

カオス複素ニューラルネットワークを用いた クラス分類法の提案

1 はじめに

本論文では、複素ニューロンにカオスを導入したカオス複素ニューロンによるネットワークの連想記憶能力を用いたクラス分類問題へのアプローチの可能性を示す。ここで言うクラス分類問題とは n 個の項目 q_1, q_2, \dots, q_n のデータ $(q_1^*, q_2^*, \dots, q_n^*)$ が与えられたとき、これが m 個のクラス (C_1, C_2, \dots, C_m) のいずれに属するデータかを判定する問題である。連想記憶による分類法としては、各クラスを記憶パターンとしてネットワークに設定し、初期パターンがどの記憶パターンを連想したかで分類を行うというものである。従来研究では、アソシエトロンや CNN (セルラ・ニューラルネットワーク) などのニューラルネットワークが用いられてきたが、これらは 3 値パターンを扱うのが限界であり、高い精度の分類ができなかった。そこで本研究では、ニューロンの状態数を複素数で記述し、これを 5 段階の出力値に対応付けし、さらに、柔軟な解の探索を可能にするカオスを導入したネットワークを構築する方法を提案する。このネットワークにより、クラス分類システムを設計し、肝臓病の診断問題と小麦種分類問題に適用し、有用性を検証する。

2 カオス複素ニューラルネットワーク

ニューロンの入出力、内部状態、重みに複素数が用いられており、連想記憶において、多値パターンを扱うことができる。また、カオスにより適切なパラメータを設定することでネットワークの局所解への収束を脱し、複数の記憶パターンの動的な想起が可能であることが知られている。

ニューラルネットワークの詳細は文献 [1] を参照されたい。

3 クラス分類問題への適用

3.1 クラス分類システムの構成

扱うパターンは 5 段階のレベルで表現され、それぞれの状態を単位円弧上に等角度に割り当てる。分類までのステップは、

1. ニューロン行列、項目の配置。
2. 各項目に対する変換関数の作成。
3. 各クラスを表す記憶パターンの作成。
4. 入力データより初期パターンを生成し、文献 [1] によるカオス複素ニューラルネットワークにより連想。
5. 完全想起すればその想起結果を分類結果とする。不完全想起の場合、ハミング距離の最も短くなるパターンのクラスを分類結果とする。

である。

3.2 肝臓病の診断問題

5 種類のクラス (C1:健康, C2:急性肝炎, C3:肝癌, C4:慢性肝炎, C5:肝硬変) の 15 項目の血液検査データがあり、これによる診断結果を検証する肝臓病問題は、従来からこの種のクラス分類法の検証に用いられてきた経緯があり、本研究でもこのデータを用いて検証を行う。

3.2.1 クラス分類システムの設計

扱うパターンの状態数は、“とても低い”、“やや低い”、“普通 (不明)”、“やや高い”、“とても高い”の 5 とし、それぞれの状態を単位円弧上の点 $\exp(0), \exp(i\frac{2}{5}\pi), \exp(i\frac{4}{5}\pi), \exp(i\frac{6}{5}\pi), \exp(i\frac{8}{5}\pi)$ (i : 虚数単位) に対応付ける。ニューロン行列は 4×5 のサイズを用い、15 個の血液検査の各項目のデータの影響の大きさを考慮し、図 1 に示す項目の配置とする。

BUN	γ GTP	AFP	Alb	Fibrino
UrA	Alp	AFP	ChE	Plt
TBil	GPT	II	GPT	DBil
LDH	GOT	GOT GPT	GOT GPT	Fibrino

図 1: 項目の配置図

次に、与えられた血液データが単位円弧上のどの点に対応するかを表す変換関数が必要である。これには文献 [2] 内で挙げられている検査値の区分表より、階段関数を作成した。その関数に各肝臓病患者の血液データの代表値を代入することで 5 種類の記憶パターンを作成した。作成した記憶パターンを図 2 に示す。

ニューロン更新は 10 ステップとし、ニューラルネットワークの各パラメータは、ランダムで選んだ各クラス 50 ずつのデータをトレーニングデータとすることで設定する。実験には各患者 50 人、計 250 人分の血液検査データに対して分類を行う。

3.2.2 肝臓病診断問題実験結果

表 1 に実験結果を示す。行は分類する 50 ずつのデータ、列は分類された結果を示している。例えば C4 の行、C5 の列の 2 とは慢性肝炎患者 50 人の内、2 人が肝硬変に誤診断されたことを表している。ここで、DSE とは正しく診断された割合 (診断率) を表している。

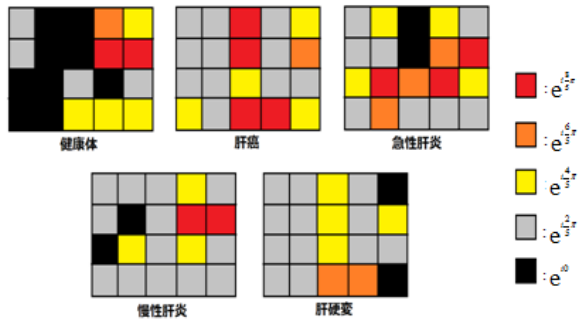


図 2: 記憶パターン

表 1: 提案法における診断結果

クラス	No.	C1	C2	C3	C4	C5	DSE(%)
C1	50	50	0	0	0	0	100
C2	50	2	39	0	0	9	78
C3	50	0	0	45	4	1	90
C4	50	8	0	1	39	2	78
C5	50	5	1	1	6	37	74

表 2 に従来手法による診断結果を示す．表より，比較的良好な診断結果が得られたことがわかる．

3.3 小麦種分類問題

3 種類の小麦 (Kama, Rosa, Canadian) があり，7 つの項目 (q_1 :面積, q_2 :周囲長, q_3 :密集度, q_4 :カーネル長, q_5 :カーネル幅, q_6 :非対称係数, q_7 :カーネル溝の長さ) から与えられるデータにより，いずれの種に属するのかを分類する問題．小麦の分類問題は非常に新しい問題であり，比較対象はあまり多くないが，データ数が豊富であること，データに欠落がないことなどからシステムの検証には優秀であると考えられる．

3.3.1 クラス分類システムの設計

扱うパターンの状態数は，肝臓病の診断問題と同じ 5 とし，それぞれの点を複素単位円弧上に割り当てる．

ニューロン行列は， 3×3 のニューロン行列を用い，各項目の影響力の大きさを考慮し，図 3 に示すような項目の配置とした．

表 2: 従来法による診断結果

クラス	DSE(診断率%)		
	アソシエトロン	CNN	ファジィ
健康	100	100	100
肝臓	100	70	50
急性肝炎		100	70
慢性肝炎	60	80	80
肝硬変	60	60	70
平均	80	82	74

変換関数は，各クラスの各項目のデータの代表値からレベル分けを行い，最小二乗法により連続値を扱うものを作成した．その変換関数より，3 つのクラスを表す記憶パターンを作成した．作成した記憶パターンを図 4 に示す．

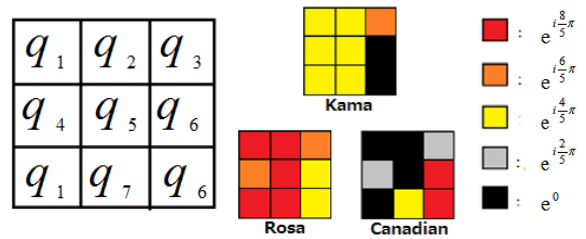


図 3: 各項目の配置 図 4: 記憶パターン

ニューロン更新は 10 ステップとし，ニューラルネットの各パラメータは，ランダムで選んだ各クラス 40 ずつのデータをトレーニングデータとすることで設定する．実験には各小麦 70 ずつ，計 210 の小麦のデータを使用した．

3.3.2 小麦種分類問題実験結果

表 3 に実験結果を示す．行は，各クラス 70 ずつ

表 3: 検証結果

	Kama	Rosa	Canadian	誤分類数	分類率 (%)
Kama	69	0	1	1	98.6
Rosa	7	63	0	7	90.0
Canadian	5	0	65	5	92.9

つの小麦を，列は分類された結果を表している．最新のクラスタリング手法の 1 つである CGCA (Complete Gradient Clustering Algorithm) 法による分類結果は，Kama が 95.7 %，Rosa が 84.2 %，Canadian が 95.7 % であることから，比較して，提案法が遜色ない結果が得られていることが分かる．

4 おわりに

本研究ではカオス複素ニューラルネットワークの連想記憶能力を用いたクラス分類法を提案し，肝臓病の診断問題と，小麦種分類問題の 2 つの問題でクラス分類システムを設計し，検証した結果，優位性を示すことができた．課題点としては，記憶パターン間の直交化，ニューロン行列，変換関数の適切な設定法の確立が挙げられる．

参考文献

- [1] 永椎・金川・滝本：“カオス複素ニューラルネットワークを用いた連想判別法”，日本経営工学会平成 24 年度秋季研究大会予稿集，pp.198-199，2012
- [2] 日本ファジィ学会：ファジィ OR，pp.231，1993