

SOINN を導入したノンパラメトリック深層畳み込み自己組織化マップの提案

日大生産工 (学部) ○高橋 秋将 日大生産工 山内 ゆかり

1. まえがき

画像認識の分野で、深層畳み込みニューラルネットワーク (Deep Convolutional Neural Networks: DCNN)[1]は近年主流の研究対象とされている。DCNNは非常に優れた認識精度を誇るが、一方で膨大な教師ラベル付きデータセットと、高い計算資源が要求される欠点があった。

これらの課題を解決するために提案されたのが、教師なし学習手法として知られる自己組織化マップ (Self-Organizing Map: SOM)[2]の活用である。Salehらは、SOMが非線形な主成分分析が可能なアルゴリズムである点に着目し、これを活用して特徴抽出を実現する深層畳み込み自己組織化マップ (Deep Convolutional Self-Organizing Map: DCSOM)[3]を提案した。

DCSOMはノイズの多い数字データを用いた実験で最先端の手法を凌駕し、優れた性能が示されている。しかし、ハイパーパラメータ数が非常に多く、運用上における障害となっている。そこで本研究では、パラメータの数が少なく、スパース表現によって計算量も少ない自己増殖型ニューラルネットワーク (Self-Organizing Incremental Neural Network: SOINN)[4]を導入する事で、運用性の改善を試みる。

2. 従来手法

2-1. BatchSOM

SOMには入力を与える毎に逐次ニューロンの更新を行うOnlineSOMと、一括で更新を行うBatchSOMがある。DCSOMにおいては、処理が高速で再現性の高いBatchSOMを用いて学習を行う。

2-1-1. 競合学習

SOMにおける各ニューロンは出力マップ上に配置されており、入力データベクトル x と等しい次元数を持つ参照ベクトル m_i に紐づけられている。入力を与えると、各参照ベクトルとのユークリッド距離を計算し、距離を最小化するニューロン c を勝者と定める。

$$c = \operatorname{argmin}_i \|x - m_i\| \quad (1)$$

このように、ニューロン同士を競合させて行う学習を競合学習という。

2-1-2. ニューロンの更新

データセットの全てのサンプルを入力として与えた後、参照ベクトルを以下の式によって更新する。

$$m_i = \frac{\sum_t h_{ci}(t) \cdot x(t)}{\sum_t h_{ci}(t)} \quad (2)$$

h_{ci} は勝者 c との距離を表現する近傍関数であり、 $i = c$ の時に最大値を取り、 i が出力マップ上で c から遠ざかるほど値が小さくなるものである。これにより、学習が進行すると各ニューロンの近傍のニューロンは似たような値を持ち、逆にマップ上で遠くに位置するニューロンとは、大きく異なる値を持つようになる。

2-2. DCSOM

第1節で述べたように、DCSOMは教師なしで特徴抽出を行う手法であり、そのネットワークは入力層、正規化層、複数の畳み込みSOM層、および出力層から構成されている。各層は前の層の出力を入力として受け取る。

2-2-1. 入力層

$n \times n$ サイズの入力画像から、 $w \times w$ サイズの局所パッチを $(n - w + 1) \times (n - w + 1)$ 枚取り出し、パッチの集合を作成する層である。このパッチを取り出す範囲は、Fig.1に示すように他のパッチと重なっており、隣り合うパッチと1ピクセルずれた範囲となっている。

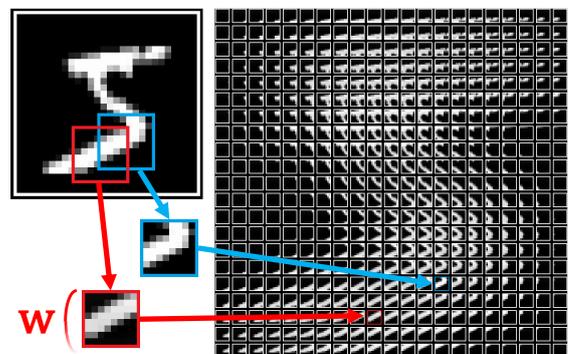


Fig.1 局所パッチ作成の例

Nonparametric Deep Convolutional Self-Organizing Map with SOINN

Akito TAKAHASHI and Yukari YAMAUCHI

2-2-2. 正規化層

作成した各局所パッチに対する前処理を行う層である。正規化の手法はいくつかあるが、DCSOMでは各パッチの平均値を減算し標準偏差で除算するZ-スコア正規化を行う。すなわち、次式の処理を行う。

$$x_i = \frac{\hat{x}_i - \text{mean}(\hat{x}_i)}{\sqrt{\text{var}(\hat{x}_i) + \epsilon}} \quad (3)$$

ここで、 x_i は正規化されたパッチ、 \hat{x}_i は正規化前のパッチを表し、また関数 mean および var は、それぞれ平均と分散を計算するものである。なお、 ϵ は0で割る事を避けるための定数であり、 10^{-8} 程度の非常に小さい値をとる。

2-2-3. 畳み込みSOM層

SOMは一般的に、2次元マップを出力とし高次元データを可視化するために用いられる。しかしDCSOMでは、視覚的特徴を効率的に抽出するために、高次元のマップを用いる。

マップ上の各ニューロンは全てのパッチの特徴を学習し、結果として各数字の局所の特徴がN次元マップ上で表現される。Fig.2に学習の例を示す。

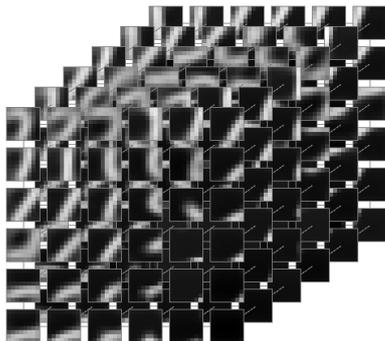


Fig.2 3次元SOMの学習例

競合学習により、類似したニューロン同士はマップ上での距離が近く、そうでないものは遠くなるように配置されるため、視覚的な特徴をマップ上での座標を用いて表現する事が可能となる。この性質を利用し、入力画像を特徴インデックス画像(Feature Index Image: FII)に変換する。FIIへの変換例をFig.3に示す。



Fig.3 N = 3とした時のFIIの例

FIIへの変換は、入力画像の各局所パッチの勝者を求めて行う。勝者のマップ上でのN次元座

標値をそのまま各次元に対応するFIIの画素値とする事で、 $(n-w+1) \times (n-w+1)$ サイズの画像をN枚出力できる。この画像は次の畳み込みSOM層への入力として渡される。

2-2-4. 出力層

DCSOMの出力は、最終的な畳み込みSOM層の出力に対してブロックヒストグラム計算を適用する事で求められる。これを連結して最終的な特徴ベクトルを計算し、サポートベクターマシンを適用する事でクラス分類を行うものである。

3. 提案手法

DCSOMは特徴抽出のために、パラメータ数が多い事で知られるSOMを高次元化、および多層化して用いている。加えて、出力層においてブロックオーバーラップ率とブロックサイズを予め決めておく必要があるため、パラメータ数が非常に多い。本研究では、SOINNを用いたクラスタリングを活用する事でパラメータの削減を試み、計算量の削減と認識精度の向上も試みる。

4. 実験

MNISTデータセットに含まれる0~9までの数字画像と、数字画像にノイズ、あるいは回転処理を加えた画像データセットを用いて、従来のDCSOMとの計算速度および認識精度を比較する。

5. まとめ

DCSOMの課題であったパラメータ量の増大をある程度解決し、かつ学習をより効率化するためにSOINNの導入を提案し、運用性の向上と計算速度、および認識精度の改善を試みた。

参考文献

- [1] A. Krizhevsky, I. Sutskever, and G. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks" In NIPS, 2012.
- [2] Kohonen, T. (1990). The self-organizing map, Proceedings of the IEEE 78: 1464-1479.
- [3] Saleh Aly and Sultan Almotairi "Deep Convolutional Self-organizing Map Network for Robust Handwritten Digit Recognition" IEEE Access, vol. 8, pp.107035-107045, 2020.
- [4] 山崎和博, 巻渕有哉, 申富饒, 長谷川修, "自己増殖型ニューラルネットワークSOINNとその実践" 日本神経回路学会誌, Vol.17, No.4, pp.187-196