

Investigation of Four-Layer Multi-Layer Perceptron with Glia Connections of Hidden-Layer Neurons

Chihiro Ikuta, Yoko Uwate, and Yoshifumi Nishio

Abstract—A glia is a nervous cell existing in a brain. This cell transmits signals by various ion concentrations. By the ion concentration, the glia correlates the neuron by ions. Thus, the glia composes the network different to the neural network. We proposed a four-layer Multi-Layer Perceptron (MLP) with glia connections of hidden-layer neurons which is inspired from characteristics of the biological glia. The proposed MLP has two hidden-layer. In this network, we connect the glia to the neurons in two hidden-layer. The glia receives the output of connecting neuron and the outputs are summed. When the summed value is over the glia excitation threshold, this glia is excited. The excited glia generates a pulse. This pulse input to the thresholds of connecting glia. We consider that the glia gives the energy to the MLP and the glia give the position relationship of neurons in the different layers. By the simulation, we confirm that the proposed MLP obtains a high solving ability. Moreover, we investigate the characteristics of the proposed MLP.

I. INTRODUCTION

BRAIN has two kinds of nervous cells which are a neuron and a glia. The neuron transmits electric signal each other. This work was found in an earlier research stage of the neuron. Thereby, the neuron has been investigated about biological characteristics and applications. On the other hand, the glia had not been investigated about details, because this cell was considered to a supportive cell of the neuron. Recently, some researchers discovered novel biological characteristics of the glia [1][2]. The novel biological characteristics that the glia can transmit signals by changing of ion concentrations such as a Ca^{2+} , an adenosine triphosphate, a glutamate acid, and so on [3]. Moreover, this cell has various receptors for these ions. Currently, the glia is considered that the glia influences neuron, thus the glia closely relates to the brain function. In these ions, we notice the Ca^{2+} . The glia generates the Ca^{2+} concentration wave according to the neuron excitation [4]. The Ca^{2+} propagates to a wide range in the brain and influence a membrane potential of neuron [5][6]. Thus, the glia composes the network different to the neural network. The neurons locally influence each other. The glia have wider relationships and compose the wide network than the neuron. We consider that the glia function can be applied to the artificial neural network.

In previous studies, we proposed some neuron-glia network [7][8]. We apply the characteristics of the biological glia to the Multi-Layer Perceptron (MLP). Because the neurons in the same layer are do not have a position relationship.

Chihiro Ikuta, Yoko Uwate, and Yoshifumi Nishio are with Department of Electrical and Electronics Engineering, Tokushima University, Japan (email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp).

In a biological system, the neuron composed of a cluster. We consider that the position relationship of the neurons is important for the artificial neural network. We connect the glia to the neurons in the hidden-layer. The glia and the neighboring glia influence each other. Moreover, the glia gives the position relationships to the connecting neurons.

In this study, we propose a four-layer MLP with glia connections of hidden-layer neurons. We give the glia to the hidden-layer neuron. In this model, we use the four-layer MLP, thus the MLP has two hidden-layer. The first hidden-layer neurons connects with every second hidden-layer neuron. The relationship of the neurons is equivalence in each position. We connect the glia to the hidden-layer neurons. In this model, the glia is connected with some neurons which exists in first hidden-layer and second hidden-layer. The glia receives the output of each neuron and these outputs are summed. When the sum of these outputs is over the excitation threshold of glia, the glia is excited. The excited glia generates the pulse. The pulse is input to the connecting neuron threshold. Thus, the neuron shares the same pulse. In a standard MLP, the neurons do not have the position relationships because these connections are equivalence. By our method, the neighboring neurons are connected by the glia pulse. The neurons obtain the position relationships of neurons in the hidden-layer. By the computer simulations, we show that the proposed model has a high solving performance than the standard model.

II. PROPOSED METHOD

The MLP is one of feed forward neural networks. It can be applied to a data mining, a pattern classification, a pattern recognition, and so on. In this model, a solution accuracy is depended on the number of neurons and a detention of nonlinear separation is depended on the number of layers. We can tune the output of this network by changing weight of connections between neuron. We often use a Back Propagation (BP) algorithm for decision of weights of connections. This algorithm is proposed by Rumelhart in 1986 [9]. In the BP algorithm, the error of network output is propagated to the backward. Then, the error decreases. However, this algorithm is often falls into the local minimum, because it is used to the steepest decent method. In general, the energy is given to the network for escaping out from local minimum. The energy helps for the escaping out from the local minimum. The energy is efficiency to the local minimum problem, however the learning of network becomes oscillatory. Moreover, the additional energy algorithm is difficult and it often needs a large computational cost.

In this study, we propose the four-layer MLP with glia connections of hidden-layer neurons. The proposed network is shown in Fig. 1. We connect the glia to the hidden-layer neurons in a cross different layers. The glia receives every connecting neuron output and these outputs are summed. When the sum of the neuron output is over the excitation threshold of the glia, the glia is excited. The excited glia generates the pulse which becomes positive response to the sum of neuron output. After that, the pulse is input to the neuron threshold because the glia influences the membrane potential of neuron in the biological system. The connecting neurons receive the same pulse at same time, thus these neurons work become similar on a short-term basis. We consider that the glia composes the local group of the neurons. In the standard model, the network learns at the same time. In our method, the network learns at different time in each group.

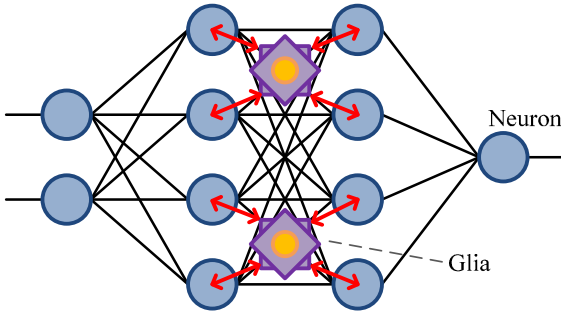


Fig. 1. Proposed MLP.

A. Updating rule of neuron

The neuron is multi-input and single output. The neuron receives the outputs of neurons in forward layer. The standard updating rule of neuron is described by Eq. (1).

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

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. In this equation, w and θ are tuned by the BP algorithm. Thus, if the network is trapped into local minimum, it cannot escape 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 neurons in the hidden-layer. It is described by Eq. (2).

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

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

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

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

where a is an inner state.

B. Back propagation algorithm

The error of MLP propagates backward in the network. The BP algorithm changes value of weights to obtain smaller error than before. The error of the network is given by Eq. (4).

$$E = \frac{1}{2} \sum_{i=1}^n (t_i - O_i), \quad (4)$$

where E is error value, t target value and O is the output of network. The error of network decrease by changing weights of connections according to the partial differentiation of error. Therefore, partial differentiation of the error is described by Eq. (5).

$$\Delta W = -\eta \frac{\partial E}{\partial w}, \quad (5)$$

where ΔW is an updated weight and η is a learning coefficient.

C. Glia

The glia is the nervous cell existing in the brain. This cell was considered to a static cell because we could not observed the work of glia. Recently, some researchers discovered that the glia transmits signals by ions concentrations according to the development observation technology. The glia uses various ions which are Ca^{2+} , glutamate acid, GABA, and so on. Among them, we notice Ca^{2+} . The glia is excited according to the neuron firing. The excited glia generates the Ca^{2+} concentration wave. The Ca^{2+} concentration wave transmit to the neighboring glia and neurons. The glia composes the network different to the neural network. We have considered that the glia can be applied to the artificial neural network.

In this study, we connect the glia to the hidden-layer neurons. We use four-layer MLP, thus this network has two hidden-layer. The glia is connected to the neurons in two hidden-layer. The glia receives the connecting neurons outputs. These outputs are summed and held according to Eq. (6).

$$g = \frac{1}{b-a} \sum_{i=a}^b (H_{1i} + H_{2i} - 1.0), \quad (6)$$

where g is a held output value, $b-a$ is the number of connecting neurons in the one layer, H_{1i} is a neuron output in the first hidden-layer, and H_{2i} is a neuron output in the second hidden-layer. When the held output value is over the threshold of excitation glia, the glia is excited. The excited glia generates the pulse which becomes the positive response. The pulse inputs to the connecting neuron threshold, thus, the connecting neurons share the same threshold change. The

glia has the period of inactivity. Although the held output value of glia is over the excitation threshold of glia, the glia cannot generate the pulse during the period of inactivity. The glia response is described by Eq. (7).

$$\psi_i(t+1) = \begin{cases} 1, & I = 0, (g > \theta_g \cap I > \theta_t) \\ -1, & I = 0, (g < -\theta_g \cap I > \theta_t) \\ \gamma\psi_i(t), & I = I + 1, \text{ else,} \end{cases} \quad (7)$$

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. In our method, the positive and negative pulses are generated according to g . The glia output is decreased in an exponential fashion. The glia response is decided by the parameters and the parameters fix during the iterations. However, the pulse is generated according to the neuron output. The neuron output is changed by learning. Thereby, the generation pattern of pulse is dynamically changed according to the learning. We show an example of the pulse generation in Fig. 2. From this figure, the neuron output is changed with time. The neuron output becomes zero when it excites the glia. We can see that the pulse is generated according to the neuron output, and the pulse generation density is changed. Because the neuron output is changed by learning. Moreover, the pulse value change from 1 to -1 . Every glia changes the pulse generation pattern similar to Fig. 2 during the iteration. The neuron and glia correlate each other. We consider that the characteristic pulse generation pattern is efficiency to the MLP learning.

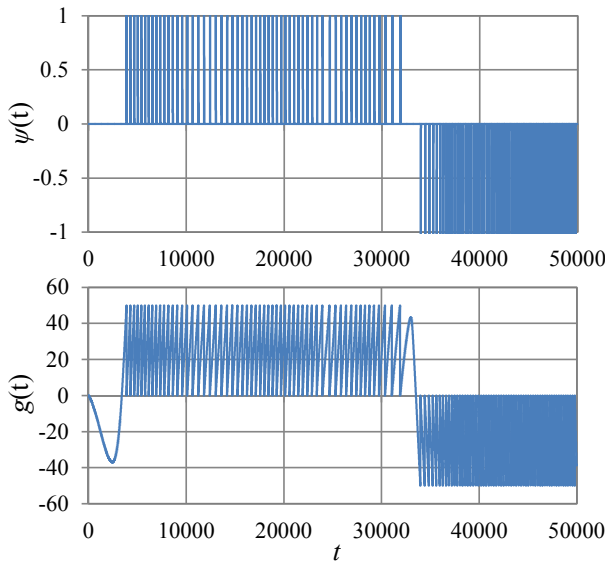


Fig. 2. An example of pulse generation.

III. SIMULATION

In this section, we show the simulation result. We use 4 kinds of the MLPs for comparison of the performance.

- (1) The standard MLP.
- (2) The MLP with random noise.

- (3) The four-layers MLP with glia connection of hidden-layer (The relationships between the neuron and glia in the same layer).
- (4) The four-layer MLP with glia connection of hidden-layer.

The standard MLP does not have the external unit. This method early converges the error, however it is often falls into local minimum. The MLP with random noise is input the uniformed random noise to every neuron in the hidden-layer. (3) and (4) uses the proposed method. In the case of (3), the glia is connected with the neuron in the same layer, thus the glia only receives neuron output from same layer and the generated pulse inputs to the neuron in the same layer. The glia which is connected with neurons in the first hidden-layer and the glia which is connected with neurons in the second hidden-layer are independent each other. (4) has the connection between first hidden-layer and second hidden-layer, thereby the neurons which connects the same glia receive the same pulse at the same time. Each MLP is composed of 4 layers. The input-layer has 2 neurons. The first hidden-layer has 10 neurons. The second hidden-layer has 10 neurons. The output-layer has 1 neuron. In the proposed MLP, the glia is connected with two neurons in one hidden-layer. Thus, one glia is connected with 4 neurons.

We evaluate the result by a Mean Square Error (MSE) which is shown as Eq. (8).

$$MSE = \frac{1}{N} \sum_{n=1}^N (T_n - O_n)^2, \quad (8)$$

where N is a number of learning data, T is a target value, and O is an output of MLP. We give 100 trials to each MLP and one trial has 500000 iterations. Every MLP has same initial conditions. From 100 trials, we obtain four kinds of indexes of MSE which are an average of error, a minimum error, a maximum error, and a standard deviation of error.

A. Simulation task

We use the Two-Spiral Problem (TSP) for a comparison of performance. The TSP is a famous task for the artificial neural network [10][11]. This task is linearly-inseparable problem, moreover this task has a high nonlinearity. In this task, the MLP is input the coordinates of spiral point and learns the classification of this point. The MLP learns every classification of spiral points. The learning spiral points shows Fig. 3 (a). Figure 3 (b) is the ideal result of TSP. It is calculated by norm between spiral points and coordinates.

B. Learning performance

First, we show the statistic learning performance of the MLPs. The results shows in Table I. The error average of standard MLP is the worst of all. Because this method often traps into the local minimum. The learning of the standard MLP is early converged. The MLP with random noise can find the best result of all. However, the error average is worth than the proposed MLP. In this model, every neuron in the hidden-layer received the uniformed random

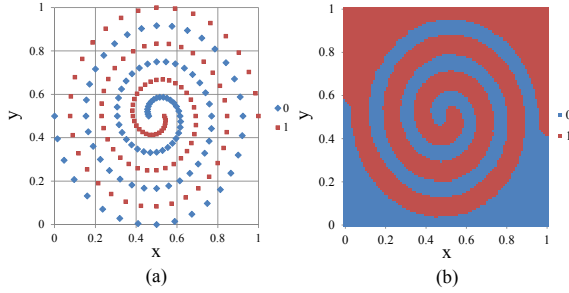


Fig. 3. Two-spiral problem. (a) Learning spiral points. (b) Ideal result of TSP.

noise. The doping noises does not have a correlations. We consider that the learning of this MLP is oscillatory. The two proposed model which receive the pulse from the glia. These performances are different. From this table, we can say that the position relationships of pulses are important for the MLP learning performance.

TABLE I
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)	0.01037	0.00002	0.07890	0.01427
(4)	0.00905	0.00001	0.07259	0.01234

We show an example of learning curve shown as Fig. 4. The learning curve of standard MLP converges earlier. However, it stops the decreasing error. The curve of MLP with random noise has small oscillations. Every neuron in the hidden-layer receives the random noise. Thus, the osculation is observed over all iterations. The proposed MLP has rapid change of error. We consider that the MLP receives the pulse at same time. The works of glias are independent each other. However, the pulses are stochastically generated at same time. Then, the MLP obtain a large energy from the glia, after that the MLP escapes out from the local minimum.

C. Classification performance

Next, we compare the generalization capability of the MLPs. The coordinates between 0 and 1 are input to the learned MLPs. We compare the ideal classification result (Fig. 3 (b)) with output of the network. The result show in Table II. In this result, the MLP with random noise is the best of all. In the classification, the MLP needs the wide area search ability. The neurons in the hidden-layer has the decorrelation noise. Thus, the MLP with random noise can search the wide area solutions. The performance of proposed MLP is worth than the MLP with random noise. The neurons receive the same pulse, thus the learning of the neurons in the hidden-layer is low variability.

Figure 5 is the classification image by each MLP. We use the best results of each MLP in Table II. The standard MLP has some deficit spaces. The MLP with random noise and

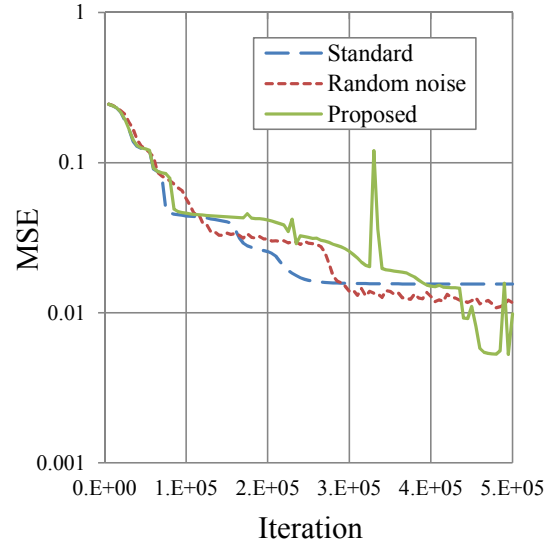


Fig. 4. Examples of learning curves.

TABLE II
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)	0.13963	0.09578	0.20769	0.02576
(4)	0.13422	0.08517	0.20657	0.02208

the proposed MLP (The relationships between the neuron and glia in the same layer) have one deficit space. In the MLP with random noise has a obvious error in a bottom part of image. The proposed MLP does not have the deficit space different from others. From these images, we can say that the MLP with random noise has a high performance in local part of image. On the other hand, the proposed MLP can classifier to two spiral in the wide space.

D. Parameter characteristics

Finally, we show the parameter characteristics of the proposed MLP. We change the parameters of the glia. The horizontal line means weight of glial effect (α). The vertical line means the MSE. We change γ from 0.7 to 0.95. Each graph shows the result by different attenuated parameter (γ). If γ is higher, the decreasing of the pulse becomes slow. Figures 6-8 have different θ_t . We cannot observe the high relationships of weight of glial effect and attenuated parameter to the learning performance. In the case of classification performance, it has a high dependency to the weight of glial effect and attenuated parameter to the learning performance.

In Fig. 6, the error of learning performance increases for $\gamma = 0.95$. $\theta_t = 30$ means that the glia can rapidly finish the period of inactivity, thus the new pulse is generated when the pulse is not enough to decrease. The classification performance improves as increasing α . When α is over 0.3, the performance is decreased. However, the error is large

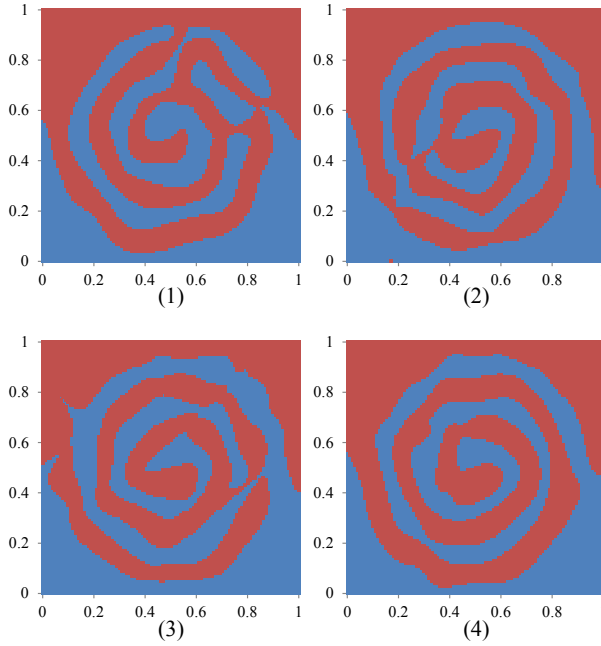


Fig. 5. Classification images.

when $\gamma = 0.95$. From two results, the new pulse should be generated when the pulse is enough decreased.

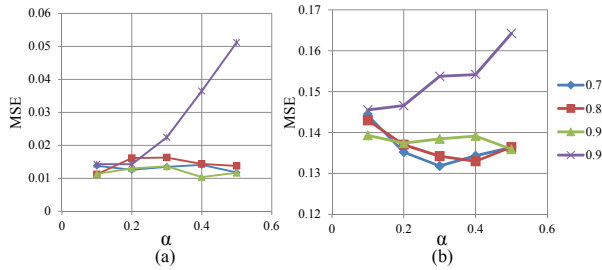


Fig. 6. Parameter characteristics between weight of glial effect (α) and the MSE ($\gamma = 0.7, 0.8, 0.9, 0.95, \theta_t = 30$). (a) Learning performance. (b) Classification performance.

Figure 7 shows the result for $\theta_t = 40$. The error of learning is increased for comparison with Fig. 6. We consider that the θ_t influence to the learning performance. In this case, dependency for α and γ is smaller than the previous result. From this result, we consider that the repetition of pulse is important for the MLP learning performance. The classification performance improves than the previous result in many conditions.

Figure 8 is the case of $\theta_t = 50$. We cannot observe the dependency to α and γ in the learning performance. However, the classification performance is similar to Fig. 7 (b). From these results, the learning performance has high θ_t dependency and the classification performance has a high α and γ dependency.

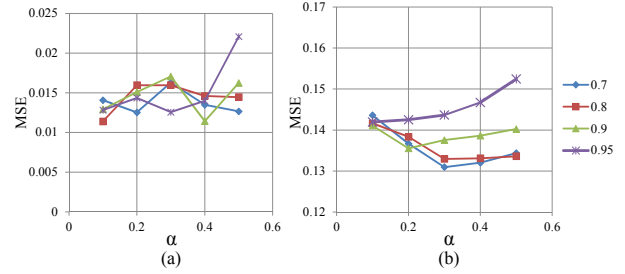


Fig. 7. Parameter characteristics between weight of glial effect (α) and the MSE ($\gamma = 0.7, 0.8, 0.9, 0.95, \theta_t = 40$). (a) Learning performance. (b) Classification performance.

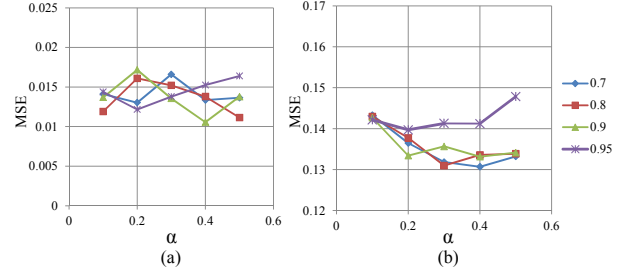


Fig. 8. Parameter characteristics between weight of glial effect (α) and the MSE ($\gamma = 0.7, 0.8, 0.9, 0.95, \theta_t = 50$). (a) Learning performance. (b) Classification performance.

IV. CONCLUSIONS

In this study, we have proposed four-layer MLP with glia connections of hidden-layer neurons. In this method, the MLP is composed of four neuron layers. We connect the glia to the neurons in two hidden-layer. The glia gives the relationships between the neurons in the first hidden-layer and the neurons in the second-hidden layer. The glia receives the outputs of connecting neurons. These outputs are summed and held by the glia. The glia is excited by the summed value of neuron output when this value is over the excitation threshold. The excited glia generates the pulse and this pulse is input to the connecting neurons thresholds. Thus, the connecting neurons have same pulse at same time. We consider that the glia pulse give the position relationships of the neurons in the hidden-layer. By solving the TSP, we confirmed that the proposed MLP has better performance than the standard MLP. Moreover, we investigated the parameter characteristics. We confirmed that the learning performance and the classification performance of the proposed MLP are depended to different parameters.

ACKNOWLEDGMENT

This work was partly supported by MEXT/JSPS Grant-in-Aid for JSPS Fellows (24-10018).

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. Kriegler and S.Y. Chiu, "Calcium Signaling of Glial Cells along Mammalian Axons," *The Journal of Neuroscience*, vol. 13, pp. 4229-4245, 1993.
- [4] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.
- [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] 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] C. Ikuta, Y. Uwate, Y. Nishio, and G. Yang, "Multi-Layer Perceptron Decided Learning Neurons by Regular Output Glia," *Proc. NOLTA'12*, pp. 719-722, Oct. 2012.
- [9] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.
- [10] J.R. Alvarez-Sanchez, "Injecting knowledge into the Solution of the Two-Spiral Problem," *Neural Computing & Applications*, vol. 8, pp. 265-272, 1999.
- [11] 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.