

4L-3

# ニューラルネットによる 手書き漢字の詳細識別

安藤 陽介 阿部 一朗  
NTTデータ通信株式会社

## 1. まえがき

ニューラルネット[1]は現在、いろいろな分野に適用されている。特に、文字や音声などの認識の分野では、その高い認識能力が注目されている。われわれは、手書き漢字の認識、特に類似した文字群の中から目的とする文字を同定する詳細識別にニューラルネットを適用することを試みている。本稿ではニューラルネットの詳細識別を実現する方法として、ネットワークを識別関数として用いる従来から使われている方法と、手書き文字の変形による特徴ベクトルの変動を吸収するためのフィルタとして用いる2つの方法について識別実験を行い、後者のような方法が識別に有効であることを示す。

## 2. ニューラルネットの詳細識別への適用

われわれは文字認識を大分類、詳細識別の2段階で行うことを前提としており、そのうちの詳細識別に注目している。そのため、われわれの目的は、少ないカテゴリながら密接した分布となっている部分を識別することにある。ニューラルネットの適用の方法はいろいろ考えられるが、今回は従来から用いられてきた特徴をあらかじめ抽出し、それをニューラルネットへの入力として識別を行うこととした。実験では以下の2つの種類のネットワークを作成し、それぞれを充分学習させ、識別率を測定した。

### [方法1]：識別型ニューラルネット

識別対象である文字から抽出した特徴を入力とし、その文字の属するカテゴリに対応する出力ユニットを1、その他の出力ユニットを0としたベクトルを理想出力とする(図1)。

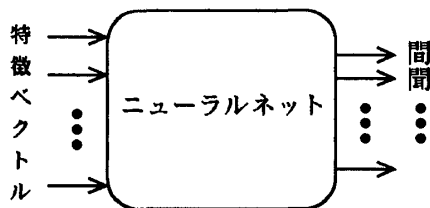


図1：識別型ニューラルネット(方法1)

これは最も一般的に使われる方法で、ニューラルネットは、従来の文字認識での識別関数の役割をする。

### [方法2]：特徴修正型ニューラルネット

識別対象である文字から抽出した特徴を入力、その文字の属するカテゴリの標準(平均)ベクトルを理想出力とする(図2)。この場合、ニューラルネットは字形の変動による特徴の変動を吸収する特徴正規化フィルタの役割をする。この場合識別は、このネットを通した後の特徴ベクトルを、辞書ベクトルと比較(距離計算)をすることで行う(図3)。ニューラルネットを識別に用いる場合、その出力ベクトルは識別結果となるカテゴリを表現している。この意味から、人為的なベクトルを出力ベクトルとする方法1よりも、カテゴリの特徴を反映している標準ベクトルを出力ベクトルとする方法2の方がより自然であると考えられる。また、今回の実験で用いるバックプロパゲーションアルゴリズムが、ネットの出力と理想出力の二乗誤差(L2距離)を最小にするように学習するため、辞書ベクトルとの距離計算で用いる距離をL2距離とすれば、つまり学習の評価関数と識別で用いられる距離関数を同一にすれば、学習が進むにつれ識別率の向上が比較的素直に現れると考えられる。

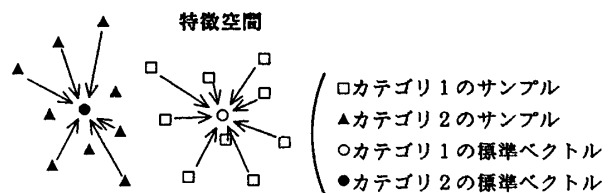


図2：方法2の概念図

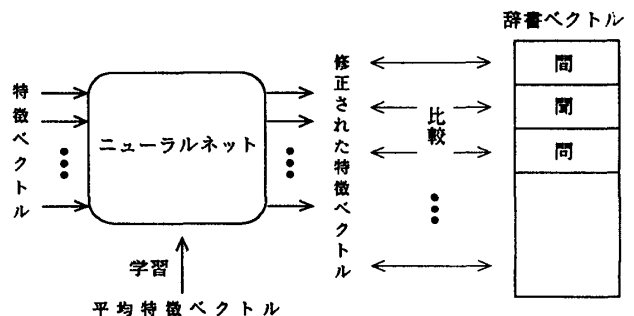


図3：特徴修正型ニューラルネット(方法2)

### 3. 識別実験

#### 3.1 実験方法

今回の実験では、類似した10カテゴリを「間」を中心に選び出し、これらを識別することとした。これらのカテゴリは特徴空間上で非常に接近しており、識別がかなり困難であることが分かっている。またニューラルネットに与える特徴として、ストローク構造集積法[2]に基づいた128次元のベクトルを用いた(表1)。

表1: 実験データ

対象カテゴリ	間間間間間 間間間間間
データ数	1800 = 10カテゴリ × 180サンプル (学習:900 未知:900)
特徴ベクトル	ストローク構造集積法 128次元 ハフマン特徴 : 40次元 外郭方向特徴 : 40次元 X方向ストローク数 : 24次元 Y方向ストローク数 : 24次元

ニューラルネットは3層のフィードフォワード型ネットワークを用い、学習アルゴリズムはバックプロパゲーションを用いた。中間層のユニットの数は試行錯誤の結果、方法1, 2共に100とした。学習は実験データ1,800サンプルのうち900サンプルを用い、これらを不規則な順序で呈示し、1000回学習する度に学習および未知サンプルの識別率を測定した。

#### 3.2 実験結果

以下に[方法1]と[方法2]の2つの実験結果、および通常の文字認識で行われる、最も簡単な識別法を用いた結果を示す(表2, 図4)。

表2: 実験データ 識別率% (誤読数)

	学習サンプル	未知サンプル
L2距離	77.0% (207)	72.8% (245)
重み付き L2距離	84.9% (136)	73.7% (237)
ニューラルネット 方法1	54.9% (406)	53.1% (422)
ニューラルネット 方法2	88.7% (102)	79.8% (182)

- ・ニューラルネットでの結果は未知サンプルの識別率の最も良いもの。
- ・L2距離 = ユークリッド距離

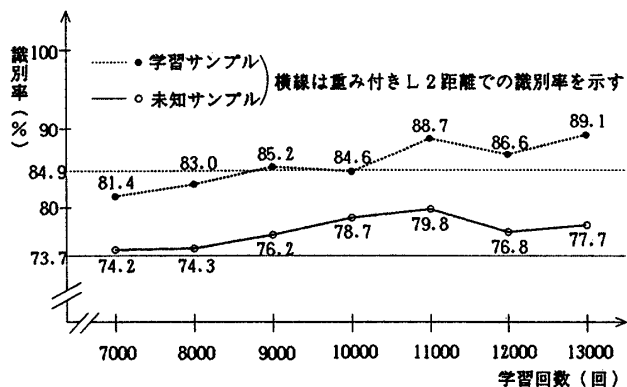


図4: 学習による識別率の推移 [方法2]

実験では、学習回数が11,000回の際に未知サンプルの識別率が最も高くなった。また、この時のネットを通した後の特徴ベクトルに対し、「間」に近い6カテゴリの辞書を用いて識別実験を行ったが、新たな誤読は起こらなかった。

今回の実験では[方法1]については十分な識別率を得ることはできなかったが、ネットワークの初期状態や中間ユニット数などのパラメータが最適であるとは言えないため、この結果が両者の能力の絶対的な差を示しているとはいえない。また収束性に関しては、実験において[方法2]が[方法1]に比べ容易に学習が進んだことから[方法2]の方が優れていると考えられる。これは[方法2]が、もともと入力と出力の差が少なく、また、入力と出力が同じ種類のベクトルであるためと考えられる。これにより、より単純な学習でニューラルネットの高い認識能力を利用することが可能となる。

### 4. あとがき

本稿では、手書き漢字の詳細識別にニューラルネットを適用する方法として2つの種類のニューラルネットを作成し識別実験を行った。この結果、特徴の変動を吸収するようなニューラルネットが識別率の面で有効であることを示した。この方法では、識別型のネットワークに比べ学習が軽くなる反面、識別結果を出すまでの処理手順が多くなるという欠点がある。今後の課題としては、①カテゴリを増やした場合の両者の識別率の測定、②特徴修正型ニューラルネットに対して、用いる特徴の性質により、変動する可能性の少ない要素の結合を減らし、より少ない学習で効果を得る方法の検討、があげられる。

#### 参考文献

[1]麻生英樹:

ニューラルネットワーク情報処理, 産業図書 (1988)

[2]赤松, 川谷, 塩, 飯田:

“手書き漢字用文字読取り装置”, 研究実用化報告, Vol. 36, No. 4 (1987).