ON-LINE RECOGNITION OF LIMITED-VOCABULARY CHINESE CHARACTER USING MULTIPLE CONVOLUTIONAL NEURAL NETWORKS

Quen-Zong Wu, Yann Le Cun**, Larry D. Jackel**, Bor-Shenn Jeng**

*Telecommunication Laboratories, Taiwan, R.O.C.
**AT&T Bell Laboratories, Holmdel, NJ 07733, U.S.A.

ABSTRACT

This paper presents a new feature extraction method together with neural network recognition for on-line Chinese characters. A Chinese character can be represented by a three-dimensional 12x12x4 array of numbers. Multiple convolutional neural networks are used for on-line small vocabulary Chinese character recognition based on this feature extraction method. We choose one hundred character classes as an example for recognition. Simulation result shows that 98.8% and 94.2% of training examples and test examples are correctly recognized respectively.

1. INTRODUCTION

The most popular input device for data processing machines is the keyboard. However, in many applications, such as the ones that involve Chinese text, keyboards are neither convenient, nor easy to use. Chinese is one of the most widespread languages in the world, yet no fully satisfactory text input device exist for it. A system that could recognize handwritten Chinese characters as they are written on "electronic paper" would be an ideal solution to the input problem.

Among the many methods that have been proposed for on-line handwritten character recognition [1, 2], two broad categories can be identified: memory-based techniques in which incoming characters are matched to a (usually large) dictionary of templates, and parameter-based methods in which preprocessed characters are sent to a trainable classifier such as a neural network [3].

Two characteristics differentiate Chinese character recognition from, say, Roman character recognition: the large number of classes, and the richness, and potentially complex internal structure of each character. Many papers [4, 5, 6] have addressed the problem of on-line Chinese character recognition. However, most of them are memory-based and make strong assumptions about the order in which the strokes are drawn.

A better long-term strategy would be to base the recognizer on *shape* information rather than dynamic information. In this paper, we introduce a small-vocabulary Chinese character recognizer which combines a shape-based preprocessor, and a neural network. The main advantage of this approach is its independence with respect to the writer, the writing speed, the stroke order, and the pen lifts.

2. PREPROCESSING

The preprocessing stage transforms the pen trajectory information into a three-dimensional 12x12x4 array of numbers using the following procedure:

(1) a rectangular grid of 12x12 points (called center points) is created so as to uniformly cover the bounding rectangle of the character. To each of these points, we associate a four-dimensional vector.

(2) Two connected adjacent points that are read by the digitizer constitute a microsegment. For each microsegment we

determine the center point which is closest to it. Components 0, 1, 2, 3 of the 4D vector associated with the center point are respectively incremented by an amount proportional to the projection of the microsegment onto the North-South, NE-SW, East-West, SE-NW directions, and weighted by a decreasing function of the distance from the microsegment to the center point.

The character is therefore represented by a 12x12 spatial map in which each location represents the density of microsegments around that location for each of the four main orientations.

3. THE NEURAL NETWORK

The classification is performed by a multiple convolutional neural network [7] trained with the backpropagation algorithm. The term "convolutional" means that some layers of the neural network are composed of several groups of units called feature maps. Units within a particular feature map are connected to a limited neighborhood in the previous layer, and are constrained to have identical weights, thereby performing the same operation on different parts of the input.

The four 12x12 feature maps are expanded to four 16x16 feature maps by adding blank boundary locations to the original feature maps. These four 16x16 feature maps are used as the input to the neural network. Each 16x16 feature map is convolutionally connected to two H1 14x14-unit groups. Every 2x2 units of all H1 groups are then locally connected to the corresponding units of both two H2 7x7-unit groups. H2 groups are then fully-connected to the 100 units in the output layer.

4. PERFORMANCE TESTING

A set of one hundred Chinese characters shown in Figure 1 was chosen to test the system. Seventeen writers were asked to write each character once. The characters produced by the first 12 writers were used for training the neural network, and the remaining 5 were used for testing. Figure 2 shows an example of unit activations in the neural network.

The performance data for training set and test set are listed in Table 1. These include the percentages that the scores of the desired class are within top 1, 2 and 3 highest scores. We also simulate the performance for recognition by using 1-layer fully connected neural network and 2-layer fully connected neural network. These performance data are listed in Table 2. Table 3 lists the comparison on numbers of weights for these three kinds of neural networks.

According to the above tables, we know that multiple convolutional neural network is superior to 1-layer fully connected network and 2-layer fully connected network.

Table 1. Performance for multiple convolutional N. N.

Set / Top	Training	Test
1	98.8%	94.2%
2	99.2%	98.4%
3	99.4%	98.8%

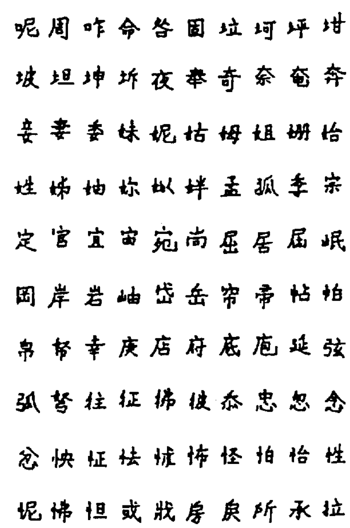

Figure 1: A subset of on-line handwritten Chinese characters.

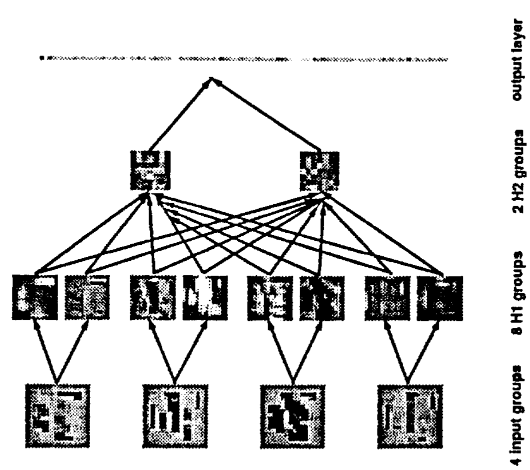

Figure 2: Neural network architecture.

Table 2. Performance for 1-layer fully connected N. N. and 2-layer fully connected N. N.

N. N. \ Set	Training	Test
1-layer fully connected	99.9%	80.6%
2-layer fully connected	99.9%	91.8%

Table 3. Comparison on numbers of weights

multiple convolutional N. N.	13214
1-layer fully connected N. N.	57700
2-layer fully connected N. N.	135500

5. CONCLUSION

This handwritten Chinese character recognizer can be used for limited-vocabulary applications. The high recognition rate for the training set shows that it is excellent for multiple writer applications. The recognition rate for test set depicts that it is also suitable for writer independent applications. In addition, this recognizer can be extended for large-vocabulary Chinese character recognition by introducing a preclassifier into it.

REFERENCES

[1] C. Y. Suen, M. Berthod and S. Mori, *Automatic Recognition of Handwritten Characters - the State of the Art, Proc. IEEE*, Vol. 68, pp. 469-487, 1980.

[2] J. Mantas, *Methodologies in Pattern Recognition and Image Analysis - a Brief Survey, Pattern Recognition*, Vol. 20, pp. 1-6, 1987.

[3] I. Guyon, P. Albrecht, Y. Le Cun, J. S. Denker and W. Hubbard, *Design of a Neural Network Character Recognizer for a Touch Terminal, Pattern Recognition*, Vol. 24, pp. 105-119, 1990.

[4] S. I. Hanaki and T. Yamazake, *On-Line Recognition of Handprinted Kanji Characters, Pattern Recognition*, Vol. 12, pp. 421-429, 1980.

[5] K. Ikeda et al., On-Line Recognition of Handwritten Characters Utilizing Positional and Stroke Vector Sequences, Pattern Recognition, Vol. 13, pp. 191-206, 1981.

[6] K. Odata, H. Arakawa and I. Masuda, *On-Line Recognition of Handwritten Characters by Approximating Each Stroke with Several Points, IEEE Trans. Syst., Man., Cybern.*, Vol. 12, pp. 898-903, 1982.

[7] Y. Le Cun, O. Matan, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, H. S. Baird, *Handwritten Zip Code Recognition with Multilayer Networks, Proc. of the International Conference on Pattern Recognition*, 1990.