

教師なし学習を導入した深層距離学習法

松岡 佑磨[†] 島田 大樹[‡] 彌富 仁[†][†]法政大学 [‡]法政大学大学院

1 背景

画像認識分野において, convolutional neural network(CNN)をはじめとした深層学習手法が一般物体認識などの様々な画像認識課題で高い性能を示している [1]. CNN は大量の教師ラベル付きデータ (教師データ) から認識に必要な画像特徴量を獲得することで, 高い認識性能を達成している. しかし, 学習に用いる大量の教師ラベル付きデータは多くの場合人手で用意する必要があり, 特に大規模データでは高コストとなるだけでなく, ラベルの一貫性が保たれなくなる可能性が高い.

教師データ生成の最も一般的な方法は, 特徴空間上でのデータ間の距離をもとに教師ラベルを伝播させていく label-spreading (LS) [2] やその改良された手法を適用することである. そのため, データの性質をよく表現する特徴を事前に抽出する必要がある.

この問題に対して我々はパラメータを共有する2つのCNNから構成される siamese network と LS を組み合わせた深層距離学習という手法を提案してきた [3]. siamese network は同じラベルに属する任意の2つのデータに対する出力の差が小さくなるようにパラメータを学習するネットワークであり, この研究において効果的な特徴表現が得られることが示唆された.

しかしながら, 従来の深層距離学習は学習に教師ラベル付きデータのみを利用するため, 本研究で扱う問題設定において識別したいデータすべてを事前に観測可能であるというメリットを活用していない. また siamese network は一部のデータの組み合わせにおいて非効率な学習を行うため, 深層距離学習に改善の余地があると考えられる.

本研究では, より効率的な学習器構造を採用したうえで, 距離学習の前処理としてラベル未付与データから教師なし事前学習を用いた特徴抽出を行う深層距離学習

法を提案する. 提案手法を教師ラベル生成問題に適用して, その有効性を確認する.

2 手法

本研究では, 画像データを対象にした教師データ生成問題を扱う. 提案する手法は, 「特徴抽出部」と, 「教師ラベル生成部」の2つから成り立つ. 「特徴抽出部」では後述する距離学習によって画像から特徴を抽出し, この特徴をもとにラベルが付与されていない画像を含む全ての画像から特徴ベクトルを作成する. 「教師ラベル生成部」では, 特徴抽出部より抽出された特徴ベクトル群に半教師あり学習によって教師ラベルを推定し, 教師データを生成する.

2.1 特徴抽出部

特徴抽出部は, 図1で示すような triplet network [4] を使用する. この特徴抽出部の学習時の入力 x_a, x_p, x_n の3つの画像データの組み合わせである. このとき x_p は x_a と同じラベルのもの, x_n は異なるラベルのものとする. これらの画像の組み合わせを (パラメータを共有する)CNN($f(\cdot)$) に入力し, 特徴ベクトル $f(x_a), f(x_p), f(x_n)$ を得る. このCNNの学習はこれらの出力を用いて, 式1で示す triplet loss 関数を最小化するように誤差逆伝播法で行われる. 式1において, m は特徴空間におけるデータ間距離を決めるハイパーパラメータであり, $d(\cdot)$ は2つの画像データに対する特徴空間上のユークリッド距離を出力する関数である.

$$\mathcal{L}_{triplet}(x_a, x_p, x_n) = \max\{0, d(x_a, x_p) - d(x_a, x_n) + m\} \quad (1)$$

2.2 教師ラベル生成部

特徴抽出部で得られた各データの特徴ベクトルと一部のデータに付与されている教師ラベルを用いて, グラフベースの半教師あり学習手法の1つである LS [2] を適用し, ラベル未付与データに教師ラベルを付与する.

Deep similarity metric learning with unsupervised learning

[†] Yuma Matsuoka, Hitoshi Iyatomi (Hosei University)

[‡] Daiiki Shimada (Graduate School of Science and Engineering, Hosei University)

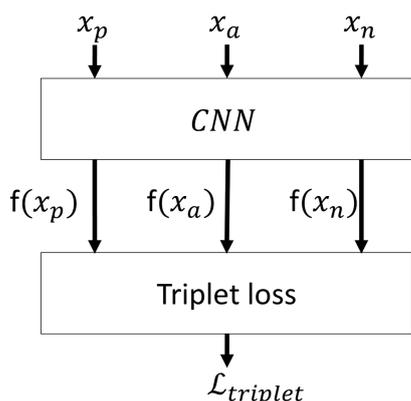


図1 本研究で使用する triplet network のモデル

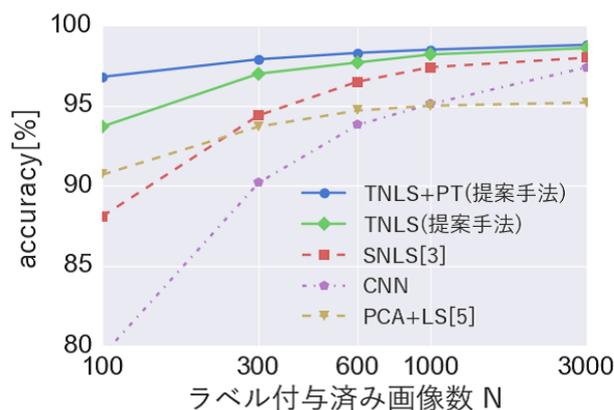


図2 MNIST データセットにおける実験結果

3 実験

提案手法の有効性を検証するために、MNIST データセットを用いて性能評価を行った。MNIST データセットは 0 から 9 までの 10 種類の教師ラベルが付与された手書き文字画像群であり、 28×28 pixel のグレースケールの 60,000 枚の学習用画像と、10,000 枚のテスト用画像から構成されている。この 60,000 枚の学習用画像群から N 枚を取り出してラベル付与済み画像として、残りの $60,000 - N$ 枚をラベル未付与画像として用いる。

次に評価実験に用いる提案手法と比較手法について説明する。提案手法は triplet network と LS を用いて教師データ生成を行う手法 (TNLS) と、TNLS に加えて図 1 の CNN の重みの初期値を教師ラベル付与、未付与両方の観測データから auto encoder による教師なし事前学習したもの (TNLS+PT) である。triplet network における CNN は畳み込み層、正規化層、プーリング層、全結合層をそれぞれ複数用い、全 9 層で構成される。最終層の全結合層から出力される特徴ベクトルの次元数は 100 次元とする。CNN の構造は予備実験の結果によって決定した。

比較としてラベル付与済み画群のみを使用して学習した CNN および、教師データ生成の先行研究である PCA+LS [5], siamese network+LS (SNLS) [3] を用いた。MNIST データセットに対して提案手法及び比較手法を適用した結果を図 2 に示す。

4 考察とまとめ

提案手法である TNLS+PT および TNLS は、既存の深層距離学習手法よりも高い教師データ生成精度を達成することができた。TNLS+PT はラベル付与済み画像数が極端に少ない場合に顕著な精度向上を果たした。教

師なし事前学習によって、ラベル未付与データの特徴抽出に有効なパラメータを CNN が事前に獲得できたと考えられる。

ベースラインとなる TNLS も既存の深層距離学習法 SNLS より高い精度を達成し、triplet network の 3 つの出力ベクトルから求める同一ラベル間、非同ラベル間距離を使用する学習が効率的な距離学習を実現していると考えられる。

これらのことから、距離学習器を教師なし事前学習することと、LS の前処理として triplet network を使用することの有効性が明らかになった。今後は他のデータセットに対しても実験を行い、提案手法が深層距離学習の性能向上に寄与するか検討する。

参考文献

- [1] A. Krizhevsky et al., Imagenet classification with deep convolutional neural networks. NIPS, pp. 1097-1105, 2012.
- [2] D. Zhou et al., Learning with Local and Global Consistency. NIPS, pp. 321-328, 2004.
- [3] 松岡ら, 深層距離学習による教師データ生成. ファジィシンポジウム講演論文集, Vol. 32, pp. 91-96, 2016.
- [4] J. Wang et al., Learning fine-grained image similarity with deep ranking. CVPR, pp. 1386-1393, 2014.
- [5] 島田ら, 半教師あり学習による聴講者の集中度推定法の可能性. ファジィシステムシンポジウム講演論文集, Vol. 31, pp. 395-398, 2015.