A LIKED – BAM NEURAL NETWORK FOR IMAGE RECOGNITION *

D.G. SHEN AND F.H. QI
Research Institute of Optical Fiber Engineering
Shanghai Jiao Tong University, 200030

ABSTRACT

A neural network model and its application to image recognition are proposed in this paper. This model consists of Mapping Network (MN) and Liked Bidirectional Associative Memory (LBAM).Invariant mapping is used in MN in order to decrease the number of dimensions of image samples and not to change the distance between them. LBAM'S structure is simple and its convergence speed is fast. Several computer simulations given to prove that the model is capable of recognizing noise – added targets and targets cut off f little part.

1.INTRODUCTION

Neural network as classifier has aroused great interest not only in theoretical research, but also for its many potential applications. It is proved to be useful in several fields. BP is used extensively because it has a automatic training method. Though there are a great many methods for accelerating training speed, the training speed of BP is still unsatisfactory. Probabistic neural network (PNN) is proposed by Donald.F. Specht to have four forward layers in 1988. Its function for classification is equal to Baye Classifier. The training of PNN is completely a forward process. So training is very simple. But it is not practical when the number of training samples increases, for PNN must store all training samples in network. In this paper, LBAM network is devised by controlling several variables to have resonable basins of attraction and resonable stable points. So LBAM network has no false stable points.

* The project supported by National Natural Science Foundation of China.

In order to reduce the scale of LBAM network, we use a invariant mapping to decrease the number of dimensions of image sample. Here invariance means that Euclidean distance between each other is keeped equal after mapping. Mapping transformation is completed in MN.
In this paper, we use MN and LBAM network to recognize image targets. MN model is given in section 2. LBAM network is described in section 3. The structure of the integrated neural network is discussed in section 4. In section 5. the experiment results are presented. The paper concludes in section6.

2.MN model

Assuming that the set of training samples consists of $X_i \in A \subseteq R^{M*M}$, $i=1,2,\cdots\cdots,N$. Here A is the subapace which is made up by $X_i, i=1,2,\cdots\cdots$, N. The number of dimensions of X_i is $M*M$. In fact, subspace A can be represented by orthonomal vectors, supposed Y_i, $i=1,2,\cdots\cdots,N$. Y_i can be obtained by the Gram – schmidt process:

$$\begin{aligned} Y_1 &= X_1/\|X_1\| \\ Y_2 &= (X_2 - C_{12}X_1)/\|X_2 - C_{12}X_1\| \\ &\cdots\cdots \\ Y_N &= (X_N - \sum_{j=1}^{N-1} C_{jN}Y_j)/\|X_N - \sum_{j=1}^{N-1} C_{jN}Y_j\| \end{aligned} \tag{1}$$

Here $C_{ji} = Y_j^T X_i$, $i,j=1,2,\cdots\cdots,N$. That is to say, C_{ji} is the projection coefficient of X_i onto Y_j. $Y_i^T Y_j = 1$ if $i \neq j$, or $Y_i^T Y_j = 0$.

Vector X_i can be combined by Y_i, $i=1,2,\cdots\cdots,N$

$$X_i = \sum_{j=1}^{i} C_{ji} Y_j, \ i=1,2,\cdots\cdots,N \tag{2}$$

Let C denotes the matrix whose i,j – th element is C_{ij}. According to Gram – shmidt process, we get:

$$C=\begin{bmatrix}1 & C_{12} & C_{13} & \cdots & C_{1N}\\ 0 & C_{22} & C_{23} & \cdots & C_{2N}\\ & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & C_{NN}\end{bmatrix} \tag{3}$$

Then equation (2) can be rewritten via matrix:

$$X=YC \tag{4}$$

Here: $X=[X_1\ X_2\ \cdots\ X_N]$
$Y=[Y_1\ Y_2\ \cdots\ Y_N]$
It is known that X_i, $i=1,2,\cdots N$, are linearly independent each other. Thus we can infer that Y, $i=1,2,\cdots,N$, are still independent each other. So vector $C_i=[C_{1i}\ C_{2i}\ \cdots\ C_{Ni}]^T$, $i=1,2,\cdots,N$, can express the difference between each other. On the basis of above defination, we can get following useful properties.

Property 1: The Euclindean between sample i and sample j is not changed by mapping transformation. That is:

$$\|X_i-X_j\|=\|C_i-C_j\|$$

Property 2:
If $\|XX-X_i\|\leqslant\|XX-X_j\|$, then $\|CC-C_i\|\leqslant\|CC-C_j\|$

Here XX is any vector in space R^{MM}. CC is the projection of XX onto Y. That is:

$$CC=P\cdot XX \tag{5}$$

Mapping matrix in MN is:

$$P=Y^T \tag{6}$$

That total mapping processing can be described as $R^{MM}\rightarrow R^N$, that is to say $XX\rightarrow CC$. In this paper, vector C_i, $i=1,2,\cdots,N$, is normalized respectively and vector CC is still normalized when vector XX is to be recognized.

3. LBAM network

Vector X_i, $i=1,2,\cdots,N$, are the training samples for MN model and C_i, $i=1,2,\cdots,N$, are the output of MN model. So C_i, $i=1,2,\cdots,N$, are the training samples for LBAM network. According to above defination, we know that CC and C_i, $i=1,2,\cdots,N$, are mormalized respectively. That is to say:

$$\|CC\|=1$$
$$\|C_i\|=1,\ i=1,2,\cdots,N \tag{7}$$

Thus Euclidean distance between C_i and CC is:

$$\begin{aligned}d_i^2=\|CC-C_i\|^2&=(CC-C_i)^T(CC-C_i)\\&=\|CC\|^2+\|C_i\|^2-2CC^TC_i\\&=2-2CC^TC_i\\&=2-2dd_i,\ i=1,2,\cdots,N\end{aligned}$$

Here $dd_i=CC^TC_i=\sum_{j=1}^{N}CC_jC_{ji}$

Output function O_i is devised for every d_i^2, $i=1,2,\cdots,N$. It is 0 if $d_i=1$ and decrease if d_i^2 increases. So Gussian function can be adopted:

$$\begin{aligned}O_i&=\exp(-d_i^2/(2\epsilon_i^2))\\&=\exp(-2-2dd_i)/(2\epsilon_N^2))\\&=\exp(-1/\epsilon i^2+1/\epsilon_i^2\sum_{j=1}CC_jC_{ji})\end{aligned} \tag{9}$$

Thus $O=-\sum_{i=1}^{N}A_iO_i\exp(-d_i^2/(2\epsilon_i^2))$ is the sum of a series of Gaussian functions. We can make affect between O_i and O_j, $i=1,2,\cdots,N$ and $i=j$, small by controlling ϵ_i^2. Gaussian O_i becomes sharp if ϵ_i^2 is small. So function O is the sum of a series of sharp Gaussian functions. In fact, these sharp functions are associated with stable points and basins of attraction. Thus:

$$\begin{aligned}\partial O/\partial CC_j&=-\sum_{i=1}^{N}A_i\partial O_i/\partial CC_j\\&=-\sum_{i=1}^{N}A_iO_iCC_{ji}/\epsilon_i^2\\&=-\sum_{i=1}^{N}B_iC_{ji}O_i\\B&=A_i/\epsilon_i^2,\ j=1,2,\cdots,N\end{aligned} \tag{10}$$

When the following iterative equation is used,

$$\begin{aligned}CC_j&=CC_j-\beta_j\partial O/CC_j\\&=CC_j+\beta_j\sum_{i=1}^{N}B_iCC_{ji}O_i\end{aligned}$$
$$j=1,2,\cdots,N \tag{11}$$

We can guarantee that O becomes small after iteration. If $B_j=B$, $j=1,2,\cdots,N$, then equation (11) can rewritten as:

$$\begin{aligned}CC_j&=\beta_jB(CC_j/\beta_jB+\sum_{i=1}^{N}C_{ji}O_i)\\&=g_i(CC_j/\beta_jB+\sum_{i=1}^{N}C_{ji}O_i)\end{aligned} \tag{12}$$

Here $g_j(x)=\beta_jBx$
Equation (9) can be rewriten like equation (12).

$$O_i=\exp(-1/\epsilon_i^2+1/\epsilon_i^2\sum_{j=1}^{N}CC_{ji}C_{ji})$$

$$= f_i(\sum_{j=1}^{N} CC_j C_{ji})$$

$$f_i(x) = \exp((x-1)/\epsilon_i^2)$$

It is known from equation (11) and function O's property that iterative process is able to converge to a globle points.

Iterative process can be represented by network illustrated in Fig.1

Parameters A_i, $i=1,2,\cdots N$, are equal if every target is equally important. In general, we use A_i, $i=1,2,\cdots N$. It is similar to adopted $\beta_j=\beta$, $j=1, 2,\cdots N$. When the training number of every class is very large, C_i represent the mean of $i-$th class and ϵi^2 is convariance of $i-$th class.

In this paper, the used network is made up by MN and LBAM. Using the property of equation (3), LBAM network in Fig.2 can be simplified, which is shown in Fig.2.

4. The structure of total network

The network used in this paper is made up by MN and LBAM. The total structure is shown in Fig.3.

The recognition procedure of network can be concluded as: the unknown target XX get through MN model and MN model output CC. CC enter LBAM network. By bidirectional association of LBAM for several times, LBAM network output the recognition result. There is only $O_{io}=1$ and $O_i=0$ if $i \neq io$.

5. Experiment result

In these experiments, neural network is used in English character recognition and plane target recognition. The targets to be recognized are bilevel images in which 1 represent target and 0 to ground. In the experiment for plane recognition. image vectors are $64*64$ and we use 10 targets to train network. It costs 14 seconds on a microcomputer AST-386 to train network. We use 100 targets, each of which is added Gaussian noise randomly. In this experiment, we use $\epsilon_i^2=0.01$, $A_i=1$, $\beta_i=0.0005$, $B_i=A_i/\epsilon_i^2=100$, $i=1,2,\cdots N$. It costs only 0.5 second to recognize unknown target. Some of the computer simulation is shown in Fig.4. On the same condition, the network is used in English character recognition. The results of expeliment are shown in Fig.5.

6. Conclusion

In this paper, A LBAM networkis devised by global method. The structure of LBAM is very simple because MN model is used. On the other hand, LBAM network is global optimun.

LBAM network as bidirectional association classifier can be used in many fields. Besides example in this paper, it can be used in vector coding too.

REFERENCE

1. Kunihiko Fukushima, "Neocognition: A self-orgnaizing Neural Network Model for a Mechanism of Pattern Recognition unaffected by Shift in Position", Biol. Cybernetics 36, 193-202(1980).
2. Widrow. B.et, "Layered Neural Nets for Pattern Recognition", IEEE Trans. on ASSP, 36,7,1988, PP.1109-1118.
3. E.Gullichsen and E. chang, "Pattern classification by neural network: An experimental system for icon recognition", IEEE First Ann. Conf. Neural Netwurks, San diego, California (June 21-24,1987).
4. D.F. Specht,"Probabilistic Neural Networks", in Neural Networks, vol.3, No.1, pp.109-118, Jan. 1990.
5. D.F. Specht,"Probabilistic Neural Networks and the Polynomid adaline as complementary Techniques for classification", in Proc. IEEE int. Conf. on Neural Networks, Vol.1, No.1, March 1990.
6. Dinggang shen and Feihui Qi, "A neural network model for image recognition", Int. conf. on Neural Network in Beijing 1992.
7. Feihu Qi and Dinggang shen,"A FDO Neural Network and Its Application in Character Recognition", Int. Conf. on Oriental Language and Chinese Character Processing in America 1992.
8. Dinggang shen and Feihu Qi, "AFDO Neural Network and Its application to image Recognition", ICCS/ISITA in SINGAPORE in 1992.

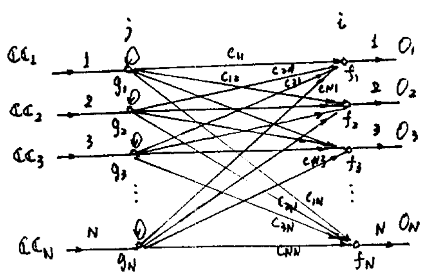

Fig.1 The structure of LBAM

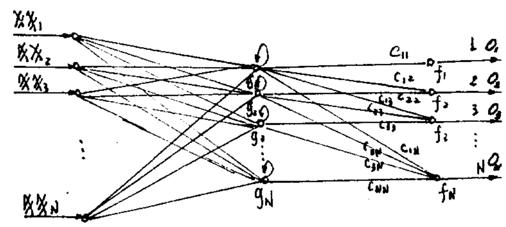

Fig.3: The network used in image recognition

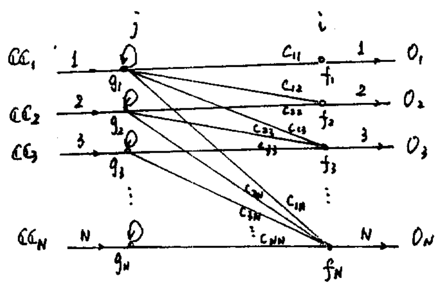

Fig.2: The structure of simplified LBAM

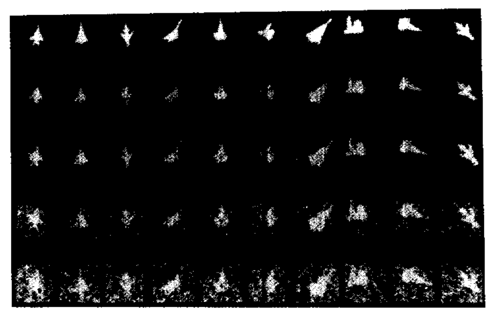

Fig.4: (a) training targets
(b) the targets can be recognized

Fig.5: (a) training characters
(b) the result of the unknown characters