

IV—9

画像データによる運転者支援システムの開発

(ニューラルネットワークモデルによる画像の識別について)

北海道大学工学部	学生員	中 井 幸 治
北海道大学大学院	学生員	野 口 智 之
北海道大学工学部	正 員	中 辻 隆
同 上	正 員	加 来 照 俊

1. はじめに

現在のハイモビリティ社会、さらに近い将来の高齢化社会の到来により、交通事故の問題はますます深刻化してくる。安全性を人・車・道路から見ると車と道路の安全設計は物理的な向上が可能であり、その方向に向けて開発が進められている。しかし、人については不特定多数が運転するため運転量や運転時の肉体・精神的状態により安全性が左右される。したがって、路上の障害物検知、車間距離推定、及び走行環境認識等、運転者の認知や判断を補助する支援システムを確立することによって運転者の錯誤や居眠り等に起因する事故の発生を未然に防ぐことは、将来における交通安全の向上に大きく寄与するものと思われる。

本研究では、ニューラルネットワークモデルの代表的な学習モデルの一つである逆伝搬（バックプロパゲーション）法の運転者支援システムにおける適用性について検討することを目的としている。まず最初に、画像データによる自動運転・運転者支援システムについて簡単に紹介する。次に、単一のニューロンレベル及び小規模なニューラルネットワークレベルでのモデルの情報処理機構と、逆伝搬法の学習メカニズムについて概説する。さらに、簡単な文字の認識を例にして逆伝搬法においてモデルの能力に大きな影響を及ぼす中間層の数とそのニューロンの数について解析を行う。

2. 自動運転システム

車両に搭載した小型のテレビ・カメラをコンピュータに接続し、そのカメラで撮った画像で道路上の白線や道路の縁を認識して道路の中央に沿って車両を誘導するのが自動運転システムである。これは、運転者—車—道路系に状況認識機能、知能処理機能、および自動操縦機能を備えることにより実現化される。アメリカでは、ペンタゴンを中心として人工知能を搭載した小型車を用いて、無人走行を行うオートノマス・ビークル・プロジェクト計画が行われている。ヨーロッパにおいても、主要自動車メーカーが中心となってプロメテウス計画が研究されている。

一方、画像データにより、前車との車間距離や自車の走行車線からの逸脱、あるいは歩行者の存在などを運転者に警告しようとするものが運転者支援システムである。画像データは、従来の距離センサーなどと比べ、持てる情報量が極めて大きいという長所がある。ところが、従来のノイマン型計算機を用いて画像データの処理を行う際にはこの計算機が直列演算を基本としているため、パターン認識能力に劣り、その処理速度も遅く、リアルタイム処理に対応できなかった。

人間の脳の機能を再現しようとするニューラルコンピュータは、従来の直列型の計算機と異なりパターン認識の能力に優れていることから、画像データによる運転者支援システムの開発に大きな力を発揮するものとして期待されている。

Application of Neural Network Model to Computer Aided Driving System,
by K. NAKAI, T. NOGUCHI, T. NAKATSUJI and T. KAKU

3. ニューラルネットワークモデル

3.1 ニューラルネットワークモデル

神経回路網（ニューラルネット）は神経細胞（ニューロン）が多数結合して何らかの信号を出す一つのシステムであり、個々のニューロンを工学的に見ると他のニューロンからの入力を受け取るシナプス結合の機能とこの受け取った出力を出す機能を持っている（図1）。また、シナプス結合を介してどれだけの入力を送るかという結合の重みはそれぞれのシナプス結合に依存する。今、シナプス結合の極性（正：興奮性、負：抑制性）と重みを含んだ結合強度を W_{ij} とすると i 番目のニューロンが受け取る入力 U_i は

$$U_i = \sum_{j=1}^N W_{ij} V_j + I_i \quad (1)$$

で与えられる。 I_i は外部からの i 番目のニューロンへの入力であり、入力 U_i を受け取ったニューロンは出力 V_i を出す。

$$V_i = f(U_i / \mu_0) \quad (2)$$

ここに関数 f は、 μ_0 の値によって傾きが変わるシグモイド型の単調増加関数である（図2）。

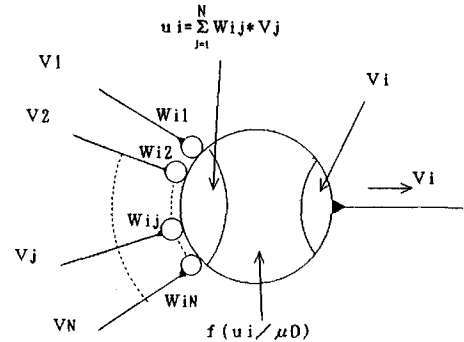


図1 神経細胞のモデル

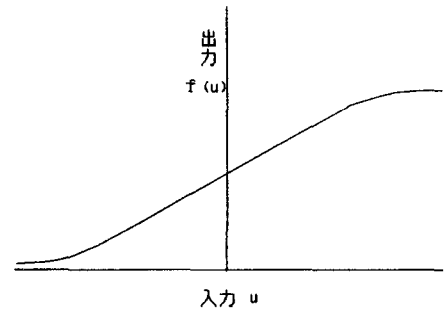


図2 シグモイド関数

3.2 逆伝搬法

図3に示した階層型モデルを例として逆伝搬法のアルゴリズムについて説明する。ここで、A層が入力層、B、C、D層が中間層、及びE層が出力層とする。学習に関係するのはB、C、D、E層の4層で、A層は入力された信号をB層に伝える役目だけを担うと仮定する。

逆伝搬法においては図3に示すように信号の流れと、その逆方向の誤差補正の流れがある。信号の流れは、入力層に、ある信号が入力されると、ニューラルネットのダイナミクスに従って中間層から出力層へと信号が順次伝達されていく。このネットワーク中で、理論的に求めることができないパラメータとして中間層の数・各層のニューロンの数などがあるが、これらは経験的に定められる。一方、誤差補正の流れについては、出力層からの信号が既知の教師信号と比較され、二乗誤差の程度に応じて各層間のシナプス荷重の大きさを入力層へ向けて順次補正していく。今、出力層の i 番目のニューロンの出力信号を y_i 、それに対応する教師信号を t_i とすると、この教師信号と実際に得られた出力信号との二乗誤差をとれば誤差関数 E は、

$$E = \sum_i (t_i - y_i)^2 / 2 \quad (3)$$

と表される。逆伝搬法の学習とは、この E を極小化するようにネットワークのシナプス荷重を変化させるこ

とにほかならない。いま、式(2)で用いたシグモイド関数を次式で定義する。

$$f(u) = 1 / \{ 1 + \exp(-u) \} \quad (4)$$

次に、シナプス荷重を各々 W 、その変化量を dW と表すものとする。まず、出力層とその直前の中間層へのシナプス荷重の変化量 dW_{kl} は、

$$dW_{kl} = \eta \delta_l y_k = y_l (1 - y_l) (t_l - y_l) y_k \quad (5)$$

$$\delta_l = y_l (1 - y_l) (t_l - y_l) \quad (6)$$

ここで、 δ_l は式(3)を W_{kl} について微分することにより求まる量であり、 y_k はニューロン k からの出力値、 η は学習定数と呼ばれる十分小さな正数である。次に中間層間のシナプス荷重の変化量についても δ を以下のように定義することにより式(6)と同様に、

$$\delta_k = y_k (1 - y_k) (\sum_l \delta_l W_{kl}) \quad (7)$$

$$dW_{jk} = \eta \delta_k y_j \quad (8)$$

さらに手前の中間層の変化量も全く同様にもとめることができる。

実際の計算においては、変化量の急変をさけるためにモーメント法による調整が行われる。

$$\Delta W(n) = dW(n) + \alpha \Delta T(n-1) \quad (9)$$

ここで、 n は収束のための演算回数で、 α は0から1の値をとる定数である。すなわち、調整後のシナプス荷重の大きさは、

$$W(n) = W(n-1) + \Delta W(n-1) \quad (10)$$

と表される。

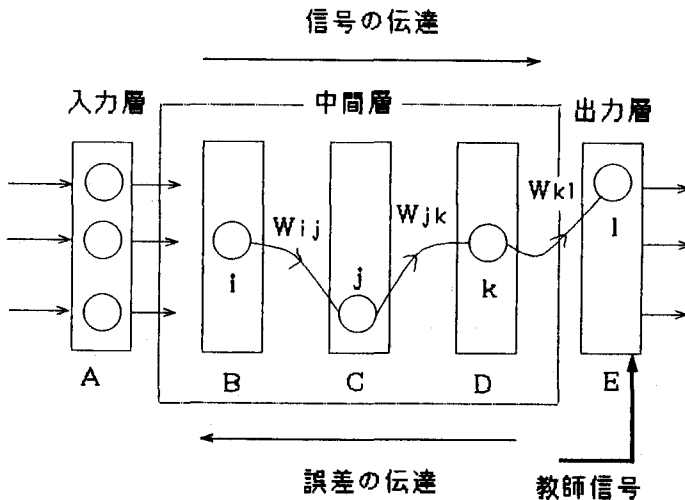


図3 逆伝搬法の学習メカニズム

4. 適用例

階層型のニューラルネットワークモデルにおいては、中間層の数とそのニューロンの数をどのように設定するかによってモデルの識別能力が大きく変わる。各ニューロンの出力値が1、0の2値をとり、出力関数がいき値関数として与えられる時には、入出力関数の線形性から論理的にそれらの影響を考察することができる。一般的に、中間層の数やそのニューロンの数が大きくなるに従って識別能力は高まると言われているが、入出力関数が式(4)のような非線形関数として与えられる時には、経験的にそれらの影響を調べる必要がある。

ここでは、かな文字「あ」「い」「う」「え」「お」の5文字を識別対象としてシミュレーションを行った。出力層のニューロン数5個、及び入力層 $16 \times 16 = 256$ 個のニューロンを固定し、中間層の数及び各中間層のニューロンの数を変化させた。その手順としては、1文字につき20パターンの文字データをランダムに15回繰り返して学習させた後に、学習に用いた5文字 $\times$ 20パターンの学習データ及び同じく5文字 $\times$ 20パターンの未学習データを入力した。中間層数1つ、及びそのニューロンの数を4つとするモデルを基準として、中間層の数とそのニューロンの数を変化させた。

学習プロセスにおいては、図4に示すように、当該文字に対応する出力ニューロンにのみ1、その他に0として教師信号を与える。そしてそれらの値と、各出力ニューロンでの出力値との比較を行い、逆伝搬法によって各層間のシナプス荷重が調整される。

図5は、中間層のニューロンの数の影響を調べたものである。ニューロンの数が4以上となると既学習データに対しては100%の、未学習データに対してもほぼ70%～80%の識別率が得られている。ニューロン数が2の場合には既学習データに対しても50%程度の識別率であった。ニューロン数が4以上

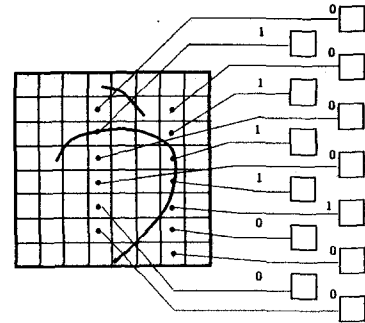


図4 手書き文字認識の例

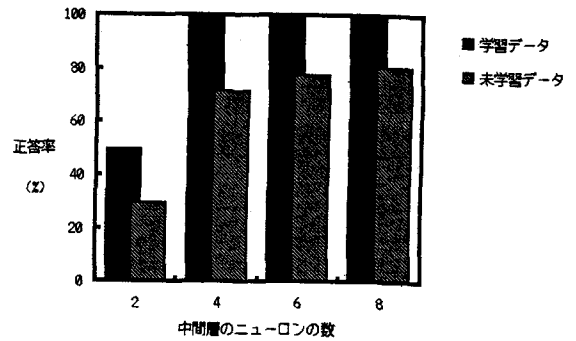


図5 中間層のニューロンの数の影響

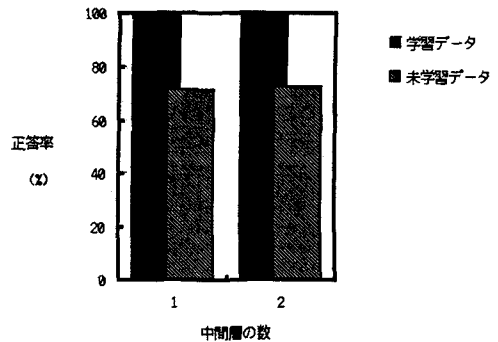


図6 中間層の数の影響

の場合には、ニューロン数が増えるに従って未学習データに対する識別率が向上している。

図6は、中間層の影響を示したものである。計算時間の制約から、中間層が1層と2層の場合のみ解析を行ったが、前者で72%、後者で73%の識別率であり若干ながら識別率が向上した。

ニューラルネットワークモデルにおいては、中間層やそのニューロンの数以上にどのようなデータでどのように学習するかによっても識別能力は変わる。今回の場合では、同じ文字を連続して学習させたときには、中間層1、ニューロン数4の基準モデルに対しても最後に学習した文字への応答が異常に高くなる傾向を示した。なお、ここでの演算はすべてパソコンによるものであり、プログラムはMS-Cで記述されている。基準とした中間層1、ニューロン層4で、1500パターンの学習に約1時間の演算時間(NEC PC-9801RA)を要した。これを、16*16の代わりに64*64の画面分割を行った場合には、1パターンの学習においても数分の演算時間を要した。

5. まとめ

ニューラルネットワークモデルを運転者支援システムに適用するための第1段階として、簡単な文字認識を例に挙げ、階層型ネットワークモデルの中間層の数とそこにおけるニューロン数の影響について調べてみた。得られた結果を要約すれば以下の通りである。

- 1) かな文字程度の認識では、縦横16*16程度の分割能によって、未学習データにたいしても70~80%の認識が可能である。
- 2) 中間層の数、あるいはニューロンの数が増えると識別能力は高まる。
- 3) 学習方法については、規則的な順番で行うよりもランダムな順番で行う方が好ましい。

ここでは、かな文字の識別という最も簡単な例を対象として、強力な学習能力とアルゴリズムの単純さからニューラルネットワークモデルの代表的学習アルゴリズムとなっているバックプロパゲーション法の適応性の検討を行い、その能力に関していくつかの知見を得た。しかしながら、それはあくまでも静止画像に対するものであり、本研究の最終目的である運転者支援システムに利用するには解決すべき課題は多い。

- 1) バックプロパゲーション法においては、静止画像でも相似形画像の認識は苦手であり、急接近を回避するための車間距離推定システム等に利用するには限界がある。
- 2) 動画画像の認識には背景画の処理などが必要となる。
- 3) パソコンレベルでの演算では、メモリ容量および演算時間の制限から64*64程度の画面が限界であり、オンライン処理を行うためにはニューロエミュレーションボードの導入などを検討する必要がある。

◆参考文献

- 1) 中野、飯沼、ニューロンネットグループ、桐谷：「ニューロコンピュータ」 技術評論社、1989
- 2) Philip D.Wasserman：「NEURAL COMPUTING」 VAN NOSTRAND REINHOLD、1989
- 3) 麻生：「ニューラルネットワーク情報処理」 産業図書、1988
- 4) 合原：「ニューラルコンピュータ」 東京電機大出版、1988
- 5) 宮沢：実践！ニューラルネット、ASCII Vol.12 #10-Vol.13 #4、1988.10-1989.4
- 6) 宇野：運動軌道の生成と学習システム、コンピュータロール No.24 pp29-37、1988