International Symposium on Nonlinear Theory and its Applications
Xi'an, PRC, October 7-11, 2002

A Neural Network Model for Character Recognition

Hirohito Shintani*, Hirofumi Nagashino†, Masatake Akutagawa††and Yohsuke Kinouchi**

* **Faculty of Engineering, The Univ. of Tokushima
2-1, Minamijosanjima,
Tokushima, 770-8506, JAPAN
Phone: +81-88-656-7475
*Email: hsintani@ee.tokushima-u.ac.jp
** kinouchi@ee.tokushima-u.ac.jp

† ††School of health Sciences, The Univ. of Tokushima
3-18-15, Kuramoto,
Tokushima, 770-8506, JAPAN
† Phone: +81-88-633-9025
Email: nagasino@medsci.tokushima-u.ac.jp
†† Phone: +81-88-633-9057
Email: makutaga@medsci.tokushima-u.ac.jp

Abstract We have developed a small scale four-layered neural network (NN) model for simple character recognition, which can recognize the patterns transformed by affine conversion. It learns by backpropagation to obtain the characteristics of the simple cell and the complex cell in the visual cortex. In this study four characters are presented as input patterns. An input pattern is divided into 64 local patterns and connected with the 1st hidden layer as in the visual cortex.

1. Introduction

A lot of applications of multilayer feedforward NNs to pattern recognition and their effectiveness have been reported [1,2]. It has been known that backpropagation (BP) neural networks are a powerful means for that purpose [3]. In the problem of classification of input data into some categories such as pattern recognition, the ability of generalization of neural networks can be regarded as extraction of some features that are embedded in the input patterns without explicit presentation. A lot of NNs with the feature extraction mechanism have so far been studied. Our previous study shows that the method of extracting the features is effective in pattern recognition[4]. It shows the usefulness of extracting the features from an input pattern. However, it is a fault that competition type NN has many binding conditions. It determines the limit of recognition capability.

When we construct a pattern recognition system by modeling the structure of the visual cortex, the complexities of the visual cortex can be simplified by understanding that the neurons of this region are distinguished by the stimulus features that each neuron detects. The three major groups of so-called feature detectors in the visual cortex include simple cells, complex cells, and hypercomplex cells. This fact shows that the whole form is recognized by unifying the easy feature of every small domain.

We propose a NN model of the simple structure referring the function and structure of the visual cortex. In the present study, we have developed a small scale four-layered NN model for simple pattern recognition, which can recognize the character patterns transformed by affine conversion. The proposed NN model is composed of two parts. One is an ordinary NN as the division for discernment, and the other is a NN which extracts the feature for divided small parts of the input pattern. They are BP-type neural networks and fully connected. After the learning, we investigated the recognition capability of the proposed NN and the performance of the feature extraction.

In this paper, we describe the model for character recognition, show the results of the evaluation of this recognition capability of the NN, and discuss the recognition mechanism of the model.

2. Network structure and learning patterns

The NN that we investigated is a feedforward network that is composed of four layers, an input layer, two hidden layers, and an output layer (shown in Fig.1). It is trained by the error backpropagation learning. The input layer and the first hidden layer are divided into 64 local parts respectivery. Only the corresponding parts of the two lay-

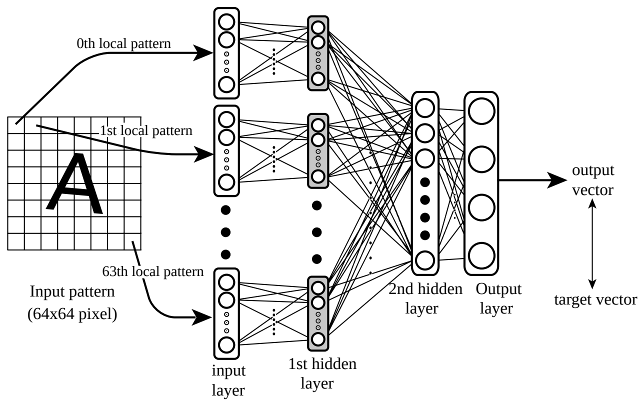

Fig.1 Strucuture of the proposed neural network

ers have connection. The number of units in the input layer is 4096(64x64). The number of units in the first hidden layer is 320(64x5). There is no connection between different parts. All the weights between 1st and 2nd hidden layer are connected. The number of units in the second hidden layer is 20. The network is presented the standard four character patterns (shown in Fig.2), and shifted, enlarged, shrinked, and rotated patterns from standard ones (shown in Fig.3) as training patterns. Rotation angle of learning patterns is in the range of 6-degrees. Scaling rate of the patterns is in the range of 80% to 120%. The amount shifted vertically or horizontally is 5 pixels. The number of the patterns generated from one standerd pattern is 175. Moreover, noise patterns of the range of 1 is also used for learning. The number of noise patterns is 400. The number of patterns used for learning is 1100 in total per iteration. The patterns are processed to 64x64 pixel images, and are divided into 64 local patterns. Each local pattern is 8x8 pixel image. The input to an input unit has the normalized intensity of a pixel in an image in the range of -1 to 1. The output function of input units is linear. The output function of hidden units and output units is a sigmoid function, which is defined as

$$f(x) = \tanh(x) \qquad (1)$$

The network is trained to classify characters, A, B, C, and D. As an input, a white square denotes a negative value (as -1) and a black one denotes a positive value (as 1). The number of units in the output layer is 4. Each unit corresponds to one of the character categories. The target value

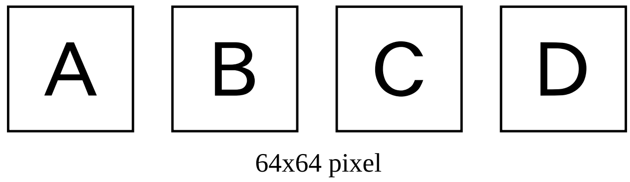

Fig.2 Standerd learning patterns

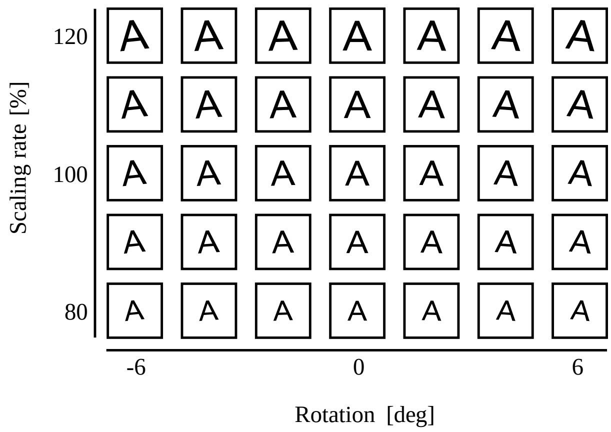

Fig.3 enlarged, shrinded, and rotated patterns

of the output unit is 1.0 for the category and –1.0 for other categories. All target values of the output units is –1.0 when a noise pattern is presented. A bias neuron is provided for each of the hidden layers and the output layer respectively.

3. Results

3.1 Recognition of test patterns

700 learning patterns and 400 noise patterns (Noise patterns are changed for each iteration.) were presented to the network. The reason of including the noise patterns in the learning patterns is to prevent the recognition of an unrelated form. The relationship between iteration and root mean square (rms) error to the learning patterns is shown in Fig.4. The dotted line shows the rms error calculated including the error when presenting the noise patterns. The solid line expresses the rms error which does not include the error for the noise patterns. After the training of 3000 iterations, the rms error of the output is less than 0.02. In the case of including noise patterns, it converged to about 0.2.

Test patterns which are transformed by different amount are presented. The network shows good performance in

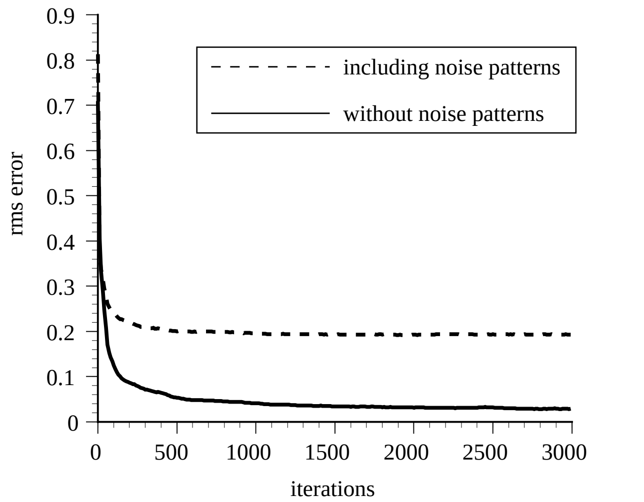

Fig.4 The relationship between iteration and rms error to learning patterns

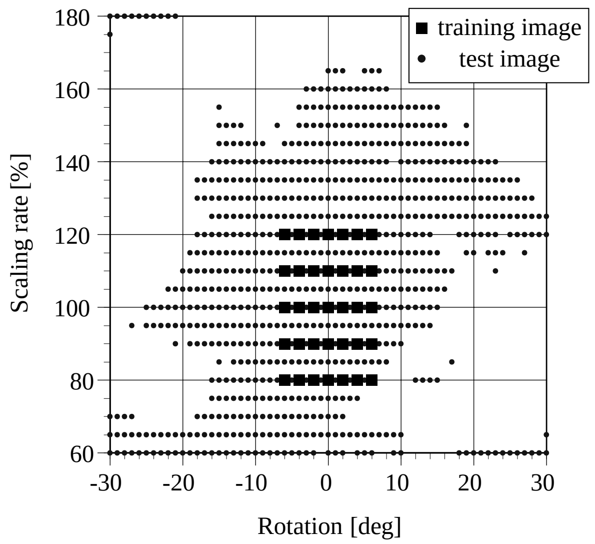

Fig.5 Distribution of recognizable test patterns

this test phase. The result of the test is shown in Fig.5. The input patterns are not shifted here. The training images are denoted by filled square, and the test images are denoted by filled circle. Recognizable patterns are distributed broadly, especially for the around the learning patterns. However, when shifted patterns are fed, recognition performance is not good.

3.2 Feature extraction

We analyzed the structure of the network after the learning was completed. We calculated the receptive fields of the 1st hidden layer units. They are 29th group of the divided input layer (shown in Fig.6). The receptive fields are represented by the coupling weight from the input layer units to the first hidden layer units. The reason for having chosen the 29th group is that the difference in the features has appeared in the 29th region when piling up all the input patterns (shown in Fig,7).

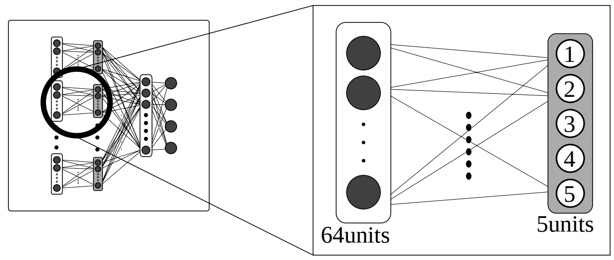

Fig.6 The portion of connections which are analyzed on feature extraction

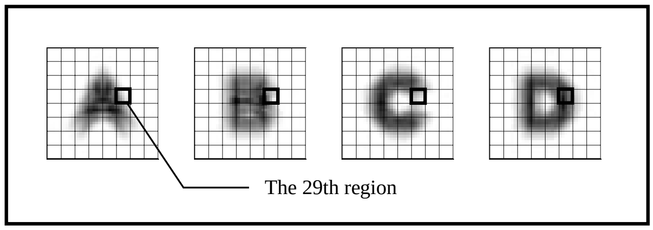

Fig.7 The 29th region

	unit1	unit2	unit3	unit4	unit5
unit1	1.000000	0.825022	0.461997	-0.003487	-0.309898
unit2	0.825022	1.000000	0.797546	0.330785	-0.161735
unit3	0.461997	0.797546	1.000000	0.742520	0.227095
unit4	-0.003487	0.330785	0.742520	1.000000	0.724712
unit5	-0.309898	-0.161735	0.227095	0.724712	1.000000

Table.1 cosθ of the weight vectors to the first hidden layer units

After the learning, we compared the receptive fields by cosθ which is defined as

$$\cos\theta_{ts} = \frac{w_t w_s}{|w_t||w_s|} \quad , \qquad (2)$$

where w_s and w_t are weight vectors to s-th unit and t-th unit respectively.

The result is shown in Table.1. The 29th group units are 5 pattern classifiers. Table.1 shows that cosθ with different t has large difference for the same s. From this result, it can be said that the units in the 1st hidden layers with the corresponding small region has classified input patterns. However, some other hidden units, for example, 0th units do not classify them.

It is suggested that the classification of patterns is performed in the local division. 1st hidden layer units extract the features of patterns.

4.Conclusions

We proposed four layer feedforward neural network model for character recognition and analyzed the performance of the model.

After the learning, test patterns transformed by affine conversion by different amount are presented. The network shows good performance in this test phase. The feature extraction part of the NN carries out the pattern classification from the divided parts of the input patterns. It was elucidated that by backpropagation learning the units in the first hidden layer acquire the capability to extract some features at a particular part of the input image as simple cells in the visual cortex.

The proposed NN model is of simple form. It has, however, good performance and the capability to extract the features of the input patterns like the visual cortex. The proposed model has potential as a general-purpose character recognition NN model.

References

[1] K. Fukushima, "Cognitron: A self-organizing multilayered neural network," Biol. Cybern., vol. 20, pp. 121-136, 1975.

[2] B. D. Ripley, Pattern recognition and neural networks, Cambridge University Press, Cambridge, 1996.

[3] D. E. Rumelhart, J. L. McClelland and the PDP Research Group, Parallel distributed processing, vol. 1, Cambridge: The MIT Press, 1986.

[4] K. Mikita, "Pattern recognition and feature extraction using neural networks (in japanese)," Thesis for Master's degree in Electrical and Electronic Engineering, The University of Tokushima, 2001.