

訓練データ拡張法に基づく Deep Learning の物体認識の精度向上

嶺友太 浅海賢一 小森望充
九州工業大学大学院 工学府 先端機能システム工学専攻

1 はじめに

Deep Learning による物体認識をうまく行うためには、認識クラス毎に分類・整理された大量の訓練データが必要となる。しかし、認識クラスに対応した様々な物体の画像を大量に撮影し、その一つ一つにラベル付けを行うには人手による膨大な労力がかかる。そこで、本研究では、Deep Learning による物体認識の精度を向上させると同時に、訓練用画像の収集に要する労力を削減することを目的としている。

実験では、訓練用画像に数パターンの処理を適用してデータを拡張し、畳み込みニューラルネットワークによる学習を行い、データの拡張を行わない場合と物体認識精度の比較を行った。

2 原理

2.1 畳み込みニューラルネットワーク

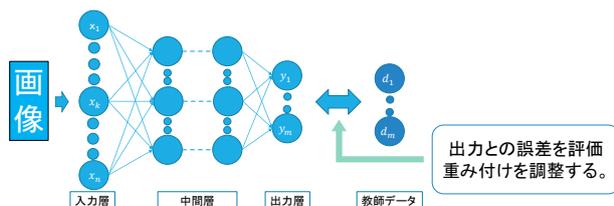


図 1. 畳み込みニューラルネットワークの基本構造

畳み込みニューラルネットワークの基本構造を図 1 に示す。学習には、正解ラベル付きの訓練データを用いて、教師あり学習を行う。Deep Learning での学習とは、訓練データをネットワークに入力したときの出力と教師信号の誤差が 0 に近づくようにユニット間の重みを調整することである。

2.2 訓練データ拡張

Deep Learning による機械学習を行うときに実行することで、汎化性能を向上する方法がいくつか知られている。機械学習では、訓練データの不足や偏りが学習の妨げになる場合がある。そのような場合、訓練データにならぬ前処理を行い、訓練データを拡張する。画像の前処理の例として、画像の正規化、拡大縮小、回転、平行移動などの幾何学的変形、ノイズの付与などがある。

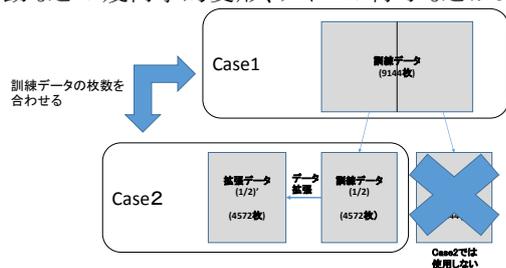


図 2. 訓練データ数の調整

一般的に、訓練データ数は多ければ多いほど精度が上がると言われているので、データ拡張を行う場合、元の訓練データの一部を用いてデータ拡張を行い、データ拡張を行った場合の訓練データ数がデータ拡張を行わない場合の訓練データ数と一致するようにしている。

3 実験結果

3.1 認識性能

本研究では、101 クラス、9,144 枚の画像から構成された Caltech101-database[1] データセットを訓練データに用いている。

評価には、訓練用画像とは異なる評価用の画像を用いた。評価のスコアが 5 番目以内のものを正解とし、それ以外を不正解としている。図 3 には、結果の一部として、蝶の評価用画像を 0~360° の間で回転させたときそれぞれの不正解率を表している。

case1 がデータ拡張無しで case2~case5 は、それぞれ異なる前処理でデータ拡張を行っている。

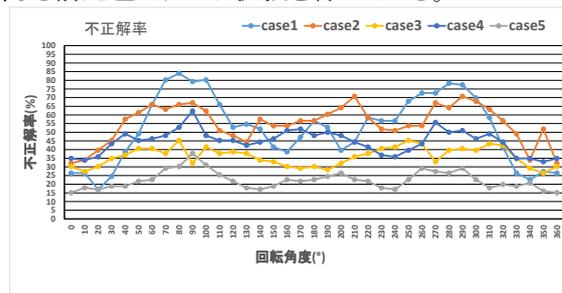


図 3. 蝶の評価用画像の不正解率

3.2 学習時間、認識時間

表 1 に学習に要した時間と、画像 1 枚の分類時間を示す。画像の認識時間は 10 回の平均をとっている。

表 1. 学習時間と認識時間

学習時間	4 時間 ± 10 分
認識時間 / 画像 1 枚	60 [ms]

4 まとめ

データ拡張を行うことで、物体認識の精度が向上できた。また、前処理の仕方によっては精度が下がってしまうこともあり、ただデータを増やせばいいということではないことが分かった。

参考文献

- [1] http://www.vision.caltech.edu/Image_Datasets/Caltech101/
- [2] 嶺ユウタ: ああああああいいいいいいいい、すごい出版社, 2011.
- [3] 山田太郎他: 深層学習, 人工知能学会, 2015.