Predicting Multiple behaviors from the activity of Deep Neural Networks: Is one visual hierarchy enough?

Aryan Zoroufi1, Aida Mirebrahimi Tafreshi2, Leslie Ungerleider3, Chris Baker3, Maryam Vaziri-Pashkam3

Department of Electrical and Computer Engineering, K.N. TOOSI University of Technology1, Department of Computer Science, Western University2, Laboratory of Brain and Cognition, National Institute of Mental Health3, *Equal Contribution

Motivation

• When we see objects we have an understanding of their category membership. We also know how we can grasp them.
• Both of these tasks require visual processing.
• Can the same visual computation support both tasks?

Method

Grasping Experiment

• 58 3D-printed models of natural objects
• Hand positions at the end of grasping movement were recorded
• Odd-One-Out task on triplets for object similarity
• Repeated 112 times for each condition.

Similarity Judgment Experiment

• 58-alternative forced-choice object categorization
• Images of the same objects in grasp experiment
• Can we predict the two behaviors from CNN activities?

Deep Convolutional Neural Networks

CNN Results

• Despite near human-level performance on categorizing natural objects most DNNs underperform in categorizing images of 3D-printed objects

Conclusion

• The low correlation between the grasp similarities and object similarities suggests that distinct object features are used for each task.
• The deep networks show better correspondence with the similarity judgment behavior in their higher layers, while middle layers have more correlation with grasp behavior.
• The distinction suggest that the underlying mechanisms for these two behaviors may diverge at mid layers of processing.
• These experimental measures can be used in future fMRI experiments to identify visual brain regions that process object features for grasp and similarity judgments.