2011 International Workshop on Nonlinear Circuits,
Communication and Signal Processing
NCSP'11, Tianjin, China, March 1-3, 2011

Multi-Layer Perceptron with Glial Network Influenced by Local External Stimulus

Chihiro Ikuta, Yoko Uwate and Yoshifumi Nishio

Tokushima University
2–1 Minami-Josanjima, Tokushima, Japan
Phone:+81-88-656-7470
Email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract

We have proposed the glial network which was inspired from the feature of glias. The glias are nervous cell existing in the brain and transmit signal each other like neurons. In the glial network, the glias connect to the neurons and other glias. The glias trade information each other by this network.

In this article, we investigate the glial network when only one glia is stimulated by an external noise. The external noise propagates the glial network and influences to the Multi-Layer Perceptron (MLP). By the effect of the one-way influence via the glial network, a kind of position-depending-feature appears in the MLP. The simulation results show that the proposed network possesses better learning performance, more biased anti-damaging property, and better generalization capability than the conventional networks.

1. Introduction

Human's brain is made by nervous cells which are almost neurons or glias. In particular, the neurons were investigated by many researchers. Because the neurons were known that they make thinking of human by connected to each other. 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].

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

In this paper, we research a learning result of the MLP when it is influenced by local stimulus. In the biological system, a neural network is not always affected by uniformed external stimulus. We assume that only one glia is affected by external stimulus and that this influence propagates in the glial network. We give random oscillation to one glia and the MLP learns data having different amplitude oscillations by the glial network. We confirm that the learning result of MLP is changed by the biased oscillations.

2. MLP with Glial Network Influenced by Local External Stimulus

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 (2-10-1) and a glial network connected to the hidden layer. Figure 1 shows an example of the MLP with the glial network.

2.1. Neuron Updating Rule

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

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

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.

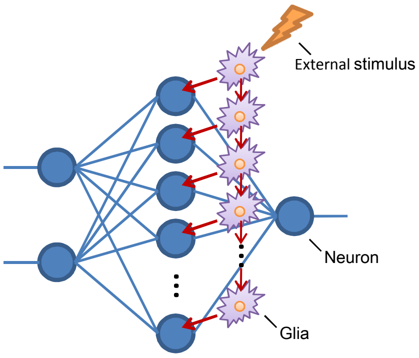

Figure 1: MLP with glial network influenced by local external stimulus.

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

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

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

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

2.2. Glial Network

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

In this study, we assume that the glias do not generate oscillations by themselves but only one glia located on an edge of the network is influenced by a local external stimulus. The external noise propagates the glial network with decaying.

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

$$\Psi_i(t) = \beta^{i-1}\psi(t-i-1) \quad (i = 1, 2, 3, \cdots 10), \tag{4}$$

where ψ is the local external stimulus given to the first glia and β denotes attenuation parameter. In this study, we input uniform random noise as external stimulus ψ. The noise effect is propagating the glial network, however, the noise is attenuated by β and delayed.

3. Simulation Results

In this section, we show the performance of the proposed MLP with the glial network by learning chaotic time series. We use the logistic map to generate chaotic time series. The logistic map is formulated by Eq. (5).

$$x(t+1) = ax(t)\{1 - x(t)\}. \tag{5}$$

We use $a = 4.0$ which makes the time series as fully developed chaotic oscillation. We compare the performances of five different MLPs, which are the conventional MLP, the MLP with the glial network ($\alpha = 0.15$, $\beta = 0.9$), the MLP with the glial network ($\alpha = 0.2$, $\beta = 0.8$), the MLP with the glial network ($\alpha = 0.4$, $\beta = 0.6$) and the MLP with inputted random noise (noise amplitude is 1.0). We decided the parameter combinations of the MLP with the glial networks as the oscillation amplitudes become around 1.0.

Figure 2 is an example of the chaotic time series (100 points) which is obtained by giving a certain initial value to the logistic map. The BP learning for the MLPs are carried out by giving two successive points of the chaotic time series as an input and the following one point as an output. The learning is repeated for 100 different sets like Fig. 3.

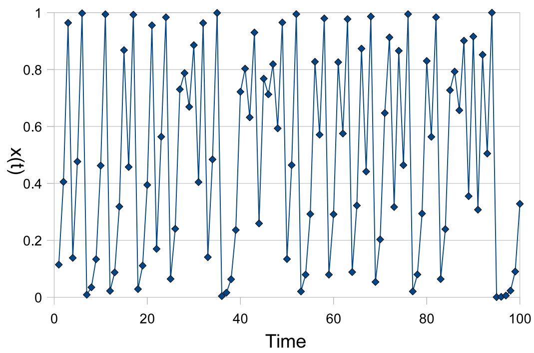

Figure 2: Chaotic time series ($a = 4.0$).

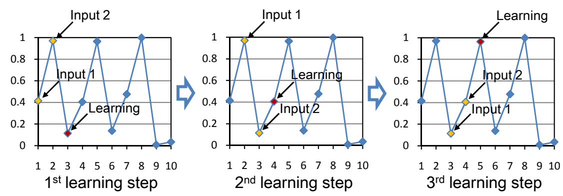

Figure 3: Learning procedure.

We evaluate three kinds of performances with this task; learning performance, anti-damaging property and generalization capability. We use the following Mean Square Error (MSE) for evaluate these performances.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - O_i)^2. \tag{6}$$

3.1. Learning Performance

First, we investigate the learning performance of the MLPs by calculating the time evolution of the errors. The number of the trials is 200 and the MLPs learn 500000 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 1 shows the obtained results. We can see that the conventional MLP is the worst. In this result, the MLP with random noise performs the best. However, the results of the MLPs with the glial networks are similar to that of the MLP with random noise. Further, the maximum error of the MLPs with the glial networks are larger than that of the MLP with random noise. We consider that the averaged values of the MLPs with the glial networks are influenced by a few very bad results.

Table 1: Learning performance (10^{-3}).

	Conv. –	Glia $\beta = 0.9$	Glia $\beta = 0.8$	Glia $\beta = 0.6$	Noise –
Avg. Err.	0.327	0.162	0.194	0.248	0.121
Min.	0.075	0.054	0.054	0.062	0.058
Max.	1.467	1.720	1.564	1.499	0.163
St. Dev.	0.478	0.204	0.304	0.382	0.024

Figure 4 shows an example of the learning curves. The error of the conventional MLP becomes constant early and is not improved any more. This is because of the local minimum problem of the optimizations. The other MLPs seem to be able to escape out from local minima. We can see that the MLPs with the glial networks can find good states earlier than the MLP with random noise in this example.

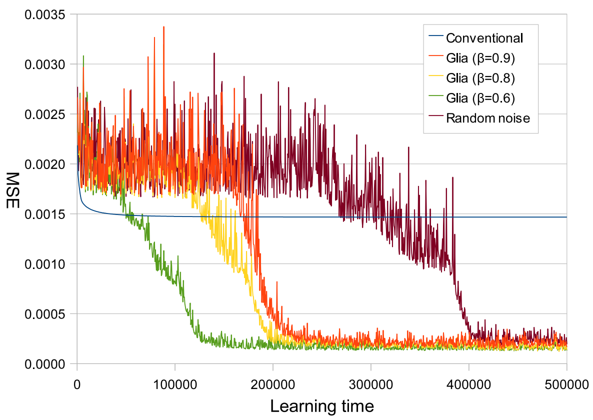

Figure 4: Learning curve of five MLPs.

3.2. Anti-Damaging Property

Next, in order to investigate the anti-damaging property of the MLPs, we cut one neuron in the hidden layer after the learning as shown in Fig. 5. This means that the MLPs must operate by only 9 neurons in their hidden layers. We change the position of the cut neuron and evaluate the errors.

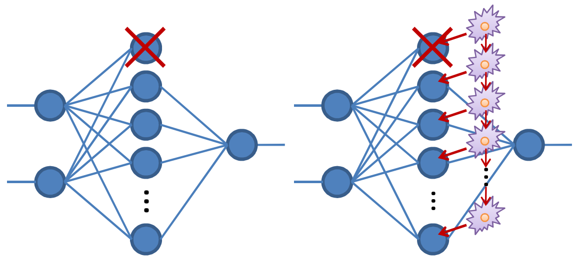

Figure 5: Examples of damaging MLPs.

Figure 6 shows the position dependence of the average error for 200 trials after cutting one neuron. We can say that a neuron is very important for the corresponding MLP if the error becomes large after cutting the neuron. On the contrary, if the performance is similar even cutting the neuron, the neuron is not useful for the MLP. In Fig. 6, the horizontal labels denote the positions of the neurons, hence, the effect of the external noise is stronger for smaller label of the positions for the MLPs with the glial networks.

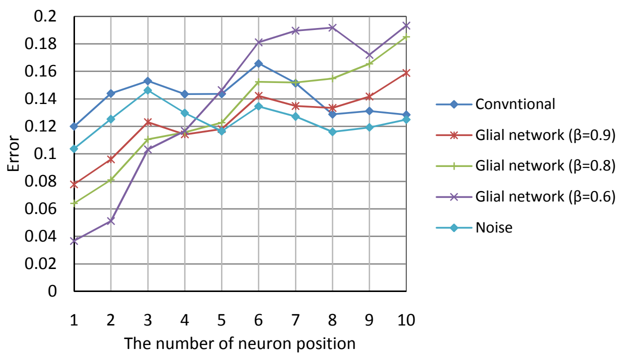

Figure 6: Performance after damaging one neuron.

From this figure, the MLP with random noise is similar to the conventional MLP. As we can expect, for the conventional MLP and the MLP with random noise, every neurons have the similar importance for the performance. However, we can see clearly from the figure that the importances of the neurons of the MLPs with the glial networks are different depending on their positions.

3.3. Generalization Capability

Finally, in order to evaluate the generalization capabilities of the MLPs, we give the MLPs unknown time series generated by the same logistic map after the learning and evaluate the errors between the outputs and the correct following points. We use ten different time series including 100 points for the test data. The learning data and one example of the test data are shown in Fig. 7.

The error performance are summarized in Table 2. The performance of the conventional MLP is the worst. The er-

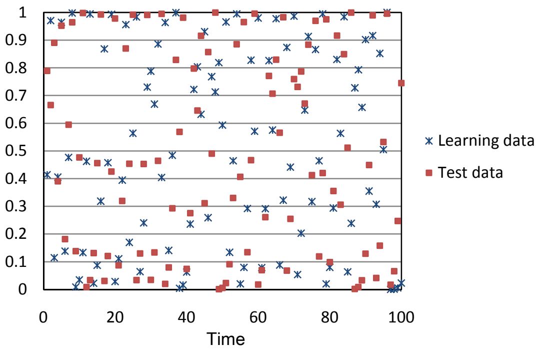

Figure 7: Learning data and test data.

ror of the MLP with random noise has the smallest standard deviation. The MLPs with the glial networks has the large difference between the maximum error and the minimum error.

Table 2: Generalization capability (10^{-3})

	Conv. –	Glia $\beta = 0.9$	Glia $\beta = 0.8$	Glia $\beta = 0.6$	Noise –
Avg. Err.	0.386	0.353	0.343	0.342	0.355
Min	0.057	0.076	0.083	0.068	0.120
Max	2.131	2.522	2.823	2.845	0.532
St. Dev.	0.550	0.230	0.280	0.418	0.271

Next, we add noise (5%) to the test patterns and investigate the denoising ability. The results are summarized in Table 3. We can see that the performance of the conventional MLP is the worst. The MLPs with the random noise has the better performance than the MLPs with the glial networks.

Table 3: Generalization capability for noisy data (10^{-3})

	Conv. –	Glia $\beta = 0.9$	Glia $\beta = 0.8$	Glia $\beta = 0.6$	Noise –
Avg. Err.	8.748	6.243	6.562	7.7729	6.000
Min	3.635	3.195	3.428	3.316	3.222
Max	58.99	20.97	17.39	42.99	17.84
St. Dev.	4.779	1.340	1.457	3.404	1.254

4. Conclusion

In this article, we have researched the MLP with the glial network when the glias do not generate any oscillations but propagate the effect of the external stimulus. By computer simulations, we confirmed that the proposed MLPs with the glial networks possessed better learning performance, more biased anti-damaging property, and better generalization capability than the conventional networks.

At the moment, we could not maintain that the proposed MLP performs much better for a certain task than the conventional one in the sense of learning performances, but the existence of the glial network produces novel features which have not realized in the conventional artificial neural networks; such as individuality of neurons. Real brains must be, of course, non-uniform and hence we believe that a kind of individuality will give wider chance to realize more intelligent artificial neural networks.

Acknowledgment

This work was partly supported by JSPS Grant-in-Aid for Scientific Research 22500203.

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] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol. 2, pp. 844-847, 2001.

[3] 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.

[4] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.

[5] 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.