

## 大規模ニューラルネット “CombNET-II”

准 員 堀田 健一<sup>†</sup>      正 員 岩田 彰<sup>†</sup>

正 員 松尾 啓志<sup>†</sup>      正 員 鈴木 宣夫<sup>†</sup>

### A Large Scale Neural Network “CombNET-II”

Ken'ichi HOTTA<sup>†</sup> Associate Member, Akira IWATA<sup>†</sup>, Hiroshi MATSUO<sup>†</sup>  
and Nobuo SUZUMURA<sup>†</sup> Members

あらまし 一般にニューラルネットが大規模になると、結合数が急速に増加し、また学習の収束が困難になるなどさまざまな問題が生じる。これを改善するために我々は、ベクトル量子化ニューラルネットと階層型ニューラルネットを組み合わせた CombNET と名づけたネットワークモデルを提案している。しかし、前段のベクトル量子化ニューラルネットにおけるカテゴリーの大分類が均等にならないという問題が生じた。本論文ではこの問題を解決する手法として、自己増殖型ニューラルネットと呼ぶベクトル量子化ニューラルネットを考案し、このネットワークを適用した CombNET-II を提案する。このネットワークの利点は二つある。まず第1に、前段で各グループに属するカテゴリーがほぼ均等になるように分割されることである。また第2には、分類するカテゴリーの増減にニューロンの数を柔軟に対応させて変化する点にある。今回はこの CombNET-II を JIS 第1, 2水準印刷漢字の識別に適用し、前段における大分類の様子、カテゴリー数の変動に対する能力および識別率の検証を行った。

キーワード：ニューラルネットワーク、文字認識、ベクトル量子化、バックプロパゲーション

#### 1. まえがき

現在までに、ニューラルネットを用いた文字認識の研究がさまざまなされてきており、英数字等の文字認識においては、手書き文字に対してまで非常に高い認識率を示している。しかし、英数字等の数十から数百のカテゴリーにおけるニューラルネットの有効性は示されているものの、日本語にある漢字のように多数のカテゴリーを識別するような問題に対しては、ネットワークが非常に大規模になることから、学習時にローカルミニマムに陥るなど収束性が悪く、また、計算量も膨大なものになる。そこで、ニューラルネットをいくつかの小規模なニューラルネットのモジュールに分けることによって、この二つの問題点を解決しようとする研究がなされてきている<sup>(1)</sup>。我々は、このようなモジュール構造をもつ大規模なニューラルネットの構築手法として CombNET を提案している<sup>(2)</sup>。

CombNET は、第1層にベクトル量子化型ニューラルネット (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) を配置した楕型の構成をしている。CombNET では、Stem Network の形成を Kohonen の自己組織化学習則<sup>(3)</sup>により行い、Branch Network の学習は、バックプロパゲーション法<sup>(4)</sup>を用いて行っている。

ここで問題となることは、Stem Network の各ニューロンと最適整合となるカテゴリーの数が1~100個と大きくばらつくことである。大分類されるカテゴリー数の大小は、後段の Branch Network の大きさ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、識別時に出力ニューロンの発火度にばらつきを生じさせる。ここで言う Branch Network の学習深度とは、学習時の学習パターンを入力したときの出力層の出力値と教師データとの誤差の収束度合と、学習されたネットワークのパターン識別汎化能力の両方を加味したものである。誤差の大きさと識別汎化能力は、誤差を小さくするようにネットワー

<sup>†</sup> 名古屋工業大学電気情報工学科, 名古屋市  
Department of Electrical and Computer Engineering, Nagoya  
Institute of Technology, Nagoya-shi, 466 Japan

クの学習を繰り返し行うほど過学習になり、識別汎化能力が低下する一般的に相反する関係にある。これまで、ネットワークの規模に応じた初期値、中間層の数、中間層ニューロンの数を最適にするさまざまな研究がなされてきたが、ネットワークの規模が極端に異なるものを同程度に学習を行うことは難しい。特にネットワークの規模が非常に大きくなるほど、ローカルミニマムに陥る可能性が高くなり、十分な識別を行えるネットワークを構築するには多くの学習回数を必要とする。すると、識別汎化能力の低下を招く。そのため学習深度をそろえるためには、ネットワーク規模をそろえることが最も効果的な手段であると言える。

また、多次元空間のベクトル量子化という観点から考えると、最適な量子化とは、限られた量子化レベル数(ニューロン数)において、量子化誤差を最小にすることである。量子化誤差を最小にするには、カテゴリーが多次元空間内で一様に分布している場合には一様量子化でよく、参照ベクトルを空間的に均等に配置すればよい。しかし、一般にはカテゴリーの分布には偏りがあるため、一様量子化は最適な量子化とはならない。このようなときには、カテゴリーの分布を一様分布に変換する写像を行った上で、一様量子化を行うことになる。写像された空間で一様に量子化をすることは、原空間ではカテゴリーの分布密度に応じて参照ベクトルを配置し、それにより分割される部分空間ごとのカテゴリーの分布密度を一様にするに対応している。すなわち、各参照ベクトルと最適整合となるカテゴリーの数が均等になるように参照ベクトルを配置したとき、最適なベクトル量子化が形成されることになる。

このような観点から本論文では、カテゴリーの分布密度に従ってニューロンの参照ベクトルを生成・配置する学習則を用いた自己増殖型ニューラルネットを提案する。そして、このネットワークを CombNET の Stem Network の学習則に取り入れ、CombNET-II と名づけた。このネットワークによって数千字種に及ぶ印刷漢字の識別実験を行い、多数のカテゴリーを識別する問題に対する有効性と、カテゴリー数の変動に対するネットワーク規模の適応能力について確かめた。

## 2. 自己増殖型ニューラルネット

ベクトル量子化とは多次元ベクトル空間内のデータを有限の参照ベクトルで代表させ量子化することである。その結果、原データは、参照ベクトルの数に等しいクラスタに分割される。クラスタリングの手法とし

ては、LBG アルゴリズム<sup>(6)</sup>がよく知られているが、この方法では、参照ベクトルの数をあらかじめ限定する必要がある。しかし実際には、必ずしもカテゴリーの数や特徴空間における分布などの状態が最初から把握できるとは限らず、はじめに決定する参照ベクトルの数が、特徴空間を量子化するのに最適な数であるように決定することが非常に困難な場合がある。もし、参照ベクトルの数が多い場合にはカテゴリーの特徴を全く反映しない参照ベクトルが生成される可能性があり(量子化効率の低下)、また少ない場合には、多くのカテゴリーの特徴を一つの参照ベクトルで表現しなければならないものが生成され、そのグループの参照ベクトルがそのグループの特徴を的確に反映しない可能性がある(量子化誤差の増大)。一方、参照ベクトルの数をデータの分布状態により対応させる方法として ISODATA<sup>(6)</sup>がある。しかし、この方法においても前章で述べた各クラスタのカテゴリーの数をそろえるという問題を解決することはできない。そこで、局所的で単純な少ない繰返し計算だけでこれらの二つの問題を解決する手法として自己増殖型ニューラルネットを考案した。この量子化の手法は、はじめから参照ベクトルの数を定めるのではなく、初期状態においては

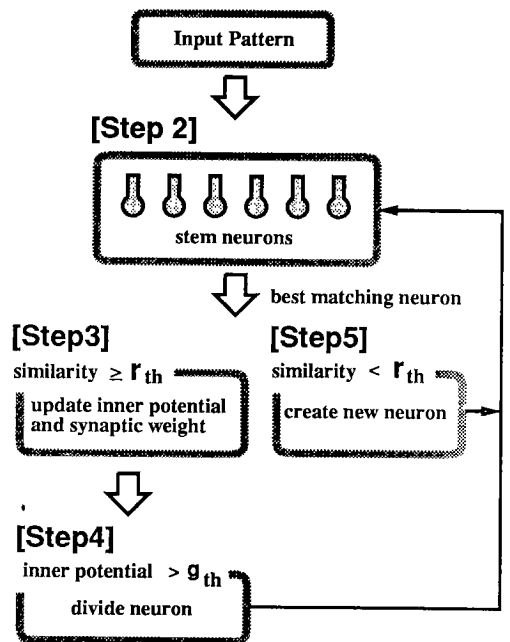


図1 自己増殖型ニューラルネットの学習過程  
Fig. 1 The self-growing neural network learning procedure.

全く何も無い状態にし、その状態から入力ベクトルの投入に応じて参照ベクトルを生成する考え方を基本にした学習則である。

自己増殖型ニューラルネットの学習には、第1過程と第2過程が存在する(図1参照)。まず第1過程では、初期状態においてニューロンは存在しない。そして、1番目の入力ベクトルをそのまま参照ベクトルとするニューロンを生成する。2番目以降の入力ベクトルは、そのときに存在しているすべてのニューロンの参照ベクトルの中で、最適整合となるニューロンを見つける。次にそのニューロンの参照ベクトルをHebb則に準拠して入力ベクトルの方向に修正する。各ニューロンには内部ポテンシャルとして、最適整合した入力ベクトルの数を記憶しておき、それがあるしきい値を超えたときに、そのニューロンを二つに分裂させる。分裂方法は、特徴空間において分裂前の参照ベクトルを通る超平面を生成し、そのニューロンと最適整合していた入力ベクトル群を2分する。そして、2分された入力ベクトル群の平均ベクトルを新しいニューロンの参照ベクトルとし、内部ポテンシャルも各々のニューロンに属する入力ベクトル数に変更する。この過程を入力ベクトルのすべてについて行う。

第2過程では、入力ベクトルの中には、第1過程でニューロンが分裂を起こしたことによって、最適整合ニューロンが以前のものとは異なるものが存在する。このような入力ベクトルに対し、以前の最適整合ニューロンの内部ポテンシャルを一つ減少させ、新しい最適整合ニューロンの内部ポテンシャルを一つ増加させる。また、更に各々の参照ベクトルも修正する。このとき、内部ポテンシャルがあるしきい値を超えれば、第1過程と同じくニューロンは分裂を起こす。この過程を内部ポテンシャルの変動がなくなるまで繰り返す。

このような自己増殖型学習則は以下のように定式化される。

ここで変数を、

$n$ : 入力ベクトルの次元数

$K$ : 入力ベクトルの個数

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

$k$  番目の入力ベクトル ( $k=1, 2, 3, \dots, K$ )

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

ニューロン  $i$  の参照ベクトル

$g_i$ : ニューロン  $i$  の内部ポテンシャル

$g_{th}$ : 分裂を起こす内部ポテンシャルのしきい値

$r_{th}$ : 新しいニューロンを生成する時の整合度のしき

い値

$v$ : 生成されたニューロンの数

と定義する。

<第1過程>

[Step 0] 初期状態は、一つもニューロンは存在しない。

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

$$\mathbf{m}_1 = \mathbf{x}_1 \quad (1)$$

$$g_1 = 1 \quad (2)$$

$$v = 1 \quad (3)$$

各入力ベクトル  $\mathbf{x}_k$  について順次 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。

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

(ニューロン  $c$ : 最適整合ニューロン)

ここで、整合度は式(4)のように入力ベクトルと参照ベクトルの相互相関によって求める。

$$\text{sim}(\mathbf{x}_k, \mathbf{m}_i) = \frac{\sum_j (x_{kj} \times m_{ij})}{\sqrt{\left(\sum_j x_{kj}^2\right) \left(\sum_j m_{ij}^2\right)}} \quad (4)$$

もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_c) \geq r_{th}$  のときは、Step 3へ進む。もし、 $\text{sim}(\mathbf{x}_k, \mathbf{m}_c) < r_{th}$  ならば、Step 5へ進む。

[Step 3] 内部ポテンシャルを更新し、 $g_{c(\text{new})} = g_{c(\text{old})} + 1$  とする。次式により参照ベクトルの修正を行う。

$$\mathbf{m}_{c(\text{new})} = \mathbf{m}_{c(\text{old})} + \frac{1}{g_{c(\text{new})}} (\mathbf{x}_k - \mathbf{m}_{c(\text{old})}) \quad (5)$$

これは、新しい参照ベクトルを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、 $g_{c(\text{new})} > g_{th}$ 、すなわち、分裂する内部ポテンシャルに達していたら、Step 4を行う。 $g_{c(\text{new})} \leq g_{th}$  の場合は、Step 2に戻る。

[Step 4] 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。

$$\mathbf{a} \cdot (\mathbf{x} - \mathbf{m}_c) + a_0 = 0 \quad (6)$$

$\mathbf{a}$ 、 $a_0$  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  $\mathbf{x}c(x_{c1}, x_{c2}, \dots, x_{cn}, \dots, x_{c(n_{\text{new}})})$  を2分する。但し、この2分された入力ベクトルの数に偏りがある場合は、2分された入力ベクトル群のカテゴリー数が、各々

$$\begin{cases} g_{upper}, (\mathbf{a} \cdot (\mathbf{x}_{ck} - \mathbf{m}_{c(oid)}) + a_0 \geq 0) \\ g_{lower}, (\mathbf{a} \cdot (\mathbf{x}_{ck} - \mathbf{m}_{c(oid)}) + a_0 < 0) \end{cases}$$

(但し,  $g_{upper} + g_{lower} = g_{c(new)}$ )

とすると,

$$|g_{upper} - g_{lower}| \leq 1 \quad (7)$$

になるまでパラメータ  $\mathbf{a}$ ,  $a_0$  を変えて分裂を繰り返す行う。

新しい参照ベクトルは, 2分された各々の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。

$$\mathbf{m}_{c(new)} = \frac{1}{g_{upper}} \sum_j \mathbf{x}_{cj} \quad (8)$$

$$m_{v+1} = \frac{1}{g_{lower}} \sum_k x_{ck} \quad (9)$$

それぞれのニューロンの内部ポテンシャルは,

$$g_{c(new)} = g_{upper} \quad (10)$$

$$g_{v+1(new)} = g_{lower} \quad (11)$$

となり, ニューロンの個数は,

$$v_{(new)} = v_{(oid)} + 1 \quad (12)$$

となる。Step 4 終了, Step 2 へ戻る。

[Step 5] 新しくニューロンを生成する。

$$\mathbf{m}_{v+1} = \mathbf{x}_k \quad (13)$$

$$g_{v+1} = 1 \quad (14)$$

$$v_{(new)} = v_{(oid)} + 1 \quad (15)$$

これは, 整合度がしきい値  $r_{th}$  以上になるニューロンが見つからない場合に, 入力ベクトルを参照ベクトルとするニューロンを新しく生成することを示している。

Step 5 終了, Step 2 へ戻る。

<第 2 過程>

各入力ベクトル  $\mathbf{x}_k$  について順次 Step 2', Step 3', Step 4, Step 5 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。

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

(ニューロン  $c'$ : 最適整合ニューロン)

ここで, 整合度は式(4)によって求める。

もし,  $\text{sim}(\mathbf{x}_k, \mathbf{m}_{c'}) \geq r_{th}$  ならば, Step 3' へ進む。もし,  $\text{sim}(\mathbf{x}_k, \mathbf{m}_{c'}) < r_{th}$  ならば, Step 5 へ進む。

[Step 3'] 以前の最適整合ニューロン  $c$  と新たな最適整合ニューロン  $c'$  とが異なった場合,

$$g_{c'(new)} = g_{c(oid)} + 1 \quad (16)$$

$$g_{c(new)} = g_{c(oid)} - 1 \quad (17)$$

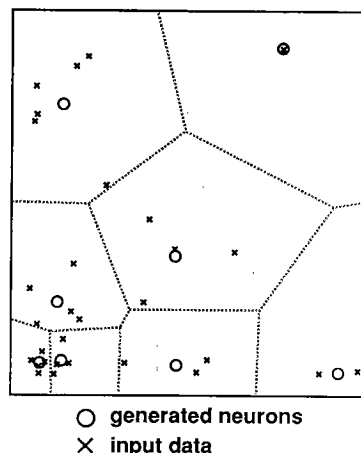


図2 2次元空間上で自己増殖型学習則によって生成された量子化ニューロン

Fig. 2 Vector quantizing neurons generated by the self-growing learning rule.

として内部ポテンシャルを更新し, 次式により参照ベクトルの修正を行う。

$$\mathbf{m}_{c'(new)} = \mathbf{m}_{c'(oid)} + \frac{1}{g_{c'(new)}} (\mathbf{x}_i - \mathbf{m}_{c'(oid)}) \quad (18)$$

$$\mathbf{m}_{c(new)} = \mathbf{m}_{c(oid)} - \frac{1}{g_{c(new)}} (\mathbf{x}_i - \mathbf{m}_{c(oid)}) \quad (19)$$

もし,  $g_{c'(new)} > g_{th}$ , すなわち, 分裂する内部ポテンシャルに達していたら, Step 4 を行う。  $g_{c'(new)} \leq g_{th}$  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。

この手法に基づいて生成された参照ベクトルは, 入力ベクトルの統計的性質を反映した値となる。そして, 各ニューロンを最適整合とするカテゴリーの数も, 分裂の内部ポテンシャルのしきい値  $g_{th}$  により上限が制限される。

また, 特徴空間でカテゴリーの分布密度が高い部分に多く参照ベクトルが生成される。このことにより, カテゴリーの分布状況にあった準最適なボロノイ分割を行う参照ベクトルの生成が可能となる。特徴も各パターンに共通の特徴を的確に表現することができる。

実際に 2次元平面におけるシミュレーションを行うと図 2 のようになる。これは,  $\times$  で表される入力ベクトルに対して, 自己増殖型学習則によって  $\circ$  で表した参照ベクトルをもつニューロンが生成された様子を示す。点線は識別境界を示す。分裂のしきい値  $g_{th}$  は 6 とし, この空間の一辺を 1.0 としたときに入力ベクトル

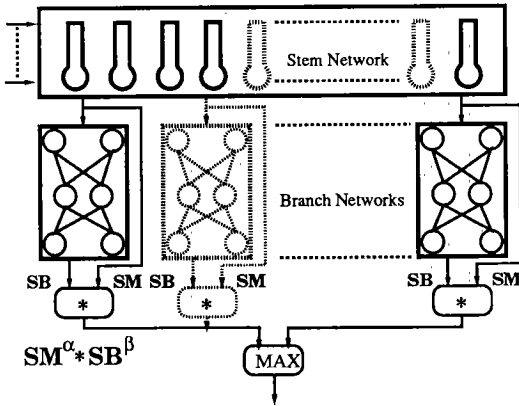


図3 CombNET-IIの構成図  
Fig. 3 The network structure of CombNET-II.

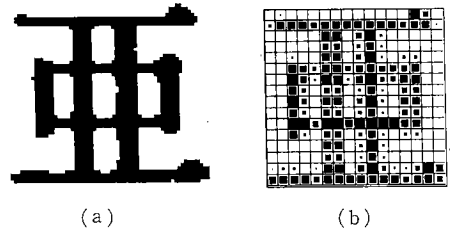
と参照ベクトルの距離が0.6以上のものは別のニューロンを生成させるという条件で行った。この図を見てもわかるように左下のカテゴリーの分布密度の高い部分には多くの参照ベクトルが生成され、それ以外の分布密度の低い部分には参照ベクトル同士の距離が比較的離れており、期待する結果が得られていることがわかる。

### 3. 大規模ニューラルネット CombNET-II

このような自己増殖型ベクトル量子化ニューラルネットをCombNETの第1層(Stem Network)の部分に用いた新しい大規模ニューラルネット“CombNET-II”を構築した。

CombNET-IIは、従来のCombNETと同じく、前段の第1層(Stem Network)にカテゴリーをいくつかのグループに大分類するベクトル量子化ニューラルネットを配置し、後段の第2, 3, 4層(Branch Network)には小グループ内のデータを分類する小規模な階層型ニューラルネットを多数並列に配置した図3のような櫛型をしたネットワークである。

CombNET-IIの学習は、まずStem Networkにおける量子化ニューロンの生成から行う。生成の手順は、前章で述べたとおりの手法で行う。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ーを代表する共通パターンが成長する。Stem Networkの学習の後、カテゴリーがStem Networkのどのニューロンと最適整合になるかを調べ、各ニューロンの分担当すべきカテゴリーを求めて、入力ベクトルをStem Networkのニューロン数と同数のグループに分割する。次に、その分割されたグループごとにその中にあるカ



(a) Binary image taken by an image scanner.  
(b) Gray level image composed of 16x16 pixels.

図4 印刷漢字データの一例  
Fig. 4 An example of printed kanji character data set.

テゴリー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、各カテゴリーがどのニューロンと最適整合となるかを求める。次にStem Networkの最適整合となったニューロンを担当するBranch Networkに入力ベクトルを入力して識別を行う。

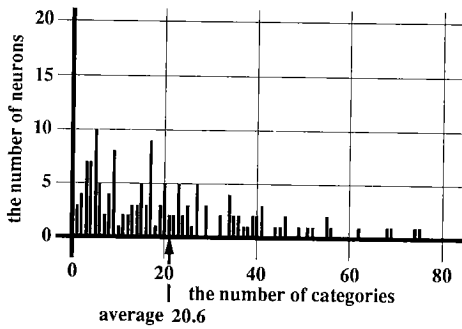
また、CombNET-IIのStem Networkは汎化能力が弱く、パターンが変動により最適整合ニューロンが変わりやすい。一方、Branch Networkは比較的パターン変動に強い性質をもつ。そこで図3に示すとおり、第1層と第4層の出力を結合したニューラルネットを構成した。すなわち、第1層に入力パターンを通し、適合度の高いものから数番目までのニューロンを選び、次に選ばれたニューロンを担当する階層型ネットワークに入力ベクトルを通し、出力層の中で最も高い発火レベルになったものの出力値を調べる。そして、(整合度)<sup>(出力値)</sup>

の値が最も高くなったものを識別結果として選ぶ。

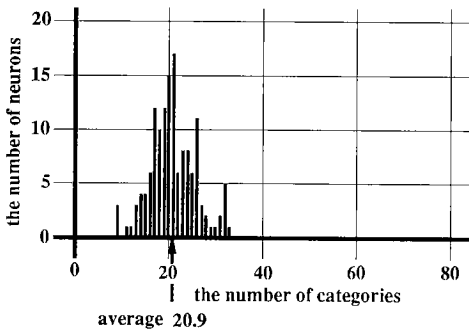
このCombNET-IIの最大の特徴は、Stem Networkにおいて、ほぼ均等なカテゴリー分割を行うことができる点にある。これによって、各Branch Networkの出力層のニューロン数がほぼ均等になり、各Branch Networkの規模もほぼ均等にすることができる。このため、各Branch Networkの学習深度がそろうことになり、CombNETにおいて問題になっていた、Branch Networkの学習の不均一性をなくし、バランスのとれた階層型ニューラルネットの出力値を得ることができる。また、識別するカテゴリーの数に応じてStem Networkのニューロン数が変化するので、CombNET-IIのネットワーク規模が適応的に変化し、最適なネットワークを自然に構築することができる。

表1 しきい値と生成ニューロン数の比較

$g_{th}$	JIS 第1水準のみ (2965文字)				JIS 第1,2水準 (6349文字)			
	生成ニューロン数	最適整合カテゴリ数			生成ニューロン数	最適整合カテゴリ数		
		最小	最大	平均		最小	最大	平均
10	456	3	10	6.50	973	3	10	6.52
15	319	4	15	9.29	661	3	15	9.60
20	240	5	20	12.35	511	4	20	12.42
25	192	7	25	15.44	414	6	25	15.33
30	159	8	30	18.64	338	6	30	18.78
35	142	9	33	20.88	299	8	35	21.23
40	125	7	40	23.72	265	10	39	23.95
60	83	16	60	35.72	173	16	60	36.70



(a) By Kohonen's self-organized feature map (144 neurons)



(b) By self-growing learning rule ( $g_{th}=35$ , 142 neurons)

図5 最適整合したカテゴリ数ごとのニューロン数の分布  
Fig. 5 The number of neurons vs. the number of categories to be classified to the same neuron.

#### 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検証するために、JIS 第1水準 2,965字種と JIS 第1, 2水準 6,349字種の印刷漢字を識別するニューラルネットの構築を試みた。

まず、今回の実験に用いたデータについて述べる。



図6 自己増殖型学習則によって生成されたニューロンの参照ベクトル ( $g_{th}=35$ )  
Fig. 6 The synaptic weight vectors of the stem neurons ( $g_{th}=35$ , 142 neurons).

学習データの作成は、次のようにして作成した。JIS 第1, 2水準漢字を印刷したもの解像度が 400 dot/inch のイメージスキャナで画像データとして取り込む。そして、イメージスキャナで取り込んだ画像をもとに、縦横方向に画素値を積分した濃度プロファイルから、しきい値処理により一文字一文字を切り出す。その結果、64ドット×64ドット程度の2値画像となって得られる(図4(a))。これを16×16の小領域にまとめて、濃度正規化したパターンを作成する(図4(b))。

ここで、識別能力を評価する前にまず、Stem Network のカテゴリ分割の様子を以下に示す。今回は、2.で示した学習則のうち Step 5の部分における整合度のしきい値  $r_{th}$  を 0.0 とした。すなわち、新たなニューロンを発生させる要因としては、ニューロンの内部ポテンシャルが分裂のしきい値を超えたときのみである。そして、分裂のしきい値  $g_{th}$  が 10, 15, 20, 25, 30, 35, 40, 60 の8例について行った。

表1は、しきい値  $g_{th}$  の値の変化による生成ニューロン数を示すものである。この表を見ると分割のしきい値  $g_{th}$  を小さい値にすると、生成されるニューロン数は増えている様子がわかる。第1水準の場合と第1, 2水準の場合を比較すると、入力カテゴリ数の増加を生成するニューロン数を増加させることによって適応的に対応している様子がわかる。

図5は、同じニューロンに最適整合するカテゴリ数によるニューロンの数の分布を Kohonen の自己組織化学習則(ニューロン数 144 個)によるものと、本論文

表2 大分類された漢字の一覧表(一部分)

No.	漢字種
0	聿盈益蓋監基宜血些左盛盜盤並豊益盟孟翌藍靈
1	卮丑筮釜球五互差三且昼登豆埋理立厘歷互亘
2	芎籜卷鼓查量壺韭魁
3	及袈兼驗姿裁竣駁裳製賤裝駱袋駐翕裂
4	雅雁股腰腸腫収腺膳脱胆膠脹腸朕膳臘版服腹膜脈履脇
5	肝既肱腔肢勝髓髓胎貯胎腦脾肺肌肥膝赴胞肪腕
6	或威域戒滅裁歲殘式城成威踐底式八武滅
7	饜穀穀饒教放最坂較汐放親政設穀殿奴數數般敏放
8	右雲翁霞看希喬禽君肴鶩煮需署省震雀雪占霜著布零露
9	安衣亥曳央去穴玄交宏史支尖失丈先尖二夫文方矢夕夕
10	応岩広寺赤走宅定店宥府庖有劣老

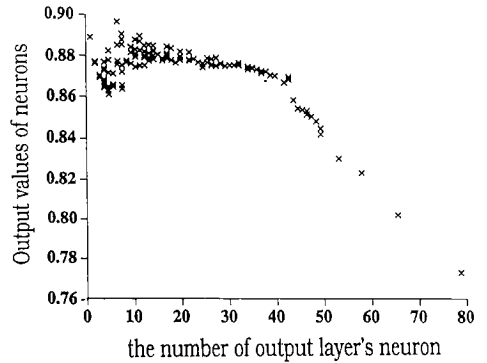
で述べている自己増殖型学習則 ( $g_{th}=35$ , ニューロン数142個) のものを比較して示したものである。各々の分布の標準偏差を求めてみると, Kohonen モデルでは16.6, 自己増殖型モデルでは4.9と自己増殖型では大きさがそろっていることがわかる。特に Kohonen モデルでは平均の数以下のものが多く, 自己増殖型では平均を中心とした分布になっており対照的である。

表1, 図5から, カテゴリー数の上限は分裂のしきい値  $g_{th}$  を下回り, 下限も  $1/2 g_{th}$  より著しく小さくなっていないことがわかる。また, カテゴリー数がしきい値の  $1/2$  を割っているものが存在することは, ランダムに生成した超平面によって入力ベクトル群を2分割し, その後2分割された各々のグループで新しい参照ベクトルを生成するが, 新しく最適整合となる入力ベクトル群と分裂前に最適整合となる入力ベクトル群が異なるために起こるものである。

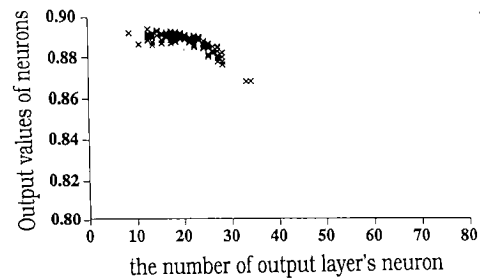
図6は, 自己増殖型学習則に先ほど述べた印刷漢字データを投入することによって生成された量子化ニューロンの参照ベクトルの値を  $16 \times 16$  の2次元マトリックス状のパターンで示したものである。参照ベクトルには, 漢字の「偏」, 「つくり」, 「かまえ」が形成されている。それ以外に, 漢字の「偏」, 「つくり」, 「かまえ」以外の特徴も参照ベクトルに表現されているものが見られる。また, 「さんずい偏」や「木偏」の偏を持つ漢字の数が多いためこれらの特徴を反映した参照ベクトルが多く生成されていることがわかる。

表2に各ニューロンが最適整合となる漢字の一覧表の一部を示す。このようにすべての漢字がその「偏」, 「つくり」, 「かまえ」などの特徴によって分類されている。

カテゴリーの分布を反映したニューロンの生成を行っても, 常に参照ベクトルと低い整合度を示すものが存



(a) By Kohonen's self-organized feature map (144 neurons)



(b) By Self-growing learning rule ( $g_{th}=35$ , 142 neurons)

図7 学習データを投入したときの Branch Network の出力値

Fig. 7 Output values of neurons in branch networks for recognizing the learning data.

在する場合がある。その例は, 「一」, 「乙」, 「心」などである。これらは, 他のカテゴリーと共通した特徴に乏しく, 参照ベクトルとの整合度も小さい。このような例については, 最適整合となる整合度のしきい値  $r_{th}$  を適当な値に設定することによって, このような特異的なカテゴリーに対して高い整合度を示すニューロンを生成させることができる。

図7は, このような自己増殖型学習則を用いてカテゴリーを大分類し, 後段の Branch Network の学習を行った後, 学習データをこのネットワークに入れて Branch Network のニューロンの出力値を示したものである。横軸は, Branch Network の出力層の大きさであり, 縦軸は, その大きさの Branch Network 平均の出力値を示した。今回の Branch Network の学習方法は, 各カテゴリーに対して出力層の発火ニューロンを一つずつ決め, その値を0.9それ以外のニューロンを0.1で発火するように教師パターンを与える一般的にパターン認識問題で用いられている方法である。そのため,

表3 CombNET-IIによる識別結果

(識別率%)

$g_{th}$	第1水準 (2965 文字)	第1,2水準 (6349 文字)
10	99.84	99.25
15	99.85	99.81
20	99.83	99.81
25	99.77	99.74
30	99.67	99.76
35	99.66	99.60
40	99.71	99.49
60	99.08	99.03

学習データを投入した場合には、0.9 付近で発火することが理想である。しかし、現実には Branch Network の規模が大きくなるに従って学習が困難になりそのため学習深度に差が出て出力値が下がってしまう。CombNET では、規模の違う Branch Network があるということは図 7 (a) のように Branch Network の出力値に差が出てくるという現象を起こすことになる。CombNET のように多数のネットワークが並列に配置される場合、これらのネットワークの学習深度に差があるということは、識別時において悪い影響を及ぼすことになる。ところが、図 7 (b) のように自己増殖型モデルを用いることにより、Branch Network の規模をそろえることが可能になり、それが Branch Network の学習深度をほぼ同程度にすることになり、これらの問題を解決することができる。

最後にこの自己増殖型学習則を用いた CombNET-II の識別能力を評価する。まず、CombNET との識別能力を評価する。CombNET のときに用いた印刷漢字データを 5 セット用意し、4 セットの平均パターンを学習データとし、残り 1 セットを識別用データとする。そして、Branch Network の選択候補数を 3 とし、CombNET のとき (第 1 水準のみ、ニューロン数 144 個) と同程度のニューロン数・結合数の場合 ( $g_{th}=35$ , 生成ニューロン数 142 個) において比較を行うと識別率の向上 (99.49% → 99.76%) が見られた。更に今回は、識別率のデータ依存性をなくすためにデータを 3 セット増やして合計 8 セット用意し、そのうち 4 セットの平均パターンを用いて学習を行い、残り 4 セットを識別実験用データとして用いた。そして、それらの学習用データと識別用データをさまざまに組合せを変えて実験を行い識別能力を評価した。

その結果を表 3 に示す。識別率は高いもので、JIS 第 1 水準のみ、第 1, 2 水準ともに 99.8% を超えた。ま

た、整合度と出力値を乗ずる際の重みの比は、 $a/\beta=2$  付近で高い識別率を示した。CombNET においては、 $a/\beta=5$  という値であったが、この値の違いは、CombNET では、Branch Network の学習深度の差があったのに対して、CombNET-II では、各々の Branch Network の学習深度がそろい、出力値に対し同等の評価をすることが可能になったため、Branch Network の出力値 SB をより強く最終出力値に反映させることになったのであろう。このことは、1 層と 4 層の結合をすることによる効果がより発揮されていることを意味するものである。

以上の結果より、多数のカテゴリーを均等分割し、後段の Branch Network の規模をそろえることで効率の良いネットワークを構成することができ、同じニューロン数・結合数という条件のもとでも CombNET-II は、CombNET より高い識別率を示すことがわかった。

## 5. むすび

我々がこれまで多数のカテゴリーを分類する手段として提案してきた大規模ニューラルネットワーク CombNET において、前段の Stem Network の学習則に Kohonen の自己組織化学習則を用いたために、大分類を行う際に均等に分けることができず、大きさのばらつきを生じていた。このばらつきは、ベクトル量子化の観点から見ても量子化誤差は大きくなり、更に Branch Network の規模に差を生じ、これが学習深度をばらつかせ、後段に多数並列にネットワークが配置されているという CombNET の構造的特徴に悪影響を及ぼしていた。この問題を解決する方法として、自己増殖型ニューラルネットワークという多数のカテゴリーを均等分割するベクトル量子化の学習則を提案した。一つのニューロンと最適整合するカテゴリーの数があるしきい値を超えたときに、そのカテゴリーを 2 分割して新たなニューロンを生成する手法を基本としたものである。本手法を用いると一つのニューロンと最適整合するカテゴリーの数を制限すると共に大分類されるデータ数のばらつきも非常に小さくすることができる。また、一般的なベクトル量子化の方法で用いられている、参照ベクトルをはじめから特徴空間上に適当に配置する方法ではなく、全く何も無い初期状態に次々とデータを投入し、投入される入力ベクトルによって必要となる数だけ参照ベクトルを生成することから、カテゴリーの増減に柔軟に対応し、効率的な量子化を行うことができる。この学習則を用いて、印刷漢字のパター



ンを投入すると、データの投入による内部ポテンシャルの上昇をきっかけに分裂するという単純な繰返しだけで、漢字のもつ「偏」、「つくり」等の共通の特徴が参照ベクトルに形成されていった。このベクトル量子化ニューラルネットワークを Stem Network として用いた CombNET-II の印刷漢字の識別実験において、各 Branch Network の学習深度がそろっている様子を確認し、99.8% を超える高い識別率をカテゴリ数増減にも影響されることなく示した。すべてのネットワークが全体としてバランス良く識別に寄与していると考えられる。この CombNET-II は、ニューラルネットワークにおいて多数のカテゴリを分類する手段の一つとして非常に有効なものと言える。

また、ネットワークは変動の大きなパターンの識別問題を新たなニューロンを生成する方法で解決することが可能であると考えられる。今後、手書き文字認識や音声認識などの問題に CombNET-II を適用し、その識別能力の検証を行う予定である。

#### 文 献

- (1) 森 吉弘：“大規模ニューラルネットワークの構築に向けて”，信学技報，PRU88-59 (1988)。
- (2) 岩田 彰，當麻孝志，松尾啓志，鈴村宣夫：“大規模4層ニューラルネットワーク“CombNET”，信学論(D-II)，J73-D-II，8，pp. 1261-1267 (1990-08)。
- (3) Kohonen T.：“Self-Organization and Associative Memory”，Springer-Verlag (1984 and 1988)。
- (4) Rumelhart, D. E., Hinton, G. E. and Williams, R. J.：“Learning Representations by Backpropagating Errors”，Nature 323, 9, pp. 533-536 (1986)。
- (5) Linde Y., Buzo A. and Gray R. M.：“An Algorithm for Vector Quantizer Design”，IEEE Trans. Commun., COM-28, 1, pp. 84-95 (1980)。
- (6) G. H. Ball and D. J. Hall：“ISODATA-Novel Method of Data Analysis and Pattern Classification”，Stanford Research Institute (1965)。

(平成3年6月26日受付，10月16日再受付)



岩田 彰

昭48名大・工・電気卒。昭50同大大学院修士課程了。同年名工大・情報・助手。昭57年4月より昭58年10月までドイツ連邦共和国ゲーセン大学医学部医用情報研究所客員研究員。昭59名工大・情報・助教授。現在、名工大・電気情報・助教授。生体情報処理、医用画像処理、ニューラルネットワークに関する研究に従事。工博。日本ME学会、情報処理学会、IEEE各会員。



松尾 啓志

昭58名工大・情報卒。昭60同大大学院修士課程了。同年松下電器産業(株)入社。平1名工大大学院博士課程了。同年名工大・電気情報・助手。現在に至る。画像処理・画像認識に関する研究に従事。工博。情報処理学会、ソフトウェア学会、IEEE各会員。



鈴村 宣夫

昭28名大・工・電気卒。民間会社勤務の後、昭38名大工・助手。以後、講師、助教授を経て、昭49名工大・情報工学科教授。学科改組により、現在、電気情報工学科教授。この間、生体信号の計測、処理、生体関連の画像処理等の研究に従事。工博。



堀田 健一

平2名工大・電気情報卒。現在、同大大学院博士前期課程(電気情報工学専攻)在学中。ニューラルネットワークに関する研究に従事。