

# 局所的な外部刺激を受ける グリアネットワークを持つ階層型パーセプトロンの研究

生田 智敬<sup>†</sup> 上手 洋子<sup>†</sup> 西尾 芳文<sup>†</sup>

<sup>†</sup> 徳島大学

〒 770-8506 徳島県徳島市南常三島町 2-1

E-mail: †{ikuta,uwate,nishio}@ee.tokushima-u.ac.jp

**あらまし** 私たちは、人間の脳内に存在するグリアと呼ばれる細胞の特徴を模したグリアネットワークについて研究を行ってきた。このグリアネットワークは、グリアがそれぞれ異なった振動を発生しており、その振動が互いに影響し合うことによって階層型パーセプトロン (MLP) の性能が向上するというものである。本研究では、各グリアは振動を発生しておらず、一個のグリアに対して外部刺激が加えられたモデルを仮定した。この外部刺激が、グリアネットワーク内を伝搬し他のグリアに伝えられるとともにニューロンに対して入力される。このように外部刺激が加えられたときの MLP の関数近似性能、被ダメージ性能、および汎化性能を示す。

**キーワード** グリアネットワーク, 階層型パーセプトロン, 外部刺激

## Investigate of Multi-Layer Perceptron with Glial Network Receiving Local External Stimulus

Chihiro IKUTA<sup>†</sup>, Yoko UWATE<sup>†</sup>, and Yoshifumi NISHIO<sup>†</sup>

<sup>†</sup> Dept. of E.E. Eng., Tokushima University, 2-1 Minami-Josanjima, Tokushima 770-8506, JAPAN

E-mail: †{ikuta,uwate,nishio}@ee.tokushima-u.ac.jp

**Abstract** We have proposed the glial network which was inspired from the feature of brain. In the glial network, glias generate independent oscillations and these oscillations propagated neurons and other glias. We confirmed that the glial network improved the learning performance of the Multi-Layer Perceptron (MLP) by connecting the neurons more effectively than the conventional networks. In this article, we investigate the glial network when the glias do not oscillate by themselves but only one glia is stimulated by an external noise. The external noise propagates the glial network and influences to the 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 networks possesses better learning performance, more biased anti-damaging property, and better generalization capability than the conventional networks.

**Key words** Glial network, MLP, External stimulus

### 1. ま え が き

現在、様々なニューラルネットワークが提案されている。中でも、階層型パーセプトロン (MLP) と呼ばれるフィードフォワードニューラルネットワークは、構造が分かりやすく扱いやすいため様々な種類の問題を解くために用いられている。MLP は、ニューロンが層を作っており、異なった層のニューロンと結合することによって出力を得ることができる。このとき、同じ層内でのニューロン同士は結合していない。

ここで、生物の脳内について考えてみる。脳内では、ニュー

ロンが複雑に結合し互いに情報交換することによって精神活動を行っている。ただし、人工ニューラルネットワークのように一様な動作ではなく、一部は非常に活発に活動するニューロン、あるいは特定の条件時だけ活動するニューロン等、様々なニューロンが存在しており、これらは、ニューロンの存在する脳内での位置に依存していると考えられる。外部に近いニューロンは常に外部刺激に影響される状態にあり、対して、内部は安定的に動作できると考えられる。つまり、実際の脳内のニューロンは不均一に動作していると仮定できる。

本研究では、外部刺激によってニューロンの学習が不均一に

なったモデルを仮定する。この外部刺激が伝搬するネットワークとして、グリアが形成したグリアネットワークに注目した。グリアとは、ニューロンと似た特徴を持った神経細胞であり、脳内に存在している [1]。このグリアは、イオン濃度を変化させることで脳の状態を保っている [2] [3]。この働きから、外部刺激が脳表面に与えられたとき、その外部刺激を内部に伝搬していくと仮定した。このようにグリアによって伝搬された外部刺激が、ニューロンの学習にどのように寄与するか調査する。

## 2. 外部刺激を受ける階層型パーセプトロン

階層型パーセプトロン (MLP) は、最も有名なフィードフォワードニューラルネットワークのひとつであり、パターン認識やデータマイニング等様々な問題を処理できることで知られている。MLP では、ニューロンが層のようになって存在しており、異なった層のニューロン同士が結合することによってネットワークを形成する。本研究では、3 層の MLP(2-10-1) を用い、その隠れ層のニューロンにグリアネットワークを結合する。そして、1 番目のグリアに外部刺激を入力し、その刺激がグリアネットワークを通じてニューロンに伝えられる図 1 に、外部刺激を受けるグリアネットワークを持った MLP の例を示す。

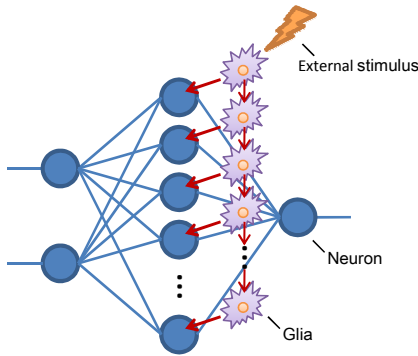


図 1 外部刺激を受けるグリアネットワークを持った MLP.

### 2.1 ニューロンの更新則

一般的なニューロンの更新則は、式 (1) のように表される。

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

ただし、 $x$  はニューロンの入力および出力、 $w$  は結合荷重、 $\theta$  はしきい値、 $f$  は出力関数である。また、結合荷重としきい値は誤差逆伝播法 (BP) によって学習する [4]。

次に、隠れ層のニューロンの更新則について示す。グリアネットワークの影響はニューロンのしきい値に対して影響することが知られているため、式 (1) のしきい値にグリアネットワークの項を加える。このグリアネットワークを結合したニューロンの更新則を式 (2) に示す。

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

ここで、 $\Psi$  はグリアの出力、 $\alpha$  はグリアの出力の振幅である。また、出力関数に用いたシグモイド関数を式 (3) に示す。

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

### 2.2 グリアネットワーク

生物学におけるニューラルネットワークでは、グリアの出力が近くのニューロンに作用するとともに、その影響が広く伝搬していくことで知られている。私たちは、このグリアの特徴からグリアネットワークを提案した [5]。

本研究では、単一のグリアに外部刺激が連続的に加えられ、その外部刺激がグリアネットワークを通じてニューラルネットワークに加えられる。以上のようなグリアネットワークの動作を式 (4) のように定義した。

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

ここで、 $\psi$  は一番目のグリアに与えられた外部刺激、 $\beta$  は減衰項である。一番目のグリアに入力された外部刺激は、次のグリアに伝搬されるたびに  $\beta$  によって減衰されるとともに 1 サイクル毎に遅れが発生する。このため、ニューロンのしきい値に入力される外部刺激は伝搬しながら徐々に小さくなり、ニューロンの動作に偏りが生まれる。

## 3. 結果

この章では、提案する MLP の特性についてカオス時系列データを学習させることによって明らかにする。なお、このカオス時系列には式 (5) で与えられるロジスティックマップを用いる。

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

ロジスティックマップは  $a$  を変更することで特性が変化することで知られており、本研究では、 $a = 4.0$  としてシミュレーションを行う。 $a = 4.0$  のとき、ロジスティックマップは、ランダム振動のように振舞う。私たちは、比較のために 5 つの異なった MLP を用いた。

- 従来法による MLP
- グリアネットワークを持った MLP ( $\alpha = 0.15, \beta = 0.9$ )
- グリアネットワークを持った MLP ( $\alpha = 0.2, \beta = 0.8$ )
- グリアネットワークを持った MLP ( $\alpha = 0.4, \beta = 0.6$ )
- ランダムノイズ注入 MLP (ノイズの振幅 = 0.1)

なお、 $\alpha$ 、 $\beta$  およびノイズの振幅は、すべてのグリアの振幅値の平均が 1.0 となるように定義した。

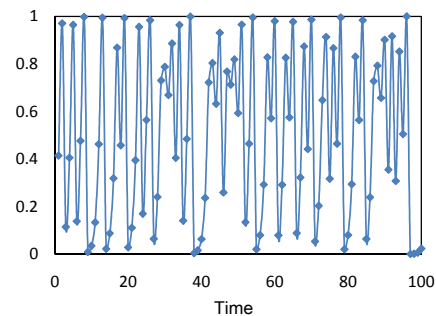


図 2 カオス時系列データ ( $a = 4.0$ ).

図 2 は、ロジスティックマップに、ある初期値を与えたとき生

成されるカオス時系列データである。図3のように、各 MLP はカオス時系列上の連続した2つの値が入力され、その値の時系列的に次の値を教師データとして受ける。その後、教師データと MLP の出力の差から誤差を求め、BP によって結合荷重を変化させることにより学習する。なお、2 入力と教師データを1つの学習データセットとし、100 個の異なる学習データセットとして学習する。

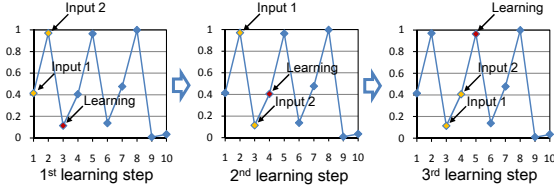


図3 MLP の学習手順。

本研究では、上記のようにカオス時系列を学習したときの、MLP の学習性能を3種類の手法によって示す。

- 関数近似性能
- 被ダメージ性能
- 汎化性能

また、比較指標として、式(6)で表される平均二乗誤差 (MSE) を用いる。

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

### 3.1 関数近似性能

私たちは、学習後の MLP の出力と教師データの差から MSE を求め、MSE の大きさから各 MLP の関数近似性能を示す。なお、各 MLP は 500000 回学習を繰り返すとともに、異なる初期値を結合荷重に与え 200 回シミュレーションを行う。また、結果の比較として全シミュレーション中の MSE の平均 (Avg. Err.), 最小値 (Min.), 最大値 (Max.), および標準偏差 (St. Dev.) を示す。

表1 関数近似性能 ( $10^{-3}$ ).

	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

表1の誤差の平均から、従来法による関数近似は他の MLP に比べて性能が劣ることがわかる。対して、ランダムノイズ注入 MLP が最も高い関数近似性能を示した。また、 $\beta$  が大きくなるに従い MSE の平均が小さくなっていることがわかる。MSE の最小値は、5 種類の MLP の結果がほとんど等しくなっている。このことから、局所解の影響が小さい場合は従来法であっても関数近似が同程度に可能であることが言える。以上の結果より、関数近似においては、MLP の学習中に加えられるノイズは一律な振幅を持っていたほうが良いことが言える。ま

た、加えられるノイズの振幅が一律でない場合、近似性能が減退することから一部のニューロンが局所解から脱出できなくなると考えられる。

### 3.2 被ダメージ性能

次に、MLP を構成する隠れ層のニューロンを1つ除去したときの関数近似性能を示す。図4は、1番目のニューロンを除去した例である。学習時、隠れ層のニューロンは上記の関数近似性能の学習と等しく10ユニット持っている。500000回学習を行った後、隠れ層のニューロンを1つ除去する。また、除去するニューロンの位置を1番目から10番目まで変化させる。以上を200回繰り返し、その平均から MLP の被ダメージ性能を示すとともに、ニューロンに与えられる振幅の変化による学習後のニューロンの状態の変化を示す。

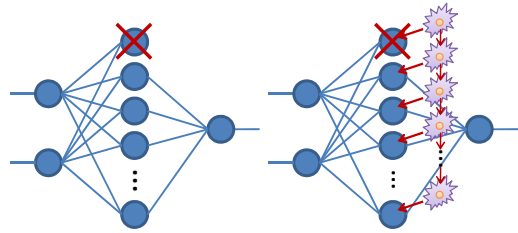


図4 一つのニューロンが除去された MLP の例。

図5は5種類の MLP に関する被ダメージ性能である。横軸は除去したニューロンの番号、縦軸は MSE を示している。図中において、MSE が大きくなるということは、除去したニューロンが MLP の出力の制御に関して重要な働きをしているということである。対して、MSE が小さくなるということは、そのニューロンが除去されたとしても学習結果に大きな影響が出ないことを示している。図5より、従来法やランダムノイズ注入 MLP はグラフが水平に近づいている。これは、各 MLP 中のニューロンが均等に学習されていることを示している。その要因として、ある程度は初期値に依存するものの、各層内のニューロンの学習は等価であることが考えられる。対して、グリアネットワークを結合した MLP においては、 $\beta$  の大きい MLP のほうがよりグラフの傾きが大きくなる傾向が見られる。また、MSE の変化の小さい除去ニューロンは1番目に偏っている。この1番目のニューロンは、グリアネットワーク中で最も大きな外部刺激を受けている部分である。この結果から、大きな外部刺激を受けたニューロンは、MLP 中での重要度が低くなると考えられる。

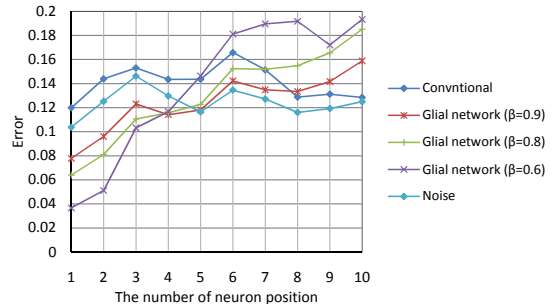


図5 一つのニューロンが除去されたときの MLP の関数近似性能。

ここで、図5の結果は、200回のシミュレーションの平均であるため結果のばらつきが考慮されていない。そこで次に、200回のシミュレーションで得たMSEの値を図6-9にヒストグラムとして表す。なお、左から順番に隠れ層の1番目のニューロン、4番目のニューロン、10番目のニューロンをそれぞれ除去したときの結果である。従来法では、3つのヒストグラムに差が見られない。この結果から、ニューロンの位置に依存せず学習していることがわかる。

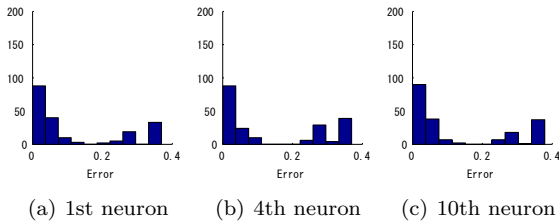


図6 従来法の被ダメージ時MSEヒストグラム。

次に、グリアネットワークを結合したMLPでは、図5と同様に、1番目のニューロンのヒストグラムがMSEの小さい範囲に集中している。対して、10番目のニューロンのヒストグラムでは、全体に分散し平均的になっている。さらにこの傾向は、 $\beta$ が大ききとき、より強くなることが図7と図8の比較からいえる。

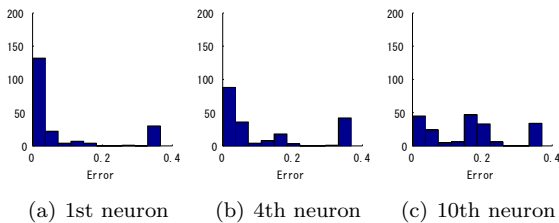


図7 グリアネットワークを持ったMLPの被ダメージ時MSEヒストグラム ( $\alpha = 0.15$ ).

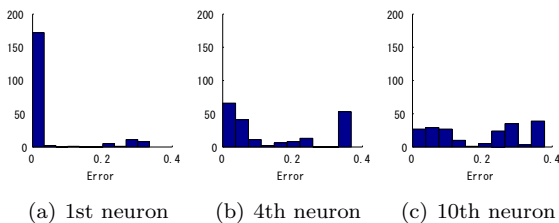


図8 グリアネットワークを持ったMLPの被ダメージ時MSEヒストグラム ( $\alpha = 0.40$ ).

ランダムノイズ注入MLPは、ほぼ従来法と近い結果となった。ただし、MSEが0.2付近での分布が少し変化している。

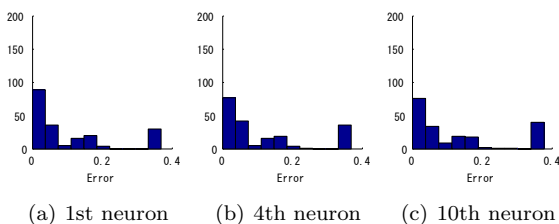


図9 ランダムノイズ注入MLPの被ダメージ時MSEヒストグラム。

### 3.3 汎化性能

最後に、MLPの汎化性能について示す。このシミュレーションでは、学習後のMLPに対して学習に用いた初期値とは異なる初期値から生成したカオス時系列データを与える。入力に用いたカオス時系列データの例と学習に用いたカオス時系列データを図10に示す。なお、このシミュレーションでは、同様にして生成したテストデータを10個用意し、200個の学習後のMLPに入力したときのMSEの平均を示す。

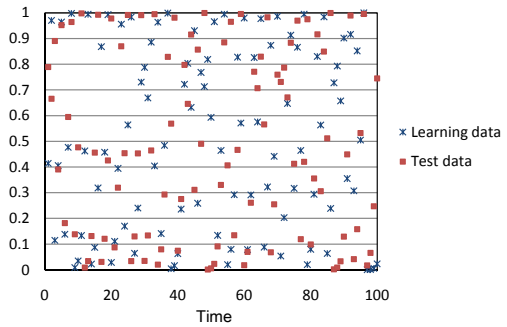


図10 学習データとテストデータ。

表2は、各MLPの汎化性能のシミュレーション結果である。関数近似性能で示した学習結果の性能と同様の傾向を示しているが、従来法と他のMLPとの差が小さくなっている。ただし、従来法では標準偏差が非常に大きくなっており、結果のばらつきが他のMLPと比して大きいと言える。

表2 汎化性能 ( $10^{-3}$ )

	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

次に、テストデータに対して最大5%のノイズを加えた場合のMSEを示す。ノイズが含まれるため入力されるテストデータは、MLPが記憶している関数とは異なったデータとなる。このシミュレーションによって、MLPのノイズ除去性能を調べることが可能である。表3に結果を示す。

上記のノイズを含まない場合と比べて性能が劣化していることがわかる。さらに、従来法は他のMLPと比較して悪い結果となっていることから、ノイズ除去性能は学習時にノイズを与えられていた場合のほうが高いと言える。また、 $\beta$ が小さくなるに従い結果が良くなるとともに、ランダムノイズ注入MLPが最も良い性能を示した。

## 4. まとめ

本研究では、グリアネットワークによって伝搬された外部刺激によって、MLPの学習がどのように変化するかを示した。1番目のグリアに入力された外部刺激は、グリアネットワークを通じて他のグリアに伝搬される。伝搬中に振動が減衰されてい

表 3 ノイズを含むデータによる汎化性能 ( $10^{-3}$ )

	Conv.	Glia	Glia	Glia	Noise
	-	$\beta = 0.9$	$\beta = 0.8$	$\beta = 0.6$	-
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

くために各ニューロンに入力される振動は振幅の異なった値となる。

シミュレーション結果より、従来法と比較して振動を入力した MLP は関数近似性能が高いことを示した。また、外部刺激を強く受けるニューロンは、MLP の出力に寄与する割合が減少する傾向が見られた。現実の脳内のニューラルネットワークの働きにおいても、それぞれのニューロンの存在する位置に対してその働きが依存すると考えらる。私たちは、人工ニューラルネットワークにおいても、このような位置依存性を考慮する必要があると考えており、本研究がその第一歩となると期待している。

#### 文 献

- [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] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.
- [5] C. Ikuta, Y. Uwate and Y. Nishio, 'ニューロ・グリアネットワークをもった階層型パーセプトロンの提案', 電気情報通信学会 非線形問題研究会 技術報告, no. NLP2010-31, pp. 13-17, Jul. 2010.