2009 Fourth International Conference on Innovative Computing, Information and Control

A Study on Japanese Historical Character Recognition using Modular Neural Networks

Tadashi Horiuchi
Department of Control Engieering
Matsue College of Technology
14-4 Nishi-ikuma, Matsue 690-8518, JAPAN
Email: horiuchi@matsue-ct.ac.jp

Satoru Kato
Department of Information Engieering
Matsue College of Technology
14-4 Nishi-ikuma, Matsue 690-8518, JAPAN
Email: kato@matsue-ct.ac.jp

***Abstract*—It is fundamental work to translate the historical characters called "kuzushi-ji" into the contemporary characters in Japanese historical studies. In this paper, we develop the japanese historical character recognition system using the directional element features and modular neural networks. Modular neural networks consist of two kinds of classifiers: a rough-classifier to find the several candidates of categories for the input pattern, and a set of fine-classifiers that determine the category of the input pattern as the final result of character recognition. We construct the rough-classifier using the Self-Organizing Maps (SOM), which can derive the multi-templates for each category from input data. The fine-classifiers are realized using multi-layered neural networks, each of which solves the two-category classification problem. We also use the rough-classifier for the selection the training samples in the learning process of multi-layered neural networks in order to reduce the learning time. Through the experiments of historical character recognition for 57 character categories, we confirmed the effectiveness of our proposed method compared with the conventional research.**

I. Introduction

It is fundamental work to translate the historical characters called "kuzushi-ji" into the contemporary characters in Japanese historical studies. There are several studies on Japanese historical character recognition system to help researchers as well as general people to read such historical characters. However, most of these have been evaluated for small number of character categories.

In this paper, we develop the historical character recognition system using the directional element features and modular neural networks. Modular neural networks consist of two kinds of classifiers: a rough-classifier to find the several candidates of categories for the input pattern, and a set of fine-classifiers that determine the category of the input pattern as the final result of character recognition. We construct the rough-classifier using the Self-Organizing Maps (SOM), which can derive the multi-templates for each category from input data. In the experiment of historical character recognition for 15 character categories and 57 character categories, we confirmed that proposed rough-classifier can achieve higher classification accuracy in multi-template matching method.

The fine-classifiers are realized using multi-layered neural networks, each of which solves the two-category classification problem. We also use the rough-classifier for the selection the training samples in the learning process of multi-layered neural networks in order to reduce the learning time. Through the experiments of historical character recognition for 57 character categories, we confirmed that proposed modular neural network can reduce learning time drastically, keeping high classification accuracy.

II. Historical Character Recognition System

The character recognition system for japanese historical documents includes the following step as well as contemporary character recognition system, 1) input of image, 2) preprocessing, 3) feature vector extraction, 4) classification, 5) output of result.

A. Preprocessing

At first, a character image is normalized and smoothed in the preprocessing stage. In normalization, a linear normalization method is employed and an input image is adjusted to 128 128 dots. As a result of smoothing, bumps and holes of strokes are patched up by using 3 3 mask.

Then, contour extraction is done. If a white pixel adjoins a black pixel to the upward, downward, left, or right direction, the black pixel is regarded as on contour. The feature vector is extracted from the pixels of contour.

B. Feature extraction

In this section, the directional element feature (DEF), which was proposed for historical handwritten Kanji character recognition, is described. The operation for extracting the DEF includes the following steps.

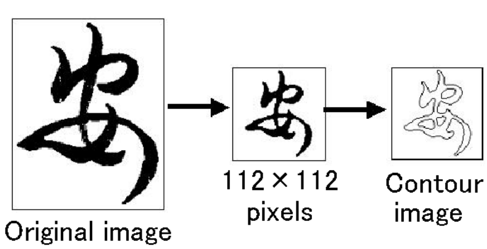
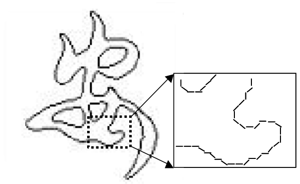
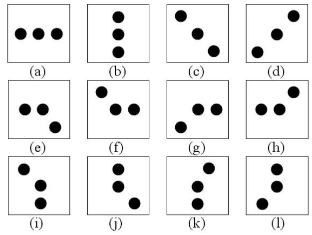


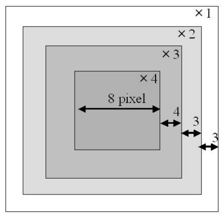
Fig. 1. Example of the preprocessing



(a) Extraction of the directional element features



(b) Line elements in the directional element features



(c) Weights in the directional element features

Fig. 2. Directional element features

1) Dot Orientation: In dot-orientation, four types of line elements, vertical, horizontal and two oblique lines slanted at 45 , are assigned to each black pixel. For a center black pixel in a 3 3 mask, two cases are considerable: One type of line element is assigned (see Fig. 2(b)-a to Fig. 2(b)-d); or if three black pixels are connected as in Fig. 2(b)-e to Fig. 2(b)-l, two types of line elements are assigned. For example, in the case of Fig. 2(b)-k, a 45 line element and a vertical line element are assigned simultaneously. Here, eight-neighbors are used to determine the direction of a black pixel.

2) Vector Construction: Consider an input pattern placed in a 128 128 mesh for which dot orientation has been completed. First, the 128 128 mesh is divided into 49, or 7 7 subareas of 28 28 pixels where each subarea overlaps eight pixels of the adjacent subareas (see Fig. 2(c)). Furthermore, each subarea is divided into four areas A, B, C, and D. A is a 8 8 area in the center. B is a 16 16 area exclusive of area A. C is a 22 22 area exclusive of areas A and B. D is a 28 28 area exclusive of areas A, B, and C. In order to reduce the negative effect caused by position variation of image, weighting factors are defined greater at the center of each subarea and decrease towards the edge. The weight of each area is 4, 3, 2, 1 for the areas A, B, C, D, respectively. For each subarea, a four-dimensional vector (x_1, x_2, x_3, x_4) is defined where x_1, x_2, x_3, x_4 represent the element quantities of the four orientations. Each element quantity is calculated by the following equation

$$x_j = 4x_j^{(A)} + 3x_j^{(B)} + 2x_j^{(C)} + x_j^{(D)} \quad j = 1, \cdots, 4 \qquad (1)$$

where $x_j^{(A)}, x_j^{(B)}, x_j^{(C)}, x_j^{(D)}$ denote the quantity of each element in A, B, C, and D, respectively. Since each subarea has four dimensions, the vector for one character is 196, or 49 4 dimensions. This vector is called directional element feature.

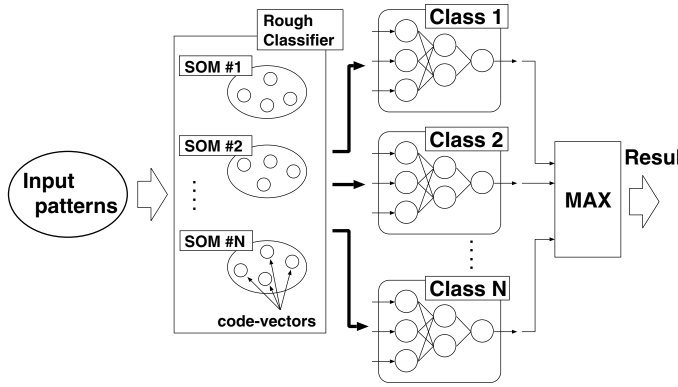


Fig. 3. Modular neural networks using SOM in rough classifier

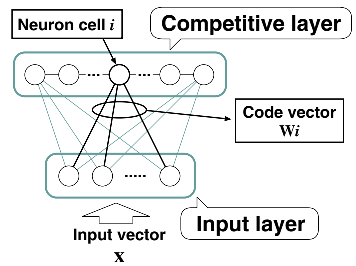


Fig. 4. Basic structure of SOM (one-dimensional SOM)

III. Modular Neural Networks Using SOM in Rough Classifier

The basic structure of our historical character recognition system is shown in Fig. 3. This consists of two kinds of classifiers: a rough-classifier to find the several candidates of categories for the input pattern, and a set of fine-classifiers that determine the category of the input pattern by multi-layered perceptrons (MLP).

The fine-classifiers are realized using a set of MLPs, each of which solves the two-category classification problem. In other words, each MLP is learned to output the value '1' only when the patterns belonging to the class for which the MLP is responsible, and otherwise to output the value '0'.

The rough-classifier is constructed using self-organizing maps (SOM), which can derive the multi-templates for each category from input data.

Basic structure of SOM is shown in Fig. 4. In the learning algorithm of SOM [6], the code-vector $\mathbf{w}$ for the winner cell which is nearest to the input vector $\mathbf{x}$ and its neighborhood cells are updated by the following equation

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t)\,\Phi(p_i)\,(\mathbf{x} - \mathbf{w}_i(t)) \qquad (2)$$

$$\Phi(p_i) = \exp\left(-\frac{{p_i}^2}{\sigma^2(t)}\right) \qquad (3)$$

where $\alpha(t)$ is learning coefficient after t learning steps. The coefficient starts from its initial value α_{ini} and then decreases monotonically as t increases, thus reaching its minimum value at the pre-set maximum number of learning steps T_{max}. $\Phi(p_i)$ is a neighborhood function with the center at winner cell

c and p_i is the distance from cell i to the winner cell c. In the equation (3), $\sigma(t)$ is a time-varying parameter that defines the neighborhood size in the competitive layer. As well as parameter $\alpha(t)$ in equation (2), this parameter decreases monotonically from σ_{ini} as t increases.

In general, SOM has significant characteristics that the distribution of code- vectors after the learning of SOM reflects the distribution of the input data. That is, code-vectors tend to converge on the area where the density of the input data is high. In this research, based on this characteristics of SOM, we construct the rough- classifier of our modular neural networks by assigning each SOM to each class category. Here, we use one-dimensional SOM for each SOM in the rough-classifier.

In the rough-classification, we select k templates which are nearest to the test sample among the all templates for all class categories. The class categories of the selected k templates result in the candidate classes in the fine-classifiers.

IV. Experiment 1

In order to verify the performance of the rough-classifier, we carried out the recognition experiments using historical character dataset HCD1 and HCD1ae which are open to be public by Historical Character Recognition Project [2].

In our recognition experiments, we use 15 character categories from HCD1 and 42 character categories from HCD1a e, which satisfies the condition that 200 character samples exist for each character category. Examples of character images in HCD1 are shown in Fig. 5. And list of 42 character categories in HCD1ae is shown in Table I.

In the experiment of historical character recognition, we obtained the recognition accuracy by calculating the average of 10 sets of 2-fold cross-validation. We changed the number of cells (the number of templates) in SOM from 5 to 20. Other learning parameters in SOM are decided as follows through the preliminary experiments: $\alpha_{ini} = 0.3$, $\sigma_{ini} = 1.0$, $T_{max} = 20,000$.

Fig.6(a) and Fig.6(b) indicate the recognition results for 15 character categories (HCD1) and 57 character categories (HCD1+HCD1ae). As comparison method, 1) select the mean vector as template (template matching method: 1-Template) and 2) randomly select 520 templates from learning samples (Random sampling) are also shown in the figures. From the figures, we can confirm that the proposed rough-classifier using SOM shows higher accuracy with small number of templates. Also we can see this tendency in case of 57 character categories compared with in case of 15 character categories.

V. Experiment 2

In this chapter, we carried out the character recognition, in order to confirm the effectiveness of the proposed modular neural networks with rough-classifier and a set of fine-classifiers.

In this experiment, we use the following parameter values: the number of cells in the competitive layer in SOM is 10, other learning parameters in SOM are same as previous section. We changed the number of candidates N in the rough-classifier for the selection of the training samples from 5 to 15. Each fine-classifier is realized by three layered perceptron with 10 units in the hidden layer. Therefore, each MLP has 196 units in the input layer, 10 units in the hidden layer and 1 unit in the output layer. We used batch learning method and set the number of learning iteration *Epoch*=100. We applied the conjugate gradient backpropagation with *Fletcher-Reeves update*. The backpropagation algorithm, which adjusts the weight vectors in the steepest descent direction, is a widely-used algorithm for multilayer perceptron. However, it turns out that, although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. In the conjugate gradient algorithms, a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions.

The parameter in the recognition after the learning, we set the value $k = 2$ (see Section III). This is decided through the preliminary experiments.

As same as the previous section, we obtained the recognition accuracy by calculating the average of 10 sets of 2-fold cross-validation. The programs for learning and recognition are implemented by using MATLAB. The computer used in the experiments has CPU of Core2Duo T7200 (2GHz) and Main Memory of 1Gbyte.

Fig.7 shows the relationship between classification accuracy and learning time in case of changing the number of candidates N in the rough-classifier for the selection of the training samples.

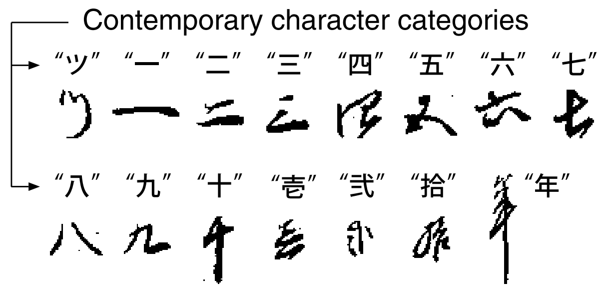
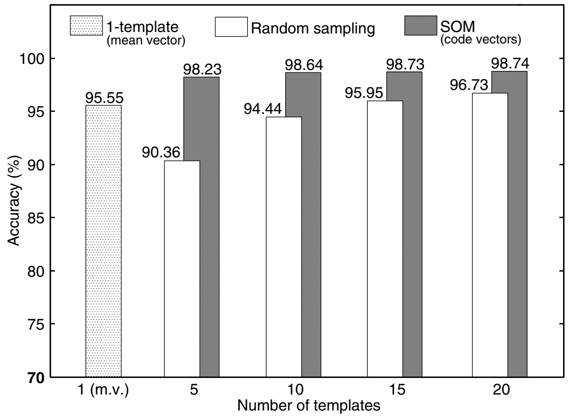


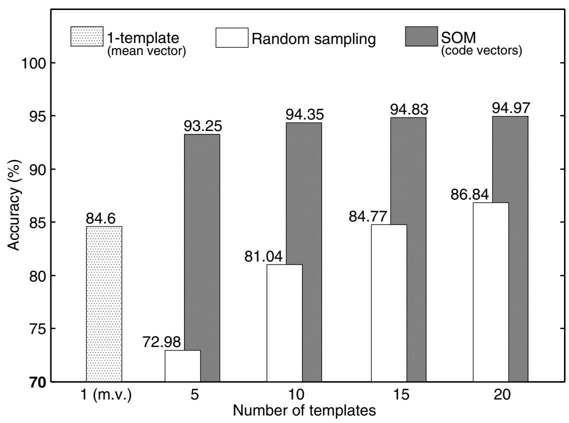
Fig. 5. Examples of character images used in the experiment

TABLE I
42 CHARACTER CATEGORIES USED IN THE EXPERIMENT (HCD1A1E)

No.	category	No.	category	No.	category	No.	category
16		17		18		19	
20		21		22		23	
24		25		26		27	
28		29		30		31	
32		33		34		35	
36		37		38		39	
40		41		42		43	
44		45		46		47	
48		49		50		51	
52		53		54		55	
56		57					



(a) In case of 15 character categories (HCD1)



(b) In case of 57 character categories (HCD1+HCD1a1e)

Fig. 6. Comparison of the experimental results of rough classifier (in case of 57 character categories)

When we select the training samples $(N = 5, 10, 15)$, we can obtain almost same recognition accuracy compared with the case without selection of the training samples ("Full (no selection)"). And we can see that great reduction of learning time is realized by the selection of the training samples.

From the above, we confirmed that we could achieve to reduce the learning time of modular neural networks without decrease of the recognition accuracy by the selection of the training samples using SOM in the rough-classifier.

VI. Conclusion

In this paper, we developed the japanese historical character recognition system using the directional element features and modular neural networks. Through the experiments of historical character recognition for 57 character categories, we confirmed the effectiveness of our proposed method compared with the conventional research. More concretely, the results of this paper are summarized as follows:

- We confirmed to obtain quite high recognition accuracy of about 95% for 57 character categories, which contain more categories than the conventional research.

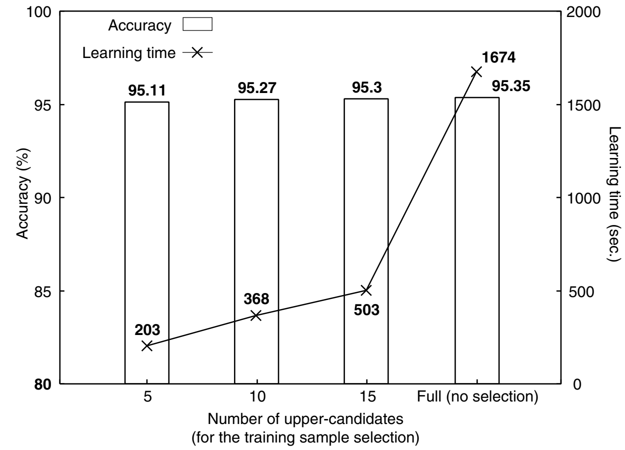


Fig. 7. Relationship between classification accuracy and learning time

- By multi-template learning using SOM in the rough-classifier, we realized to select the training samples in the learning process of multi-layered neural networks and to reduce the learning time without decrease of the recognition accuracy.

There are several future works including the comparison of recognition performance of support vector machines (SVM) and nearest-neighbors classifier with Mahalanobis distance and other classification methods. Moreover, we are going to apply the recognition methods to the reading support system for japanese historical documents.

Acknowledgment

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 20560389, 2008.

References

[1] S. Yamada and M. Shibayama, Outline of Historical Character Recognition Project, *Journal of IPSJ*, Vol.43, No.9, pp.950–955 (2002)

[2] Historical Character Recognition Project
http://www.nichibun.ac.jp/~shoji/hcr/index.html

[3] R.A. Jacobs and M.I. Jordan: Learning Peicewise Control Strategies in a Modular Neural Network Architecture, *IEEE Trans. on Systems, Man and Cybernetics*, Vol.23, pp.337–345 (1993)

[4] H.C. Fu, Y.P. Le, C.C Chiang and H.T. Pao, Diveide-and-Conquer Learning and Moduler Perceptron Networks, *IEEE Trans. on Neural Networks*, Vol.12, No.2, pp.250-263 (2001)

[5] K. Saruta, N. Kato, M. Abe and Y. Nemoto: High Accuracy Recognition of ETL9B Using Exclusive Learning Neural Network-II (ELNET-II), *IEICE Trans. Inf. and Syst.*, Vol.E79-D, No.5, pp.516–522, (1996)

[6] T. Kohonen: Self-Organizing Maps, Springer-Verlag (1995)

[7] N. Sun, M. Abe and Y. Nemoto, A Handwritten Character Recognition System by Using Improved Directional Element Feature and Subspace Method, *IEICE Trans.*, Vol.J78-D-II, No.6, pp.922–930 (1995)