

# Gradient-Based Learning Applied to Document Recognition の和訳

## 著者

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner

## 概説

逆誤差伝搬法で学習した多層ニューラルネットワークは、勾配に基づく学習法の最たる成功例である。適切なネットワークアーキテクチャがあれば、勾配に基づく学習アルゴリズムは手書き文字などの高次元パターンを分類するような複雑な決定超平面を、最小限の前処理で作成することができる。本論文では、手書き文字の認識に適用される様々な方法を振り返り、標準的な手書き整数認識タスクを用いて比較する。その結果、二次元構造の多様性を取り扱えるように設計された CNN が他の手法に比べて良い性能を示した。

実際の文章認識システムは、フィールド抽出、分割、認識、言語モデリングなど複数のモジュールにより構成されている。Graph Transformer Networks (GTN) と呼ばれる新しい学習の枠組みにより、このような多くのモジュールからなるシステムを勾配に基づく学習アルゴリズムを用いてパフォーマンス指標を最小化するように大域的に学習することが可能となった。

本論文では、オンライン手書き文字認識の 2 つのシステムを紹介する。実験により、大域的な学習の利点と GTN の柔軟性が示された。

また、本論文では銀行小切手を読み取る GTN も紹介する。この手法では、CNN の文字認識と、商業や個人利用の小切手の正確さを記録するための大域的な学習方法を組み合わせている。商業的に利用されており、1 日あたりに数百万の小切手の認識をしている。

キーワード – ニューラルネットワーク, OCR, 文章認識, 機械学習, 勾配に基づく学習, CNN, Graph Transformer Networks, Finite State Transducers.

## 略称

- GT Graph transformer

- GTN Graph transformer network.
- HMM Hidden Markov model.
- HOS Heuristic oversegmentation
- K-NN K-nearest neighbor
- NN Neural network.
- OCR Optical character recognition.
- PCA Principal component analysis.
- RBF Radial basis function.
- RS-SVM Reduced-set support vector method.
- SDNN Space displacement neural network.
- SVM Support vector method.
- TDNN Time delay neural network.
- V-SVM Virtual support vector method.

## 1 序論

ここ数年、機械学習の技術、それらを適用したニューラルネットワークがパターン認識システムの設計において、非常に重要な役割を果たしている。実際に、近年の連続的な音声の認識、手書き文字の認識といったパターン認識アプリケーションの成功にはこのような学習法の利用が重要な要因とも言える。

本論文の主な主張は、人間の経験的な知識に頼らず、自動的な学習に依存することで、より良いパターン認識システムを作成することができるということだ。これは、近年の機械学習や情報工学の進歩により可能となった。文字認識を例として、本論文ではこれまでの手作業の特徴抽出が、ピクセル画像を直接操作する注意深い学習機械に置き換えることができることを示している。

## 参考文献