

二分木を利用した大分類・詳細認識型 文字認識ニューラルネットワークの作成*

4 G-1

篠沢 佳久†

大駒 誠一†

慶應義塾大学理工学部管理工学科§

1 はじめに

大規模な文字認識を行なう場合まず始めに文字パターンをいくつかのグループに分類しそして次に各グループ内で詳細に小規模な認識をするという2段階による認識方法が一般的となっている。しかし大分類部において一回の処理で文字パターンをいくつかのグループに分類することは複雑な分類問題となり結果的に分類率の低下につながる。そこで本報告では大分類に文字パターンを次々と二分していく方法を試みた。文字パターンを二分する小規模なニューラルネットワークを二分木型に配置し適切なグループに分類するのである。一方詳細認識部では従来のバックプロパゲーションアルゴリズムを用いてグループごとに認識を行なう。そして大分類部と詳細認識部を統合化し認識を行なった結果良好な結果を得ることができた。

2 大分類部の二分木化による利点

大分類部を二分木化することによる一番の利点は特徴空間を超平面で切るだけであり、複雑な識別面を設定せずに済むということである。また認識すべき文字数が2倍になっても一つの分類を行なうニューラルネットワークが分類率99%のものを構築できれば全体的な分類率は0.99倍しか低下せずに済むのである。

3 文字パターンを二分する方法

二値化した文字画像に細線化処理、正規化処理を施し特徴の抽出を行なう。本報告で用いた特徴は文字濃度特徴（メッシュパターン：256次元）、端点交点特徴（256次元）、背景特徴（ペリフェラルパターン：400次元）、方向成分特徴（400次元）である。これらの特徴を用いて文字パターンを二つのグループに分類する。文字パターンの分類は次の2段階の操作で行なわれる。

1. 個々の文字パターンが二つのグループのうちどちらに属するのかを定める。

2. 前の処理で決定したグループに属するように文字パターンを正確に分類するニューラルネットワークを構築する。

まず、個々の文字パターンが二つのグループのうちどちらに属するのかを決定する。前記の4つ抽出した特徴の中からどれか一つを選択し学習ベクトル量子化(LVQ)を用いて文字パターンを二つに分類する。図1に示すようにLVQネットは入力層と出力層から構成される2層のネットワークである。LVQネットへの入力を選択した特徴ベクトルであり、選択した特徴ベクトルの次元数だけ入力層のニューロンを用意する。出力層は2個である。一種類の文字パターンに対してN個特徴ベクトルを入力する。LVQにより二つに分類し

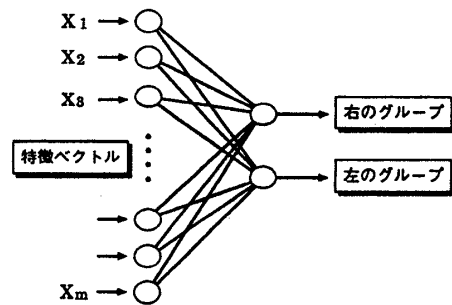


図1: LVQ ネット

た結果まず左右に分かれたグループの各カテゴリ数のバランスがとれているかを調べる。すなわち100個の文字パターンを二分した場合右のグループに1個、左のグループに99個といったように分かれてしまっている場合は二つに分類したとは言えないからである。左右のバランスがとれていない場合は学習係数を変え再び分類を行なわせる。そして次に個々の文字パターンに対してきちんと分類されているかを調べる。一つの文字につきN個のサンプルを使用しているとすると手書き文字のように変形の激しいパターンでは左のグループにN個、右のグループに0個といったように分類されることはまれである。そこでN個中、左のグループにp個、右のグループにq個分類されたとする。p>qの場合p/Nが一定の値以上ならこの文字パターンは左のグループに属するものとし、p/Nが一定の値未満ならばこの文字パターンは左右両方のグループに属するも

*The Construction of Neural Network for the Classification and Detail Recognition using a Binary Tree.

†Yoshihisa SHINOZAWA

‡Seiichi OKOMA

§Faculty of Science and Technology Keio University

のとする。一方 $p < q$ の場合も同様に q/N が一定の値以上ならこの文字パターンは右のグループに属するものとし、 q/N が一定の値未満ならばこの文字パターンは左右両方のグループに属するものとする。このように文字パターンを二分したならば次に右左のグループで同様のことを行なわせ、各グループで処理するカテゴリ数が一定以下になるまで分類を行なう。一つの特徴のみで分類を行なわずと徐々に二分することが困難となるので4つの特徴を繰り返し使用する。

このままではある文字パターンでは N 個のうち p 個はきちんと左のグループに分類されるのだが、残り q 個は右のグループに分類されてしまったままである。そこで N 個全てを左のグループに分類させるネットワークを新たに構築する。図2に示す3層型のバック

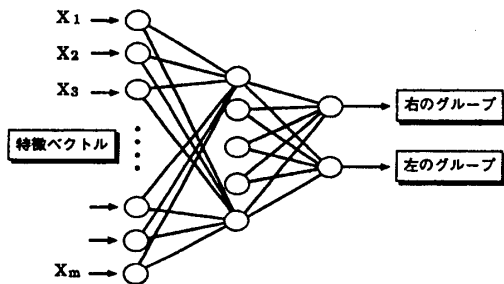


図2: BP ネット

プロパゲーションネットワーク (BP) を利用して前の処理で決定したように分類が可能なネットワークを作成した。このようにまず始めに LVQ によって特徴空間を超平面で二つに切り、次に BP によって詳細に識別面を形成し、文字パターンを適切に分類できるようにしている。

4 詳細認識部の作成

次々と文字パターンを二分していき末端のグループでのカテゴリ数が一定以下になるまで分類した後で各グループ内で詳細に認識を行なう。詳細認識は局所結合型の3層バックプロパゲーションネットワークを用いた。

5 認識結果

個別に作成した大分類部と詳細認識部を統合化し認識を行なわせた。対象とした文字は平仮名 (71 文字)、片仮名 (71 文字)、漢字 (330 文字) を混ぜた計 572 文字である。一文字につき 24 個学習に使用し 8 個未学習文字として認識を行なった。前述した方法で二分木を生成した結果大分類部では 21 のグループに分類された。その分類結果を表1に示す。この結果から図3のように大分類部の第3候補までのグループを選び、選択された3つの各グループごとに詳細認識を行ない、上位候補を選択した。そしてそれらの結果を統合認識して認

識結果を出力するようにした。その結果を表2に示す。認識数は 4576 個である。

表1: 大分類部の分類能力

候補	認識数	正解数	分類率
1	4576	4375	95.6%
2	4576	4523	98.8%
3	4576	4562	99.7%
4	4576	4570	99.8%
5	4576	4572	99.9%

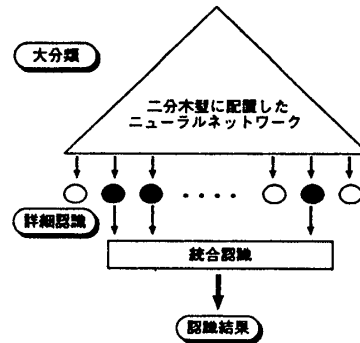


図3: 統合化したネットワーク

表2: 統合化したネットワークの認識率

候補	正解数	認識率	候補	正解数	認識率
1	4310	94.2%	6	4530	98.9%
2	4379	95.6%	7	4532	99.0%
3	4463	97.5%	8	4533	99.0%
4	4516	98.7%	9	4534	99.1%
5	4528	98.9%	10	4534	99.1%

6 まとめ

本報告では大分類部を二分木化することで分類率を向上を試み良好な結果を得ることができた。今後の課題としては二分木生成においてバランス良く文字パターンを分類していくこと、効率良く分類するにはどのような特徴を用いるべきかといったような問題が残されており、今後これらの問題に取り組むと同時にさらに大規模なネットワークの構築する予定である。

参考文献

[1] 岩田 彰他: 大規模ニューラルネット CombNET, 電子情報通信学会 D-II Vol.J73-D-II No.8 pp.1261-1267(1990年8月)

[2] 岩田 彰他: 大規模ニューラルネット CombNET-II, 電子情報通信学会 D-II Vol.J75-D-II No.3 pp.545-553(1992年3月)