

ディープラーニングを用いた 日本語のオフライン手書き文字認識

1 はじめに

新しく提案された機械学習の手法であるディープラーニングは、現在、音声認識や画像認識をはじめとした様々な分野において注目されている。その特徴としてあげられるのが、学習による複雑な内部表現の獲得である。

本論文ではディープラーニングを用いたモデルの1つである Deep Belief Network を実装し、複雑な識別問題である日本語のオフライン手書き文字認識に関して従来手法との比較を行った。

2 オフライン手書き文字認識

手書き文字認識は使用するデータによって二種類に分類される。オンライン手書き文字認識と呼ばれるものは、文字の形の他に、筆順・筆の方向など文字が書かれる過程もデータとして用いる。一方、オフライン手書き文字認識は文字の光学データなど文字の形のみを入力として識別を行う。上記のように特徴抽出が困難で、こと日本語に関しては常用漢字だけでも 2136 字と文字数が膨大であるため、現在でも特徴量の研究が行われている。

3 制限付きボルツマンマシン

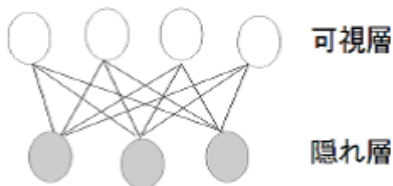


図 1: 制限付きボルツマンマシン

制限付きボルツマンマシン (RBM) は、ディープラーニングを実装する際のレイヤーとしてよく用いられる学習器である。その構造は可視層と隠れ層の2つの層を持ったボルツマンマシンとして見ることができる。ただし RBM は接続が制限されており、同層内に結合はなく、可視層と隠れ層の間のみ結合が存在している。

4 Deep Belief Network

Deep Belief Network (DBN) は多層構造のニューラルネットワークであり、ディープラーニングを用いたモデルの1つである。各層は RBM による学習を積み重ねて作られる。この学習を事前学習と呼ぶ。事前学習は、通常の学習に入る前に行われる。ここで特徴抽出を行うことにより、効率のよい学習が可能となる。

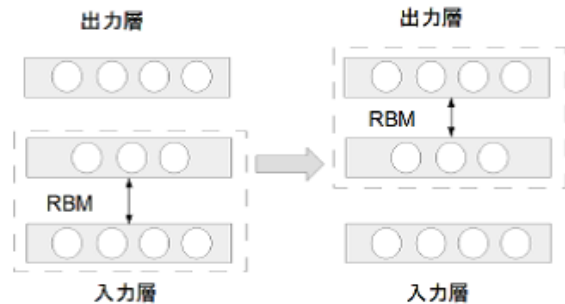


図 2: Deep Belief Network

5 実験

DBN を実際に作成し、オフライン手書き文字認識を行う。本実験では比較する従来手法として、ディープラーニングを用いていない多層パーセプトロン (MLP) を用意した。実験ではこの2つのネットワークにそれぞれ同じ入力データを与えて 150 回学習させ、文字の識別を行う。

また使用するデータとして、ETL 文字データベースの ETL9B より、任意の 2 値手書き文字データを 100 字種 (200 パターン) 書き出し、本実験用の 3 つのデータセットを作成した。

表 1: 手書き文字の識字率

	データセット A	データセット B	データセット C
DBN	83.01	84.67	82.10
MLP	67.07	69.57	65.23

6 おわりに

本論文では、新たな機械学習の手法であるディープラーニングを用いて、日本語のオフライン手書き文字認識を行い、その有効性について調査を行った。

残念ながら、実用できるような高い識字率には達しなかったが、ディープラーニングを用いていないニューラルネットワークとの比較を行うことで、その性能を示すことができた。

今後の課題として、その問題点を克服すると共に、より精度が高く、より多くの文字に対応したネットワークを検討していきたい。

参考文献

- [1] Geoffrey E. Hinton, Simon Osindero, Yee-Whye Teh : "A fast learning algorithm for deep belief nets", Neural Computation V18, pp 1527 - 1554 2006
- [2] Ruslan Salakhutdinov, Geoffrey Hinton : "Deep Boltzmann Machines" 12th International Conference on Artificial Intelligence and Statistics 2009.