

画像認識のための畳み込みニューラルネットワーク (CNN) の設計, 学習および評価

永田研究室 F115107 山本 さと子

1. 目的

近年, 工学上の応用技術として「ニューラルネットワーク」が広く使われるようになってきた. ニューラルネットワークとは人間の脳の仕組みをコンピュータで模倣しようとしたものであり, 例えば, 人の大脳皮質には 140 億を超える膨大な数のニューロン (神経細胞) が様々な結合し, 情報を伝えたり処理したりすることで, 記憶したり計算したりしていると考えられている. 特に, 画像認識の分野に特化した畳み込みニューラルネットワーク (CNN) は機械学習の中で最も注目を集めている技術の一つである. 研究室では, C++や Python といったプログラム言語を使用することなく CNN を設計できるツールを MATLAB 上で開発している. 15 層からなる CNN を設計し, 公開された画像データセットの訓練データを用いて学習させた後にテストデータを用いて認識率の評価を行い, 開発したツールの有用性を確認する.

2. 研究内容

設計した 15 層の CNN 用いて図 1 のような 3 種類の画像データセットを使用し, 訓練とテストを通じて認識率の検証を行った. 用いた画像データセットは MNIST, Fashion MNIST, CIFAR10 であり多くの CNN の研究で利用されている公開データを使用した. 実験では訓練データで CNN を学習させたあとテストデータを用いて画像分類精度を検証し CNN の汎化性を評価した. CNN の代表的なパラメータのうち, Max Epochs, Mini Batch Size, Initial Learn Rate, Accuracy を試行錯誤的に変化させて比較評価した. このとき, 学習の停滞が観察された場合は過学習とならないように学習のプロセスを終了させて対処した. なお, Fashion MNIST にはグレースケール 28×28 ピクセルの訓練データ 5851 枚とテストデータ 892 枚の画像が, MNIST にはグレースケール 28×28 ピクセルの訓練データ 6000 枚とテストデータ 1000 枚の画像が, さらに CIFAR10 にはカラー画像 32×32 ピクセルの訓練データ 5000 枚とテストデータ 1000 枚が含まれている.

3. 結果

各画像データセットを用いて複数回の訓練と検証を行い, 最も高い認識率で比較した. MNIST, Fashion MNIST の評価の際に設定した Max Epochs, Mini Batch Size, Initial Learn Rate, Accuracy のパラメータはそれぞれ 30, 100, 0.001, 0.99, CIFAR10 では同様のパラメータをそれぞれ 5, 100, 0.001, 0.999 に設定し, 検証を行った. その結果, MNIST, Fashion MNIST, CIFAR10 では 99.7 %, 87.5 %, 70.4 % の認識率が得られた. なお, 先行の研究論文では MNIST, Fashion MNIST それぞれの数値が 93.3 %, 86.6 % であったため, それらと比較しても同等の性能が発揮できたことが確認できた. しかし, CIFAR10 を使った最近の研究論文では 80~90 % の結果が出ているのに対し, 実験結果では低い認識率となった. 原因としては, 先行論文では新しいデータセットを作成し, 公開データに加え, 分類精度を上げる学習方法を使用していたことが考えられる.

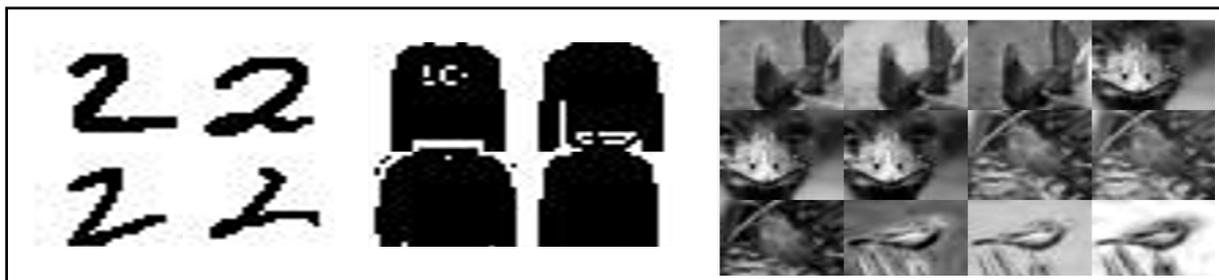


図 1 実験で使用した画像データセット (左から MNIST, Fashion MNIST, CIFAR10)