

Guided Grad+Score CAMを用いた fMRI 画像分類用 3D-CNN の視覚化の検討

1275102 坂口 白磨 【コンピュータ構成学研究室】

Visualization of 3D-CNN for fMRI Image Classification Using Guided Grad+Score CAM

1275102 Hakuma Sakaguchi 【Advanced Computer Engineering Lab.】

1 はじめに

近年、脳の神経活動のメカニズムを、機械学習を用いた fMRI 画像のデコーディングによって解明を目指す研究が進められている。先行研究 [1] では、異なる視覚刺激画像が呈示された際の脳の賦活状況が記録された fMRI 画像を分類する 3DCNN モデルに対して Guided Grad CAM(GGC) を適用することで、3DCNN の分類根拠を視覚化し、脳の賦活領域を推定している。その結果、単変量解析で賦活が示された脳領域に加えて、賦活が示されなかった一部の脳領域についても賦活していると示唆される結果が得られた。GGC における Grad CAM(GC) 部分は、勾配マップに基づいて特徴マップの分類貢献度を算出し、分類根拠の視覚化を実現しているが、分類への貢献度が低い特徴マップに対しても高い分類貢献度を算出し、正しく分類根拠を視覚化できない場合が示されている。これに対して、ある特徴マップが示す領域でマスクされた元画像（マスク後画像）の分類確率に基づいて特徴マップの分類貢献度を算出し、分類根拠を視覚化する Score CAM が提案されている [2]。

本研究は、GC の各チャンネルの視覚化結果に対して、Score CAM による分類貢献度の算出結果を統合する手法 Guided Grad + Score CAM を提案し、その視覚化により推定された脳領域について評価・考察する。

2 Guided Grad CAM

GGC は、GuidedBackPropagation(以降 GBP) と GC を組み合わせた手法である。GC は、勾配マップに基づいた特徴マップの分類貢献度によって分類根拠を視覚化するが、この分類貢献度が必ずしも対象クラス分類確率と一致しないことが問題である [2]。図 1 に、各チャンネルの特徴マップで得られた領域（頭部のみ、頭部+胴体、胴体のみ）でマスクした後の画像例を示す。下欄に各マスク後画像の GC に基づく分類貢献度、および、各マスク後画像を CNN により分類した確率を示す。この例では、勾配マップに基づく頭部の特徴が最も分類貢献度が高いが、頭部だけのマスク後画像の分類確率は著しく低いため、実際には分類に貢献しない。

元画像	各特徴マップでマスクした画像(マスク後画像)		
	頭部	頭部+胴体	胴体
			
GCに基づく分類貢献度	高い	やや高い	やや高い
マスク画像の分類確率	0.003	0.999	0.997

図 1 GC の分類貢献度とマスク後画像の分類確率

3 Guided Grad+Score CAM

図 2 は提案手法の構成図である。提案手法は GC で計算される勾配マップに基づく分類貢献度を、Score CAM で計算されるマスク後画像の分類確率に基づく分類貢献度で補正する手法となっている。

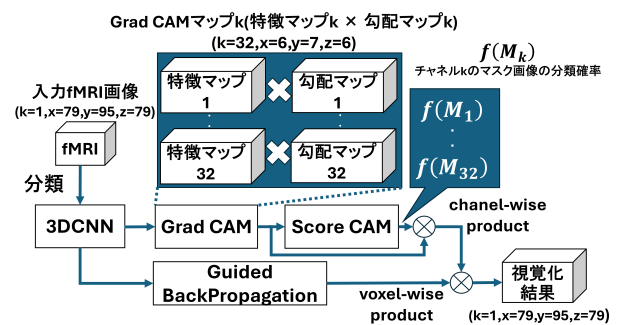


図 2 提案手法の全体図

まず GC を用いて各特徴マップ k の分類貢献度である勾配マップ k を重みづけした GC マップ k を取得する。次に Score CAM 部分では、GC マップ k を入力 fMRI 画像と同じサイズに拡大し、それぞれ入力 fMRI 画像にマスクした 32 枚 (チャンネル数) のマスク後画像を生成する。次にこれらを対象に 3D-CNN で分類した確率 $f(M_k)$ を求める。この値を GC マップ k の分類貢献度として、GC マップ k をチャンネル毎に補正し、最後に、GBP とボクセル単位でかけ合わせる。

4 提案手法の評価

顔カテゴリと物体カテゴリの2つの視覚刺激画像によるfMRI画像を分類する3DCNNモデルを学習し、テストデータの分類結果に対して、提案手法による分類根拠の視覚化、先行研究の再現と提案手法との比較のために、GGCについても視覚化を行った。視覚化結果については、まず並べ替え検定($p < 0.05$)を行い、各視覚刺激画像によって有意に賦活したと考えられるボクセルを推定した。次にROI分析を行い、各脳領域内に含まれるボクセル値について、単変量解析で有意に賦活が示された領域、GGCのみ有意に賦活が示された領域について評価を行った。

4.1 評価条件

使用したデータセットは、被験者50名による顔カテゴリと物体カテゴリの視覚刺激画像分類時の脳の賦活状況を記録したfMRI画像である。記録されたfMRI画像の前処理は、被験者ごとの脳の形状の違いを考慮し、MNI(Montreal Neurological Institute)座標に基づいて正規化した。また、fMRI画像のノイズ低減のため、各被験者ごとにfMRI画像の平均化を行った。

モデルの学習は、被験者45名分のデータを学習・検証データ、5名をテストデータとして9-fold交差検証を行った。学習時のハイパーパラメータは、バッチサイズ100、エポック500、学習率0.0001と設定し、学習を行った結果、分類精度は約89%となり、先行研究と同程度であることを確認した。本評価ではこの分類モデルを使用し、視覚化を行った。

4.2 推定された脳領域の評価

まず視覚化結果の定性的な評価として、FSLのrandomise関数を使用し、並べ替え検定を行った。並べ替え検定の結果、提案手法は、顔カテゴリと物体カテゴリともに脳の全域が賦活していると考えられるような結果で、先行研究のGGCと同様の結果となり、単変量解析の結果と比較しても、より多くの脳領域が賦活していると考えられる結果となった。

次に、ROI分析による定量的な評価を行った。ROI分析は、HarvardOxfordアトラスの計110脳領域内に含まれるボクセル値について、学習モデルと非学習モデルによる視覚化結果に有意差があるかウィルコクソンの符号順位検定($FDR\ q < 0.05$)によって評価を行った。ROI分析の結果、GGCによる視覚化では、顔カテゴリについて69領域、物体カテゴリについて48領域、提案手法による視覚化では、顔カテゴリについて68領域、物体カテゴリについて47領域が有意に賦活を示し、GGCと提案手法の賦活領域数はほぼ一致した。

表1は、顔カテゴリにおける単変量解析で有意な賦活の有無、非学習モデルに対してGGCと提案手法で有意な賦活の有無の8通り($a \sim h$)の組合せについて、ROI数を示している。組合せ a より、単変量解析で有意で

表1 提案手法およびGGCによる視覚化と単変量解析との関連

	a	b	c	d	e	f	g	h
univariate	**	**	**	**	-	-	-	-
GGC	*	*	-	-	*	*	-	-
Proposed	*	-	*	-	*	-	*	-
#ROI	10	0	1	2	57	2	0	38

*: $FDR\ q < 0.05$, **: $p < 0.05$

あった13領域のうち10領域(LOC, TOF, OCPなど)がGGCおよび提案手法で視覚化できていることから、3D-CNNの視覚化によっても脳領域の賦活状況のある程度推定できることが判る。

組合せ b が0領域であるのに対して組合せ c では1領域(L.PcC)であることから、提案手法がGGCよりも視覚化能力がやや優れていると言える。組合せ d では2領域について視覚化できていない。この点については、単変量解析に課題があるか、視覚化に課題があるかは不明であり、今後の課題として残されている。組合せ e では57領域であり、単変量解析では特定できないROIが視覚化により推定できる可能性を示唆している。組合せ f では2領域(R.TFCp, L.FOC)であったことから、GGCがノイズを誤推定している可能性がある。組合せ g が0領域であることから、提案手法がノイズに伴う誤推定を回避できていると考えられる。物体カテゴリについても顔カテゴリの結果と同様の傾向が見られた。

5 まとめ

本研究では、GCの各チャネルの視覚化結果に対して、Score CAMの分類貢献度の算出結果を統合する手法Guided Grad+Score CAMを提案した。ROI分析によって、GGCのみで有意な賦活が示された領域について、提案手法でも有意に賦活が示されたことから、深層学習モデルの視覚化が脳の賦活領域の説明性の向上に貢献できる潜在性を示唆する結果となった。

本提案手法は、特徴マップの各チャネルごとにマスク後画像の分類確率に基づく分類貢献度を導入したが、一方で、この分類貢献度に大きなばらつきが見られる特徴マップがいくつかあった。ばらつきの大きい特徴マップは分類にあまり貢献しないと考えられるため、そのような分類貢献度を低減することが今回の提案手法の改善点になると考える。

参考文献

- [1] N.Watanabe, et al., "Multi-modal deep neural decoding of visual object representation in humans," NeuroImage, Vol.~275, 120164, July 2023.
- [2] H.Wang, et al., "Score cam: Score-weighted visual explanations for convolutional neural networks," CVPRW 2020, pp.~111-119, 2020.