

深層学習による3次元MRI脳画像分類モデルの説明可能性に関する研究

1255115 藤沢 元 【知能情報学研究室】

A Study on Explainability of Deep Neural Network-Based Classification Models for 3D Brain MRI

1255115 FUJISAWA, Moto 【Intelligent Informatics Lab.】

1 はじめに

畳み込みニューラルネットワーク (Convolutional Neural Network:CNN) による画像診断支援が活発に研究開発されている。ニューラルネットのブラックボックス性の問題のため、どのような特徴や情報に基づいて結論を得ているかが不明である。そのため、判定の根拠が解釈可能な説明可能 AI が注目されている。MRI 脳画像を用いた CNN の説明性は寄与部位の可視化や部位ごとによる精度差などがあげられ、寄与部位をヒートマップを用いた可視化や画像の前処理ごとのヒートマップの違いによって識別領域の説明性を求める。本研究では、レジストレーション・セグメンテーション等の変換処理を行った画像や変換処理の過程で生じる変位場等の変換情報に対して 3DCNN を行い、Grad-CAM[1] による男女脳識別の可視化、精度やセグメンテーションによって 3DCNN モデルの説明性とする。

2 関連研究

文献 [2] などでは 3DCNN を用いた男女識別は、頭蓋あり画像・頭蓋あり線形レジストレーション・頭蓋なし全脳画像で高精度である。しかし、脳領域抽出やレジストレーションを行わない場合、頭蓋やその他の領域などの脳以外の部位や男女差があるとされる脳の大きさによる識別の可能性が残る。

また文献 [3] では、画像の前処理方法による識別精度の違いは、頭蓋あり (レジストレーション画像を含む) \geq セグメンテーション \geq マスク \geq 変位 $>$ チャンスレベルの順であることが確認されている。非線形レジストレーション画像は、式 (1) のアフィン変換行列 A を用いて平行移動、回転、拡大縮小、角度変更の線形変換を行い、非線形変換の位置 (x, y, z) における微小変位 $d(x, y, z)$ を加算することで得られる画像である。図 2 は式 (1) の $d(x, y, z)$ である。この画像を用いて男女識別が可能であることから、非線形レジストレーションの変位には男女識別を行える情報があると考えられている。変位画像は脳の形がはっきりわかる画像ではないため、脳の局所的な形状情報があると考えられ、解析対象とする。

文献 [2][3] より、画像の前処理として脳領域抽出とレジストレーションの両方を行った画像の作成が必要であ

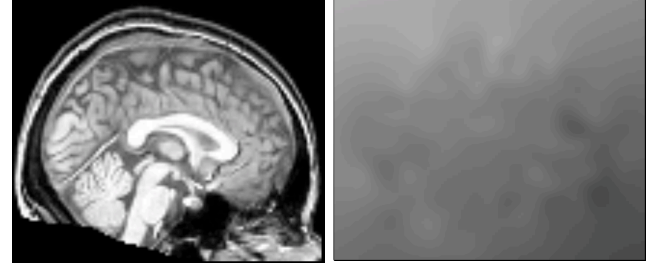


図 1 非線形標準脳

図 2 非線形変換の変位

ると考え、新たな前処理画像によって脳領域以外の影響を抑えても識別可能であるかを確認し、Grad-CAM によるヒートマップによって CNN モデルの説明性を探る。

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} + \begin{bmatrix} d_x(x, y, z) \\ d_y(x, y, z) \\ d_z(x, y, z) \\ 0 \end{bmatrix} \quad (1)$$

3 実験内容

データセットはロンドンの異なる 3 箇所の病院で収集された 581 人の T1 強調画像が含まれる IXI-Dataset を使用する。画像に対する前処理としてレジストレーションとセグメンテーションを行い、A. 線形レジストレーション、B. 線形レジストレーション&脳領域抽出、C. 全脳セグメンテーション、D. 白質セグメンテーション、E. 灰白質セグメンテーション、F. 脳髄液セグメンテーション、G. 非線形レジストレーション、H. 非線形レジストレーション変位、I. 非線形レジストレーション&脳領域抽出の 9 種類の画像を作成する。性能評価は Accuracy を用い、チャンスレベルは女性 66 人男性 53 人の 55% である。男女識別タスクを行い、識別結果に対する寄与部位の可視化を行うため、Grad-CAM を用いた寄与部位の計算を行い、脳画像の平均、Grad-CAM の計算結果の平均を重ねた画像を作成する。

4 結果・考察

表 1 より、各画像での男女識別の結果を示し、画像 I を除き各結果が 55% より高いことから構築した 3DCNN

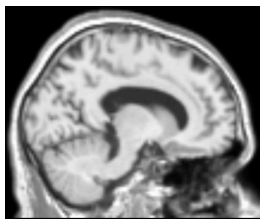


図3 画像A

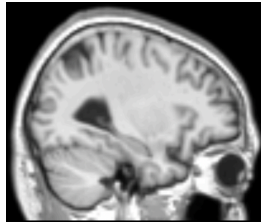


図4 画像A

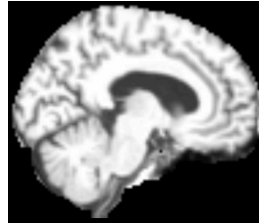


図5 画像B

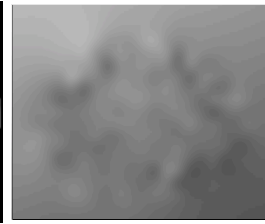


図6 画像H

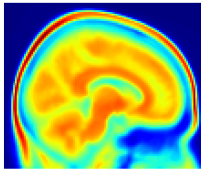


図7 画像A 小脳周辺

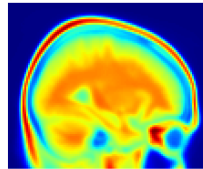


図8 画像A 視神経

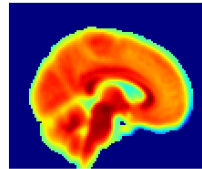


図9 画像B 小脳周辺

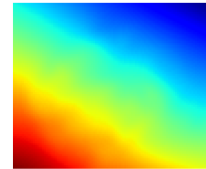


図10 画像H 後頭部

表1 3DCNN モデルの各画像での識別結果

Image	A	B	C	D	E	F	G	H	I
Acc	85	71	76	82	75	63	83	73	55

を用いて男女識別タスクを解くことが可能である。

4.1 画像Aの結果・考察

画像Aは、図7から大脳よりも頭蓋、小脳、脳梁や視床などが寄与していること、図8から視神経が寄与していることが確認できる。これらより、これまで画像Aは脳領域以外の影響も受けていると考えられており、実際に頭蓋やその他部位が寄与していることが見てとれる。識別精度だけは脳領域抽出前の画像によって高精度な判断をCNNモデルから得ることが可能である。

4.2 画像Bの結果・考察

画像Bの結果は、男女差がある脳の大きさを合わせ、頭蓋の影響を取り除いた上で男女識別が行えていることがわかる。図9から白質や灰白質といった脳全体、特に小脳や脳梁や視床などから推定していると確認できる。また、セグメンテーション画像と脳領域抽出後の画像による精度の比較として、 $D > C \geq E \geq B > F$ であることからヒートマップによる可視化以外にも白質が寄与していると考えられる。

4.3 画像Hの結果と考察

画像Hは、図10より、首部分の脳領域がない部分が寄与していることが確認できる。また、非線形レジストレーションを行うことで得られる画像Gと画像Hでは男女識別が可能であるが、画像Gから脳領域抽出を行うことで得られる画像Iでは識別不能である。これより、画像Gの頭蓋など、画像Hの脳領域以外の寄与部位の情報が男女識別に寄与しており、画像Iから脳領域

には識別可能情報が残っていないことが考えられる。

4.4 全体の考察

男女識別を行う上で、精度に影響を与える部位は頭蓋と小脳や脳梁や視床、視神経だと可視化によって判断可能である。

5 おわりに

MRI脳画像を用いた機械学習タスクは様々あり、男女識別は複数の前処理方法の画像によって高精度である。しかし、前処理方法やCNNモデルの説明可能性については未解決であるため、Grad-CAMによるヒートマップの作成と画像ごとの精度やヒートマップの違いなどによって本研究では説明性とする。可視化や判別確率、セグメンテーションごとの精度差によって判別可能なことは頭蓋と白質が精度へ寄与していることである。

IXI-Datasetsにおいては、頭蓋あり画像の場合、頭蓋や頭部の表面など脳以外の情報を用いて、識別に高く寄与していることがわかった。脳領域抽出画像では、脳領域が識別に高く寄与していることがわかった。変位画像では、後頭部の脳領域以外の部位が識別に高く寄与していることがわかった。

参考文献

- [1] Ramprasaath R. Selvaraju et al, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp618-626.
- [2] 笹谷, "機械学習による脳構造MRIからの個人属性の推定," 高知工科大学大学院修士論文, 2019.
- [3] 西森, "MRI脳画像の前処理方法の違いによる男女識別精度の比較," 高知工科大学修士学位論文, 2021.