Deep Image Reconstruction from Natural Scenes Dataset

Students: James Du, Michael La, Connor Levenson, Elise Nguyen
Sponsor: Dr. Thomas Sprague, Sikun Lin

UC SANTA BARBARA

SPRAGUE LAB

Data Science Capstone

Introduction / Background

- Goal: Gain insight into how different visual stimuli produce activity into different regions of the brain
- Task: Create a model in which fMRI activity can be reconstructed into the original visual stimulus image
- Functional Magnetic Resonance Imaging (fMRI): measures brain activity associated with blood flow

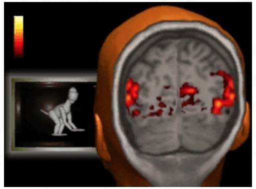

Figure 1. Example of fMRI scan and visual stimulus.

Natural Scenes Dataset

- Large-scale fMRI dataset including fMRI scans, physiological & eye tracking data, anatomical & functional neuroimaging measures resting-state data, ROIs, diffusion data,, and behavioral measures
- 8 adult subjects viewed 9,000 - 10,000 color natural scenes over the course of 30-40 scan sessions
- Responses to a total of 70,566 distinct natural scene images

References	Dataset	Number of Subjects	Number of Image Stimuli
VanRullen and Reddy (2019)	*Faces*	4	108
Kay et al. (2008)	*vim-1*	2	1,870
Horikawa and Kamitani (2017)	*Generic object decoding*	5	1,250
Shen et al. (2019)	*Deep image reconstruction*	3	1,250
Allen et al. (2021)*	*Natural Scenes Dataset*	8	70,566

Figure 2. Datasets used for the Deep Image Reconstruction task. Note that ours (*NSD*) is much larger and has not yet been used for DIR.

Methodology

- From the NSD Dataset, we built a joint embedding model by encoding the fMRI vectors to CLIP Space, which is a feature space where similar concepts are mapped to the same general regions, and different concepts are mapped far apart.
- From this shared feature space, we used LAFITE, which includes a generative adversarial network, which reconstructs an image based on the encoded feature space.
- From the shared feature space, we map images back to fMRI vector space to gain a better understanding of which voxels are responsible for certain features in images.

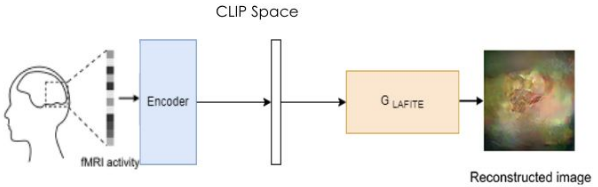

Figure 3. Model Architecture (only the blue portion in trainable)

Conclusions and Future Work

- Current text-to-image generation methods can be extended to the Deep Image Reconstruction task
- In the future, we would like to add a term to the loss function which explicitly uses the output from some Generator
- We would also like to use ideas from current state of the art text-to-image models such as Imagen or Dalle

Sample Final Results

- We find that we can generate images with high accuracy for general features, but lose precision in the finer details.
- We typically find that our generated images maintain the same subject feature as the ground truth images.

Figure 4. Sample ground truths (left) and outputs (right)

References / Acknowledgements

1. Rakhimberdina, Zarina, et al. "Natural Image Reconstruction From FMRI Using Deep Learning: A Survey." Frontiers,
2. Allen, E.J., St-Yves, G., Wu, Y. et al. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. Nat Neurosci 25, 116–126 (2022).

This work was partially supported by NSF Award #1924205.