

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/220587524>

# 手書き数字認識のための多層パーセプトロンと最近傍分類器の比較。

2005年11月ジャーナル・オブ・インフォメーション・サイエンス・アンド・エンジニアリングに掲載: DBLP

## CITATIONS

10

5 authors, including:



**Kaushik Roy**

西ベンガル州立大学

307の出版物 3,471の引用

[SEE PROFILE](#)



**Mahantapas Kundu**

Jadavpur University

133 PUBLICATIONS 3,161 CITATIONS

[SEE PROFILE](#)

## READS

2,192



**Chitrita Chaudhuri**

Jadavpur University

13 PUBLICATIONS 237 CITATIONS

[SEE PROFILE](#)



**Mita Nasipuri**

Jadavpur University

649 PUBLICATIONS 9,102 CITATIONS

[SEE PROFILE](#)



## Short Paper

---

### 手書き数字認識のための多層パーセプトロンと最近傍分類器の比較

K. ROY\*, C. CHAUDHURI, M. KUNDU, M. NASIPURI AND D. K. BASU

*Computer Science and Engineering Department*

*Jadavpur University*

*Kolkata 700032, India*

*\*C.V.P.R. Unit*

*Indian Statistical Institute*

*Kolkata 700108, India*

この論文は、手書き数字認識問題に対する多層パーセプトロン（MLP）と最近傍（NN）分類器の性能を比較するために実施された調査の結果を示している。比較は、認識性能と個々の分類器の計算要件の観点から行われる。その結果、2層パーセプトロンは、制約のない手書き数字を認識する際に、標準的なパターン分類器のようなNNに匹敵する性能を示す一方、計算コスト的にはより効率的であることが示された。この研究は、96特徴という大きな特徴セットを持つ手書き数字を認識するための標準パターン分類器としてのMLPの有用性を示すものである。

キーワード：人工ニューラルネットワーク、多層パーセプトロン、パターン認識、最近傍分類器、学習、汎化、訓練

## 1. INTRODUCTION

多層パーセプトロン（MLP）は、フィードフォワード人工ニューラルネットワーク（ANN）の重要なクラスであり、生物学的ニューラルネットワークの機能をモデル化する試みにより、人間の学習能力や汎化能力を再現するために開発された。人工知能（AI）やパターン認識（PR）の分野で多くの応用が期待されている。手書き数字認識は、パターン認識のベンチマーク問題である。商業的な重要性が明確に定義されており、難易度が高いが、完全に解決不可能というほど大規模な問題ではない。手書き数字の光学的文字認識（OCR）は、銀行の小切手から金額を読み取ったり、記入されたフォームから数字データを抽出したり、郵便物から手書きのピンコードを読み取ったりなど、多くの商業アプリケーションの中心となっている。ここで紹介する研究は、パターン分類器としてのMLPの有用性を、従来の分類器として最適でなかった最近傍（Nearest Neighbor：NN）分類器と比較することを目的としています。

Received March 13, 2003; revised August 11, 2003 & March 15, 2004; accepted April 29, 2004.  
Communicated by Kuo-Chin Fan.

## 1.1 Previous Work

以前、Burr氏[1]はMLPを用いた手書き数字の認識に、7つのシャドーコードからなる小さな特徴セットを使用していました。Burrの研究の後、Weidemanら[2]は、手書き数字認識問題に対するMLPのようなフィードフォワードニューラルネットワークとNN分類器の認識性能と計算要件を比較するための大規模な研究を行った。この研究では、手書き数字を表現するための98の特徴セットを検討した。この特徴量セットには、36の幾何モーメント特徴量、18の位相特徴量、28の2次元高速フーリエ変換（FFT）特徴量の特別に選択された成分、16の影特徴量が含まれる。これらについては後で説明する。Weidemanらの研究は、最終的に、複雑な問題に対してニューラルネットワークはNN分類器と同等の性能を発揮し、同時に費用対効果も大幅に向上すると結論付けている。

## 1.2 Motivation

Weidemanらの研究のために設計されたMLPのようなフィードフォワード・ニューラルネットワークは3層であった。入力層に加え、3層パーセプトロンは2つの隠れ層と1つの出力層から構成される。Weidemanらによって設計されたネットワークのシナプス結合は、通常のMLPで見られるものよりも複雑であった。層から次の層へのシナプス接続に加えて、2つの隠れ層のそれぞれから次の層への接続も設計されている。しかし、もし隠れ層を1つ減らしても（つまりネットワークの複雑さを減らしても）、少なくとも認識率は変わらないまま、層から次の層へのシナプス結合だけを維持できるのであれば、よりコスト効率の良いネットワークの設計が可能になる。本研究の動機はこの点にある。

### 1.3現在の作業

この研究では、手書き数字を認識するための2層MLPの費用対効果の高い設計を紹介します。また、MLPとNN分類器の性能の比較評価も行う。手書き数字を表現するために選択された特徴セットには、Weidemanらによって選択された2つの特徴（数字画像にループがある場合、ループの上端と下端のエネルギーの比と、数字画像内のループの数）を除くすべての特徴が含まれている。これらの特徴量を除外しても、本研究のネットワークの認識性能に影響はない。

## 2. 特徴ベクトルの設計

前述したように、ここでは、特徴空間における各数字画像を表現するために、96成分の特徴ベクトルを考慮する。数字画像から特徴を抽出するために、まず画像を最小の矩形で囲み、それぞれを32×32ピクセルのサイズにスケールアップする。こうして得られた拡大縮小画像は、最後に閾値処理によって2値化される。これは、数字画像をサイズ不変にし、同時に並進を避けるためである。数字画像から抽出されたすべての特徴は、実際に使用する前に正規化が必要である。

### 2.1 影の特徴

影特徴量は、図1に示すように、最小サイズの長方形の4辺と8分辺における数字画像の投影の長さを考慮して計算される。本研究で考慮した特徴ベクトルには、16個の影特徴が含まれている。

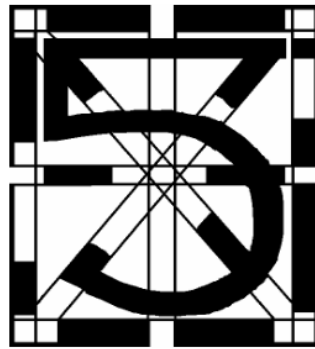


図1. 数字画像の影の特徴。

### 2.2 トポロジー特徴

トポロジー的特徴は、数字パターンの輪郭に関する情報を提供する。例えば、画像パターン「0」の最小外接矩形の4つの角点からの最初の黒画素の対角線方向の距離は、画像パターン「1」に対して同じ方法で計算された距離とは異なる。本研究では、このような16個の位相的特徴[2]を考慮し、特徴ベクトルを形成する。

### 2.3 2次元FFT特徴量

2Dデジタル画像からFFT特徴を抽出するためには、画像を画像平面上のx-y座標の離散関数として表現し、それらの点におけるグレーレベルの値を与える必要がある。このように表現された画像は、FFTを適用することで、重み付けされた多数の周期関数（正弦波）に分解することができます。これらの関数に関連する重みは、複素数であるフーリエ係数（FC）の集合を構成する。こうして画像から生成されるFCの数は、ピクセルの数に等しい。係数の大きさの2乗はエネルギーと呼ばれる。

この特徴ベクトルには、数字画像から抽出した24個の低周波FFT係数が含まれる。これらの係数は、高周波の係数に比べ、回転やスケーリングの影響を受けにくい。また、中心極限定理[2]により、ガウス分布の確率密度を持つ。

特徴ベクトルには、画像平面の対角線の上半分にある、隣接する4つの領域のFFT係数を用いて計算された周波数領域のエネルギーも含まれる。

これらの4つの領域はそれぞれ、対角線の中点で $45^\circ$ の角度を持つ。これらの領域から、数字画像中の水平、垂直、斜めの線の存在を検出するための4つの特徴量を得る。

#### 2.4 幾何モーメント

2次元幾何モーメントは、画像から物体の形状を記述するためによく使われる特徴である。数学的には、モーメントの集合は、2変数の2値画像関数の2変数多項式の集合への投影と見なすことができる。モーメントは平行移動、回転、拡大縮小、反射に対して不変である。本研究では、4つの画像象限で計算された36の正規化モーメント（4次まで）を用い、象限ごとに9つのモーメントを計算する。

### 3. DESIGN OF THE MLP

MLPは一般に層状のフィードフォワードネットワークであり、図2に示すように有向非循環グラフで表現される。グラフ内の各ノードは MLP の人工ニューロンを表し、各有向アークは 2 つのニューロン間のシナプス結合と、MLP 内の信号の流れの方向を表します。グラフのアークで使用されているラベルは、MLP のシナプス結合の強さを表し、重みとも呼ばれます。MLPの各層は特定の数のニューロンで構成され、各ニューロンは直後の層のすべてのニューロンと接続され、生物学的ニューラルネットワークのシナプス結合を再現します。生体ニューロンの機能は、MLPの各人工ニューロンに対して微分可能な非線形関数（シグモイドなど）を計算することによってモデル化されます。このようなシグモイド関数の使用は、生体ニューロンの不応期を考慮しようとするものであり、生物学的な動機付けにもなっています。

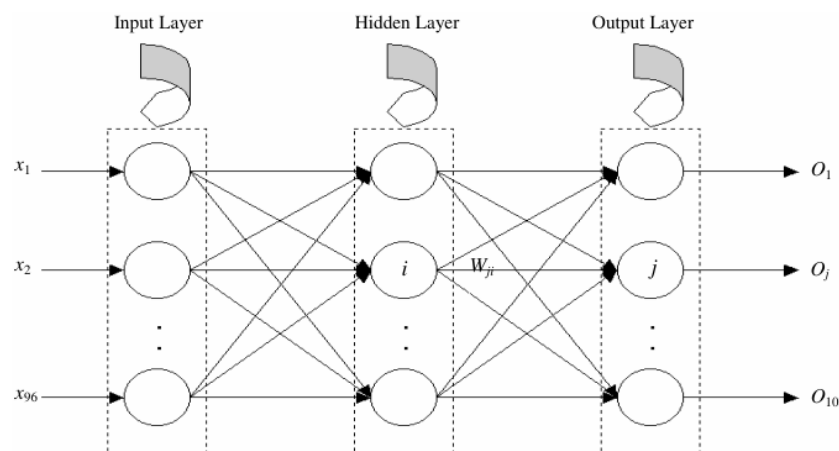


図2. 図2. 2層MLPのグラフ。

パターン分類の場合、MLPの入力層のニューロン数は、特徴空間において関連するパターンを表現するために選択される特徴の数によって決定される。隠れ層と出力層の各ニューロンは、入力値と対応する接続の重み値の積の和のシグモイド関数を計算する。ここで、 $net_j = \sum_i W_{ji} O_i$ 、 $W_{ji}$  = i-jニューロン間の接続の3番目の重みである。

MLPの出力層に含まれるニューロンの数は、興味のある問題で扱われる可能性のあるパターンクラスの数によって決定される。最も高い出力値を生成する出力ニューロンに割り当てられるクラスラベルによって、MLP に供給される入力パターンのクラスが決定される。

MLPの学習プロセスでは、主にシナプス結合の強さの調整、隠れ層の数およびその中のニューロン数の調整が行われます。MLPには学習能力が組み込まれています。同じMLPの汎化能力は、トレーニング後に、トレーニングに使われなかった入力パターンに対する応答をチェックすることでテストされる。

### 3.1 逆伝播アルゴリズム

教師あり学習法であるバックプロパゲーション (BP) アルゴリズムは、既知のクラスパターンを用いたMLPの学習に使用される。このアルゴリズムは、すべての学習パターンで生成される2乗誤差の合計を最小化するパーセプトロンの重み値のセットを見つけることを目的としています。その際、BPアルゴリズムはパーセプトロンの誤差面を重み空間で勾配降下探索します。誤差の二乗和を最小化するために必要な重みの変化量  $\Delta W_{ji}$  は以下の通りであることが証明できる。

$$\Delta W_{ji} = \eta \delta_j O_i \quad \text{where } \eta = \text{learning rate parameter } (0 < \eta < 1),$$

$$\delta_j = \text{the error gradient of } j^{\text{th}} \text{ neuron}$$

$$\delta_j = \begin{cases} O_j(1 - O_j)(d_j - O_j) & \text{if } j^{\text{th}} \text{ neuron belongs to output layer} \\ O_j(1 - O_j) \sum_k \delta_k W_{kj} & \text{if } j^{\text{th}} \text{ neuron belongs to a hidden layer} \end{cases}$$

運動量項  $\alpha$  ( $0 < \alpha < 1$ ) を導入することで、BPアルゴリズムが使用する重みを更新する式は次のようになります。

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_j O_i(t) + \alpha(W_{ji}(t) - W_{ji}(t-1)), \text{ where } t \text{ stands for time.}$$

本研究の停止基準は、900 個の学習パターンすべてについて、二乗誤差の和が 0.6 以下の正の数になる条件として選択される。

### 3.2 入力層、出力層、隠れ層の設計

Weidemanらが手書き数字認識のために設計した3層のニューラルネットワークと比較して、本研究では2層のパーセプトロンを選択した。普遍近似定理[3]によれば、入力 $x_1, \dots, x_n$ と目標出力関数 $f(x_1, \dots, x_n)$ の集合で表される与えられた学習集合に対して、一様な $\varepsilon$ 近似を計算するには隠れ層が1層あれば十分である。ここで選択したパーセプトロンの入力層と出力層のニューロン数は、特徴量セットの特徴数が96であり、手書き数字のクラス数が理想的には10であることから、それぞれ96と10とした。最適な隠れ層のニューロン数を選択するために、個々の学習パターンに対する大域誤差の総和の隠れユニット数に対する変化を示す曲線を図3にプロットする。大域誤差の和は、パーセプトロンを1000エポック訓練した後に毎回記録される。大域誤差の総和を記録する際、パーセプトロンは訓練セットでテストされる。カーブプロット実験は、 $\eta$ と $\alpha$ の値をそれぞれ0.65と0.7に選び、100個の数字からなる小さな学習セットを用いて行われる。この段階で使用される学習セットは、10個の各クラスから同数ずつランダムに選ばれた100個のサンプルパターンから構成される。 $\eta$ と $\alpha$ は、実験の初期段階で学習アルゴリズムの収束を早めるために高い値が選ばれる。しかし、このような選択はオーバーシュートの危険性がある。設計したMLPの汎化能力をテストするため、大規模なデータベースから各クラス10個のサンプルを任意に選び、100個の手書き数字からなるテストセットを作成します。隠れユニットの数は、図3に示す曲線が最小の誤差をもたらすように、最終的に17の値に設定される。

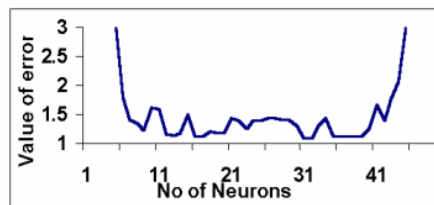


図3. 図3. 2層パーセプトロンの隠れニューロン数の増加に伴うグローバル誤差総和の変化。(  $\eta = 0.65$ ,  $\alpha = 0.7$ , 学習セットサイズ = 100, エポック数 = 1000)

### 3.3 $\eta$ と $\alpha$ の値の選択

$\alpha$

BP学習では、学習率 $\eta$ は、重み空間においてネットワークに生成される誤差面の大域的極小値を探索する際の網羅性の程度を制御するために選択される。 $\eta$ の値が小さいほど学習は遅くなり、探索はより網羅的になる。しかし $\eta$ の値が大きいと、収束が速くなり、特に誤差曲面上の曲率の高い点では発散的な振る舞いになる可能性がある。

モーメントパラメータ $\alpha$ は、改良された加重値に前回の加重変化の割合を加えるもので、各反復における偏差の程度を制御する。

$\alpha$ の値は、全体的な探索プロセスにおける $\eta$ の影響を補正するために使用することができる。 $\eta$ と $\alpha$ の値の適切な選択は、学習速度と解の質のトレードオフを必要とする。

学習速度と解の質に対する $\eta$ と $\alpha$ の値の異なる組み合わせの効果を観察するために、 $\eta = 0.45, 0.55, 0.65, 0.75, 0.85$ と $\alpha = 0.6, 0.7, 0.8, 0.9$ に対して、エポック数に対する誤差値を示す曲線が図4(a-e)に描かれている。曲線に示された誤差は、トレーニングセットで計算されたものである。

曲線から、 $\eta$ の値が高い場合、エポック数が増加するにつれて誤差が急激に減少することが観察される。これは、大域的極小値を探索する際に誤差面がオーバーシュートする危険性があります。そこで $\eta$ は $\{0.45, 0.55\}$ から選択する。これらの $\eta$ の値について、エポック数に対する誤差値を示す曲線を図4(a-b)に示す。これらの曲線のうち、 $\eta = 0.45$ と $\alpha = 0.7$ に対して描かれた曲線は、エポック数の増加とともに誤差値が着実に減少していることを示している。したがって、 $\eta$ と $\alpha$ の値は最終的にそれぞれ0.45と0.7に選ばれた。

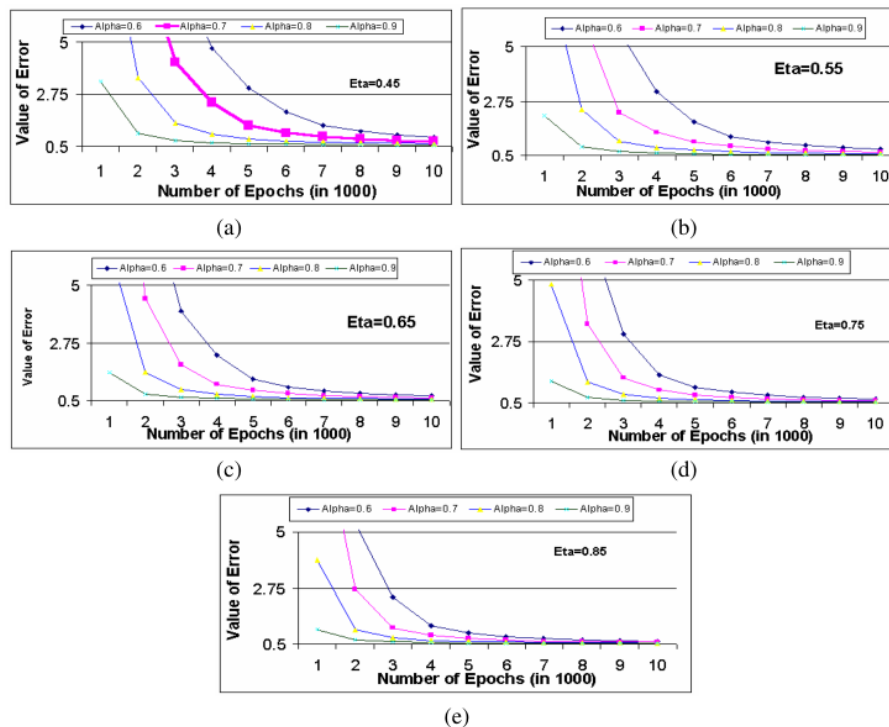


図4. エポック数の増加に伴う大域誤差総和の変化。

### 3.4 隠れニューロンの数の最終選択

隠れニューロンの数をより細かくチューニングするために、トレーニング中のエポック数を1000から10000まで1000ステップで増加させ、



テストセット上で異なる隠れニューロン数を持つMLPのパーセント認識率を記録した。実験中、 $\eta$  と  $\alpha$  の値はそれぞれ0.45と0.7に固定した。結果を表1に示す。

表1. 隠れ層のニューロン数を変えた場合の、エポック数の増加に伴う認識率の変化。

Epochs No of H U	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
7	86	92	94	95	95	95	95	95	95	95
10	91	93	94	95	95	95	95	96	96	96
13	94	96	97	97	97	97	97	97	97	97
15	95	96	97	97	97	97	97	97	97	97
17	94	95	96	96	97	97	97	97	97	97
23	94	95	96	96	96	96	96	96	96	96
30	94	95	95	96	96	96	96	96	96	96
37	94	95	95	95	96	96	96	96	96	96
45	94	95	96	96	96	96	96	97	97	97
96	92	95	96	96	96	96	96	96	96	96

表から、達成可能な最高認識率は97%であることがわかる。また、隠れユニットが15の場合、より少ないエポックで最高の認識率に達している。これが、MLPの隠れニューロン数が最終的に  $\eta = 0.45$ 、 $\alpha = 0.7$  で15と選ばれた理由です。

### 3.5 MLP の棄却基準

上述のように、 $\eta$ 、 $\alpha$ 、および本研究の収束基準の値を選択した後、900 サンプルのトレーニングセットで BP アルゴリズムを実行することにより、2 層MLP (96-15-10) が設計されます。Weideman らの研究と同様、MLP にもリジェクト機能が導入されている。リジェクト基準は以下の通りである。

- i) 出力ニューロンの活性化レベルの最高値がある閾値（この場合0.7）を下回ると、入力パターンは拒絶される。あるいは
- ii) 出力ニューロンの1番目と2番目に高い活性化レベルが非常に近い場合（つまり、この場合2つの差が0.2以下である場合）、入力パターンは棄却される。

## 4. NN分類器の設計

NN 分類器は、参照集合と呼ばれる既知のクラスのサンプルパターン集合を必要とする。未知のパターンXを分類するためには、まず特徴空間内でXに最も近い参照パターンを見つけ、同じクラスラベルをXに割り当てます。

### 4.1 なぜNN分類器なのか？

性能的には、ベイズ分類器は分類誤りを犯す確率が最も低いので最適である。しかし、その設計には、すべてのクラスの確率と密度、

および判断のコストに関する知識が必要であり、これは今回の研究で使用したサンプルでは利用できない。そこで、ノンパラメトリック・アプローチに従うNN分類器が選択される。その誤り率は、低確率の誤りに対して、ベイズ誤り率の2倍を超えることはありません[4]。NN分類器は、特にベイズ誤差率がゼロに近づくか、ランダム推測の誤差率に近づく場合に、ベイズ分類器の良い近似であると考えることができる。

#### 4.2 k-最近傍分類器

未知のパターン $X$ 进行分类するために、k-最近傍(k-NN)分類器は、まず $X$ に最も近い参照パターンを $k$ 個見つけ、その $k$ 個の参照パターンの中で最大のサンプル数を持つクラスに $X$ 进行分类する。k-NN分類器の誤り率は、 $k$ の値と、 $k$ の値に対する参照サンプルの総数の比率の両方が無限大に近づくにつれて、ベイズ分類器の誤り率に近づく。

この研究では、未知パターンの $k$ 個の最近傍それぞれに重み付き投票[4]を用います。ここで、 $d_e$ は未知パターンとその特徴空間内の $k$ 個の最近傍の1つとの間のユークリッド距離である。

#### 4.3 参照集合の設計

本研究では、マキシミン距離クラスタリング[5]を採用し、MLPの学習に使用した900サンプルと同じ学習セットからクラスタ中心を求める。こうして得られたクラスタ中心は一次参照集合を形成する。この参照集合は、訓練集合の非中心サンプル进行分类する1-NN分類器に使用されます。非中心サンプルは、このようにしてクラスタ中心とグループ化されます。50%以上の誤分類サンプルに関連するクラスタ・センターは、k-NN分類器を使用するために一次参照集合から除去されます。最終的に、参照セットは179のサンプルパターンから構成される。

#### 4.4 kの値の選択

ここでの作業に使用された参照集合では、クラスが持つプロトタイプの最小数は3であり、クラスが持つプロトタイプの最大数は38である。この情報に基づいて、本作業では $k$ の値を3とした。

#### 4.5 NN分類器の棄却基準

NN分類器に組み込まれる棄却判定基準は以下の通りである。入力パターン $X$ は、その最も近い参照サンプルとの距離が、特徴空間内の異なるクラスの2番目に近い参照サンプルとの距離の3%未満である場合に棄却される。

## 5. EXPERIMENTAL RESULTS

本研究の学習セットとテストセットは、それぞれ900個と100個のランダムに選ばれた手書き数字画像のサンプルから構成される（図5）。前述したように、これらのサンプルはすべて32×32ピクセルサイズの2値画像である。



図5. 各数字の最初の5サンプルのBMP画像。

### 5.1 認識性能

MLP (96-15-10) 分類器の認識性能を、学習セットとテストセットで記録したものを表2 (a) に示す。179のクラスタ中心を持つ3-NN分類器の認識性能を表2 (b) に示す。その

表 2. (a) 2層パーセプトロンの認識性能。

Training Set			Test Set		
% of Classification	% of Miss-Classification	% of Rejection	% of Classification	% of Miss-Classification	% of Rejection
99.78	0.00	0.22	97.00	0.00	3.00

表2. (b) 3-NN分類器の認識性能。

Training Set			Test Set		
% of Classification	% of Miss-Classification	% of Rejection	% of Classification	% of Miss-Classification	% of Rejection
95.00	4.13	0.86	90.00	8.00	2.00

### 5.2 計算量

計算量を調べるには、アルゴリズムに必要な乗算、加算、関数評価の回数を比較するのが合理的である。階層ネットワークは $N_h$  ( $N_i + N_o$ ) 個の重みを必要とし、添え字はそれぞれ隠れユニット、入力ユニット、出力ユニットの数を示す。つまり

96-15-10ネットワークは、2つのバイアスニューロン（1つは隠れ層への入力用、もう1つは隠れ層から出力層への出力用）の出力に関連するものを含め、1,615個の重みを持つ。これは、各文字の決定に対して、ここで選んだMLPネットワークは1,615回の乗算、1,615回の加算、ロジスティック関数の25回の評価を必要とすることを意味する。このため、1回の決定につき合計3,255回の演算が必要となる。一方、179のパターンからなる参照集合を用いてここで設計された3-NN分類器は、17,184の値を格納する。3-NN分類器による各文字の決定には、17,187回の乗算、34,371回の加算／減算、179回の平方根関数の評価が必要である。これは1回の判定につき、合計51,737回の演算を必要とすることになる。つまり、3-NN分類器は1つの手書き数字画像を分類するのにMLP（96-15-10）の15.9倍の計算を必要とする。これに加えて、3-NN分類器は、特徴セットに対してMLPの10.64倍のストレージを必要とします。

## 6. CONCLUSIONS

上記の結果は、手書き数字認識において、96の特徴量を持つ2層パーセプトロンを設計することで、NN分類器に匹敵する認識性能を示すことができることを示唆している。

MLPの性能がk-NN分類器よりも優れている理由は、以下のように説明できる。NN分類器もMLPも、特徴空間内の未知のパターンを分類するために、任意の決定領域を近似することができます。ここで使用されるk-NN分類器の性能は、参照集合と作業に選択されるkの値に依存します。参照集合は、類似性の尺度としてユークリッド距離を使用し、クラスタ探索アルゴリズムで最大値によって識別されたクラスタ中心の一部を選択的に収集することによって形成される。クラスタ探索アルゴリズムの性能は、やはりそのために選択された類似性尺度に依存する。したがって、k-NN分類器の性能を微調整できる範囲は、ここではかなり限定される。しかし、MLPの隠れ層ニューロンの数と接続重みは調整可能であるため、決定領域の微調整の機会が広がります。このため、MLPの認識性能は、手書き数字パターンの分類に使用されたNN分類器の認識性能よりも向上させることができる。

1992年の第1回国勢調査光学式文字認識システム会議[6]で選ばれた上位10分類器のうち、手書き数字認識問題で観察された最高の認識性能は98%でした。これらの分類器はすべて、ある種の多層フィード・フォワード・ネットワークかNN分類器に基づいていた。これらの分類器で観察された最高の性能と比較すると、本研究で設計されたMLPの認識性能は97パーセントであることが観察された。この値は、第1回国勢調査光学式文字認識システム会議の上位10人の最高性能に近い。

Weidemanらの研究と比較して、本研究のために設計されたMLPの計算量は、層と層のシナプス接続のみを持つ2層パーセプトロンの選択により、さらに削減されています。このことは、Weidemanらの3層MLPニューラルネットワークで観察された、層-次層シナプス結合を追加した場合と比較して、認識性能に悪影響を及ぼさない。

最後に、ここで紹介した研究は、大規模な特徴セットを用いた手書き数字の認識のような複雑な実世界の問題に対する BP アルゴリズムの有用性と、標準的なパターン分類器としての MLP の有用性を確立するための裏付けとなる。

### ACKNOWLEDGMENTS

筆者らは、本研究の進行中に必要なインフラ設備を提供してくれたジャダヴプール大学コンピュータ理工学部マイクロプロセッサ応用センターおよびマルチメディア・ビデオ・プロジェクトに感謝する。

### REFERENCES

1. D. J. Burr, "Experiments on neural net recognition of spoken and written text," *IEEE Transactions on Neural Networks*, ASSP, Vol. 36, 1988, pp. 1162-1168.
2. W. E. Weideman, M. T. Manry, H. C. Yau, and W. Gong, "Comparisons of a neural network and a nearest-neighbor classifier via the numeric handprint recognition problem," *IEEE Transactions on Neural Networks*, Vol. 6, 1995, pp. 1524-1530.
3. S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed., Pearson Education Asia, 2001, pp. 208-209.
4. E. Gose, J. Richard, and S. Jost, *Pattern Recognition and Image Analysis*, Prentice-Hall of India Pvt. Ltd., 1999, pp. 167-169, pp. 174-175.
5. J. T. Tou and C. G. Rafael, *Pattern Recognition Principles*, Addison-Wesley Publishing Company, 1974, pp. 92-94.
6. R. A. Wilkinson, et al., "First census optical character recognition system conference," Technical Report, NISTIR 4912, US Dept. Commerce, National Institute of Standards and Technology, Gaithersburg, MD, 1992.

カウシク・ロイ (Kaushik Roy) : 1998年にアッサム大学 (インド、シルチャール)、2002年にジャダヴプール大学 (J.U.) でコンピュータ科学・工学の学士号および修士号を取得。2003年よりコルカタのI. S. I. でプロジェクト・リンク要員として勤務。現在の研究分野は、パターン認識、画像処理、人工知能。

1980年にジャダヴプール大学で電気工学学士号を、1982年に電気工学修士号を取得。2001年より同大学講師。現在の研究分野は、パターン認識、画像処理、マルチメディア技術、データマイニング。

マハントパス・クンドゥは、1983年にジャダヴプール大学で電気工学学士号、1985年に電気工学修士号、1995年に工学博士号を取得。1988年に同大学に講師として着任し、現在は同大学で講師を務める。現在の研究分野は、パターン認識、画像処理、マルチメディア・データベース、人工知能など。

ミタ・ナシプリは1979年、1981年、1990年にそれぞれジャダヴプール大学で電気工学学士号、電気工学修士号、工学博士号を取得。1987年より同大学教員。現在の研究テーマは画像処理、パターン認識、マルチメディア・システム。米国IEEE 上級会員、インドI.E. フェロー、インド・コルカタW.B.A.S.T. フェロー。

1964年、1966年、1969年にそれぞれジャダヴプール大学で電気工学学士号、電気工学修士号、工学博士号を取得。1968年より同大学の教員。現在の研究分野は、パターン認識、画像処理、マルチメディア・システム。IEEE（米国）上級会員、I.E.（インド）フェロー、W.B.A.S.T.（インド、コルカタ）フェロー、アレクサンダー・フォン・フンボルト財団（ドイツ）元フェロー。