Improvement of Learning Performance of Multi-Layer Perceptron by Two Different Pulse Glial Networks

Chihiro Ikuta†, Yoko Uwate†, Yoshifumi Nishio†, and Guoan Yang‡
†Dept. of Electrical and Electronics Eng., Tokushima University
2–1 Minami–Josanjima, Tokushima Japan
Email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp
‡Dept. of Automation Science and Technology,
School of Electronic and Information Eng., Xi'an Jiaotong University
No.28 Xianning West Road, Xi'an City, Shaanxi Province, China
Email: gayang@mail.xjtu.edu.cn

Abstract—A glia is the most number of nervous cells in a brain. The glia is investigated in a medical field, because the glia correlates to neuron works and composes a human cerebration. We consider that the glia function can be applied to an artificial neural network. In this study, we propose the Multi-Layer Perceptron (MLP) with the two different pulse glial networks. The proposed MLP has the glial network which is inspired from biological functions of the glia. One neuron is connected with two glias. Two glias generate the pulse depending on the output neurons. One glia connects the neuron for increasing the output of neuron. On the other hand, the glia connects the neuron for decreasing the output of neuron. Both glias composes the glial networks. These effects are propagated into the networks. The glial effects become complexity and affects the MLP learning performance. By the computer simulation, we confirm that the learning performance of the proposed MLP is better than the conventional MLP.

I. INTRODUCTION

Applications of neurons have been investigated for a long time. However, the glias have not attracted attentions, because the glias can not use an electric signal. Recently, some researchers discovered that the glias transmitted signals by using the ions such as Ca^{2+}, Glu, ATP, and so on [1][2]. The glias are considered to be important for the brain works. Especially, Ca^{2+} can change a membrane potential of the neuron [3][4]. Moreover, the Ca^{2+} affects the states of the neighboring glias. The Ca^{2+} effects are propagated to wide range in the brain [5][6]. We consider that the glias can be applied to the artificial neural network.

In the previous study, we proposed the positive and negative pulse glial chain [7]. In this network, the glias are connected with the neurons one by one. When the neuron has large output, the connected glia generates the positive value pulse. When the neuron has small output, the connected glia generates the negative value pulse. The generated pulses influence the neuron threshold and neighboring glial state The positive and negative pulse glial chain influences the neuron threshold for increasing the neuron output.

In this study, we propose the Multi-Layer Perceptron (MLP) with two different pulse glial networks. One neuron connects with two glias. Two glias have different excitation thresholds and inverse value output. $Glia_1$ gives the pulse to the neuron for increasing the neuron output. $Glia_2$ gives the pulse to the neuron for decreasing the neuron output. When the neuron has large the output, the glia is excited and generate the pulse. First, $glia_1$ is excited because the $glia_1$'s excitation threshold is smaller than the $glia_2$. By the glial effect, the neuron output becomes larger more. The $glia_2$ is excited when the neuron output is over the excitation threshold of the $glia_2$. We consider that the glial effect give the energy to the MLP for escaping out from the local minimum. By the computer simulation, we confirm that the proposed MLP obtain the better learning performance than the conventional MLP.

II. PROPOSED MLP

The MLP is a famous feed forward neural network. It is composed of the layer of neurons. It can solve various problem such as a classification problem, a pattern recognition, a data mining, and so on. We can change the output of the network by tuning the weight of connections. We often use a Back Propagation (BP) algorithm for the deciding method of the weight of connections [8]. However, the BP algorithm often falls into a local minimum, because it use the steepest decent method. For the local minimum problem, we exploit the noise. The noise gives the energy for escaping out from the local minimum.

In this study, we propose the MLP with two different pulse glial networks. The construct of the proposed MLP is illustrated in Fig. 1. The two different pulse glial networks are composed of the glias. In this network, we connect two glias to one neuron. The glias generate the pulse depending on the neuron output. The glial pulse is inputed to the neuron threshold. We consider that the glial pulse give the energy to the MLP for escaping out from the local minimum. Moreover, the pulses are propagated to the neighboring glias.

We consider that the amount of pulse is changed little by little in the neurons. We believe that this work gives the novel relationships in the hidden layer of the MLP.
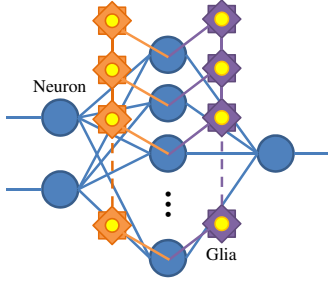

Fig. 1. The MLP with two different pulse glial networks.

A. Updating rule of neuron

The neuron has multi-inputs and single output. We can change the neuron output by the tuning weights of connections. The standard updating rule of the neuron is defined by Eq. (1).

$$y_i(t+1) = f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t)\right), \qquad (1)$$

where y is an output of the neuron, w is a weight of connection, x is an input of the neuron, and θ is a threshold of neuron. We connect the two different pulse glial networks to the hidden-layer of the MLP. Two glial pulses input to the neuron threshold, because the biological glial effect affects to the membrane potential of neuron. We define the proposed neuron updating rule in Eq. (2).

$$y_i(t+1) = $$
$$f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t) + \alpha(\psi_{i1}(t) + \beta\psi_{i2}(t))\right)(2)$$

where Ψ_1 and Ψ_1 are the outputs of the glias, α is a weight of the glial effect, β is a weight of Psi_2. By β, the amount ratio $\psi_i 1$ and $\psi_i 2$ are changed.

B. Two different pulse glial networks

The glia is one of nervous cells in the brain. It transmits signals by ion concentrations. We note this function and we have proposed the glial network. In this study, we extend the glial network to two different pulse glial networks. The glial network is connected with the neurons in the hidden layer of the MLP. It gives the energy for increasing the neuron output. In this model, we sometimes observed that the MLP learning is early fixed. We hope that the glial network help the MLP, however the glial network gives constant amount of pulse. Thus, the MLP cannot escape out from the local minimum. However, the MLP cannot learn when the glial effect is strong. From this result, we consider that the MLP needs more energy for escaping out from the local minimum.

Moreover, the glial network should change the work depending on the MLP learning.

In two different pulse glial networks, we connect two glias to one neuron. Two glias make different networks. One glia can easily excite by the neuron output. On the other hand, the glia need the large neuron output for excitation.

$$\psi_i(t+1) = $$
$$\begin{cases} 1, & \{(\theta_{ni} < y_i \cup \psi_{i+1,i-1}(t - i*D) = 1) \\ & \qquad\qquad\qquad \cap (\theta_g > \psi_i(t))\} \\ \gamma\psi_i(t), & else, \end{cases} , (3)$$

$$\psi_i(t+1) = $$
$$\begin{cases} -1, & \{(1 - \theta_{ni} < y_i \cup \\ & \psi_{i+1,i-1}(t - i*D) = 1) \cap (\theta_g > \psi_i(t))\} \\ \gamma\psi_i(t), & else, \end{cases} , (4)$$

where ψ is an output of a glia, γ is an attenuated parameter, y is an output of a connecting neuron, θ_n is a glia threshold of excitation, θ_g is a period of inactivity, and D is a delay time of a glial effect. The glia does not learn, it depend on the output of connected neurons. However, the neurons are learned by BP algorithm, thus the generation pattern of glia output can dynamically change during the learning. In the two different pulse glial networks, two glial networks have different θ_n. We define $\theta_{n1} < \theta_{n2}$ in every simulation. Thus, the connected two glias are excited different conditions. We believe that the pulse glial network which has θ_{n2} help the proposed MLP when the proposed MLP is trapped into local minimum. We present an example of the pulses by two glias in Fig. 2. In this case, the glia 1 only generates the negative pulses. However, the glia 2 generates the positive and the negative pulses. Thus, the neuron receives the positive and the negative pulses. In the previous glial network, the pulse often becomes like glia 1. We consider that the pulse pattern become complexity by connected two glial network.

Figure 3 is an example of the inputed glial effect which is added two glial pulses. We show the minimum output to the black, the large output is drawn in the white, and the gray is near 0. The generated glial effect patterns are dynamically change, because two glias have the inversed work for the neuron output. Thus, neuron output is changed by the glial effects. Moreover, the neurons are learned by the BP algorithm and the neurons' outputs are changed. We can see that the pulses are propagated into the glial network. The pulses are sometime disappeared, because two connected glias annihilate each other. In the case of dark black or light white, two glias generate same vector values. Thus, they are turned up each effect.

III. SIMULATION

In this simulation, we use Two-Spirals Problem (TSP) to the benchmark shown in Fig. 4. The TSP is a famous benchmark for the artificial neural network. Moreover, the TSP is known the high non-linearity problem, thereby the standard MLP cannot easily solve this problem. The MLP learns two different spirals coordinates and classifications of spirals. Moreover,
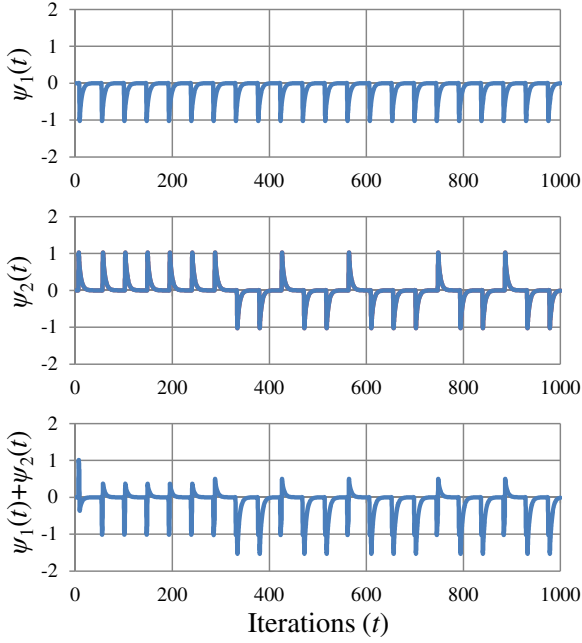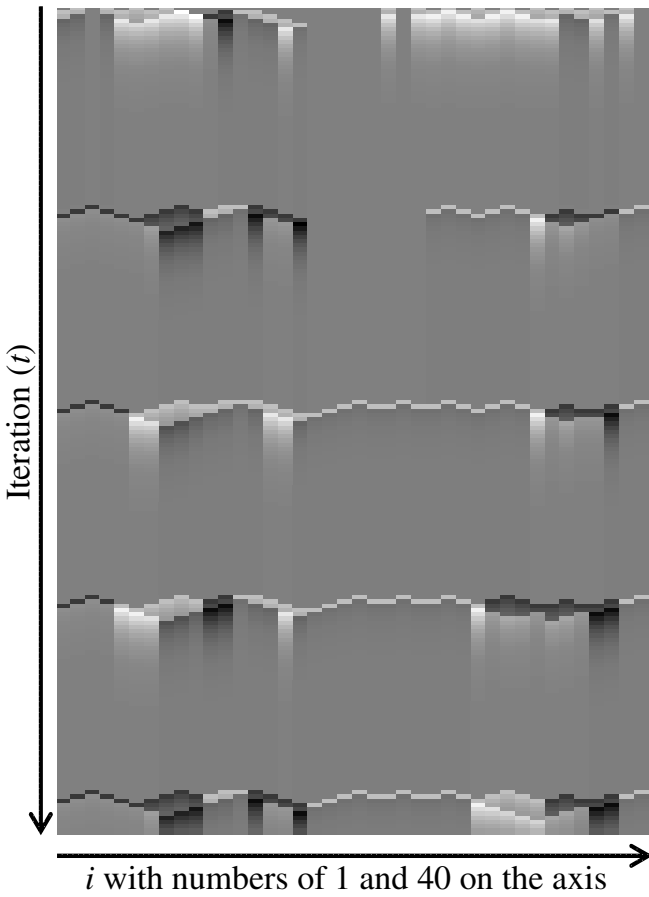
Fig. 2. An example of the glial pulse.

i with numbers of 1 and 40 on the axis

Fig. 3. The inputted glial effects into 40 neurons.

the learned MLP can classier non-learned coordinates to two spirals. The ideal classification result is presented in Fig. 4 (b). We compare the MLP with two different pulse glial networks
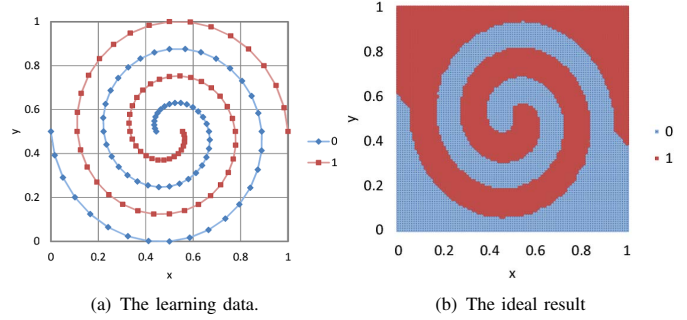

(a) The learning data.

(b) The ideal result

Fig. 4. Two-spirals problem.

to the MLP with pulse glial chain (two directions). The MLPs are composed 2-40-1 neurons. The MLPs learn 50000 times in 1 trial. We obtain the average of error from 100 trials. For the measure, we use the Mean Square Error(MSE) which is described by Eq. (5).

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (T_n - O_n)^2, \tag{5}$$

where T is a target value, O is an output of the MLP.

Figure 5 shows the approximation results for different parameters. In the this figure, the horizontal axis is weight of glial effect α and the vertical axis is the MSE. We decide $\theta_{n1} = 0.8$ which is not changed and we change $\theta_{n2} = 0.9, 0.95 and 0.99$. In the case of $\beta = 0.5$, the inverse effect is small, thus the curves are similar to the previous MLP. And also, the three curves of the proposed MLP are similar. Because, the influence of the inverse glial network is smaller then another glial network. However, when β becomes high, the approximation performance of proposed MLP is better than the previous MLP. The proposed MLP has a better performance than the previous MLP at $\beta = 1.0$. Then, the curves of proposed MLP are similar trends to the previous MLP, however the errors are reduced in every α. In Fig 5 (c), the approximation performance decreased when β is higher than 0.8 Because the glial effect become too large by β. We consider that the MLP cannot learn the ideal point by the glial effects.

Next, we present the classification performance in each condition. In this result, we input the coordinates in $x - y$ field to after learning MLP. After that the MLP obtain the output. We compare the ideal classification (Fig. 4 (b)) to the MLP output. The classification performances have similar trend to the approximation performances. In $\beta = 0.5$, the classification performances of the proposed MLP are similar to the previous MLP. In $\beta = 1.0$, the classification performances of the proposed MLP are better than the previous MLP. The proposed MLP is worth classification performance than the previous MLP when the glia has $\beta = 2.0$ and α is high. From these results the two pulse glial network give the good influence to both abilities of the MLP.

358

(a) $\beta = 0.5$.

(b) $\beta = 1.0$.

(c) $\beta = 2.0$.

Fig. 5. Approximation results.

(a) $\beta = 0.5$.

(b) $\beta = 1.0$.

(c) $\beta = 2.0$.

Fig. 6. Classification results.

IV. CONCLUSIONS

We have proposed MLP with two different pulse glial networks. The proposed network is inspired from the functions of the biological glia. In the proposed network, we connected two glias to one neuron. And also, two glias make the different glial networks which have different parameters. The glia is fired by the connecting neuron output, after that the fired glia generates the pulse. The pulse affects the neuron threshold and the state of neighboring glias. Thus, the glial pulses are propagated into the glial network. We considered that the pulses are affected each other, thereby the pulses pattern become more complexity than the previous network. We confirmed that the proposed MLP has the better performances than the previous MLP in some conditions.

REFERENCES

[1] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol. 2, pp. 844-847, 2001.
[2] S. Koizumi, M. Tsuda, Y. Shigemoto-Nogami and K. Inoue, "Dynamic Inhibition of Excitatory Synaptic Transmission by Astrocyte-Derived ATP in Hippocampal Cultures," *Proc. National Academy of Science of U.S.A*, vol. 100, pp. 11023-11028, 2003.
[3] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.
[4] G. Perea and A. Araque, "Glial Calcium Signaling and Neuro-Glia Communication," *Cell Calcium*, vol. 38, pp. 375-382, 2005.
[5] S. Kriegler and S.Y. Chiu, "Calcium Signaling of Glial Cells along Mammalian Axons," *The Journal of Neuroscience*, vol. 13, pp. 4229-4245, 1993.
[6] M.P. Mattoson and S.L. Chan, "Neuronal and Glial Calcium Signaling in Alzheimer's Disease," *Cell Calcium*, vol. 34, pp. 385-397, 2003.
[7] C. Ikuta, Y. Uwate and Y. Nishio, "Multi-Layer Perceptron with Positive and Negative Pulse Glial Chain for Solving Two-Spirals Problem," *Proc. IJCNN'12*, pp. 2590-2595, Jun. 2012.
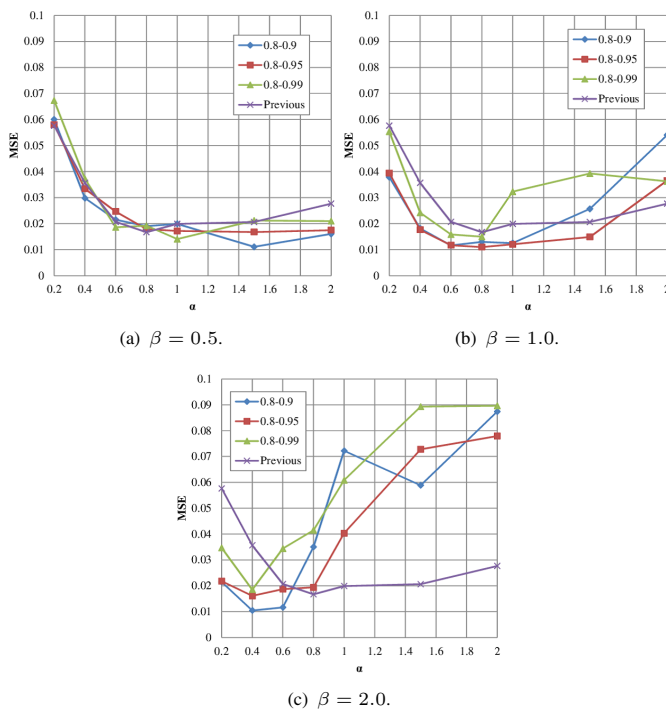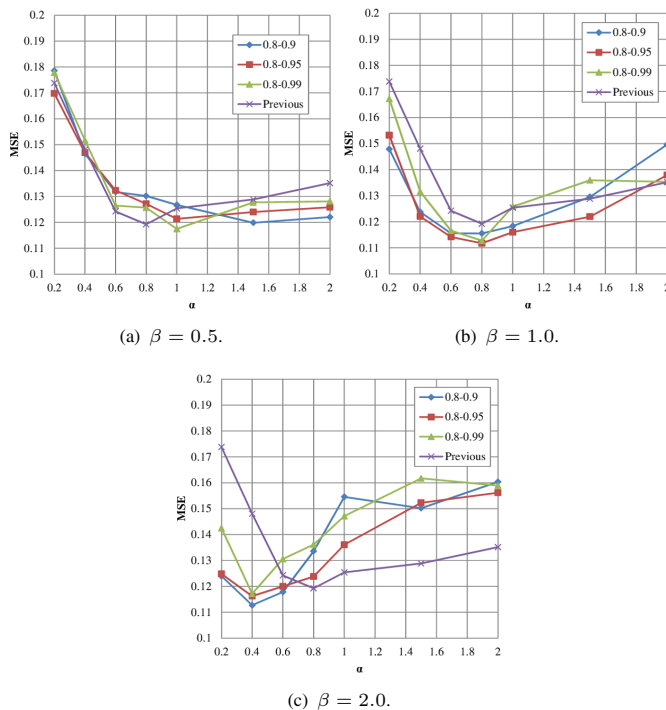[8] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.