2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Investigation of Multi-Layer Perceptron with Pulse Glial Chain Based on Individual Inactivity Period

Chihiro Ikuta, Yoko Uwate, and Yoshifumi Nishio

***Abstract*— In this study, we propose a Multi-Layer Perceptron (MLP) with pulse glial chain based on individual inactivity period which is inspired from biological characteristics of a glia. In this method, we one-by-one connect a glia with neurons in the hidden-layer. The connected glia is excited by the connecting neuron output. Then, the glia generates the pulse. This pulse is input to the connecting neuron threshold. Moreover, this pulse is propagated into the glia network. Thus, the glia has a position density each other. In this network, a period of inactivity of the glia is dynamically changed according to pulse generation time. In the previous method, we fix the period of inactivity, thus the pulse generation pattern is often fixed. It is similar to the local minimum. By varied the period of inactivity, the pulse generation pattern obtains the diversity. We consider that this diversity of the pulse generation pattern is efficiency to the MLP performance. By the simulation, we confirm that the proposed MLP improves the MLP performance than the conventional MLP.**

I. Introduction

A Human brain has two kinds of nervous cells which are the neuron and the glia. We have considered that a human cerebration is only made by the neurons. Because the neuron can transmit an electric signal each other and this phenomenon was found at an earlier stage of a research. Actually, the transmission of the electric signal has a high relationship for the human cerebration and it achieved some positive results. On the other hand, we considered that the glia was a support cell for the neuron. However, some researchers discovered that the glia has novel glia functions [1][2]. The glia can transmit signal by using ions concentrations which are a glutamate acid, an adenosine triphosphoric acid (ATP), calcium (Ca^{2+}), and so on [3][4]. These ions are also used in a gap junction of the neuron. Among them, the Ca^{2+} is important for the transmission of information between the glia. The concentration change of the Ca^{2+} induces the stimulus from the neuron. The Ca^{2+} propagates to the other glias. The glia is considered that the glia and the neuron closely related. Moreover, the glia makes the different network from the neuron. Currently, we should consider to a network between the neuron and the glia.

The glia-neural network is important for a detailed investigation of the brain works. However, the brain research is mainly about the neuron. Especially, the application of the glia has not almost investigated. We therefore applied the glia characteristics to a Multi-Layer Perceptron (MLP) for the application of the glia. The MLP is a famous artificial neural network. This network is composed of layers of neurons. The MLP is generally learned by a Back Propagation (BP) algorithm [5]. By this learning, the MLP can be applied to a pattern learning, a data mining, and so on. However, the BP algorithm has a local minimum problem because this learning algorithm uses the steepest decent method. The MLP does not have the connections in the same layer. The neurons connect to different layer of neurons thus the neurons do not correlate in the same layer. In the previous study, we proposed the MLP with pulse glial chain in IJCNN'12 [6]. In the previous model, we connect the glia with the neurons for solving these problems. The glia is connected with the neurons in the hidden-layer and the neighboring glias, and it generates the pulse according to the connecting neuron output. The generated pulse is propagated to the connecting neuron and the other glias. We consider that the glia pulse gives position relationships of the neuron in the hidden-layer and an energy for escaping out from the local minimum. From the previous study, we confirmed that the previous model has a better performance than the standard MLP. However, the previous model has a problem. This problem is that the pulse generation pattern is often converged in the previous model. Every glia has the same parameters. Thereby whole pulse generation pattern is depended on the one glia influence.

In this study, we propose the MLP with pulse glial chain based on individual inactivity period. We introduce the individual period of inactivity to each glia. If the glia is excited by the connecting neuron output, the glia cannot be excited again during the period of inactivity. The previous model has same time length of the period of inactivity, thereby the generation pulse pattern becomes the same cycle. In this method, the time length of the period of inactivity is varied to a short when the glia is continuously excited. The glia which is excited at short interval, obtains different pulse generation cycle. We consider that the varying the period of inactivity breaks the periodic pulse generation. The network learning obtains the diversity. By the computer simulation, we show that the pulse generation pattern becomes the diversity. Moreover proposed network has a better performance than the conventional method.

II. Proposed Method

In this study, we propose the MLP with pulse glial chain based on individual inactivity period as shown in Fig. 1. We connect the glias to the neurons in the hidden-layer. The glia makes the different network from the neural network.

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).

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

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

Firstly, the glia receives the connecting neuron output. If it is over the excitation threshold of the glia, the glia is excited. The excited glia generates the pulse. This pulse can have a negative value and a positive value. It is depended on the connecting neuron output. After that, the pulse is input to the connecting neuron threshold. Moreover, the pulse influences to the neighboring glias. The neighboring glias are also excited by this pulse independent from the connecting neuron output. Thus, the pulse is propagated into the glia network. The pulse gives the energy to the network, because the glia pulse is independent from the network learning. Moreover, pulse propagation gives the position relationship with each neuron in the hidden-layer. The pulse generation time is similar each other. In the previous method, we fix the period of inactivity. The period of inactivity decides the cycle of the pulse generation. Then the pulse generation often became the periodic. We consider that it reduces the possibility of escaping out from the local minimum. In the proposed method, we vary the period of inactivity according to the glia excitation. When the same glia is continuously excited by the connecting neuron, the period of inactivity of this glia becomes a short. The glia obtains the different period of inactivity each other with time. Thus, this glia exits the periodic pulse generation because the neighboring gila does not finish the period of inactivity when this glia finishes the period of inactivity.

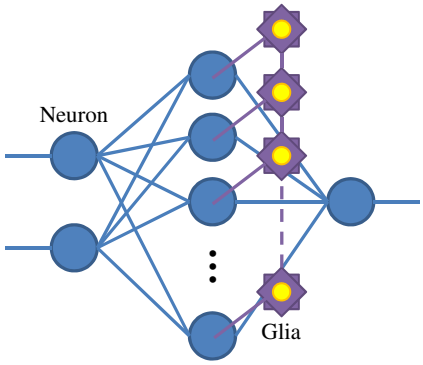


Fig. 1. MLP with pulse glial chain based on individual inactivity period.

A. Glia response

The glia has two different states which are the positive response and the negative response. We define the output function as the positive response of the glia in Eq. (1).

$$\psi_i(t+1) = \begin{cases} 1, & \{(\theta_n < y_i \cup \psi_{i+1,i-1}(t - i * D) = 1) \\ & \cap (\tau_i \geq \theta_{gi})\} \\ \gamma\psi_i(t), & else, \end{cases} \quad (1)$$

where ψ is an output of a glia, i is a position of the glia, θ_n is a glia threshold of excitation, y is an output of a connected neuron, D is a delay time of a glial effect, τ is local time of the glia during a period of inactivity, θ_g is a length of the period of inactivity, γ is an attenuated parameter. In the proposed method, the length of the period of inactivity is varied according to the pulse generation. If the glia is continuously excited by the connecting neuron output, the length of the period of inactivity becomes a short. Moreover, if the glia is excited by the neighboring glia pulse, the period of inactivity of this glia returns to original time length of the period of inactivity. Figure 2 shows the two different pulse generation. In the upper figure, the pulse generation cycle becomes short with time. The bottom figure has periodic pulse generation. In the glia network, the both glias exist at one time. Thereby, we consider that the pulse generation pattern is dynamically changed in the proposed method.

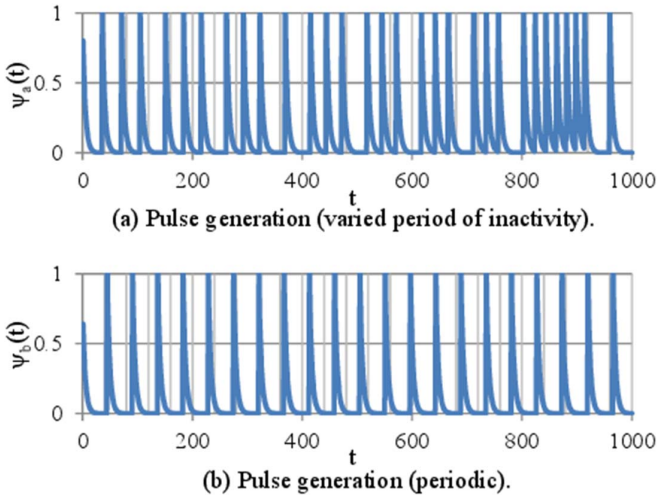


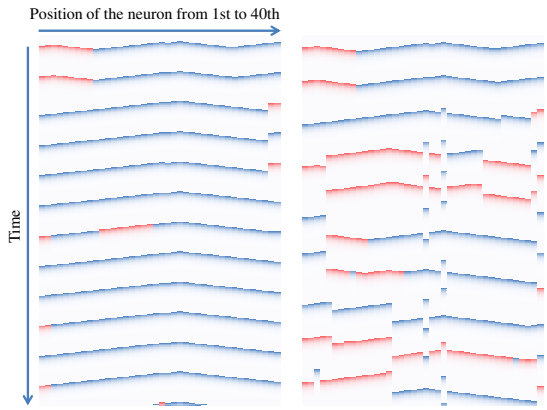
Fig. 2. Varying period of inactivity. (a) The length of the period of inactivity becomes short with time. (b) Periodic pulse generation.

B. Pulse propagation

Figure 3 shows an example of the pulse generation and a propagation. In this figure, some glias are excited and pulse generates. If the glia receives the large output of the connecting neuron, this glia generates the positive pulse. If the glia receives the small output of the connecting neuron, this glia generates the negative pulse. The red part shows the negative value pulse, the blue part shows the positive value pulse. After that this pulses are propagated to the other glias. Both pulse generations are similar pattern at first. In the case of (a), we can observe a small change of the pulse generation pattern. The pulse generation pattern is fixed with time. On the other hand, the pulse generation pattern (b) piecemeal varies from (a). Moreover, the pulse generation pattern (b) varies for a long time than (a). From the figure, the proposed network breaks the periodic pulse generation and makes the diversity.

C. Updating rule of neuron

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



(a) Previous pulse generation (b) Proposed pulse generation

Fig. 3. Pulse generation and propagation. (a) The pulses are generated by the previous glia network. (b) The pulses are generated by the proposed glia network.

defined by Eq. (2).

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

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, the weight of connection and the threshold of the neuron are learned by BP algorithm. Thus, the neuron output is depended on the BP learning. Next, we show a proposed updating rule of the neuron. We add the glial pulse to the threshold of neuron. We use this updating rule to the neurons in the hidden layer. It is described by Eq. (3).

$$y_i(t+1) = f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t) + \alpha\psi_i(t)\right), \tag{3}$$

where α is a weight of the glial effect. We can change the glial effect by change of α. In this equation, the weight of connection and the threshold are changed by BP algorithm as same as the standard updating rule of the neuron. However, the glial effect is not changed. It is updated by Eq. (1).

Equations (2) and (3) are used a sigmoidal function to an activating function which is described by Eq. (4).

$$f(a) = \frac{1}{1+e^{-a}} \tag{4}$$

where a is an inner state.

III. Simulations

We compare five kinds of the MLPs;

(1) The standard MLP.
(2) The MLP with random noise.
(3) The MLP with pulse glial chain.
(4) The MLP with pulse glial chain based on individual inactivity period (The period of inactivity is random.).
(5) The MLP with pulse glial chain based on individual inactivity period (The period of inactivity is varied according to the pulse generations.).

The network of (1) does not have the external unit, thus this network is often falls into local minimum. The network of (2) noise has an uniformed random noise. The network of (3) has same period of inactivity in every glia. In the (4), every glia has different the length of the period of inactivity which is decided at random. Every MLP has same number of neurons and layers. The MLP is composed of 2-40-1 neurons. We obtain the experimental result from 100 trials. Every trial has different initial conditions. One trial has 50000 iterations. We use Mean Square Error (MSE) for the error function. The MSE is described by Eq. (5).

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

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

A. Simulation task

We use a Two-Spiral Problem (TSP) for the simulation task which is shown in Fig. 4. The TSP is famous task for the artificial neural network [7][8]. It has high nonlinearity. Thus, the standard MLP often falls into the local minimum. In this task, we input the spiral coordinates to the MLP (shown as Fig. 4). After that the MLP learns the classification of the spiral. We change the number of spiral points (98 and 130 points) and obtain the result from two simulations. Figure 5 shows that a classification of $x-y$ plane which is obtained from a norm of each coordinate.

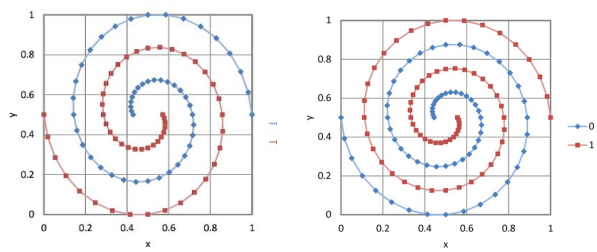


Fig. 4. Two-Spiral Problem.

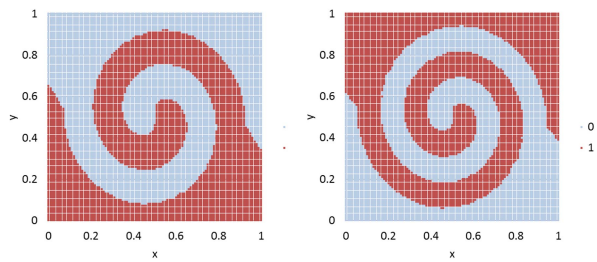


Fig. 5. Classification of Two-Spiral Problem in $x-y$ field.

B. Simulation results

1) The number of spirals are 98: Firstly, we use the 98 spiral points to the learning of the MLP. The learning performance means the fitting between the output of the MLP and the supervised classification. Table I shows the experimental result of the learning performance. Every method improves the performance than the standard MLP. From this result, we can see that the proposed MLP has a three times better performance than the MLP with pulse glial chain. The MLP with pulse glial chain and proposed MLP have the better performance than the MLP with random noise, thus the pulse is efficient to the MLP learning. Moreover, we consider that the pulse generation pattern is important to the MLP learning.

TABLE I
LEARNING PERFORMANCE OF SPIRAL OF 98 POINTS.

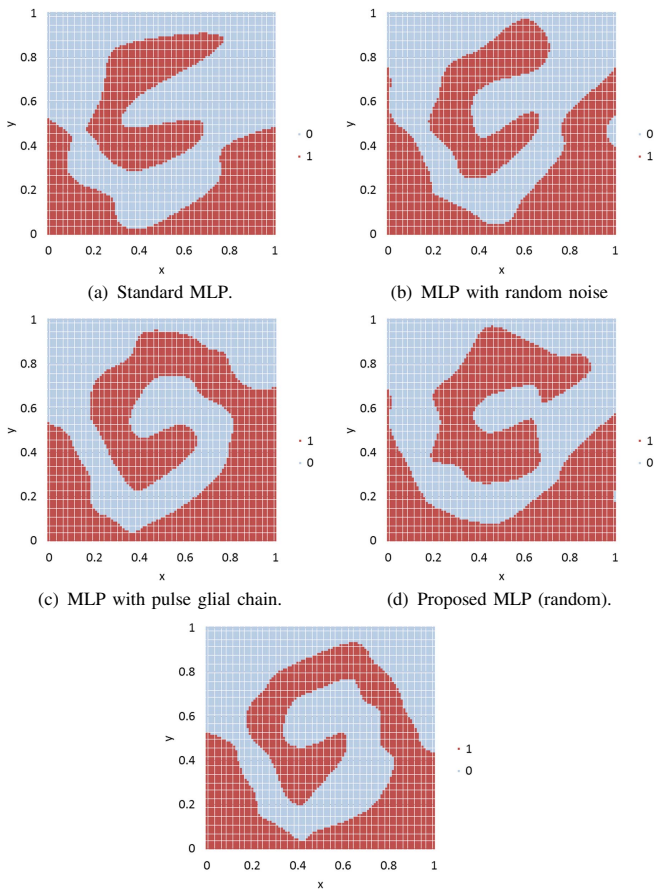
	Average	Minimum	Maximum	Std. Dev.
(1)	0.04153	0.00017	0.18387	0.02637
(2)	0.03711	0.00006	0.17352	0.02946
(3)	0.01531	0.00009	0.06157	0.01636
(4)	0.01791	0.00016	0.18380	0.02415
(5)	0.00444	0.00016	0.04151	0.00956

Table II shows the classification performance. We input the unlearning coordinates to the MLP which finishes the learning. After that we obtain the output of the MLP in correspondent of the input coordinates. We compare the true classification and the output of the MLP. The true classification is obtained from norm between the input classification and the learning spiral coordinate. The trend of the results is similar to the learning performance. We can see that the proposed MLP is only under 0.1 in the average of error. In the learning performance, the proposed MLP has a high ability. In general, the MLP becomes the over fitting when it has too much learning, because the MLP falls the deep local optimum solution. However the proposed MLP can classify the unknown data than the others, it means that the proposed MLP has a high generalization capability. From this result, our method can find a better solution. Moreover it can search a wide range of a solution space.

TABLE II
CLASSIFICATION PERFORMANCE OF SPIRAL OF 98 POINTS.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.15029	0.08085	0.21127	0.02434
(2)	0.13966	0.08083	0.20378	0.02879
(3)	0.10980	0.06408	0.15069	0.01902
(4)	0.11647	0.07176	0.17159	0.02310
(5)	0.09565	0.06188	0.17970	0.01773

Figure 6 shows the classification of unknown coordinates when the MLP learns the 98 spiral points. We obtain these figures from the near average result in Table II. The standard MLP, the MLP with random noise, and the MLP with pulse glial chain based on individual inactivity period (The period of inactivity is random.) cannot draw the spirals. These MLPs have crack at the periphery of $(x, y) = (1, 0.5)$. The MLP with pulse glial chain can draw the two spirals, however it also has some errors at the periphery of $(x, y) = (1, 0.5)$. A border value of the two spirals becomes about $(x, y) = (1, 0.7)$. On the other hand, our proposed MLP can obtain the two spirals in the field, moreover it does not have the large error in every area.



(a) Standard MLP. (b) MLP with random noise (c) MLP with pulse glial chain. (d) Proposed MLP (random). (e) Proposed MLP.

Fig. 6. Classification of two spirals of 98 points for unknown coordinates.

2) The number of spirals are 130: Secondly, we show the learning performance of the spirals of 130 points. Of course, the TSP becomes difficult by increasing the number of the spiral points. In this case, the number of turns is also improves, thus this task has stronger nonlinearity than the previous task. The statistic result shows in Table III. We can see that the standard MLP often traps into the local minimum. Thereby, the average of error is the worst of all. The result of the MLP with random noise is similar to the standard MLP. From this result, the uniformed random noise is not efficient to the TSP. Other three MLPs improve the performance from the result of the standard MLP. Especially, the MLP with pulse glial chain and the proposed MLP have a good learning performance. Moreover, the maximum error of the proposed MLP is the best of all. From this result, we can say that the proposed MLP has a high ability for escaping out from the local minimum. Thereby, our proposed MLP reduces an initial valued dependence. It means that we can stably obtain

the better result.

TABLE III
LEARNING PERFORMANCE OF SPIRAL OF 130 POINTS.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.12269	0.00831	0.23857	0.05554
(2)	0.10847	0.00047	0.24278	0.05742
(3)	0.01990	0.00067	0.11664	0.02226
(4)	0.05546	0.00134	0.14481	0.03608
(5)	0.01414	0.00052	0.04851	0.01313

Next, we show the classification performance of the MLPs. The classification results show in Table IV. The trend of the simulation results is similar to the learning performance. The standard MLP and the MLP with random noise are worse results. The classification performance of the proposed MLP is the best of all in every index.

TABLE IV
CLASSIFICATION PERFORMANCE OF SPIRAL OF 130 POINTS.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.21782	0.10565	0.29477	0.03858
(2)	0.19278	0.10460	0.33065	0.04434
(3)	0.12538	0.08027	0.19639	0.02625
(4)	0.15334	0.09368	0.24328	0.02948
(5)	0.11857	0.06876	0.19142	0.02473

Figure 7 shows learning curves of each MLP. The error reduction of the standard MLP converges at 25000. It is trapped into the local minimum. The convergence of the error in the MLP with random noise is a slower than the others. However, it reduces the error than the standard MLP. The uniformed random noise has a small efficiency to the learning of the MLP. On the others, these curves have a oscillation during the iterations. Moreover the performance of the error reduction improves. The pulse locally gives the large energy to the network. The pulse helps escaping out from the local minimum. The glia has the period of inactivity. During the period of inactivity, the glia does not generate the pulse again. Thereby, the MLP can search the better solution during the period of inactivity. The error reduction of the proposed MLP is earlier than the others. Thus, the pulse generation pattern influences the learning of the MLP.

By using this result, we compare the proposed MLP with the MLP with pulse glial chain. The proposed MLP has a better performance than the MLP with pulse glial chain, however a difference of superiority in the statistic result is not observed. Here, we show the error reduction curves of the proposed MLP and the MLP with pulse glial chain (shown as Fig. 8) which is obtained from an average error at each iteration. We can see that the error of the proposed MLP rapidly decreases from a start of the learning, moreover the error converges the learning earlier than that of the MLP with pulse glial chain. We consider that it is an influence of changing period of inactivity, because the proposed MLP can vary the pulse generation pattern by changing the period of inactivity. We consider that the influence of the pulse glial chain becomes small by convergence of the pulse generation

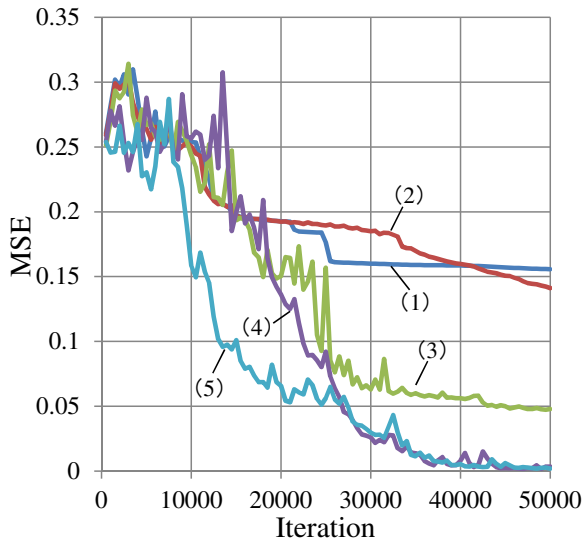


Fig. 7. Learning curve.

pattern. Actually, the MLP with pulse glial chain early fix the pulse generation pattern, thereby the error reduction becomes gradual with temporal progress.

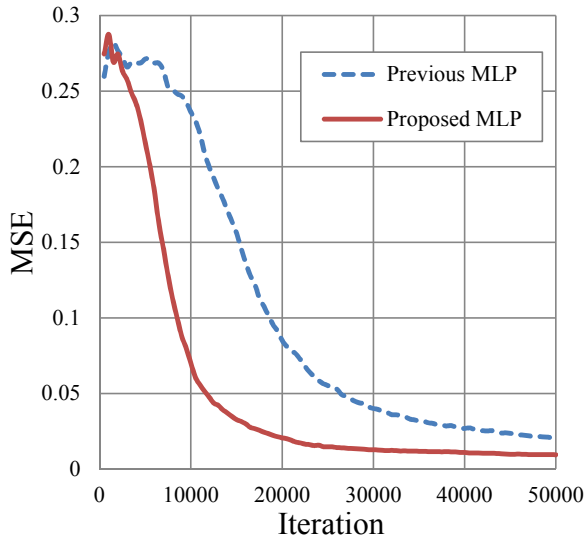
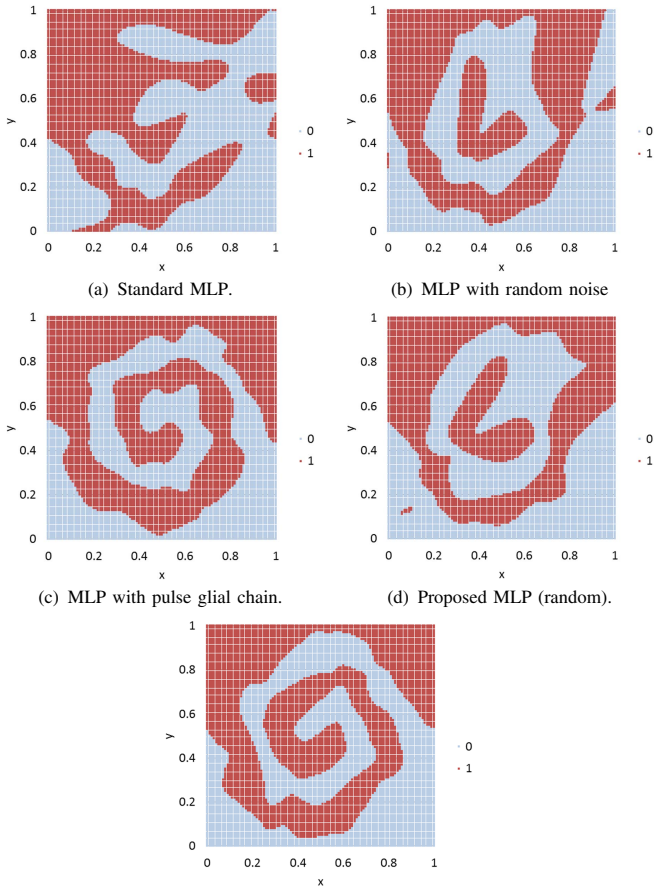


Fig. 8. Comparison of the convergence of the proposed MLP and the MLP with pulse glial chain.

Finally, we show the classification image of the TSP as shown in Fig. 9. We obtain the classification image from average result in Table IV. The standard MLP cannot draw the spirals, thus the solving ability of the standard MLP is unsatisfactory for the TSP. We can observe the outside circle of the spiral in the MLP with random noise however it has any cracks. The MLP with pulse glial chain can classifier the spirals. The MLP with pulse glial chain based on individual inactivity period (the period of inactivity is random) has two cracks. This model is similar to the proposed MLP however its performance is worse in every result. The proposed MLP can also separate the two spirals. Moreover, the outside curve

is better than the MLP with pulse glial chain. Actually, the MLP with pulse glial chain has error near coordinates (0.0, 0.5).



(a) Standard MLP.

(b) MLP with random noise

(c) MLP with pulse glial chain.

(d) Proposed MLP (random).

(e) Proposed MLP.

Fig. 9. Classification of two spirals of 130 points for unknown coordinates.

IV. Conclusions

In this study, we have proposed the MLP with pulse glial chain based on individual inactivity period. We connect the glia to the neuron in the hidden-layer. The glia receives the connecting neuron output. The glia generates the pulse when the neuron output is over the excitation threshold of the glia. This pulse is input to the connecting neuron threshold and moreover it is propagated to the neighboring glias. In this method, the period of inactivity is varied according to the pulse generation time. If the pulse generation continuously occurs by the connecting neuron output, the period of inactivity becomes short. By this influence, the pulse generation pattern is dynamically changed because the period of inactivity of the glia is different each other. We consider that the glia pulse improves the MLP performance. Actually, we confirm that the proposed MLP has a better performance than the conventional MLP by the computer simulation.

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, 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, 375-382, 2005.

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

[6] C. Ikuta, Y. Uwate, and Y. Nishio, "Multi-Layer Perceptron with Positive and NegativePulse Glial Chain for Solving Two-Spirals Problem," *Proc. IJCNN'12*, pp.2590-2595, Jun. 2012.

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

[8] 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.