

複写学習モデルによる文字列の復元

Restoration of Character Strings with a Copy-Learning Model

金 万福† 岡部 成将† 須崎 健一†
 Mannhuku Kin Sigemasa Okabe Kenichi Suzuki

1. まえがき

文字列を認識するニューラルネットとしてエルマンネットワーク[1]が提案されているが、文字欠落や汚れに対する認識率の低下などの問題が多い。本稿は3層誤差逆伝搬法(以下BP)による基本文字の学習結果を利用して、複数の文字列の複写学習法[2]を行い、文字列を認識する複写学習型のニューラルネットを提案する。本稿は文字列の欠落・汚れが複写行列に及ぼす影響を調べ、欠落・汚れ文字列に対する汎化能力について考察する。

2. 複写学習モデル

提案ネットの構造を図1に示す。複写学習の一例として「ATM」、「COP」の2文字列を取り扱う。文字はすべて正方メッシュ上で表す。O、I、LとA、T、Mの文字列(列ベクトル)を I_o 、 I_i 、 I_l と I_a 、 I_t 、 I_m 、文字列(列ベクトル) OILと ATM を I_{oil} と I_{atm} で表す。 I_o と I_{oil} に着目し、両文字の関係を表す。

$$I_o = M \cdot I_{oil} \quad (1)$$

ここで、 M を変換行列と呼ぶ

複写学習法とは、 I_o を基本ネットの中間層 H_o (図1)への入力積和 S_o と I_{oil} を部分ネットの中間層 H_i の入力積和 S_i を同じにするために、以下の各式で各重み、各バイアス WH_1 、 BH_1 、 BO_1 、 WO_1 を求めることである。

$$WH_1^t \cdot I_{oil} = WH_0^t \cdot I_o \quad (2)$$

$$BH_1 = BH_0, BO_1 = BO_0, WO_1 = WO_0 \quad (3)$$

ここで、文献3により、次式を求める。

$$M = I_o \cdot I_{oil}^+ \quad (4)$$

$+$ は擬似逆行列[4]を表す。この M から WH_1 を求めれば、

$$\begin{aligned} WH_{11}^t &= WH_0^t \cdot M_1 \\ WH_{12}^t &= WH_0^t \cdot M_2 \\ WH_{13}^t &= WH_0^t \cdot M_3 \end{aligned} \quad (5)$$

となる。次に I_t と I_{atm} に着目し、次式から WH_1 を求める。

$$\begin{aligned} WH_{11}^t &= WH_0^t \cdot M_1 \\ WH_{12}^t &= WH_0^t \cdot M_2 \\ WH_{13}^t &= WH_0^t \cdot M_3 \end{aligned} \quad (6)$$

認識の際には、 M を次式のように分解して、GATE2を開いて、文字列の数だけ複写と認識を繰り返すことにより1文字目から順にもとめる。

$$\begin{aligned} M_1 &= \begin{bmatrix} I_o & I_a \\ I_i & I_t \\ I_l & I_m \end{bmatrix} \cdot \begin{bmatrix} I_{oil} & I_{atm} \end{bmatrix}^+ \\ M_2 &= \begin{bmatrix} I_i & I_t \\ I_l & I_m \end{bmatrix} \cdot \begin{bmatrix} I_{oil} & I_{atm} \end{bmatrix}^+ \\ M_3 &= \begin{bmatrix} I_l & I_m \end{bmatrix} \cdot \begin{bmatrix} I_{oil} & I_{atm} \end{bmatrix}^+ \end{aligned} \quad (7)$$

式(7)の複写学習を順に行えば I_o 、 I_i 、 I_l および I_a 、 I_t 、 I_m の認識時の O_0 の出力値と I_{oil} および I_{atm} の認識時の O_1 の出力値は一致する。これにより、文字列を認識することができる。

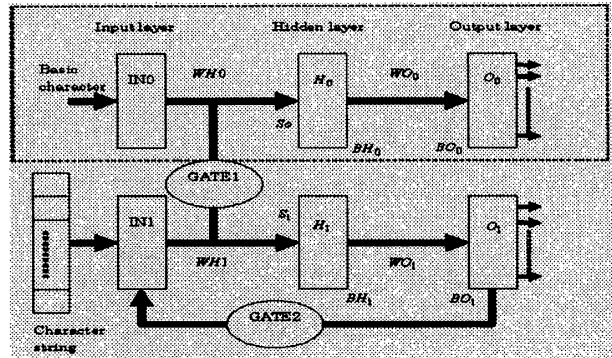


図1. 提案ネットの基本構造

しかし、これは入力パターンに汚れがある場合は、認識率は低下する。そこで閾値 θ を設けて、時刻 $d=0,1,2,\dots$ における入力ユニットへの入力を $U(t)$ で、 M_i との積を A_i とし、次式で表す。 S_0 と θ の関係を式(9)と(10)で表す。

$$A_i = M_i \cdot u(t) \quad (8)$$

$$S_0 = WH_0^t f_i(t) \quad (9)$$

$$f_i(i) = \sum (A(i) - \theta) \quad (10)$$

認識結果が判定基準に満たない場合、

$$f_i(t+1) = \sum f_i(t) - \theta \quad (11)$$

として、判定基準を満たすまで繰り返す。ここで、 d を式(12)で求め、

$$d = \sqrt{\sum (f_i(t) - f_i(t-1))^2} \quad (12)$$

d の値が基準値より大きい場合は $\theta=0$ で、小さくなると $f_i(t)$ に変化を付けるために θ を増す。ただし、繰り返す回数に上限を設けて上限を超えると、認識不能とみなす。

† 福岡工業大学大学院情報工学専攻, 福岡市

3. 実験概要

複写学習モデルの有効性を考察するために、欠落(劣化)・汚れを受けた低品質文字列の認識実験を行う。4-(1)1文字目から4文字目の間で順番に1文字を完全欠落させた場合、4-(2)文字列全体(入力メッシュ全体)に劣化が生じる場合、4-(3)文字列全体に汚れが生じる場合の認識実験を行う。実験条件は、各層の重みとバイアスを乱数により値を決め、学習係数 $\epsilon=0.25$ 、慣性係数 $\alpha=0.80$ とし、基本文字をBP法により出力層の各ユニットの(最小二乗)出力誤差 0.001 以下になるまで学習を行った。基本文字は、 20×20 の正方メッシュ上に 0.1 の 2 値で表現された手書き漢字文字 18 個である。学習した文字列は長さを 4 とし、文字列数を最大 8 個とした。認識判定条件は、望ましい出力ユニットの出力値が 0.7 以上、かつ、二番目に大きい出力値との差が 0.3 以上のとき文字を認識したとする。そして、文字列のすべての文字が同条件を満たしたとき文字列を認識したとする。

4. 低品質文字列認識実験

(1) 文字が完全欠落した場合の認識実験



図 2. 先頭文字が欠落した文字列

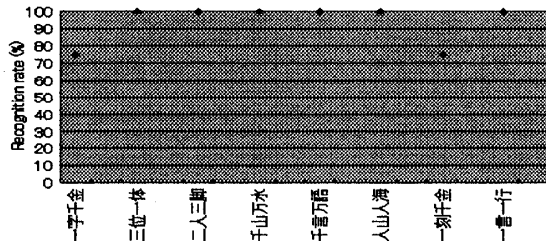


図 3. 文字列数別の平均認識率

(2) 文字列全体に欠落が生じた場合の認識実験



図 4. 30%の欠落文字列

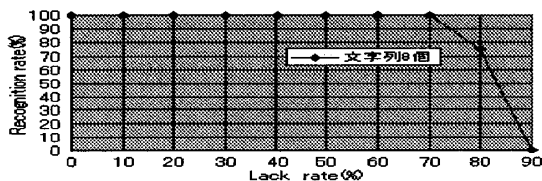


図 5. 文字列の欠落率と平均認識率

(3) 文字列全体に汚れが生じた場合の認識実験

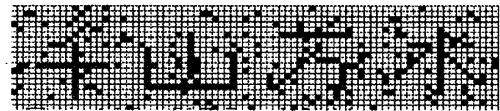


図 6. 10%の汚れ文字列

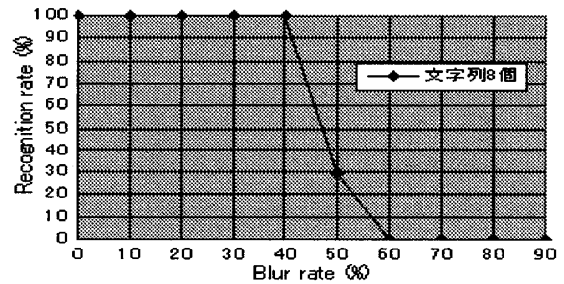


図 7. 文字列の汚れ率と平均認識率

図 3 から、文字列に依存するがほぼ 100% の認識率が得られた。また、図 5、7 より文字列全体の欠落や汚れに対しても高い認識能力があることがわかった。

5. まとめ

本研究では、低品質文字列の認識を行う複写学習法を提案し、その手法を導入したニューラルネットワークを構築し学習認識実験を行った。その結果、(1) 1文字の完全欠落は 100% に近い復元能力を確認したが、2文字の完全欠落の場合は大幅に低下した。(2) 文字列全体の欠落や汚れに対する認識実験で、高い認識能力をもつネットワークであることを確認した。現在、文字列の復元や耐汚れなどに対する他のネットワークを調査中であるが、提案ネットより優れた報告は見当たらない。今後は文字列の欠落や汚れ変化に対する復元アルゴリズムを改良する予定である。

参考文献

- [1] J.L. Elman: Distributed representations, simple recurrent networks, and grammatical structure, *Machine Learning*, Vol.7, pp.195-225, 1991.
- [2] 須崎他: 重ね文字を認識する複写学習モデル, *情報処理*, Vol.44, No.4, 2003.
- [3] 井上他: 単語を認識する複写学習ニューラルネットワーク, *電気関係学会九州支部連合大会*, 13-2A-14, 第 57 回(2004).
- [4] A.Albert: *Regression and the Moore-Penrose Pseudo Inverse*, Academic Press, New York, 1972.