

5E-5

ニューラルネットワークによる
手書き類似漢字認識について

武中 裕司 塩野 充
岡山理科大学工学部電子工学科

[1] まえがき

ニューラルネットワークの形態は、フィードバック結合を含むものと含まないものとに大別できる⁽¹⁾。本研究では、前者に属するタイプの1つである層内結合を持つ階層的なネットワーク(3層)を使用し、手書き類似漢字の認識実験を行った。尚、学習は中間層に直接教師信号(例えば、言偏(ごんぺん)を学習せよ、等)を与え、前向きに行った。

[2] 特徴の抽出

手書き文字認識では、一般的に構造解析的な手法(例えば、ストロークあるいはセグメントの抽出に基づく手法)が有効である。しかし、複雑な手書き漢字においては、つぶれなどによりその抽出処理は一般に困難であり、常に良好な抽出結果を得る事は難しい。しかし、認識手法としてニューラルネットワークを用いる場合、その性質(かすれやつぶれに強い)により、不完全な抽出においても特徴量として十分利用できると思われる。本研究では、単純な手法で抽出した4方向のセグメントを用い、認識実験を行った。まず原画像(画面次数は63×64)の黒領域(画素値1)の各画素について、図1-aに示す8方向に黒点を追跡し、最も長い方向をその画素が属する方向であると決定する。次に、手書きによる癖の吸収を目的として方向の合併を行う(図1-b)。

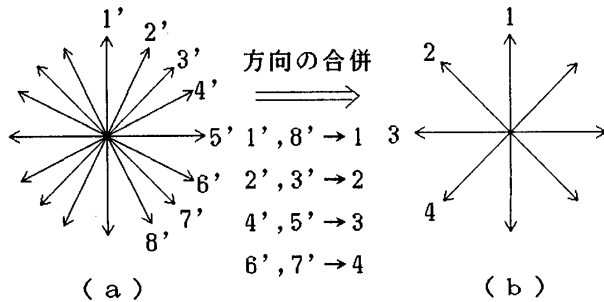


図1. セグメントの方向

また、微少な位置ずれやかすれを吸収するため2段階の画面次数縮小を行った(2×2画素→1画素: 4×4画素→3×3画素)。それぞれの段階における画素値は、式(1)に従って計算される。

$$P_k = \sum_{i=1}^m \sum_{j=1}^n f_k(i, j) - \frac{1}{2} \times \sum_{k=1}^4 \sum_{i=1}^m \sum_{j=1}^n f_k(i, j) \quad (1)$$

但し、 P_k : 画素値(0以下は0)
 f_k : 画面次数縮小を行う前の画素値
 k : セグメントの方向に対応
 第2項: 小数点以下切り捨て
 m, n : $m=n=2$

[3] ニューラルネットワーク
 本実験に用いたニューラルネットワークを図2に示す。擬似ニューロンとして、線形しきい値ユニットを使用した。層内結合は全て抑制結合であり、競合作用を表わしている。

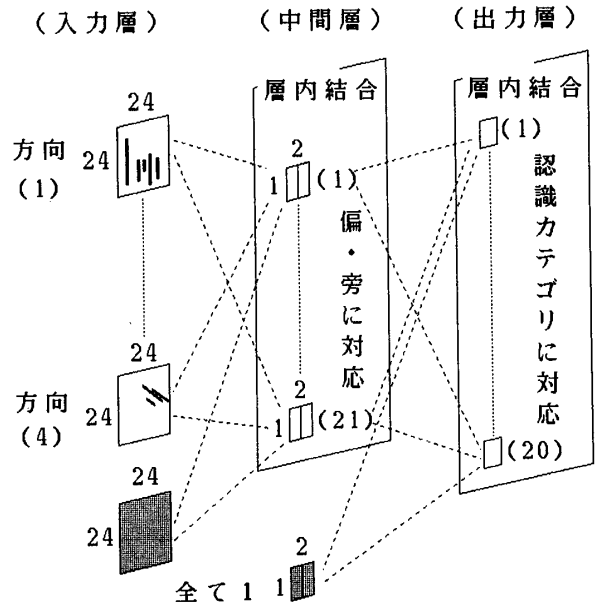


図2. 本方式におけるニューラルネットワーク

(3-1) 学習

ニューラルネットワークを用いた、文字認識についての種々の研究によって、ある条件下において中間層(内部表現)は入力特徴空間の偏(へん)あるいは旁(つくり)を学習しているということが報告されている⁽²⁾。そこで、中間層に認識を行うカテゴリの偏、旁に対応する数だけのユニットを用意して、それぞれのユニットに何を学習すべきかを直接教える。学習は、前向きに行い、初期には興奮性の学習を中心に、徐々に抑制性の学習を強めることとした。学習(シナプス結合係数の変化)は、式(2)に従う。

On recognition of similar handprinted KANJI characters using neural networks.

Yuji TAKENAKA, Mitsuru SHIONO
 Okayama University of Science.

$$W_{NEW} = \alpha \times W_{OLD} + \beta \times W' \pm 0.01 \times PIX \quad (2)$$

但し、 W_{NEW} : 学習により変化した新しいシナプス結合係数 (しきい値を含む)

W_{OLD} : 学習前のシナプス結合係数 (しきい値を含む)

W' : W_{OLD}' (W_{OLD} の0以下を0としたもの) $\times PIX'$ (入力信号を2値化したもの)

α : 過去の重み (学習が進むにつれ大きくする)

β : 学習速度 (学習が進むにつれ小さくする)

\pm : 興奮性学習のときは+
抑制性学習のときは-

PIX : 入力信号

第3項: 荷重0からの脱出

学習は、教師信号との差が許容誤差内の値になるまで同じデータを使って繰り返し学習する。一定回数の繰り返しを過ぎても許容誤差内に収束しない時は、ノイズを発生させ状態の脱出を試みる。また、1つのデータにおける最大のループ回数を決めて、切り出し、正規化の失敗など変形の大きな文字による過度学習を避ける。

[4] 認識実験

(4-1) 実験概要

本実験に用いたデータは、電総研JIS第1水準手書き漢字データベースETL-9(B2)であり、1カテゴリにつき200サンプルのデータが格納されている。前半の50個、あるいは100個のサンプル(学習サンプルとする)を使って学習を行い、後半100個のサンプル(未知サンプルとする)について認識実験を行う。実験は、表1に示す20カテゴリの類似漢字について行った⁽³⁾。使用した計算機は本学FACOM/M380、言語はFORTRANである。

表1. 使用カテゴリ

維轄幹騎験講購債情績
難備偏編愉論輸倫輸論

(4-2) 実験結果

表2に平均認識率を示す。また、比較のため手書き類似漢字の認識実験によく用いられる4カテゴリ(①論、輸、輸、論 ②詰、結、紹、詔 ③渴、掲、湯、揚)のみについて、それぞれ学習、認識を行ったときの結果も表3に示しておく。

表2. 平均認識率(20カテゴリ)

学習サンプル数(個)		認識率(%)	
中間層	出力層	学習サンプル	未知サンプル
50	50	99.8	79.4
100	100	98.6	81.2

表3. 平均認識率(4カテゴリ)

学習サンプル数: 中間層, 出力層とも100個

認識対象カテゴリ		認識率(%)	
		学習サンプル	未知サンプル
①	論 輸 輸 論	100.0	85.7
②	詰 結 紹 詔	100.0	94.2
③	渴 掲 湯 揚	100.0	95.7
平均		100.0	91.9

(4-3) 考察

表2において、学習サンプル数を100個に増やしたとき、未知サンプルにおける認識率は、わずかに向上したものの、1.8%に留まった。また、そのときのリジェクト(認識棄却: 出力層において出力1を出すユニットが存在しないとき、または2個以上のユニットが出力を出した場合)率は、1.6%減少しており、単に学習サンプル数を増やすというだけでは、誤読が減少したというより、リジェクトされていた文字が認識できるようになったという表現の方が正しいと言えるだろう。また、表3の結果から認識手法としてニューラルネットを用いる場合、抽出不完全な方向セグメントでも特徴量として利用できると思われる。

[5] むすび

層内結合を持つ階層的なニューラルネットワークを用い、手書き類似漢字の認識実験を行った。20カテゴリに対し平均81.2%、4カテゴリに対しては平均91.9%の認識率を得た。まだまだ実用になる数字は得られなかったが、認識手法としてニューラルネットを用いる場合、抽出不完全な方向セグメントでも特徴量として利用できること、また中間層に学習すべき偏や旁を直接教える前向きの学習によっても正しい学習ができる(ある条件下において中間層は入力特徴空間の偏あるいは旁を学習している⁽²⁾)ということが確認された。今後は、中間層と出力層における学習サンプル数の変化や平均パターンを使った学習による認識率の検討、改善に努めたいと考える。

[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) 森, 横澤, 梅田: "PDPモデルによる手書き漢字認識", 信学技報, MBE87-156, Vol. 87, No. 428, pp407-414, 1988年3月.
- (3) 梅田三千雄: "文字認識と単語辞書", 信学技報, PRU88-17, pp9-16, 昭63-06.