many ways to combine different classifiers such as Boosting, Mixture of Experts, ECOC noted. They each work differently. Boosting classifier based on the distribution of the input samples for training classifiers in the direction of change is more difficult to focus on the specimen. In Mixture of Experts during training classifiers try each of them part of the input space and the particular practice expert. One of the ways to overcome the complexity of combining ECOC label when the number of classes is large and good results for synthetic and real data. In this way, a complex multi class problem under the simple binary matrix-matrix code is called, it will break down. The following two-class problem according to the 0 and 1 in each column of the matrix code defined by the classifier trained accordingly. Measured for each prototype vector is formed by base classifiers arrays, the output of the base classifiers. Then specify the class sample should be similar or close to a row of a matrix-vector code words is called is measured. Many methods have been proposed to measure the similarity of the most important and easiest to measure "minimum distance" is. Despite the many advantages that this measure has widespread use it, there are some weaknesses that caused the error is classified. All these measures are calculated from dichotomies have the same effect, while these dichotomies facing problems due to different binary logic that has the impact and importance are different in the distance calculation. In recent years a lot of attention in research areas related to face recognition biometrics, pattern recognition and computer vision is presented. Security and commercial applications as well as some methods are used to identify the faces. These applications include the control of security, access control, identification of the offender (for example, passport control), reconstruction of the face, orbits are the interface between humans and computers. In addition, there are other detection methods for face recognition with one of them can be identified through fingerprints, hands, voice, eyes, ears and signature touches. For different authentication methods require different data collection methods is that they can be based on factors such as a person's level of cooperation, accuracy, cost, measured easily categorized. Identified with the help of face images because it requires little cooperation from others and nature of non-interference is considered a security company has been addressed. For years, researchers have studied the problem of face recognition algorithms have been proposed classifier is much to it. None of these methods alone cannot solve this problem, and each category has its own strengths and weaknesses. So as a solution for combining classifiers utilizes the strengths of different classifiers are used thereby increasing the accuracy of

classification. In this paper, as one of the important ECOC algorithm combining classifiers, used for solving this problem based on the performance model for the problem of face identification has been proposed. Then introduce a measure of the reliability of the proposed model and compare it with other existing methods deals. The rest of the paper is as follows: In the next part of the ECOC algorithm and different stages of its implementation are described in detail. In the third part of our proposed method for face recognition problem based on standard algorithms to predict the system reliability has been identified and Part IV compares the results of empirical proposed algorithm and the standard has model. Finally, this paper summarizes the results are presented.

ECOC algorithm

As a general method for solving multi-class problems arose, and the main idea of the transfer of information and communication signals using the code error correction arising from the transfer of the property have been taken. The main features of this code to create the replicated data should be sent. So if some errors occur due to noise in the signal transmission channel, needs to be properly signal found in the destination.

In this section we use a simple example to explain the method in classification problems.

Suppose we want a four-class classification problem that A·B·C·D is a class label. To start each label a different code word length n is assigned to do so $n \geq \log_2(C)$.

Table 1. Using a binary matrix encoding j 10 for four classes.

Lable	Code word
A	0110110001
B	0001111100
C	1010101101
D	100011010

In other words, we work with each of the classes with over 10 unique word was code. Because of the code words, said code matrix. The next step consists of creating ECOC classifier based binary classifier for each column of the matrix code. So in this example, we create a binary classifier 10. I have one class classifier j j-th column of the matrix code examples are classes that are in rows 1 and -1 are the rest of the class. For example, the third paragraph of class differentiation examples classes are A·C and B·D classes. Thus, as shown in Figure 1, is comprised of 10 of the 10 binary classifier trained with them. After training classifiers to classify each test sample output classifiers as a vector of length 10 [0011110100] form. For measurements near the samples of different classes, different methods are the most commonly used Hamming distance of the code is the vector of each class, respectively, four, two, five and seven is. Minimum distance marker closest sample to the corresponding class and the second class is a sample belongs to.

ECOC algorithm

On the training:

For each column of the matrix code, forming the corresponding binary sequence of zeros (0) Howe 1 for training samples (decomposition).

Basic binary classifier trained according to a defined problem.

Stage test: All output is calculated based binary classifier for each sample and the test set consists of the $\bar{Y} = [y_1, y_2, \dots, y_n]^T$

output vector, Y_j is the j-th class.

In order to classify each sample (reconstruction):

Distance of the output vector from all rows of the matrix are calculated Code:

$$(1) L_i = \sum_{j=1}^n |z_{ij} - y_j|$$

i-th row and j-th column of the array Z_{ij} is the matrix code.

Sample belongs to the class with the minimum distance.

So overall ECOC system combines two phases are formed in the following section for more details on them:

Analysis of ECOC (encoding)

At this stage, a multi-class problem according to matrix binary code is converted into a problem. Each of the binary instruction set defines a learning algorithm for the training set consists of a binary decision model. Below are a number of methods to produce matrices refers to the code:

A-Generating a random code: Many papers on the effectiveness of the method were ECOC. Matrix produced by this method have $10 \log_2(C)$ column for a class C is the problem. The matrix elements of the set {0, 1} are randomly chosen.

B-BCH method: In this way, the relationships and rules of matrix algebra to generate code with minimum distance between rows is used to add binary values based on BCD numbers are proven to. Number of columns and rows are always a power of two, so in many classification problems, the number of columns and rows of the matrix to remove.

C-1-vs-All: This code is a square matrix whose main diagonal corresponding to -1, and the remaining members of the array matrix are -1. Each classifier consists of the code deals with the compression of a class than other classes.

D-1-vs-1: This method, which extends from the code matrix is binary, with 0 array, the two arrays of 1 - and 1 + is. In this case, the i-th row of an array 0 Hurston, indicating a lack of training classifier corresponding column i is in class. Each classifier corresponding to the code to compare the two classes without regard to the other classes is.

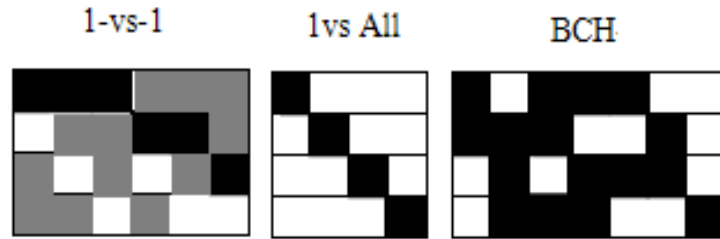


Figure 1. The structure of the above three types of matrix displays for the four classes.

All these methods have two major aims in the following code to generate the matrix. In other words, a good code matrix must have the following characteristics:

Resolution of Rows: Each class label corresponding code word from other classes code word Hamming distance is sufficient. This property makes up office network is more error correction properties.

Resolution columns: each column of a matrix must also be far enough with the other columns. This feature also reduces the correlation is caused by binary problems that led to independence in decision-making is of dichotomies.

Reconstruction of ECOC (decoded)

After decomposing the original problem into binary problems and create algorithms based binary classifiers decisions are in the testing phase for each sample of the class to be determined. This step must be set using the solutions of the binary classifiers is the main class of the sample, and set reconstruction. The easiest and most effective methods of reconstruction or decoding codes, Hamming and Margin distance measure can be calculated by Equation 1, the observational points are defined as follows:

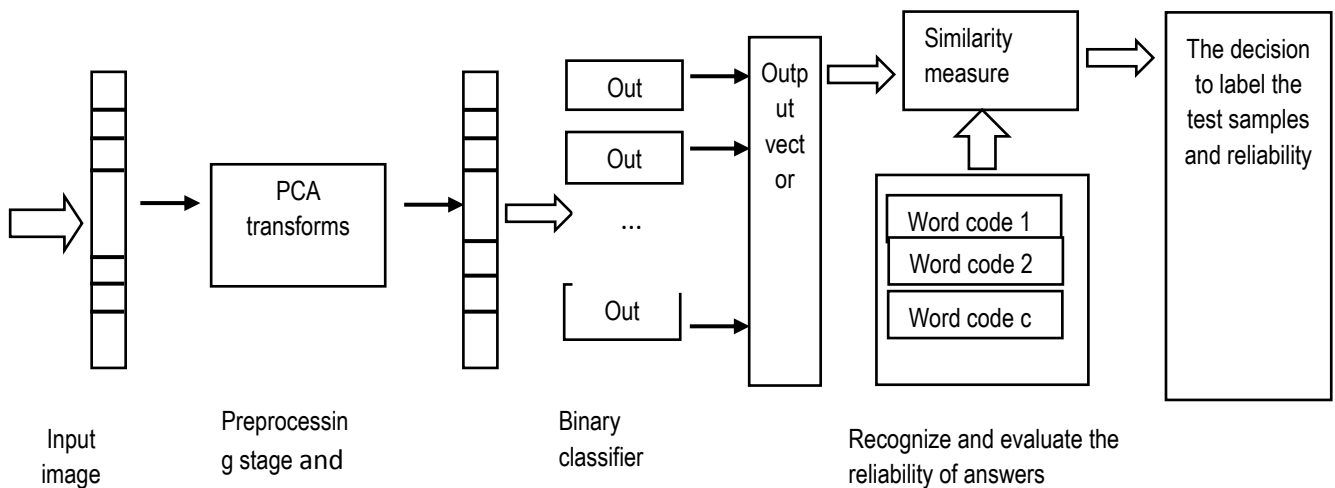


Figure 2. The structure of the matrix to form a classification system based on ECOC algorithm. Home black and white array 1 array, 1 -, and gray represents 0.

Results

In this section, to evaluate the performance of the proposed model for the problem of face identification, we use the ORL dataset. This data set consists of 40 two-dimensional images and image for each of 10 different states face different view angles (total 400 samples) is. A few sample images are shown in the following figure. In this experiment, we first five images per person to train the model to remove redundant information and images for more efficient network, as described in the previous section, the first component into 50 PCA use.

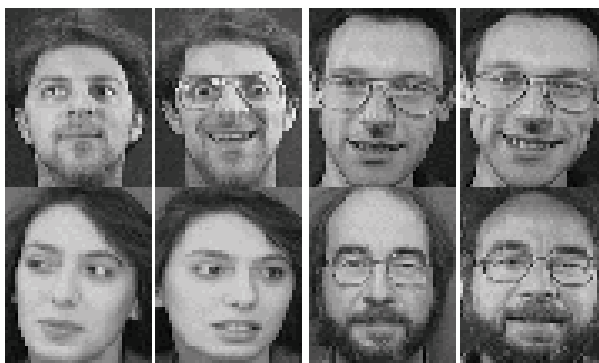


Figure 3. Example of image datasets with different conditions ORL face.

A-hamming Decoding: This method is based on binary classifiers output, y_j as the binary values of zero (0) or 1 are.

B-Margin Decoding: Binary classifiers based on standard output, y_j , the actual values are displayed. Therefore, the network output vector is a sample of the actual values.

The proposed method in the reconstruction

At this point, the proposed model for the problem of face recognition based on ECOC algorithm explains.

This model consists of two stages: the stage show and pre-processing of input images and the recognition and evaluation of the reliability of the answers given. During pre-processing, after normalization of images and resize those to 32x32, the conversion of PCA to reduce the use said. Another advantage of using this conversion eliminates the extra features, and noise on the input network, most of the learning network will not help, in the recognition of the algorithm ECOC as a decision model we use has a high accuracy for the problem is.

In the ECOC decomposition methods, we generate a random matrix codes disposal. Matrix for the ten classes of binary code in accordance with 10 rows and $10 \log_2(10) \sim 33$ columns is the number of classes that represent the decomposition of the binary is 33. To create binary dichotomies of Multilayer Perceptron Neural Network MLP with back-propagation learning rule was used. According to the structure of the network input and output layer neurons is 16 and 1 respectively. The optimum number of neurons in the middle layer also tested experimentally and the number of neurons 10 and 40 reps of 100 learning algorithm, respectively. As Table 2 shows, a total of 30 neurons in the intermediate layer have a lower error rate than the other values.

Table 2. Recognition error rate ECOC networks to changes in the number of neurons in the middle layer of dichotomies.

The number of neurons in the middle layer	20	25	30	35	40	45	50
Recognition error rate	14.13	13.88	13.33	11.81	11.37	13.56	12.35

The learning algorithms for finding the optimal number of repeat value of 200 to 1000 in accordance with Table 3 were tested. As shown in the table are the first to recognize the error rate is reduced by increasing the number of learning to reach its minimum extent Over fitting then due to the phenomenon of excessive learning algorithm occurs repeating Rate network Recognizing will be reduced.

Table 3. Recognition error rate of model-based face detection algorithm to changes in the number of occurrences ECOC learning algorithm based on binary classifiers.

The occurrence	200	400	600	800	1000
Recognition error rate	11.37	9.44	8.01	6.92	7.7

After finding the optimized parameters of dichotomies and training with training samples, the proposed method with two different reconstruction methods commonly used in the previous sections we introduced, as well as two other methods Weighted proposed in (Allwein et al., 2000) that the algorithm Adaboost to uses a binary dichotomies, we review and compare.

Table 4. Recognition error rate ECOC network with a variety of reconstructing methods in the face recognition problem.

Various methods of reconstruction	Hamming decoding	Margin decoding	Loss-based Decoding	Loss-based Decoding
Recognition error rate	7.21	6.92	8.8	8.1

Table 5. Comparison of proposed model with the conventional algorithms to identify two-dimensional figures.

Various algorithms for face recognition	Support Vector Machine	Decision Tree	Bayesian Classifier	The proposed method
Recognition error rate	7.06	10.80	8.09	6.92

The next step is to compare the proposed model with the conventional algorithm, we identify the face. As shown in Table 5, despite the simplicity of the proposed method, a fairly substantial reduction in the recognition error for the problem of recognizing faces.

Conclusion

One of the most commonly used methods for combining classifiers for multi- class problems (ECOC) in identifying two-dimensional figures are explored. ECOC decomposition methods implemented in two stages and each stage is Bazzaz was introduced in a variety of ways .compared with conventional approaches and algorithms for identifying faces , better performance and high reliability in the classification error rate was reduced enjoy .

Acknowledgement

This study supports Engineering and Technology Research Center, Shahed University, Iran Telecommunication Research Center conducted.

References

Allwein EL, Shapire RE Singer Y, 2000. Reducing Multiclass to Binary: A Unifying Approach for Margin Classifiers. Journal of Machine Learning Research. 1: 113-141.

Breiman L, 1996. Bagging predictors. Machine Learning. 26(2): 123-140.

Dietterich TG, Bakiri G, 1995. Solving multiclass learning problems via error correcting output codes. J. of Artificial Intelligence Research. 2: 263-286.

Gong S, McKenna SJ, Psarrou A, 2000. Dynamic Vision: from Images to Face Recognition, Imperial College Press and World Scientific Publishing.

Haykin S, 1999. Neural Networks: A Comprehensive Foundation, Prentice Hall International, Inc, New York, NY, USA.

Masulli F, Valentini G, 2000. Comparing decomposition methods for classification. In Proc. International Conference on knowledge-Based Intelligent Engineering Systems and Applied Technologies (Kes 2000), Brighton, UK.

Rennie JDM, 2001. Improving Multi-class Text Classification with Naive Bayes. M.S Thesis, Massachusetts institute of technology, September 2001.

Wechsler H, Phillips P, Bruce V, Soulie F, Huang T, Face Recognition: From Theory to Applications, Springer-Verlag.

Yao J, Antani S, Long R, Thoma G, Zhang Z, 2006. Automatic Medical Image Annotation and Retrieval Using SECC, Proceedings of 19th International Symposium on Computer-Based Medical Systems (CBMS 2006), Salt Lake City, Utah, June 2006