

## 5 R-4

## パターン認識用ネットからのアルゴリズムの抽出法

阿部重夫 鹿山昌宏 武長寛 北村忠明  
(株)日立製作所 日立研究所

## 1. はじめに

教師データを用いて逆伝播法<sup>1</sup>で学習することにより、自動的にパターン判別ネットを構成できることから、パターン認識問題に多層ネットの適用が進められている。しかしながら多層ネットはブラックボックスとしてしか扱えず、誤認識したときネットの解析が行えないという問題がある。

これに対して、我々は文献2)で、あるクラスに属する教師データが、入力と中間層の間の重みが生成する超平面で他のクラスから単一の領域に分離されるときは、3層で、そうでないときは4層で合成できることを証明した。このように合成しておくことにより、誤認識が生じたときもその解析が可能となった。

本論文では上記の理論を基にして、学習済みのネットワークからパターン判別アルゴリズムを抽出する方式を提案し、次いでこの方式を数字認識の例に適用して、その妥当性を検証する。

## 2. ネットワークの合成

あるクラスに属する入力教師データが  $k$  個の超平面の同じ側にあり、他のクラスの教師データが分割された領域に存在しないとき、このクラスは  $k$  個の超平面により単一領域に分割されるという。もしあるクラスがいくつかの単一領域の和であらわされるとき、このクラスは複数領域に分割されるという。またこのときの超平面を分離超平面ということにする。このとき次の定理が証明できる。<sup>2</sup>

定理1  $n(1)$ 次元のデータを  $n(3)$ 個のクラスに分類するとき、 $n(1)$ 次元の  $n(2)$ 個の超平面の部分集合により 全てのクラスの教師データが単一領域に分離できるとき、 $n(1)$ 個の入力、 $n(2)$ 個の中間層ニューロン、および  $n(3)$ 個の出力ニューロンで構成される3層ネットでパターン分離ネットが合成できる。

定理2 もしあるクラスが複数領域に分離されるとき、3層目で各々の単一領域を分離し、4層目で3層目のニューロンの出力の和をとることにより、4層ネットで認識ネットを合成できる。

出力ニューロン数が  $n(3)$ の3層ネットは4層目に

$n(3)$ 個の出力ニューロンを設定して各々の出力ニューロンを元の出力ニューロンに接続して3層目と同じ出力が出るように重みを調整することにより4層ネットに容易に拡張できる。したがってもし合成されたネットが3層と4層よりなるときは4層ネットにすることができる。このとき得られたネットでは同一のクラスに属する  $z(3)$ は3層目と4層目の重みにより単一領域に分離されている。(ただし、 $z(3)$ は3層目のニューロンの出力ベクトルとする。)したがって、次の系が得られる。

系 パターン分離用ニューラルネットを4層で合成したとき、全ての同一クラスの出力  $z(3)$ が単一領域に分離できるように合成できる。

## 3 分離超平面の抽出

2節の方法で認識ネットが合成できるためには、分離超平面を学習する必要があるが、効率のよい方法が分かっていないため<sup>3</sup>、ここでは逆伝播法により分離平面を学習することにする。

逆伝播法により  $l$ 層 ( $l=3, \text{ or } 4$ )のネットが得られたとする。このとき次のようにして分離平面を求めることができる。

アルゴリズム (i) 教師データ  $x_k(l), k=1, \dots, m$  に対して中間層出力  $z_j(l-1), j=1, \dots, n(l-1)$ を調べる。もし同一クラスに属する教師データに対して  $z_j(l-1) > 1/2$ が成立するとき  $z_j(l-1)$ を1とみ、 $z_j(l-1) < 1/2$ のとき0とみる。そうでないときは0あるいは1とみる。ここで0あるいは1を  $dc$ と表現することにする。同一クラスの  $z_j(l-1)$ が同一の離散化された数字のときは、これらは  $j$ 番目の中間層ニューロンに対応する超平面の同一の側にあり、 $dc$ のときは両方の側にある。もしクラス毎の離散化された  $z(l-1)$ が独立のとき、分離平面が得られたことになる。そうでないときは、(ii)に進む。

(ii) このステップでは分離平面により分離されていない教師データの数が最小になる条件のもとで  $z(l-1)$ が独立となるように  $dc$ を1か0に変える。ここでクラス  $i, j$ にたいする離散化された  $z(l-1)$ が従属とする。クラス  $i, j$ に対する  $z_j(l-1)$ が  $dc$ となるものを探

す。各々の  $k$  に対してクラスあたりの分離されていない教師データの数を計算する。これを  $\text{diff}_k$  とする。ここで両方のクラスの  $z_k(l-1)$  が dc のときの計算法を示す。  $n_i$  をクラス  $i$  の教師データで  $z_k(l-1) > 1/2$  を満足するものの数とすると

$$\text{diff}_k = \min \{ (n_i + m_j - n_k)/2, (m_i - n_i + n_j)/2 \}$$

で与えられる。ただし  $m_i$  はクラス  $i$  に対する教師データの数である。

$\text{diff}_k$  が最小となる  $k$  を求めクラス  $i$  あるいは  $j$  の  $z_k(l-1)$  を変える。このときどちらでも dc で  $(n_i + m_j - n_k)/2$  のほうが小さいときはクラス  $i$  に対する  $z_k(l-1)$  を 1 に変え、クラス  $j$  に対する  $z_k(l-1)$  を 0 に変える。もしそのような  $k$  が存在しないときは分離平面の抽出に失敗したことになる。

もしステップ(ii)を用いて、分離平面を求めたとすると、分離されていない教師データが存在する。またステップ(i)だけによって分離平面が求まったとしても中間層出力は、0.5から1,あるいは0から0.5に散らばっている可能性がある。このため、教師データを分離し、また教師データの中間層出力をある範囲以内に収めるために逆伝播法で求めた重みをチューニングする方法を次に述べる。

#### 4. ネットのチューニング

分離平面を抽出したため重みのチューニングは各層毎に行なうことができる。

それは  $m$  個の  $n$  次教師データを  $x_{ji}$ ,  $i=1, \dots, n$ ,  $j=1, \dots, m$ , 重みを  $w_{11}, \dots, w_{mn}$  とし、最初の  $k$  個の教師出力を  $1-\epsilon$  残りを  $\epsilon$  とすると次の不等式を解けばよい。

$$\sum_{j=1}^{n+1} w_j x_{ji} \geq \alpha \quad i=1, \dots, k \quad (3)$$

$$\sum_{j=1}^{n+1} w_j x_{ji} \leq -\alpha \quad i=k+1, \dots, m \quad (4)$$

ただし  $\alpha = -T \log(1/(1-2\epsilon)-1) = T \log(1/2\epsilon - 1) > 0$  である。上記は2層のBPを用いることにより解くことができる。

#### 4. 数字認識の例

12特徴量入力、10出力の数字認識ネットに対してアルゴリズムの抽出及びチューニングの効果を調べた。

教師データとして300のデータを選び、テストデータとし1330のデータをとった。中間層のニューロン数を4から10に変え、3層で評価した。-0.5から0.5に分布する乱数で重みを初期化して、同一構成で

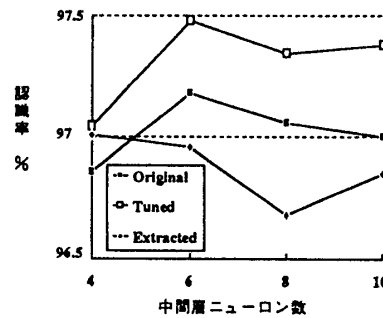


図1 1330テストデータの平均認識率  
(100回試行、3層、300教師データ、 $\epsilon=0.1$ )

100回学習した。その後各々のネットから分離平面を抽出して、重みのチューニングを入力側から  $\epsilon=0.1, 0.01$  として行なった。テストデータに対する認識率は、もともとのネット、チューニングされたネット、抽出されたアルゴリズムに対して行なった。判定は、ニューロンの入力側で0出力側では0.5で行なった。

$\epsilon=0.1$  で80,000回の学習では収束しなかったので、1000回で打切ることにした。このときの最大認識率は99.67%で1つの教師データが分離できなかった。平均の教師データの認識率は99.14%から99.66%まで変化し、中間層ニューロン数が増えるほど認識率が上がった。4及び6中間層ニューロンのときは3者の認識率はほぼ同じであったが、8及び10中間層ニューロンのときはもとのネットの認識率が0.2%から0.3%良かった。

図1にテストデータの認識率を示す。図1において、チューニングされたネットの認識率が最も良く、最大の認識率は8中間層ニューロンのときで99.02%となった。4中間層ニューロンのとき抽出アルゴリズムの認識率がもとのネットの認識率より良くなった。100回の試行のうち、55回抽出アルゴリズムがもとのネットより認識率がよかった。

#### 5. おわりに

パターン認識用ニューラルネットの構成理論に基づいてニューラルネットから認識アルゴリズムを抽出する方式を開発した。これを数字認識システムで評価して、アルゴリズムが抽出できることを示した。

#### 参考文献

1. D.E. Rumelhart et al., "Parallel Distributed Processing", Vols. 1, 2, MIT Press, Cambridge, Mass.
2. S. Abe, M. Kayama, H. Takenaga, "How Neural Networks for Pattern Recognition Can Be Synthesized," Journal of Information Processing, Vol. 14, NO. 3, 1991.
3. A.G.アルカデーエフ、電子計算機とパターン認識、ラティス社、昭和43年6月。