

エルマンネットを拡張した文字認識ニューラルネットワークの作成と評価

篠沢 佳久[†]慶應義塾大学 理工学部 管理工学科[†]

1. はじめに

筆者らはテンプレートマッチング方式による文字画像の認識をニューラルネットワークによって実現する手法を提案してきた[1]. これは入力層にテンプレートの文字パターンと未知文字パターンを入力し, 同じ文字種の場合は発火し, 異なった文字種の場合は発火しないように動作するニューラルネットワークを用いて実現した.

本研究においては上記の方式を拡張し, 単純再帰結合型ネットワーク(エルマンネット)を利用することによって局所領域間ごとの類似性から認識結果を求める手法を提案し, 文字画像データベース ETL9B による評価実験を行なう.

2. テンプレートマッチング方式の実現

ニューラルネットワーク(以下ネットワークと略す)を用いたテンプレートマッチング方式による文字パターンの認識方法について説明する. 図 1 に示すようにネットワークの構造は, 入力層, 中間層, 出力層の 3 層から構成される階層型のフィードフォワード構造とする. 出力層のニューロンの個数は 1 個, 中間層のニューロンの個数は任意とする. 認識対象の文字種が N 個の場合, N 個のネットワークを用意し, 各ネットワークと任意の一字種との対応づけをする.

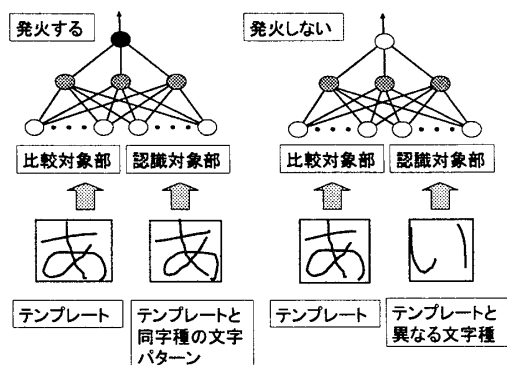


図 1 ニューラルネットワークの構造

Neural Networks for Character Recognition by using Improved Simple Recurrent Networks

[†]Yoshihisa Shinozawa, Faculty of Science and Technology, Keio University

入力層には, そのネットワークで処理する文字種のテンプレートの文字パターンから抽出した特徴と認識すべき未知文字パターンの特徴を同時に入力する. 入力層においてテンプレートの文字パターンの特徴を入力する部分を比較対象部, 未知文字パターンの特徴を入力する部分を認識対象部と呼ぶ. 従って入力層のニューロンは二つの文字パターンの特徴数分必要となる.

そして認識対象部に入力した未知文字パターンがテンプレートと同じ文字種である場合は, 出力層が発火し, 異なった文字種である場合は, 発火しないようにネットワークを動作させる. 文字種ごとで上記のような動作をするように各ネットワークを学習する. 未知文字パターンを認識する場合, 各ネットワークにおいて対応した文字種のテンプレートと未知文字パターンの特徴を入力層に入力し, 同時に全ネットワークを動作させ, 出力層の発火が最大のネットワークで対象としている文字種を認識結果とする.

3. 局所領域間でのマッチングへの拡張

本研究においては, 前節で述べたテンプレートマッチング方式を拡張し, テンプレートの文字パターンと未知文字パターンの全領域をマッチングするのではなく, 局所領域における特徴に考慮するためテンプレートと未知文字パターンを局所的な領域に分割し, 対応する領域ごとでマッチングを行ない, 各領域間ごとの類似性から認識結果を求める手法を提案する.

図 2 に示すように文字パターン x を任意の大きさの局所的な領域(ブロックと呼ぶ) $x_1, x_2, \dots, x_d$ に分割する. 図 2 の場合, 分割数 d は 4 とし, その大きさは文字パターンの半分としている. 文字種 i のテンプレート T_i と未知文字パターン x においても, それぞれを同じ分割方法でブロック $\{T_{i1}, T_{i2}, \dots, T_{id}\}$, $\{x_1, x_2, \dots, x_d\}$ に分割し, 各ブロック T_{ij} と x_j ($j=1, 2, \dots, d$) との類似性から認識結果を求めることができるように以下の点からネットワークの改良を試みる.

- 出力層のニューロンの増加

提案モデルにおいては, 位置的に対応したブロックを組合せ, $(T_{i1}, x_1), (T_{i2}, x_2), \dots, (T_{id}, x_d)$ という順番にネットワークに入力する(図 3).

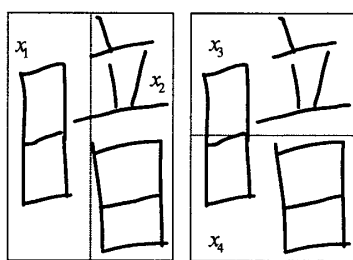


図2 文字パターンの分割

出力層については、分割数 d 個分のニューロンを用意し、 j 番目のニューロンは j 番目に入力する二つのブロック T_{ij} と x_j との類似性を処理する。すなわち文字種 i を対象としたネットワークにおいて、 j 番目に入力するブロック T_{ij} と x_j が同一の構成要素である場合は、出力層の j 番目のニューロンのみを発火させ、それ以外は発火させないように教師信号 t_{ij} を与え、ネットワークを学習する。認識の際は、出力値と教師信号との差の自乗和が最小のネットワークを認識結果とする。

- ネットワークの内部結合の再帰結合化
ネットワークの入力部には、中間層と同数のニューロンを文脈層に配置した再帰結合型とし、文脈層への入力は直前の中間層の状態をコピーする。これはこの文脈層へのコピーによって、以前の入力情報が蓄積されることになり、 x_j を学習する上で x_1 から x_j は同字種のブロックという情報が残ることになると考えたからである。

4. ネットワークの学習方法

個々のニューラルネットワーク（エルマンネット）の学習アルゴリズムを以下に示す。

Step1（非教師パターンの選択）

学習文字パターンを用いて大分類を行ない、各文字種 i ごとに第 P 候補（実験では $P=10$ ）までに出現する文字候補の集合 C_i を求め、文字種 i を認識対象としたネットワークにおける非教師パターンとする。文字候補数を c_i 個とする。

Step2（文字パターンの分割）

文字種 i のテンプレート T_i と非教師パターン x （ C_i に属する）を d 個のブロック $\{T_{i1}, T_{i2}, \dots, T_{id}\}$ 、 $\{x_1, x_2, \dots, x_d\}$ に分割する。

Step3（教師パターンの学習）

入力層の比較対象部にテンプレートの文字パターンのブロック T_{ij} 、認識対象部にテンプレートと同字種の文字パターンのブロック T_{ij} を入力する。文脈層へは直前の中間層の値をコピーする。

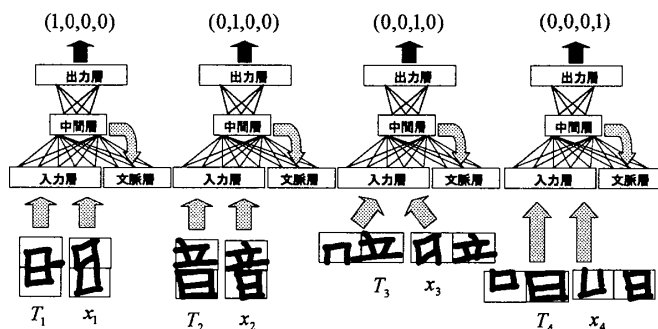


図3 提案するネットワーク

出力層の j 番目のニューロンのみを発火させ、他のニューロンは発火しないように教師信号 t_{ij} を与え、誤差逆伝播則によって学習する。

Step4（非教師パターンの学習）

非教師パターンのブロック x_j （ $x_j \in C_i$ ）を認識対象部に入力し、他はStep3と同様に出力層の全てのニューロンが発火しないように学習する。

5. 認識実験

認識実験のため産業技術総合研究所提供のETL9Bの3036文字を利用した。学習文字として最初の一セット40パターンを利用し、認識用の未知文字として二セット目の40パターンを用いた。文字パターンには非線形化処理を施し、 48×48 の大きさにした後、196次元の輪郭線特徴を抽出し、図2に示す112次元のブロックを入力とした。ネットワークの構造については、入力層のニューロンは112次元のブロックを2個組み合わせているため224個、文脈層と中間層は32個、出力層は分割数と同数である。分割数 d を変えて、学習した結果を表1に示す。

表1 分割数を増やした時の認識率

分割数	学習文字	未知文字
4個	99.39%	92.63%
8個	99.40%	93.75%
12個	99.41%	94.09%
ユークリッド距離	95.08%	92.47%

表1よりブロック数の増加にともない、入力層のニューロン数を増やすことなく認識率の向上が確認できた。

参考文献

[1] 篠沢, 大駒, テンプレートマッチングによるオフライン手書き文字認識ニューラルネットワークの作成, 情報処理学会論文誌, Vol.42, No.1, pp.16-25 (2001)