

H-031

## 重ね文字列を認識するニューラルネットワーク

A Neural Network for Recognizing Superimposed Character String

金岡 哲正† 中川 勝博† 井上 慎太郎† 須崎 健一†  
 Akimasa Kanaoka Katuhiro Nakagawa Shintaro Inoue Kenichi Sauzaki

## 1. まえがき

現在の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である。更に、最近の装飾文字やロゴマークなどには、故意に文字列の一部を重ねた文字列も見受けられる。このような特殊な文字列は最新の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、他のニューラルネットワークによる報告例はない。また、重ね文字列を認識することができるニューラルネットワークも報告されていない。

本稿では重ね文字列を認識することができる複写型のエルマンニューラルネットワークを提案し、その有効性について報告する。

## 2. 提案ネットと学習認識法

文字列を学習認識する手法の一つとしてエルマンニューラルネットワークが提案[2]されている。このネットワークは通常の3層BP(誤差逆伝播)ネットにリカレント結合を有したものである。このネットは入力文字を処理する入力層と直前の中間層の状態を入力とする状態層、そして出力層から構成される。また、中間層と状態層間の重みは1であり、重みの学習は順方向の結合についてだけ行われるので、通常のBP法がそのまま使える。

提案ネットは図1のようなネットの一部(点線内)を用いて、2値で表された基本文字列を学習し、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複写し、ネット全体で重ね文字列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 $Out_i$  ( $i=0,1,2$ )のユニット数は、基本文字列数に等しい。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。

本稿では $n$ 文字数からなる文字列を同じカテゴリとして学習するため1文字として取り扱う。ここでは最も簡単な例として1文字同士の重ね文字を取り扱う。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 $I_h$ と $I_d$ で表す。

次に、重ね文字 $I_{hd}$ を式(1)のように表す。 $\text{sgn}(\cdot)$ 関数は括弧の中が1以上であれば1となる関数である。

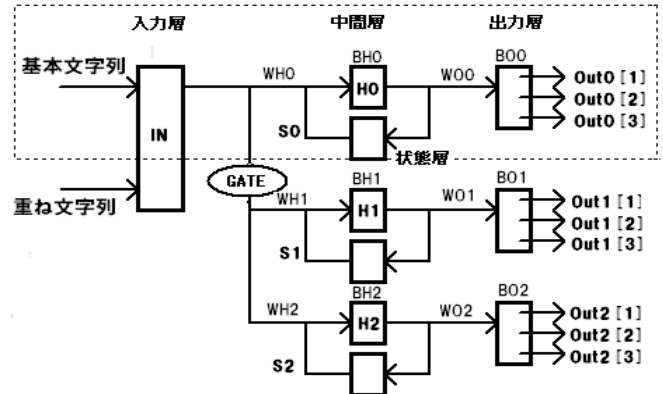


図1. 提案ネットの基本構造

$$I_{hd} = \text{sgn}(I_h + I_d) \quad (1)$$

$I_h$  と  $I_{hd}$  の関係を次式で表す。

$$I_h = M \cdot I_{hd} \quad (2)$$

ここで、 $M$ を重ね行列と呼ぶ。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、 $I_h$ を認識テストしたときの $H_0$ の入力積和を $S_0$ とする(式(3))。また、 $I_{hd}$ を認識テストしたときの $H_1$ の入力積和を $S_1$ とする(式(4))。

$$S_0 = WHO^t \cdot I_h \quad (3)$$

$$S_1 = WH1^t \cdot I_{hd} \quad (4)$$

ここで、 $S_0=S_1$ 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、 $I_h$ の認識時の $Out_0$ の出力値と $I_{hd}$ の認識時の $Out_1$ の出力値は一致する。

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

$$BH1 = BH0, \quad B01 = B00, \quad W01 = W00 \quad (6)$$

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 $I_{hd}$ は式(2)から $M^{-1} \cdot I_h$ と表されるので次式のようなになる。

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

式(7)は、次式のと看、右辺と左辺が等しくなる。

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

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

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

\*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は非正則となるため、Moore-Penroseの擬似逆行列[3]を用いた。

次に、IdとIhdに着目すれば、Rは(IId · Ihd<sup>+</sup>)<sup>t</sup>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の認識結果から重ね文字の判定ができる。

ここまで、1文字同士の重ね文字の例を示したが、n個の文字同士の重ね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”の重ね文字の縦ベクトルである。

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

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

### 3. 重ね文字列の認識実験

文字列の長さに対する学習認識はネット構造に依存する。ここでは基礎実験として、3個の基本文字列の重ね文字列を取り扱う。ネットの基本構造は図1と同じで、入力層のユニット数256、中間層Hiの各ユニット数50、状態層の各ユニット数50、出力層Outi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個の基本文字列(ABC, DEF, GHI)の学習完了後、図2のように重なった文字列のR (IH1, IH2 に対するR1, R2)を作成し、未使用の部分ネットに重み変換を行い、重ね文字列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(重ね文字列判定用)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、3文字列同士の重ね文字列と認識できた。

次に、未学習の重ね文字列の認識テストを行った。この実験では未学習文字列も含んだ認識テストになるため、出力値が0.5以上の最大出力値により認識判定を行った。結果を図3に示す。横軸の上1とは、一方の文字列を1メッシュ上に位置ずれたことを示す。図3の実線は基本文字列同士の重ね文字列を複写学習しただけで、一方の文字列を上下1, 2メッシュ位置ずれた場合の平均認識率の変化を示す。また、点線は基本文字列と上下2メッシュ位置ずれたものと同じカテゴリになるように複写学習し、上下1メッシュ位置ずれた重ね文字列の補間力を示したものである。

なお、提案ネットは基本文字列が増えても、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。



図2. 基本文字列同士の重ね文字列.

表1 基本文字列と重ね文字列の認識時出力値.

		Out0	Out1	Out2
基本文字列	ABC	0.995	0.005	0.003
	DEF	0.005	0.995	0.003
	GHI	0.005	0.005	0.994
重ね文字列	ABC+DEF	0.005	0.995	0.005
		0.005	0.005	0.995
		0.003	0.003	0.003
	ABC+GHI	0.005	0.995	0.005
		0.005	0.005	0.005
		0.003	0.003	0.994
DEF+GHI	0.005	0.005	0.005	
	0.005	0.995	0.005	
	0.003	0.003	0.994	

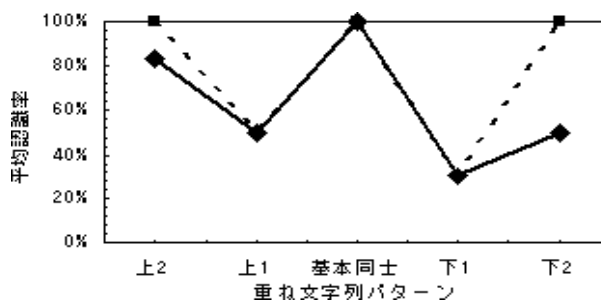


図3. 重ね文字列と認識率の関係

### 4. まとめ

本稿は基本文字列の学習で重ね文字列が認識できる複写型のエルマンニューラルネットワークを提案し、基礎実験を通して有効性について考察した。なお、本稿は重ね文字列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列にも適用できるので、変形(位置ずれ・回転・拡大縮小)重ね文字列の認識について研究を行う予定である。

#### 参考文献

- (1)K.Fukushima.:Neural Network model for selective attention in visual pattern recognition and associative recall, Applied Optics, 26 [23], pp.4985-4992 (Dec.1987).
- (2) Elman, J.L. Finding structure in time. *Cognitive Science*, 14, 179-211. (1990).
- (3)A.Albert.: Regression and the Moore-Penrose Pseudo Inverse, Academic Press, New York,1972.