

Performance and Features of Multi-Layer Perceptron with Impulse Glial Network

Chihiro Ikuta, Yoko Uwate and Yoshifumi Nishio

Abstract—We have proposed the glial network which was inspired from the feature of the brain. The glial network is composed by glia connecting each other. All glia generate oscillations and these oscillations propagate in the glial network. We confirmed that the glial network improved the learning performance of the Multi-Layer Perceptron (MLP).

In this article, we investigate the MLP with the impulse glial network. The glia generate only impulse output, however they make the complex output by correlating with each other. We research the proposed networks' parameter dependency. Moreover, we show that the proposed network possess better learning performance and better generalization capability than the conventional MLPs.

I. INTRODUCTION

CURRENTLY, various artificial neural network models are proposed. They mostly imitate a biological brain system and are useful to solve difficult tasks. The Multi-Layer Perceptron (MLP) is one of feed forward neural networks and is useful to perform several tasks, for example, pattern recognition, pattern classification, data mining and so on. Back Propagation (BP) algorithm is a learning algorithm for the MLP using the steepest decent method [1]. The MLP can solve the linearly-inseparable problem by using this learning algorithm. However, when we use this learning algorithm, we must be careful to the local minimum problem. The MLP can not often find the global solution from the influence of this problem.

We have proposed the glial networks to improve the performance of MLP [2][3]. The glial network was inspired from the feature of the glia which are existing in the brain. The glia were not attract researchers' attentions for a long time, because the glia have been believed not to use electrical signals. Recently, some researchers discovered that the glia transmit signals by using ions [4]. In the brain, the glia influence each other and the ions affect neurons' thresholds [5][6]. We tried to exploit the glia network's behavior such that another network located close to the neural network helps the functions of the neural networks. In [2], we proposed the MLP with the glial network whose glia generated independent oscillations and these oscillations propagated neurons and other glia. We confirmed by computer simulations that the glial network improved the learning performance of the MLP by connecting the neurons more effectively than the conventional networks.

Chihiro Ikuta, Yoko Uwate and Yoshifumi Nishio are with the Department of Electrical and Electronic Engineering, The Tokushima University, 2-1 Minami-Josanjima, Tokushima, Japan, (email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp).

In this paper, we propose a MLP with the impulse glial network. In the biological glia, the glia generate Ca^{2+} density wave which is known to like impulse [7]. If the glia are excited, they rapidly change the Ca^{2+} density in the brain. After that, the Ca^{2+} density return to a normal state with time [8][9]. In order to confirm the relationship between glia and neurons, we investigate the impulse glial network. In our network, all glia generate impulses when the glia is excited by the connecting neuron. The glia connects the neuron and watch to neuron's output. If the neuron's output becomes to high, the glia generates an impulse. After that, this impulse is propagating in the glial network, this glia can not generate the impulse during constant time like biological neuron's refractory period. The glia generate simplex impulse output, however, they have the glial network. They can make the complex output by correlating each other. We consider that all glia generate the impulses at random times each other, the glia's effects give good influence to the MLP. We show the performance and features of proposed MLP with the impulse glial network by the computer simulations.

II. MLP WITH IMPULSE GLIAL NETWORK

The MLP is the most famous feed forward neural network. Several methods using the MLP have been proposed for solving many kinds of tasks. This network has some neuron layers and the weights between the layers are learned by the BP algorithm. In this study, we use the MLP with three layers (4-10-1) and a glial network connected to the hidden layer. Figure 1 shows an example of the MLP with the glial network.

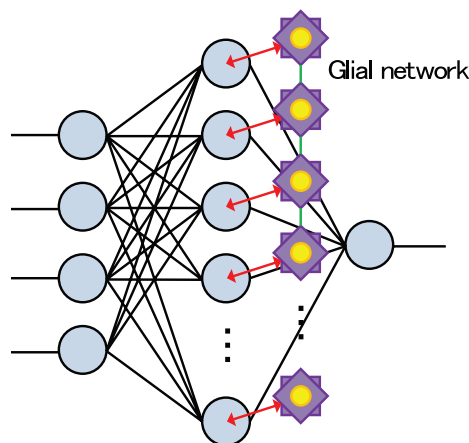


Fig. 1. MLP with impulse glial network.

A. Neuron Updating Rule

The standard neuron updating rule is given by Eq. (1).

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

where x : input or output, w : weight parameter, θ : threshold and f : output function. The parameters w and θ are learned by using the BP algorithm.

The updating rule of the hidden layer's neurons of the proposed neural network with the glial network is modified as Eq. (2).

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

where Ψ : output of the glia, α : weight of glia outputs. We use the sigmoid function for the output f as Eq. (3).

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

B. Impulse Glial Network

In the biological neural network, it is known that the glia affect the neighboring neurons over a wide range by making their outputs to be propagated in the network.

The output of the glia can be given as Eq. (4).

$$\Psi_i(t) = \sum_{k=-m}^m \beta^{|k|} \psi_{i+k}(t - |k|), \quad (4)$$

where ψ is the glia making impulse, β denotes attenuation parameter and m is propagating range of glia's impulse. We defined the glia's impulse, it is given by Eq. (5).

$$\begin{aligned} \psi_i(t+1) &= 1, \quad (\theta_n < O_i) \cap (\theta_g > \psi_i(t)) \\ \psi_i(t+1) &= \gamma\psi_i(t), \quad \text{else,} \end{aligned} \quad (5)$$

where θ_n is a threshold of existing the glia, O is an output of each neuron, θ_g is a threshold of glia's refractory period and γ is an attenuated parameter of the glial output. The glia generates the impulse output as exciting by neuron's output. However, if the neuron's output is not over the threshold or into the refractory period, the glia can not generate the impulse and this glia's output is attenuated by β . In the Fig. 2, we show the generating impulses by the three different positions' glia. $\psi_1(t)$ and $\psi_2(t)$ are generated by neighborhood glia. $\psi_6(t)$ is generated by the far glia from $\psi_1(t)$.

We consider that each glia can make only impulse, however the glial network can make complex oscillations by the glia connecting each other. One example of the glial network's oscillation is shown in Figs. 3 and 4. $\Psi_1(t)$ is similar to $\Psi_2(t)$, because they are correlation each other by the glial network. Between $\Psi_1(t)$ and $\Psi_6(t)$, their correlation is weak, thus, their oscillations become different patterns.

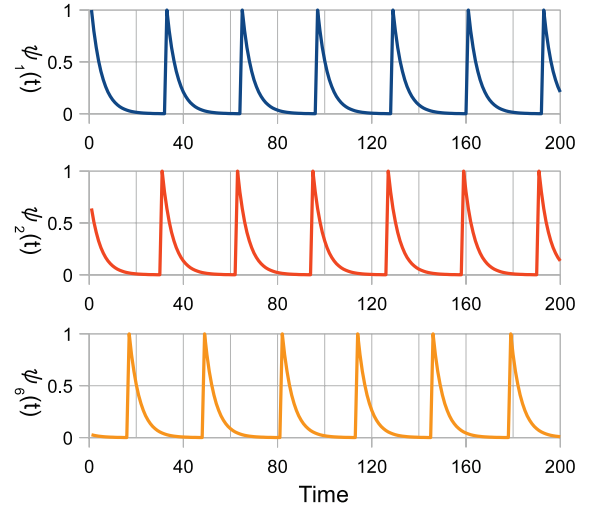


Fig. 2. Outputs of the glia in three different positions.

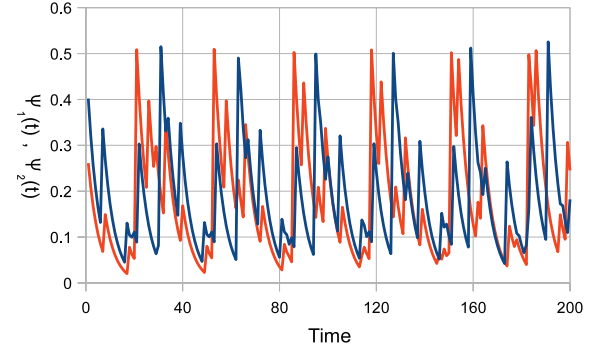


Fig. 3. Outputs of the impulse glial network as first glia and second glia.

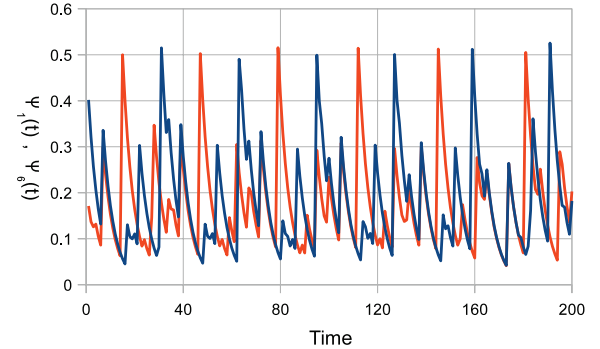


Fig. 4. Outputs of the impulse glial network as first glia and sixth glia.

C. Back Propagation

The error of the MLP propagates backward in the feed forward neural network. The BP algorithm changes the values of the weights between the neurons to obtain smaller error than before. The error of the network is given by Eq. (6).

$$E = \frac{1}{2} \sum_{i=1}^n (t_i - O_i)^2, \quad (6)$$

where E : error value and t : target value.

III. SIMULATION RESULTS

In this section, we show the performance of the proposed MLP with the impulse glial network by learning two kinds of chaotic time series. We use the skew tent map as different A to generate chaotic time series. The skew tent map is formulated by Eq. (7).

$$\phi_i(t+1) = \begin{cases} \frac{2\phi(t)+1-A}{1+A} & (-1 \leq \phi(t) \leq A) \\ \frac{-2\phi(t)+1+A}{1-A} & (A < \phi(t) \leq 1) \end{cases} \quad (7)$$

Figure 5 is skew tent map which is given by Eq. (7).

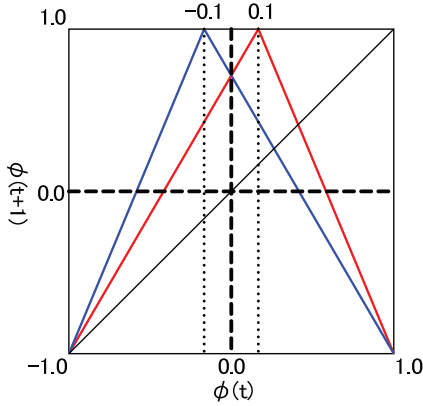


Fig. 5. Skew tent map ($A = -0.10, 0.10$).

The skew tent map generates chaotic time series between -1 and 1 , thus, we transform chaotic time series which become between 0 and 1 . The MLP learns to two different chaotic time series of classification. If the MLP takes $A = -0.10$ chaotic time series, the MLP learns to 1 and if the MLP takes $A = 0.10$ chaotic time series, the MLP learns to 0 . Figure 6 shows example of learning chaotic time series.

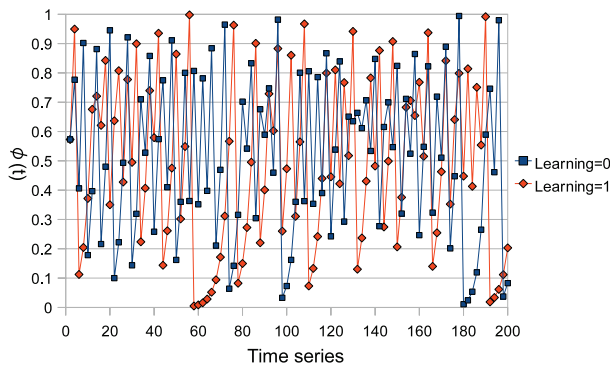


Fig. 6. Chaotic time series ($A = -0.1, 0.1$).

The BP learning for the MLPs is carried out by giving four successive points of the chaotic time series as an input and the following 0 or 1 as an output. The learning is repeated for 200 different sets which are included two different chaotic time series like Fig. 7.

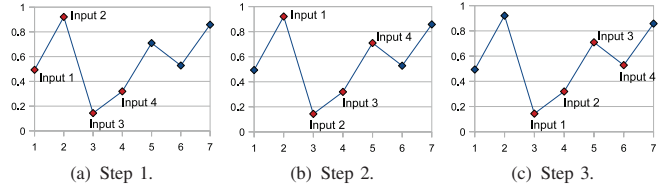


Fig. 7. Learning procedure.

We compare the performances of four different MLPs, which are the conventional MLP, the MLP with the impulse glial network, the MLP with the glial network and the MLP with random noise. We use Mean Square Error (MSE) for evaluate these performances.

$$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - O_i)^2, \quad (8)$$

where N is the number of trials.

A. Parameter Dependency

In this simulations, we confirm the feature of the MLP with the impulse glial network by researching parameters dependency. In each simulation, we obtain the average error from the 100 trials.

1) α dependency: First, we show that the learning performance of the MLP with impulse glial network by using different α in Eq. (5). α is the amplitude of the oscillations, and influences directly to the MLP learning. We change the amount of the α from 0.1 to 0.9 and obtain the MSE as each α . For comparing the results, we show the results of the MLP with the glial network and the MLP with random noise.

Figure 8 is amount of error as each α . The error distribution of the MLP with random noise become flat. From this result, we can say that the α dependency of the MLP with random noise is weaker than the others. The error of the MLP with the glial network becomes small around $\alpha = 0.15$, and become large value in larger α . In the case of the MLP with the impulse glial network, the error become small value in larger α . We consider that the MLP with the impulse glial network can use large value of α because the MLP has the period of inactivity. The MLP can learn true target during the period of inactivity if the MLP is given the large input by the impulse glial network.

2) θ dependency: Next, we show the results as two kinds of thresholds. Figure 9 is the result of changing threshold which is limit to be excited glia as Eq. (5). If this threshold is increased, the glia need large amount of the neurons' output for generating the impulse outputs. In this figure, we change the threshold from 0.50 to 0.95 because we consider that the glia become exciting by positive output of the neuron. We use sigmoidal function for the neuron's output function, thus, $\theta_n = 0.50$ means that the neurons have positive inner state.

In Fig. 9, the errors become small on the whole. The results between 0.60 and 0.90 especially become small error. We consider that the MLP with the impulse glial network does

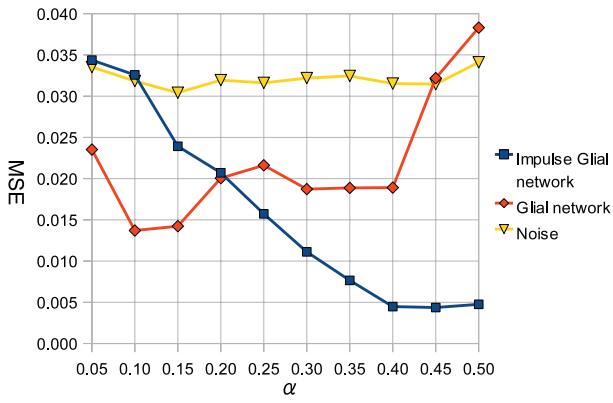


Fig. 8. α dependency of the MLPs.

not have highly dependency from this result, however, we can see that the result become bad around the 0.05. This because of the glias are excited by the small outputs of the neurons and the glias generate positive value output. Thus, the entire neurons' work become positive direction.

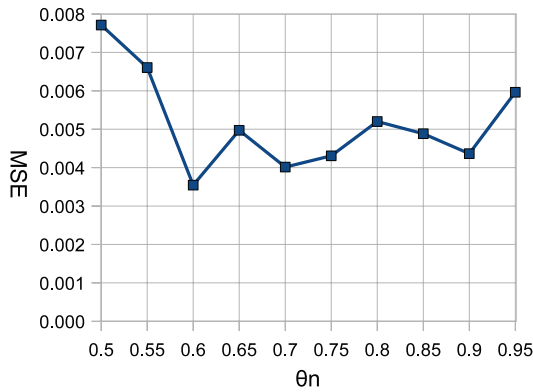


Fig. 9. θ_n dependency of the MLP with the impulse glial network.

Figure 10 is error as the thresholds of glia's refractory period. θ_g operates time length of the period of inactivity. If the output of each glia must become smaller than this threshold, the glias can not generate next impulse output. The glias' outputs are decreased value in an exponential fashion. Thereby, we change the threshold from 10^{-5} to 10^{-1} in an exponential fashion.

This result shows that the error as large θ_n is larger than the error as small θ_n . In larger θ_n , the glias generate the impulse output at short intervals. We consider that the performance of the MLP become well when the glias' outputs are like more impulse from this figure. In fact, the glia's exciting intervals are longer than neurons' work in the biological brain.

3) γ dependency: Finally, we show γ density of the MLP with the impulse glial network as Eq. (5). γ is the attenuated parameter of the glia's output. If γ becomes large, to be attenuated value of the glia's output become small it the one time.

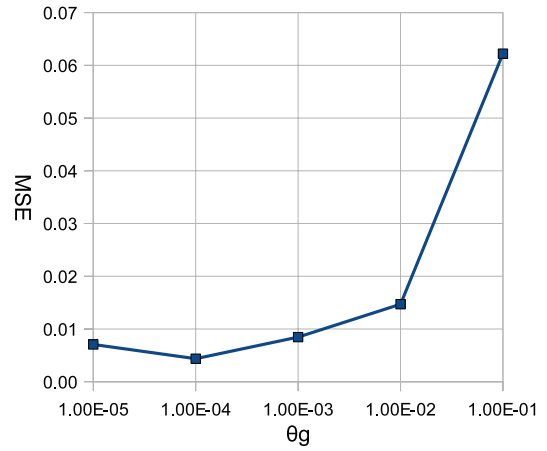


Fig. 10. θ_g dependency of the MLP with the impulse glial network.

The graph inclination is that when γ is larger, the error is small in Fig. 11. As $\gamma = 0.1$, the performance of the MLP with the impulse glial network is the worst. This is because of the glias' output is rapidly decreased by γ . Thereby, we consider that the glias can not influence to the MLP learning. We hold on amount of β as Eq. (4), the glia generating impulse affect around glia, however, it can not affect by itself.

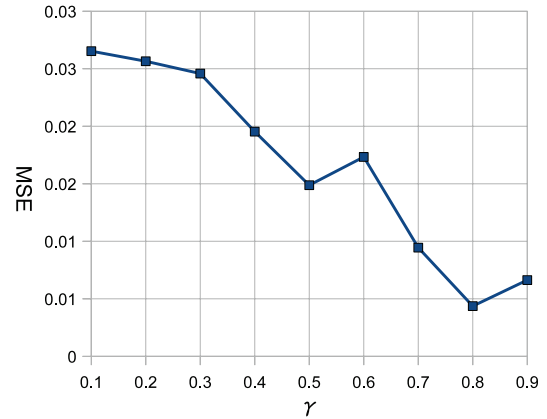


Fig. 11. γ dependency of the MLP with the impulse glial network.

Figure 12 is a graph of the relationship between γ and β . Both γ and β are attenuated parameters. From this figure, we can see that the error is small value as $\beta = 0.4$ or $\beta = 0.6$ which are independent from the value of γ . When both γ and β are small value, the errors are large. In this case, the glias' output reduce rapidly by these two parameters. We consider that the glias must keep up the output value for certain time. And also, when the ratio of γ with β is high, the error become large.

B. Learning Performance

In this paragraph, we investigate the learning performance of the MLPs by calculating the time evolution of the errors. The number of the trials is 100 and the MLPs learn 25000

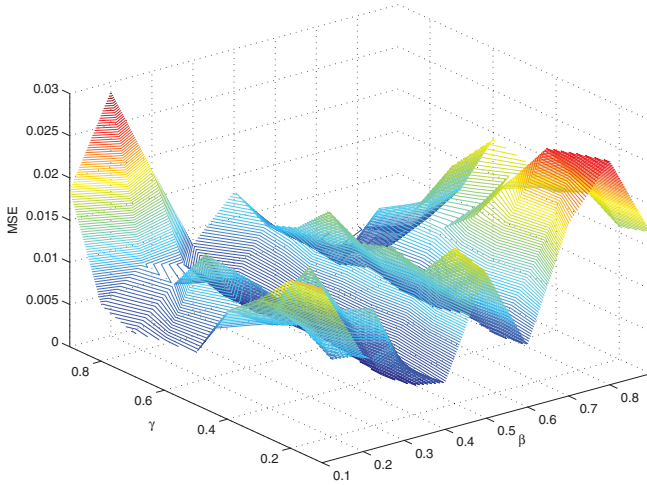


Fig. 12. γ - β characteristic.

times during one trial. We calculate the average of error (Avg. Err.), the minimum error (Min.), the maximum error (Max.) and the standard deviation (St. Dev.). Table I shows the obtained results.

TABLE I
LEARNING PERFORMANCE.

	Avg. Err.	Min.	Max.	St. Dev.
Conventional	0.0351	0.0011	0.1179	0.0434
Glia	0.0074	0.0016	0.0644	0.0125
Impulse Glia	0.0028	0.0004	0.0610	0.0062
Noise	0.0111	0.0009	0.1098	0.0190

In the average of error, we can see that the conventional method is the worst. The MLP with the impulse glial network and the MLP with the glial network reduce more the maximum error than the conventional MLP and the MLP with random noise. We consider that the MLP with the glial networks could escape out from the largest local minimum. The MLP with the impulse glial network is better than the MLP with the glial network. We consider that the MLP with the impulse glial network has the period of inactivity. In the glial network, all glias make random value output, thus, the MLP is possible learned to false value over long periods of time. The MLP with the impulse glial network can learn to true value during the period of inactivity. Because the if the glias make large impulse, the output is decreased and is not generated impulse during this period.

Figure 13 shows an example of the learning curves. The error of the conventional MLP converges constant value and is not improved any more. The MLP with the impulse glial network reduced the error earliest of all. This because of neighborhood neurons are inputted similar oscillations by the impulse glial network. We assume that the similar oscillations give good influence to the MLP learning.

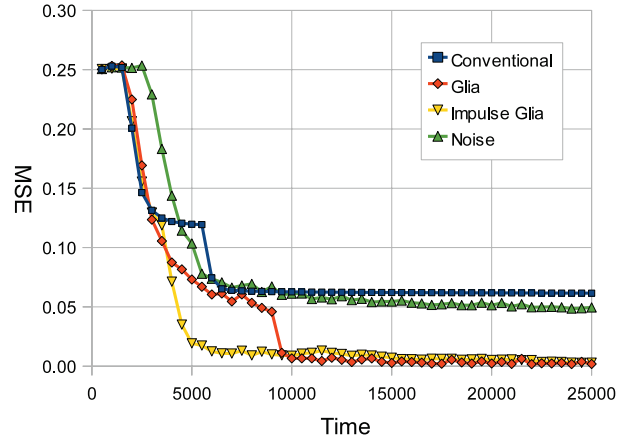


Fig. 13. Error curve of four different MLPs.

C. Generalization Capability

Finally, in order to evaluate the generalization capabilities of the MLPs, we give the MLPs unknown time series generated by the same skew tent map after the learning and evaluate the errors between the outputs and the correct value. We use ten different time series including 200 points for the test data. The learning data and one example of the test data are shown in Fig. 14.

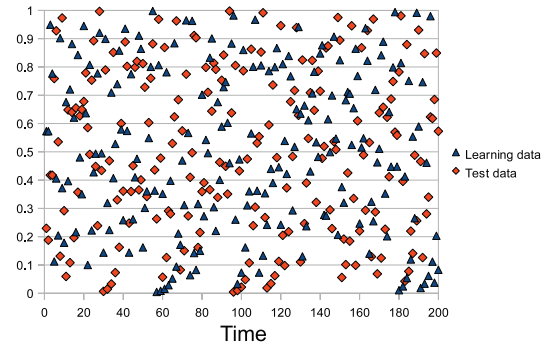


Fig. 14. Learning data and test data.

This result is summarized in Tab. II. From Tab II, we can see that the MLP with the impulse glial network is the best as average of error. However, this network's maximum error is the worst, this is because of this network having the oscillations by the glial network, thus this network took bad states at the finishing learning.

TABLE II
GENERALIZATION CAPABILITY.

	Avg. Err.	Min.	Max.	St. Dev.
Conventional	0.0496	0.0006	0.1684	0.0468
Glia	0.0345	0.0060	0.1380	0.0216
Impulse Glia	0.0326	0.0004	0.2156	0.0232
Noise	0.0443	0.0008	0.1744	0.0419

IV. CONCLUSION

In this article, we have proposed the MLP with the impulse glial network. The glia generate the impulse output by the MLP's output, moreover the outputs are propagating to other glia, thus the glial network can generate complex oscillations. We could see that the glia's outputs were similar patterns as local positions in the impulse glial network.

By computer simulations, we showed that the changing performance of MLP by the each parameter as the impulse glial network and found the good parameters for learning MLP. Moreover, we confirmed that the proposed MLPs with the impulse glial networks is better learning performance and better generalization capability than the conventional networks. The impulse glial network has the period of inactivity, thus, this network can have large amplitude oscillations. We consider that this behavior gave good influence to the MLP learning.

In the Future works, we will particularly research about outputs of the impulse glial network. Neighborhood glia generating outputs looks like synchronizing each other. We assume that the correlation of the glia is important for the MLP learning.

REFERENCES

- [1] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.
- [2] C. Ikuta, Y. Uwate and Y. Nishio, "Chaos Glial Network Connected to Multi-Layer Perceptron for Solving Two-Spiral Problem," *Proc. ISCAS'10*, pp. 1360-1363, May 2010.
- [3] C. Ikuta, Y. Uwate and Y. Nishio, "Multi-Layer Perceptron with Impulse Glial Network," *Proc. NCN'10*, pp.9-11, Dec. 2010.
- [4] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol. 2, pp. 844-847, 2001.
- [5] 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.
- [6] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.
- [7] G. Perea and A. Araque, "Glial Calcium Signaling and Neuro-Glia communication," *Cell Calcium*, vol. 38, 375-382, 2005.
- [8] S. Kriegler and S.Y. Chiu, "Calcium Signaling of Glial Cells along Mammalian Axons," *The Journal of Neuroscience*, vol. 13, 4229-4245, 1993.
- [9] M.P. Mattoson and S.L. Chan, "Neuronal and Glial Calcium Signaling in Alzheimer's Disease," *Cell Calcium*, vol. 34, 385-397, 2003.