

# 画像刺激下の脳活動状態における推定精度向上への取り組み

理学専攻・情報科学コース

2140664

田口 遥香

## 1 はじめに

Yamins ら [1] によって視覚情報を処理する深層学習モデルの層と脳内の視覚情報処理の階層的処理の層の間に相関性があることが示されている。畳込みニューラルネットワーク (CNN) を用いて物体識別を行う際に、識別根拠となる物体の特定の箇所に焦点を当てる注意機構を導入することにより、視覚情報の一部によるカテゴリの識別を高精度で実現している。

このことを踏まえ、本研究では、CNN を用いて、画像刺激が与えられた際のヒト脳の状態を推定することを目的とし、特に脳活動状態を推定するための作業モデルとして、どのような画像処理深層学習モデルを利用することが効果的であるかについて脳活動状態の推定精度の観点から検証を行う。また、CNN を用いて画像刺激が与えられた際のヒト脳の状態を推定する際に、注意機構 [2] を強化した際の画像特徴量としない場合のものから予測されるそれぞれの予測脳活動を比較することにより、符号化モデルの入力情報となる属性の特徴量が直接識別に関わる場合とそうでない場合とでどのような違いがあるのかについて考察を行い、その考察の正当性を検証する。

## 2 画像特徴量に基づく脳活動状態の推定

脳活動測定時の刺激動画を静止画像として切り出したものを各種 CNN に入力として与え画像特徴量を抽出したもの、および、ResNet50 を用いたエンコーダ・デコーダで構成されるオートエンコーダを用いて抽出した画像特徴量から Ridge 回帰を用いて脳活動状態を推定する。

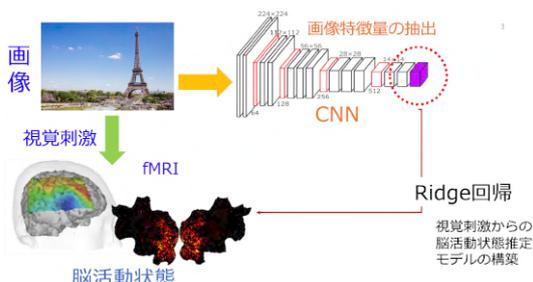


図 1: 概要

### 2.1 実験設定

ResNet-50 および ABN において ImageNet を使って事前に学習したモデルを使用する。脳活動データとしては、BOLD 信号を fMRI を用いて記録した脳神経活動データの皮質に相当する 70,933 次元のデータを使用した。画像特徴量は、各 CNN モデルの基づいて設定した。

### 符号化モデル

画像特徴量と脳活動の対データ (学習データ 7,200 対、評価データ 600 対) を用いて各ボクセルについて Ridge 回帰を行い、脳活動状態を推定した。推定され

た脳活動状態と計測した脳活動状態のピアソンの積率相関を調べ、帰無仮説「推定値と評価データの間には相関性は存在しない」を設定し、両側検定にて  $p$  値が 0.01 以下のものを対象として相関係数の算出および可視化を行なった。

### 2.2 結果と考察

表 1: 画像識別精度と脳活動状態推定精度

モデル名	Top-1 エラー率	相関値の平均	$p < 0.01$ の割合
AlexNet	43.45	0.269	0.140
ResNet-50	23.85	0.264	0.148
VGG19	27.62	0.263	0.132
VGG16+bn	26.63	0.262	0.130
ResNet-152	21.69	0.260	0.146
Densenet-161	22.35	0.258	0.147
Densenet-201	22.80	0.257	0.152
VGG16	28.41	0.257	0.136
VGG19+bn	25.76	0.256	0.135
Densenet-121	25.35	0.255	0.148
GoogleNet	30.22	0.248	0.132
AE	-	0.229	0.073

CNN 精度の Top1 エラー率の良さと有効なボクセルの相関値の平均との間に明確な相関性が存在しないことがわかった。これは画像識別能力の高さが必ずしも脳活動状態推定の精度とは強い相関があるとは言えないことを示唆していると考えられる。AlexNet においては、画像識別の精度に比べて脳活動状態推定において高い精度を出すことがわかった。また、上位 4 位までの画像識別モデルは、他のモデルと比較すると、層がそれほど深くない CNN モデルとなっており、脳活動状態推定においては比較的良好な結果を出していることから、「ヒト脳内では画像の特徴抽出においては必要以上の深層構造は取っていない」という仮説が考えられる。また、同様にヒト脳における画像特徴量抽出のメカニズムは多層によって表現される深層学習モデルほど複雑な工夫によって構築されているものではない可能性も示唆される。ROI<sup>1</sup> ごとの結果について着目すると、モデルごとに少し差異はあれど、強く反応している部位はほとんど共通しており、視覚野である後頭葉の部位である。一方で良い反応をしている ROI は同じであっても、モデルによって全て同じように反応しているわけではなく、ROI の良い成績の順位は異なっている。このことから 1 つのモデルで比較するよりも、より多くのモデル、被験者で比較することでより脳の解析に有用な情報を得られると考える。

### 3 注意機構を用いた脳内状態の推定

脳活動測定時の刺激動画を静止画像として切り出したものを、ResNet-50 で訓練された ABN に入力と

<sup>1</sup>ROI の分け方は [3] を参照した。

して与え、Attention Branch Network(ABN) [4] の存在があり/なしのそれぞれのものから抽出した画像特徴量から Ridge 線形回帰を用いて脳活動状態を推定する<sup>2</sup>。

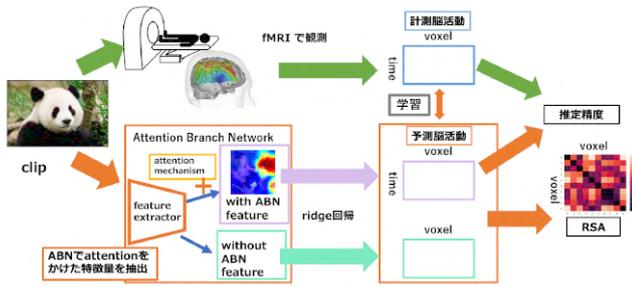


図 2: 実験概要

### 3.1 実験設定

ResNet-50 および ABN において ImageNet を使って事前に学習したモデルを使用する。脳活動データとしては、BOLD 信号を fMRI を用いて記録した脳神経活動データの皮質に相当する 70,933 次元のデータを使用した。画像特徴量は、ABN を通して attention あり/なしのものを抽出した。

#### 符号化モデル

画像特徴量と脳活動の対データ (学習データ 7,200 対, 評価データ 600 対) を用いて各ボクセルについて Ridge 回帰を行い、脳活動状態を推定した。推定された脳活動状態と計測した脳活動状態のピアソンの積率相関を調べ、帰無仮説「推定値と評価データの間には相関性は存在しない」を設定し、両側検定にて  $p$  値が 0.01 以下のものを対象として相関係数の算出および可視化を行なった。

#### RSA の設定

符号化モデルにより予測された脳活動状態が (評価データ)  $\times$  (ボクセル) として得られる。これに対して、(ボクセル)  $\times$  (ボクセル) の RDM を作ることで、予測脳活動状態に対するボクセル間の関係を捉える。

### 3.2 結果と考察

図 3 より attention あり/なしにおける UMAP の可視化では差が見られた。attention なしのものに関しては情報表現において偏りが見受けられないことに対して、attention ありのものは、偏りが見られた。また、flatmap の局在性の可視化の図より attention あり/なしのどちらも脳の領域ごとに類似性を持つことがわかった。attention なしのものは特に視覚にまつわる領域内の細かい類似性が見られ、attention ありのものに関してはより広範囲にわたった領域での類似性を見ることができる。

相関係数が高いほど視覚野などの青い部分が、相関係数が低いほど円環状の部分が反応していることがわかった。ABN を通すことで刺激を受け、反応している

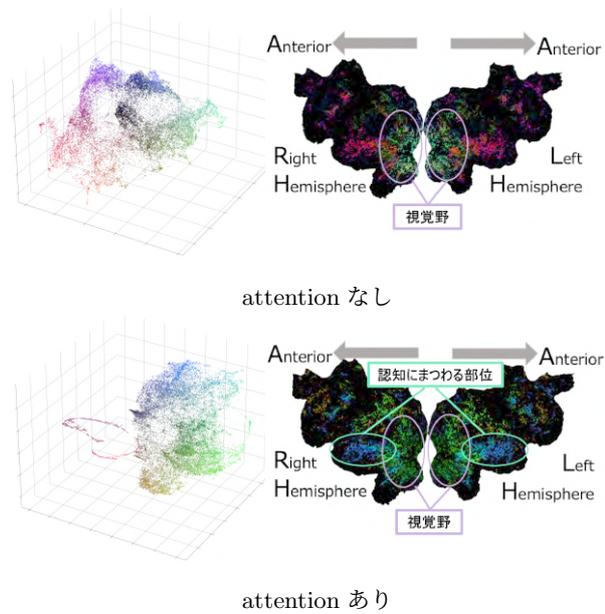


図 3:  $p < 0.01$  で棄却した予測脳活動に基づく分析結果の可視化

部分とそうでない部分とが UMAP で可視化した際に分かれている可能性があると考える。

## 4 おわりに

本研究では、被験者に画像刺激を与えた際の脳活動を fMRI で観測し、同じ画像を複数の画像識別深層学習モデルによって抽出された画像特徴量及び ABN を用いて抽出した attention あり/なしの画像特徴量と脳活動データとの対応関係をとらえる回帰モデルを構築した。また、各モデルにおける ROI ごとの相関性を調査したところ、ほとんどのモデルにおいて共通した視覚野に関わる ROI に対する推定精度が高くなっていった。一方で僅かではあるが、相関性の高い ROI が異なることも確認できた。

ABN によって重みづけされた画像特徴量から回帰することで脳活動状態推定においてどのように作用するのかについての実験を行い、その結果の正当性の検証を行なった。その結果、UMAP による三次元空間への可視化結果では領域における情報表現の偏りの違いを捉えることができた。

## 参考文献

- [1] D. L. K. Yamins, H. Hong, C. F. Cadieu, E. A. Solomon, D. Seibert, and J. J. DiCarlo. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, pp. 8619 – 8624, 10/2014 2014.
- [2] Dzmitry Bahdanau, Kyung Hyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. January 2015. 3rd International Conference on Learning Representations, ICLR 2015 ; Conference date: 07-05-2015 Through 09-05-2015.
- [3] Christophe Destrieux, Bruce Fischl, Anders M. Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, Vol. 53, No. 1, pp. 1–15, 2010.
- [4] Hiroshi Fukui, Tsubasa Hirakawa, Takayoshi Yamashita, and Hironobu Fujiyoshi. Attention branch network: Learning of attention mechanism for visual explanation. *CoRR*, Vol. abs/1812.10025, , 2018.

<sup>2</sup>本研究では、ABN の公開コード [https://github.com/machine-perception-robotics-group/attention\\_branch\\_network](https://github.com/machine-perception-robotics-group/attention_branch_network) に基づき、Pytorch の事前学習済の ResNet - 50 をベースラインモデルとし、再実装を行い、それによって得られたモデルを用いた。