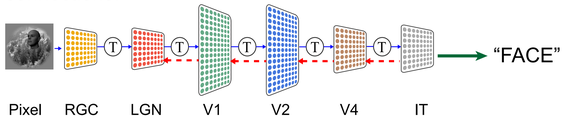
Evidence that feedback is required for object identity inferences computed by the ventral stream

II-37
COSYNE 2017

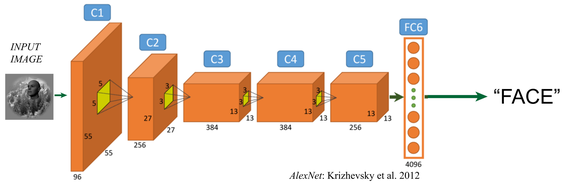
Kohitij Kar, Jonas Kubilius, Elias B Issa, Kailyn M Schmidt, and James J DiCarlo
McGovern Institute for Brain Research, Department of Brain and Cognitive Science, MIT

Introduction

- The primate ventral visual stream for object recognition contains prominent corticocortical feed-forward and feedback connections.



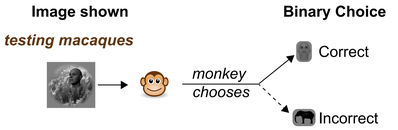
- However, most accurate models of online, rapid (<200 ms) inference in the ventral stream are largely feed-forward (hierarchical convolutional neural networks, HCNN).



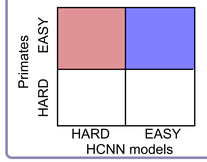
- Are computations along the feedback lines required for rapid object inferences?

Approach

1 Compare image-by-image difficulties between primates and HCNN models

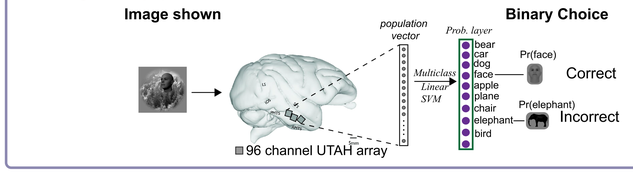


Working hypothesis: *Images that are difficult for HCNNs but easy for primates might be benefitting from feedback computations available in primates and not in the HCNNs.*

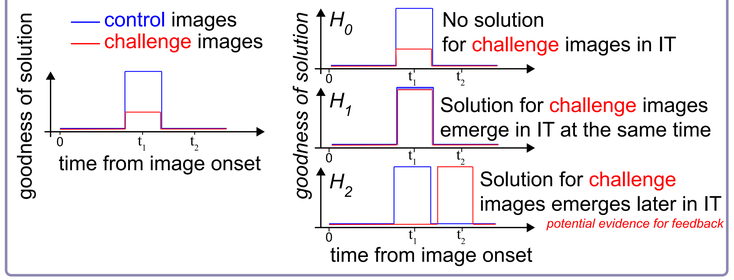


- Control images
  H_0: do not need feedback
- Challenge images
  H_1: requires feedback

2 Measure strength of solutions for image-by image object identity in IT

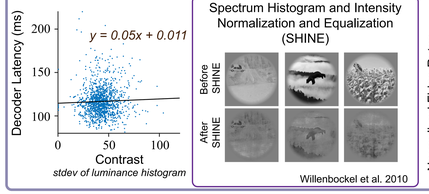


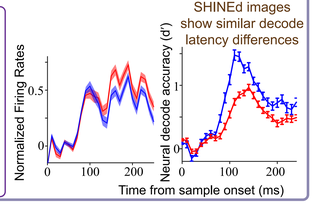
3 Compare IT solution timings between control and challenge images



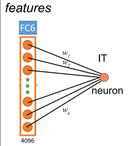
Results

Low level image properties do not acount for the decoder latencies





AlexNet model features explain ~50% of early IT neural variance but does much worse at later time points (> 100 ms)



Randomly split the imageset into two groups (70% train and 30% test)

Use partial least square regression to estimate the mapping weights (w) (use training images and corresponding neural responses averaged across a specific time interval)

Given the weights, predict the neural responses of the test images.

Yamins et al. 2014

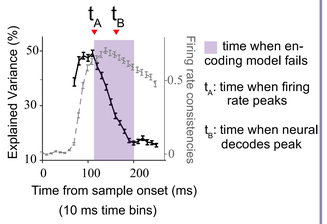
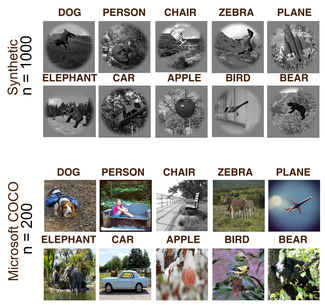
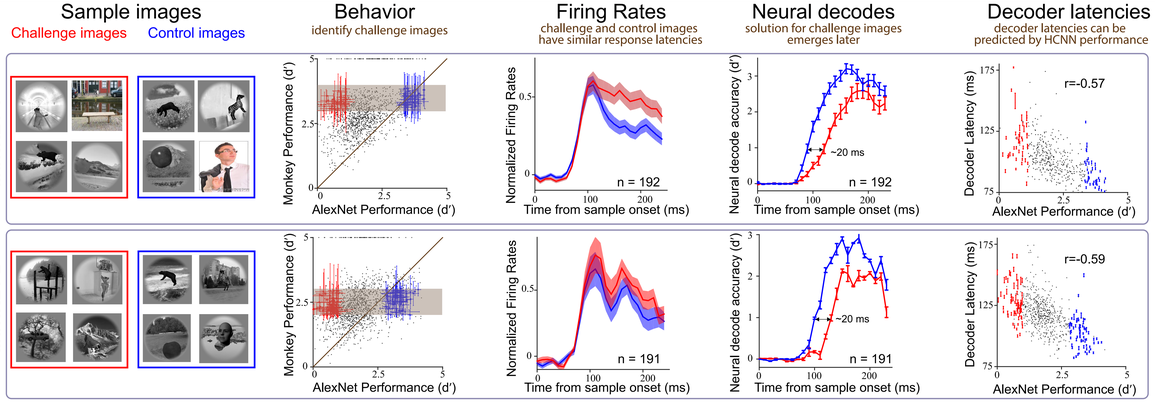


Image-set

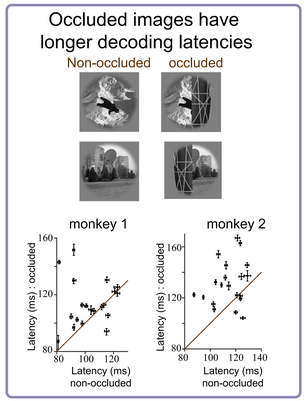
We have tested a mixture of synthetic images and photographs



Results



Occluded images have longer decoding latencies



Conclusions

- Images that are hard for current HCNNs but relatively easier for primates are likely candidates for which the ventral stream uses feedback computations during rapid online inference.
- Neural recordings of the brain's (IT) solution representation of the challenge images reveal that the solution takes longer to emerge, suggesting the possibility of recruiting feedback computations (currently not in use in the HCNN models).
- The failure of the feedforward HCNNs to encode later parts of IT time course is also consistent with the hypothesis that feedback processes are recruited during this time. Future work will address how the inclusion of such feedback connections improve the model's explanatory power.

This research was supported by the Office of Naval Research, Multi-University Research Initiative-60744752-3551877 and NIH-NEI- R01-EY014970-11A1.