

注意機構を導入した特徴量抽出に基づく画像刺激下の脳内状態推定

田口遥香[†]西田知史^{‡,*}西本伸志^{*,‡}小林一郎[†][†] お茶の水女子大学 [‡] 情報通信研究機構 CiNet * 大阪大学

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 画像刺激下の脳内状態推定

画像識別深層学習モデル ResNet-50 に、画像中の物体に対して焦点を当てる注意機構を追加するネットワークとして独立して学習する Attention Branch Network (ABN) [3] を導入し、抽出した画像特徴量から脳活動状態を予測する。この際、ABN の有無による画像特徴量からの脳活動状態予測における精度の変化について検証を行う。

ABN は、Attention Map と呼ばれる画像中の事物への注意 (Attention) を強化した領域図を生成することにより、認識結果の可視化を可能にし、重みを与えることで認識性能の向上を実現する。

画像刺激下の脳内状態を推定するために、画像特徴量から脳内状態を予測する符号化モデル [4] を構築する。予測した脳内活動の類似性を捉えるために Kriegeskorte ら

[5, 6] によって提案された表象類似解析 (Representation Similarity Analysis: RSA) を用いて解析を行う。RSA において、RDM (Representational Dissimilarity Matrix) という、対象とする項目の全組み合わせの非類似度を示す行列を作成する。RDM を作成することで、対象項目同士の相関を計算し様々な表現を比較することが可能になる。

3 実験

3.1 実験概要

脳活動測定時の刺激画像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 に入力として与え、ABN の介在があり/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*。

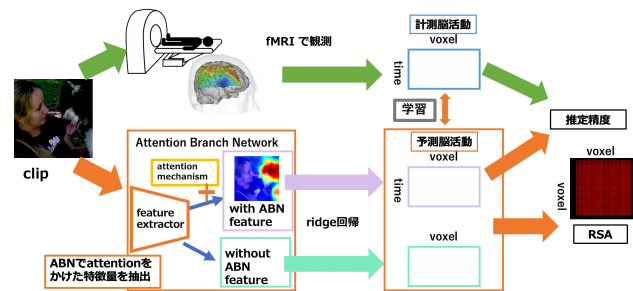


図 1: 実験概要

3.2 実験設定

使用データ

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。

An Approach to Estimation of Brain State evoked by Visual Stimuli by using the Features extracted through Attention Mechanism

[†]Haruka TAGUCHI (1720527@is.ocha.ac.jp)

[‡]Satoshi NISHIDA (s-nishida@nict.go.jp)

[†]Shinji NISHIMOTO (nishimoto.shinji.fbs@osaka-u.ac.jp)

[†]Ichiro KOBAYASHI (koba@is.ocha.ac.jp)

*本研究では、ABN の公開コード 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。

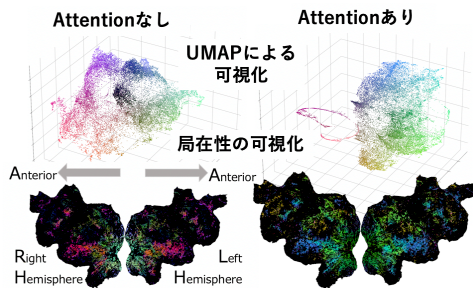


図 2: $p < 0.01$ で棄却した予測脳活動に基づく分析結果の可視化

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。

4 実験結果

UMAP を用いた三次元空間への描画した図、並びにその結果を Pycortex* の脳の flatmap にマッピングした結果を図 2 に示す。

attention が各ボクセルの相関係数によってどのような影響を及ぼすのかを調べるために、相関係数の値が 0 以上 0.1 以下, 0.1 以上 0.2 以下, 0.2 以上 0.3 以下, 0.3 以上 にわけて UMAP を用いた三次元空間への描画した図、その結果を Pycortex の脳の flatmap にマッピングした結果を図 3 に示す。

5 考察

図 2 において 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、局在性の可視化の図より attention あり/なしのどちらも脳の領

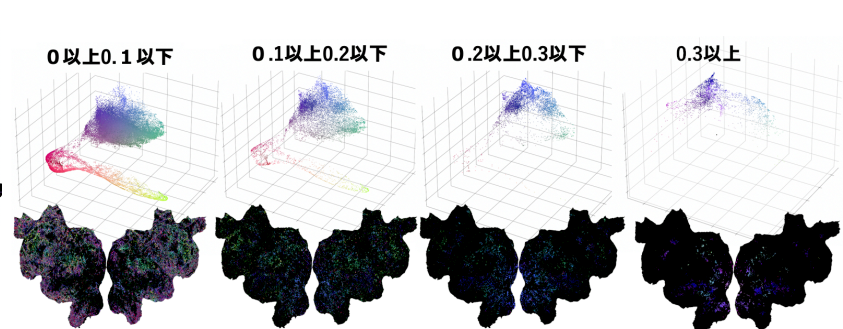


図 3: 相関係数による予測脳活動の可視化結果

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野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。

図 3 において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える。

6 まとめ

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 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。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野における情報表現の偏りの違いを捉えることができた。この結果は ABN が作用している可能性が高いと考える。

今後の課題として、この結果を踏まえて、さらに検証を進めることで正当性を深めていきたい。

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] Hiroshi Fukui, Tsubasa Hirakawa, Takayoshi Yamashita, and Hironobu Fujiyoshi. Attention Branch Network: Learning of Attention Mechanism for Visual Explanation. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2019.
- [4] Thomas Naselaris, Kendrick N. Kay, Shinji Nishimoto, and Jack L. Gallant. Encoding and decoding in fmri. *NeuroImage*, Vol. 56, No. 2, pp. 400–410, 2011. Multivariate Decoding and Brain Reading.
- [5] Nikolaus Kriegeskorte, Marieke Mur, and Peter Bandettini. Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, Vol. 2, p. 4, 2008.
- [6] Nikolaus Kriegeskorte and Rogier A Kievit. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci*, Vol. 17, No. 8, pp. 401–412, Aug 2013.

*<https://gallantlab.github.io/pycortex/>