

氏名(本籍)	猿 田 和 樹 (秋田県)
学位の種類	博士(情報科学)
学位記番号	情博第25号
学位授与年月日	平成8年3月26日
学位授与の要件	学位規則第4条第1項該当
研究科, 専攻	東北大学大学院情報科学研究科(博士課程)システム情報科学専攻
学位論文題目	ニューラルネットワークを用いた手書き文字の高精度認識に関する研究
論文審査委員	(主査) 東北大学教授 根元 義章 東北大学教授 牧野 正三 東北大学教授 阿曾 弘具 東北大学助教授 安倍 正人 (工学研究科)

## 論文内容要旨

### 第1章 序 論

情報化社会の発達により、コンピュータによる情報管理・文書電子化の必要性が増大している。情報の保存や処理という観点において、これらの果たす役割は非常に大きい。しかし、多くの場合はキーボードからの入力となるため、利用者の負担は大きく、コンピュータに対する嫌悪感へと発展してしまう可能性もある。よって、より快適な作業環境を提供することは、情報化社会において極めて重要な課題であると言える。文字の入力をキーボードを介することなく行うことができれば、作業環境は快適なものとなり、作業効率の向上は計り知れない。この問題の解決の大きな鍵を握っているのが文字認識技術である。

近年の高精度な認識手法の提案により、手書き日本語文字の認識率は98%を超えており、文字認識は確立された研究分野であるように扱われる場合もある。しかし、記入枠や文字の制限、あるいは低品質文字の認識や確実なリジェクト等残された課題は多く、より高精度で柔軟な認識を行うためには、従来のアルゴリズムである論理的記号の情報処理だけではなく、人間の脳機能のような柔軟性と適応性を有する直感的パターンの情報処理との融合が必要であるとも言われている。

人間の脳に近い情報処理の実現を目的とした研究分野にニューラルネットワークがある。ニューラルネットワークは多数のユニット(ニューロン)の結合により構成される。ユニットは非線形な入出力写像を有しており、高精度な認識システムの実現が期待されている。実際、手書き数字認識等の比較的小規模な問題では、これまでに非常に高精度な認識結果が報告されている。しかし、漢字のように数千もの字種を対象とする問題では、学習の収束・計算量等の問題もあり、ニューラルネットワークを用いた認識方式の報告自体が少なく、大規模問題への適用方法が確立されていない。そのため長年研究されてきたパターン整合法等と比べ、認識率は劣っているのが現状である。

以上の背景により、本論文ではニューラルネットワークを用いた大規模かつ高精度な文字認識アルゴリズムの実現を目的とし、大規模問題の細分類部に適したニューラルネットワークである排他的学習ネットを中心に、いくつかの提案を行った。その結果、国内最大の手書き文字データベース ETL9B (3036字種) に対する認識実験では、ニューラルネットワークとして初めて96%の認識率に到達した。

本論文は全体で6章から構成されている。

## 第2章 係数変化型学習法 (DCLA法)

大規模な認識系を構築する場合は、複数の小規模ネットワークを組み合わせることが有効であるとされている。この小規模ネットワークの学習アルゴリズムとして、最も広く利用されている手法がBP (Back Propagation) 法である。しかし、ローカルミニマムによる学習の停止及び学習に時間がかかる、といった大きな問題点を有している。これらの問題点に対し、高速化に関する報告は多いものの、学習停止の回避に関する報告は少ない。そこで小規模ネットワークの学習性能の向上を目的とし、BP法における学習停止の改善を図った。まず、ネットワークが誤った状態に収束した場合、ユニット(ニューロン)の出力が飽和領域にあるときのユニット間の結合荷重の修正量が小さいために、その状態から抜け出せなくなり、学習停止となることを解析により明らかにした。この解析結果をもとに、BP法において従来固定パラメータとして用いられてきた学習係数をユニットの出力値に応じて変化させ、荷重の修正量を柔軟に制御する、係数変化型学習法(DCLA法)を提案した。DCLA法の提案により、上記の問題点は改善され、学習の収束率はBP法と比較して大きく向上した。また、ニューラルネットワークによる文字認識固有の問題として、学習停止により未収束パターンが残った場合、そのパターンに類似した未学習パターンは正しく認識不可能であること、さらに、認識時に誤りが生じた場合、それが未収束データが残ったことに起因するのか、ニューラルネットワークの性能に起因するのか判断不可能であることが挙げられる。本論文で提案したDCLA法により、これらの問題に対しても改善がなされたと言える。

## 第3章 排他的学習ネットワーク (ELNET)

従来の大規模認識問題へのニューラルネットワークの適用方法は、まず大分類部でクラスタリングによりグループを決定し、細分類部では各グループ(複数字種)を1つのニューラルネットワーク(モジュール)に対応させ、グループ内で認識を行うというものである。しかし、1つのモジュールに対応する字種数が多くなれば、学習・認識の効率は低下する。一方、これまでの生理学・解剖学の成果として、人間の形態識別では、下側頭皮質中における特定のパターンのみ反応する細胞の存在が明らかにされている。これらを考慮し、大規模問題の細分類部に適用するニューラルネットワークとして、多数のモジュールによって構成される排他的学習ネットワーク(ELNET)を提案した。ELNETは従来とは異なり、1モジュールを1字種に対応させ、モジュール毎に発火用データと抑制用データを用いて排他的に学習を行うという特徴を有する。なお、抑制用データには特徴空間上における距離が小さい字種のデータを用いる。

従来認識が難しいとされている手書き類似文字を対象として行った基礎実験では、1モジュールに複数字種に対応させる従来型の構成を上回る認識性能をELNETが有することを明らかにした。さらに、大分類部にパターンマッチング法を用いた認識システムを構築し、ETL9B全字種を対象とした認識実験を行った。その結果、未学習データに対する認識率として94.54%が得られ、大分類認識率と比較して2.12%(6436文字)の向上となり、ELNETの有効性を示すことができた。また、学習後の各モジュールの分析により、抑制用データを増加することでモジュールの抑制効果が強化され、特定の1字種にのみ反応するモジュールを実現できることも示すことができた。以上により、大規模問題の細分類部へのニューラルネットワークの適用を容易なものとした。

## 第4章 候補出現頻度を考慮した排他的学習ネットワーク (ELNET-II)

前章で提案したELNETでは、特徴空間上における字種間の相関のみが考慮され、実際の学習データ間の相関は考慮されていなかったため、学習に使用したデータは有効な識別境界の形成に必ずしも適切であるとは限らない。また、全モジュールを同じデータ数で学習を行ったが、認識に影響を及ぼす類似文字の個数は字種によりばらつきがあると考えられるため、モジュールによっては学習データの不足や剰余も生じている可能性がある。以上の考察に基づき、大分類結果における各字種の上位候補への出現頻度を考慮した排他的学習ネットワーク(ELNET-II)を提案した。ELNET-IIでは、あるデータで大分類を行った場合、その上位候補字種に対応するモジュールにそのデータを抑制用データとして割り当てる方式であるため、データ間の距離分布も考慮され、各字種の境界付近のデータをより多く学習に使用できる。これはニューラルネットワークの認識性能を高めるうえで非常に重要である。さらにこの方式では、類似文字の多い字種、つまり上位候補に出現する頻度の高い字種に対応するモジュールほど多くの抑制用データが与えられるという利点もある。また、これを利用して、各モジュール毎の中間層ユニット数をデータ数に応じて可変にすることができる特徴も有する。ETL9B全字種に対する認識実験では、ELNET-IIの有する各特徴が認識率の向上に寄与し、ELNETより少な

い学習データでも高い認識率が得られることを示した。最高では95.84%の認識率となり、本章で提案した改良方法の有効性を示すことができた。また、さらに提案法の有効性をネットワークの振る舞いの点からも考察するために、各モジュールの内部表現に対し主成分分析による解析を行った。その結果、ELNET-IIではより正確な内部表現が獲得できている、これが改良における認識率向上の要因であることを明らかにした。

## 第5章 荷重逆投影を用いた詳細識別法

ここでは学習終了後のELNET-IIが獲得した特徴抽出能力を利用した詳細識別法を提案し、さらに高精度化を図った。提案法は、まず学習後のモジュールの荷重の積和により荷重逆投影値を求め、文字のどの部分の特徴が各モジュールの出力に強く反映されるかを定直し、各字種の特徴的な部分領域を自動抽出する。これにより、従来の詳細識別法のようにあらかじめ抽出する特徴領域の位置を定める必要がなく、各字種毎に自動決定できるという柔軟性を有する。さらに、抽出した領域に対応した入力層-中間層間の荷重をモジュールの出力傾向に合わせて調整する。これにより、部分領域内への入力が強調され、それがモジュールの出力値に反映される。

類似文字に対して行った予備実験では、文字の特徴が顕著である領域を抽出できることを示すことができた。さらに、これらの認識システムの詳細識別部として採用し、ETL9B全字種に対する認識実験を行った。その結果、詳細識別を行わない場合に比べ約600文字について改善がなされ、96.06%の認識率が得られた。また、細分類部での認識率の低い字種ほど提案法が有効であるという傾向が分析により明らかとなった。以上により、文字の特徴領域の自動抽出法、及び新たな詳細識別法とその有効性を示すことができた。

## 第6章 結 論

本論文では、ニューラルネットワークを用いた大規模かつ高精度な文字認識システムの実現を目的とし、小規模ニューラルネットワークの学習性能の向上、大規模問題に適したニューラルネットワーク、及びその高精度化について研究を行った。

まず、従来広く応用されているBP法の問題点である学習停止の原因を示し、学習の収束性の向上に有効な手法を明らかにした。続いて、従来型の構成を上回る認識能力を有するニューラルネットワークを構築し、大規模問題の細分類部へのニューラルネットワークの適用を容易なものとした。また、このニューラルネットワークの高精度化を図るための手法を提案し、実験により有効性を明らかにすると共に、その振る舞いについても明らかにした。さらに、学習終了後のニューラルネットワークが有する特徴抽出能力を利用し、文字の特徴的な部分の自動抽出と詳細識別の手法を提案し、その有効性を示した。これらの提案法とパターンマッチング法とを組み合わせた認識システムによる認識実験では、従来提案されているニューラルネットワークを上回る、96.06%の認識率が得られることが示された。

以上が本研究の成果である。

## 審 査 結 果 の 要 旨

コンピュータの普及により文書処理の自動化が重視され、文字認識の高精度化が重要視されているが、手書き文字認識は癖字などによるパターンとしての自由度が多く高精度な認識に課題を残している。手書き文字認識の有望な実現手法として、ニューラルネットワークを用いることがあげられる。しかし、学習停止の回避および多くの字種を対象とする認識手法についての研究はまだ十分には行われていない。そこで著者は、学習性能を改善する方法を提案し、手書き文字の大規模認識の高精度化について研究を行った。本論文は、その成果をまとめたもので、全編7章よりなる。

第1章は序論である。

第2章では、ニューラルネットワークにおけるBP法の学習停止問題について検討し、学習停止の要因を明らかにするとともに、これを改善するためユニットの出力値に応じて学習係数を動的に制御する方式を提案している。シミュレーションにより本方式を用いることで学習の収束性が大きく改善されることを示している。これは重要な成果である。

第3章では、認識システムの細分類部のネットワークとして1モジュールを1字種だけに対応させて構成される排他的学習ネット(ELNET)を提案している。これにより類似文字の誤認識が軽減でき、従来より優れた認識率を得ることができることを手書き文字データベース ETL9B を対象とした大規模認識実験により明らかにしている。これは実用上有用であり高く評価できる。

第4章では、パターン空間上での各字種の分布を考慮し、字種の上位候補としての出現頻度に応じて学習データ数、モジュールサイズを可変とする学習アルゴリズムを導出し、ELNETの高性能化を図っている。本アルゴリズムに従えば、ELNETと比較し、少ない学習データ数で高い認識性能が得られることを大規模認識実験により明らかにしている。

第5章では、学習済みのモジュールに対し、学習により得られた特徴抽出能力を利用して、文字の特徴的な領域の抽出を行い、この領域を更に強調して認識を行う荷重逆投影を用いた詳細識別法を導出している。本手法を利用することで認識率がさらに向上することを示している。

第6章は結論である。

以上要するに本論文は、ニューラルネットワークを手書き文字認識に応用するうえで重要な、学習性能を向上させる学習アルゴリズムおよび大規模な認識を実現するネットワーク構成法を与えたもので、システム情報科学の発展に寄与するところが少なくない。

よって、本論文は博士(情報科学)の学位論文として合格と認める。