

Enhanced Fuzzy Single Layer Perceptron

Gyoo-Yong Chae, Sang-Hee Eom and Kwang-Baek Kim, Member, KIMICS

Abstract—In this paper, a method of improving the learning speed and convergence rate is proposed to exploit the advantages of artificial neural networks and neuro-fuzzy systems. This method is applied to the XOR problem, n bit parity problem, which is used as the benchmark in the field of pattern recognition. The method is also applied to the recognition of digital image for practical image application. As a result of experiment, it does not always guarantee convergence. However, the network showed considerable improvement in learning time and has a high convergence rate. The proposed network can be extended to any number of layers. When we consider only the case of the single layer, the networks had the capability of high speed during the learning process and rapid processing on huge images.

Index Terms—XOR problem, learning speed, convergence rate, pattern recognition

I. INTRODUCTION

In the recent trend of image recognition research, there is a vigorous interest to apply artificial networks and fuzzy theory beyond simple image recognition. However, we may face the problems-oscillation at local minima, high price in learning speed, misidentification to degrade the efficiency of recognition etc. as well as inability to regulate rules using respectively existing artificial networks and fuzzy theory.

In the conventional single layer perceptron, it is inappropriate when a decision boundary for classifying input pattern does not composed of hyperplane. Moreover, the conventional single layer perceptron, due to its use of unit function, was highly sensitive to change in the weights, difficult to implement and could not learn from past data[1]. Therefore, it could not find a solution of the exclusive OR problem, the benchmark.

There are a lot of endeavor to implement a fuzzy theory to artificial neural network[2]. Goh et al.[3] proposed the fuzzy single layer perceptron algorithm, and advanced fuzzy perceptron based on the generalized

delta rule to solve the XOR problem, and the classical problem[3]. This algorithm guarantees some degree of stability and convergence in application using fuzzy data, however, it causes an increased amount of computation and some difficulties in application of the complicated image recognition. However, the enhanced fuzzy perceptron has shortcomings such as the possibility being located in local minima and slowness learning time[4].

In this paper, we propose a enhanced fuzzy single layer learning algorithm. We construct, and train, a new type of fuzzy neural net to model the linear function. Properties of this new type of fuzzy neural net include: (1) proposed linear activation function; and (2) a modified delta rule for learning. We will show that such properties can guarantee to find solutions for the problems-such as exclusive OR, 3 bit parity, 4 bit parity and digit image recognition on which simple perceptron and simple fuzzy perceptron can not.

II. A FUZZY SINGLE LAYER PERCEPTRON

The learning algorithm for our single layer perceptron will be proposed. Before we discuss the new learning algorithm, we introduce the proposed learning architecture. Fig.1 shows the architecture of the new learning algorithm.

A. A Fuzzy Single Layer Learning Algorithm

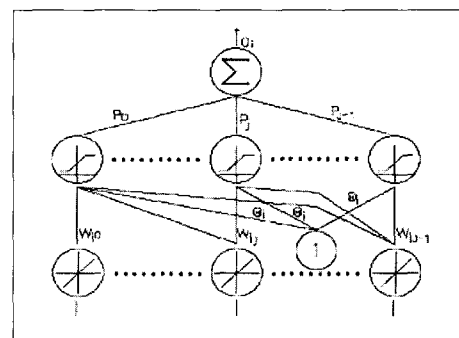


Fig.1 A fuzzy single layer perceptron model.

A proposed learning algorithm can be simplified and divided into four steps. For each input, repeat step 1, step 2, step 3, and step 4 until error is minimized

Step 1 : Initialize weight and bias term.

Define, W_{ij} ($1 \leq i \leq I$) to be the weight from input j to output i at time t , and θ_i to be the bias term in the output soma. Set $W_{ij}(0)$ to small random values, thus initializing all the weights and bias term.

Manuscript received March 25, 2004.

G. Y. Chae is with the IT Design Research Center, Silla University, Busan, Korea, (e-mail: cgy1234@hanmail.net)

S. H. Eom is with School of Computer Information and Communications, Dongju College, Busan, Korea, (e-mail: sheom@dongju.ac.kr)

K. B. Kim is with the Computer Engineering Department, Silla University, Busan, Korea, (e-mail: gbkim@silla.ac.kr)

Step 2 : Rearrange A_j according to the ascending order of membership degree m_j and add an item m_0 at the beginning of this sequence.

$$0 = m_0 \leq m_1 \leq \dots \leq m_j \leq m_{j-1} \leq 1.0$$

Compute the consecutive difference between the items of the sequence.

$$P_k = m_j - m_{j-1} \quad \text{Where } k = 0, \dots, n$$

Step 3 : Calculate a soma(O_j)'s actual output.

$$O_i = \sum_{k=0}^{J-1} P_k * f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right)$$

Where $f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right)$ is linear activation function

Where $i = 1, \dots, I$.

In the sigmoid function if the value of $\left(\frac{1.0}{1.0 + e^{-net}}\right)$ is between 0.0 and 0.25, $\left(\left(\frac{1.0}{1.0 + e^{-net}}\right) * \left(1 - \frac{1.0}{1.0 + e^{-net}}\right)\right)$ is very similar to $\left(\frac{1.0}{1.0 + e^{-net}}\right)$.

If the value of $\left(\frac{1.0}{1.0 + e^{-net}}\right)$ is between 0.25 and 0.75, $\left(\left(\frac{1.0}{1.0 + e^{-net}}\right) * \left(1 - \frac{1.0}{1.0 + e^{-net}}\right)\right)$ is very similar to 0.25. If the value of $\left(\frac{1.0}{1.0 + e^{-net}}\right)$ is between 0.75 and 1.0, $\left(\left(\frac{1.0}{1.0 + e^{-net}}\right) * \left(1 - \frac{1.0}{1.0 + e^{-net}}\right)\right)$ is very similar to $\left(1 - \frac{1.0}{1.0 + e^{-net}}\right)$.

Therefore, the proposed linear activation function expression is represented as follows:

$$f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) = 1.0 \quad \text{where } \left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) > 5.0$$

$$f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) = \rho * \left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) + 0.5$$

$$\text{where } -5.0 \leq \left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) \leq 5.0, \rho \in [0.1, 0.4]$$

$$f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) = 0.0 \quad \text{where } \left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) < -5.0$$

The formulation of the activation linear function is follow.

$$f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) = \left(\frac{1}{range * 2}\right) * \left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) + 0.5$$

Where the range means monotonic increasing internal except for interval except for the interval between 0.0

and 1.0 of value of the $f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right)$

Step 4: Applying the modified delta rule. And we derive the incremental changes for weight and bias term.

$$\Delta W_{ij}(t+1) = \eta * E_i * \sum_{k=0}^J P_k * f\left(\sum_{j=k}^{J-1} W_{ij} + \theta_i\right) + \alpha_i * \Delta W_{ij}(t)$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}(t+1)$$

$$\Delta \theta_i(t+1) = \eta * E_i * f\left(\theta_i\right) + \alpha_i * \Delta \theta_i(t)$$

$$\theta_i(t+1) = \theta_i(t) + \Delta \theta_i(t+1)$$

Where η is learning rate α_i is momentum.

Finally, we enhance the training speed by using the dynamical learning rate and momentum based on the division of soma.

$$\begin{aligned} &\text{if } (Inactivation_{totalsoma} - Activation_{totalsoma} > 0) \\ &\text{then } \Delta \eta_i(t+1) = E^2 \\ &\quad \eta_i(t+1) = \eta_i(t) + \Delta \eta_i(t+1) \\ &\text{if } (Inactivation_{totalsoma} - Activation_{totalsoma} > 0) \\ &\text{then } \Delta \alpha_i(t+1) = E^2 \\ &\quad \alpha_i(t+1) = \alpha_i(t) + \Delta \alpha_i(t+1) \end{aligned}$$

B. Error Criteria Problem by Division of Soma

In the conventional learning method, learning is continued until squared sum of error is smaller than error criteria. However, this method is contradictory to the physiological neuron structure and takes place the occasion, which a certain soma's output is not any longer, decreased and learn no more[5]. The error criteria were divided into activation and inactivation criteria. One is an activation soma's criterion of output "1", the other is an inactivation soma's of output "0". The activation criterion is decided by soma of value "1" in the discriminated problem of actual output patterns, which means in physiological analysis that the pattern is classified by the activated soma. In this case, the network must be activated by activated somas. The criterion of activated soma can be set to the range of [0,0.1].

In this paper, however, the error criterion of activated soma was established as 0.05. On the other hand, the error criterion of inactivation soma was defined as the squared error sum, the difference between output and target value of the soma. The degree of activation and inactivation was equally set up on the basis that activation and inactivation is the same[5,6]. Fig.2 shows the proposed algorithm.

```

while ((Activation_no == Target_activated_no)
    && (Inactivation_error <= Inactivation_area))
do {
    for (i=0; i<Pattern_no; i++)
        for(j=0; j<Out_cell_no; j++) {
            Forward Pass; Backward Pass;
            if(Out_cell==Activation_soma&&|error|<=

```

```

    Activation area)
    Activation_number++;
    if (Out_cell==Inactivation_soma)
    Inactivation_error += error * error;
}
}

```

Fig.2 Learning Algorithm by division of soma

III. SIMULATION & RESULT

We simulated our method on IBM PC/586 with C++ language. In order to evaluate the proposed algorithm, we applied it to the exclusive OR, 3 bit parity using benchmark in neural network and pattern recognition problems, a kind of image recognition. In the proposed algorithm, the error criteria of activation and inactivation for soma were set to 0.09."

A. Exclusive OR and 3 bit parity

Here we set up initial learning rate 0.5 initial momentum 0.75 respectively. Also we set up the range of weight [0,1]. In general, the range of weights were [-0.5,0.5] or [-1,1].

As shown in Table 1 and Table 2, our model showed higher performance than fuzzy perceptron in convergence epochs and convergence rates of the three tasks.

Table 1 Convergence rate in initial weight range

Appled Problem	Fuzzy Perceptron	Proposed Algorithm	Initial Weight Range
Exclusive OR	100%	100%	[0.0, 1.0]
	89%	99%	[0.0, 5.0]
3 bit parity	83%	97%	[0.0, 1.0]
	52%	96%	[0.0, 5.0]
4 bit parity	0%	96%	[0.0, 1.0]
	0%	95%	[0.0, 5.0]

Table 2 Comparison of step number

Epoch No	Fuzzy Perceptron	Proposed Algorithm
Exclusive OR	8 (converge)	4 (converge)
3 bit parity	13 (converge)	8 (converge)
4 bit Parity	0 (not converge)	15 (converge)

B. Digit image recognition

The procedure of image pre-processing is presented in Fig. 3 and the example images we used are shown in Fig. 4

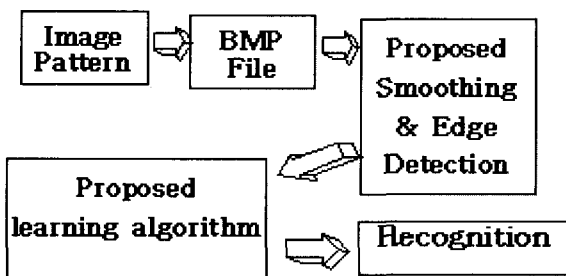


Fig. 3 Preprocessing diagram.

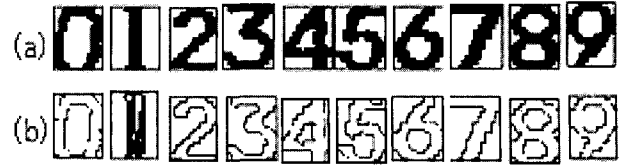


Fig. 4 (a) Digit Image and (b) Training image by edge detection.

We carried out image pre-processing in order to prevent high computational load as well as loss of information. If the original images were used as training patterns, it requires expensive computation load. In contrast skeleton method causes loss of important information of images. To overcome this trade-off, we used edge information of images.

The most-frequent value method we had developed was used for image pre-processing. This method was used because blurring of boundary when a common smoothing method was used. Thus, it degrades both color and contour lines[7,8]. The new method replaced a pixel's value with the most frequent value among specific neighboring pixels. If the difference of absolute value between neighborhoods is zero in a given area, the area was considered as background. Otherwise, it was considered as a contour. This contour was used as a training pattern.

The input units were composed of 32 * 32 array for image patterns. In simulation, the fuzzy perceptron was not converged, but the proposed method was converged on 70 step at image patterns. Table 3 is shown the summary of the results in training epochs between two algorithms.

Table 3 The comparison of epoch number

Image Pattern	Epoch Number
Fuzzy perceptron	0 (not converge)
Proposed algorithm	70 (converge)

IV. CONCLUSIONS

The study and application of fusion fuzzy theory with logic and inference and neural network with learning ability have been actually achieving according to expansion of automatic system and information processing, etc.

We have proposed a fuzzy supervised learning algorithm, which has greater stability and functional varieties compared to the conventional fuzzy perceptron, and a new learning algorithm for the network with enhanced fuzzy learning algorithm that for the first time, allows automatic extraction of fuzzy relations from data.

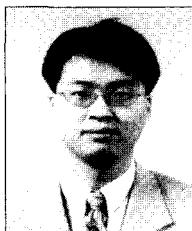
The proposed network is able to extend the arbitrary layers and has high convergence in case of two layers or more. When we considered only the case of the single layer, the networks had the capability of high speed during the learning process and rapid processing on huge image patterns.

REFERENCES

- [1] F. Rosenblatt, "The perceptron : A perceiving and recognizing automaton," Cornell Univ. Ithaca, NY, Project PARA Cornell Aeronaut Lab, Rep., 85-460-1, 1957.
- [2] M. M. Gupta and J. Qi, "On Fuzzy Neuron Models," Proceedings of IJCNN, Vol.2, pp. 431-435, 1991.
- [3] TH Goth, PZ Wang, and HC Lui, "Learning Algorithm for Enhanced Fuzzy Perceptron," Proceedings of IJCNN, Vol.2, pp. 435-440, 1992.
- [4] Kwang Baek Kim, Eui Young Cha, "A New Single Layer Perceptron using Fuzzy Neural Controller," Simulators International XII, Vol.27, No.3, pp.341-343, 1995.
- [5] Kwang Baek Kim, Ji Ae Jun, and Eui Young Cha, "The Solution Method of Learning Time and Local Minima of Error Backpropagation by the Nervous System," Proceedings of JCEANF, pp.592-601, 1992.
- [6] Judith E. Dayhoff, "Biological Synapse," Neural Network Architectures, pp.136-162, 1989.
- [7] D. G. Lowe, "Organization of Smooth Image Curves at Multiple Scales," International Journal of Computer Vision, 1:119-130, 1989.
- [8] R. B. Paranjape, R. N. Rangayyan, W. M. Morrow, and H. N. Nguyen, Adaptive neighborhood image processing, CVGIP Graphical Models and Image Processing, 54(3): 259-267, 1992.

**Gyooyong Chae**

He received the M.S. and the Ph.D. degrees in the Dept. of industrial engineering, Konkuk University. At present, he is a researcher of IT design research Center, Silla University His research interests fuzzy Logic, Artificial Intelligence, Data Mining etc.

**Sang-Hee Eom**

He received the B.S. and M.S. degrees in the Electrical Engineering from Donga University. He received Ph.D. degree in 2000 at the Electronic Engineering of Busan National University. He is currently with School of Computer Information & Communications, Dongju College as a Full-Time Lecture. His research interests are Signal Processing, Pattern Recognition, Artificial Intelligence, Biomedical Engineering etc.

**Kwang-Baek Kim**

He received the M. S. and the Ph.D. degrees in the Department of Computer Science from Busan National University, Busan, Korea, in 1993 and 1999, respectively. From 1997 to present, he is an Associate Professor, Department of Computer Engineering, and Silla University in Korea. His research interests include Fuzzy Neural Networks and Application, Image Processing, Biological Signal Processing and Biomedical System.