Abstract:

Deep Learning introduced powerful research tools for studying visual representation in the human brain. Here, we harnessed those tools for two branches of research: 1. The primary branch focuses on brain decoding: reconstructing and semantically classifying observed natural images from novel (unknown) fMRI brain recordings. This is a very difficult task due to the scarce supervised “paired” training examples (images with their corresponding fMRI recordings) that are available, even in the largest image-fMRI datasets. We present a self-supervised deep learning approach that overcomes this barrier. This is obtained by enriching the scarce paired training data with additional easily accessible “unpaired” data from both domains (i.e., images without fMRI, and fMRI without images). Our approach achieves state-of-the-art results in image reconstruction from fMRI responses, as well as unprecedented large-scale (1000-way) semantic classification of never-before-seen classes. 2. The secondary branch of research focuses on face representation in the human brain. We studied whether the unique structure of the face-space geometry, which is defined by pairwise similarities in activation patterns to different face images, constitutes a critical aspect in face perception. To test this, we compared the pairwise similarity between responses to face images of human-brain and of artificial Deep Convolutional Neural Networks (DCNN) that achieve human-level face recognition performance. Our results revealed a stark match between neural and intermediate DCNN layers' face-spaces. Our findings support the importance of face-space geometry in enabling face perception as well as a pictorial function of high-order face-selective regions of the human visual cortex.