A class-modular GLVQ ensemble with outlier learning for handwritten digit recognition

Katsuhiko TAKAHASHI and Daisuke NISHIWAKI
Multimedia Research Laboratories, NEC Corporation
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki, Kanagawa 216-8555, Japan
{katuhiko,daisuke}@ccm.cl.nec.co.jp

Abstract

A class-modular generalized learning vector quantization (GLVQ) ensemble method with outlier learning for handwritten digit recognition is proposed. A GLVQ classifier is one of discriminative methods. Though discriminative classifiers have remarkable ability to solve character recognition problems, they are poor at outlier resistance. To overcome this problem, a GLVQ classifier trained with both digit images and outlier images is introduced. Moreover, the original 10-classification problem is separated into ten 2-classification problems using ten GLVQ classifiers, each of which recognizes its corresponding digit class. Experimental results of handwritten digit recognition and outlier rejection reveal that our method is far more superior at outlier resistance than a conventional GLVQ classifier, while maintaining its digit recognition performance.

1 Introduction

It is very important for optical character reader (OCR) systems to reduce recognition errors because it is very expensive to employ operators to correct misclassified characters. Moreover, errors lead to user distrust and dissatisfaction. Therefore, uncertain recognition results should be rejected.

Earlier models of Japanese OCRs read characters written in separated character frames, one by one. These OCRs mainly rejected results corresponding to ambiguous input images. However, the latest OCRs, intending to ease some writing style limitations, have to reject other types of input images. In handwritten string recognition, mis-segmented character images generated due to the ambiguity of each character boundary, should be rejected. Input images with additional lines for cancellation and images with noise due to the incompleteness of the line removal process should also be rejected.

These cause a problem for discriminative character recognition methods, e.g. multilayer neural networks and LVQs. Though discriminative methods are efficient for character classification problems when each test image is reasonably assumed to correspond to one object class, they are very poor at rejecting garbage images that correspond to no object class, like mis-segmented or cancelled images. In this paper, we call this garbage "outliers."

There is little research on rejecting outliers [1]-[5], and some of it is on discriminative methods [1]-[3]. The basic idea is to train classifiers with both character images and outlier images. Bromley et al. [1] proposed a digit recognition method using a feedforward backpropagation networks trained with mis-segmented digit images. The network is trained in such a way to keep low output values for outlier images. Liu et al. [2] proposed a new MCE criterion which consists of the misclassification measure and the false rejection measure. Classifiers are trained based on the criterion of minimizing empirical loss for whole training samples (both digit images and mis-segmented images). Their methods have common ground in that they use classification functions only for object classes (digits) and train classifiers to output low values for outlier input. Kim et al. [3] proposed a method which rejects outliers with a pre-processor and a multilayer neural network that utilizes a discriminant function corresponding to an outlier class.

Though these methods improved the rejection accuracy on their test data, which consisted of a limited variety of outlier images, either mis-segmented digit images or cancelled images, problems remain. In practical OCR use, the prior probability of an outlier class, or the probability with which outlier images are input into a character classifier, is usually unknown. Therefore, classifiers which recognize character images as correctly as conventional methods and reject outlier images more precisely than conventional methods are desired. The above three methods, however, mainly focus on outlier resistance and have some drawbacks with respect to maintaining the high digit recognition accuracy that conventional methods have. Since the

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 IEEE

11-classification problem is more complicated than the conventional 10-classification problem, adding the 11th outlier class lowers the digit recognition accuracy. A rejection rule based on a top candidate's confidence value [2] is not preferred to one based on the difference between two top candidate's confidence values in digit recognition.

We focus on the handwritten digit recognition problem and propose a class-modular generalized LVQ (GLVQ) ensemble method for rejecting outliers more accurately than conventional methods, while maintaining high digit recognition performance. A GLVQ classifier is a discriminative method, its efficiency in solving character classification problem having previously been shown [6]-[8]. To avoid making the problem more complicated by adding an outlier class to a conventional 10-classification problem, we decompose the 10-classification problem into ten 2-classification problems and utilize ten GLVQ classifiers, each of which discriminates its corresponding digit class from the other classes (both 9 digit classes and outlier class). Then ten outputs from each classifier are combined in a combiner to derive a final recognition result.

Experimental results using handwritten digits and outlier image databases showed that outlier error using our method is less than half that of a conventional GLVQ method without outlier learning, while our method matches the conventional GLVQ's digit recognition performance.

Section 2 briefly outlines the class-modular recognition concept. In section 3, the classification and learning steps of the class-modular GLVQ ensemble are presented. In section 4, the experimental results of digit recognition and outlier rejection are evaluated, and performance by our method is compared to that of a conventional GLVQ.

2 Class-modular recognition

The framework of class-modular character recognition is common in methods that represent class distribution in feature space. For example, this framework is used in subspace methods, k-nearest neighbour, maharanobis-distance-based methods and modified quadratic discriminant functions. Figure 1 shows the architecture of a class-modular recognizer. Each classifier, $M_c, c = 1, 2, ..., k$, is trained independently and is responsible only for its corresponding digit class, ω_c. The k denotes the number of the object class. In digit recognition, k is equal to 10. The combiner generates a final decision, Z, based on the classifier's output, z_cs. Character recognition methods that represent the distribution of each character class are superior to discriminative methods in outlier resistance, even if they are trained without outlier data [9]. However, they tend to misclassify input images which are located near the ideal boundary of character classes in feature space. This is a serious problem for these methods.

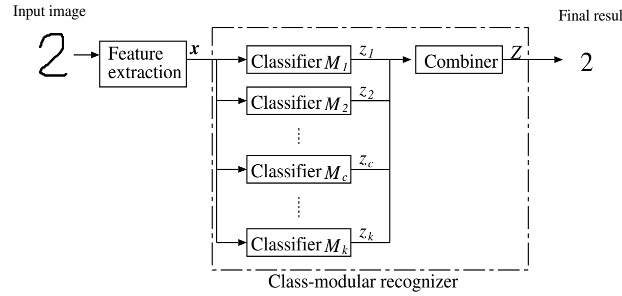

Figure 1. Architecture of a class-modular recognizer

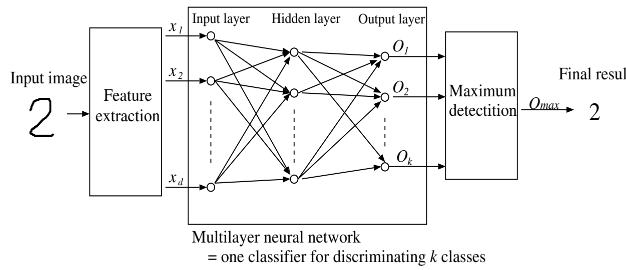

Figure 2. Architecture of conventional multilayer neural network

Oh et al. proposed a class-modular feedforward neural network architecture [10]. Conventional neural networks are classified as discriminative methods and usually recognize characters using a single network, as shown in figure 2. Since conventional neural networks did not work well for large class problems, they decomposed a k-classification problem into k 2-classification problems and trained each neural classifier to distinguish a specific class from the other $k - 1$ classes. Though this method is superior to a conventional neural network in character recognition, we suppose that it does not have high outlier rejection performance because a neural network without outlier learning are poor at outlier resistance.

3 Class-modular GLVQ ensemble with outlier learning

To achieve excellent digit recognition and outlier rejection ability, a discriminative method trained with outliers without an increase in the problem's complexity is necessary. In this section, we present a method consisting of multiple GLVQs, each of which is trained with outlier patterns and is responsible for discriminating a specific digit class from the other character classes and outlier class. In other words, denoting outlier class ω_0, a GLVQ classi-

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 IEEE

fier, M_c, is trained to discriminate the two meta classes, Ω_{c0} from Ω_{c1}, where $\Omega_{c0} = \{\omega_c\}$ and $\Omega_{c1} = \{\omega_i | i = 1, 2, ...c-1, c+1, .., k, 0\}$. The combiner decides the final recognition results based on the k confidence values, z_c, output by each classifier. The classification and training phases are described below.

3.1 Classification phase

Each GLVQ classifier holds reference patterns corresponding to two meta classes and calculates Euclidean distances between the input pattern and reference patterns. Then, classifier M_c outputs a confidence value:

$$z_c = \frac{(d_{c0})^{-2}}{(d_{c0})^{-2} + (d_{c1})^{-2}}, \tag{1}$$

where d_{c0} and d_{c1} are the distances between input pattern $\boldsymbol{x}$ and its nearest reference pattern belonging to meta classes Ω_{c0} and Ω_{c1}, respectively. The confidence value is $0 \leq z_c \leq 1$ and takes 0.5 when d_{c0} equals d_{c1}. Additionally, z_c tends to exceed 0.5, if the class of the input pattern is ω_c.

Since an ideal output for an input pattern belonging to ω_c is one in which only z_c exceeds 0.5, the combiner generates a final output based on the following rule:

$$Z = \begin{cases} \omega_{1st} & if(z_{1st} > \theta_a \ \& \ \theta_b > z_{2nd}), \\ \omega_0 & otherwise \end{cases} \tag{2}$$

where z_{1st} and z_{2nd} are the smallest and the second smallest confidence values among the z_c. The θ_a and θ_b are threshold values, and $\theta_a > 0.5 > \theta_b$.

3.2 Training phase

The k GLVQ classifiers are trained independently. Though each classifier is trained using a conventional GLVQ learning algorithm, the target values and the way training samples are learned are slightly modified.

Target values for training classifier M_c are redefined as either Ω_{c0} or Ω_{c1}. A GLVQ classifier is trained to classify these two meta classes. Since this redefinition of target values is done by a GLVQ learning program, training samples with redefined target values do not have to be stored on a hard drive.

Training samples belonging to Ω_{c0} and Ω_{c1} are input alternatively into a GLVQ learning program, and their number is balanced. To do this, each training sample belonging to Ω_{c0} is trained more than nine times as often as one of Ω_{c1}.

4 Experiments of handwritten digit recognition and outlier rejection

To confirm the efficiency of our method, experiments on handwritten digit recognition and outlier rejection were performed.

4.1 Experimental procedure

Two experiments were performed. In our first experiment, error vs. reject characteristics on handwritten digit recognition were examined using a digit test set (TEST SET (A)), while changing thresholds (rejection parameters). The digit test set consisted of about 400,000 digit images and no outlier images.

In our second experiment, rejection performance was examined using an outlier test set (TEST SET (B)), while changing thresholds. The outlier test set consisted of about 15,000 outlier samples, which included mis-segmented digit images (vertically or horizontally) and digit images with cancelled lines (shown in figure 3). They are artificial images generated by combining two digit images or a digit image and a line image for cancellation. The relative location and size of the two images were set randomly. Pairs of unconnected digits were also used because character segmentation process for numeral string recognition sometimes generates such patterns. These patterns had large variety compared to outlier images used in previous research [1, 2, 4, 5, 11]. The result was evaluated based on a graph of trade-off between the rejection rate of digit samples (obtained in the first experiment) and the error rate of outlier samples.

Note, in the second experiment, an outlier sample correctly rejected was counted as 'reject' (not 'correct'). Therefore, ideal recognition performance for experiment 2 is correct 0%, reject 100%, and error 0%.

For comparison, four recognition methods including the proposed one (d) were used in each experiment, as listed below.

- method (a)...Conventional GLVQ (trained with digit samples)
- method (b)...Conventional GLVQ (trained with both digit and outlier samples, outlier class was trained as eleventh class)
- method (c)...Class-modular GLVQ ensemble (trained with digit samples)
- method (d)...Class-modular GLVQ ensemble (trained with both digit and outlier samples, proposed)

The training samples used for methods (a) and (c) consisted of about 300,000 digit samples. The training samples used for methods (b) and (d) consisted of about 300,000 digit and 30,000 outlier samples. These training sets were different from the learning sets. 400 dimensional weighted direction histogram feature [12] was used.

The final result Z of conventional GLVQ (methods (a) and (b)) was obtained based on the following rejection rule:

$$Z = \begin{cases} \omega_{1st} & if\ (\frac{d_{2nd} - d_{1st}}{d_{1st} + d_{2nd}} > \theta_c), \\ \omega_0 & otherwise \end{cases} \tag{3}$$

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 IEEE

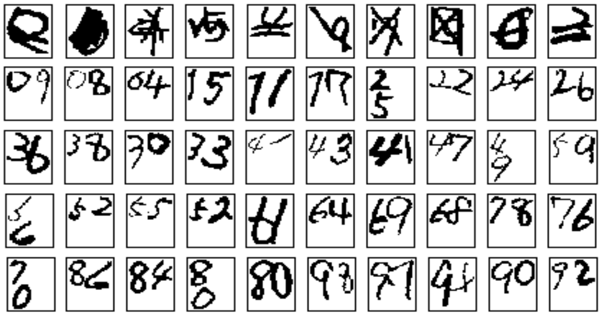

Figure 3. Example of outlier images

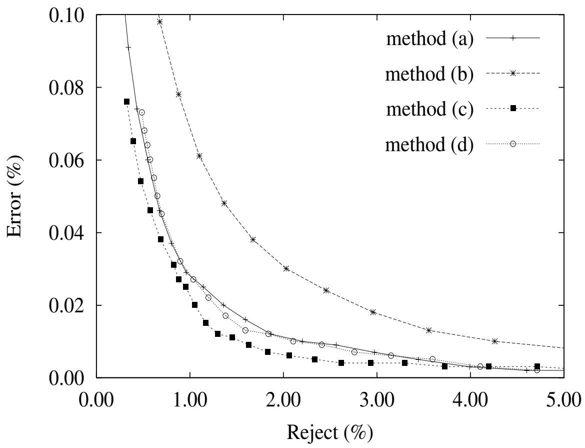

Figure 4. Error vs. reject characteristics on digit recognition (TEST SET (A))

where d_{1st} and d_{2nd} are the distance of the top two classes, and θ_c is the threshold value. When GLVQ classifier classified the input pattern as an outlier class, it was also rejected. The final result Z of methods (c) and (d) was obtained using rule (1).

4.2 Experimental results for digit recognition

Figure 4 shows the error-reject rate characteristics for digit images. Note that the proposed method (d) performed as well as the conventional GLVQ (method (a)). Additionally, class-modular GLVQ ensemble methods (methods (a) and (b)) was superior to conventional GLVQ methods (methods (c) and (d)), provided the training set was the same. Most notably, the performance of method (d) exceeded that of method (b), while the difference in performance between methods (a) and (c) was slight.

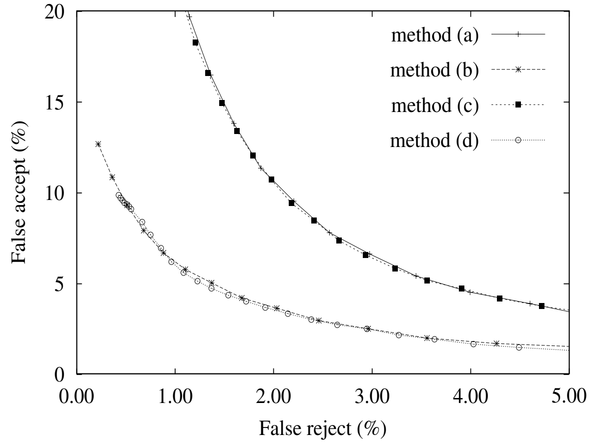

Figure 5. False-rejection vs. false acceptance ratio on outlier rejection (TEST SET (B))

4.3 Experimental results for outlier rejection

The experimental results for outlier rejection is shown in figure 5. Note that the horizontal axis represents the rejection rate of digit samples and the vertical axis the error rate of outlier samples. This figure shows that the error rates of methods (b) and (d), which were trained with outlier samples, were less than half that of methods (a) and (c). A comparison of figure 4 with 5 shows that the error rates for outliers were 10 to 100 times that for digit images. Figure 6 shows the outlier samples which were rejected correctly by method (d) but mis-recognized by method (a). This figure shows method (a) tends to judge input patterns (outliers) as digits.

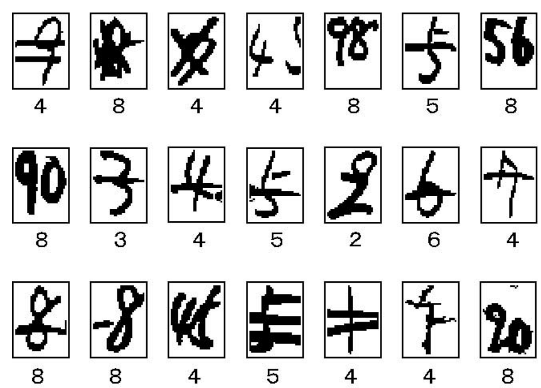

Figure 6. Outlier samples rejected by method (d). Digit shown under each sample pattern indicates the recognition result by method (a).

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**

4.4 Comprehensive evaluation

In the first experiment, methods (a), (c), and (d) performed well on digit recognition. But in the second experiment, the performance of methods (a) and (c) was obviously worse than that of method (d). As mentioned before, the prior probability for outlier samples is usually unknown when using practical OCR applications. Therefore, a method whose performance is excellent on both digit samples and outliers is desirable. From this point of view, we conclude that the proposed method (d), which performed well in both experiments, is the best among the four methods.

5 Conclusion

In this paper, a class-modular GLVQ ensemble with outlier learning was proposed. Though discriminative methods are poor at rejecting outliers, our method improves outlier resistance while maintaining the performance on digit samples. Experimental results confirmed that the proposed method works better than a conventional GLVQ. We are planning to apply our method to the digit string recognition problem, in which outlier images are processed frequently.

References

[1] J. Bromley, J. S. Denker, "Improving rejection performance on handwritten digits by training with "Rubbish"", Neural Computation, Vol. 5, pp. 367-370 (1993).

[2] C.-L. Liu, H. Sako, "Discriminative learning to improve the outlier resistance of neural classifiers", Technical report of IEICE, PRMU 2001-99, pp. 1-8 (2001).

[3] S. Kim, S. You, Y. Choi, "Rejection of garbage patterns for improving reliability in continuous handwritten numeral recognition", Methodologies for the conception, design and application of intelligent systems, Vol. 2, pp. 778-781 (1996).

[4] T. Kato, S. Omachi, H. Aso, "Unsupervised learning for single hypothesis problem via mixtures of modified quadratic discriminant functions", Technical report of IEICE, PRMU 2001-103, pp. 31-38 (2001) (in Japanese).

[5] S. Naoi, "Toward personal optival character reader form enterprise special use", Technical report of IEICE, PRMU 97-249, pp. 63-72 (1998) (in Japanese).

[6] A. Sato, K. Yamada, "Generalized learning vector quantization", Advances in neural information processing systems 8, MIT press, Cambridge, MA, pp. 423-429 (1996).

[7] C.-L. Liu, M. Nakagawa, "Evaluation of prototype learning algorithms for nearest-neighbor classifier in application to handwritten character recognition", Pattern Recognition, Vol. 34, pp. 601-615 (2001).

[8] A. Sato, "A learning method for definite canonicalization based on minimum classification error", Proc. of ICPR 2000, pp. 199-202 (2000).

[9] C.-L. Liu, H. Sako, H. Fujisawa, "Performance evaluation of pattern classifiers for handwritten character recognition", International journal on documnent analysis and recognition, Vol. 4, No. 3,pp. 191-204 (2002).

[10] I.-S. Oh, C.Y. Suen, "A class-modular feedforward neural network for handwriting recognition", Pattern Recognition, Vol. 35, pp. 229-244 (2002).

[11] C.Y. Suen, K. Liu, N.W. Strathy, "Sorting and recognizing cheques and financial documents", Proc. of third IAPR workshop on document analysis systems, pp. 1-18 (1998).

[12] T. Wakabayashi, S. Tsuruoka, F. Kimura, Y. Miyake, "Accuracy improvement through increased feature size in handwritten numeral recognition", Transactions of IEICE(D-II), Vol.J77-D-II, No.10, pp. 2046-2053 (1994)(in Japanese).

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**