ADVERSARIAL POLICY GRADIENT FOR LEARNING GRAPH-BASED REPRESENTATION IN HUMAN VISUAL PROCESSING

Anonymous authors
Paper under double-blind review

ABSTRACT

This article discusses the challenges in modeling the neural mechanisms underlying human visual processing and the use of graph-based representations to capture inter-region relationships in visual processing. While graphs have shown promise in analyzing neural responses, learning an optimal graph representation from limited data is challenging, as there is no ground truth to learn from. To address this, the authors propose a novel approach to graph-based representation using Adversarial Policy Gradient, which involves an adversarial game between the Policy Network and the Reward Network to generate a better graph representation.

1 INTRODUCTION

Compared to computer vision technologies, humans possess superior abilities in recognizing faces and interpreting scenes. The underlying reasons for the difference in performance between humans and artificially intelligent systems (AI) have been a topic of significant interest. Researchers in both neuroscience and AI have undertaken extensive efforts to investigate the neural mechanisms involved in human visual processing and how they differ from state-of-the-art AI(Kamitani & Tong, 2005; Du et al., 2018; Palazzo et al., 2020). In computer vision models, a latent representation space is typically generated, then mapped to the task at hand(Robert, 2019). In neuroscience, computer vision modeling techniques have been utilized to elucidate the neuronal mechanisms underlying visual processing (Encoding models) (Agrawal et al., 2014; Cui et al., 2021) and facilitate neuronal decoding (Decoding models) (Meng & Ge, 2022; Du et al., 2022). However, the generation of the representation space for human visual processing remains an area that requires further exploration.

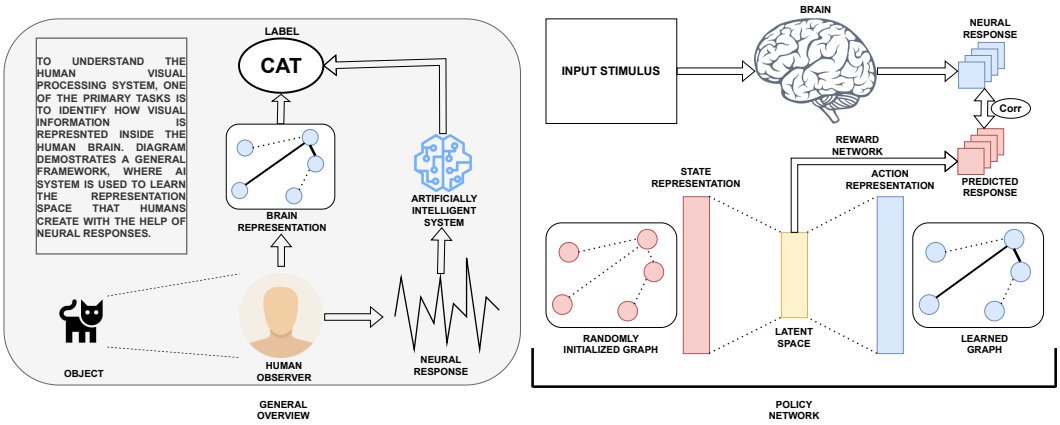

Figure 1: The General Architecture of our proposed methodology.

Computational modeling of neural responses, such as functional Magnetic Resonance Imaging (fMRI), has been widely employed to investigate the human visual system during visual tasks(Haynes & Rees, 2005; Thirion et al., 2007). Encoding models estimate neural responses from stimuli, while decoding models predict stimuli from neural responses(Naselaris et al., 2011;

Kriegeskorte, 2011; Haxby, 2012; Kay, 2018). In both cases, feature representations are generated to capture the stimulus-response relationship. Handcrafted (Wu et al., 2006; Kay et al., 2008) and learnable computer vision models(Agrawal et al., 2014; Güçlü & van Gerven, 2014; Wen et al., 2018a; Han et al., 2019; Wen et al., 2018b; Cui et al., 2021) have been used in encoding models. Similarly, learned feature spaces have been employed in decoding models to regenerate image stimuli or predict image stimulus categories (Du et al., 2018; Mozafari et al., 2020; Qiao et al., 2022; Du et al., 2022). However, slight alterations in the region of acquired neural response can significantly impact the functionality and behavior of the visual system. Recent studies have suggested that non-Euclidean data structures like graphs may more accurately represent this characteristic.

Graph representations have been widely employed for fMRI data analysis since the late 1990s (Conturo et al., 1999). Graph structures, such as functional connectivity(Dipasquale et al., 2017; Mohanty et al., 2020) and brain pathway analyses(Gilson et al., 2019; Deshpande & Wang, 2022), have successfully analyzed neural responses. Graph representations have recently been used for decoding neural data, including stimulus categorization and image reconstruction. In 2021, an interpretable Graph Neural Network (GNN) model, named BrainGNN, was proposed, successfully identifying neurological biomarkers for Autism Spectrum Disorder(Li et al., 2021). Another connectivity-based Graph Convolutional Network (cGCN) was proposed in 2021 for neural decoding. This architecture integrated functional connectivity within the GCN structure to extract spatial features from connectomic neighborhoods, outperforming Convolutional Neural Network (CNN) models(Li et al., 2022). In visual decoding, a GCN model was proposed in 2022, which selected Regions of Interest (ROIs) responsible for visual processing and estimated an optimal functional connectivity matrix among the ROIs, achieving comparable classification accuracy to prior state-of-the-art models(Meng & Ge, 2022). However, the available models are data-driven and, therefore, constrained by data availability.

In neural mechanism modeling, the latent space is crucial in encoding and decoding models. Graph-based representations have been proposed to capture inter-region relationships in visual processing better. However, learning an optimal graph representation from limited data is challenging, especially when there is no ground truth. As labeled data is unavailable, supervised learning algorithms are infeasible. This paper presents a new method for graph-based representation of human visual processing using Adversarial Policy Gradient. The approach involves using a randomly initialized graph structure to represent neural response space and two neural networks: a Policy Network and a Reward Network. The Policy Network is updated using the Policy Gradient algorithm based on rewards generated by the Reward Network. The Reward Network cannot be trained using Supervised Learning since the input graph is updated at every iteration. Hence, the proposed framework is trained using an adversarial game between the two networks to generate an optimal graph representation.

2 METHODOLOGY

The proposed methodology is shown in Figure 1. A Graph $\mathcal{G}$ is represented by a vector S_t, obtained by flattening the upper triangular part of the adjacency matrix of $\mathcal{G}$. We begin with a random graph $\mathcal{G}_0$ (V_0, E_0), where E_0 is the initial edges. We seek to find a graph $\mathcal{G}_f$ with final edges E_f, represented by the vector S_f, that accurately captures the behavior of inter-ROI neural responses. A vanilla neural network called the Policy Network, parameterized by random weights θ, takes in the state representation S_t of the graph $\mathcal{G}_t$ at instance t and outputs a new state representation S_{t+1}, considered an improved representation G_{t+1}. Policy Gradient algorithm was used to train the Policy Network. The Reward R is defined as the correlation between the predicted neural response from the state S_{t+1} and the original response. The Reward R_t at instance t is computed by another learnable neural network called the Reward Network. However, the Reward Network cannot be trained using supervised learning with the original responses as the target label since with every new instance $t + n$, the input state to the Reward Network changes to S_{t+n}. We address the problem with Adversarial training, where the Policy Network and the Reward Network play an adversarial game. The entire framework can be formulated as a minimax optimization problem to maximize the Reward and minimize the loss of the predicted neural response.

URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of ICLR 2023 Tiny Papers Track.

REFERENCES

Pulkit Agrawal, Dustin Stansbury, Jitendra Malik, and Jack Gallant. Pixels to voxels: Modeling visual representation in the human brain. 07 2014.

Thomas E. Conturo, Nicolas F. Lori, Thomas S. Cull, Erbil Akbudak, Abraham Z. Snyder, Joshua S. Shimony, Robert C. McKinstry, Harold Burton, and Marcus E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences*, 96(18): 10422–10427, 1999. doi: 10.1073/pnas.96.18.10422. URL https://www.pnas.org/doi/abs/10.1073/pnas.96.18.10422.

Yibo Cui, Kai Qiao, Chi Zhang, Linyuan Wang, Bin Yan, and li Tong. Gabornet visual encoding: A lightweight region-based visual encoding model with good expressiveness and biological interpretability. *Frontiers in Neuroscience*, 15, 02 2021. doi: 10.3389/fnins.2021.614182.

Gopikrishna Deshpande and Yun Wang. Noninvasive characterization of functional pathways in layer-specific microcircuits of the human brain using 7t fmri. *Brain Sciences*, 12:1361, 10 2022. doi: 10.3390/brainsci12101361.

Ottavia Dipasquale, Arjun Sethi, Maria Marcella Laganá, Francesca Baglio, Giuseppe Baselli, Prantik Kundu, Neil A. Harrison, and Mara Cercignani. Comparing resting state fmri de-noising approaches using multi- and single-echo acquisitions. *PLoS ONE*, 12, 2017.

Changde Du, Changying Du, Lijie Huang, and Huiguang He. Reconstructing perceived images from human brain activities with bayesian deep multiview learning. *IEEE Transactions on Neural Networks and Learning Systems*, PP:1–14, 12 2018. doi: 10.1109/TNNLS.2018.2882456.

Changde Du, Kaicheng Fu, Jinpeng Li, and Huiguang He. Decoding visual neural representations by multimodal learning of brain-visual-linguistic features, 10 2022.

Matthieu Gilson, Nikos Kouvaris, Gustavo Deco, Jean-François Mangin, Cyril Poupon, Sandrine Lefranc, Denis Rivière, and Gorka Zamora-López. Network analysis of whole-brain fmri dynamics: A new framework based on dynamic communicability. *NeuroImage*, 201:116007, 07 2019. doi: 10.1016/j.neuroimage.2019.116007.

Umut Güçlü and Marcel A. J. van Gerven. Deep neural networks reveal a gradient in the complexity of neural representations across the brain's ventral visual pathway. 11 2014. doi: 10.1523/JNEUROSCI.5023-14.2015. URL http://arxiv.org/abs/1411.6422http://dx.doi.org/10.1523/JNEUROSCI.5023-14.2015.

Kuan Han, Haiguang Wen, Junxing Shi, Kun Han Lu, Yizhen Zhang, Di Fu, and Zhongming Liu. Variational autoencoder: An unsupervised model for encoding and decoding fmri activity in visual cortex. *NeuroImage*, 198:125–136, 9 2019. ISSN 10959572. doi: 10.1016/j.neuroimage.2019.05.039.

James V. Haxby. Multivariate pattern analysis of fmri: The early beginnings, 8 2012. ISSN 10538119.

John-Dylan Haynes and Geraint Rees. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature neuroscience*, 8:686–91, 06 2005. doi: 10.1038/nn1445.

Yukiyasu Kamitani and Frank Tong. Kamitani, y. & tong, f. decoding the visual and subjective contents of the human brain. nature neurosci. 8, 679-685. *Nature neuroscience*, 8:679–85, 06 2005. doi: 10.1038/nn1444.

Kendrick Kay, Thomas Naselaris, Ryan Prenger, and Jack Gallant. Identifying natural images from human brain activity. *Nature*, 452:352–5, 04 2008. doi: 10.1038/nature06713.

Kendrick N. Kay. Principles for models of neural information processing, 10 2018. ISSN 10959572.

Nikolaus Kriegeskorte. Pattern-information analysis: From stimulus decoding to computational-model testing. *NeuroImage*, 56:411–421, 5 2011. ISSN 10538119. doi: 10.1016/j.neuroimage.2011.01.061.

Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence Staib, Pamela Ventola, and James Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 09 2021. doi: 10.1016/j.media.2021.102233.

Yu Li, Xin Zhang, Jingxin Nie, Guowei Zhang, Ruiyan Fang, Xiangmin Xu, Zhengwang Wu, Dan Hu, Li Wang, Han Zhang, Weili Lin, and Gang Li. Brain connectivity based graph convolutional networks for infant age prediction. *IEEE Transactions on Medical Imaging*, PP:1–1, 05 2022. doi: 10.1109/TMI.2022.3171778.

Lu Meng and Kang Ge. Decoding visual fmri stimuli from human brain based on graph convolutional neural network. *Brain Sciences*, 12:1394, 10 2022. doi: 10.3390/brainsci12101394.

Rosaleena Mohanty, William Sethares, Veena Nair, and Vivek Prabhakaran. Rethinking measures of functional connectivity via feature extraction. *Scientific Reports*, 10, 01 2020. doi: 10.1038/s41598-020-57915-w.

Milad Mozafari, Leila Reddy, and Rufin van Rullen. Reconstructing natural scenes from fmri patterns using bigbigan. *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2020.

Thomas Naselaris, Kendrick N. Kay, Shinji Nishimoto, and Jack L. Gallant. Encoding and decoding in fmri, 5 2011. ISSN 10538119.

Simone Palazzo, Concetto Spampinato, Isaak Kavasidis, D. Giordano, Joseph Schmidt, and Mubarak Shah. Decoding brain representations by multimodal learning of neural activity and visual features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP:1–1, 05 2020. doi: 10.1109/TPAMI.2020.2995909.

Kai Qiao, Jian Chen, Linyuan Wang, Chi Zhang, Li Tong, and Bin Yan. Reconstructing natural images from human fmri by alternating encoding and decoding with shared autoencoder regularization. *Biomedical Signal Processing and Control*, 2022.

Thomas Robert. *Improving Latent Representations of ConvNets for Visual Understanding*. PhD thesis, 10 2019.

Bertrand Thirion, Edouard Duchesnay, Edward Hubbard, Jessica Dubois, Jean-Baptiste Poline, Denis Lebihan, and Stanislas Dehaene. Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, 33:1104–16, 01 2007. doi: 10.1016/j.neuroimage.2006.06.062.

Haiguang Wen, Junxing Shi, Wei Chen, and Zhongming Liu. Deep residual network predicts cortical representation and organization of visual features for rapid categorization. *Scientific Reports*, 8, 12 2018a. ISSN 20452322. doi: 10.1038/s41598-018-22160-9.

Haiguang Wen, Junxing Shi, Yizhen Zhang, Kun Han Lu, Jiayue Cao, and Zhongming Liu. Neural encoding and decoding with deep learning for dynamic natural vision. *Cerebral Cortex*, 28:4136–4160, 12 2018b. ISSN 14602199. doi: 10.1093/cercor/bhx268.

Michael Wu, Stephen David, and Jack Gallant. Complete functional characterization of sensory neurons by system identification. *Annual review of neuroscience*, 29:477–505, 02 2006. doi: 10.1146/annurev.neuro.29.051605.113024.

A ADVERSARIAL POLICY GRADIENT

The Modified Policy Gradient algorithm is used to update the Policy Network by maximizing the expected reward. The Policy Network generates an updated state representation given the current state representation, which is used to generate the graph. Let θ be the parameters of the Policy Network, s_t be the current state representation and s_{t+1} be the next state representation. The Policy Network can be represented as a function $\pi(s_{t+1}|s_t, \theta)$, where s_{t+1} is the next state derived at time t given state s_t.

The reward is generated by a separate neural network called the Reward Network, which predicts the neural response based on the current graph representation. Let W be the parameters of the Reward Network, and r_t be the reward at time t. The Reward Network can be represented as a function $R(s_t, W)$.

The goal is to maximize the expected reward by updating the Policy Network parameters θ. The update rule can be written as:

$$\Delta\theta = \alpha \nabla_\theta \log \pi(s_{t+1}|s_t, \theta) r_t$$

where α is the learning rate and ∇_θ is the gradient with respect to θ. The gradient of the expected reward with respect to θ is calculated using the Policy Gradient theorem:

$$\nabla_\theta J(\theta) = E_\pi[\nabla_\theta \log(s_{t+1}|s_t, \theta) r_t$$

where $J(\theta)$ is the objective function to be maximized. The expectation is taken over all possible current state-next state pairs under the policy π.

To update the Reward Network, we use the mean squared error loss between the predicted neural response and the actual neural response as the loss function. Let y_t be the actual neural response and y_hat_t be the predicted neural response. The loss function can be written as:

$$L(W) = (y_t - \hat{y}_t)^2$$

The parameters of the Reward Network, W, are updated using stochastic gradient descent by minimizing the loss function:

$$\Delta W = -\alpha \nabla_W L(W)$$

where α is the learning rate and ∇_W is the gradient with respect to W.

The minimax optimization problem for the proposed Adversarial Policy Gradient approach can be formulated as follows:

$$min_W max_\theta Z(W, \theta)$$

where θ represents the parameters of the Policy Network, W represents the parameters of the Reward Network, and $Z(W, \theta)$ is the objective function defined as:

$$Z(W, \theta) = E_\tau[\sum_{t=1}^{T} R(s_t, s_{t+1}) - \beta KL(\pi_\theta || \pi_W | s_t)]$$

where τ represents a trajectory of states and actions, β is a hyperparameter that controls the strength of the KL divergence regularization term, $KL(\pi_\theta || \pi_W | s_t)$ represents the KL divergence between the Policy Network and the Reward Network's policies, and π_θ and π_W represent the probability distributions over actions or states, parameterized by θ and W, respectively.

Overall, the Adversarial Policy Gradient algorithm involves an adversarial game between the Policy Network and the Reward Network, where the Policy Network generates an updated state representation to improve the graph representation, and the Reward Network generates a reward based on

the current graph representation. The Policy Network is updated using the Policy Gradient algorithm with the reward generated from the Reward Network. The adversarial game continues until the graph representation adequately captures the behavior of inter-ROI neural responses.