

じん肺画像判定 CNN の最適化アルゴリズムの比較と特徴獲得

1240325 椎葉 啓介 【 知能情報学研究室 】

1 はじめに

CNN を利用した画像判定において、学習アルゴリズムと学習率の選定は重要である。そこで本研究では、じん肺の画像判定 CNN において、学習アルゴリズムと学習率がどの程度正解率に影響を及ぼすのか、またどの学習アルゴリズムと学習率の組み合わせにおいて最も正解率が高いか比較を行う。さらに、モデルにより獲得した特徴についても GradCAM にて可視化し、比較する。

2 実験

2.1 データセット

National Institute of Occupational Safety and Health(NIOSH), 高知大学医学部 (KM), National Institutes of Health Clinical Center(NIHCC) の胸部 X 線画像を元に、肺領域のみを切り取った計 238 枚の画像を使用した。各画像は、それぞれ NF(No Finding) 画像あるいはじん肺画像のどちらかにラベル付けされている。

2.2 実験内容

Keras に実装されている 8 つの学習アルゴリズムに対し、学習モデル VGG16 で学習を行う。バッチサイズは Nadam では 1, RMSprop では 4 とし、それ以外においては 16 とする。学習率は 10^{-1} , 10^{-2} , 10^{-4} , 10^{-5} , 10^{-6} を使用し、epoch 数は予備実験で検討した結果、全て 100 としている。また、各学習アルゴリズムにおける最高正解率とその学習率を求め、その組み合わせの学習済みのモデルに対し GradCAM を用いることで、CNN における画像判定の根拠を確認する。

3 結果・考察

表 1 に各学習アルゴリズムにおける最高正解率とその学習率を示す。正解率が最も高いのは Root Mean Squared Prop(RMSprop) の学習率が 10^{-5} の場合であった。また、8 つの学習アルゴリズムの内、6 つは最高正解率が 90% を超えたが、Nesterov-accelerated Adaptive Moment Estimation(Nadam) と Follow The Regularized Leader(Ftrl) ではそれぞれおよそ 73% と 49% であった。

図 1 に精度が最も高いモデルであった RMSprop と、精度が最も低いモデルであった Ftrl に対し、GradCAM を用い画像の判定根拠を可視化した画像を示す。精度が最も高いモデルでは肺領域を中心に判定の根拠としているのに対し、精度が最も低いモデルでは肺領域に関係なく画像全体を根拠としていることから、精度が低いモデルでは過学習や学習不足が起きているのではなく、適切な特徴が獲得できていないと考える。

最急降下法から発展した Adam などの学習アルゴリ

表 1: 各学習アルゴリズムの最高正解率

最適化法	学習率	正解率
Adadelta	10^{-1}	90.24
Adagrad	10^{-2}	90.24
Adam	10^{-5}	90.24
Adamax	10^{-4}	92.68
Ftrl	10^{-2}	48.78
Nadam	10^{-2}	73.17
RMSprop	10^{-5}	92.68
SGD	10^{-2}	90.24

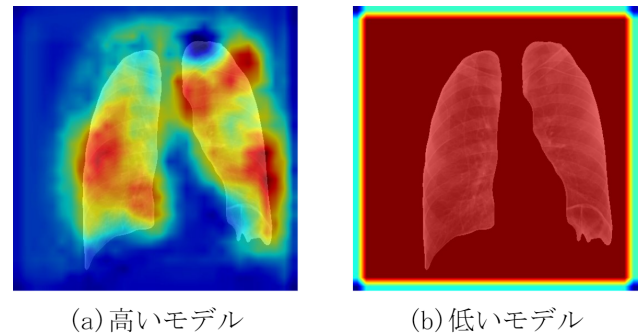


図 1: 正解率の違いによる画像判定根拠の比較

ズムにおいては高い精度を示したが、Nadam の精度が十分でなかった理由は分かっていない。一方で、Ftrl は広告のクリック率の予測のために開発されたアルゴリズムであるとされるため、一般の画像認識に対しては不適切であると考ええる。

4 まとめ

じん肺の画像判定 CNN において学習アルゴリズムと学習率を変更した結果、正解率が低くなる学習アルゴリズムが存在するだけでなく、特徴の獲得が十分でない学習アルゴリズムが存在することが判明した。そのため、じん肺の画像判定 CNN に対し、学習アルゴリズムおよび学習率の選定は重要であると考えられる。また、GradCAM を用いることでモデルの根拠が可視化されることにより、そのモデルの正解率を裏付けるものとなった。

参考文献

- [1] 東, “胸部 X 線画像からの機械学習を用いた塵肺検出に関する研究”, 高知工科大学修士論文, 2021.