

シヨートノート

位置ずれ・回転パターンを認識するニューラルネットワーク

須崎 健一[†] 荒屋 眞二[†] 中村 良三^{††}

本稿は位置ずれと回転パターンを認識する3層のニューラルネットワークモデルを提案する。このモデルは、従来モデルとは全く異なる複写という学習機構を導入している。本モデルは三つの処理からなる。初めに、3層ネットワークの一部分だけを用いて基本パターンだけを学習する。次に、学習で得られた重みとバイアスの値を位置ずれ・回転させ、未使用ネットに複写する。複写終了後、ネット全体で入力パターンを認識する。提案モデルは、従来モデルと比べ、学習認識時間が大幅に短縮され、かつ従来モデルでは不可能だった回転角と位置ずれの大きさが得られる。本モデルの有効性が実験を通して確かめられる。

A Neural Network for Recognizing Translated and Rotated Patterns

KENICHI SUZAKI,[†] SHINJI ARAYA[†] and RYOZO NAKAMURA^{††}

This paper proposes a 3-layer neural net model that can recognize translated and rotated patterns by learning only standard patterns. This model introduces a new copy-learning mechanism different from conventional models. The proposed model combines three types of processing: first, it learns only the standard patterns by using a part of the 3-layer neural net; secondly, it translates and rotates the values of obtained weights and biases and copies them into an unused part of the net. After copy learning, all parts of the network are used to recognize unknown input patterns. In the proposed model, the following merits are found: its performance is superior and cost is lower than the conventional models. The proposed model can obtain information of rotation angle and translation that the conventional models can not obtain. The function and features of the proposed model is confirmed by simulation experiment.

1. ま え が き

最近、ニューラルネットワークによる位置ずれ・回転文字認識に関する研究が多く報告されている。ニューラルネットワークで位置ずれ・回転を認識するモデルには、Widrowらの多層ネットワークモデルの手法^{1),2)}と、数学的な前処理を施し、位置ずれ・回転に不変な特徴量を抽出し、それをネットに記憶させる手法³⁾に大別できる。本稿では、従来モデルとは全く異なる手法、つま

り取り扱うすべてのパターンを学習させるようなWidrowらの手法や面倒な数学的な前処理も要せず、小規模なネットと簡素な学習法で位置ずれ・回転を認識する手法を扱う。すでに著者らは、新しい学習機構を導入した、次のようなモデルに取り組んできた。(1) 制約付きBP (誤差逆伝播法) モデル⁴⁾, (2) パターン回転モデル⁵⁾, (3) 複写モデル^{6),7)}。これら3種のモデルは、回転パターンだけの認識を扱ったものであるが、簡単に位置ずれの認識にも適用できる。

本稿では、位置ずれと回転の認識に複写モデルを用いる。複写モデルとは、通常の3層BPネットの一部分だけを用いて、基本パターンだけを学習し、得られた重みとバイアスの値を、未使用の部分ネットに、ある規則に基づいて複写し、ネット全体で入力パターンを認識するものである。複写モデルは従来モデルと比べ、格段に学習並びに認識時間が短縮され、かつ、

[†] 福岡工業大学工学部情報工学科
Department of Communication and Computer Engineering, Faculty of Engineering, Fukuoka Institute of Technology

^{††} 熊本大学工学部電気情報工学科
Department of Electrical Engineering and Computer Science, Faculty of Engineering, Kumamoto University

従来モデルでは不可能だった回転角度と位置ずれの大きさも得られる。

2. 複写モデル

2.1 ネットワーク構造

本モデルの構造を図1に示す。入力層は N^2 個のユニットからなり、入力パターンを表現するための $N \times N$ の正方メッシュに対応する。隠れ層は、想定する位置ずれ数と回転パターン数を合わせた数 k 個のユニットグループに分かれる。隠れユニットグループのユニット数 L は記憶させたいパターン数に依存する。出力層のユニット数 M は基本パターン数に等しい。入力層から隠れ層の一つのユニットグループへの重みを IH_i とする。 IH_i は $N^2 \times L$ 次元のベクトルとなる。隠れ層 H_i から出力層 OUT への結合の重みを HO_i ($i=1, 2, \dots, k$) とする。 HO_i は $L \times M$ 次元のベクトルとなる。隠れユニットグループのバイアスを BH_i ($i=1, 2, \dots, k$) とする。

2.2 複写学習と認識

複写の手法をわかりやすくするため、基本パターンが上下左右に1ユニットだけの位置ずれと90度ごとの回転を扱う。したがって、 k は20となる。学習時には、図1の $IH_2 \sim IH_{20}$ に設けたゲートが閉じられ、 IN, H_1 , および OUT からなる部分ネットが、BP法により基本パターンだけを学習する。この部分ネットの学習が完了すると、次の(S1)~(S3)のような複写を実行する。なお、説明に図2に示すような、 $N=3, L=1$ の場合の IH_i 相互関係を用いる。

(S1) IH_1 の要素 ($W_1 \sim W_9$) を IH_5 に、図2に示すように、上に1ユニットだけずれ込んだように複写する。同様に、 IH_1 の要素を IH_{13} に、左に1ユニットだけずれ込んだように複写する。すなわち、1ユニットだけ上と左に対して位置ずれしたとき、対応する要素の値が等しいことを記号 # (translationally equal) で表せば、(1)式ようになる。

$$\left. \begin{array}{l} IH_1 \# IH_5 \# IH_9 \# IH_1 \\ IH_1 \# IH_{13} \# IH_{17} \# IH_1 \end{array} \right\} \quad (1)$$

(S2) $IH_1 \sim IH_4$ 間の要素 ($W_1 \sim W_9$) は、互いに右90度の回転に応じて、図2のように、ずれ込んだように複写する。同様の方法で、 $IH_5 \sim IH_8, IH_9 \sim IH_{12}, IH_{13} \sim IH_{16}, IH_{17} \sim IH_{20}$ 間も複写する。すなわち、右90度回転しとき対応する要素の値が等しいことを記号 @ (rotationally equal) で表せば、(2)式ようになる。

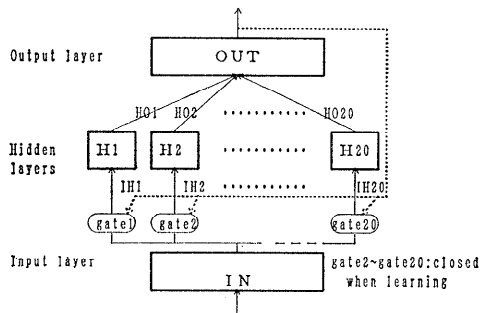


図1 位置ずれ・回転パターンを認識する複写学習 (逐次認識) モデル

Fig. 1 Copy-learning (sequential-recognizing) models for recognizing translated and rotated patterns.

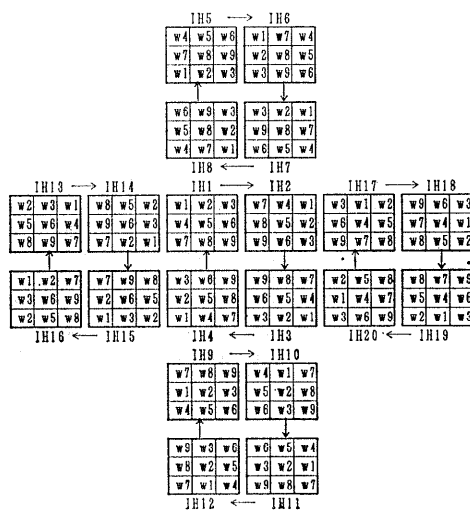


図2 入力・隠れ層間の位置ずれ・回転等価性

Fig. 2 Translational and rotational equality of input-hidden connections ($N=3, L=1$).

$$\left. \begin{array}{l} IH_1 @ IH_2 @ IH_3 @ IH_4 @ IH_1 \\ IH_5 @ IH_6 @ IH_7 @ IH_8 @ IH_5 \\ IH_9 @ IH_{10} @ IH_{11} @ IH_{12} @ IH_9 \\ IH_{13} @ IH_{14} @ IH_{15} @ IH_{16} @ IH_{13} \\ IH_{17} @ IH_{18} @ IH_{19} @ IH_{20} @ IH_{17} \end{array} \right\} \quad (2)$$

(S3) HO_1 と BH_1 を(3), (4)式のように複写する。

$$HO_i = HO_1 \quad (i=2, 3, \dots, 20) \quad (3)$$

$$BH_i = BH_1 \quad (i=2, 3, \dots, 20) \quad (4)$$

一方、認識時は部分ネット単位で逐次的に認識処理を行う。つまり、 $j=1, 2, \dots, 20$ の順で、ゲート j だ

け開かれ、IN, H_j, OUT からなる部分ネットが認識を行う。出力層の一つのユニット出力値が1に近く、残りすべてのユニット出力値が0に近い値が出力された時点で認識処理を停止する。どのゲートの切り替えで認識に成功したかによって、入力パターンの位置ずれの大きさと回転角度がわかる。

3. 実験の目的と内容

実験の目的は、提案モデルの複写効果の検証と未学習パターンに対する汎化力を明示化することである。また、性能比較のため、従来モデルの実験も行う。

まず、複写モデルは、複写を行った後、位置ずれと90度ごとの回転パターンを認識させた。また、5度ごとに回転させた未学習パターンも認識させた。

さらに、複写モデルと従来モデルを公平に比較するため、パターンを直接に扱っている Widrow らのモデルと通常の3層 BP モデルを取り上げた。Widrow らのモデルには、基本パターンだけを提示し、それらの位置ずれと90度ごとのスクランブルパターン^{4), 5)}を学習並びに認識させた。通常の3層 BP モデルには、取り扱うすべての位置ずれと90度ごとのパターンを学習並びに認識させた。

各実験では、まず、各モデルに対して、8×8の正方メッシュで表現された2値化パターンを用いて位置ずれと回転パターンの学習並びに認識テストを行った。次に、複写モデルの未学習パターンに対する汎化力を調べるために、パターンの値を多値化(m)し、正方メッシュ($N \times N$)の大きさも変更した。なお、パターンは基本が英字26文字で、それらを位置ずれ・回転させたものである。未学習パターンに対する認識テストは、1文字でも位置ずれと回転を行うと膨大な数となるため、任意の10文字を選択した。そのときの認識判定は入力パターンの回転角度に最も近い出力ユニットに最大値が現れた場合を認識、そうでない場合を誤認識とした。実験はニューラルネットワークシミュレータ SunNct を用いて S-4/LC 上で行った。

パラメータとしては、学習率0.25、モーメントム0.9、重みとバイアスの初期値の分散0.5とした。学習終了判定は、全訓練パターンに対して平均誤差自乗が0.001以下に達したときとした。

4. 実験結果と考察

複写モデルと Widrow らのモデルに回転だけを適用した出力結果はすでに報告^{6), 7)}済みである。本稿では、その結果に位置ずれを加えただけであるので省略する。また、Widrow らのモデルのネット規模や性能などの詳細な説明も文献^{2), 4), 5)}に譲る。本稿では、本実験で得られた結果を基に要点を述べる。

複写モデルは学習パターンと位置ずれ・回転したパターン間の重なり具合によっては、誤認識が発生し、認識率は約70%まで低下した。その要因は、ネットに正しく認識すべきパターンと、誤認識したパターンとの区別が十分に学習されていないからである。そこで、誤認識が発生すれば、そのネットワーク状態から次のような追加学習^{6), 7)}を行った。

認識すべきパターンと誤認識したパターンのどちらか一方を基本パターンに追加して、IN, H1, OUT からなる部分ネットで追加学習を行う。このとき、追加したパターンに対する教師パターンは、出力層の複写ユニットに対して、なるべく0に近い値を与える。もし追加パターンが複数個あれば、教師パターンが重ならないように与える。追加学習が完了したなら、再び複写を行う。追加学習で完全に誤認識は解消され^{6), 7)}、Widrow らのモデルと同様の認識率を達成することができた。追加学習に要したステップ数は、基本パターンの学習に比べ、約1/10程度であった。なお、次回からの学習の際には、パターンのフォントやネットワークの初期条件を変えなければ、初めから追加すべきパターンを基本パターンの中に入れて学習すればよいから追加学習は不要となる。誤認識が完全に解消された複写モデルは、どのゲートを切り替えたとき、1に近い出力値が得られたかによって位置ずれの大きさと回転角度を知ることができた。

次に、複写モデルの未学習パターンに対する認識率を図3に示す。認識率はネットワークの初期値を5回変更して、その平均値とした。 m と N の値は図上に

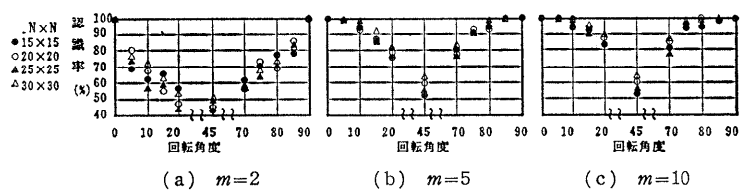


図3 未知(回転)パターンと認識率の関係
Fig. 3 Relationship between unknown (rotated) patterns and recognition rate.

示した。同図(a)から、 m が2で、5度以下の回転では、パターンの移動はほとんどなく、100% の認識であったが、5度以上の回転では認識率は大幅に低下した。これは5度付近でパターンの移動が生じ、基本パターンとの重なりが減少し、汎化力が弱まったためである。 N と認識率の関係は、本実験で用いた N 程度ではほとんどなかった。同図(b)(c)から、 m が5以上になれば、多値化の影響が大きく現れ、認識率は全般的に向上した。また多値化されたパターンも、 N の値にはほとんど影響を受けない結果となった。これらの結果から、 m を大きくすれば複写の効果が高まり、10度の回転までなら約95~100%の認識率を達成できることがわかった。なお、認識率が45度を中心に左右対象となっていないのは、用いたパターンのフォントの影響のためである。図3の認識率は、基本と90度の両パターンを学習並びに複写学習しているため、基本パターンだけの学習に対する回転パターンの認識率⁹⁾と比べると、全般的に向上している。もし、通常の3層BPモデルで任意の回転パターンを認識できるようにするためには、約10度ごとの回転パターンを同じカテゴリとして学習させなければならぬ報告⁹⁾がある。これはパターン作成の手間や学習

に膨大な時間を要する。これに比べ、複写モデルは、複写角度を小さくし⁹⁾各部分ネットの汎化力を高めれば、その間の回転パターンも認識できるため、パターン作成や学習時間が大幅に短縮される。

次に、“H”、“I”、“O”、“N”、“X”、“Z”の6個のパターンは、パターンのフォントによっては、位置ずれや回転を受けると、自ら、または他の文字と完全に重なる場合がある。このようなパターンに対する学習は、Widrowらのモデルでは、大きな誤差が残り、認識時の出力値は大幅に低下する。すなわち、カテゴリの列挙すら不可能となる^{4),5)}。一方、複写モデルでは、基本パターンだけの学習で済むため、誤差は残らず収束する。しかし、位置ずれと回転角度の認識は判定できない。これは人間でも同様である。

5. 従来モデルとの比較

複写モデルと従来モデルのネット規模と性能比較を表1に示す。各ネットの規模は、8×8の正方メッシュで表現された2値化の回転と位置ずれパターンを扱ったときのものである。表1で示した隠れユニット数は、各モデルの訓練パターン数が異なるため、学習が最も早く完了したときのものを用いた。なお、複写モデルで逐次認識とは、図1のモデルであり、並列認識^{6),7)}とは、出力層を k 個用意し、認識時のゲート切り替えをなくし、認識時間を短縮させたものである。

表1から、複写モデルは従来モデルに比べ、ネット規模、学習並びに認識時間が格段に改善されている。また、表には示していないが、複写モデルは、著者が提案したモデル^{4),5)}の中でも最も優れている。

表1 各ニューラルネットワークモデルの比較
Table 1 Comparison of respective neural network models.

項目	複写学習モデル		Widrow らの スクランブルモデル	通常の3層 BP モデル
	並列認識	逐次認識		
ユニット数				
入力層	$N^2: 64$	$N^2: 64$	〈固定ネット部〉 $N^2: 64$	$N^2: 64$
隠れ層	$20L: 400$	$20L: 400$	〈可変ネット部〉 4層 $L: 20$	$20L: 20$
出力層	$20M: 420$	$M: 26$	5層 $M: 36$	$M: 26$
ユニット合計	984	490	1,464	490
結合数	36,000	36,000	119,200	36,000
訓練パターン数	26	26	26	520
追加パターン数	27	27	—	—
平均学習時間 (s)	40	40	1,800	11,940
追加学習時間 (s)	6	6	—	—
複写時間 (s)	4	4	—	—
追加複写時間 (s)	4	4	—	—
平均学習ステップ数	145	145	246	104
追加学習ステップ数	10	10	—	—
平均認識時間 (ms)	4	40	35,000	103

6. むすび

本稿では、従来モデルと全く異なった複写学習という新しい学習機構を導入して、位置ずれと回転パターンを認識するモデルを提案した。また、従来モデルと比べ、性能的に格段に優れていることも示した。

現在のところ、正方メッシュ

で表現された任意の回転パターンを数学的な前処理なしで認識できるモデルは実現されていない。複写モデルは、基本的には入力メッシュを回転させ、回転前と後の重なり具合によって、重みを修正する簡単な手法であるため、任意の回転パターンにも対応できる。現在、複写モデルに回転マトリックス⁹⁾を導入し、90度倍数以外の角度を複写学習するネットモデルを構築している。基礎実験の段階では、良好な結果が得られている。今後は、何度刻みの複写を行えば、その間の回転パターンが認識できるかについて検討する予定である。

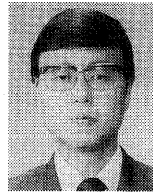
謝辞 本研究は平成4年度文部省科学研究費の補助によった。

参 考 文 献

- 1) Widrow, B., Winter, R. G. and Baxter, R. A.: Layered Neural Nets for Pattern Recognition, *IEEE Trans. Acoust., Speech & Signal Process.*, Vol. ASSP-36, No. 7, pp. 1109-1118 (1988).
- 2) 福見, 細川, 大松: ニューラルネットワークによる位置ずれ・回転に不変なパターン認識システム, *電気学会論文誌C*, Vol. 110-C, No. 3, pp. 148-155 (1990).
- 3) 勘解由, 大西, 杉江: ニューラルネットによる回転と大きさに対して不変な手書き数字認識, *信学論*, Vol. J75-D-II, No. 11, pp. 1900-1908 (1992).
- 4) 荒屋, 須崎, 麻生: 回転パターンを学習認識するニューラルネットワーク, *電気学会論文誌C*, Vol. 111-C, No. 5, pp. 202-208 (1991).
- 5) 須崎, 荒屋, 中村: ニューラルネットワークによる回転パターンの学習認識, *電気学会論文誌C*, Vol. 112-C, No. 9, pp. 561-567 (1992).
- 6) 須崎, 荒屋, 中村: BP学習と複写学習の組み合わせによる回転パターンと回転角の学習認識, *情報処理学会人工知能研究会報告*, 77-18, pp. 159-167 (1991.7).
- 7) Suzuki, K., Araya, S. and Nakamura, R.: Three-layer Neural Network Models for Rotated Patterns Recognition, *Proceedings of the IFIP 12th World Computer Congress*, Vol. 1, pp. 667-673 (1992).
- 8) 須崎, 荒屋, 中村: 回転パターンに対するBPネットワークの汎化能力に関する実験, *電気学会論文誌C*, Vol. 113-C, No. 5, pp. 366-367 (1993).
- 9) 須崎, 荒屋, 中村: 任意に回転したパターンを認識する複写学習モデル, *信学会春季大会論文集*, D-41 (1993).

(平成5年1月20日受付)

(平成5年9月8日採録)



須崎 健一 (正会員)

昭和47年日本大学理工学部電気工学科卒業。昭和52年東京電機大学大学院修士課程修了。昭和47年防衛大学校助手。昭和55年福岡工業大学講師。平成元年同助教授。ニューラルネットの研究に従事。電子情報通信学会、電気学会各会員。



荒屋 眞二 (正会員)

昭和47年東北大学工学部通信卒業。同年三菱電機(株)入社。昭和60年福岡工業大学助教授。平成元年同教授。工学博士(東大)。人工知能の研究に従事。昭和57年電気学会論文賞受賞。電気学会、電子情報通信学会、人工知能学会、IEEE各会員。



中村 良三 (正会員)

昭和39年防衛大学校応用物理専攻卒業。昭和43年熊本大学大学院電気工学専攻修士課程修了。中部電力(株)を経て、昭和50年熊本大学工学部勤務。現在、同工学部電気情報工学科教授。工学博士。データ構造とアルゴリズムの設計と解析、計算機言語、推論処理などに興味を持つ。電子情報通信学会、電気学会各会員。