Predicting perceptual representations from brain activity

Object representations in human high-level visual cortex are at the interface between perception and cognition. What is the nature of these representations, and how are they computed? Furthermore, can they predict human perception?

I will address these questions using representational similarity analysis, an experimental and data-analytical framework for relating brain activity data, behaviour, and computational models. I will focus on experimental data acquired in healthy human participants while they were viewing object images from a wide range of natural categories, including faces and places. The data consist of functional magnetic resonance imaging (fMRI) data from visual cortex and object-similarity judgments, which were acquired outside the fMRI scanner.

I will show that object representations in high-level visual cortex are at once categorical and continuous, and can be explained similarly well by category labels, visual features of intermediate complexity, and deep convolutional neural networks. Among the visual features, it is those correlated with category membership that explain the high-level visual object representation. I will further show that the high-level object representation predicts human object-similarity judgments reasonably well, but fails to capture evolutionary more recent category divisions present in the judgments. These more recent category divisions are human-related and reflect the distinction between humans and nonhuman animals, and between manmade and natural objects.

Together, these findings suggest that high-level visual cortex has developed feature detectors that distinguish between categories of long-standing evolutionary relevance, and that other brain systems might adaptively read out or introduce category divisions that serve current behavioural goals.