

ニューラルネットワークを用いた濁点、半濁点の有無による手書き平仮名の分類

大倉 充・塩野 充・橋本 禮治*

岡山理科大学工学部情報工学科

*岡山理科大学工学部電子工学科

(1994年9月30日 受理)

1. まえがき

手書き平仮名文字を認識対象として場合、認識を困難とする要因の一つとして、濁点、半濁点を有する字種（以後、カテゴリーと記す）の存在が挙げられる。例えば、カテゴリー「あ」と「い」を識別することと、「は」と「ば」あるいは「は」と「ぱ」を識別することの間には大きな違いがあると考えられる。前者は構造的に大きな差異があるものの識別であり、後者はある特定領域における特定パターンの有無によって識別が可能となるものである。このため、これら両者が混在したカテゴリー群を単一の認識手法によって認識処理することはかなり困難である。

一般的な手書き文字認識システムにおいては、分類及び識別の段階的な手順が採用されている¹⁾。分類は、入力パターンに対する複数の候補カテゴリーの検出処理であり、その後続く識別処理に大きな影響を与える。しかし、これまでの認識システムにおける分類処理は、カテゴリー数が多いという問題を解消する、すなわち処理時間の短縮という目的のために考えられていたという側面を持つ。そのためこの場合の分類処理では、ある特徴空間において基準となるパターンの周辺に存在するカテゴリー群が集められていた。従ってその後続く識別処理では、カテゴリー数は減少しているものの、上記の問題は本質的には解決されていない。識別に有効に作用する分類とはどのようなものであろうか。一つの考え方として、門構えを持つ漢字群、濁点、半濁点を有する平仮名群のように、分類されたカテゴリー群がある特定の共通した部分構造を有する分類²⁾が挙げられる。この場合、識別処理では共通の部分構造以外の構造の比較になるため、識別率の上昇が期待される。

本論文では、上述した分類の試みの一つとして、濁点、半濁点を有する平仮名群とそれ以外の平仮名群の分類を3層階層型ニューラルネットワークを用いて行った結果を示す。3層階層型ニューラルネットワークによる分類に関する研究例として、梅田による手書き漢字と平仮名の分類が挙げられる³⁾。梅田は、漢字と平仮名を構成するストロークの直線性と曲線性の違いに着目し、それらを表す特徴量として局所所方向寄与度特徴と Hough 変換に基づく特徴量の2つを採用し、ニューラルネットワークへの入力データとした。本論

文では、濁点、半濁点の有無を判別させるという目的のために、最も単純な特徴量である局所領域における粗いメッシュ・パターン¹⁾をニューラルネットワークへの入力データとした。

2. 実験に使用したデータ

実験には、工業院電子技術総合研究所作成の手書き漢字データベース ETL-8 (B2)⁴⁾及び ETL-9⁵⁾に含まれる平仮名71カテゴリーを用いた。これらのデータベースに含まれるサンプル群は位置と大きさの正規化が成されており、画面次数63×64の2値データである。各カテゴリーには、ETL-8 (B2) で160サンプル、ETL-9 で200サンプル含まれている。本実験では、サンプルの中に雑音のために位置と大きさの正規化が失敗しているものがあるため、全てのサンプルに対して、5×5マスクによる雑音除去処理を行い、一次変換の拡大・縮少の式を用いて、再度、位置と大きさの正規化を行い、最終的に画面次数32×32に縮少した。

表1にグループ分けした平仮名を示す。カテゴリー数は、濁点、半濁点を有するカテゴリー（グループ#1）が25、それ以外のカテゴリー（グループ#2）が46である。

ネットワークへの入力データとした局所領域における粗いメッシュ・パターンは、次に示す方法で作成した（図1参照）。まず文字パターンを16×16の小領域に分割する（各小領

表1 カテゴリーのグループ分け

グループ#1					グループ#2									
が	ぎ	ぐ	げ	ご	あ	い	う	え	お	か	き	く	け	こ
ざ	じ	ず	ぜ	ぞ	さ	し	す	せ	そ	た	ち	つ	て	と
だ	ぢ	づ	で	ど	な	に	ぬ	ね	の	は	ひ	ふ	へ	ほ
ば	び	ぶ	べ	ぼ	ま	み	む	め	も	や	ゆ	よ	ら	り
ぱ	ぴ	ぷ	ぺ	ぽ	る	れ	ろ	わ	を	ん				

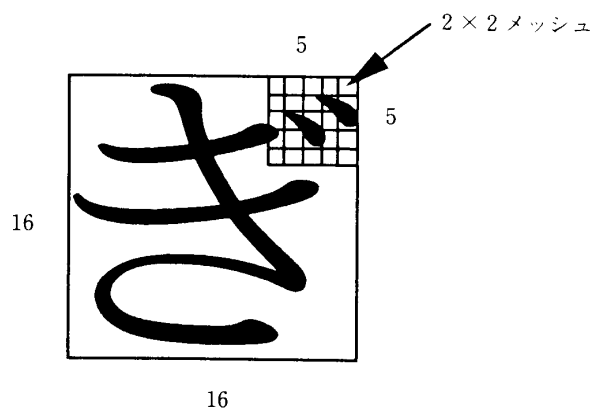


図1 局所領域における粗いメッシュ・パターン

域は 2×2 メッシュとなる)。次に、濁点、半濁点が存在すると思われる文字パターンの右上 5×5 領域に着目し、各小領域における黒点数を計数し、小領域の面積(=4)で除した値を特徴ベクトルの要素とする。したがって本論文における特徴ベクトルの次元数は25である。

3. ニューラルネットワーク

本論文では、ネットワークとして3層階層型ニューラルネットワークを用いた。学習則は、バックプロパゲーション則⁶⁾である。またユニットの応答関数は、シグモイド関数を用いた。ネットワークの構成ユニット数及び各種パラメータの値を表2に示す。ここで入力層ユニット数は、上述した特徴量(特徴ベクトル)の次元数と等しい。なお、結合荷重及び各ユニットのしきい値の初期値は乱数で与えた。またネットワークの学習は、グループ#1に含まれるサンプルを全てネットワークに入力した後にグループ#2に含まれるサンプルを入力するといったグループ単位で行うのではなく、グループ#1に含まれる1サンプルを入力した後はグループ#2に含まれる1サンプルを入力する(サンプルを交互に切り替える)方法で行った。

4. 実験

実験は、ワークステーション NWS-5000WI (SONY) を用いて行った。プログラミング言語はCである。

4.1 単純類似度法による識別実験

分類実験結果を示す前に、濁点、半濁点を有するカテゴリーとその他のカテゴリーに分類した場合の識別実験を行い、平仮名全カテゴリーを用いた場合との識別率の比較を行い、分類することの有効性を検証する。識別手法としては、最も基本的な重ね合せ的手法である単純類似度法¹⁾を採用した。

類似度とは、各カテゴリーに用意する標準パターンと入力パターン間の近さを測る尺度の一つであり、次式で定義される。

$$r^{(i)} = \frac{(x, T^{(i)})}{\|x\| \cdot \|T^{(i)}\|} \quad (1)$$

表2 ネットワークの構成ユニット数及び各種パラメータ

入力層ユニット数	25
中間層ユニット数	25
出力層ユニット数	2
結合荷重修正定数	0.4
しきい値修正定数	0.5
シグモイド関数の傾き	0.65
学習回数の上限	10,000

ここで、 x は入力パターン、 $T^{(i)}$ は i 番目のカテゴリーに用意された標準パターンを表す。識別は、入力パターンと全ての標準パターン間の類似度を計算し、最大の類似度値を示した標準パターンの所属するカテゴリーを入力パターンの決定カテゴリーとすることで行われる。標準パターンは、各カテゴリーにおいて標準パターン作成用サンプル（学習サンプル）の対応する要素を全て加え合わせ平均することで得た。標準パターン（平均ベクトル）の次元数は $32 \times 32 = 1024$ である。

識別実験は、

- (i) 平仮名全71カテゴリー
- (ii) 濁点、半濁点を有する25カテゴリー（グループ# 1）
- (iii) 濁点、半濁点を有しない46カテゴリー（グループ# 2）

の3つの場合について行った。実験は学習サンプルの内容を変えて3種類行っている。表3に各実験における学習サンプルの内容とサンプル数を示す。なお未知サンプルは、ETL-9の後半100サンプルに固定した。実験では棄却処理を設定せず、全てのサンプルに対して、正読あるいは誤読の判定を行った。

表4に得られた識別実験結果を示す。表中の値は平均識別率である。本実験では ETL-9に含まれるサンプルを未知サンプルとしたために、ETL-9に含まれるサンプルを学習サンプルとして採用した場合に最も高い識別率が得られている。また3つの実験結果より、グループ# 2の場合に最も良好な識別率が得られており、グループ# 1の場合に最も低い識別率が得られたことがわかる。したがってこの結果からは、濁点、半濁点の有無によってカテゴリーを分類することの有効性が明確ではない。そのため、分類した後の2つの識別実験で識別されたサンプル数をトータルのサンプル数（未知サンプルの場合7100サンプル）で除して識別率を求めた結果を表5に示す。分類前で示される識別率は、表4に示さ

表3 識別実験における学習サンプルの内容とサンプル数

データNo.	学習サンプルの内容とサンプル数
1	ETL-8 (B2) : 前半80サンプル
2	ETL-8 (B2) : 160サンプル
3	ETL-9 : 前半100サンプル

表4 識別実験結果（平均識別率（%））

データ No.	平仮名71カテゴリー		グループ# 1		グループ# 2	
	学習	未知	学習	未知	学習	未知
1	85.51	67.77	84.35	63.72	89.54	74.39
2	82.49	69.75	81.28	67.00	87.28	75.85
3	81.66	74.80	80.72	73.52	85.78	81.22

表5 識別実験結果 (平均識別率 (%))

データ No.	分類前		分類後	
	学習	未知	学習	未知
1	85.51	67.77	87.71	70.63
2	82.49	69.75	85.17	72.73
3	81.66	74.80	84.00	78.51

表6 分類実験における学習、未知両サンプルの内容とサンプル数

グループ# 1	ETL-8 (B2) に含まれるカテゴリの先頭サンプルから138サンプル 合計：3450サンプル	ETL-9に含まれるカテゴリの先頭サンプルから100サンプル 合計：2500サンプル
グループ# 2	ETL-8 (B2) に含まれるカテゴリの先頭サンプルから75サンプル 合計：3450サンプル	ETL-9に含まれるカテゴリの先頭サンプルから54ないし55サンプル 合計：2500サンプル
サンプルの合計	6900	5000

れた平仮名71カテゴリーの場合の識別率と同じものである。分類した場合、未知サンプルでは3～4 (%)の識別率の上昇が得られている。この結果より、濁点、半濁点の有無によってカテゴリーを分類することの有効性が示されたと思われるが、識別率の上昇自体は大きなものではない。この点に関しては、識別手法を分類されたグループごとに変えるという方法で対処することが考えられるが、これは今後の課題の一つとしたい。

4.2 分類実験

以下に示す2つの分類実験においては棄却処理は導入されておらず、全てのサンプルに対して正分類あるいは誤分類の判定が下されている。表6に分類実験における学習、未知両サンプルの内容とサンプル数を示す。なおニューラルネットワークによる分類実験では学習サンプル数を変化させて実験を行っており、表6に示した学習サンプルの内容は、最もサンプル数を増やした場合のものである。

4.2.1 特徴整合法による分類実験

比較のために、サンプルから抽出した25次元の特徴ベクトルを用いて特徴整合法による分類実験を行った。識別関数は上述の単純類似度を採用した。各グループの標準パターンは、学習サンプルから抽出された特徴ベクトルの対応する要素を重ね合わせ平均することで得た。入力サンプルのグループ決定は、より大きな類似度値を示す標準パターンの所属するグループに判定することで行った。

表7に得られた分類率を示す。同表には、表6に示した学習サンプルを用いた場合のニューラルネットワークにより得られた分類率も示してある。特徴整合法を用いた場合、学習、未知両サンプル共にグループ# 1に入力サンプルが判定される傾向が見られ、グルー

表7 分類実験結果 (分類率 (%))

	特徴整合法		ニューラルネットワーク	
	学習	未知	学習	未知
グループ# 1	85.51	84.24	98.32	92.96
グループ# 2	63.57	64.60	99.74	94.80
グループ# 3	74.54	74.42	99.03	93.88

プ# 2 の分類率との差が大きい。原因としては、例えば、“お”、“で”のようにストロークの一部が文字パターンの右上領域に存在するカテゴリーがグループ# 2 に数多く存在することが挙げられる。またグループ# 1 では、手書きのため濁点、半濁点の位置が変動してサンプルの誤分類が生じたと推測される。この結果から、単純に標準パターンを作成しマッチングを行う方法では、濁点、半濁点の有無の判別はかなり困難であることがわかる。

4. 2. 2 ニューラルネットワークによる分類実験

図2に学習サンプル数を変化させた場合の分類率を示す。学習サンプル数は、100, 200, 300, 500, 1000, 1500, 2000, 2500, 3000, 3450の10種類である。サンプル数3450の場合の内容は表6に示してあり、その他の場合もカテゴリーが偏らないようにサンプルを選出した。また未知サンプルに関しては、学習サンプル数に関係なく表6に示した未知サンプルの内容に固定した。

学習サンプルに関しては、サンプル数の増加に対する分類率の低下は見られるものの、平均では99(%)程度の分類率は維持している。また両グループの分類率の差もかなり小さい。これらは、良好なネットワークの学習が行えたこと及びネットワークの中間層ユニット数が妥当なサイズであったことを示していると思われる。未知サンプルでは、学習サンプル数500程度までは平均分類率の上昇は大きいですが、それ以後は緩やかな上昇となっており、ほぼ分類率の限界を示していると考えられる。また学習サンプルの場合と比較すると両グループの分類率の差が大きく、特徴整合法による実験結果とは逆に、グループ# 2 に入力サンプルが判定される傾向が見られる。ニューラルネットワークでは、濁点、半濁点部分のみの学習を行ったため、手書きのために濁点、半濁点の位置が大きく変動しているサンプルには対処できなかったと考えられる。表8に誤分類の多かった未知サンプルを含むカテゴリーの例を示す。グループ# 1 では、上述したように、濁点、半濁点の位置が大きく変動しやすいカテゴリーが多く、グループ# 2 では、濁点、半濁点位置にストロークの一部が記入されやすいカテゴリーが多い。これらのことから、文字パターンの右上領域を固定してネットワークを学習させる方法では、未知サンプルの分類率をこれ以上引き上げることは困難と思われる。分類率向上のためには、領域を固定しないか、あるいは領域をより広くとることが考えられ、特徴量に関しても、濁点、半濁点であるか、あるいはストロークの一部であるかを判別可能な文字パターンの構造に関する情報を持つ特徴量の使用

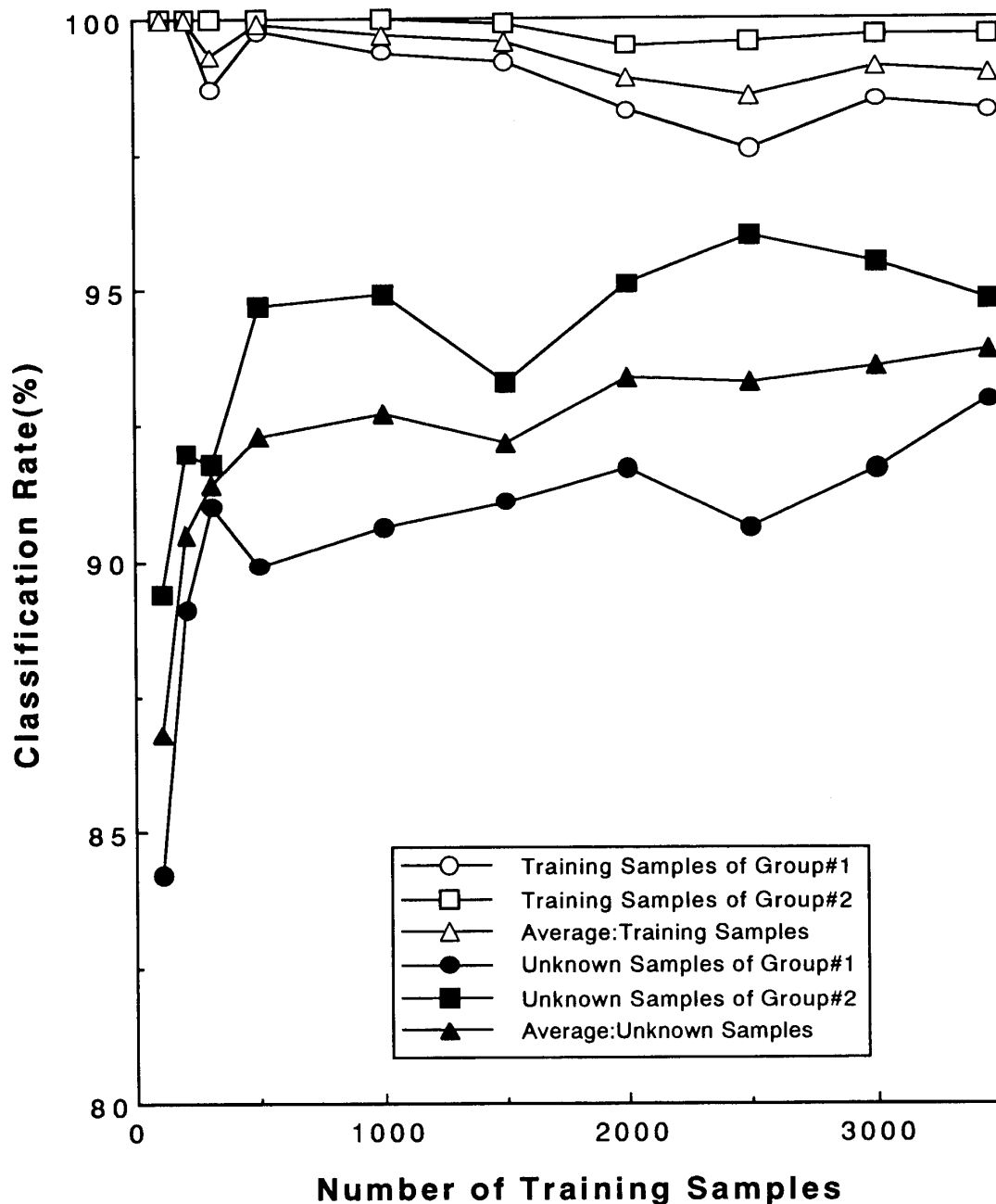


図2 ニューラルネットワークによる分類実験結果

表8 誤分類の多かった未知サンプルを含むカテゴリーの例

グループ#1					グループ#2						
ぐ,	じ,	ず,	で,	ど	お,	き,	く,	け,	こ,	さ,	て
び,	ふ,	ふ,	べ,	ぺ	な,	は,	ほ,	む,	や,	ら,	り

が必要と考えられる。

5. む す び

本論文では、3層階層型ニューラルネットワークにより手書き平仮名の濁点、半濁点の有無によるカテゴリーの分類を試みた。ネットワークへの入力データとして局所領域における25次元の粗いメッシュ・パターンを用い、6900サンプルを学習させたネットワークにより未知の5000サンプルに対して94(%)程度の平均分類率を得た。分類率向上のためには、濁点、半濁着位置の固定の緩和及び特徴量の再考の必要性があることが明確になった。また、分類後のグループに対する識別手法の検討も必要である。

最後に、本実験を行うにあたって御助力を頂いた大倉(充)研究室の平成5年度卒業生大久保智裕氏(現在、(株)松芝エンジニアリング)、佐藤 薫氏(現在、(株)難波プレス工業)、藤山武久氏(現在、(株)松芝エンジニアリング)、真野義和氏(現在、(株)富士技術開発)に厚く感謝する。

参 考 文 献

- 1) 橋本新一郎編著：文字認識概論，オーム社（昭57）。
- 2) 大倉 充，吉田 誠，原 卓宣，塩野 充，橋本禮治：ニューラルネットワークによる同じ部首を有する手書き漢字の抽出実験，第47回情処全大，1 L-5，pp. (2-83)－(2-84) (平05)。
- 3) 梅田三千雄：PDP モデルによる手書き漢字と平仮名の区別，信学技報，PRU91-125，pp. 95－102 (1992)。
- 4) 斎藤泰一，山田博三，山本和彦：手書文字データベースの解析 (V)，電総研彙報，45，1，2 (昭56)。
- 5) 斎藤泰一，山田博三，山本和彦：手書文字データベースの解析 (VIII)，電総研彙報，49，7，pp. 487－525 (昭60)。
- 6) 中野 馨監修，飯沼一元編：入門と実習ニューロコンピュータ，技術評論社 (平01)。

A Classification of Handwritten HIRAGANA Characters by 3-Layer Neural Networks

Mitsuru OHKURA, Mitsuru SHIONO and Reiji HASHIMOTO*

Department of Information & Computer Engineering,

**Department of Electronic Engineering,*

Okayama University of Science,

Ridai-cho 1-1, Okayama 700, Japan

(Received September 30, 1994)

Handwritten HIRAGANA characters are classified according to whether they have DAKUTEN (or HAN-DAKUTEN) or not by the 3-layer neural networks. The input data to the networks is 25-dimensional local mesh-feature extracted from an original character pattern. The numbers of units of three layers are 25(input-layer), 25(hidden-layer) and 2(output-layer). The numbers of training and unknown samples used in a classification experiment are 6900 and 5000, respectively. The average classification rate of 94(%) for the unknown samples is obtained.