ISSN 2186-5647

6-1

正常数字を学習したネットワークによる欠損数字の認識

日大生産工（院）　○三好　亮輔
日大生産工　原　一之

1. はじめに

コンピュータを用いた画像認識において、画像に局所的な欠損が生じたとき、その判別能力はどれほど悪化してしまうのか検証する。そこで今回は、手書き数字画像集（MNIST）の中から任意に抽出した画像にマスキングを施し、意図的に数字領域の一部を欠損させた画像を 100 枚程度作成したのち、これらを無欠損の正常な数字画像で学習させたネットワークに認識させて、判別精度を確認する実験を行った。

2. 研究方法

まず、Mac OS を用いてこの中に仮想環境を展開し、ゲスト OS「Ubuntu」をインストールした。さらにこの上にフレームワーク「Caffe」を構築し、環境とした。

今回、判別に使用した画像は 1 枚 28×28 ピクセルの 0～9 の手書き数字の画像を集めた「MNIST」と呼ばれる画像集である。MNIST には、あらかじめ正解ラベルが付与された学習用画像 60,000 枚と、これを検証するためのテスト用画像 10,000 枚が用意されている。

テスト用画像の中から 100 枚（0～9 各々20 枚ずつ）を任意に選び、全画像に対して Fig.1 のようにマスキング加工を行ったものを用意した（以下、実験用画像）。マスキング加工を施すにあたっては、画像のある一定の座標エリアに対してのみ黒画素を配置するプログラムを作成した。なお、マスク部分の幅については変更できるよう、プログラム中に可変箇所を設け、100 枚の画像に対し、一括で同じ処理を施せるようにした。また、マスクしない内部の正方形部分（以下、窓）の一辺は 14×14 ピクセルとした。

Fig.1 The process of masking the image

判別に用いたネットワークには、予め Caffe に学習用画像を学習させ、出来上がった学習済モデルを使用した。判別の際は、これをインポートして実験用画像を判別させた。

画像を CNN に読ませて判別を行う際、局所領域からフィルタを通して検出していき、抽出された特徴を「特徴マップ」と呼ばれる情報に変換するが、マスキングによって数字の特徴が失われた場合、本来の画像の特徴マップと異なってくる。マスクなしの画像と比較して取得可能な特徴量は限られるため、判別の精度は下がることは予想できる。

今回の判別実験で実際にどの程度分類能力が落ちるのか、どの数字により顕著に表現するのかを確認した。Caffe の分類機能には第 3 候補まで可能性を表示するようになっている。例えば、Fig.2 のように、「0」を分類したとき、0 という可能性の他に、5 や 7 である可能性も表示されている。

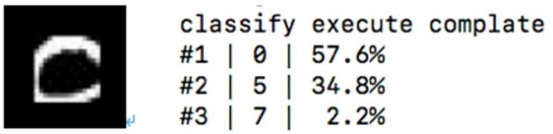

Fig.2 [An example] When classifying the number “0”

3. 結果

それぞれの数字の判別結果の平均値を 10 個の数字ごとにとり、まとめたものが Fig.3 のグラフである。

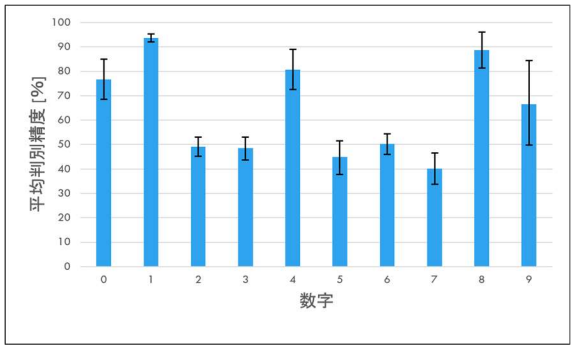

Fig.3 Accuracy of classifying numbers

棒グラフが高いほど判別精度が高く、頂点の黒線が長いほど判別結果にバラつきがあることを示す。

このグラフを見ると、「0」、「1」、「4」、「8」、「9」については誤りが少ないと言える。しかし、精度が高く

Study on recognize masked numbers using network learned correct numbers

Ryosuke MIYOSHI

見える「0」、「4」、とくに「9」などは、10 枚の画像の判別結果のバラつきが顕著であり、これらの数字の画像の中には誤判別されたものも多く含まれることが分かった。これは、元の数字の位置や形状、さらには書く人のクセに左右されるからであるとも考えられる。例えば、同じ数字をマスキングしても、”窓”の部分に収まるような小さく中央寄りの画像もあれば、画面いっぱいに大きく書かれたもので、マスキングで数字の大部分が失われてしまうような画像であれば精度は下がると予測できる。また、画像の中には分類しづらいデータ（例えば「6」の丸い部分を「0」のように書かれた画像など）もあるためであると考察した。下図のように、人が見てもどの数字なのか紛らわしい画像もある。Fig.4 はどちらも「6」としてラベル付けされた画像であるが、左は「0」、右は「0」ないし「4」に見えてしまう。

Fig.4 Misleading numbers

この 10 個のマスク数字の分類実験は、MNIST の 10,000 枚に及ぶテスト用画像を、任意に抽出して加工した結果であり、すべての画像をテストしたものではないため、完全とは言えない。
また、今回の実験で得られたもので、画像の分類において、何らかの類似する特徴をもつ要素に対しては、ある一定の割合で誤って分類されてしまうという傾向があることを示した結果もある。これらの問題点は、学習用画像を増やして、様々なシチュエーションに柔軟に対応できるよう、学習精度を向上させてゆくことで、ある程度改善されるものと考えられる。

4. 参考文献

[1] UC Berkeley BVLC「Caffe Installation」
http://caffe.berkeleyvision.org/installation.html

[2] 斎藤 康毅
「ゼロから作る Deep Learning」
オライリー・ジャパン（2016 年）

[3] 内橋 堅志
「Caffe で Deep Learning
つまずきやすいところを中心に
https://qiita.com/uchihashi_k/items/8333f80529bb3498e32f」
（2014 年）

[4] Ry0_Ka
「Caffe を使って Deep Learning するためのデータセット作り
https://ry0.github.io/blog/2015/08/22/create-dataset/」
（2015 年）