Detection of ambiguous patterns in a SOM based recognition system: application to handwritten numeral classification

Leticia M. Seijas and Enrique C. Segura
Departamento de Computación, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires
BUENOS AIRES, ARGENTINA
{lseijas, esegura}@dc.uba.ar

Keywords: self-organising maps, pattern recognition, bayesian statistics

Abstract— This work presents a system for pattern recognition that combines a self-organising unsupervised technique (via a Kohonen-type SOM) with a bayesian strategy in order to classify input patterns from a given probability distribution and, at the same time, detect ambiguous cases and explain answers. We apply the system to the recognition of handwritten digits. This proposal is intended as an improvement of a model previously introduced by our group, consisting basically of a hybrid unsupervised, self-organising model, followed by a supervised stage. Experiments were carried out on the handwritten digit database of Concordia University, which is generally accepted as one of the standards in most of the literature in the field.

1 Introduction

Visual interpretation of scenes (the process of vision), while effortless for humans and animals, represents one of the greatest challenges to machine intelligence. Indeed, no automated vision system has so far been obtained whose performance can be compared even to that of the simplest animals. At the same time, relatively little is known about how biological vision operates. Great part of the difficulty in understanding and realizing vision derives from the lack of general concepts and approaches to represent, characterize and model the intricate data and processes underlying visual processing.

Optical character recognition (OCR) is one of the most traditional topics in the context of Pattern Recognition and includes as a key issue the recognition of handwritten characters. One of the main difficulties lies in the fact that the intra-class variance is high, due to the different forms associated with the same pattern, because of the particular writing style of each individual. No mathematical model is presently available being capable of giving account of such pattern variations [1]. Many models have been proposed to deal with this problem, but none of them has succeeded in obtaining levels of response comparable to human ones.

The use of neural networks has provided good results in handwritten character recognition. Most of the existing literature on this matter applies classical methods for pattern recognition, such as feed-forward networks (multi-layer perceptrons) trained with the backpropagation algorithm [1][2][3][4][5]. This architecture has been acknowledged as a powerful tool for solving the problem of pattern classification, given its capacity to discriminate and to learn and represent implicit knowledge. Competitive results can also be obtained in handwritten character recognition by means of unsupervised learning techniques such as Kohonen self-organising maps (SOM), even combined with other techniques [6][7][8].

This work proposes a SOM based pattern recognition system with a probabilistic strategy in order to classify, detect ambiguous patterns and explain answers. We apply the system to the recognition of handwritten digits. Our experiments were carried out on the handwritten digit database of Concordia University, which is generally accepted as one of the standards in most of the literature in the field.

This work is organized as follows: in section 2 the recogniser structure is explained, and some examples based on our implementation are given. In section 3 we present implementation details and experimental results. Concluding remarks are presented in section 4.

2 Recognition System

The general structure of the system is depicted in Figure 1.

The first stage of the process consists of a pre-processing of input data in which relevant features are extracted from the patterns. This provides a more general and simple structure for the system, specifically oriented to classification rather than feature selection that is therefore independent from system architecture. The recogniser is composed of two levels. The first one is formed by a collection of two-dimensional and independent self-organising maps, each one specialised in a different feature extracted from the input pattern. The second level consists of an analysing module in charge of defining and explaining the output of the system. This module is integrated by the following elements: the table of reliability and two parameters adjustable while running the system. Each SOM in the first stage produces a response to an input pattern, as a judge who, only based on the analysis of the corresponding fea-

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

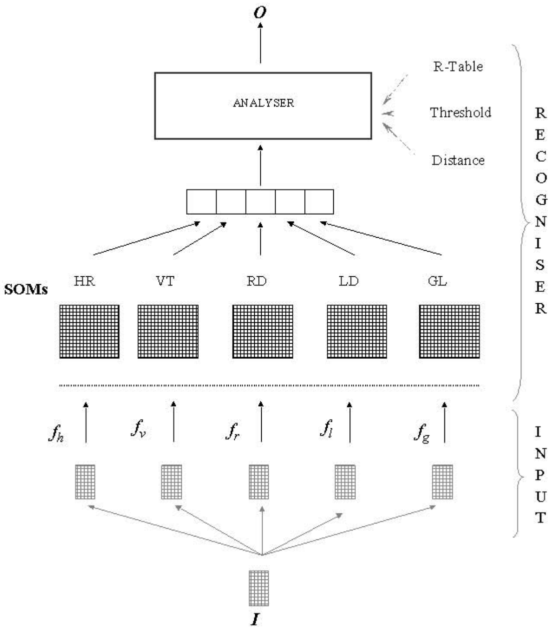


Figure 1: Architecture of the proposed recognition system. Each SOM is associated with a feature. In the analyser: table of reliability (R-Table), threshold of reliability and minimal distance parameter.

ture, decides which class the pattern belongs to. The connection between the first and second layers of the system is performed through this new representation of the pattern, formed with the answers of the "judges". The purpose of the table of reliability is to represent how trusty the "vote" of each SOM is. Using these elements, the module of analysis of the second layer has to produce the final answer. The system is able to explain its responses, indicating which class is most similar to the input pattern respecting each particular feature, on the base of the vote of each map/judge and the weight assigned to each one. As a part of the explanation, if a pattern is ambiguous for the system, we can know which other digits it could be identified with (i.e. which classes it has more features in common). Another element taking part in the output of the system is the graphical representation of the distribution of the input patterns (those used to train the system), by means of the topological maps in the first stage. This similarity-based grouping or clustering is performed separately for each feature and also for the complete pattern, which permits interesting comparison besides evaluation of the incidence of each feature in the definition of patterns.

2.1 SOMs Level

The performance of a character recognition system strongly depends on how the features that represent each pattern are defined. In the first layer of the recognition system, each SOM is trained on the base of a certain feature previously defined in the pre-processing stage according to the problem to deal with. In the context of handwritten numeral recognition, local detection of line segments and global detection of line structures seem to be an adequate feature extraction method. Kirsch masks [9] have been used as directional feature extractors by several authors [10][7][11], as they allow local detection of line segments. We used these masks to extract four directional features from the set of patterns: horizontal, vertical, right diagonal, left diagonal. In addition, we also considered the complete (original) pattern, which we call global feature. Hence we defined five SOMs for the first layer of the recogniser, each one dedicated to a particular feature. Each SOM is trained independently according to the Kohonen learning algorithm [12][13][14]. In our implementation each SOM consists of a 30 x 30 neuron square array, associated with an input of 16 x 16 pixels (the size of the input patterns). These values have proven adequate; however, in the present model it is not necessary to use the same parameter values for all maps. In order to assign labels to the output units, each neuron was associated with the class for which it was more frequently activated in the training stage. Figure 2 depicts the trained map associated with global feature. Note cluster formation for each digit, taking account that the map was treated as a toroid for learning.

From this first stage, topological maps are obtained representing input data distribution on a two-dimensional space. These maps enable an analysis of the relation between each directional feature and global feature (complete pattern). Moreover, the responses/votes of these maps represent each input pattern for both the construction of the classifier and recognition stage.

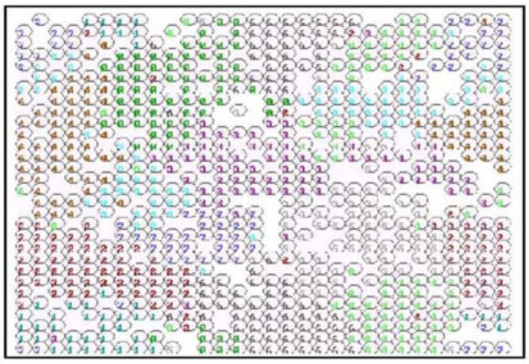


Figure 2: SOM associated with global feature. Each color represents a different class.

2.2 Analyser Level

Once the maps of the first layer have been trained, the second level of the recogniser for detection of ambiguous patterns and decision of the answer has to be constructed, be-

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

ing its parts: a reliability table and the parameters ”confidence threshold” and ”minimal distance”.

Upon analysis of different strategies for constructing the table that expresses how reliable the answer of each map is, a Bayesian probabilistic approach was chosen because of its better results. Using the definition of conditional probability and the multiplication rule, we define the probability of pattern p belonging to class C given that feature map f has responded C for input p as:

$$P(\{p, C\}/\{f, p, C\}) = \frac{P(\{f, p, C\}/\{p, C\})P(\{p, C\})}{P(\{f, p, C\})} \quad (1)$$

where
$P(\{p, C\})$: probability of input pattern p belonging to class C, estimated from the labeled training set.
$P(\{f, p, C\})$: This probability is estimated from trained feature map f. We assume that the response C of a SOM f given an input pattern p is independent from responses of other SOMs.
$P(\{f, p, C\}/\{p, C\})$: this probability is estimated from correct outputs of map f given input patterns of class C.

Table 1 shows the reliability table for the recogniser. Values in this table represent probabilities calculated with (1) for each feature map and each class.

Table 1: Reliability Table: values that represent how trusty a map response is for a given input. Maps are related to horizontal(HR), vertical(VT), right diagonal(RD), left diagonal(LD) and global(GL) features.

Class	HT	VT	RD	LD	GL
0	0.925	0.938	0.932	0.958	0.964
1	0.955	0.929	0.964	0.957	0.973
2	0.903	0.931	0.936	0.920	0.963
3	0.850	0.868	0.857	0.887	0.926
4	0.972	0.916	0.970	0.941	0.968
5	0.926	0.898	0.856	0.943	0.933
6	0.967	0.970	0.970	0.968	0.983
7	0.933	0.851	0.955	0.912	0.943
8	0.919	0.966	0.905	0.934	0.964
9	0.919	0.849	0.915	0.914	0.964

In the classification stage, once the input pattern has been represented by the k votes of the SOM’s, a score is computed for each voted class, on the base of the reliability of each map, according to the table. For this sake, such values are added for each class, so that a class with a greater score implies more reliable answers and more votes for the same class. In our previous works [15] [16] the values of the reliability table were calculated on the base of the relative error committed by each map for each class. The score s for each class was computed as:

$$s_C = \sum_{f \in F_C} r_{C,f} \quad (2)$$

where C indicates the selected class, f indicates feature map, F_C the feature maps that voted class C, $r_{C,f}$ reliability value taken from table for class C and feature f.

Some feature maps could be less reliable than others; however, all of them make their contribution that may be relevant for certain classes. This fact has been verified by computing separately the efficiency of each SOM and comparing it with the total performance of the system which has been better as regards both a correctly classify percentage of patterns and its remarkable properties.

One of the main difficulties for classification is dealing with outliers, ”ambiguous patterns”, since the distortions they exhibit make difficult their correct classification (being far away from the mean value of its class, they could be incorrectly associated with another class closer in average). In this work we have considered the distance from the pattern represented by its feature vector to the mean value (or centroid) of the class assigned by the map: if the pattern is close to that mean value, we assume that it is well defined and belongs to that class. A pattern far away from the centroid might be consider as an outlier (according to the variance of the class) and hence a candidate to ”ambiguous pattern” for the output. This information is used as a reinforcement factor for the score given by the table. Each probability is divided by the normalized distance between the pattern and the mean value; thus, nearby patterns will increase the score and those far away will not be so favored. Then, the score s assigned to each voted class C is calculated as

$$s_C = \sum_{f \in F_C} r_{C,f} \frac{1}{d(p_f, \mu_{f,C})} \quad (3)$$

where C indicates the selected class, f indicates feature map, F_C the feature maps that voted class C, $r_{C,f}$ reliability value taken from table for class C and feature f, p_f pattern represented by its feature vector associated with feature map f, $d(p_f, \mu_{f,C})$ normalized distance between the feature vector of the pattern p and the mean value of class C for feature f.

As an example, the following vector V shows votes for HR, VT, RD, LD and GL map respectively for a given test input: V = (0 5 0 2 2). Score associated with each voted class is computed using (3). In this case, scores are 1.81, 2.24 and 0.99 for classes ”0”, ”2” and ”5” respectively. Class ”2” obtains the higher score.

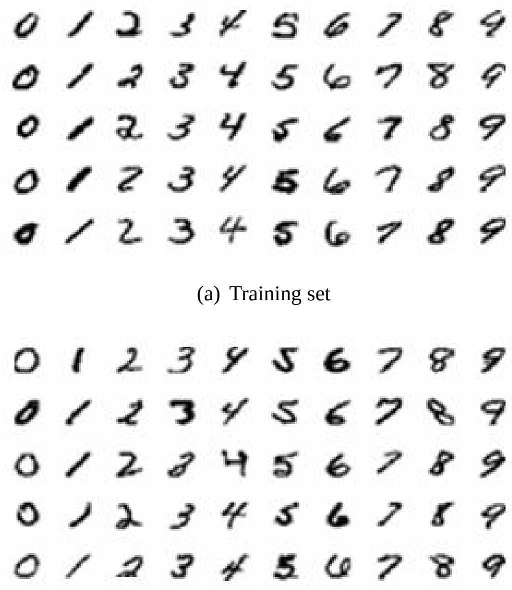
Once the score has been computed for each voted class for a given input pattern, the output of the system has to be defined. In order to this, the class with the higher score is identified; this score is compared with the reliability threshold that determines which patterns are considered as am-

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

biguous and which are not. If the total score for the winning class surpasses the threshold, then the system considers that pattern as well defined and the answer is that class. On the other hand, if the cumulative score is lower than the threshold, the system decides that the pattern is ambiguous; this can be due simultaneously to several causes: the fact of having few votes per class and many candidate classes; the votes of certain feature maps not being reliable enough for the produced answer; a pattern far away from the mean value of the proposed class. When an ambiguous pattern is detected (i.e. one not well defined), it is necessary to determine the class it might be confused with. From all voted classes, the one closest to the winner is selected if distance between it and the winning class is lower than minimal distance parameter.

Values for the threshold of reliability and for the minimal distance are chosen empirically, on the base of information provided by the training set in the stage of adjustment of the classifier. Variation of these parameters permits to adjust the output of the system without need of a new training of the SOM's.

Thus the output of the system detects and reports which patterns are ambiguous, which class they would belong to according to the classification and which class they could be confused with.



(a) Training set

(b) Test set

Figure 3: Handwritten digits from CENPARMI database, normalized in size.

3 Experiments

3.1 The Data Set

Our experiments were performed on the handwritten numeral database from the Centre for Pattern Recognition and Machine Intelligence at Concordia University (CENPARMI), Canada. This database contains 6000 unconstrained handwritten numerals originally collected from dead letter envelopes by the U.S. Postal Service at different locations in the United States. The numerals in the database were digitized in bilevel on a 64 x 224 grid of 0.153 mm square elements, given a resolution of approximately 166 ppi. The digits taken from the database present many different writing styles as well as different sizes and stroke widths. Some of the numerals are very difficult to recognize even with human eyes. Since the data set was prepared by thorough preprocessing, each digit is scaled to fit in a 16 x 16 bounding box such that the aspect ratio of the image is preserved. Then we apply Kirsch masks on each image, as mentioned in Section 2.1. The training and test sets contain 4000 and 2000 numerals from the database (400 / 200 by digit) respectively. Figure 3 shows samples from both sets.

The CENPARMI database is widely accepted as a standard benchmark to test and compare performances of the methods of pattern recognition and classification.

3.2 Recognition Results

We have implemented a pattern recognition system for classification of handwritten characters using the database described in Section 3.1. The general structure of the system and the SOMs used were described in Section 2. The table of reliability (see Table 1) was constructed on the base of (1) and the strategy used to compute the scores associated with each voted class was presented in (3). This strategy permitted improvement of the percentage of patterns correctly classified for the test set obtained in previous works [15] [16]. Present results for different values of reliability threshold and minimal distance are shown in Table 2.

Table 2: Recognition results (%) - RT: reliability threshold - MD: minimal distance parameter

RT	MD	Correct (includes ambiguous)	Correct (unique response)	Error
2.0	0.5	91.00	90.50	9.00
6.0	3.0	94.50	80.60	5.50
6.0	2.0	94.20	84.20	5.80
9.0	3.0	94.50	80.60	5.50
15.0	1.5	93.65	86.20	6.35

In the experiment for threshold 2.0 and distance 0.5, the value of threshold is low as compared to the scores obtained by the winning classes, hence there is practically no detec-

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

tion of ambiguous patterns.

As threshold increases, patterns associated with greater values of scores of winning classes will result well defined, and the rest will be considered as patterns with a certain degree of similarity with elements of other classes. Using the minimal distance enables to introduce a second class of output for these patterns, as long as such class has a score near enough to that which obtained the maximum. If that is not the case, the output for this not clearly defined pattern is unique. It can be noticed in the second and third rows in Table 2 that for the same threshold value the minimal distance has been decreased and that, as a consequence, the number of patterns associated with a unique class has increased ("Correct (unique response)" column).

Table 3: Some results of recognition over the testing set for reliability threshold 6.0 and minimal distance 3.0, with indication of which patterns were considered as ambiguous and the class voted by each feature map (HR - horizontal, VT- vertical, RD - right diagonal, LD - left diagonal and GL - global)

Class	Sys. Out.	Ambig	Vote HR	VT	RD	LD	GL
0	0	No	0	0	0	0	0
0	0 or 6	Yes	0	0	6	0	6
0	0 or 2	Yes	0	5	2	0	1
2	2	No	2	2	2	2	2
2	2 or 6	Yes	2	6	2	2	2
2	0 or 2	Yes	0	5	0	2	2
5	5	No	3	5	5	5	5
5	3 or 5	Yes	3	3	3	5	5
5	6 or 5	Yes	5	6	5	6	6

Figure 4: Test patterns that were correctly classified. Each row shows examples of classes "0", "2" and "5" respectively.

Table 3 shows some data forming the output of the system for each digit in Figure 4, grouped by class. Digits in the first column are well defined, i.e. they are not ambiguous for the system. The first row of each group in Table 3 shows that almost all maps voted for the same class; however, as explained in Section 2.2, in the definition of the output not only the number of votes takes part. The second and third columns in Figure 4 show patterns that the system considers as ambiguous. It can be observed in Table 3 that the output indicates two possible classes for the pattern, and one of them is the right one. The votes are distributed between different classes, hence the score of winning classes is lower than scores of well defined patterns. Visual analysis shows that the second "0" is in fact similar to a "6"; the third "0" is noisy; for the rest of the ambiguous patterns, forms can be observed that are not associated with a unique class. The second pattern labelled as a "5" might well be an incomplete "3"; and the third "5" fairly resembles an incomplete "6".

Table 4 shows error rates of different methods on CENPARMI database. The error rate obtained with our method is high, and further investigation should be made to improved it. On the other hand, our method presents remarkable properties that permit an analisys of the system response as mentioned above.

Table 4: Error rates of different methods on CENPARMI database

Method	Raw Error rate without rejection(%)
SVC-rbf [17]	1.10
Virtual SVM [18]	1.30
Local Learning Framework [19]	1.90
S. W. Lee [11]	2.20
S. B. Cho [20]	3.95
Proposed method	5.50

4 Conclusions

We have presented a system for pattern recognition that combines the use of Kohonen self-organising maps with a probabilistic bayesian approach. This proposal basically consists of a hybrid unsupervised self-organising model, followed by a supervised stage. Besides class identification, the system is able to detect ambiguous patterns and explain its answers. The classifier was applied to the recognition of handwritten digits. Experiments were carried out on the handwritten digit database from CENPARMI.

One of the most innovative aspects of the proposal is the way in which the unsupervised level interacts with the supervised one. The strategy used by the supervised level in order to define the system output and explain answers is the core of the interaction. The "reliability table" -estimating how reliable is the answer of each map for an input pattern- and a few parameters are used to decide when a pattern is considered as ambiguous, and which class or classes it might be confused with. Other remarkable feature is the bayesian strategy for constructing the table of reliability; in fact, this approach has produced the best results. Different criteria were also tried in order to define the winning class or classes, aimed to obtain a tradeoff between the number of votes per class and the quality of them. These criteria make use of information provided by the table, the distance of the pattern to the mean value of each class (the "centroid" of the class) and the number of votes per class.

Finally, our method presents remarkable properties that permit an analisys of the system response; the use of SOMs contributes to this task. However, further investigation should be made in order to improve error rate percentage.

References

[1] P. Wunsch and A. Laine,"Wavelet Descriptors for Multiresolution Recognition of Handprinted Characters", *Pattern Recognition*, Vol. 28, No. 8, pp. 1237-1249, 1995.

[2] Le Cun Y., B. Boser, J. S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition", *Neural Computation*, Vol. 1, pp. 541-551, 1989.

[3] A. Krzyzak, W. Dai and C.Y. Suen, "Unconstrained handwritten character classification using modified backpropagation model", Proceedings of 1st. International Workshop on Frontiers in Handwritten Recognition", Montreal, Canada, pp. 155-166, 1990.

[4] G. Y. Chen, T. D. Bui and A. Krzyzak,"Contour-Based Handwritten Numeral Recognition using Multiwavelets and Neural Networks", *Pattern Recognition*, Vol.36, No.7, pp.1597-1604, 2003.

[5] D. Romero, L. Seijas and A. Ruedin, "Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks", *JCST*, Vol. 7, No. 1, pp. 66-71, 2007.

[6] Z.-B. Liu,"The Training Strategy for Creating Decision Tree", Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xi'an, 2003.

[7] S.-B. Cho,"Ensemble of structure-adaptive self-organizing maps for high performance classification", *Information Sciences*, Vol. 123, pp. 103-114, 2000.

[8] M.-C. Su, E. Lai, C.-Y. Tew,"A SOM-based Fuzzy System and Its Application in Handwritten Digits Recognition", Proceedings International Symposium on Multimedia Software Engineering, Taipei, Taiwan, pp. 253-258, 2000.

[9] W. Pratt: *D*igital Image Processing, New York, Wiley, 1978.

[10] D. Gorgevik, D. Cakmakov, "An Efficient Three-Stage Classifier for Handwritten Digit Recognition", Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), Vol.4, pp. 507-510, 2004.

[11] S. Lee, "Multilayer Cluster Neural Network for Totally Unconstrained Handwritten Numeral Recognition", *Neural Networks*, Vol. 8, No. 5, pp. 783-792, 1995.

[12] T. Kohonen,"Self-Organized formation of topologically correct feature maps", *Biological Cybernetics*, Vol. 43, pp. 59-69, 1982.

[13] J. Hertz, A. Krogh, R. Palmer: *Introduction to the Theory of Neural Computation*, Santa Fe Institute Editorial Board, 1990.

[14] S. Haykin: *Neural Networks A Comprehensive Foundation*, Prentice Hall, 1999.

[15] L. M. Seijas and E. C. Segura, "A Hybrid Neural Network Model for Pattern Recognition: Solving Ambiguities and Explaining Answers", Proceedings ASIS 2005 - 34 JAIIO, Rosario, Argentina, 2005.

[16] L. M. Seijas and E. C. Segura, "Un clasificador neuronal que explica sus respuestas: aplicación al reconocimiento de dígitos manuscritos", Proceedings IX Congreso Argentino de Ciencias de la Computación (CACIC 2003), La Plata, Argentina, 2003.

[17] C. Liu, K. Nakashima, H. Sako and H. Fujisawa, "Handwritten Digit Recognition Using State-of-the-art Techniques," Proceedings of the eighth International Workshop on Frontiers in Handwriting Recognition, pp. 320–325, Niagara-on-the lake, Canada. August 6–8, 2002.

[18] J. X. Dong, A. Krzyzak, and C. Y. Suen, "Fast SVM Training Algorithm with Decomposition on Very Large Datasets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 4, pp. 603–618, April 2005.

[19] J. X. Dong, A. Krzyzak. and C. Y. Suen, "Local Learning Framework for Handwritten Character Recognition", *Engineering Applications of Artificial Intelligence*, Vol. 15, No. 2, pp. 151–159, April 2002.

[20] S. B. Cho and J. H. Kim, "Combining multiple neural network by fuzzy integral for robust classification," *IEEE Trans. Systems, Man and Cybernetics*, Vol. 25, No. 2, pp. 380–384, 1995.