Research Statement

Zijin Gu

January 2, 2023

Neuroscience data, like brain connectivity and activity, are usually big in terms of both scale (number of neurons and nodes) and complexity (communications between neurons). It's natural that we need to apply models with more flexibility to those data in order to get a clearer picture of what's going on inside the brain. Probabilistic and statistical models, as well as machine learning (ML) based artificial intelligence (AI) approaches have progressed from purely theoretical models to real implementations, making it a great tool to handle such big data. This opens up possibilities to revisiting old problems and exploring new strategies on a broader canvas.

Standing at the intersection of the machine learning and neuroscience, my past research mostly focuses on characterizing the unity and variety of the human brain, and particularly, I am interested in understanding and interpreting neural activation and connectivity patterns with the help from ML and statistical tools. For future research, I plan to develop better ML algorithms based on principles of the brain and focus more on the medical and health application side.

Past Research

"*The beauty of the universe consists not only of unity in variety, but also of variety in unity*", said by Italian philosopher Umberto Eco. Unique brains and unique minds characterize each and every human being. At the same time there's brain and mind unity in human diversity. Connectivity and activity are two properties of human brain that can be viewed as "finger-prints" to identify individual from a large population. Brain connectivity refers to a pattern of either anatomical links, e.g. structural connectivity, or statistical dependencies, e.g. functional connectivity, between distinct units within a nervous system. The units can have different scales such as individual neurons, neuronal populations, or anatomically segregated brain regions. Brain activity refers to the signals neurons send and receive when performing a given cognitive or physiological task, which are largely constrained by connectivity.

Heritability and interindividual variability analysis of brain structural and functional coupling. The relationship between structure and function is one of the key concepts in many fields of study. In neuroscience particularly, studies of form and function, e.g., whole-brain, anatomical (white matter) and physiological (functional co-activation) brain networks, have led to fundamental discoveries in human populations [22, 13, 16]. Joint structure-function modeling, e.g. neural mass models, network diffusion models, graph theoretical and statistical approaches, can help understand the way of communications between neural populations and spreading of functional activation via the structure backbone, and identify function-specific subnetworks or multi-modal mechanisms of recovery after injury [5, 1, 15, 24].

Given that most modeling approaches treated structure and functional connectivity globally which ignore the regional variability, we analyzed MRI data from Human Connectome Project [23] and came up with the regional structure-function coupling, computed as the region-wise Spearman rank correlation between structural connectivity and function connectivity. Using a generalized linear model (GLM) approach, we quantified regional associations between structure-function coupling and age, sex and cognition. We also assessed the heritability of structure-function coupling used a linear mixed effects (LME) model developed to disentangle inter- versus intra-subject variation. Through rigorous statistical tests with careful measurements of significance and several replication study and sensitivity analyses, we found that regional structure-function coupling has an idiosyncratic characteristic and may be influenced by genetic factors. This work was published in *Nature Communications* [11].

Construction of personalized neural encoding and decoding models. As a fundamental area in neuroscience, the neural coding [20] problem asks two questions: 1) can we predict the response of a neural population to a given stimulus? and 2) is it possible to decode the neural activity, e.g. reconstruct the external stimuli using only a limited number of observed neural responses [18]? Neural coding of vision in human studies, which try to approximate brain responses to images, have become an important subarea with the increased use of non-invasive functional MRI (fMRI) techniques [17] and recent advances in large-scale fMRI datasets with brain responses to visual stimuli [2].
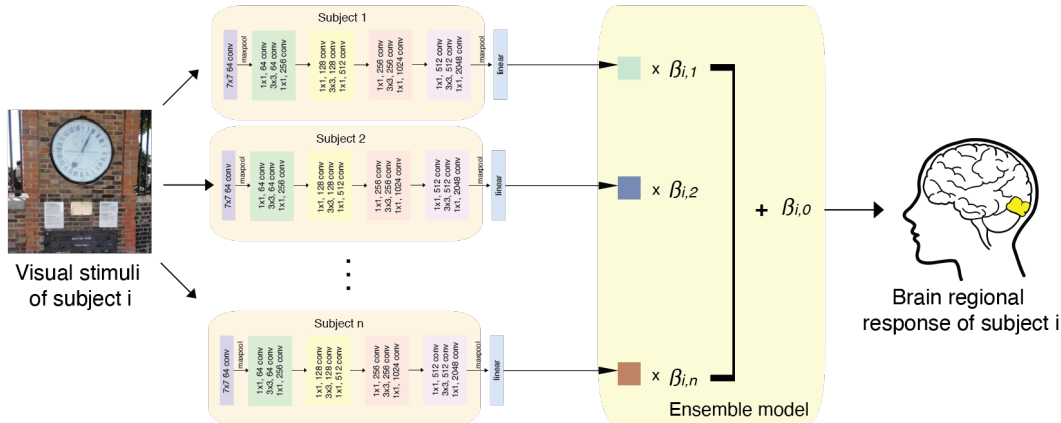

Figure 1: The proposed linear ensemble encoding model architecture.

1) Neural encoding. Neural encoding studies try to capture the relationship between stimuli and brain response by predicting brain response to stimuli. With thousands of training data from densely-sampled fMRI datasets, accurate individual-level encoding models can be constructed [25, 12, 4, 14]. However, obtaining large data from one individual usually requires excessive resources, e.g. ~ 40 hours of fMRI. Such experiments are restricted thus it's difficult to predict the response given a novel individual. Our work aims to solve this problem and provided a way to easily construct personalized encoding model with small data. e.g. ~ 40 minutes of fMRI. The method, named linear ensemble shown in Figure 1, utilizes a group of n existing pretrained encoding models to obtain a set of predicted activations, which are combined via a linear model fitted via ordinary least squares to optimally predict the query subject's regional brain response. In addition, we brought up a new evaluation metric as a complement to accuracy, called prediction consistency, to quantify the model's ability to preserve inter-individual variability. The proposed linear ensemble approach, simple but effective, has the best balance of accuracy and consistency comparing with several other approaches. Importantly, trained with only 1.5% the size of the large model's training data, it is able to achieve similar accuracy as the large model. The results show the potential to use existing densely-sampled data from a single individual, to efficiently create accurate, personalized encoding models for new individuals scanned under different experimental conditions. This work was published in *Communications Biology* [9].

2) Neural decoding. As opposed to neural encoding, neural decoding studies approach the relationship between stimuli and brain response by reconstructing stimuli from brain response. This is a more challenging task due to the low signal-to-noise ratio (SNR) and limited resolution of functional MRI data, and the high complexity of natural images. One observation from previous work is that they usually handle fMRI data by selecting voxels that belong to certain visual regions-of-interest (ROIs) and then flattened them into 1D vectors [21, 3, 6, 19, 7]. The weaknesses of such approach are the inconsistency across subjects resulted from the subjectivity in ROI definitions and selection and the ignorance of the spatial topology of the 2D cortical areas. We addressed these weaknesses by proposing a shared decoding framework called Cortex2Image, shown in Figure 2, which uses surface convolution to exploit spatial information in brain activity patterns and is combined trained with IC-GAN end-to-end to improves computational efficiency comparing with previous works. Image reconstructions from Cortex2Image achieve state-of-the-art accuracy in matching high-level semantics, while capturing much of the image details. This work is currently under review, but a preprint is publicly available [8].
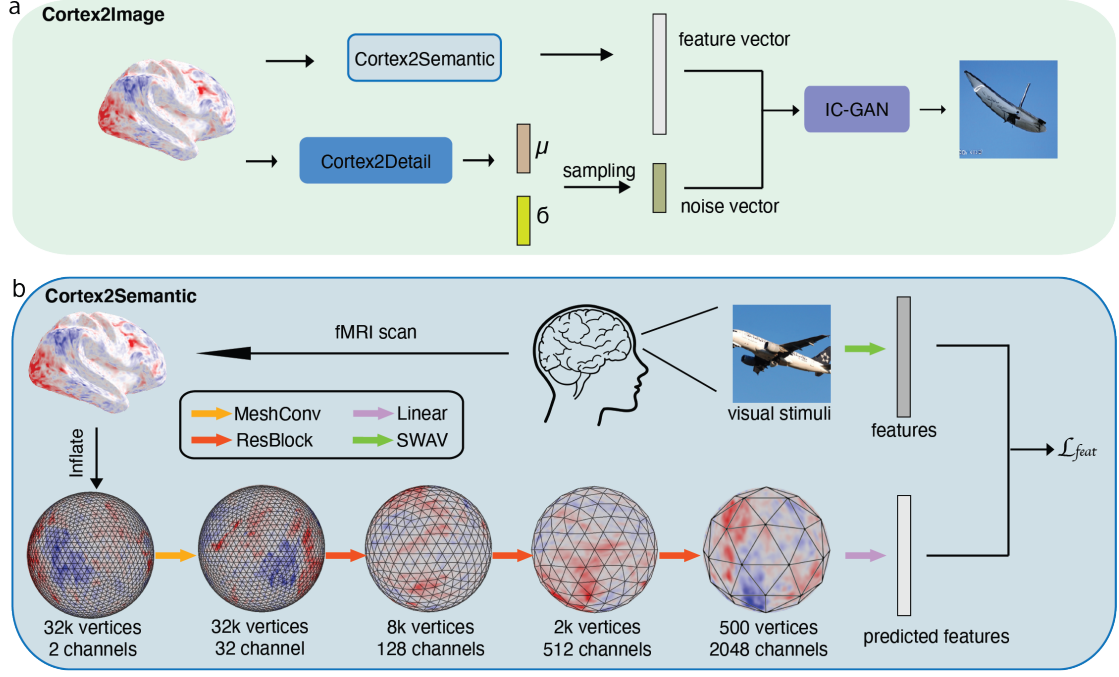
Figure 2: The proposed Cortex2Image decoding framework.

NeuroGen: an optimized image synthesis framework   Optimal stimulus design is a natural extension of the neural encoding problem where we want to adaptively select or generate the optimal stimuli for characterizing the response function. To this end, we proposed a generative framework called NeuroGen, which allows synthesis of images that are optimized to achieve specific, predetermined brain activation responses in the human brain, shown in Figure 3. In NeuroGen, the encoding model is concatenated with a pretrained generative network (we used BigGAN-deep in our paper but it can be replaced by any state-of-the-art GAN/VAE/diffusion models) to synthesize images that are predicted to optimally match a desired response pattern (e.g. maximizing predicted activation in the fusiform face area). We demonstrated that the NeuroGen framework is able to detect and amplify differences in regional and individual human brain response patterns to visual stimuli, and create synthetic images predicted to achieve response patterns not achievable by the best-matching natural images. This work was published in *NeuroImage* [10].
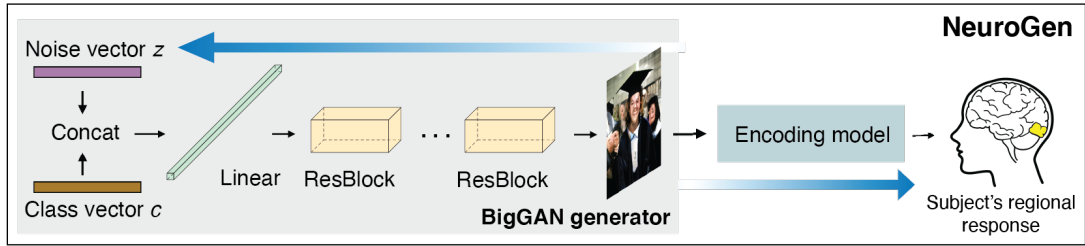

Figure 3: The proposed NeuroGen framework.

We also did follow-up study to test whether the selected and generated images can achieve desired fMRI activation where we recruited 6 subjects and showed them the natural and synthetic images while they were in the scanner. We used LME models to examine the statistical difference between several image conditions, e.g., images designed to achieve maximum activation and average activation on group level, images designed to achieve maximum activation specifically for one individual and for other individuals. We found that the ability of obtaining significant differences between the conditions we designed is largely correlated with the encoding model accuracy, and the human brain regional level response can be driven by the generative networks. A manuscript related to this work is in preparation.

3

Future Research

While my past work is mainly about applying ML and statistics to neuroscience problems, I would like to explore more in the reverse direction, where developing better ML/AI algorithms based on principles of the brain is one of my research interests. Neuroscience, or broadly biology, no doubt has the potential to push the boundaries of what is currently possible with AI by developing algorithms and systems that can learn and adapt like the advanced biological entities, most notably humans. With the study of human intelligence, I would like to 1) recognize the limitations of current ML and identify potential areas for improvement, such as robustness to adversarial attacks and spurious features; 2) boost the performance of ML algorithms to be more efficient and effective; and 3) build multi-task AI systems for handling a wider range of tasks with more adaptability.

On another front, AI is in a state of rapid development and expansion. Although significant progresses have been made in neuroscience over the past few decades, there remains much about the brain that is yet to be understood, an area in which AI/ML can be of great assistance. Besides analyzing large scale neuroimaging data for neuroscientific discoveries that I did during my Ph.D, I also plan to focus more on medical and healthcare sides which have potential application scenarios, such as developing new diagnostic tools and treatments for diseases, building AI systems to improve access to care, and optimizing the design of human-computer interfaces.

I plan to devote myself into one or more research directions that are mentioned above. To expand upon the first line of work, I plan to work closely with AI scientists/engineers and neuroscientists to identify the current weaknesses of ML algorithms and seek possible solutions from human intelligence. To advancing the second line of work, I plan to search opportunities to collaborate with hospitals and clinicians to learn more about the characteristics of different types of diseases and develop suitable AI algorithms to help with prevention, diagnoses, and treatment.

References

[1] Farras Abdelnour, Henning U Voss, and Ashish Raj. Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *Neuroimage*, 90:335–347, 2014.

[2] Emily J Allen, Ghislain St-Yves, Yihan Wu, Jesse L Breedlove, Jacob S Prince, Logan T Dowdle, Matthias Nau, Brad Caron, Franco Pestilli, Ian Charest, et al. A massive 7t fmri dataset to bridge cognitive neuroscience and artificial intelligence. *Nature neuroscience*, pages 1–11, 2021.

[3] Roman Beliy, Guy Gaziv, Assaf Hoogi, Francesca Strappini, Tal Golan, and Michal Irani. From voxels to pixels and back: Self-supervision in natural-image reconstruction from fmri. *Advances in Neural Information Processing Systems*, 32, 2019.

[4] Radoslaw Martin Cichy, Aditya Khosla, Dimitrios Pantazis, Antonio Torralba, and Aude Oliva. Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific reports*, 6(1):1–13, 2016.

[5] Gustavo Deco, Viktor K Jirsa, and Anthony R McIntosh. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12(1):43–56, 2011.

[6] Tao Fang, Yu Qi, and Gang Pan. Reconstructing perceptive images from brain activity by shape-semantic gan. *Advances in Neural Information Processing Systems*, 33:13038–13048, 2020.

[7] Guy Gaziv, Roman Beliy, Niv Granot, Assaf Hoogi, Francesca Strappini, Tal Golan, and Michal Irani. Self-supervised natural image reconstruction and large-scale semantic classification from brain activity. *NeuroImage*, 254:119121, 2022.

[8] Zijin Gu, Keith Jamison, Amy Kuceyeski, and Mert Sabuncu. Decoding natural image stimuli from fmri data with a surface-based convolutional network. *arXiv preprint arXiv:2212.02409*, 2022.

[9] Zijin Gu, Keith Jamison, Mert Sabuncu, and Amy Kuceyeski. Personalized visual encoding model construction with small data. *Communications Biology*, 5(1):1382, 2022.

[10] Zijin Gu, Keith Wakefield Jamison, Meenakshi Khosla, Emily J Allen, Yihan Wu, Ghislain St-Yves, Thomas Naselaris, Kendrick Kay, Mert R Sabuncu, and Amy Kuceyeski. Neurogen: activation optimized image synthesis for discovery neuroscience. *NeuroImage*, 247:118812, 2022.

[11] Zijin Gu, Keith Wakefield Jamison, Mert Rory Sabuncu, and Amy Kuceyeski. Heritability and interindividual variability of regional structure-function coupling. *Nature Communications*, 12(1):1–12, 2021.

[12] Umut Güçlü and Marcel AJ van Gerven. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *Journal of Neuroscience*, 35(27):10005–10014, 2015.

[13] Ann M Hermundstad, Danielle S Bassett, Kevin S Brown, Elissa M Aminoff, David Clewett, Scott Freeman, Amy Frithsen, Arianne Johnson, Christine M Tipper, Michael B Miller, et al. Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proceedings of the National Academy of Sciences*, 110(15):6169–6174, 2013.

[14] Meenakshi Khosla, Gia H Ngo, Keith Jamison, Amy Kuceyeski, and Mert R Sabuncu. Cortical response to naturalistic stimuli is largely predictable with deep neural networks. *Science Advances*, 7(22):eabe7547, 2021.

[15] A Kuceyeski, S Shah, JP Dyke, S Bickel, F Abdelnour, ND Schiff, HU Voss, and A Raj. The application of a mathematical model linking structural and functional connectomes in severe brain injury. *NeuroImage: Clinical*, 11:635–647, 2016.

[16] Bratislav Mišić, Richard F Betzel, Marcel A De Reus, Martijn P Van Den Heuvel, Marc G Berman, Anthony R McIntosh, and Olaf Sporns. Network-level structure-function relationships in human neocortex. *Cerebral Cortex*, 26(7):3285–3296, 2016.

[17] Thomas Naselaris, Kendrick N Kay, Shinji Nishimoto, and Jack L Gallant. Encoding and decoding in fmri. *Neuroimage*, 56(2):400–410, 2011.

[18] Liam Paninski, Jonathan Pillow, and Jeremy Lewi. Statistical models for neural encoding, decoding, and optimal stimulus design. *Progress in brain research*, 165:493–507, 2007.

[19] Ziqi Ren, Jie Li, Xuetong Xue, Xin Li, Fan Yang, Zhicheng Jiao, and Xinbo Gao. Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning. *NeuroImage*, 228:117602, 2021.

[20] Fred Rieke, David Warland, Rob de Ruyter Van Steveninck, and William Bialek. *Spikes: exploring the neural code*. MIT press, 1999.

[21] Guohua Shen, Tomoyasu Horikawa, Kei Majima, and Yukiyasu Kamitani. Deep image reconstruction from human brain activity. *PLoS computational biology*, 15(1):e1006633, 2019.

[22] Lucina Q Uddin. Complex relationships between structural and functional brain connectivity. *Trends in cognitive sciences*, 17(12):600–602, 2013.

[23] David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, Wu-Minn HCP Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.

[24] Bertha Vázquez-Rodríguez, Laura E Suárez, Ross D Markello, Golia Shafiei, Casey Paquola, Patric Hagmann, Martijn P Van Den Heuvel, Boris C Bernhardt, R Nathan Spreng, and Bratislav Misic. Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, 116(42):21219–21227, 2019.

[25] Daniel LK Yamins and James J DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature neuroscience*, 19(3):356–365, 2016.