

PCANet のアンサンブル学習への適用

An application of PCANet to Ensemble Learning

石田 岳志
Takeshi Ishita

山下 晃弘
Akihiro Yamashita

松林 勝志
Katsushi Matsubayashi

東京工業高等専門学校 情報工学科
National Institute of Technology, Tokyo College

PCANet[Chan et al., 2015]は、カーネルの重みを PCA で計算することで学習における計算量を抑えた CNN(Convolutional Neural Network)の一種である。PCANet は学習が非常に高速である一方、モデルが単純な演算のみで構成されるため、表現力が限られる。我々はこの弱点を克服するため、PCANet を弱学習器とみなしてアンサンブル学習に適用し、手書き文字認識データセットである MNIST¹を用いて識別性能を比較した。

1. はじめに

画像認識においては CNN がその性能から最も一般的なモデルになりつつある。多くの CNN はバックプロパゲーションによる勾配降下法を用いて学習を行うが、学習に非常に時間がかかってしまうという欠点がある。

PCANet は PCA(Principal Component Analysis)によって重みを計算する CNN の一種である。PCANet は学習が高速であり、ハイパーパラメータの調整も容易である。しかし一方で、モデル全体が単純な演算のみで構成されるため、表現力が低いという問題がある。

我々は PCANet の弱点を克服するため、PCANet を弱学習器とみなしてアンサンブル学習を行い、手書き文字認識のデータセットである MNIST¹を用いて識別性能を比較した。

2. 関連技術

2.1 PCANet

(1) 概要

PCANet は CNN の一種であり、重みを PCA で計算するという特徴がある。PCANet は学習が非常に高速であり、調整の難しいハイパーパラメータも無い。また、入力データの主成分をネットワークの重みとするため、学習は教師ラベルを必要としない。なお、学習された PCANet は入力画像を特徴ベクトルに変換するための特徴抽出器として動作する。

(2) 構造

PCANet は 2 つの Convolution 層と、1 つの Binary-hashing 層から成り立つ。第一層の重みを求める手順を次に示す。

1. 全ての入力画像から CNN のフィルタ(カーネル)領域に相当する部分画像を取得する。
2. 各部分画像をベクトル化し、正規化する。
3. 2で得られたベクトルの集合の主成分を CNN の重みとする。

同様の手順を第一層の出力に適用すれば、第二層の重みが求まる。PCANet は部分画像の主成分を CNN の重みとするため、主成分の数が各層の出力マップ数となる。Chan らの論文[Chan et al., 2014]では、第二層の出力に対し binary hashing という量子化処理を行い、量子化された出力を重み付け和したもののヒストグラムを特徴ベクトルとしている。なお、PCANet は単体では特徴抽出器として振る舞うため、識別を行う際には PCANet の出力を識別器にかけてデータのラベルを求める。

(3) 欠点

PCANet は Convolution 層に非線形な活性化関数を備えておらず、モデル全体が単純な演算だけで構成されている。このため、PCANet は他の非線形な CNN と比較すると、表現力が劣る。例えば一般物体認識のベンチマークである CIFAR10 データセットに対して、PCANet の識別精度は 78.67%である。[He, 2016]などの非線形な CNN が 90%以上の精度を達成していることを考えると、PCANet はやや表現力が劣っていると考えられる。

我々は PCANet の欠点を克服するため、PCANet を弱学習器とみなしてアンサンブル学習に適用した。また、MNIST データセットを用いた評価では、オリジナルの PCANet に比べてモデルの表現力が向上することを確認した。

2.2 アンサンブル学習

アンサンブル学習とは、多数の弱学習器を組み合わせることによって 1 つの強い学習器を構築する手法である。例えば、決定木は単体で識別問題の学習を行うと簡単に過学習してしまい、また分離平面も単純なものしか得ることができない。しかし決定木を弱学習器とみなしてアンサンブル学習を行うと、過学習せず、表現力も豊かな識別器を得ることができる。

アンサンブル学習には主に Bagging と Boosting という 2 つの種類がある。

(1) Bagging

Bagging は入力データをランダムにサンプリングし、データのそれぞれの部分集合を別々の弱学習器に入力する[Breiman, 1996]。弱学習器は完全に独立に学習を行う。推論の際はそれぞれの弱学習器の出力の多数決を取ることで、多数の弱学習器を 1 つの大きな強い学習器に見せかけている。Bagging は実装が容易であり、完全に並列に学習することができる。また、ハイパーパラメータは生成する弱学習器の数とデータのサンプリング割合のみであり、これらは精度に対して鈍感であるため調整が容易である。

(2) Boosting

Boosting は、ある弱学習器の学習結果を用いて次に学習するデータサンプルの出現確率をコントロールし、効率的な学習を図る。Boosting は弱学習器を完全に並列に学習させることはできない。一方で、ある弱学習器がうまく分類できなかったデータサンプルを重点的に別の学習器に与えるため、効率よく学習を行えるという利点がある。

連絡先: 石田岳志, 東京工業高等専門学校, ishitah.takeshi@gmail.com

¹ <http://yann.lecun.com/exdb/mnist/>

3. Ensemble PCANet

非線形な活性化関数を備えた CNN は一般的に学習に時間がかかり、また学習データの数が十分に無いとモデルが過学習してしまうため、アンサンブル学習には適していない。一方で、PCANet は表現力が強すぎず学習も高速であるため、アンサンブル学習の弱学習器として用いるのに適している。

本論文ではアンサンブル学習のモデルとして、実装がしやすくハイパーパラメータの調整も容易である Bagging を選択した。PCANet を Bagging の弱学習器とすることで、扱いやすく表現力も高いモデルを得ることができる。また、PCANet で抽出した特徴量を識別するためのモデルとして線形ソフトマージン SVM (Linear SVC)[Cortes et al., 1995]を用いた。Linear SVC は非線形 SVM や Random Forest と比べると表現力が弱いモデルであるため、特徴量の質を評価するのに向いている。我々の提案する Ensemble PCANet を図 1 に示す。

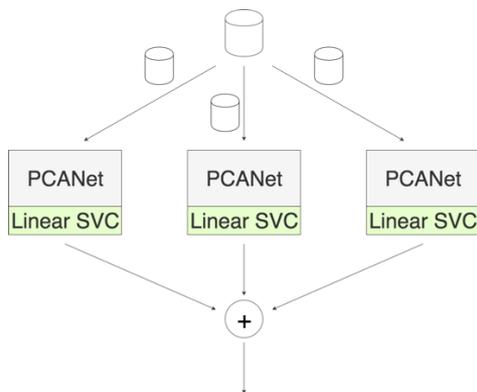


図 1 Ensemble PCANet

4. 実験

4.1 識別器の設定

我々は Ensemble PCANet とオリジナルの PCANet の精度と学習時間の比較実験を行った。いずれのモデルについても識別器として Linear SVC を用いた。また、Ensemble PCANet における弱学習器としての PCANet と、オリジナルの PCANet は全く同じハイパーパラメータを与えた。実験に用いたハイパーパラメータの値を表 1 に示す。

表 1 実験に用いた PCANet のパラメータ

入力画像サイズ	28×28
Convolution 第 1 層のフィルタサイズ	4×4
Convolution 第 1 層のフィルタステップ幅	2
Convolution 第 1 層の出力マップ数	3
Convolution 第 2 層のフィルタサイズ	4×4
Convolution 第 2 層のフィルタステップ幅	1
Convolution 第 2 層の出力マップ数	3
プーリング層のフィルタサイズ	5×5

4.2 データセット

実験では MNIST データセットを用いた。MNIST は 10 クラスの手書き文字認識データセットであり、サイズ 28×28 の手書き文字画像を学習用に 60000 枚、テスト用に 10000 枚含んでいる。なお、実験では交差検定は行わず、60000 枚の学習画像を全て学習のみに用いた。

4.3 実験結果

まず精度評価の基準として、オリジナルの PCANet の精度評価を行った。識別精度は 92.02%であった。

次に、Ensemble PCANet の評価実験として、弱学習器の数を変えながら識別精度を計測した。図 2 に計測結果を示す。図に示されているとおり、サンプリング割合が十分に大きいとき、弱識別器数が 100 程度で精度が飽和する。この実験では、サンプリング割合 0.07、弱識別器数 120 のときに最も高い精度を記録し、その値は 94.47%であった。

精度比較の結果を表 2 に示す。Ensemble PCANet はオリジナルの PCANet よりも高い識別を達成していることがわかる。

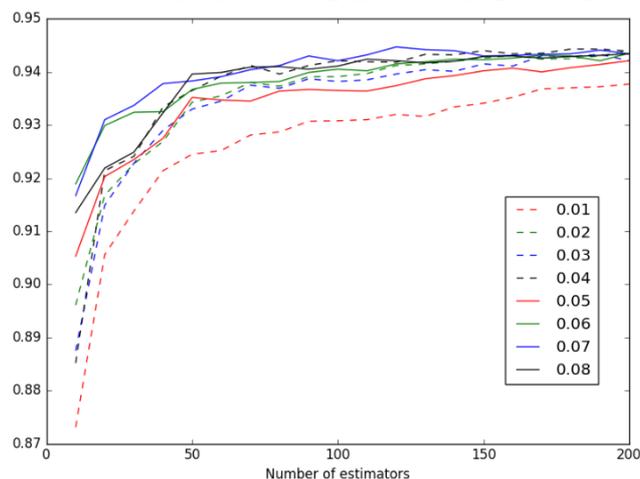


図 2 弱学習器の数と識別精度. それぞれの線は別々のサンプリング割合に対する結果を示している。本実験ではデータのサンプリング比を 0.01 から 0.08 まで変化させ、それぞれについて精度を計測した。

表 2 オリジナルの PCANet と Ensemble PCANet の精度の比較

	精度 [%]
Normal PCANet	92.02
Ensemble PCANet	94.47

5. 結論

PCANet は高速に学習可能であり、かつハイパーパラメータの調整が容易な CNN である。一方で、非線形な CNN と比較するとモデルの表現力が弱い。我々は PCANet を Bagging における弱学習器とみなして学習を行うことで、PCANet の利点を活かしつつより表現力の高いモデルを得ることができた。

参考文献

- [Cortes, 1995] Cortes, Corinna, and Vladimir Vapnik: "Support-vector networks.", Machine learning, 1995.
- [Chan, 2015] Chan, Tsung-Han, Kui Jia, Shenghua Gao, Jiwen Lu, Zinan Zeng, and Yi Ma: "Pcanet: A simple deep learning baseline for image classification?.", IEEE Transactions on Image Processing, 2015.
- [He, 2016] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun: "Deep residual learning for image recognition.", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [Breiman, 1996] Breiman, Leo: "Bagging predictors.", Machine learning, 1996.