

5N-5 並列多段ニューラルネットによる手書き漢字の認識*

荒井正之†

江刺家公也‡

王晋申‡

奥田健三‡

†帝京大学

‡宇都宮大学

1 はじめに

ニューラルネットを用いた文字認識の研究は今日盛んであるが、比較的少数の文字カテゴリを対象としたものが多い。しかし、手書き漢字を対象とした場合には、数千字種の認識が可能でなくてはならない。そこで、我々は小学校学習漢字996文字を対象として、ベクトル量子化型ニューラルネット(VQネット)を用いて大分類の後、誤差逆伝播法を用いた階層型ニューラルネット(階層ネット)により認識を行うネットワークを構築し、実験、考察を進めてきた。VQネットを2段構成にすることにより、処理効率が大幅に向上し、分類数の均等化に寄与することを確認した[1]。今回は階層ネットを結合するとともに、認識率を向上させるため、さらにこれらを並列にし、それぞれに相補的な特徴を入力するネットワークを構築した。その結果について報告する。

大分類、階層ネットについては、[2][3]で詳述する。

2 ネットワーク構成

ネットワーク構成を図1に示す。まず最初に、VQネットを2段設け大分類を行い、比較的少数のカテゴリの文字に分類する。分類後、入力層と中間層を部分結合させた階層ネットにより認識を行う。さらに今回は、このような構造のネットワークを2つ(LLネット, GLネット)並列に用意した。

LLネットにおいては、局所的方向寄与度密度特徴(LDCD)[4]128次元を入力し、大分類、認識ともにこの特徴を用いる。GLネットにおいては、LDCD128次元と大局的方向寄与度密度特徴(GDCD)[4]64次元、合計192次元の特徴ベクトルを入力する。実際には、大分類にGDCDを使用し、認識にはLDCDを用いる。2つの並列に配置されたネットワークに、相補関係の強い特徴を用いることにより、相互に補完しあい認識率の向上が期待される。

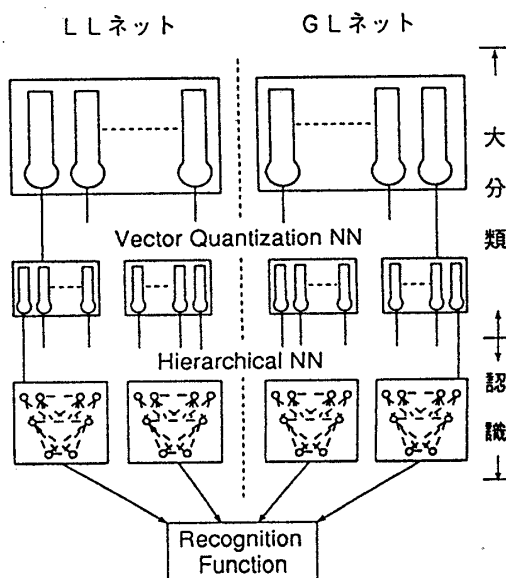


図1 ネットワーク構成

3 実験と結果の考察

電総研の手書き文字データベース ETL9 [5] から、学習データ 30 サンプル、未学習データ 10 サンプルを用いて実験を行った。計算機はワークステーション(NEWS 3460)を使用した。

3.1 階層ネットの学習

階層ネットにおいては、2層にて分類されたカテゴリを学習させる。2層 VQ ネットでは、平均して 10 カテゴリに分類されるようにニューロンを配置したが、実際に学習するカテゴリ数は、表 1 に示すとおりとなった。階層ネットの学習に要した時間を表 2 に示す。学習回数は上限を 300 回とした。中間層ニューロン数は、出力層ニューロン数によって、12, 16, 20 個と変化させた。

通常の階層ネットでは、入力層-中間層-出力層のそれぞれの層のニューロン間は全て結合されている。我々は、認識率を低下させずにシナプス数を削減できる図 2 に示す入力層-中間層を部分結合させた階層ネットを用いた [3]。この部分結合法をとると、入力特徴ベクトルが 2 分割され中間層ニューロ

*Recognition of Handwritten Kanji by Parallel Multi-layered Neural Network

†Masayuki ARAI: Teikyo University

‡Kimiya ESASHIKA, Wang Jin Shen, Kenzo OKUDA: Utsunomiya University

表1 カテゴリ数と階層ネット数

カテゴリ数	6以下	7-13	14以上
LL ネット	24	56	19
GL ネット	17	65	17

表2 階層ネット学習時間 (CPU 時間)

LL ネット	GL ネット	計
6h03m	5h35m	11h38m

ンのそれぞれ半数が分担することになる。全結合に比較し、部分結合では表3に示すとおり、シナプス数は半減する。

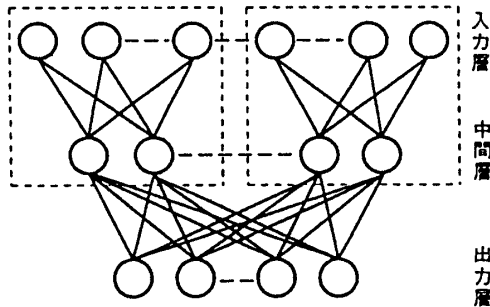


図2 階層ネットの構造

3.2 認識実験

認識においては、まずLL,GL ネットともに1層の整合度上位3位、2層整合度上位3位、合計3×3=9個の2層ニューロンを担当する階層ネットの出力値が最大のものを候補文字とする。9つの候補の中から、1,2層の正規化した距離と5層の出力値を結合させた下記の識別関数やその他数種類の識別関数を用い、認識実験を行った。

$$(1 \text{ 層距離} + 2 \text{ 層距離})^\alpha \times (5 \text{ 層出力値})^\beta \quad (1)$$

$$(1 \text{ 層距離})^\alpha \times (2 \text{ 層距離})^\beta \times (5 \text{ 層出力値})^\gamma \quad (2)$$

表3 シナプス数 (中間層 16 個, 出力層 10 個の場合)

	入力-中間	中間-出力	計
全結合	2048	160	2208
部分結合	1024	160	1184

表4 α, β, γ と認識率

		α	β	γ	認識率
LL ネット	学習	1.0	1.5	4.1	81.90%
	未学習	1.0	2.7	0.5	58.60%
GL ネット	学習	1.0	1.6	12.6	73.77%
	未学習	1.0	2.1	2.0	52.17%

1,2層で誤分類された文字は、正認識されることはないことを考慮すると、1,2層のウェイトが大きくなることを予想していた。しかし、(2)式で最も認識率の高い α, β, γ は表4に示すとおりとなった。正規化された1,2層距離や出力値はともに0-1の間の値であるから、LL ネット未学習のみ第1,2層の距離よりも第5層の出力値をより強く反映しているという結果になった。

上記の方法では、LL,GL ネットから1つずつ計2つの認識結果が得られる。1つの認識文字に絞り込むため、LL,GL ネットそれぞれの識別関数(2)式の値にさらにウェイトをつけて認識実験を行った。その結果、最も高いもので学習 84.68%、未学習 62.80%になった。

4 おわりに

相補的な特徴を用いた並列ネットは、個々のネットの認識率を上回り、学習 84.68%、未学習 62.80%という結果になった。

しかし、本方式はLL,GL ネットで合計18個の候補文字を選んでおり、累積分類率にすると学習 96.89%、未学習 89.43%になる。さらに、分類、識別方法を検討することにより、認識率の向上を図りたい。

参考文献

- [1] 荒井,他:多段ニューラルネットワークによる手書き漢字の分類,情処第43回全大,G-8,1991
- [2] 江刺家,他:並列多段ニューラルネットワークによる手書き漢字の大分類,情処第44回全大,5N-6,1992
- [3] 王,他:階層型ニューラルネットワークによる手書き類似漢字の認識,情処第44回全大,5N-7,1992
- [4] 萩田,他:大局的・局所的方向寄与度密度特徴による手書き漢字認識方式,信学論 D,J66-D,6,pp722-729,1983
- [5] 斎藤,他:JIS第1水準手書き漢字データベース ETL9 とその解析,信学論 D ,J68-D,4,pp758-764,1985