

修士論文

SpikeProp ネットワークの
汎化能力の向上
—不連続な出力変化をもたらす入力の検出—

平成26年度修了

三重大学大学院 工学研究科

博士前期課程 電気電子工学専攻

北川 陽一郎

目次

第1章	はじめに	1
1.1	背景	1
1.2	研究の概要	2
第2章	SpikeProp ネットワーク	4
2.1	ユニットのモデル	4
2.2	ネットワークの構成	5
2.3	学習法	6
2.4	不連続な入出力関係	7
第3章	不連続な入出力関係をもたらす入力の判定	10
3.1	不連続な入出力関係を軽減するための基本方針	10
3.2	問題がある入力の検出法	11
3.2.1	基本方針	11
3.2.2	問題がある入力に対して現れる内部電位の特徴	12
3.2.3	問題がある入力の判定方法	14
第4章	実験	16
4.1	実験条件	16
4.2	評価方法	18
4.3	実験結果と考察	18
4.3.1	θ_{1+} についての検討	19
4.3.2	θ_{1-} についての検討	20
4.3.3	θ_{2+} についての検討	21
4.3.4	θ_{2-} についての検討	22
4.3.5	t_θ についての検討	24
4.3.6	判定精度	25
4.4	興味深い事柄	26

第5章 まとめと今後の課題	27
謝辞	28
参考文献	29
発表	32

表 目 次

2.1 SpikeProp における学習成功率と不連続な入出力関係を示す割合 . . .	9
4.1 時間版 XOR 問題	16
4.2 実験での基本的なパラメータ表	17
4.3 パターンに対して提案法を使用したときの四分表	26
4.4 判定法の陽性検出率, 陰性検出率	26

目 次

2.1	スパイクングニューロンモデル	4
2.2	ユニット j の内部電位	5
2.3	SpikeProp ネットワークの構造	6
2.4	時間版 XOR 問題で学習したときの SpikeProp の入出力関係を示す カラーマップ	8
3.1	学習点対策のフローチャート	11
3.2	問題がある入力に対して現れる特徴 (ケース 1)	13
3.3	問題がある入力に対して現れる特徴 (ケース 2)	13
3.4	ケース 1 に対する判定法	15
3.5	ケース 2 に対する判定法	15
4.1	θ_{1+} に関する ROC 曲線	19
4.2	θ_{1-} に関する ROC 曲線	20
4.3	θ_{2+} に関する ROC 曲線 (θ_{2-} 変化)	21
4.4	θ_{2+} に関する ROC 曲線 (t_θ 変化)	22
4.5	θ_{2-} に関する ROC 曲線 (θ_{2-} 変化)	23
4.6	θ_{2-} に関する ROC 曲線 (t_θ 変化)	23
4.7	t_θ に関する ROC 曲線 (θ_{2+} 変化)	24
4.8	t_θ に関する ROC 曲線 (θ_{2-} 変化)	25

第1章 はじめに

1.1 背景

昨今の計算機の性能の向上はめざましく、それに伴い、計算機の役割は非常に身近で大きなものとなっている。計算機により、さまざまな情報処理が正確・高速に行われるようになってきているが、一部の処理はいまだ人間に及ばない。そのような処理の一つが、文字、音声、画像などの認識である。このような処理を、人間などの情報処理手法を模倣することにより獲得する試みが広くなされている [1, 2].

人工ニューラルネットワークは、このような試みの一つである [3, 4, 5, 6]. これは、人間の脳が、何千億個もの神経細胞のネットワークにより行われていることを模倣した計算モデルである。各神経細胞は高い計算能力を持たないが、これが多数集まりネットワークを構成することで、学習能力・汎化能力といった能力を獲得する。学習能力とは、ある入力に対して期待する出力を得ることができるようにネットワークを自動的に調整する能力である。また、汎化能力とは未知の入力に対してもっともらしい出力を出力できる能力である。この能力により、雑音などの外的要因で入力が学習したものから変動しても、期待した動作を続けることができる。

これまでの人工ニューラルネットワークに関する研究は大きく三段階に分けることができる。

第一段階は、「パーセプトロン」に関する研究である。これは、1958年に計算機科学者の Rosenblatt によって発表された [7]. これは、線形分離問題のように簡単なパターン認識の問題を識別できるネットワークである。複数のパターン（パターンセット）により表現される入出力関係を、パーセプトロンは学習できる。しかし、1969年に Minsky と Papert によって、XOR 問題などの線形分離不能なパターンセットを学習できないという指摘 [8] がなされ、多くの研究者が興味を失った。

第二段階は、「BP (Back Propagation) モデル」に関する研究である。これは、

1986年にRumelhartらによって提案された、多層パーセプトロンとその学習法である誤差逆伝搬法（BP法）である[9]。線形分離不能なパターンセットを含む任意のパターンセットにより表される入出力関係を、BPモデルは学習できる。そのためニューラルネットワークは認知分野、工学分野で応用されている。

第三段階は、スパイクニューラルネットワークに関する研究である。スパイクニューラルネットワークは、スパイクニューロンによって構成されたネットワークである。パーセプトロンやBPモデルでは、各ユニットは入力を入力に変換する一種の関数として働いていた。スパイクニューロンは、神経細胞がスパイク（パルス）の列を受け取り、それに応じてパルス列を出力する動作をモデル化したものである。これを用いてネットワークを構成することで、時間を含んだ情報を自然な形で処理できる[10, 11, 12]。スパイクニューラルネットワークの研究は古くから行われてきたが、半導体技術の発展・計算機的能力向上などにより、これをパターン認識などに応用する研究が行われるようになってきた。

スパイクニューラルネットワークのモデルは情報をどのように表現するかで大きく2つに分けられている。第1はニューロンの平均発火率によって情報を表現するモデルである[13, 14]。第2はニューロンの発火時刻によって情報を表現するモデルである[15, 16, 17]。これは、一つのスパイクの観測により情報を得ることができるため、素早く情報を得ることができる[18, 19]。

本研究では、後者のような情報の表現方法にスパイクの発火タイミングを用いるモデルの一つであるSpikePropネットワークを対象とする。SpikePropは計算能力が高く[20]、実装が容易であるという利点がある。これはパターン認識などの分野での応用が期待されている。

1.2 研究の概要

これまでの研究でSpikePropネットワークを詳細に調べた結果、ニューラルネットワークに必要な「汎化能力」がしばしば低くなることがわかった[21]。具体的には、連続的に入力を変化させても、出力が不連続に変化する場合があった。以下、このような入出力関係を「不連続な入出力関係」と呼ぶ。このような入出力関係のもとでは、わずかな入力の変化で、大きな出力の変化がもたらされる。これは、

同一の入力パターンに対しても、わずかな雑音により全く異なる値が出力されることを意味しており、応用の際の問題となる。

本研究では、不連続な入出力関係を解消することで、SpikeProp ネットワークの汎化能力を向上させることを目的とする。不連続な入出力関係の遠因は、スパイクングニューロンがスパイクを出力する機構にある。しかし、この機構はスパイクングニューラルネットワークの好ましい特徴ももたらす。そこで、スパイクングニューロンの発火のメカニズムを変更することなく、この不連続な入出力関係を解消することをめざす。ここでは、汎化能力の向上を目的としているため、不連続な入出力関係を完全に解消するのではなく、学習した入力付近のみで解消することにする。

この目的に向けて、これまでにいくつかの研究が行われてきた [21, 22, 23]。吉川らは、不連続な入出力関係が生じにくいネットワーク構造を獲得するように、本来の SpikeProp の学習法に荷重減衰法 (Weight Decay [22, 24]) を適用し、不連続な入出力関係を減らすことに成功した [23]。李らは、この研究を発展させ、有効な荷重減衰の方法について検討し、二乗型のペナルティ項を用いた荷重減衰法が、不連続な入出力関係の解消に最も効果があることを示した [21]。しかし現時点では、学習後の約半数のネットワークは、学習したいずれかの入力付近で不連続な入出力関係を示しており、さらなる改善が必要である。

本研究では、これまでの不連続な入出力関係が生じにくいネットワーク構造を獲得する方針のままではこれ以上の改善は困難であると考え、問題がある入力に対して個別に対策するための方法について検討する。ここで、問題がある入力とは、その近傍に不連続な入出力関係が生じている入力のことである。本研究では、その第一歩として、問題がある入力を簡便に検出する方法について検討する。

第2章 SpikeProp ネットワーク

SpikeProp ネットワークとは, Bohte らにより提案されたスパイク型ニューラルネットワークのモデルの一つで階層型ニューラルネットワークである. 本章では, 積分発火型ニューロンモデル, SpikeProp ネットワークの構造, 学習, ネットワークの不連続な入出力関係と汎化能力について説明する.

2.1 ユニットのモデル

スパイク型ニューロン (以下ユニット) とは, ニューラルネットワークの構成要素の一つであり, 信号を変換する機能を担う. SpikeProp の場合, そのモデルは積分発火モデルに基づいている. その概要を図 2.1 に示す. 各ユニットは結合によって接続され, その接続点をシナプスと呼ぶ. 各ユニットの出力は, そのユニットへと接続するユニットが出力するスパイクの時間により定まる.

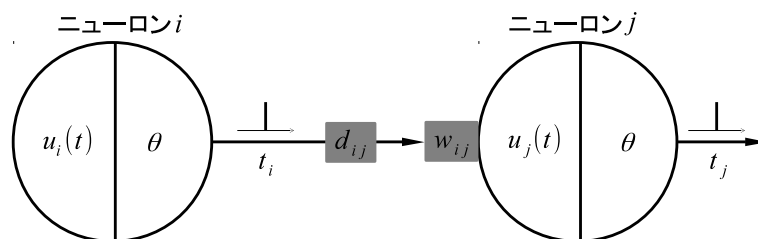


図 2.1: スパイク型ニューロンモデル

ここで d_{ij} , w_{ij} はユニット i からユニット j への結合のシナプスに付与された「時間遅れ」と「結合荷重」である. また $u_i(t)$, $u_j(t)$ は時刻 t でのユニット i , j の内部電位を表す. まずユニットにスパイクが入力されると, シナプスにおいて

時間遅れの演算が施される．次にスパイクに対する応答が式 (2.1) のスパイク応答関数 $\varepsilon(t)$ により与えられ，この応答を結合荷重により重みづけしたうえで合計することで，ユニットの内部電位をえる．

$$\varepsilon(t) = \begin{cases} 0 & (t < 0) \\ \frac{t}{\tau} \exp\left(1 - \frac{t}{\tau}\right) & (t \geq 0) \end{cases} \quad (2.1)$$

ここで τ は時定数である．ユニット内部では，すべての結合からの入力値を合計することで内部電位を得る．式で表すと式 (2.2) のようになる．

$$u_j(t) = \sum_{i \in \Gamma_j} w_{ij} \varepsilon(t - t_i - d_{ij}) \quad (2.2)$$

ここで t_i , t_j はそれぞれユニット i , j の発火時刻を表しており， Γ_j はユニット j の前のユニットの集合を表している．ユニット j の内部電位の状態は例えば図 2.2 のようになる．この図の時刻 t_j のように，内部電位が最初に出力しきい値 θ を超えたときにユニット j は発火する．ただし SpikeProp ネットワークでは，ユニットは最初に θ を超えたときにのみ発火する．

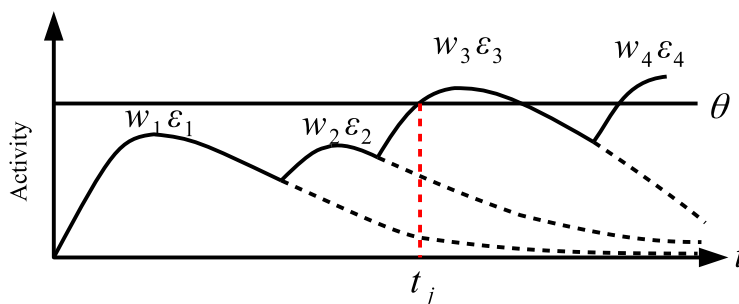


図 2.2: ユニット j の内部電位

2.2 ネットワークの構成

SpikeProp ネットワークは，階層型の構造をもつ (図 2.3)．ユニットはネットワークの中間層，出力層のいずれかに配置されており，各ユニットは前の層から接続され，次の層へと接続している．ネットワークの入力は入力層に与えられ，ネットワークの出力は出力層から得る．各結合は複数の副結合で構成されており，そ

それぞれの副結合には時間遅れと結合荷重が付与されている。簡略化のため時間遅れは副結合番号のみにより決まるものとする。

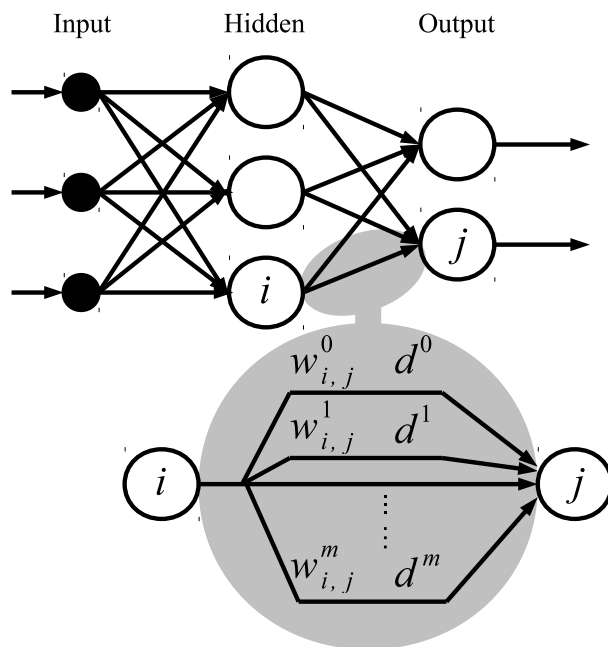


図 2.3: SpikeProp ネットワークの構造

これをふまえると、式 (2.2) は次のように書き直される。

$$u_j(t) = \sum_{i \in \Gamma_j} \sum_k w_{ij}^k \varepsilon(t - t_i - d^k) \quad (2.3)$$

ここで、 d^k は k 番目の副結合の時間遅れを、 w_{ij}^k はユニット i からユニット j への k 番目の結合荷重を表す。

2.3 学習法

SpikeProp ネットワークは結合荷重を調整することで、所望の入出力関係を実現できる。すなわち、ある時刻のネットワークへの入力スパイクに対して、所望の時刻にネットワークの出力スパイクを得ることができる。なお、この関係を複数組学習できる。これを行うための手法のひとつが Bohte らによって提案された誤差逆伝搬法に基づいた手法である。この手法は誤差の勾配に基づいて結合荷重を繰り返し調整することで学習を行う。SpikeProp の学習では、ネットワークの出力

層の出力，すなわちネットワークの出力スパイクの時刻と，あらかじめ与えられた希望の出力時刻との差の二乗を誤差とする（式(2.4)）。

$$E = \frac{1}{2} \sum_{j \in J} (t_j^a - t_j^d)^2 \quad (2.4)$$

ここで， t_j^a は出力層Jに属するニューロン j の実際の出力時刻， t_j^d は希望する出力時刻である．このとき，一回の結合荷重の更新量 Δw_{ij}^k は式(2.5)のようになる．

$$\Delta w_{ij}^k = -\eta \frac{\partial E}{\partial w_{ij}^k} \quad (2.5)$$

ここで， η は学習率であり，結合荷重の更新量の大きさを表す正の定数である．出力の計算，結合荷重の更新を誤差が十分に小さくなるまで繰り返すことで，所望の動作をするネットワークを得ることができる．

2.4 不連続な入出力関係

不連続な入出力関係とは，入力の変動的な変化にもかかわらず，出力が不連続に値が大きく変化する関係のことである．この変化が大きい時 SpikeProp の汎化能力が低下する．また，SpikePropにおける汎化能力の低下は，わずかな入力変化に対して連続的だが大きな出力変化が生じる場合にももたらされる．本論文では，このような入出力関係も含めて「不連続な入出力関係」と呼ぶ．数学的な不連続の定義とは異なるので注意する．

実際の SpikeProp において，不連続な入出力関係が生じている例を図 2.4 に示す．これは，2入力1出力の SpikeProp の入出力関係をカラーマップとして示したものである．縦軸，横軸は2つの入力を意味し，各入力に対する出力が色により表されている．入力，出力の値はともにスパイクの時刻である．図 2.4 において，色が大きく変化している部分が，不連続な入出力関係を示している．この図においては，左下角にある赤と紫の境界において，不連続に出力が変化している．

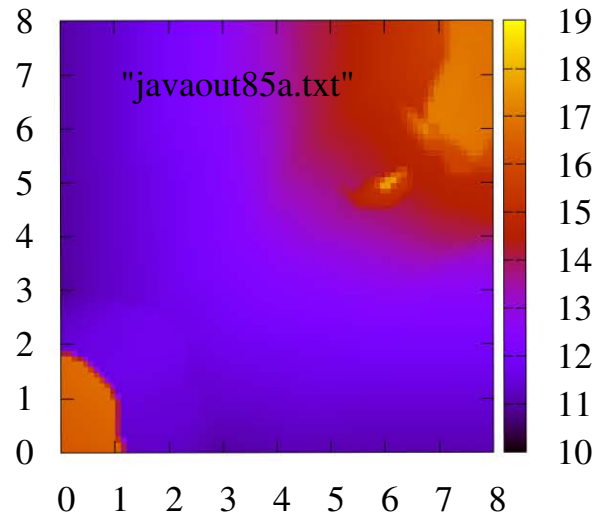


図 2.4: 時間版 XOR 問題で学習したときの SpikeProp の入出力関係を示すカラーマップ

この SpikeProp は 4 章の実験でも使用する時間版 XOR 問題を学習したネットワークであり、その学習した入力は、 $(1,1)$, $(1,7)$, $(7,1)$, $(7,7)$ である。これを「教師入力」と呼ぶ。教師入力 $(1,7)$, $(7,1)$, $(7,7)$ の近傍には色が大きく変化している所はないが、教師入力 $(1,1)$ の近傍では、色が大きく変化している。実際、教師入力 $(1,1)$ に対する出力は 16.75 であるが、わずかにずれた教師入力ではない入力 $(1.2,1)$ に対する出力は 11.53 である。このネットワークは、学習した入力とわずかに異なる入力に対して、所望の動作をせず汎化能力が低いと言える。本研究では、ある入力の近傍に不連続な入出力関係が生じている場合、その入力を「問題がある入力」と呼ぶ。今回の例では 4 つの教師入力のうち $(1,1)$ の近傍で先ほどの不連続な入出力関係の条件にあてはまる入出力関係が存在したため、入力 $(1,1)$ は問題がある入力である。逆に教師入力 $(1,7)$, $(7,1)$, $(7,7)$ の近傍には不連続な入出力関係が存在しなかったため問題がない入力である。また教師入力全てのうち、いずれか 1 つでも問題がある入力を含んだネットワークのことをここでは「問題があるネットワーク」と呼ぶ。今回の例のネットワークは、問題がある教師入力 $(1,1)$ があるので、問題があるネットワークである。教師入力に問題がある入力を含まないようにすることが、本研究の目的である。ある入力「問題がある入力」であるか否かは、入力の「わずかな」変化に対して出力が「大きく」変化するかどうかにより定まる。これらの変化の大きさの許容値は、パターンと用途により決まる。

李らは、不連続な入出力関係を学習した入力付近で生じないようにする学習法を提案した [21]。これにより Bohte らによる学習法では、問題があるネットワークが学習後に 94.5% あったのを、51.0% にまで減らした。ここで、李らの手法により学習したネットワークの入出力関係を分析したところ、表 2.1 に示す結果を得た。これは、問題があるネットワークと教師入力、学習に成功したネットワークに対して占める割合を表している。学習に成功したかどうかの判断は文献 [21] と同様に行った。

表 2.1: SpikeProp における学習成功率と不連続な入出力関係を示す割合

	オリジナル	Weight Decay 二乗型
学習成功率	99.5	92.0
問題があるものの割合		
ネットワーク	94.5	51.0
教師入力	58.9	22.4

[%]

表 2.1 の問題がある教師入力の割合は、問題があるネットワークの割合に比べて少ない。これは「ネットワーク」では 4 つの教師入力のうち 1 つでも問題があれば、それは問題があるネットワークと判定される。しかしほとんどの場合、全ての教師入力に問題があるというわけではない。つまり 4 つの教師入力のうち 3 つは問題がない教師入力であるにもかかわらず、残りの 1 つの教師入力の原因でネットワーク全体として問題ありと判定されている。そのため、表 2.1 のように問題があるネットワークの場合と比べて問題がある教師入力が少なくなる。図 2.4 の例においても、問題があるネットワークの 4 つの教師入力のうちただ 1 つ (1,1) の教師入力だけに問題がある。そこで、従来の方法のネットワーク構造を好ましいものにするという方針では、問題があるネットワークの割合をこれ以上減らすことは難しいと考えた。

以上より、不連続な入出力関係を解消するためには、ネットワーク構造を好ましいものにするアプローチだけではなく、個別に問題がある教師入力 (図 2.4 の例では教師入力 (1,1)) に対処する必要がある。次章で、このような手法を検討する。

第3章 不連続な入出力関係をもたらす入力の判定

この章では、個別に問題がある教師入力に対処することで、SpikeProp ネットワークの不連続な入出力関係を解消する手法について検討する。特に、問題がある教師入力を検出する手法について検討する。

3.1 不連続な入出力関係を軽減するための基本方針

2.4節で述べたように、本研究では、不連続な入出力関係を解消する基本方針として、個別の問題がある教師入力に対処するアプローチをとる。すなわち、従来法により不連続な入出力関係が出にくくなるようにし、それでも現れる不連続な入出力関係に対して個別に対処する。このアプローチの基本的なアルゴリズムを図3.1に示す。

青色の部分の手続きが従来法（李らの手法）に新たに加えられた手続きである。

この手法の提案には、(1) 問題がある教師入力の判定 (2) それに対する対処法に関する検討が必要である。本論文では、(1) について検討する。図3.1に示すとおり、学習ステップごとに各教師入力に対してこの判定を行う。そのため、この新しい手続きによる追加の計算量が少ないことが求められる。この手法では、問題があるとされた入力に対してのみ対処をする。そのため、問題がある入力を問題がないと判定すること（偽陰性）は好ましくない。逆に、問題がない入力を問題があると判定すること（偽陽性）は対処のコストが増加するだけなので、多少は許容してもよい。これらをまとめると、提案する手法に対する要件は次の2つである。

- (A) 追加の計算量が少ないこと
- (B) 偽陰性が極力少ないこと（多少の偽陽性は許容する）

これらを満たす検出法について、次節で検討する。

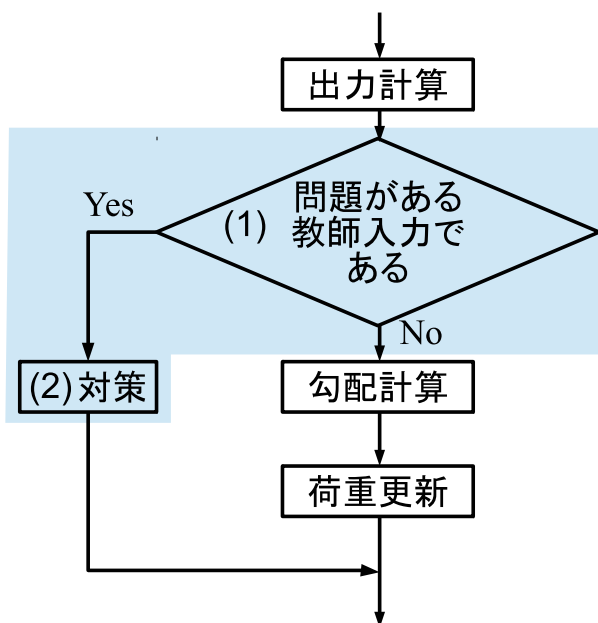


図 3.1: 学習点対策のフローチャート

3.2 問題がある入力の検出法

3.2.1 基本方針

教師入力、問題がある入力かどうかを判定する安直な方法は、教師入力の近傍における入出力関係を多数調べ、大きく異なる出力がされているかどうかにより判定する方法である。しかし、入力ユニット数が増加すると、この数は指数関数的に増加するため現実的な手法とは言えない。例えば、李らは2入力のネットワークにおいて、399点の近傍（各入力について20種の値）をとり、これらに対するネットワーク出力に基づいて判定していた。

本研究では、ユニットの内部電位に着目する。すなわち、検査入力に対する各ユニットの内部電位を調べ、問題が発生する兆候の有無により判定を行う。ユニットの出力は内部電位が発火しきい値を超えることで行われる。そのため、ユニットの内部電位を観測することで、その出力が変化する兆候をとらえることができると考えた。この方法は、教師入力に対するネットワークの挙動のみから判断で

き、教師入力の変動を検査する必要はない。また、教師入力に対する挙動は学習するために必要なものであるため、判定のために追加で求める必要はない。そのため、3.1節に示した要求(A)を満たす。

次節で、問題がある入力に対して、内部電位に現れる特徴について調査する。

3.2.2 問題がある入力に対して現れる内部電位の特徴

本節では、実際の学習後のネットワークについて、さまざまな教師入力に対して、各ユニットの内部電位の時間変化を観測した。ネットワークは、4章の実験で使用する教師パターンセットを李らの手法により学習したものである。観測の結果、問題がある教師パターンを入力した時には、以下に示す2つの代表的な挙動が見いだされた。

第1の挙動は図3.2のように発火しきい値と近い極値を持つ極大が存在するケースである。入力のわずかな変化により、内部電位がわずかに変化する。このとき、この極大の極大値が発火しきい値をまたいで変化することで、発火時刻が大きく変化する。図3.2では、時刻 t_0 で発火していたユニットが、内部電位がわずかに変化することで、時刻 t'_0 で発火する。極大値が発火しきい値に近いほど、問題が発生しやすい。

第2の挙動は図3.3のように発火しきい値に近い値を内部電位がとり続けるケースである。入力のわずかな変化により、内部電位がわずかに変化する。この場合、内部電位の変化がわずかであっても、発火時刻は大きく変化してしまう。図3.3では、時刻 t_0 で発火していたユニットが、内部電位がわずかに変化することで、時刻 t'_0 で発火する。内部電位が発火しきい値に近い値をとり続ける時間が長いほど、発火時刻の差が大きくなる。

以上の特徴は、ユニットの発火メカニズムから考えても妥当であり、XOR問題に限らないものであろう。

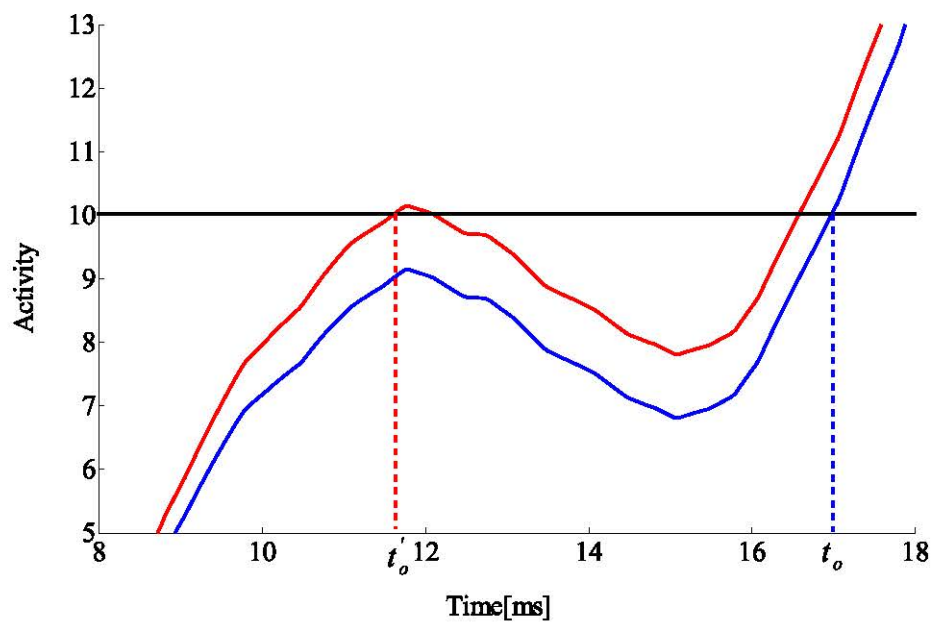


図 3.2: 問題がある入力に対して現れる特徴 (ケース 1)

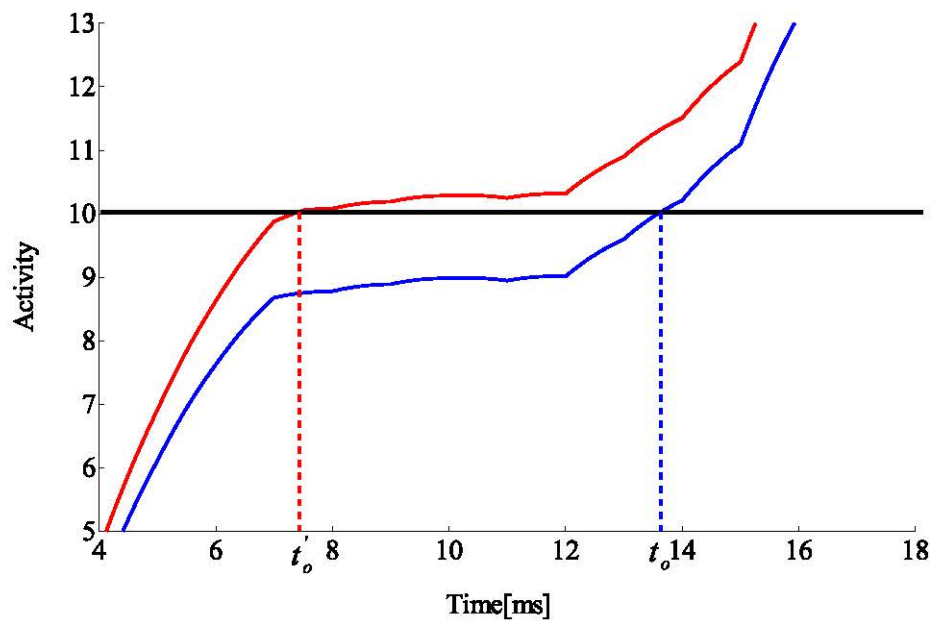


図 3.3: 問題がある入力に対して現れる特徴 (ケース 2)

3.2.3 問題がある入力の判定方法

前項の分析をふまえ、与えられたネットワークの入力（検査入力）が、問題のある入力であるかどうかを判定する方法を提案する。この手法は、前節で述べた2つの挙動のいずれかが、ネットワーク内のいずれかのユニットで観測された場合、そのときの入力を問題がある入力（陽性）として判定する。以下が具体的な手順である。これは、図3.1における青色の(1)の部分に対応している。

- (1) 検査入力に対する、各ユニットの内部状態の時間変化を調べ、以下に該当するかどうかを判定する。
 - (A) $\theta + \theta_{1+}$ 以下かつ $\theta - \theta_{1-}$ 以上内部電位が存在する。
(図3.4の赤い領域の部分に極大（赤い円部分）が存在する.)
 - (B) 連続して $\theta + \theta_{2+}$ 以下かつ $\theta - \theta_{2-}$ 以上となる最大の区間が t_θ 以上の長さを持つ。
(図3.5の青い領域の部分に、連続して入っている時間（青い矢印部分）の長さが t_θ 以上)
- (2) 該当すれば、検査入力は問題があると判定する。

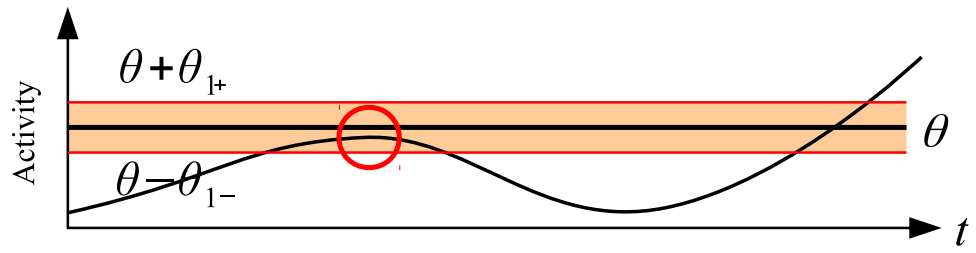


図 3.4: ケース 1 に対する判定法

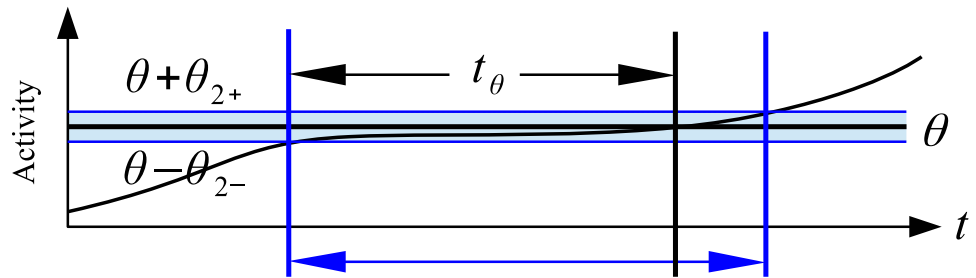


図 3.5: ケース 2 に対する判定法

第4章 実験

この章では、3章で提案した判定法で連続・不連続を検出できるかどうかを検討する。また判定法の5つのパラメータ ($\theta_{1\pm}$, $\theta_{2\pm}$, t_θ) と判定精度との関係を議論する。

4.1 実験条件

本節では、実験で扱う時間版 XOR 問題、時間版 XNOR 問題とネットワーク構造について説明する。

排他的論理和 (XOR) は線形分離不可能なパターン分類の問題の中で最も簡単なパターンであるため、これを学習するためには、単純な入出力関係ではあっても非線形なパターン変換能力を必要とする。このため、ニューラルネットワークの学習用ベンチマークとしてしばしば用いられる。今回は文献 [17] で用いられた時間版 XOR 問題を修正して用いる。表 4.1 に、具体的な入出力を示す。表中の数値はすべて発火時刻を表す。いずれも早い発火時刻を論理値の 1 に、遅い発火時刻を論理値の 0 に対応させている。

表 4.1: 時間版 XOR 問題

入力時刻 1	入力時刻 2	希望出力時刻
1	1	17
1	7	11
7	1	11
7	7	17

[ms]

また、時間版 XNOR 問題は表 4.1 における希望出力時刻の 11 と 17 を入れ替えた問題である。

実験で用いた SpikeProp ネットワーク 基本的な設定は、文献 [17] での実験に準じた条件で行った。

表 4.2: 実験での基本的なパラメータ表

ユニット数	
入力層	3 個
中間層	5 個
出力層	1 個
ニューロンのパラメータ	
スパイク応答関数の時定数	7ms
しきい値	10
結合のパラメータ	
最小時間遅れ	1ms
最大時間遅れ	16ms
副結合数	16 本
学習のパラメータ	
学習率	0.01
最大学習回数	50,000 回
学習終了条件	$E < 0.1$
Weight Decay の減衰率	0.03

実験では、文献 [21] で用いられている学習法により、時間版 XOR 問題で 50 種類、時間版 XNOR 問題で 50 種類、計 100 種類の異なる初期荷重からネットワークをそれぞれ学習し、学習に成功したネットワークの入出力関係を調査した。

今回の実験では、学習後のネットワークにおいて 4 つの学習点近傍が連続か不連続かを判定する。まず各教師入力に問題があるかどうかを文献 [21] の手法にならって判定し、これを正しい判定結果とみなす。この手法は、1 つの学習点につき学習点とその周囲 399 点の出力値を計算し、全ての点でその出力値が学習した出力と比べ、 ± 3.0 であれば「連続」な入出力関係、そうでなければ「不連続」な入出力関係と判断している。そのため、1 つの学習点につき 1 つで、時間版 XOR 問題では 4 つの学習点が存在するため、検査領域は 4 つである。したがって 50 種類の異なる初期荷重で時間版 XOR 問題と時間版 XNOR 問題を学習するため、最終

的な入力の数 は $50 \times 2 \times 4 = 400$ パターンである。

4.2 評価方法

この実験の目的は、3.1節で示した要求(B)について今回提案した判定法が有効であるかどうかを調査するためである。判定法が有効であるかどうかを、陽性検出率、陰性検出率、偽陽性率、偽陰性率の4種類により評価した。

陽性検出率は、問題がある入力に対し、提案法で正しい判定が行えた割合であり、陰性検出率は、問題がない入力に対し、提案法で正しい判定を行えた割合である。陽性検出率と陰性検出率は「正解をどれだけ検出できたかの割合」であるため、この2つの指標は高いほど性能が良いと言える。

偽陽性率は、問題がない入力に対し、提案法で間違った判定をした割合であり、偽陰性率は、問題がある入力に対し、提案法で間違った判定をした割合である。偽陽性と偽陰性は「不正解の割合」であるため、この2つの指標が低いほど性能が良いと言える。また、将来のことを見据えると、偽陰性に至っては、問題がある入力であるにも関わらず、対策が行われないことになる。そのため、偽陽性よりも偽陰性が高くなることの方が深刻である。

4.3 実験結果と考察

まず判定法の性能が5つのパラメータのどれに依存しているかを調べるため、 $\theta_{1\pm}$ に関してはケース1のみを用いて、 $\theta_{2\pm}$ 、 t_θ に関してはケース2のみを用いて、陽性検出率、陰性検出率、偽陽性率、偽陰性率を調べた。そして、横軸を偽陽性率、縦軸を陽性検出率としたROC(Receiver Operating Characteristic) 曲線を使い、判定法の性能の各パラメータへの依存性を調べる。

4.3.1 θ_{1+} についての検討

この節では、パラメータ θ_{1+} と判定精度の関係について検討する。図 4.1 は 11 種の θ_{1+} に対する ROC 曲線群の集合である。各 θ_{1+} において θ_{1-} を 0~5 に変化させた結果が一本の ROC 曲線に対応する。なお、議論を簡単にするため、判定はケース 1 に該当するかどうかのみにより行った。

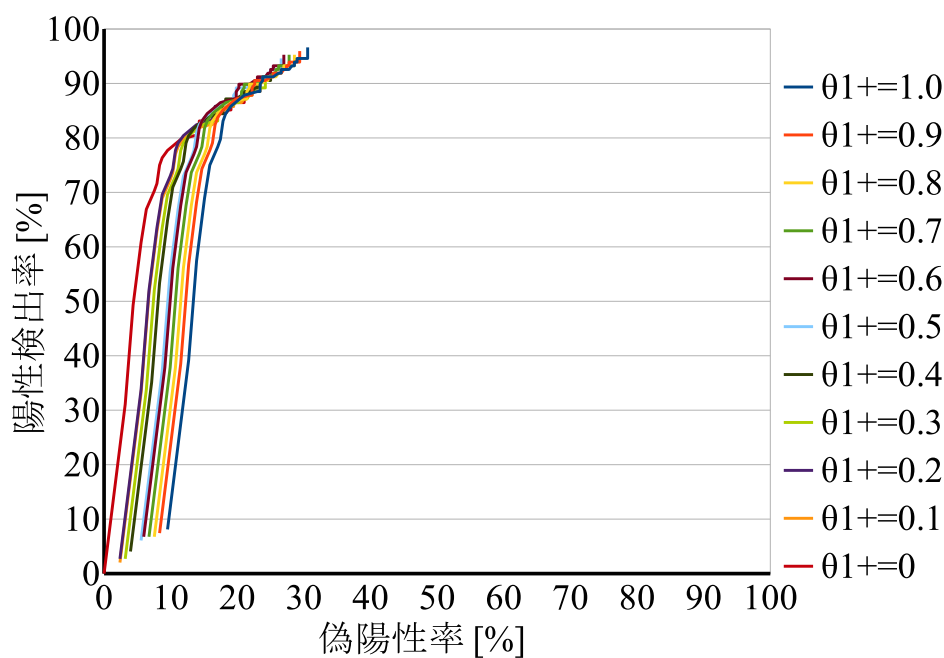


図 4.1: θ_{1+} に関する ROC 曲線

図 4.1 では、 θ_{1+} の値が下がると、ROC 曲線は左に向かう。よって θ_{1+} が下がると横軸「偽陽性率」の値は低くなる。そのため「 θ_{1+} を小さくすると偽陽性が減少する」といえる。実験結果より、ROC 曲線が最も左となったのは、 $\theta_{1+} = 0$ のときであった。

4.3.2 θ_{1-} についての検討

この節では、パラメータ θ_{1-} と判定精度の関係について検討する。図 4.2 は 11 種の θ_{1-} に対する ROC 曲線群の集合である。各 θ_{1-} において θ_{1+} を 0~5 に変化させた結果が一本の ROC 曲線に対応する。なお、議論を簡単にするため、判定はケース 1 に該当するかどうかのみにより行った。

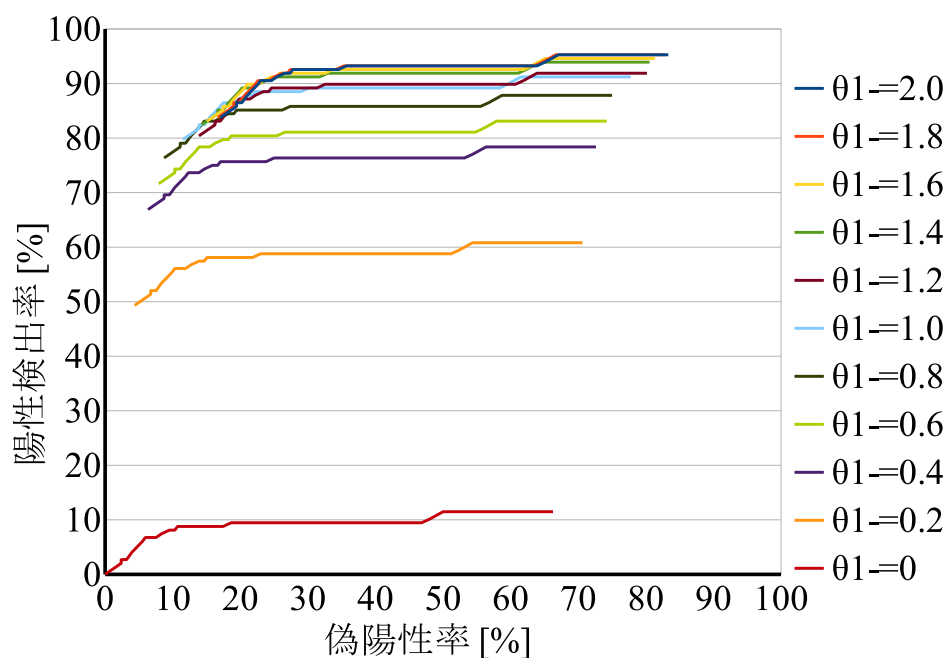


図 4.2: θ_{1-} に関する ROC 曲線

図 4.2 では、 θ_{1-} の値が下がると、ROC 曲線も下に下がる。よって θ_{1-} が下がると縦軸「陽性検出率」の値が低くなる。そのため「 θ_{1-} を小さくすると陽性検出率が減少する」といえる。実験結果より、ROC 曲線が最も上となったのは、 $\theta_{1-} = 2.0$ のときであった。

4.3.3 θ_{2+} についての検討

この節では、パラメータ θ_{2+} と判定精度の関係について検討する。図 4.3、図 4.4 は 11 種の θ_{2+} に対する ROC 曲線群である。各 θ_{2+} において、図 4.3 では θ_{2-} 、図 4.4 では t_θ を 0~5 に変化させた結果が一本の ROC 曲線に対応する。なお、議論を簡単にするため、判定はケース 2 に該当するかどうかのみにより行った。

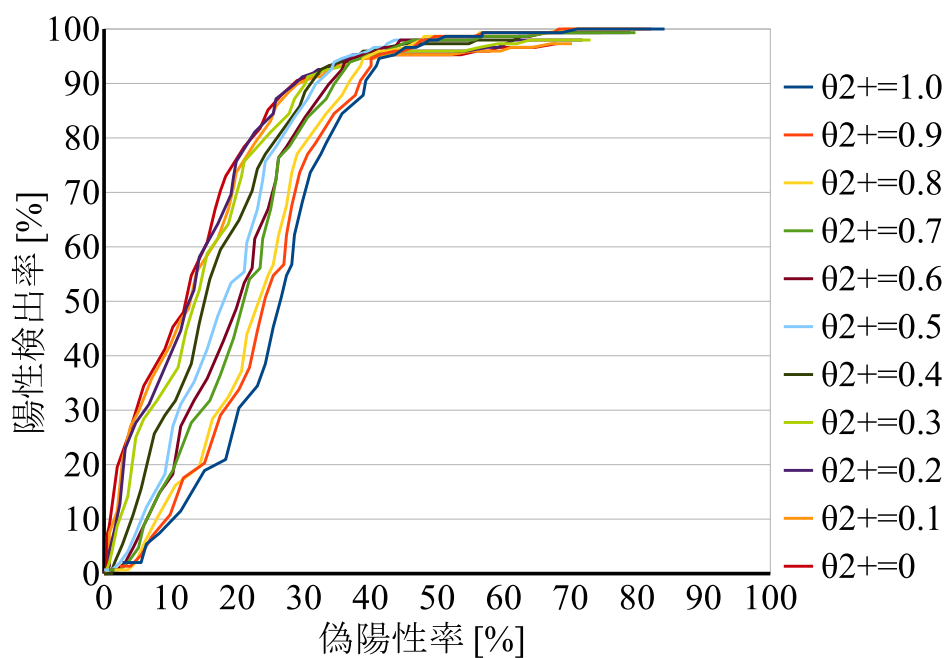


図 4.3: θ_{2+} に関する ROC 曲線 (θ_{2-} 変化)

図 4.3、図 4.4 では、 θ_{2+} の値が下がると、ROC 曲線全体が左に向かい、偽陽性率が下がりやすくなる。そのため「 θ_{2+} を小さくすると偽陽性率が減少する」といえる。実験結果より、ROC 曲線が最も左となったのは、 $\theta_{2+} = 0$ のときであった。

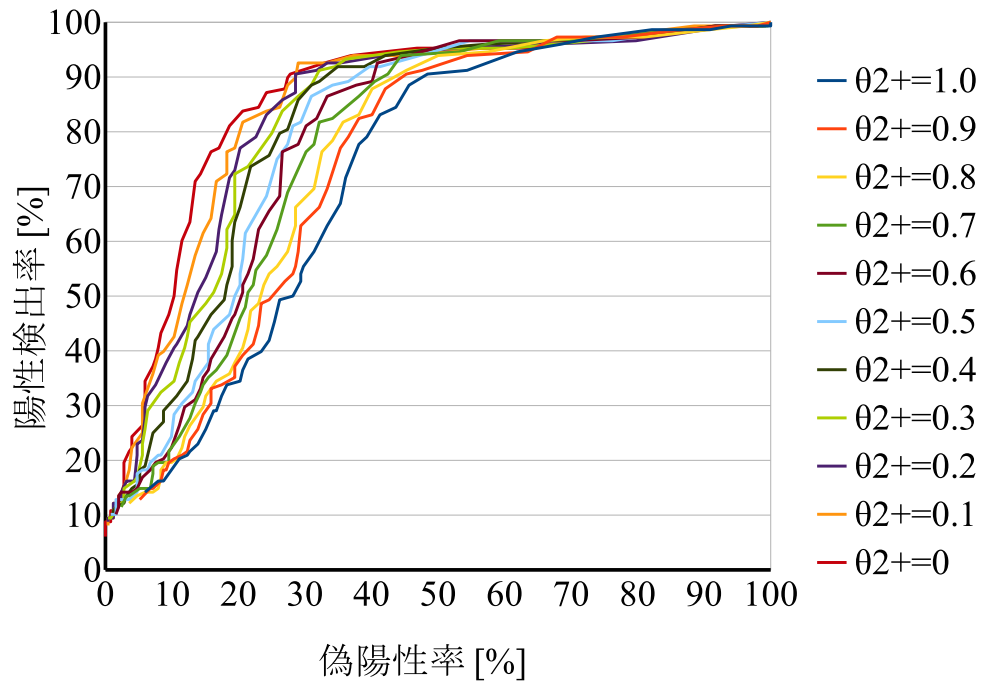


図 4.4: θ_{2+} に関する ROC 曲線 (t_θ 変化)

4.3.4 θ_{2-} についての検討

この節では、パラメータ θ_{2-} と判定精度の関係について検討する。図 4.5, 図 4.6 は 11 種の θ_{2-} に対する ROC 曲線群である。各 θ_{2-} において、図 4.5 では θ_{2+} , 図 4.6 では t_θ を 0~5 に変化させた結果が一本の ROC 曲線に対応する。なお、議論を簡単にするため、判定はケース 2 に該当するかどうかのみにより行った。

図 4.5, 図 4.6 では、 θ_{2-} の値が下がると、ROC 曲線全体が下に下がり、陽性検出率が上がりにくくなる。そのため「 θ_{2-} を小さくすると陽性検出率が減少する」といえる。実験結果より、ROC 曲線が最も上となったのは、 $\theta_{2-} = 2.0$ のときであった。

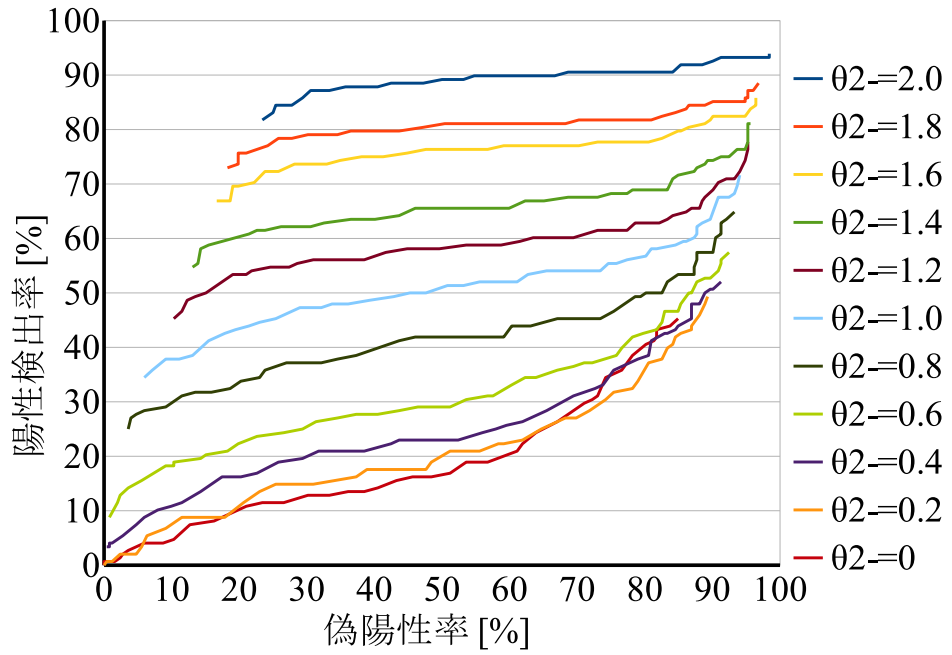


図 4.5: θ_{2-} に関する ROC 曲線 (θ_{2-} 変化)

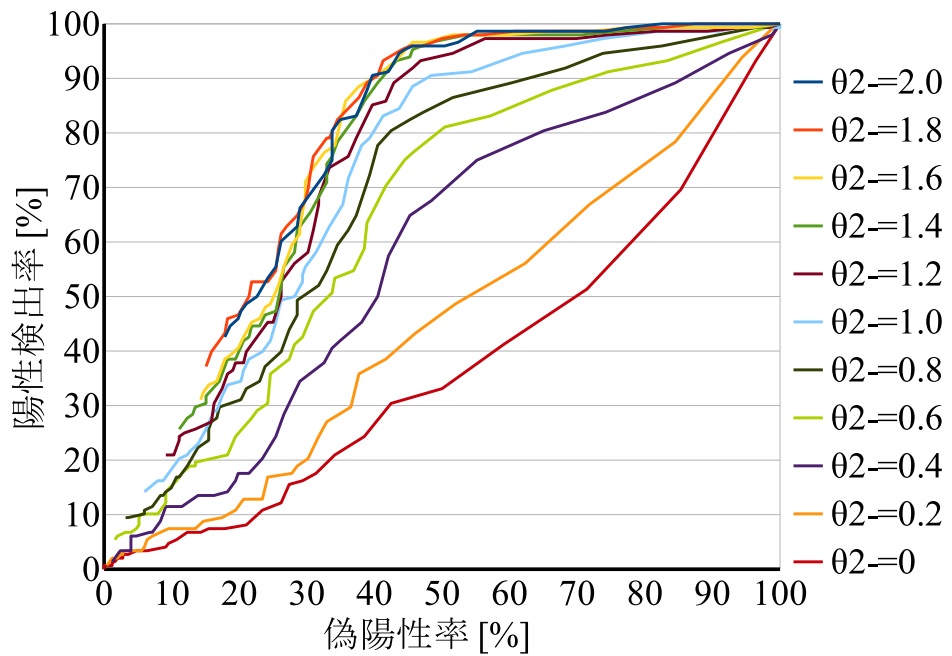


図 4.6: θ_{2-} に関する ROC 曲線 (t_{θ} 変化)

4.3.5 t_θ についての検討

この節では、パラメータ t_θ と判定精度の関係について検討する。図 4.7, 図 4.8 は 11 種の t_θ に対する ROC 曲線群である。各 t_θ において、図 4.7 では θ_{2+} , 図 4.8 では θ_{2-} を 0~5 に変化させた結果が一本の ROC 曲線に対応する。なお、議論を簡単にするため、判定はケース 2 に該当するかどうかのみにより行った。

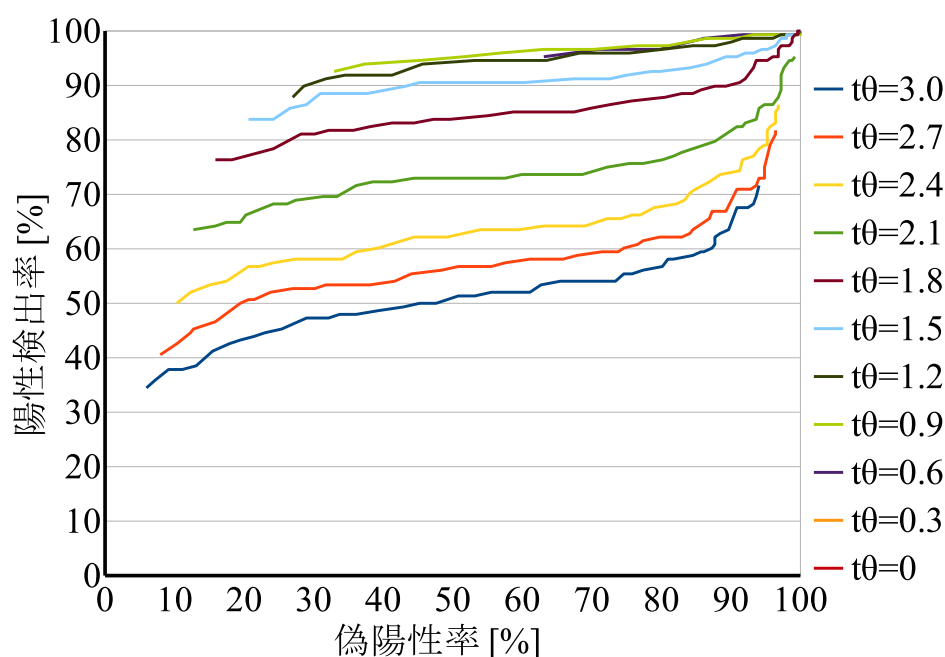


図 4.7: t_θ に関する ROC 曲線 (θ_{2+} 変化)

図 4.7 では、 t_θ の値が下がると ROC 曲線全体は上に向かう。図 4.7 は θ_{2-} を固定した結果が ROC 曲線に対応しているので、「 θ_{2-} を固定したとき t_θ を小さくすると陽性検出率が増加する」といえる。図 4.8 では、 t_θ の値が下がると ROC 曲線全体は右に向かう。図 4.8 は θ_{2+} を固定した結果が ROC 曲線に対応しているので、「 θ_{2+} を固定したとき t_θ を小さくすると偽陽性率が増加する」といえる。しかし 3 つのパラメータが同時に動いたときに関しては今後の課題として検討が必要である。なお、実験結果より、 θ_{2-} を固定したとき ROC 曲線が最も上となったのは、 $t_\theta = 0$, θ_{2+} を固定したとき ROC 曲線が最も左となったのは、 $t_\theta = 3.0$ のときであった。

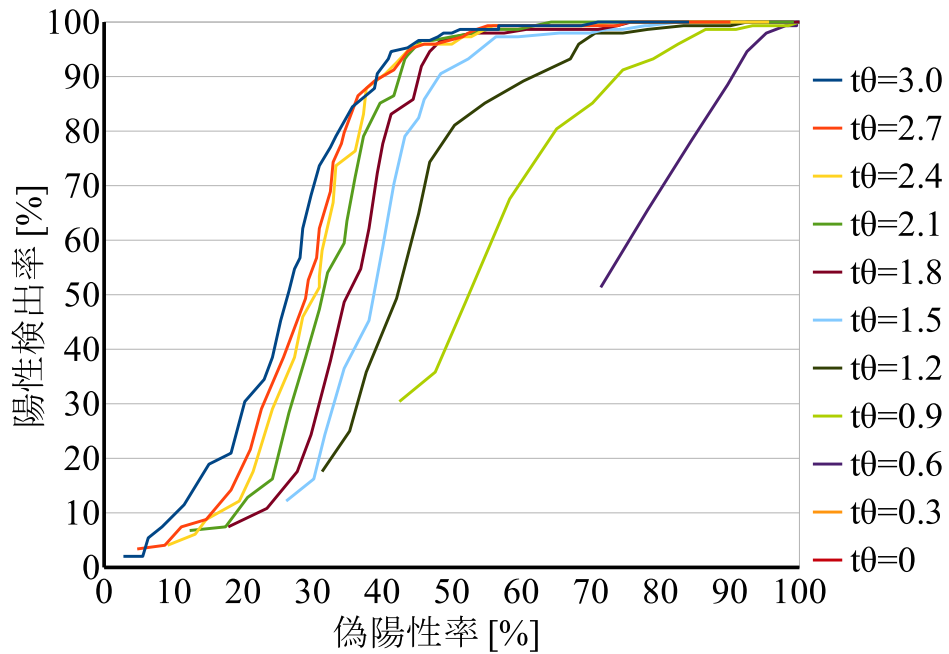


図 4.8: t_{θ} に関する ROC 曲線 (θ_{2-} 変化)

4.3.6 判定精度

本節では、前節までの実験により得られた適切なパラメータ値 ($\theta_{1+} = 0$, $\theta_{1-} = 1.5$, $\theta_{2+} = 0$, $\theta_{2-} = 1.3$, $t_{\theta} = 3.0$) を用いた場合の判定精度を評価する。表 4.3 に前節までと同じく 400 の入力について、それぞれ判定を行った結果を示す。判定精度の検討に関しては、ケース 1, ケース 2 両方を使用して調べた。陽性検出率, 陰性検出率, 偽陽性率, 偽陰性率は下の表 4.4 のようになった。表 4.4 より偽陽性率は 20% 未満, 偽陰性率は 10% 未満で 3.1 節の要求 (B) について満たしていると考えられる。また提案法の計算量は、判定そのものは内部電位の値に基づいて行い、追加で計算した内部電位はないため要求 (A) について満たしていると考えられる。したがってこの判定法は有効であるといえる。

表 4.3: パターンに対して提案法を使用したときの四分表

四分表		提案法	
		問題あり	問題なし
教師入力	不連続	135	13
	連続	47	205

表 4.4: 判定法の陽性検出率, 陰性検出率

陽性検出率	91
陰性検出率	81
偽陽性率	19
偽陰性率	9

[%]

4.4 興味深い事柄

そのほかにも, 内部電位を調査することでいくつかの興味深い結果を得た.

第1は, 中間層でも出力層でも極大値が存在しない場合は, 連続な入出力関係となることである. これは, 問題がない入力の特徴として考えられるので判定精度が向上できる.

第2は, 中間層の極大値は時定数 τ の少し後の時刻にできることが多々あることである. これは極大値が存在しやすい場所がわかることで問題がある入力への対策に役立つのではと考えられる.

しかしこれらの仮説は十分に検証されていないため, さらなる調査が必要である.

第5章 まとめと今後の課題

SpikeProp ネットワークには不連続な入出力関係によって、汎化能力が低下する可能性があるという問題がある。これまでに、Weight Decay を用いて不連続な入出力関係が生じにくいネットワークを構築する方法が提案されてきた。しかし、いまだ十分にこの問題を解消できていない。従来法では、ネットワークの構造が不連続な入出力関係を引き起こしにくくなるよう学習点（教師パターン）全体に対策していた。これに対して、本研究では個々の学習点ごとに対策することで、より効果を高めることができると考えた。このアプローチでは、対策すべきパターンの判定と、それに対する対策の2点について検討する必要がある。本研究では、対策すべき（問題がある）入力を判定する方法について検討した。

ここで、この判定には、(A) 追加の計算量が少ないこと、(B) 偽陰性（検出漏れ）が極力少ないことの2つが要求される。そこでユニットの内部電位の時間変化により判定する。具体的には(1)(2)のいずれかを示すユニットが存在する入力に対して、その入力を「問題がある入力」と判定する。

- (1) 内部電位が、しきい値に近い極大値を持つ場合
- (2) 内部電位が、しきい値と近い値を長時間とる場合

簡単な実験により提案法の判定の有効性を確かめたところ、偽陽性率は20%未満、偽陰性率は10%未満であり、追加の計算量も少なく行えるので要求(A)(B)を満たすことができた。そのため、この提案法は有効であるといえる。

今後の課題としては、対策すべき入力とそうでない入力を80%以上判定できるため、不連続な入出力関係のさらなる軽減を見込めるが、図3.1の(1)の部分に関して、さらなる精度の向上を図る必要がある。また、図3.1の(2)の部分に関して、検出法を用いて発見した入力に対し、具体的にどのような対策をするのかを検討しなくてはならない。その方法と本論文の提案法を組み合わせることで、不連続な入出力関係を軽減できる。

謝辞

本論文は、著者が三重大学大学院工学研究科電気電子工学専攻の情報処理研究室にて行った研究をまとめたもので、その際には多くのご支援をいただきました。本論文を進めるにあたり、丁寧なご指導を賜った三重大学の鶴岡教授、高瀬准教授、川中助教、北准教授、および博士、修士の先輩方に感謝いたします。特に高瀬准教授をはじめとする高瀬グループの皆様方には、3年間を通して研究のアイデアや指針の提案、発表方法や発表スライドに対する改善、資料や論文に対する修正など、さまざまな面でご指摘、ご指導いただきまして、本当にありがとうございました。この3年間で多くのことを学ぶことができました。また、日頃熱心に討論していただいた情報処理研究室の皆様方に厚く御礼申し上げます。

最後に、本論文をまとめるにあたり、助言、討論、その他お世話になったすべての方々に感謝いたします。

参考文献

- [1] 谷口忠大, 岩橋直人: 複数予測モデル遷移の N-gram 統計に基づく非分節運動系列からの模倣学習手法, 日本知能情報ファジィ学会ファジィシステムシンポジウム講演論文集, G-R03, 2009
- [2] 田淵一真, 谷口忠大, 榎木哲夫: 模倣学習と強化学習の調和による効率的行動獲得, 人工知能学会全国大会論文集, 3C1-2, 2006
- [3] 地村弘二, 松岡隆史, 駒崎久明, 中川正宣, 楠見孝: カオスニューラルネットワークによる漢字記憶検索プロセスのダイナミカルなモデル化, *Cognitive Studies*, Vol.6, No.1, pp.44–54, 1999
- [4] 西村治彦, 合原一幸: ニューラルネットワークとカオスの共働, 計測と制御, Vol.39, No.3, 2000
- [5] 柴田克成, 岡部洋一, 伊藤宏司: ニューラルネットワークを用いた Direct-Vision-Based 強化学習 -センサからモータまで-, 計測自動制御学会論文集, Vol.37, No.2, pp.168–177, 2001
- [6] 市村直幸, 竹内俱佳, 永井豊: パターン認識のためのクラスタ分析に基づく 3 層ニューラルネットワークの重み初期値の一設定法, 電子情報通信学会論文誌 D-2, Vol.J77-D-2, No.2, pp.301–310, 1994
- [7] F. Rosenblatt: The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review*, Vol.65, No.6, pp.386–408, 1958
- [8] M. L. Minsky, S. A. Papert: An Introducton to Computational Geometry, The MIT Press, 1969

- [9] David E. Rumelhart, James McClelland: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, MA: MIT Press, 1986
- [10] W. Maass, C. M. Bishop: *Pulsed Neural Networks*, The MIT Press, 1998
- [11] 廣瀬哲也, 浅井哲也, 雨宮好仁: 磁束量子回路によるスパイクニューロン回路とその応用, 電子情報通信学会技術研究報告. SDM, シリコン材料・デバイス, Vol.106, No.521, pp.41–45, 2007
- [12] 木村一郎, 黒江康明, 小寺広倫, 村田智哉: スパイクングニューラルネットワークを用いた視覚系の特徴抽出モデルの構築, 計測自動制御学会論文集, Vol.45, No.11, pp.612–619, 2009
- [13] Gray D. Johnston, Wu S. Miao-Sin: *Foundations of cellular neurophysiology*(3rd.ed), The MIT Press, 1997
- [14] Anthony N. Burkitt, Hamish Meffin, David B. Grayden: Spike-Timing-Dependent-Plasticity: The Relationship to Rate-Based Learning for Models with Weight Dynamics Determined by a Stable Fixed Point, *Neural Computation* 16, pp.885–940, 2004
- [15] Andrzej KASINSKI and Filip PONULAK: Comparison of Supervised Learning Methods for Spike Time Coding in Spiking Neural Networks, *International Journal of Applied Mathematics and Computer Science*, Vol.16, No. 1, pp.101–113, 2006
- [16] 黒江康明: スパイクングニューラルネットワーク —学習法の中心として, システム／情報／制御, Vol.48, No.2, pp.57–62, 2004
- [17] S. M. Bohte, J. N. Kok, and J. A. La Poutre: Error-backpropagation in temporally encoded networks of spiking neurons, *Neurocomputing* Vol.48, pp.17–37, 2002
- [18] S. J. Thorpe, J. Gautrais: Rapid visual processing using spike asynchrony, M. C. Mozer, M. I. Jordan, T. Petsche(Eds.), *Advances in Neural Information Processing Systems*, Vol.9, Cambridge: MIT Press. MA, p.901, 1997

- [19] Sander M. Bohte: The evidence for neural information processing with precise spike-times, A survey, *Natural Computing*, Vol. 3, pp. 195–206, 2004
- [20] W. Maass: Noisy spiking neurons with temporal coding have more computational power than sigmoidal neurons, in : M.C. Mozer, M.I. Jordan, T. Petsche (Eds.), *Advances in Neural Information Processing Systems*, Vol.9, The MIT Press, Cambridge, MA, p. 211, 1997
- [21] 李彦: SpikeProp ネットワークの不連続な入出力関係を改善するための Weight Decay 法に関する研究, 三重大学大学院工学研究科電気電子工学専攻修士論文, 2012
- [22] David C. Plaut, Steven J. Nowlan, Geoffrey E. Hinton: Experiments on Learning by Back Propagation, Technical Report CMU-CS-86-126, Carnegie-Mellon University, 1986
- [23] Yuya Yoshikawa, Haruhiko Takase, Hiroharu Kawanaka, Shinji Tsuruoka: Effectiveness of Weight Decay to improve discontinuous output in SpikeProp, *International workshop on regional innovation studies*, pp.65–68, 2011
- [24] Anders Krogh, John A. Hertz: A Simple Weight Decay Can Improve Generalization, in: J.E. Moody, S.J. Hanson and R.P. Lippmann (Eds.), *Advances in Neural Information Processing Systems*, Vol. 4, Morgan Kauffmann Publishers, San Mateo CA, pp. 950–957, 1992

発表

- [A1] 北川陽一郎, 高瀬治彦, 川中普晴, 鶴岡信治: 内部状態に着目した SpikeProp の出力の不連続な変化の改善に関する一考察, 電気関係学会東海支部連合大会講演論文集, I3-4, 2013
- [A2] 北川陽一郎, 高瀬治彦, 川中普晴, 鶴岡信治: SpikeProp の不連続な出力変化をもたらす入力判定法, 電気・電子・情報関係学会東海支部連合大会講演論文集, J5-2, 2014
- [A3] Yoichiro Kitagawa, Haruhiko Takase, Hiroharu Kawanaka, Shinji Tsuruoka: Improvement for Input-Output Characteristic of SpikeProp Networks —Input Patterns that Cause Discontinuous Output Change—, Proceedings of the Sixth International Workshop on Regional Innovation Studies, pp.63–66, 2014[Outstanding Paper Award 受賞]