

Deep learningによる眼底 画像診断に関する研究

平石研究室

S18105 島野 郁人

・はじめに

- ディープラーニングはニューラルネットワークを多層に結合することで高度な学習を可能にしたものであり，自然言語処理，音声認識，画像認識などの分野で活躍している。
- 中でも画像認識技術は産業界で実用化されるほど活躍している。
- 本研究では診断が難しい場合が存在する緑内障診断について画像認識を扱い，どれほどの精度で認識できるかを検証する。

<使用するライブラリについて>

- 画像解析をするツールとしてwekaを扱い，wekaに含まれているディープラーニングのアルゴリズムを幅広くサポートしたフレームワークであるdeep learning4jを使用する。

・ 研究内容について

- 健康な目と緑内障の眼底画像をweka内のdeep learning4jに含まれる様々なネットワークに学習させ、学習後評価し、ネットワークとの関係性を考察する.

<データの詳細について>

健康な目の眼底画像 55枚

緑内障の眼底画像 104枚

・ 実行したネットワークについて

ネットワーク名	実行が出来たかどうか
AlexNet	○
Darknet19	×
FaceNetNN4Small2	○ (同じ結果)
inceptionResNetV1	○ (同じ結果)
ResNet50	○ (同じ結果)
LeNet	○
SqueezeNet	○ (同じ結果)
VGG16	×
VGG19	○
Xception	○

- ・ 実行可能であった8種類ネットワークの中でも有名なアーキテクチャでありかつ数値として比較しやすかったAlexNet, inceptionResNetV1, LeNetの3種類のネットワークで比較する.

・ 学習方法と評価方法について

• 10-fold cross validation (10-分割交差検証)

訓練データセットを10個のサブセットに分割し、そのうち9個で学習し残り1個で検証するサイクルを10回繰り返すものである。

< 評価指標 >

• 分類精度

正しい分類が出来ているかを調べるものである。

• 二乗平均平方根誤差

0に近い（誤差が小さい）ほど精度が良いと判断できる。

• K統計量

値が高いほど信頼性が高いと判断できる。

• 混合行列

予測の結果と実際の結果の関係を示すものである。

・比較するネットワークの説明

・ AlexNet (2012年)

AlexNetは人が特徴量を設定しなくても十分なデータさえ存在すれば、機械自身が特徴量を見つけ出すことを証明した。

Weka内のAlexNetの構造は畳み込み層が5層，サブサンプリング層が3層，全結合層が2層，出力層という11層から成り立っている。

・ inceptionResNetV1 (2015年)

ResNetはshortcut connectionという機能を導入し，手前の層の入力を後ろの層に直接足し合わせることで勾配消失問題を解決した。

weka内のinceptionResNetV1の構造は畳み込み層が127層，activation層が41層，サブサンプリング層が4層，全結合層が1層，出力層という174層から成り立っている。

・ LeNet (1998年)

LeNetは畳み込みニューラルネットワークの最初の提案および実装である。

weka内のLeNetの構造は畳み込み層が2層，サブサンプリング層が2層，全結合層が1層，出力層という6層から成り立っている。

・ AlexNetの実行結果

AlexNetの混合行列表

実際の画像 \ 予測内容	緑内障	健康な目
緑内障	94(Ture Positive)	10(False Negative)
健康な目	49(False Positive)	6(Ture Negative)

分類精度：62.8931%

二乗平均平方根誤差：0.6091

K統計量：0.0155

この表から緑内障を緑内障と判断できた割合が94/104であり、健康な目を健康な目と判断できた割合が6/55であるため、後者の割合が低すぎる微妙な結果となった。また、二乗平均平方根誤差の値が大きい点からも誤差が大きいモデル(精度が悪いモデル)ということが分かった。

・ inceptionResNetV1の実行結果

inceptionResNetV1の混合行列表

実際の画像 \ 予測内容	緑内障	健康な目
緑内障	104(Ture Positive)	0(False Negative)
健康な目	55(False Positive)	0(Ture Negative)

分類精度：65.4088%

二乗平均平方根誤差：0.4806

K統計量：0

この表より眼底画像をすべて陽性と判断しているため、正しく評価ができなかったことが分かる。観察者間の結果の一致度を評価するk統計量の値が0という数値は、観察者同士の一致がない、もしくは観察者同士の一致率を偶然の一致率が超えてしまう場合しかないため、この評価は信頼性がないものだと考えられる。また、ネットワークの大きさに対して、入力したデータが少ないことから学習結果側に問題があったと考えている。

・ LeNetの実行結果

LeNetの混合行列表

実際の画像 \ 予測内容	緑内障	健康な目
緑内障	91(Ture Positive)	13(False Negative)
健康な目	13(False Positive)	42(Ture Negative)

分類精度：83.6478%

二乗平均平方根誤差：0.3682

K統計量：0.6386

この表から緑内障を緑内障と判断できた割合が91/104であり，健康な目を健康な目と判断できた割合が42/55であるため良い結果となった．また二乗平均平方根誤差の値が一番低い点やk統計量の値が6割を超えている点から精度が良く，観察者間の一致度が高い，つまり評価への信頼性が高い値にあることが分かる．

・まとめと考察

- 本研究では緑内障診断を様々なネットワークを利用してどのくらいの精度を出すことが出来るのか検証を行った。
- 結果、一番古いネットワークであり、層の数が少ないLeNetが一番良い結果となった。これは階層数とデータ数が比例関係にあることからデータ数が少なかったため、低階層のネットワークで良い結果が出たのではないかと考えられる。
- データ数に依存する結果となってしまったが、精度を上げるにはデータ数を増やし、層ごとの与える影響を理解し適切な畳み込み層数などを考える必要があるのではないかとと思われる。