Self-Supervised Natural Scene Reconstruction and Rich Semantic Classification from Brain Activity

Guy Gaziv1,*, Roman Belyi1,*, Niv Granot1,*, Assaf Hoogi1, Francesca Strappini2, Tal Golan3, Michal Irani 1

1. Dept. of Computer Science and Applied Math, The Weizmann Institute of Science
2. Dept. of Neurobiology, The Weizmann Institute of Science
3. Zuckerman Institute, Columbia U.

*Equal contributors
Corresponding author: michal.irani@weizmann.ac.il

Problems:
• Insufficient training data (~1200 pairs from 150 semantic classes) 
  → does not span natural images and their rich semantic space
• Differs between brains 
  → cannot be combined across subjects

Goal:
Reconstruct images and decode semantic categories from fMRI recordings

We propose:
Training on many additional “unpaired” data:
(i) Images without fMRI recordings 
  → Adapts to statistics of Natural Images
(ii) fMRI recordings without images
  → Adapts to statistics of input test data

Adding self-supervision on unpaired data improves reconstruction

Two training phases:
E-D: trained on unpaired images (with no fMRI)
  (50,000 natural images from 1000 rich ImageNet classes without their class labels)
  → Adapts to the statistics of natural images and novel semantic classes
D-E: trained on unpaired fMRI (unknown image)
  (Specifically, the unpaired test fMRI, without any corresponding images)
  → Adapts to the statistics of the new (unpaired) test data

Our Results (for ‘fMRI on ImageNet’ [Horikawa et al. 2017]):
Semantic classification results – predicting class label out of 1000+ classes

Adding self-supervision on unpaired data allows classifying against a gallery of class representatives

Classifying against a gallery of class representatives

Exploring visual features and cross-subject classification in rich and novel semantic categories

Reconstruction results for the entire test data (50 images)