

実数型 PAN(Periodic Alternate Neuron)モデル を用いたニューラルネットワーク

Neural Network Constructed with Real Type PAN Model

○陳 玉民, 下河原 力, 吉田 等明, 恒川 佳隆
三浦 守 (岩手大学)

Yumin Chen, Riki Simokawara, Hitoaki Yoshida, Yoshitaka Tsunekawa
and Mamoru Miura (Iwate University)

キーワード : ニューラルネットワーク(neural network), ニューロンモデル(Neuron model), 学習(learning),
シグモイド関数(sigmoid function), 非線形分離問題(nonlinearly separable problem)

連絡先 : 〒020 岩手県盛岡市上田 4-3-5 岩手大学工学部 情報工学科
陳 玉民, E-mail: yumin@cis.iwate-u.ac.jp

1 まえがき

近年ニューラルネットワークの持つ学習性や並列性及び一般性等から応用研究が盛んに行われている。具体的応用に当たっては、学習回数の多さやネットワークの肥大化などの問題があり、これらは長年の課題とされていた。

このような課題を解決する一手法として、Martinelli 等は剰余数系に基づくニューロンモデル(RNS ニューロンモデル)を提案した[1]。これは、剰余数の周期性をニューロンの出力関数に適用し、非線形分離問題への対応やネットワークの単純化を図ったものである。また、この

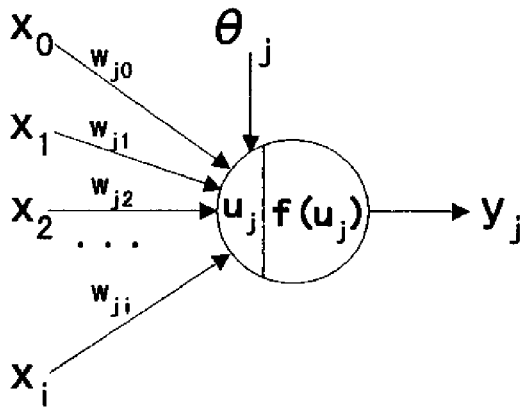
モデルは一般的なニューロンモデルと比較し、ニューロン数、学習速度及び学習の収束率等の性能を向上されることが可能となるが、その反面学習過程や出力を得る過程が複雑化するという欠点も有している[2]。

著者らは、このような RNS ニューロンモデルの欠点を改善するため出力関数を実数型の周期関数に拡張した PAN(Periodic Alternate Neuron)モデルを先に提案した[3][4]。本稿では、文字パターンの学習を通して、実数型 PAN モデルに基づくニューラルネットワーク(RPNN)と一般的なシグモイド関数ニューロンモデルに基づくニューラルネットワーク(SNN)の性能を比較・検討したので報告する。

2 シグモイド関数ニューロン (SN)モデルと PAN モデル

2.1 シグモイド関数ニューロン (SN)モデル

一般的に用いられているニューロンモデルを図1に示す。図2に示すようなシグモイド



- x_i : ニューロン i からの入力値
- w_{ji} : ニューロン i からニューロン j への結合荷重
- θ_j : ニューロン j の閾値
- u_j : ニューロン j の内部状態
- y_j : ニューロン j の出力
- $f()$: ニューロン j の出力関数

図1 一般的なニューロンモデル

関数を用いる場合を、ここではシグモイド関数ニューロン (SN) と呼ぶことにする。SN モデルにおけるニューロン j の内部状態 u_j 及び出力 y_j はそれぞれ(1)式及び(2)式で与えられる。

$$u_j = \sum x_i w_{ji} + \theta_j \quad (1)$$

$$y_j = f(u_j) = \frac{1}{1 + \exp(-u_j)} \quad (2)$$

SN を用いて入力パターンの分類などを行う場合、図2のシグモイド関数においてニューロンの内部状態が0の位置、すなわち出力が0.5

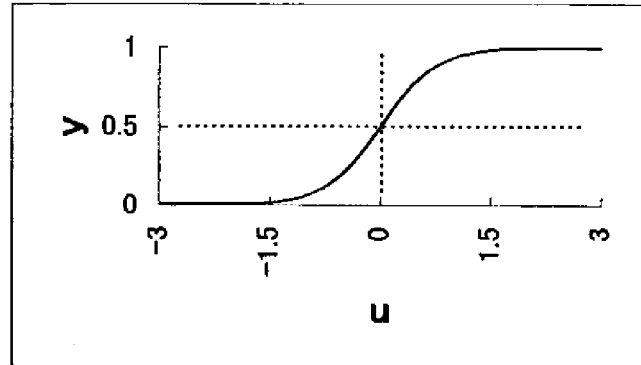


図2 シグモイド関数

の位置を分類の基準に設定する。今、ニューロンの入力数を n 個とすれば、ニューロンの内部状態が0となる入力値 x_i の集合は n 次元空間における $(n-1)$ 次元の超平面となり、これは入力パターンを分類する境界超平面を表わすことになる。超平面のある側を1領域、反対側を0領域とよぶ。図3は、入力数が2の場合の境界超平面(境界線)を示している。この場合の

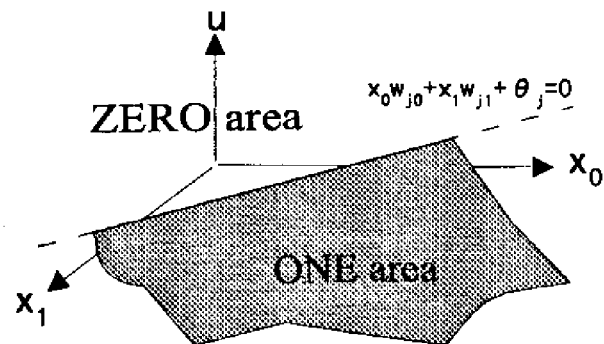


図3 一個のSNの識別領域

境界線の方程式は、次式のようにになる。

$$x_0 w_{j0} + x_1 w_{j1} + \theta_j = 0 \quad (3)$$

SN単体の学習は、例えば(3)式を満足するように結合荷重 w と閾値 θ を調整することによ

って行われる。従って、SN 単体の学習能力は、図 4 に示すような線形な問題(線形分離問題)に限定される。

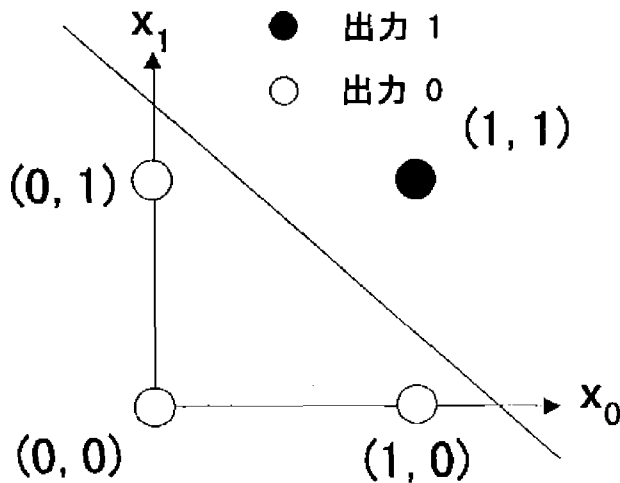


図 4 SN 単体の学習能力

2.2 XOR 問題

XOR 論理関係を認識する問題はニューラルネットワーク研究の中で典型的な問題である。その入力パターンの集合を図 5 に示す。その中で、●は出力 1, ○は出力 0 である。これから分かるように、一本の直線で●の入力パターンの集合 $\{(0, 1), (1, 0)\}$, と○の入力パターンの集合 $\{(0, 0), (1, 1)\}$ を分離することは不可能である。

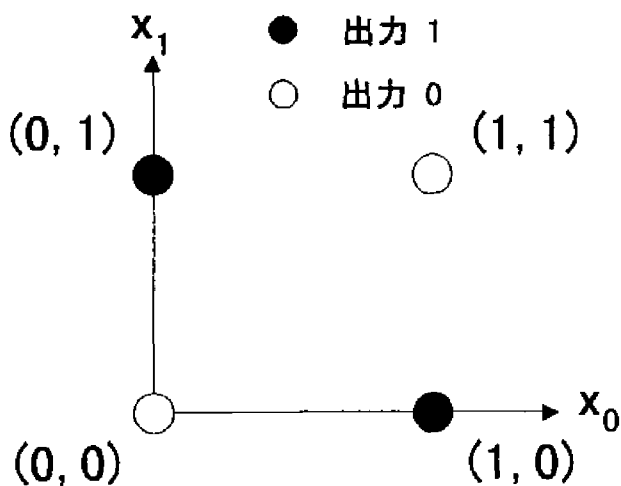


図 5 XOR 問題

すなわち、XOR 問題は非線形分離問題であり、単体の SN では XOR 問題の入力パターンを認識することが出来ない。この XOR 問題を解決するために、SN を最低三つ使って図 6 のようなニューラルネットワーク (SNN) を構成することが必要となる。

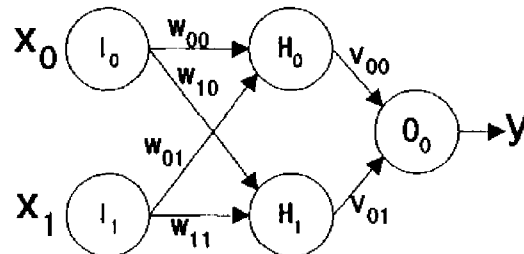


図 6 XOR 問題を解くための最小な SNN

2.3 実数型 PAN モデル

PAN モデルは、出力関数が周期的に変化するニューロンモデルの総称であり、出力関数として実数型の周期関数を用いる場合を実数型 PAN モデルと呼ぶ。ここでは、図 7 に示す正弦波関数を用いる場合について考える。これは、図 1 の一般的なニューロンモデルと同様に表現することが出来るが、この時のニューロン j の内部状態 u_j 及び出力 y_j は、それぞれ(4)式及び(5)式のように与えられる。

$$u_j = \frac{2\pi (\sum x_i w_{ji} + \theta_j)}{T_j} + 2\pi\theta_j \quad (4)$$

T_j : 正弦波出力関数の周期

$$y_j = f(u_j) = \frac{(\sin(u_j) + 1)}{2} \quad (5)$$

実数型 PAN 単体の識別領域は図 8 に示すように 0 領域と 1 領域が交互に存在する。学習によって、その識別領域の方向や広さを調整することが出来るので、前節で説明した XOR 問題

の入力パターンは一つの実数型 PAN で認識することが可能となる。

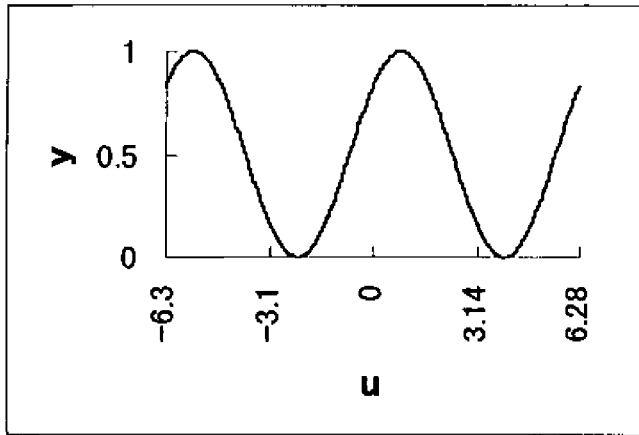


図7 PAN 出力関数

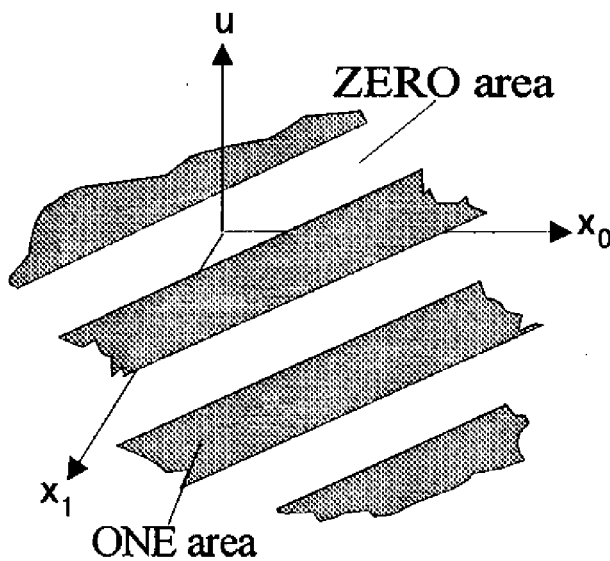


図8 一個の PAN の識別領域

2.4 XOR 論理問題の学習結果

シグモイド関数モデルに基づくネットワーク SNN と実数型 PAN モデルに基づくネットワーク RPNN で、XOR 問題を学習した結果を表 1 に示す。

表 1 SNN と RPNN で XOR 問題を学習する結果

	SNN	RPNN
最小セル数	3	1
平均学習回数	87.3	16.9

3 手書き文字の学習実験

PAN モデルと SN モデルの性能を比較するために、RPNN と SNN を構成して、両方のネットワークで手書き数字と手書き漢字を学習した。実験結果について以下に述べる。

3.1 ネットワークの構成

ネットワークの構成は、いずれも図 9 で示しているような三層ネットワークにしている。SNN では、隠れ層と出力層はすべて SN で構成しているのに対して、RPNN の隠れ層は実数型 PAN を使用し、出力層は図 10 に示す準シグモイド関数ニューロンモデルで構成している。学習アルゴリズムはいずれもバックプロパゲーション (BP) 法を使用し、ネットワークの入力層のセル数は手書き文字の点数と同じで、出力層セル数は学習する文字数と同じにしている。

学習定数等は予備実験を行い、それぞれのネットワークにおいて最も適した値を選んだ。RPNN の隠れ層における実数型 PAN モデルの周期 T_j は、次式のように設定した(理由は後述する)。

$$T_j = \frac{3\pi(j+1)}{10} \quad (6)$$

このように構成したニューラルネットワークを用いて、手書き数字と手書き漢字を学習し、隠れ層のセル数を変えて実験を行った。

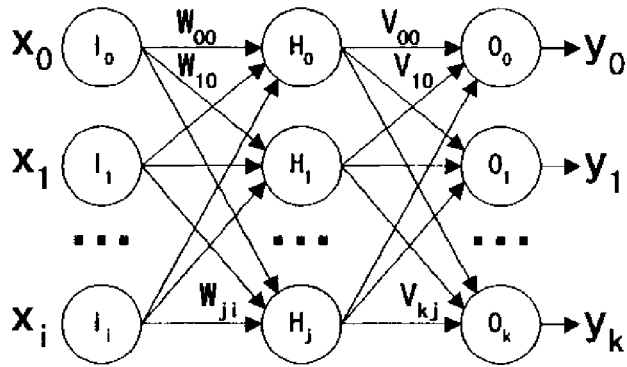
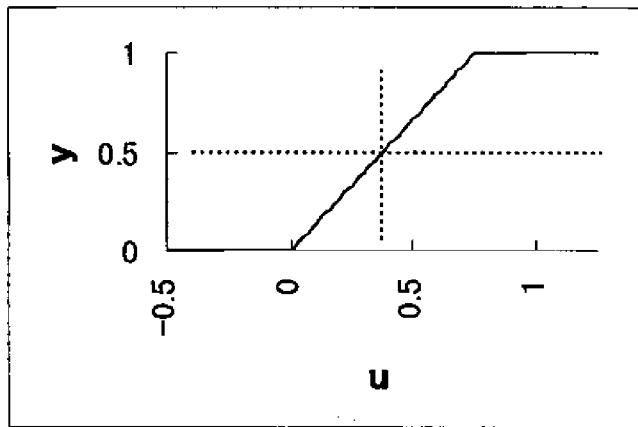


図9 ネットワークの構成



$$y = u / (u + 0.75) \quad (0 \leq u \leq u_0)$$

$$y = 1 \quad (u > u_0)$$

$$y = 0 \quad (u < 0)$$

$$u_0 = 0.75$$

図10 準シグモイド出力関数

3.2 手書き数字の学習

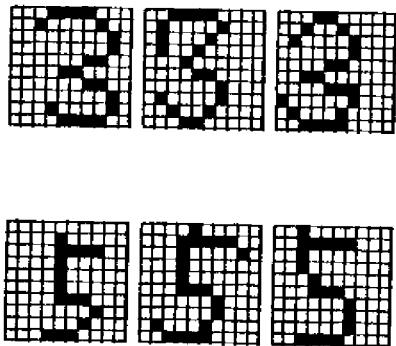


図11 手書き数字学習パターン

3.2.1 手書き数字学習パターン

数字1~5までの2値手書きパターンを、それぞれ10個ずつ、全部で50個用意して、学習パターンとして使用した。図11には数字の3と5の場合の入力パターン例を示す。

3.2.2 学習

隠れ層のセル数を2~10個のそれぞれの場合で比較した。学習の終了条件は学習が300回以内で、誤差自乗和は0.05以下になったときとした。

3.2.3 学習結果

学習結果を表2に示す。

表2 平均学習回数と収束率の比較

隠れ層セル数	RPNN		SNN	
	平均学習回数	収束率 (%)	平均学習回数	収束率 (%)
2	-	0.0	168	5.00
3	102	53.3	67.1	71.6
4	17.7	100	60.7	81.7
5	14.3	100	53.9	88.7
6	13.2	100	45.7	92.7
7	13.1	100	38.4	93.7
8	12.7	100	34.9	97.7
9	12.6	100	34.8	95.3
10	10.8	100	31.2	97.7

これからわかるように、隠れ層内のニューロン数(セル数)が極端に少ない時、RPNNはSNNより性能が劣っている。しかし、セル数が4

個以上では学習率，収束率とも RPNN の方が性能は上回っている。

3.3 手書き漢字の学習

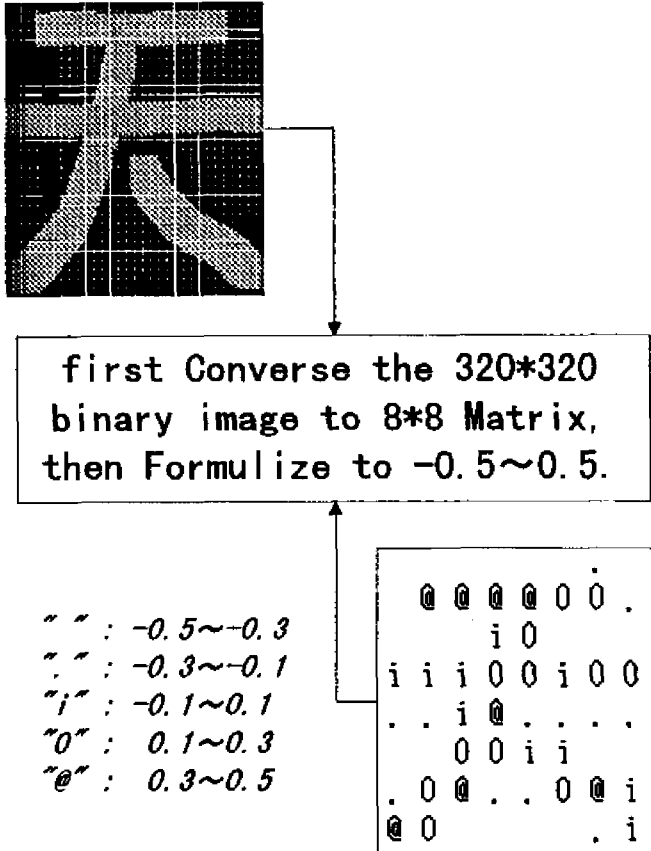


図 12 学習パターンの作成

表 3 学習パターンとして使っている漢字

上(5)	中(5)	下(5)	天(5)	地(5)	大(3)	国(3)
日(3)	本(3)	月(3)	三(4)	四(3)	五(3)	六(2)
品(3)	向(6)	省(3)	力(3)	化(3)	全(4)	合(3)
介(3)	人(5)	入(5)	浦(2)	子(3)	広(4)	学(3)
安(2)	多(3)	用(3)	実(3)	北(3)	由(3)	甲(2)
申(3)	田(3)	山(3)	東(3)	南(3)	西(3)	年(3)
示(3)	出(3)	一(3)	二(3)	七(3)	八(3)	九(3)
十(3)	廿(3)	自(3)	白(3)	目(3)	生(3)	万(1)
台(3)	央(3)	会(3)	先(1)	吉(3)	方(3)	同(3)
回(3)	円(3)	加(3)	号(3)	石(3)	足(3)	了(1)

注：()中の数字はこの漢字に対応するパターン数
漢字数：70，学習パターン数：220。

3.3.1 手書き漢字学習パターン

図 12 で示しているように，コンピュータ画面から入力した 320*320 の 2 値手書き文字を 8*8 のマトリックスに変換して，各点の値を -0.5 から 0.5 までの間に規格化する．このようなパターンを 70 文字に対して，全部で 220 個用意して，学習パターンとして使った(表 3)。

3.3.2 学習

収束条件としてネットワークが全ての学習パターンを認識できる場合とした．学習する漢字数を，それぞれ 24,51,62,70 にして RPNN と SNN で学習し，比較，検討した。

表 4 手書き漢字を学習する結果の比較

学習パターン数	カテゴリ数	隠れ層セル数	平均学習回数	
			RPNN	SNN
220	70	16	912	-
195	62	16	779	-
166	51	16	510	870
		12	1413	496
		10	1736	2261
		8	7129	-
87	24	16	429	208
		14	430	284
		12	343	168
		10	415	522
		8	700	-

注 1) 平均学習回数：収束した場合の学習回数の平均

2) - : 1 万回の学習で収束しなかった場合。

3.3.3 学習結果

学習結果は、表 4 に示す通りである。

表 4 から、隠れ層のセル数が極端に少ないときは SNN では収束しなくなるが、RPNN ではまだ収束できることが分かった。

3.4 実数型 PAN に関する考察

3.4.1 各モデルの周期値の設定

RPNN の隠れ層中の各ニューロンにおける周期は、固定するか、乱数で設定する方法なども考えられるが、ここでは、(6)式で示したように各ニューロンの周期に若干のずれを与える方法を採用した。これによって収束は速くなることが分かった。

3.4.2 修正量

結合荷重と閾値を一回で多く修正すると、収束しなくなる場合があるので、BP 法の中の学習パラメータを適当に設定する必要がある。このパラメータの設定はセル数と関係があり、セル数が多くなるに従って、学習パラメータを小さく設定する方が良い結果が得られた。

4 結び

本研究では、PAN モデルを実数型へ拡張し、そのネットワーク (RPNN) の構成法および学習アルゴリズムについて検討した。ソフトウェアシミュレーションによって、一般的なニューラルネットワーク (SNN) と比較し、性能評価を行った。実数型 PAN モデルの特徴を挙げると、次のようになる。

1) 一般的なニューロンモデルより、少ない隠

れ層ニューロンで高い性能を得ることができるので、ネットワークの縮小を図ることができる。

- 2) 出力関数に周期を持たせることで、PAN 単体でも非線形分離問題に対応可能となる。
- 3) 実数型 PAN モデルに基づくネットワーク (RPNN) は、一般的なネットワーク SNN に比較し、学習能力が優れていることが分かった

参考文献

- [1] Martinelli, G. and Perfetti, R., "RNS NEURAL NETWORKS", Proc. of ISCAS, PP2955~2958, (1990)
- [2] 瓜田, 三浦, 吉田, 千葉, "剰余数系に基づくニューラルネットワーク", 計測自動制御学会東北支部研究集会資料, 140-10, pp.1~9 (1993)
- [3] 三浦, 瓜田, 吉田, 千葉, "周期的出力関数を有するニューロンモデル", 電気関係学会東北支部連合大会論文集, 1D9 (1993)
- [4] 瓜田, 吉田, 三浦, 千葉 "周期的連続値出力関数を有するニューロンモデルとその応用", 信学技報, CAS 93-122, pp53~58 (1994)
- [5] 甘利 俊一(編), "ニューラルネットワークの新展開", サイエンス社 (1993)
- [6] 瓜田 健司, "PAN モデルとニューラルネットワークへの応用", 岩手大学院修士論文 (1995)