

解 説

大規模樹型ニューラルネット CombNET-II による文字認識

岩 田 彰

名古屋工業大学電気情報工学科 〒466 名古屋市昭和区御器所町

(1991年8月2日受理)

Character Recognition by a Large Scale Neural Network CombNET-II

Akira IWATA

Department of Electrical and Computer Engineering,
Nagoya Institute of Technology,
Gokiso-cho, Showa-ku, Nagoya 466

1. ニューラルネットとは

1兆個もの神経細胞（ニューロン）をもつわれわれの脳神経回路網は、外界から入ってくる文字、音声、イメージなど極めて数多くのパターン情報を的確に素早く判断・識別する能力を持っている。このような脳神経回路網のパターン識別能力は、長い年月をかけて経験や学習により培われたものである。経験や知恵は神経細胞と神経細胞の結合関係の中に蓄えられている。そのような神経細胞のネットワークを神経回路網（ニューラルネットワーク）という。本当に成功しているパターン認識システムは人間の脳だけであるといわれるよう、脳神経回路網によってさまざまな情報処理を行うことができる。（人工）ニューラルネットワークは、このような脳神経回路網に類似した並列処理機構を用いた情報処理様式に基づいて高度な情報処理を行おうとするものである（図1）。

ニューラルネットワークの特徴は、学習と並列処理である。学習とは、あらかじめ用意された学習データに基づいて、ある望ましいネットワークの入出力関係を実現することである。学習によってネットワークの結合重みを決定するため、どのように学習するかという学習方式があれば学習データを構造のあるいはアルゴリズム的に分析しなくてもよいため、同じソフトウェア・ハードウ

エアをさまざまなアプリケーションに適用できる利点がある。

また、ニューロンにおける演算は非常に単純であり、その出力はそのときのニューロン内部に記憶されている値（結合重み）とそのニューロンへの入力値のみに依存するので、それぞれのニューロンにおける演算は独立に並行して行うことができる。すなわち本質的に並列処理であることからマルチプロセッサ方式のコンピュータに對して優れた適合性を持ち、高速演算が可能である。数個のプロセッサを持つシステムからニューロンの数と同じ数のプロセッサを持つ超並列システムまで考えられている。

ニューラルネットワークの応用分野は広いが、その中でも文字認識や音声認識などのパターン認識は最もよく適用される分野の一つである。ニューラルネットを用いると、学習によって非線形な識別境界が獲得でき、従来技術によるパターン認識能力を越える認識機械を構成することができると期待されている。ここでは、われわれが考案した大規模樹型ニューラルネットワークモデル CombNET-II のネットワーク構造と学習方式について解説し、このネットワークが、従来からよく用いられている BP 学習則による階層型ネットに比べて優れたパターン認識能力や学習安定性をもつことを示すとともに、CombNET-II を用いて印刷文字や手書き文字の認

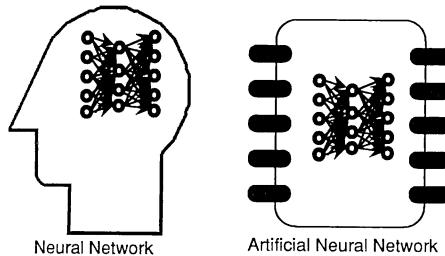


図1 脳神経回路網と人工ニューラルネット

識実験を行った事例について紹介する。

2. バックプロパゲーションは万能ではない

文字認識や音声認識などのパターン認識問題とは、入力データから得られる各種特徴量に基づいてそのデータの属するカテゴリを推定する問題である。これは、入力データが分布している入力特徴空間をカテゴリごとに分割する識別境界を決定する問題と考えられる。3層以上の階層型ニューラルネットによって任意の連続写像が実現可能¹⁾であることが証明されており、ニューラルネットによる学習によって任意の非線形な識別境界を獲得できることになる（図2）。

しかし、分類カテゴリ数が増加するなどして識別境界が複雑になると、階層型ニューラルネットの学習法としてよく用いられるバックプロパゲーション学習²⁾では識別境界を獲得することは実際には多くの困難を伴う。そのような大規模なネットワークの学習では、ローカルミニマムに陥る可能性も高く、たとえ収束するにせよ膨大な計算量を費すことになる。BP学習則はネットワークの出力を望ましい値（教師信号）に少しずつ近付けるべく、ニューロン間の結合重みを修正していく方式であり、原理的には最急降下法である。分類すべきカテゴリ数が増えネットワークが大規模になると、誤差曲面が複雑になり、最適な結合重みを見つけるまでに局所的な最適解（ローカルミニマム）にトラップされてしまい、それ以上学習が進まなくなってしまう（図3）。最適解（グローバルミニマム）は必ず存在するのにそこに到達できなくなる。結合重みの初期値を変えればうまくいくこともあるが、ネットワークが大規模になればなるほど最適解を見つける確率が低くなる。実際にこれまでに行われたニューラルネットワークに関する研究では、分類カテゴリ数の少ない比較的小規模なニューラルネットを取り扱っており、実用的で大規模なニューラルネットにおける成果をそのまま拡張できるかは疑問である。

われわれは、このような大規模ニューラルネットワー

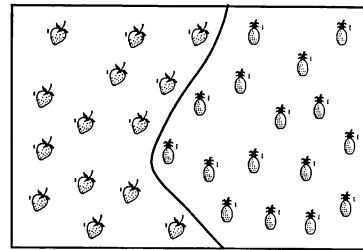


図2 ニューラルネットによって任意の非線形な識別境界を獲得

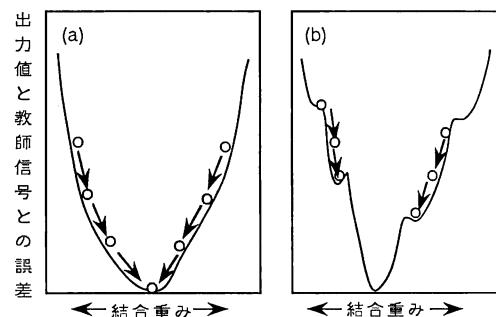


図3 結合重みの逐次修正による出力値と教師信号の誤差の減少

(a)誤差曲面が单峰性（小規模なネットワーク）、(b)誤差曲面が多峰性（大規模なネットワーク）

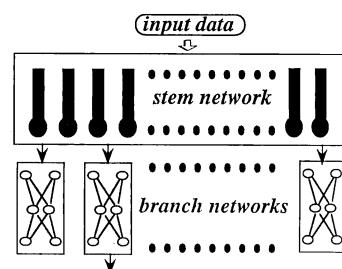


図4 CombNET-II のネットワーク構造

クの構築手法に関してひとつの手法を提案している³⁾。これは前段に入力データを大分類するためのベクトル量子化ニューラルネット（stem network）を配置し、後段にはグループ内のデータを細分類する階層型ニューラルネット（branch network）を多数並列に配置したものである。これによって前段で大まかなグループ分けを行い、後段でグループ内での識別を行うネットワークを構築する。この大規模樹型ニューラルネットワークモデルをCombNETと名付けている（図4）。

CombNETにおけるstem networkは、それぞれのニューロンの結合重みによって表される参照ベクトルを

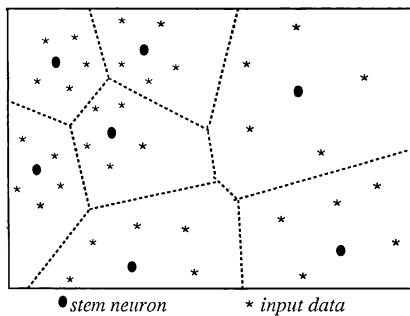


図 5 Stem network は入力特徴空間を大分割

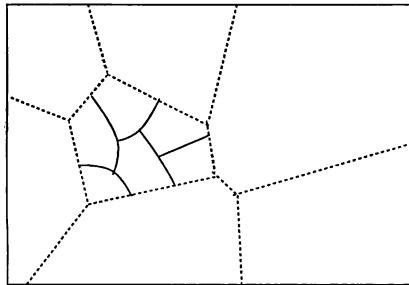


図 6 Branch network は部分空間内を細分割

核として、入力特徴空間を大分割（ベクトル量子化）する役割を担う（図5）。Branch network は、stem network により大分割された部分空間内を細分割し、そこに存在するカテゴリを個別に識別する境界を形成する（図6）。それぞれの branch network が担当するカテゴリ数はそれほど多くならず、したがって branch network の規模は小さく学習の収束は容易になり、全体として多数のカテゴリを識別する大規模なネットワークを安定に構築することができる。問題は、大分類を行う stem network を学習によりいかにして構築するかである。

3. CombNET-II は大分類のあと細分類する 2 段構え

最初に発表した CombNET-I³⁾ では、stem network の形成を Kohonen の自己組織化アルゴリズム⁴⁾により行い、branch network の学習は、バックプロパゲーション法を用いて行っていた。しかし、Kohonen の自己組織化アルゴリズムでは、大分類された各グループのメンバー数を制御することができず、その数が各グループごとに大きくばらついてしまう問題があった。大分類される入力データ数の大小は、後段の branch network のサイズにそのまま反映される。このことは、各 branch

network の学習深度に影響を与え、認識時に出力ニューロンの発火度にばらつきを生じさせる。われわれは、このような観点から独自のベクトル量子化ニューラルネットとして自己増殖型ニューラルネットを新しく考案した。そしてこれを CombNET の stem network の学習アルゴリズムに取り入れ CombNET-II と名付けた⁵⁾。

自己増殖型ニューラルネットは、一つのニューロンと最適整合する入力データ数がある閾値を超えた時に、ニューロンを分裂させ、分裂前のニューロンと最適整合していた入力データ群を、分裂後二つのニューロンのどちらかに最適整合する二つのグループに 2 分割する手法を基本としたもので、本手法を用いると一つのニューロンと最適整合する入力データ数を制限することができる。

自己増殖型学習則は次のように定式化される（図7 参照）。ここで

$$\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn})$$

（第 k 番目の入力ベクトル， $k=1 \dots K$ ）

$$\mathbf{m}_i = (m_{i1}, m_{i2}, \dots, m_{in})$$

（ニューロン i の参照ベクトル）

n ：入力ベクトルの次元数

g_i ：ニューロン i の内部ポテンシャル

g_{th} ：細胞分裂を起こす内部ポテンシャルの閾値

r_{th} ：新しいニューロンを生成する時の整合度の閾値

v ：生成されたニューロンの数

と定義する。

3.1 第 1 過程

[Step 0] 初期状態は、一つもニューロンは存在しな

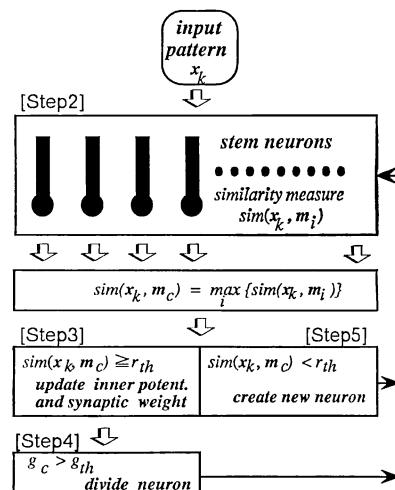


図 7 自己増殖型ニューラルネットにおける学習則

い。

[Step 1] 第1番目の入力ベクトル \mathbf{x}_k を参照ベクトル \mathbf{m}_1 とするニューロンを生成する。

各入力ベクトル \mathbf{x}_k について Step 2, Step 3, Step 4 を繰り返す。

[Step 2] 入力ベクトルと最も高い整合度を示す最適整合ニューロン c を求める。

$$\text{sim}(\mathbf{x}_k, \mathbf{m}_c) = \max_i \{\text{sim}(\mathbf{x}_k, \mathbf{m}_i)\} \quad (1)$$

(ニューロン c : 最適整合ニューロン)

ここで、整合度は(2)式のように入力ベクトルと参照ベクトルの相互相關によって求める。

$$\text{sim}(\mathbf{x}_k, \mathbf{m}_i) = \frac{\mathbf{x}_k \cdot \mathbf{m}_i}{|\mathbf{x}_k| \cdot |\mathbf{m}_i|} \quad (2)$$

もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_i) \geq r_{\text{th}}$ のときは、Step 3 へ進む。

もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_i) < r_{\text{th}}$ ならば Step 5 へ進む。

[Step 3] 内部ポテンシャルを更新し、 $g_c = g_c + 1$ とする。次式により参照ベクトルの修正を行う。

$$\mathbf{m}_{c(\text{new})} = \mathbf{m}_{c(\text{old})} + \frac{1}{g_c} (\mathbf{x}_i - \mathbf{m}_{c(\text{old})}) \quad (3)$$

これは、新しい参照ベクトルがそのニューロンに最適整合する入力ベクトル群の平均値になるように修正することを示している。

もし、 $g_c > g_{\text{th}}$ 、すなわち、内部ポテンシャルが細胞分裂する閾値に達していたら、Step 4 を行う。

$g_c \leq g_{\text{th}}$ の場合は、Step 2 に戻る。

[Step 4] 細胞分裂を行う。まず、ニューロン c の参照ベクトル \mathbf{m}_c を通る次式で示される超平面を生成する。

$$\mathbf{a} \cdot (\mathbf{x} - \mathbf{m}_c) + a_0 = 0 \quad (4)$$

\mathbf{a} , a_0 は超平面のパラメータであり乱数で与える。この超平面によってニューロン c に最適整合している入力ベクトル群 $\{\mathbf{x}_k\}$ を2分する。ただし、この2分された入力ベクトルの数に偏りがある場合は、等分割されるまでランダムに超平面を生成し、分割の繰り返しを行う。

新しい参照ベクトルは、それぞれの最適整合入力ベクトル群の平均とする。これは、Hebb 学習に基づく安定平衡解を最初から与えることに相当する。

$$\mathbf{m}_{c(\text{new})} = \frac{1}{g_c/2} \sum_j \mathbf{x}_j \quad \text{for } \mathbf{a} \cdot (\mathbf{x}_j - \mathbf{m}_{c(\text{old})}) + a_0 \geq 0 \quad (5)$$

$$\mathbf{m}_{c+1} = \frac{1}{g_c/2} \sum_k \mathbf{x}_k \quad \text{for } \mathbf{a} \cdot (\mathbf{x}_k - \mathbf{m}_{c(\text{old})}) + a_0 < 0 \quad (6)$$

それぞれのニューロンの内部ポテンシャルは、

$$g_{c(\text{new})} = g_{c+1} = \frac{1}{2} g_{c(\text{old})} \quad (7)$$

とし、ニューロンの個数は、

$$v = v + 1 \quad (8)$$

とする。

[Step 5] 新しくニューロンを生成する。すなわち、整合度が閾値 r_{th} 以上となるニューロンが見つからない場合、入力ベクトルそのものを参照ベクトルとするニューロンを生成する。

$$\mathbf{m}_{v+1} = \mathbf{x}_k, \quad g_v = g_v + 1, \quad v = v + 1 \quad (9)$$

3.2 第2過程

各入力ベクトル \mathbf{x}_k について順に Step 2', Step 3', Step 4 を内部ポテンシャルの変動がなくなるまで繰り返す。

[Step 2'] 入力ベクトルと最適整合となるニューロン c' を求める。

$$\text{sim}(\mathbf{x}_k, \mathbf{m}_{c'}) = \max_i \{\text{sim}(\mathbf{x}_k, \mathbf{m}_i)\} \quad (10)$$

もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_{c'}) \geq r_{\text{th}}$ のときは、Step 3' へ進む。

もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_{c'}) < r_{\text{th}}$ ならば Step 5 へ進む。

[Step 3'] 今までの最適整合ニューロン c と新たなる最適整合ニューロン c' とが異なった場合、内部ポテンシャルと参照ベクトルの更新を行う。

$$g_{c'} = g_{c'} + 1, \quad g_c = g_c - 1$$

$$\mathbf{m}_{c'(\text{new})} = \mathbf{m}_{c'(\text{old})} + \frac{1}{g_{c'}} (\mathbf{x}_i - \mathbf{m}_{c'(\text{old})}) \quad (11)$$

$$\mathbf{m}_{c(\text{new})} = \mathbf{m}_{c(\text{old})} - \frac{1}{g_c} (\mathbf{x}_i - \mathbf{m}_{c(\text{old})}) \quad (12)$$

もし、 $g_{c'} > g_{\text{th}}$ 、すなわち、内部ポテンシャルが細胞分裂する閾値に達していたら、Step 4 を行う。

$g_c \leq g_{\text{th}}$ の場合は、Step 2' に戻る。

Step 4 と Step 5 は、第1過程における処理と同様に行う。

このような方法に基づいてニューラルネットの形成をしていくと、特徴空間で入力データの分布密度が高い部分において細胞分裂が多くなり、入力データの分布密度に適合したニューロンの生成が行われる。このことにより入力データの分布状況に適合した準最適なボロノイ分割を行う量子化ニューロンの生成が可能となり、特徴空間におけるニューロン位置に対応する参照ベクトルもそのニューロンに最適マッチングする入力データ群の共通の特徴を表現するものになる。そして、一つのニューロンを最適整合とする入力データの数も、細胞分裂の内部ポテンシャルの閾値により上限が制限され、下限もその2分の1よりも著しく小さくなることはない(図5)。

CombNET-II の学習は、まず stem network ニューロンの生成から行う。これにより生成されたニューロンの参照ベクトルには近隣のカテゴリを代表する共通パターンが成長する。Stem network 学習の後、入力データの各カテゴリが stem network のどのニューロンの参照ベクトルに最も類似しているか（最適整合）を調べ、入力データを stem network のニューロン数と同数のグループに大分類する。次に、その大分類されたグループごとにその中にあるデータを識別するための階層型ニューラルネットの学習をバックプロパゲーション法を用いて行う。

識別時には、まず、stem network によって入力データがどのニューロンと最適整合となるかを求める。次に、最適整合となったニューロンが担当するカテゴリグループを分類する後段の階層型ネットワークに入力データを入力し、最も強く発火したニューロンに対応するカテゴリに入力データを識別する。

この CombNET-II の最大の特徴は、stem networkにおいて、ほぼ均一な入力データ分割を行うことができる点にある。これによって、各 branch network の出力層の数をほぼ均一な大きさにでき、各 branch network の規模もそろえることができる。このため、各 branch network の学習深度をほぼ均一にすることができ、CombNET-Iにおいて問題になっていた branch network の学習の不均一性をなくし、バランスのとれた階層型ニューラルネットの出力値を得ることができる。

4. JIS 第1, 2水準印刷漢字識別実験

CombNET-II のパターン識別能力を示すため、JIS 第1, 2水準の印刷漢字（6,349 文字）を識別する実験を行った⁶⁾。イメージスキャナから読み取ったデータから 16×16 の正規化濃度情報に変換した文字パターンを作成した。今回、これを 8 セット作成し、そのうちの 4 セットの平均パターンを学習データにした。また残り 4 セットで未学習パターンに対する識別能力の評価を行った。

図 8 は、この方式によって生成された stem network のニューロンの参照ベクトルの値を 16×16 の 2 次元マトリックス状のパターンで示したもので、各参照ベクトルがどのようなテンプレートを形成しているかが明らかとなっている。参照ベクトルには、漢字の「へん」、「つくり」、「かまえ」などの部首が形成されている。例えば、左上のニューロンには「門」構えを表す参照ベクトルが得られている。「宀」、「辶」など多くの文字で使われ

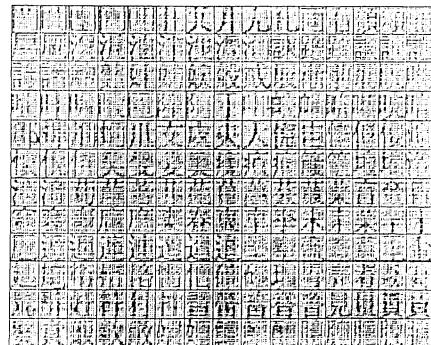


図 8 自己増殖型学習則によって生成された量子化ニューロンの参照ベクトル (JIS 第1, 2水準印刷漢字)

る部首は分裂回数が増えるため複数の参照ベクトルに表れている。Stem network の自己増殖型学習則によってニューロンの参照ベクトルに漢字の部首に相当するテンプレートが自律的に形成されたことは注目に値する。

この結果は、特定の視覚的特徴に選択的に反応する脳細胞がかたまって存在するという Hubel, Wiesel の報告⁷⁾と対応づけると興味深い。すなわち、漢字の部首などの特徴に注目して反応するニューロンがあり、さらにそのニューロンの周りには似たような特徴を持つ漢字に反応するニューロンが集まっていると考えることもできる。われわれが、漢字を識別するとき、いきなり特定の文字を認識するのではなく、まず、部首などの特徴に注目して識別対象候補を絞り込み、それから特定の文字を認識していると思われるが、これは、漢字の部首などの特徴に注目して反応するニューロンが発火することによって大分類を行い、その後に続くニューラルネットによって細分類を行っていると考えることもできる。CombNET はまさにこの形態を具現化したものと考えられ、学習も安定に行うことができることから、極めて合理的なネットワークモデルといえる。

このようにして構築された CombNET-II を用いて JIS 第1, 2水準印刷漢字 6,349 文字の識別実験を行ったところ 99.8% 前後という高い識別率が得られた。

5. 手書き文字認識実験

CombNET-II は大規模なカテゴリを識別するさまざまな問題に適用することができる。次に、手書き文字（現段階では数字）の認識実験を行った結果について示す⁸⁾。手書き文字は、人によって字体が大きく異なり、同一人物の文字でもひとつひとつ変形を伴うため、同じ文字でも多くの異なるパターンがある。CombNET-II

は、このように同一カテゴリに属するパターン内で大きく変動するような問題についてもうまく適応したネットワークを形成することができる。

ここでは、ある程度の文字の変形を吸収するため、2層からなる前処理層を設けた。前処理第1層では各方向ごとに局所的な直線成分を抽出し、第2層ではある程度の位置ずれを吸収するため圧縮を行う。この結果、 8×8 のメッシュ特徴からなる入力データは、256個の8方向局所直線成分特徴量が変換される。この特徴量がCombNET-IIの入力データとなる。

今回、電総研から提供されている文字データベースのうち、ETL 6に含まれている手書き数字を識別対象に実験を行った。図9に正しく認識された文字例を示す。図10に認識率を示す。この図の横軸は学習時に用いた学習データの数(人数)を示し、縦軸は認識率を示している。認識率を求めるときに用いたのは、学習サンプルとは別の750人分のデータ(7,500文字)である。この図に示されているように、学習データ数を増加させると認識率も向上している。すなわち、ニューラルネットが

多くの学習データを学習すればするほど、未知データに対する認識率、すなわち汎化能力が向上している。600人分のデータを学習したときには、未知データに対する認識率が99.4%に達している。

Stem networkの自己増殖型学習則は教師なし学習であるため、入力データの属するカテゴリとは関係なくニューロンの生成と分裂が行われる。このため、手書き文字のように個々のデータのばらつきが非常に大きい場合、同じカテゴリの入力データでも別々のニューロンと最適整合し別々のグループに大分類されることがある。この場合、複数のbranch networkに同じカテゴリに応じた出力層ニューロンを用意する。この結果、変動の大きなカテゴリについては複数の正解ニューロンが存在することになり、パターン変動に強い汎化能力の高いネットワークを構築することができる。CombNET-IIは、このように個々のデータの変動にも柔軟に対処することができるネットワークとなっている。

今回は、手書き数字だけを対象にしたが、今後、アルファベット、ひらがななど識別対象を拡大していく予定であるが、識別対象の拡大はCombNET-IIにおけるstem networkのニューロンが自動的に増加し、それに対応してbranch networkの数も増加させることによって対処できると考えている。

6. 汎用的なパターン認識学習機械としてのCombNET-II

多数のカテゴリを分類するための大規模ニューラルネットワークモデルCombNET-IIについて、その学習方式と印刷文字・手書き文字認識に適用した例を示した。CombNET-IIを用いることによって、BP学習則では困難であった大規模なニューラルネットワークを安定的に構築できる。印刷文字や手書き文字の認識実験において十分に実用化できる認識率を得ることができた。このほか、1,000単語にも及ぶ大語彙の単語音声認識にも適用し、たいへんよい結果を得ている⁹⁾。このようにCombNET-IIは、分類カテゴリ数の多いパターン認識の識別関数の学習方式として、単純な階層型ニューラルネットワークの限界を超える能力をもっており、さまざまなパターンや信号の認識問題に汎用的に用いることができる。

ニューラルネットの各ニューロンにおける演算は、シグモイド関数の計算（これはデーブル参照方式でできる）を除けば、単純な積和演算であり、また各ニューロンの演算は互いに独立で並列に処理することがで

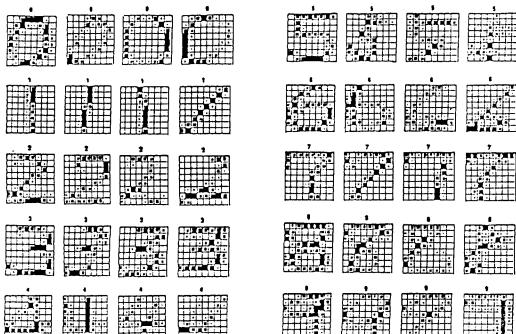


図9 正しく認識された文字例

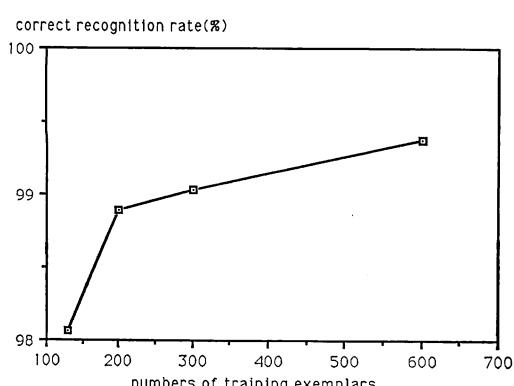


図10 手書き数字認識率

きる。したがって、このような処理機構をもつマルチプロセッサ方式のハードウェア（ニューロコンピュータ、あるいは、ニューロチップ）は比較的単純な構成で実現できる。今後、CombNET-II のハードウェア化についても検討していく予定である。

文 献

- 1) 舟橋賢一：“ニューラルネットワークの capability について”，電子情報通信学会技術研究報告，ME とバイオサイバネティックス，MBE 88-52 (1988) 127-134.
- 2) D. E. Rumelhart, G. E. Hinton and R. J. Williams: “Learning representations by back-propagating errors,” Nature, 323 (1986) 533-536.
- 3) 岩田 彰, 畠麻孝志, 松尾啓志, 鈴村宣夫：“大規模4層ニューラルネット“CombNET””，電子情報通信学会論文誌, J 73-D 2 (1990) 8, 1261-1267.
- 4) T. Kohonen : *Self-organization and Associative Memory* (Springer-Verlag, 1984 and 1988).
- 5) 堀田健一, 岩田 彰, 松尾啓志, 鈴村宣夫：“大規模ニューラルネット CombNET-II”，電子情報通信学会技術研究報告, NC 90-34 (1990) 29-36.
- 6) 堀田健一, 岩田 彰, 諏訪喜久, 鈴村宣夫：“CombNET-II による多数カテゴリの分類”，1991年電子情報通信学会春季全国大会論文集, D-64, p. 6-6-64.
- 7) D. H. Hubel and T. N. Wiesel: “Receptive fields, binocular interaction and functional architecture in the cat's visual cortex,” J. Physiol., 160 (1962) 106-153.
- 8) A. Iwata, H. Kawajiri and N. Suzumura: “Classification of hand-written digits by a large scale neural network “CombNET-II,”” Proc. of Int. Joint Conf. on Neural Networks (IJCNN91-Singapore), 印刷中, 1991.
- 9) 魏 回, 北村 正, 岩田 彰, 鈴村宣夫：“CombNET-II による 1000 単語音声認識”，電子情報通信学会技術研究報告, ニューロコンピューティング NC 90-141 (1991) 175-180.