接触文字を認識するBPネットの一提案

非会員 井上 充裕* 正員 須崎 健一*

非会員 百原 武敏* 正員 荒屋 眞二*

A Proposal of Back-Propagation Network for Recognizing Contact Character

Mituhiko Inoue*, Non-member, Kenichi Suzuki*, Member,
Takatoshi Momohara*, Non-member, Shinji Araya*, Member

This study proposes a neural network model which can recognize contact characters by only learning the standard characters. This paper describes how to experimentally recognize contact character on standard character and recognize the standard characters from the contact character.

キーワード：接触文字，誤差逆伝播，文字認識

keywords：contact character, back-propagation, character recognition

1. まえがき

現在，最新のOCRソフトは，認識評価データベースの適用と統計的な手法により下線接触文字も認識可能となった。更に，最近の装飾文字やロゴマークなどには，故意に文字の一部を重ねた文字も見受けられる。このような特殊な文字は最新のOCRソフトでも認識した報告例はない。現在，重ね文字を分離できるニューラルネットにネオコグニトロン⁽¹⁾が提案されている。このネットは異種細胞間の多層構造であり，競合学習で重ね文字の分離を行っている。しかし，誤差逆伝播（BP）法による報告例はない。

本稿は特殊な文字の認識例として，文字の一部を故意に重ねた接触文字を認識することができるBPネットを提案する。

2. 提案ネットと学習認識法

本ネットは図1のような3層BPネットの一部（点線内）を用いて，2値で表された基本文字を通常のBP法で学習し，得られた重みとバイアスを未使用の部分ネットに，ある規則に基づいて変換し，ネット全体で接触文字を認識するものである。ネットの各層間は完全結合である。INは入力層で，隠れ層 H_i ($i=0,1,2$)のユニット数 L は，記憶させたい文字数に依存する。出力層 Out_i ($i=0,1,2$)のユニット数 u は，基本文字数に等しい。INから H_i 間の結合の重みを WH_i ($i=0,1,2$)とする。 H_i から Out_i への結合の重みを WO_i ($i=0,1,2$)とする。 H_i と Out_i の各バイアスを BH_i ($i=0,1,2$)， BO_i ($i=0,1,2$)とする。

図1 提案モデルの構造

Fig. 1. Structure of a proposed neural network.

接触文字の例として2文字を取り扱う。いま，基本文字HとDを正方入力メッシュ上で表す。ここで，HとDの文字（列ベクトル）をそれぞれ Ih と Id で表す。次に，位置ずれ接触文字 Ihd を式(1)のように表す。 $\text{sgn}()$ 関数は括弧の中が1以上であれば1となる関数である。

$$Ihd = \text{sgn}(Ih + Id) \quad (1)$$

Ih と Ihd の関係を次式で表す。

$$Ih = M \cdot Ihd \quad (2)$$

ここで， M を位置ずれ接触行列と呼ぶ。図1を用いて学習法を説明する。まず，gateを閉じ，点線で囲んだ部分ネットを用いて，基本文字だけを学習する。学習完了後，gateを開け，ネット全体で認識テストを行う。ここで， Ih を認識テストしたときの $H0$ の入力積和を $S0$ とする（式(3)）。また， Ihd を認識テストしたときの $H1$ の入力積和を $S1$ とする（式(4)）。

$$S0 = WH0' \cdot Ih \quad (3)$$

* 福岡工業大学
〒811-0295 福岡市東区和白東3-30-1
Information Engineering, Fukuoka Institute of Technology

$$S1 = WH1^t \cdot Ihd \dots\dots\dots(4)$$

ここで、 $S0=S1$ となるように $WH1$ を式(5)で求め、隠れ層から出力層までの部分ネットをすべて同じ (式(6)) にすれば、 Ih の認識時の $Out0$ の出力値と Ihd の認識時の $Out1$ の出力値は一致する。

$$WH1 = R \cdot WHO \dots\dots\dots(5)$$

$$BH1 = BH0, BO1 = BO0, WO1 = WO0 \dots\dots\dots(6)$$

ここで、 R を変換行列と呼ぶ。 R は式(3)=式(4)から求める。式(4)の Ihd は式(2)から $M^{-1} \cdot Ih$ と表されるので次式のようにになる。

$$WHO^t \cdot Ih = (R \cdot WHO)^t \cdot M^{-1} \cdot Ih \dots\dots\dots(7)$$

式(7)は、次式のと看、右辺と左辺が等しくなる。

$$R^t \cdot M^{-1} = E \quad (E: \text{単位行列}) \dots\dots\dots(8)$$

式(8)から $R = M^t$ となる。ここで、 M を式(2)から求めると R は次式のようにになる。

$$R = (Ih \cdot Ihd^+)^t \dots\dots\dots(9)$$

$+$ は擬似逆行列を表す。この R を式(5)の右辺に代入し、 $WH1$ を求めれば、 $S0=S1$ となる。ここで、 Ihd は非正則となり、Moore-Penrose の擬似逆行列⁽²⁾を用いた。

次に、 Id と Ihd に着目すれば、 R は $(Id \cdot Ihd^+)^t$ となり、式(5)~(9)の手法を適用し $WH2$, $BH2$, $BO2$, $WO2$ を求めれば、 $S0=S2$ となり、 Id の認識時の $Out0$ の出力値と、 Ihd の認識時の $Out2$ の出力値は等しくなる。つまり、 $Out0$, $Out1$, $Out2$ の認識結果から接触文字の判定ができる。

ここまでは、1組の接触文字の例を示したが、 n 組の接触文字の変換も可能である。式(10),(11)は $n=3$ のときの $R1$ と $R2$ の例である。但し、 Ink , Ihn , Ikh は、それぞれ文字“ N と K ”, “ H と N ”, “ K と H ”の位置ずれ接触文字の縦ベクトルである。

$$R1 = (|In Ih Ik| \cdot |Ink Ihn Ikh|)^t \dots\dots\dots(10)$$

$$R2 = (|Ik In Ih| \cdot |Ink Ihn Ikh|)^t \dots\dots\dots(11)$$

3. 接触文字の認識実験

両文字位置ずれ接触文字の特殊なケースとして、図2のような文字の一部が完全に重なった文字の学習認識実験を行った。ネットの基本構造は図1と同じで、入力層のユニット数400、中間層 H_i の各ユニット数3、出力層 Out_i の各ユニット数3、学習率0.25、慣性係数0.8とした。

まず、3個の基本文字の学習完了後、図2のような両文字とも位置ずれした接触文字の R ($Ih1$, $Ih2$ に対する $R1$, $R2$) を作成し、未使用の部分ネットに3組の重み変換を行い、接触文字の認識テストを行った。結果の一例を表1に示す。表1から、部分ネット(接触文字判定用)の $Out1$, $Out2$ の出力値は $Out0$ の出力値と一致しており、かつ他の出力ユニットの出力値は0.1以下となり認識精度が高く、基本文字同士の接触文字と認識できた。なお、この認識精度は基本文字数を増やし10組の接触文字に対しても同じ傾向が見られた。

また、図2の3組接触文字の上下左右2,4メッシュ位置ず

図2 基本文字と接触文字

Fig. 2. Standard character patterns and contact character patterns.

表1 基本文字と特殊な接触文字の認識テスト

Table 1. Recognition test of standard character patterns and contact character patterns.

test characters	output value of Outi(i=0) layer		
	Out0(1)	Out0(2)	Out0(3)
N	0.944	0.025	0.041
D	0.035	0.951	0.042
H	0.039	0.049	0.943
test characters	output value of Outi(i=1,2) layers		
	Out1(1)	Out1(2)	Out1(3)
N and D	0.944	0.025	0.041
	0.035	0.951	0.042
H and D	0.039	0.049	0.943
	0.035	0.951	0.042
H and N	0.039	0.049	0.943
	0.944	0.025	0.041

れ接触文字(以下、上下左右とメッシュの記述を略する)に対する R を作成し、1,2,3,4位置ずれ接触文字の認識実験を行った。この実験は未学習文字も含んだ認識テストになるため、最大出力値により認識判定を行った。ネットの初期値を5回変更し、学習後に48個(3組×16個)の位置ずれ接触文字に対して認識テストを行い、平均認識率を求めた。その結果、重み変換を行った2,4位置ずれ接触文字は100%、また、未学習の1,3位置ずれ接触文字で95%の平均認識率が得られ、ネットの補間力が大きく現れた。なお、提案ネットは基本文字が増えても、重みとバイアスを1組の部分ネットに変換するだけであるため、小規模で構造も簡素である。

4. ま と め

本稿は基本文字の学習で接触文字が認識できるBPネットを提案し、有効性を示した。なお、本稿は接触文字の認識手法に焦点を絞って議論したので正方メッシュ上の文字に限定した。しかし、事前処理を行い、変換行列を作成すれば変形文字にも適用可能である。また、本手法は一般の重ね文字にも適用できるので、変形(位置ずれ・回転・拡大縮小)重ね文字の認識について研究を行う予定である。

(平成14年11月25日受付)

文 献

- (1) K.Fukushima: "Neural Network model for selective attention in visual pattern recognition and associative recall", Appl. Opt., Vol.26, No.23, pp.4985-4992 (Dec.1987)
- (2) A.Albert: Regression and the Moore-Penrose Pseudo Inverse, Academic Press, New York (1972)