Recurrence processing during visual object recognition in the human brain

Pablo Oyarzo1, Kohitij Kar2, Radoslaw Martin Cichy1

1Freie Universität Berlin, 2Massachusetts Institute of Technology

Summary

Recurrence computations may be necessary to model visual object recognition at very fast timescales (~200 ms) and could play a critical role in the observed temporal variability of neural responses to visual stimulation. Here, we compared the performance of humans and a feedforward CNN to identify a group of putative “recurrence-dependent” and “control” images (details in Kar et al., 2019; 2021). We presented these images to 25 human participants (N=121 each group) while simultaneously measuring their EEG responses across 64 channels. Multivariate pattern analysis revealed that object information can be approximated later for the “recurrence-dependent” images and that their signal trajectories differ rapidly after stimulus presentation (~150 ms) compared to the control images.

Stimulus set and Experimental design

Comparison of human and DNN performance

10 object exemplars

132 images per object

10 object exemplars

DNN object classification

EEG paradigm

Example of images

Selection of images

EEG task

100 ms

1500 ms

200 ms

200 ms

100 ms

14 images per object

Object discrimination task

Kar et al., 2019

Results: time series of decoding accuracies

Temporal cross-generalization analysis

- Cross generalization showed that decoders trained with recurrence-dependent images tend to exhibit a higher decoding accuracy than those trained with control images. This effect is reflected by a lower decoding accuracy off-diagonal in the first group compared to the other (upper-left and lower-right panels).

- The maximum amplitude of the object-decoding curves was different for both categories of images. Control images peaked higher than recurrence-dependent images (left).

- Within-group generalization analysis revealed that information available during recurrence-dependent processing is more delayed than during the processing of control images. This is reflected by a lower decoding accuracy off-diagonal in the first group compared to the threshold set at 0.535 of accuracy (center).

- Representation similarity analysis between decoding and behavioral performance in object recognition revealed that the representation of objects correlates more than 60 ms later for recurrence-dependent images compared to control (right).

Discussion

We contribute with evidence that variability in early stages of visual processing is consistent with stages of recurrent processing of information, and that this can be experimen tally assessed without manipulation of explicit features in the stimulus. Our findings show that a feedforward CNN can be used as behavioral model to track differential dynamics in brain activity using whole-brain, non-invasive recordings in humans. These results contribute to understanding the adaptive computations performed by the human brain during the processing of objects in the visual ventral stream. In addition, they present new opportunities for using human-brain imaging to expand the existing knowledge of the network architecture that supports these processes.