Multi-Layer Perceptron with Local Glia Connection

Chihiro Ikuta, Yoko Uwate, and Yoshifumi NIshio

†Department of Electrical and Electronics Engineering, Tokushima University
2-1, Minami-Josanjima, Tokushima, Japan
Email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract—In a biological system, a glia and a neuron are correlated each other. These cells closely relate and compose a higher brain function. In this study, we propose a Multi-Layer Perceptron (MLP) with local glia connection. We connect the glia to the neurons in two hidden-layer. The glias receive the outputs of the connecting neurons and the glias sum the neuron outputs. When the summed value is over the excitation threshold of the glia, the glia generates a pulse. After that, the pulse is input to the neuron threshold. Moreover, the excitation of the glia decreases the excitation threshold of neighboring glias. Thus, the neighboring glias are excited at a similar timing. We consider that the position relationship between the glia and the neuron is important to the MLP performance. By simulations, we confirm the influence of the glia for the MLP performance.

1. Introduction

The glia is one of nervous cells in the brain. This cell was considered to the static cell because the glia works in the brain was not observed. However, some researchers discovered that the glia has important functions for the brain [1]. The glia can transmit signals to the neurons and other glias by using various ions such as a glutamate acid, an adenosine triphosphate, a GABA, a Ca^{2+}, and so on [2][3]. Thus, the glia and the neuron are closely related each other. Currently, we consider that the glia is the important cell for the brain functions. In these ions, we notice the Ca^{2+}, because the glia generates the change of the Ca^{2+} concentration depending on the neurons. The change of the Ca^{2+} concentration is propagated to the neighboring glias, moreover it influences a membrane potential of neurons [4]-[6]. Thus, the glia and the neuron are closely related by the ions. From these characteristics, we have proposed the applications of the relationships between the neuron and the glia to an artificial neural network.

In this study, we propose the Multi-Layer Perceptron (MLP) with local glia connection which is inspired from the characteristics of the biological glia. This MLP is composed of four layers. We connect the glias to the neurons in two hidden-layer. One glia is connected with the neighboring three neurons in two hidden-layer. In this network, the glias is connected with same two neurons for the neighboring glias. The glia receives the output of the connecting neurons and charges these outputs. When the charged value

is over an excitation threshold of the glia, the glia generates the pulse. The pulse is input to the threshold of the nearest neuron in both layers. The glias shares the outputs of part of neurons with the neighboring neurons, thus the excitation timing of glias becomes similar to neighborhood. Moreover, the excited glia decreases the excitation thresholds of the neighboring glias. By these influences, the timing of the pulse generation is similar to each other. We consider that the glia gives the position relationships to the neurons. In the same layer, the timing the pulse generation is different little by little. Between the first-hidden-layer and the second-hidden-layer, the neurons have the same pulse in the same positions. By the simulations, we confirm that a solving ability of the proposed MLP.

2. Proposed method

The MLP is a famous feed forward neural networks. It is applied to a pattern classification, a pattern recognition, a data mining, and so on. The output of network is decided by the weights of the connections of the neurons. In general, we use a Back Propagation (BP) algorithm for determining weights [7]. In the standard MLP, the neurons do not have the connections in the same group, moreover the neurons connect with every neuron in the forward-layer. Thus, the connections of the neurons are equivalent each other.

In this study, we propose the MLP with local glia connection which is inspired from the biological characteristics of glia. The proposed model is shown in Fig. 1. We connect the glias to the neurons in the first-hidden-layer and the second-hidden-layer. One glia is connected with three neurons in each layer. The glias receive the outputs from the connecting neurons and sum these value. When the summed value is over the excitation threshold of glia, the glia generates the pulse. The pulse is input to the center neurons in the connecting neurons. The pulse influence to the threshold of neuron because the glia influence to the membrane potential in the biological system. Thus, the neurons in the first-hidden-layer and the second-hidden-layer have same pulse. Moreover, the excited glia decreases the excitation threshold of neighboring glia. The glias shared a part of neuron outputs as shown in Fig 1 and the glia directly influences to the excitation threshold of neighboring glia. Thus, the pulse generation pattern becomes similar in the neighborhood glias. By the

pulses which have the position relationships each other, the neuron works obtain the position relationships in the both hidden-layer.
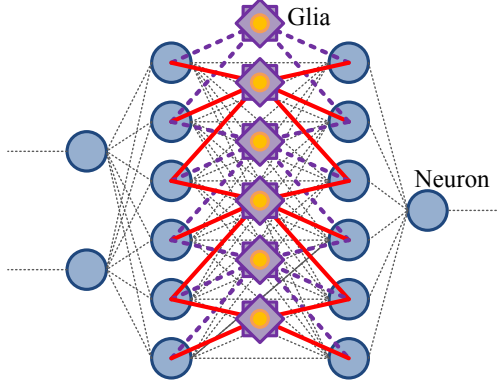

Figure 1: Proposed MLP.

2.1. Time flow of glia excitation

We show the detail of the glia works in Fig. 2. In this figure, we only show the neurons in one hidden-layer. Actually, the glias are connected with the neurons in two hidden-layer. The glias receives the outputs from two or three neurons. The neighboring glia receives the output from same neurons. At t=2, one glia is excited. Then, the glia generates the pulse and it input to the nearest neurons. Moreover, this glia influence to the neighboring glias. This influence decreases the excitation threshold of the glias. By this influence, the neighboring glias are excited at t=3. These glias generates the pulses and input to the nearest neurons. The excitation of the glia is propagated into the network. Thus, the generations of pulses are different little by little.
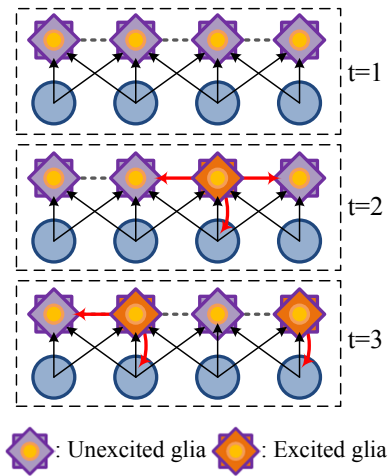

: Unexcited glia : Excited glia

Figure 2: The generation and propagation of glial effect.

2.2. Updating rule of neuron

The neuron is the multi-input and the single output. The neuron receives the outputs of neurons in forward layer. The standard updating rule of neuron is described 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 the connection, x is an input of the neuron, and θ is a threshold of the neuron. In this equation, w and θ are tuned by the BP algorithm. Thus, if the network is trapped into local minimum, it cannot escape out from there. Next, I show a proposed updating rule of the neuron. We add the glial effect to the threshold of neuron. This updating rule is used to the neurons in hidden layer. It is described by Eq. (2).

$$y_i(t+1) = f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t) + \alpha\psi_i(t)\right), \qquad (2)$$

where α is a weight of the glial effect and ψ is a pulse of the glia. We can control the glial effect by the change of α. In this equation, the weight of connection and the threshold are learned by the BP algorithm as same as the standard updating rule of neuron.

Eqs. (1) and (2) are used a sigmoidal function to an activating function which is described by Eq. (3).

$$f(a) = \frac{1}{1 + e^{-a}} \qquad (3)$$

where a is an inner state.

2.3. Glia

The glia is one of nervous cells. This cell transmits signal by using Ca^{2+}. The Ca^{2+} influences to the state of the neighboring glia and the membrane potential. In this study, the glia is excited by the neurons' outputs. One glia is connected with three neurons in each hidden-layer. The glia sums the received neurons' outputs and holds this value which is described by Eq. (4).

$$g_i(t) = g_i(t-1) + \frac{1}{3}\sum_{i=i-1}^{1+1}\left\{H_{1i}(t) + H_{2i}(t) - 1.0\right\}, \qquad (4)$$

where g is a hold value of the glia, H_{1i} is an output of the connecting neurons in the first-hidden-layer, and H_{2i} is an output of the connecting neurons in the second-hidden-layer. When g is over the excitation threshold of the glia, the glia generates the pulse. g can obtain the positive value and the negative value. If g is positive value, the glia generates the positive pulse. On the other hand, the glia generates the negative pulse. The excitation condition of the glia is defined in Eq. (5).

$$\psi_i(t+1) = \begin{cases} 1, & (g_i(t) > \theta_g \cap I_i > \theta_t) \\ -1, & (g_i(t) < -\theta_g \cap I_i > \theta_t) \\ \gamma\psi_i(t), & else, \end{cases} \qquad (5)$$

where ψ is a glia output, γ is an attenuated parameter, θ_g is an excitation threshold of the glia, I is a local time of the glia, and θ_t is a time length of a period of inactivity. The generation pulse is attenuated in an exponential fashion. The glia has the period of inactivity. The period of inactivity starts when the glia generates the pulse. During this period, the glia cannot have the response. In the proposed method, θ is decreased in the neighborhoods of the excited glia.

3. Simulations

In this section, we show the performance of the proposed MLP. We compare three kinds of MLPs which are;

(1) The standard MLP.

(2) The MLP with random noise.

(3) The MLP with local glia connection.

The standard MLP is a basic model. It can early converges the error, however it often trapped into a local minimum. The MLP with random noise has an uniformed random noise in the hidden-layer neurons. Every MLP is composed of four layers. The number of neurons in each layer are 2-10-10-1. We obtain the statistic result from 100 trials in each MLP. One trial has 5000000 iterations. We use the Mean Square Error (MSE) for an error index. The MSE is described by Eq. (6).

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (T_n - O_n)^2, \qquad (6)$$

where N is a number of learning data, T is a target value, and O is an output of MLP. From 100 trials, we obtain four kinds of the results which are an average of error, a minimum of error, a maximum error, and a standard deviation.

In this simulation, we use a Two-Spiral Problem (TSP) for a simulation task. The TSP is a famous task for the artificial neural network because this task cannot be a linear separation [8][9]. The TSP is composed of two different spirals. We input the coordinates of the spirals to the MLP and the MLP learns the corresponding classification. The learning spirals and the ideal result are shown in Fig. 3. The spirals are composed of the 194 points. The MLP learns the 194 data sets.

3.1. Learning performance

Firstly, we show the learning performance in Table 1. The minimum error becomes the same result for every MLP. However, the average error of the standard MLP is worst of all. Because it is often falls into the local minimum. The random noise improves the performance than the standard MLP. In this case, the maximum error becomes high. We consider that the random noise sometimes inhibits the learning. In the average of the error and the
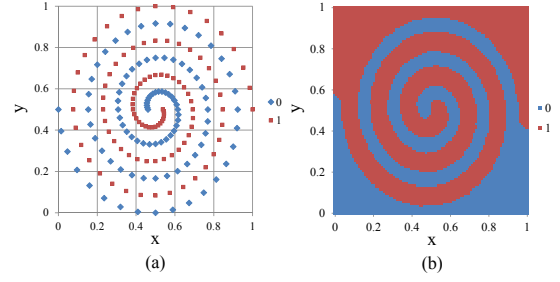
Figure 3: Two-spiral problem. (a) Learning spiral points. (b) Ideal result of TSP.

maximum error, the proposed MLP is the best of all. From this result, we can say that the proposed method gives the energy for escaping out from the local minimum. In our method, the timing of the pulse generation is similar to each other. Thus, the pulses locally give the large energy. However, the glias have period of inactivity. During this period, the pulse is not generated. The MLP locally increases the error, however it can rapidly converge the error during period of inactivity.

Table 1: Learning performance.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.01384	0.00001	0.08687	0.01959
(2)	0.01220	0.00000	0.09470	0.01772
(3)	0.00883	0.00001	0.07979	0.01463

Figure 4 is an example of learning curves. The standard MLP early converges the error. The oscillation of curve of MLP with random noise becomes a large. The random noise helps for escaping out from the local minimum, however the noise inhibits the convergence error. The proposed MLP rapidly decreases the error. When decreasing error becomes slower, the oscillation happens. Moreover, the error reducing is continued.
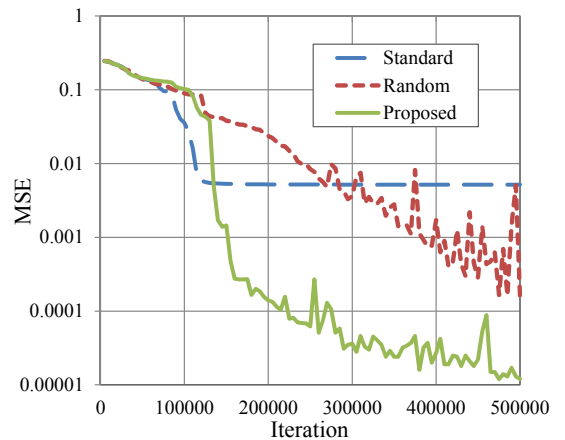
Figure 4: Examples of learning curves.

3.2. Classification performance

Finally, we show the classification performance in Table 2. The standard MLP is the worst of all. In this case, the MLP with random noise has a better performance than the proposed MLP. The MLP needs the generalization capability. In the proposed MLP, the pulse locally influences to the learning. During the period of inactivity, this learning becomes similar to the standard MLP. Thus, the proposed MLP lost the generalization capability than the MLP with random noise.

Table 2: Classification performance.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.14629	0.09879	0.22581	0.02469
(2)	0.12915	0.08278	0.18220	0.02132
(3)	0.13111	0.08725	0.18420	0.02008

4. Conclusions

In this study, we have proposed the MLP with local glia connection. The glias are connected with the neurons in the first-hidden-layer and the second-hidden-layer. The standard MLP does not have the relationships in the same layer, moreover the connections of the neurons are equivalence in the different layer. In our method, the glia generates the pulse according to the outputs of the connecting neurons. The glias share the outputs of part of neurons each other. The pulse is input to the neuron threshold. Moreover, the excited glia decreases the excitation threshold of neighboring glia. We consider that the sharing neurons' outputs and decreasing the excitation threshold gives the position relationships. By solving the TSP, we confirm that the proposed method is efficiency to the MLP performance. Moreover, we can see that the proposed MLP has a high convergence performance and a classification performance.

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, Mar. 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] S. Kriegler and S.Y. Chiu, "Calcium Signaling of Glial Cells along Mammalian Axons," *The Journal of Neuroscience*, vol. 13, pp. 4229-4245, 1993.

[5] M.P. Mattoson and S.L. Chan, "Neuronal and Glial Calcium Signaling in Alzheimer's Disease," *Cell Calcium*, vol. 34, pp. 385-397, 2003.

[6] G. Perea and A. Araque, "Glial Calcium Signaling and Neuro-Glia Communication," *Cell Calcium*, vol. 38, pp. 375-382, 2005.

[7] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.

[8] J.R. Alvarez-Sanchez, "Injecting knowledge into the Solution of the Two-Spiral Problem," *Neural Computing & Applications*, vol. 8, pp. 265-272, 1999.

[9] H. Sasaki, T. Shiraishi and S. Morishita, "High precision learning for neural networks by dynamic modification of their network structure," *Dynamics & Design Conference*, pp. 411-1–411-6, 2004.