2018 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2018, Tarragona, Spain, September 2-6, 2018

Learning Method Using Acetylcholine among Neurotransmitter in Feed Forward Neural Networks

Kazuki Nagao, Yoko Uwate and Yoshifumi Nishio

Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima 770-8506, Japan
Email: {k-nagao, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract—In this study, we propose back propagation learning in feed forward Neural Networks using characteristics of acetylcholine. Acetylcholine is a major neurotransmitter in the brain. The main work of acetylcholine is to excite the nerves and enhance motivation for learning and memory. In Neural Networks, we bring it closer to human information processing by using this function. Our purpose is to reduce learning error rate.

1. Introduction

Neural Networks (NN) and Deep Learning (DL) are made to imitate the human nervous system. Such systems learn to do tasks by considering examples, generally without task-specific programming. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the analytic results to identify cats in other images. They have found most use in applications difficult to express in a traditional computer algorithm using rule-based programming [1].

The human brain processes many and various kinds of information got on the basis of the five senses (sight, hearing, smell, taste and touch). And we can judge things and put us into action. Thereby, the human brain has high bandwidth, high speed parallel processing and flexibility. So, artificial intelligence was created to do the same work as the human brain on a computer. Typical examples are NN and DL [2]. In recent years, they have been applied in various fields and used for applications. For example, smartphone, car, stock price forecast, robot, consumer electronics and so on.

There are the following mainly substances constituting the human neural circuit. First, there are many nerve cells (neuron) in the brain, and information is passed on next neurons through the connection, and we can store and judge things. Second, each connection (synapse) between neurons can transmit a signal to another neuron. The receiving neuron can process the signals and then signal downstream neurons connected to it. Neurons may have stated, generally represented by real numbers, typically between 0 and 1. Between neurons may also have a weight that varies as learning proceeds, which can increase or decrease the strength of the signal that it sends downstream. Third, neurotransmitter (chemical messenger) is substances generated from synapse. It excites and suppresses neurons. It has a role to promote learning. Other substances include glial cells. Glial cells include several types of cells such as astrocytes, and oligodendrocytes apart from the neurons in the brain. In particular, astrocytes are known to be important in higher function and are therefore sometimes simply called glial cells. Now, neuron, synapse and glial cell are used in NN [3] and DL. However, neurotransmitter is not used at all in NN and DL.

The original goal of NN was to solve problems in the same way that the human brain would. Over time, attention focused on matching specific mental abilities. In this study, we propose a system of Feed Forward NN (FFNN) having artificial neurotransmitter. We use characteristic of acetylcholine among neurotransmitter. Acetylcholine has been reported to be related to diseases and memory [4], [5]. We use this system and improve learning error rate.

2. Acetylcholine among Neurotansmitter (Chemical Messenger)

Neurotransmitter, also known as chemical messenger, is endogenous chemicals that enable neurotransmission. They transmit signals across a chemical synapse, such as a neuromuscular junction, from one neuron to another neuron. Neurotransmitter is released from synaptic vesicles in synapses into the synaptic cleft, where they are received by receptors on the target cells. Many neurotransmitters are synthesized from simple and plentiful precursors such as amino acids, which are readily available from the diet and only require a small number of biosynthetic steps for conversion. Neurotransmitter plays a major role in shaping everyday life and functions. Their exact numbers are unknown, but more than 100 chemical messengers have been uniquely identified. For example, there are acetylcholine, adrenaline, glutamate, dopamine and so on.

In this study, we focused on acetylcholine among neurotransmitter. Acetylcholine was discovered by Henry Halletdale in 1914 and revealed to be a neurotransmitter by Otto Levy. It is early detection among neurotransmitters. It has many features. First, it has a relationship with illness. It is known that too much acetylcholine secretion leads to Parkinson's disease. It is also known that too little acetylcholine secretion leads to Alzheimer's disease. Second, it has a relationship with memory. In the human brain, there is a part that plays a central role in memory formation called hippocampus. It is studied as acetylcholine as a substance that promotes memory of the hippocampus. Third, it is secretion amount changes within 1 day. Secretion tends to decrease throughout the day. As a result, the learning function declines and you lose motivation and feel sleepy. Seafood is known to help secrete acetylcholine. However, you can restore the secretion of acetylcholine by having a meal. However, because there are individual differences, specific amount of secretion is unknown.

3. Proposed System

We propose a system using artificial characteristics of acetylcholine for the only middle layer of FFNN. This is because the middle layer plays a very large role in FFNN.

In this study, we propose two methods adding acetylcholine. The first one is "Method Using Acetylcholine for Input Information of Middle Layer". As a characteristic of acetylcholine, it excites and inhibits the next neuron during signal transduction. This is modeled between input layer and middle layer to make it the first one.

The second one is "Method Using Acetylcholine for Output Information of Middle Layer". Like the characteristic of acetylcholine used earlier, it excites and inhibits the next neuron during signal transduction. This is modeled between middle layer and output layer to make it the second one.

3.1. Method Using Acetylcholine for Input Information of Middle Layer

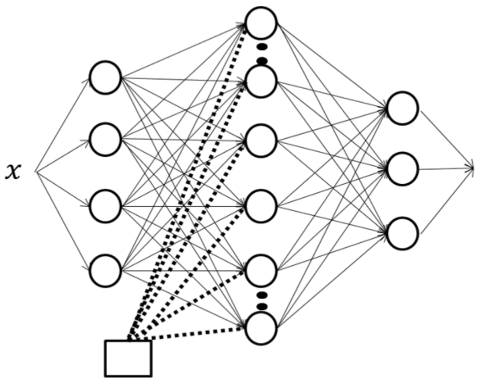
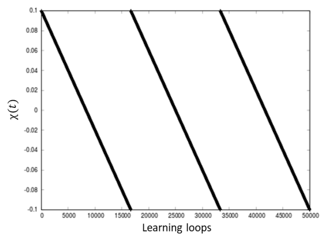


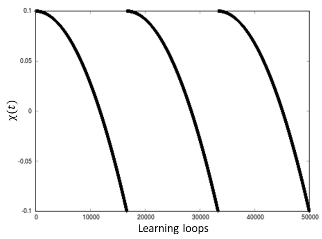
Figure 1: A Schematic diagram of method using acetylcholine for input information of middle layer in FFNN.

Characteristics of acetylcholine are not fully understood yet. We use two reported characteristics of acetylcholine in this method. First, the amount of acetylcholine secretion decreases in a daily life. Second, the amount of secretion recovers by having three meals (breakfast, lunch and dinner). For using these characteristics, we propose three patterns which artificial characteristics of acetylcholine as shown in Fig. 2. We consider the number of trials to be one day. So we decide to recover the secretion amount three times within learning loops. The range of acetylcholine is changed from -0.1 to 0.1 in all patterns.

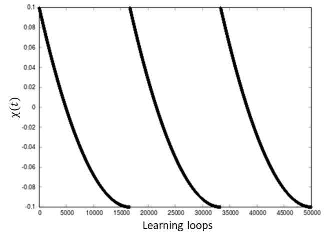
Pattern 1 shows that the amount of secretion decreases linearly at each learning loop. Pattern 2 shows that the amount of secretion decreases curvilinearly at each learning loop. The angle of the tangent becomes sharply. Pattern 3 also shows that the amount of secretion decreases curvilinearly at each learning loop. The angle of the tangent becomes gentle. Patterns 2 and 3 are represented by quadratic functions.



(a) Pattern 1.



(b) Pattern 2.



(c) Pattern 3.

Figure 2: Three patterns which artificial characteristics of acetylcholine.

Equation (1) shows the propagating equation of the conventional neuron system at input layer to hidden layer. It does not have the acetylcholine's system.

$$h_{ij}(t+1) = f\left\{\sum w_{ij}(t)x(t) - \theta(t)\right\} \quad (1)$$

x is input data. w is connection weight of neurons. θ is threshold. h is output value.

Equation (2) shows the propagating equation of the neuron having the acetylcholine's system at input layer to hidden layer.

$$h_{ij}(t+1) = f\left\{\sum w_{ij}(t)(x(t) + \chi(t)) - \theta(t)\right\} \quad (2)$$

$\chi(t)$ denotes the acetylcholine's value. It also includes the characteristics of Fig. 2.

And Eqs. (1) and (2) using sigmoid function Eq. (3) for activation function [6].

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

3.2. Method Using Acetylcholine for Output Information of Middle Layer

In machine learning you may fall into the local solution. The local solution is a system that converges to a partial optimal solution and does not reach a real optimal solution. In order to solve this problem dropout and others are proposed [7]. In this method, We propose a new one that turns into dropout. And it improves learning error rate.

We use one reported characteristic of acetylcholine in this method. It is secretion amount changes within one day. Secretion tends to decrease throughout the day. As the result, the learning function declines and you lose motivation and feel sleepy. We use this feature to increase or decrease the amount of output information in input layer to middle layer of FFNN.

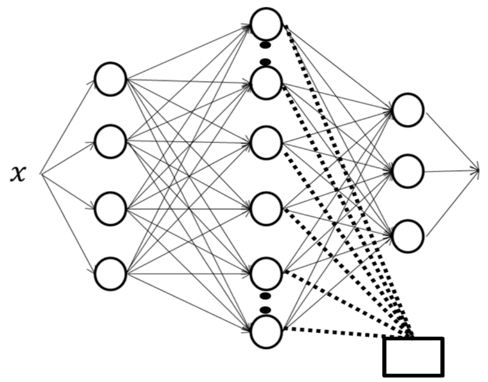


Figure 3: A Schematic diagram of method using acetylcholine for output information of middle layer in FFNN.

Cases (0.5 - 1.5) and (0.5 - 2.0) are shown how much to increase or decrease the amount of output information at input layer to hidden layer. For example, in case (0.5 - 2.0), it means changing the information volume of neurons from 1/2 to 2 times. The amount of change is shown in Fig. 4.

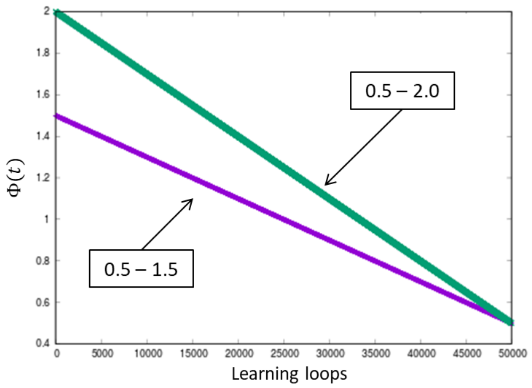


Figure 4: Changing the information volume of output.

Equation (4) shows the propagating equation of the neuron having the acetylcholine's system at input layer to hidden layer.

$$h_{ij}(t+1) = f\left\{\sum (w_{ij}(t)x(t) - \theta(t))\Phi(t)\right\} \qquad (4)$$

$\Phi(t)$ denotes the acetylcholine's value. Equation (3) is used as the activation function.

4. Simulation Result

We compared with the minimum value in the average of learning error rate of 10 consecutive times. Learning error rate is determined by using root mean squared error. Root mean square error is expressed as Eq. (5).

$$A(y) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{learn} - y)^2} \qquad (5)$$

y_{learn} is derived based on learning. y is the answer we have prepared. The weight correction constant is set to 0.1 in all systems.

4.1. Method Using Acetylcholine for Input Information of Middle Layer

We verify the proposed method using three data sets. First, iris data set [8]. It has 150 data and uses the network of the number of neurons in the input layer is 4, the number of neurons in the middle layer is 10, the number of neurons in the output layer is 3. Second, we use a data set with 1728 data. It is one that evaluates six to four features of cars (car data set) [8]. The network of the number of neurons in the input layer is 6, the number of neurons in the middle layer is 10, the number of neurons in the output layer is 4. Third, we use a data set with 201 data. It lets you learn the point of [-1, 1] of the function of $y = x^2$.

The results when using three data sets are shown in Tables 1, 2 and 3 below. In Tables 1, 2 and 3, "Conventional" shows the conventional system that does not have acetylcholine's system. Patterns 1, 2 and 3 show the proposed system that has features of acetylcholine. These features are shown in Fig. 2.

Table 1: Learning error rate of using iris data set.

System	Learning error rate
Conventional	0.05011
Pattern 1	0.04246
Pattern 2	0.04496
Pattern 3	0.04659

In Table 1, all proposed patterns obtain higher learning accuracy than the conventional system. In pattern 1, learning accuracy is improved by about 18%.

Table 2: Learning error rate of using car data set.

System	Learning error rate
Conventional	0.00001
Pattern 1	0.00001
Pattern 2	0.00002
Pattern 3	0.00002

In Table 2, pattern 1 has the same learning accuracy as the conventional system, and patterns 2 and 3 have poor error rate. As a cause of this, because the number of data sets is large. Very high learning accuracy is obtained also in the conventional system, so the proposed method of this time seems to have failed.

Table 3: Learning error rate of using $y = x^2$ data set.

System	Learning error rate
Conventional	0.00559
Pattern 1	0.00562
Pattern 2	0.00531
Pattern 3	0.00520

In Table 3, patterns 2 and 3 were able to reduce learning error rate compared with the conventional system. In pattern 3, learning error rate is reduced by about 7%.

4.2. Method Using Acetylcholine for Output Information of Middle Layer

The results when using one data sets are shown in Fig. 5 and Table 4 below. In Table 4, "Conventional" shows the conventional system that does not have acetylcholine's system.

Figure 5 shows two graphs when the information amount of neurons is changed. From Fig. 5, we can see that we are escaping from the local solution by using this learning method.

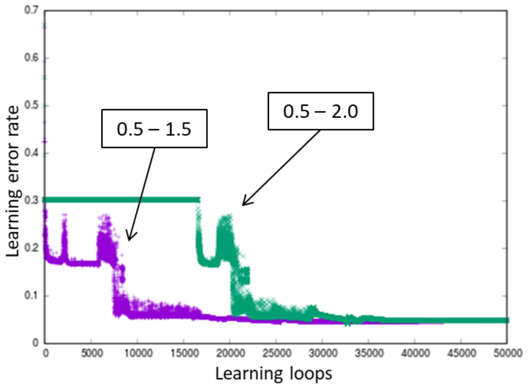


Figure 5: Changing in error rate when changing information amount of neurons.

Table 4: Learning error rate of iris data set.

System	Learning error rate
Conventional	0.05011
0.5 - 1.5	0.04641
0.5 - 2.0	0.04893

Next, minimum values of learning error rate of the conventional system and proposed systems are compared. From Table 4, proposed two types are able to reduce learning error rate over the conventional system.

5. Conclusion

First, new methods were proposed for FFNN using the property of acetylcholine, which is a neurotransmitter. Then, we used the amount of acetylcholine secretion decreases in a daily life and the amount of secretion recovers by having meals (breakfast, lunch and dinner). For using these characteristics, we made three patterns. As the results of verification in three data sets, it was possible to reduce the error rate as compared with the conventional system. However, it was difficult to say that it was suitable for all patterns. It will be necessary to verify using more data sets in the future.

Second, we used that it is secretion amount changes within one day. Secretion tends to decrease throughout the day. As the result, the learning function declines and you feel sleepy. For using this characteristic, we proposed the method to change to dropout. As the results of this study, it is possible to reduce the error rate as compared with the conventional system and get out the local solution.

We believe that the versatility of the neural networks can be improved by using two methods for one system. We would like to challenge more realistic problems in the future.

Acknowledgment

This work was partly supported by JSPS Grant-in-Aid for Scientific Research 16K06357.

References

[1] Hinton G, Vinyals O and Dean J, "Distilling knowledge in a neural network," In Deep Learning and Representation Learning Workshop, NIPS, 2014.

[2] Gurney and Kevin, "An introduction to neural networks," UCL Press, ISBN 1857286731, OCLC 37875698, 1997.

[3] Chihiro Ikuta, Yoko Uwate, Yoshifumi Nishio and Guoan Yang, "Multi-Layer Perceptron with Pulse Glial Chain," IEICE Transactions on Fundamentals, vol. E99-A, no. 3, pp. 742-755, Mar. 2016.

[4] Lombardo S and Maskos U, "Role of the nicotinic acetylcholine receptor in Alzheimer's disease pathology and treatment," Neuropharmacology 96(ptB), 255-62, 2015.

[5] Dai Mitsushima, Akane Sano and Takuya Takahashi, "A cholinergic trigger drives learning-induced plasticity at hippocampal synapses, "Nature Communications 4, 2760, 2013.

[6] Cybenko adn G.V, "Approximation by Superpositions of a Sigmoidal function," In van Schuppen, Jan H. Mathematics of Control, Signals, and Systems, Springer International, pp. 303314, 2006.

[7] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," In ICASSP, 2013.

[8] http://archive.ics.uci.edu/ml/index.php, "Center for Machine Learning and Intelligent Systems"