

大脳視覚野の情報処理に基づく階層型 SOM の パターン認識問題への適用

加藤 聡[†] 堀内 匡[†] 下平 博[‡] 堀口 進[‡]
[†] 松江工業高等専門学校 [‡] 北陸先端科学技術大学院大学

1. はじめに

大脳視覚野の第 1 次視覚野 (V1 野), および下部側頭葉皮質 (IT 野) における生理学的知見⁽¹⁾⁽²⁾ から, 視覚野における物体の形状分類は, V1 野が抽出した線分の情報を IT 野が統合することによって行なわれていると考えられている.

我々はこれらの知見に基づき, V1 野から IT 野に至る神経結合モデルである V1-IT SOM を提案している⁽³⁾. このモデルは, Kohonen の自己組織化マップ (SOM)⁽⁴⁾ を階層的に用いたものである. 本稿では, パターン認識問題の 1 つである手書き数字認識を取り上げ, V1-IT SOM の適用を行ない, V1 野から IT 野への階層的な視覚情報処理に基づいた階層型 SOM の有効性について議論する.

2. V1-IT SOM の概要

2.1 ネットワーク構造

V1-IT SOM の構造を図 1 に示す. 入力層には $n \times n$ 個の入力ユニットが 2 次元格子状に配置され, 第 1 競合層および第 2 競合層には $m \times m$, $l \times l$ 個のノードがそれぞれ 2 次元格子状に配置されている. 入力層では, 各ユニットが 0.0~1.0 までの値をとることで $n \times n$ 画素の視覚イメージを表現する. 第 1 競合層は V1 野に, そして第 2 競合層は IT 野にそれぞれ相当し, 第 1 競合層を V1 層, 第 2 競合層を IT 層と呼ぶ.

V1 層の各ノードは, それぞれ部分的に重複する局所的な受容野を持っている. これは, V1 層に対して入力パターンの部分的な特徴を学習させるためである. 入力層のユニット i と V1 層のノード j には結合重み w_{ij} が与えられている. V1 層と IT 層は完全結合しており, V1 層のノード j と IT 層のノード k には結合重み w_{jk} が存在する.

2.2 学習方式

V1-IT SOM における入力パターンの学習は, 以下に示す 2 つの段階に分けられる. このとき, V1 層あるいは IT 層の各ノードにおける結合重み w_{ij} および w_{jk} の

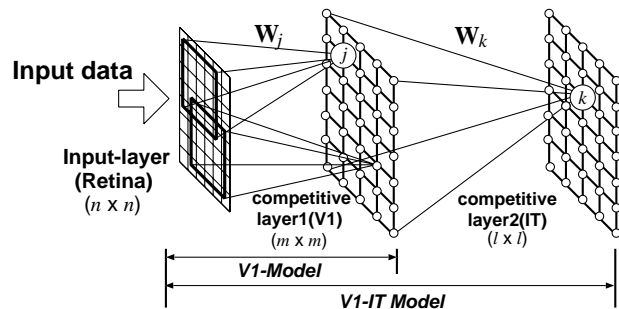


図 1: V1-IT SOM の構造

更新には, Kohonen の自己組織化マップにおける学習アルゴリズム⁽⁴⁾ を用いる.

第 1 段階

V1 層に対して, 入力パターンが持つ局所的な特徴の抽出機能を持たせるためのものであり, 図 2(a) に示す V1 層に対して行なわれる. V1 層の各ノードは, 局所的な受容野を通して呈示された入力パターンの部分的構造に応じて, 結合重み w_{ij} を適切に調節し, 入力パターンの部分的特徴を学習する.

第 2 段階

IT 層に対して, 個々の入力パターンに対する選択性を持たせるためのものであり, 図 2(b) に示す 3 層構造の V1-IT SOM に対して行なわれる. このとき, V1 層は特徴抽出層として用いられ, 新たな学習をしない.

V1 層の役割は, 入力層に与えられた入力信号から何らかの出力信号を生成し, IT 層に送り出すことである. V1 層が IT 層に与える入力信号は, 図 2(b) に示すように, 入力層に手書き数字が呈示された時の, V1 層各ノードの反応値からなる反応パターンである. IT 層は, V1 層の反応パターンを入力ベクトルとして w_{jk} を更新し, 学習を行なう.

3. V1-IT SOM による手書き数字認識

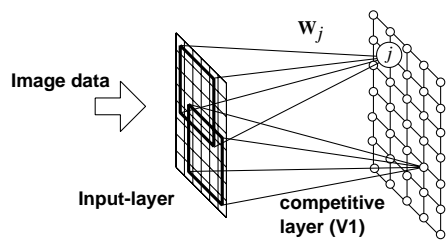
3.1 手書き数字データベースおよび SOM の構成

認識実験には, 手書き数字データベースである MNIST handwritten digits database⁽⁵⁾ を用いた. 60,000 字分の訓練用サンプルから, 10,000 字を学習用としてランダムに抽出し, 同様に 10,000 字をテスト用として抽出

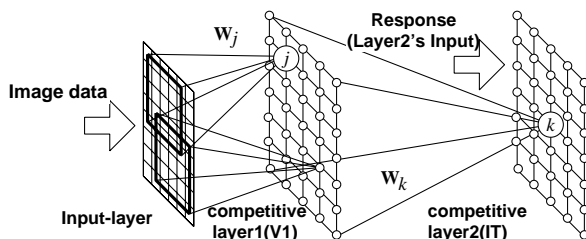
A Study on Pattern Recognition by Hierarchical SOM derived from Information Processing of Human Visual System

[†] Matsue National College of Technology

[‡] Japan Advanced Institute of Science and Technology



(a) 第1段階：V1層における入力パターンの部分的特徴の学習



(b) 第2段階：IT層における入力パターンの学習

図 2: V1-IT SOM の学習方式

した。

V1-IT SOM では、入力層に 30×30 個のユニット、中間競合層に 10×10 個、出力競合層に 20×20 個のノードをそれぞれ格子状に配置した。V1-IT SOM の場合、実際に手書き数字の学習を行なうのは IT 層であり、その際の特徴空間の次元数は V1 層のノード数と等しく 10×10 次元である。従って、パターン認識に用いられる競合層における特徴空間の次元数を揃えるために、基本 SOM では入力層のサイズを 10×10 、競合層のサイズを 20×20 とした。基本 SOM による学習時には、 10×10 画素に解像度を落とした文字パターンが入力層に呈示される。基本 SOM および V1-IT SOM の第 1 段階と第 2 段階における学習回数はいずれも 1,000,000 回とした。また、学習時の近傍半径 d および学習率 α の初期値は、基本 SOM および V1-IT SOM の第 2 段階においては $d_{ini} = 2$ 、 $\alpha_{ini} = 0.2$ とし、V1-IT SOM の第 1 段階では $d_{ini} = 5$ 、 $\alpha_{ini} = 0.2$ とした。

表 1: 手書き数字に対する認識率の比較

	Training accuracy(%)	Testing accuracy(%)
V1-IT SOM	93.28	92.98
Basic SOM	92.39	91.89

3.2 学習終了後の文字認識手法および認識実験結果

学習およびラベル付けが終了した後に、テスト用の文字データに対する認識率を求める。認識には、最近傍決定則による認識手法 (1-NN 法) を用いた。このとき、基本 SOM の競合層、あるいは V1-IT SOM の IT 層の各ノードが持つ重みベクトルが、認識クラス決定のためのプロトタイプとなる。今回の実験では、いずれの場合もプロトタイプ数は 400 個である。

表 1 に、基本 SOM および V1-IT SOM それぞれにおける、学習用およびテスト用サンプルに対する認識率を示す。V1-IT SOM では、基本 SOM の場合に比べて若干良い認識率が得られている。このことは、V1-IT SOM において入力信号の部分的特徴を利用することがパターン認識に有効に作用している可能性を示唆している。

4. むすび

本稿では、視覚情報の部分的特徴を学習する V1 層と、それらを統合する IT 層から構成される V1-IT SOM に対して、手書き数字認識問題への適用を試みた。認識実験の結果から、V1-IT SOM は基本 SOM よりも若干ではあるが精度の高い認識が可能であることが分かった。

関連研究として、Wu らによる階層型 SOM の提案および手書き数字認識への適用があげられる⁽⁶⁾。Wu らの手法は、第 1 層の SOM で特徴空間を粗く分割しておき、第 2 層の SOM で適応的に細かい分割を行なうことで、入力パターンの効率的な学習を行ない、認識精度の向上や、認識処理の高速化を図るものである。

現在、Wu らの階層型 SOM と V1-IT SOM に対して、パターン認識問題における基本性能の比較を行っており、その結果については当日発表する予定である。

文 献

- (1) Hubel, D.H. and Wiesel, T.N. "Receptive fields and functional architecture of monkey striate cortex", *J. Physiol., Lond.*, Vol.195, pp.215-243, 1968.
- (2) Tanaka, K. "Neural Mechanisms of Object Recognition", *Science*, Vol.262, pp.685-688, 1993.
- (3) 加藤 聡, 堀内 匡, 下平 博, 堀口 進, "大脳視覚野の情報処理に基づいた階層型 SOM に関する考察", 電気学会 電子・情報・システム部門大会講演論文集, pp.1024-1029, 2003.
- (4) T. Kohonen, "Self-Organizing Maps" Springer-Verlag Berlin Heidelberg, 1995.
- (5) Yann LeCun, "The MNIST database of handwritten digits" <http://yann.lecun.com/exdb/mnist/>
- (6) Jing Wu, Hong Yan and Andrew Chalmers, "Handwritten digit recognition using two-layer self-organizing maps." *International Journal of Neural Systems*, Vol.5, No.4, pp.357-362, 1994.