

ディープラーニングによる医療画像の分類の基礎研究

永田研究室 F115087 本田 修平

1. 目的

近年、ディープラーニングをはじめとする AI の研究・活用が盛んに行われている。その中でも自動運転や ILSVRC(ImageNet Large Scale Visual Recognition Challenge)の結果に代表されるように特に視覚の分野で大きな成果を挙げている。そこで、これらのディープラーニングの視覚に関する技術 (CNN) を医療分野に活用することを考えた。本研究では CNN を用いてレントゲン写真の中から健康体と肺炎患者を見分ける研究に取り込み、分類精度の向上を目的とした。

2. 研究内容

カリフォルニア大学の医師である D. Kermay, K. Zhang, M. Goldbaum の 3 氏により公開されている胸部レントゲン写真のデータセット (Fig. 1) を用いた。データセット内にある 5,856 枚の健康体と肺炎患者の胸部レントゲン写真を学習データ、検証データ、テストデータに分割し、学習データと検証データを用いて構築した CNN をトレーニング、ハイパラメータチューニングを十分な結果が得られるまで繰り返した。その後、モデルの汎化能力を試すためにテストデータを用いて評価し、ILSVRC で優勝経験のモデルを同じデータで学習させて分類精度を比較した。さらに高い精度での分類を目指して層の追加や新たな CNN の構築、データ拡張などの試行錯誤と評価、比較を繰り返し行った。最終的に、Inception モジュール、Shortcut Connection, Depthwise Separable Convolution, Batch Normalization, Global Average Pooling などの技術を用いて全結合層を用いない 575 層から成るネットワークを構築した。

3. 結果

ILSVRC で優勝経験のある VGG16 をベースにした CNN では正解率が 79.4%, Inception ResNet v2 では正解率が 93.8%であるのに対して、本研究で開発したのネットワークでは正解率が 98.1%であった。また、実際の肺炎患者の内、モデルが肺炎患者と予測した割合を示す再現率は Inception ResNetv2 が 97.2%, 提案するネットワークでは 99.7%と両方の指標において後者の方が精度が高かった。本研究で用いたデータは健康体より肺炎患者の方が多く、誤認識したもののほとんどが健康体を肺炎患者と誤認識しているものであった。このことからデータの偏りが精度に影響を及ぼすことが確認できた。しかし、この影響を最小限に抑えるために施したデータ拡張と ImageNet で学習した重みの移植が大きく結果に寄与したものと考えられる。

今回は健康体と肺炎患者というバイナリクラス問題であったが、今後はマルチクラスにした研究に拡張するなど、広範囲でさらに高い精度での分類を目指して研究を続けていきたい。



Fig. 1 研究に用いた胸部レントゲン写真の一例
(出典: <https://data.mendeley.com/datasets/rsbjbr9sj/2>)