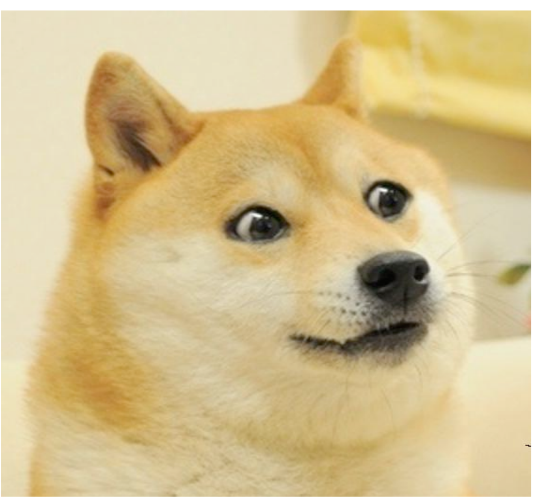
A recurrent neural network for visual object recognition:
An interactive account for rapid superordinate categorization & the basic-level advantage

Qihong Lu & Timothy T. Rogers
Department of Psychology
University of Wisconsin-Madison

Background

Empirical findings: people can categorize objects at the superordinate level by 120ms post visual stimuli onset [1], which has been taken as an evidence for a feed-forward view of visual recognition [2].

We assess whether an interactive recurrent neural network model can explain ultra-rapid superordinate classification in behavior, EEG, and ECoG, while also explaining the basic-level categorization advantage.



Terminology:

- Superordinate: "Animal"
- Basic: "Dog"
- Subordinate: "Shiba Inu"

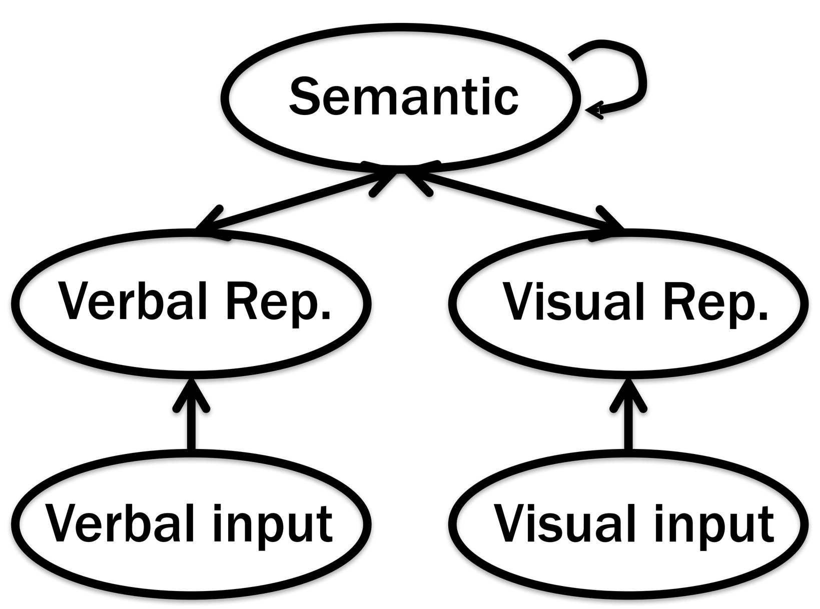
The PDP Modeling Framework

Information processing occurs through propagation of activations among neuron-like processing units.

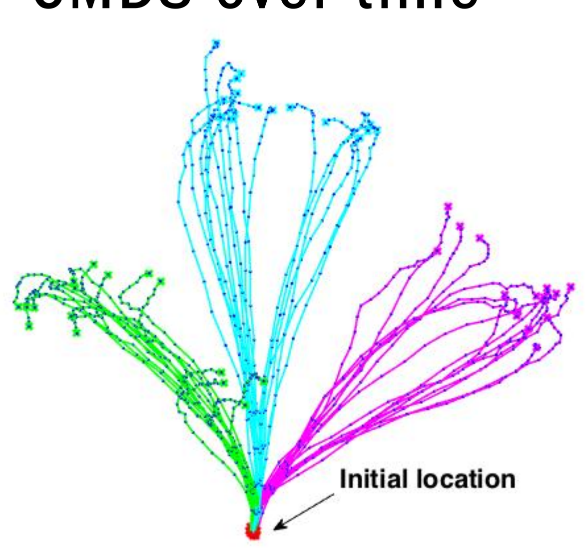
Similar concepts are represented by similar distributed patterns of activation over units [3], which is illustrated by the temporal MDS plot below.

Details: We used a recurrent neural network model to simulate temporal dynamics in a visual recognition task. The semantic layer learns cross-modality representations from the environment. Superordinate and basic-level structure is captured by the similarity structure of individual patterns. The model is trained to produce basic-level names more often than superordinate names.

Model architecture

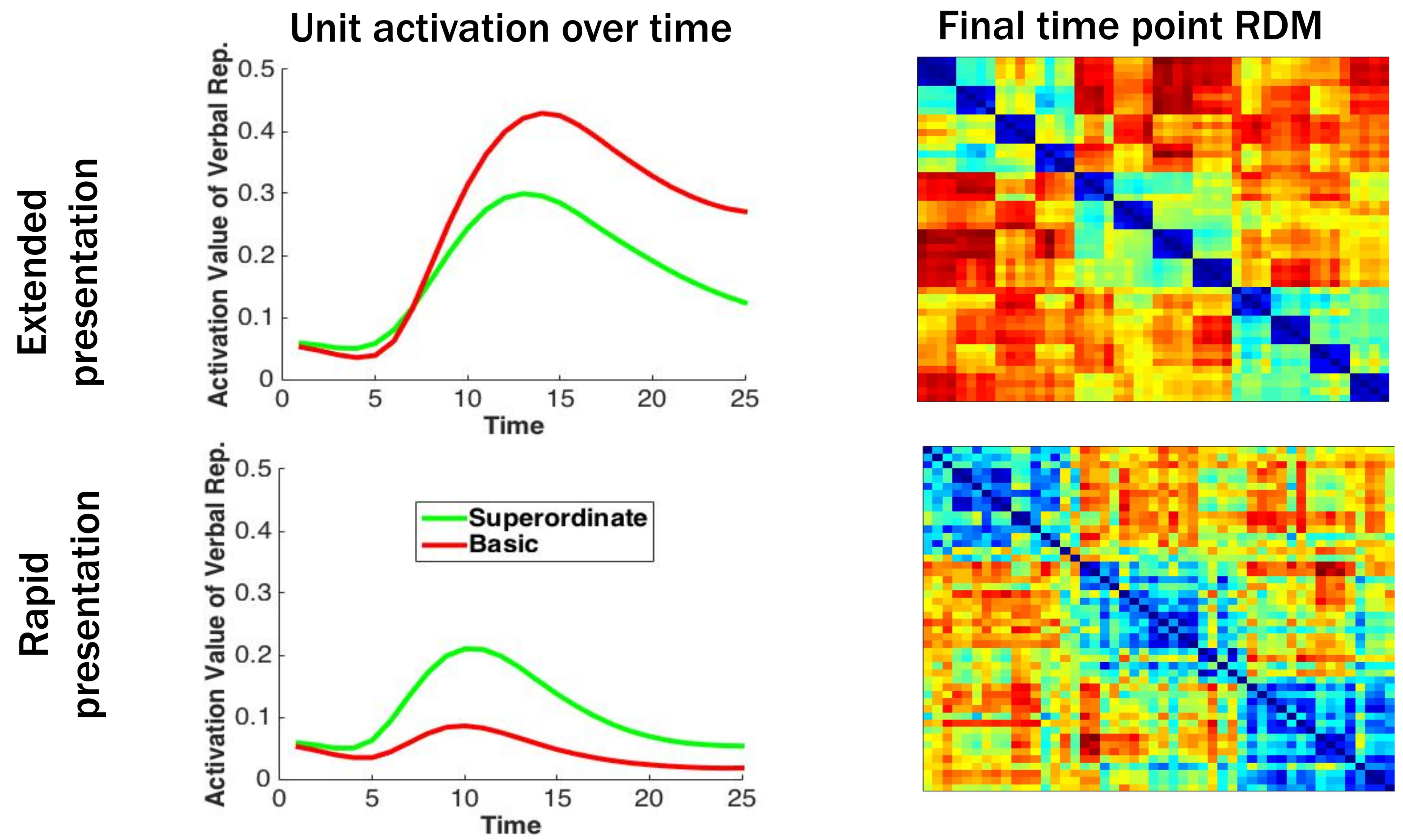


CMDS over time



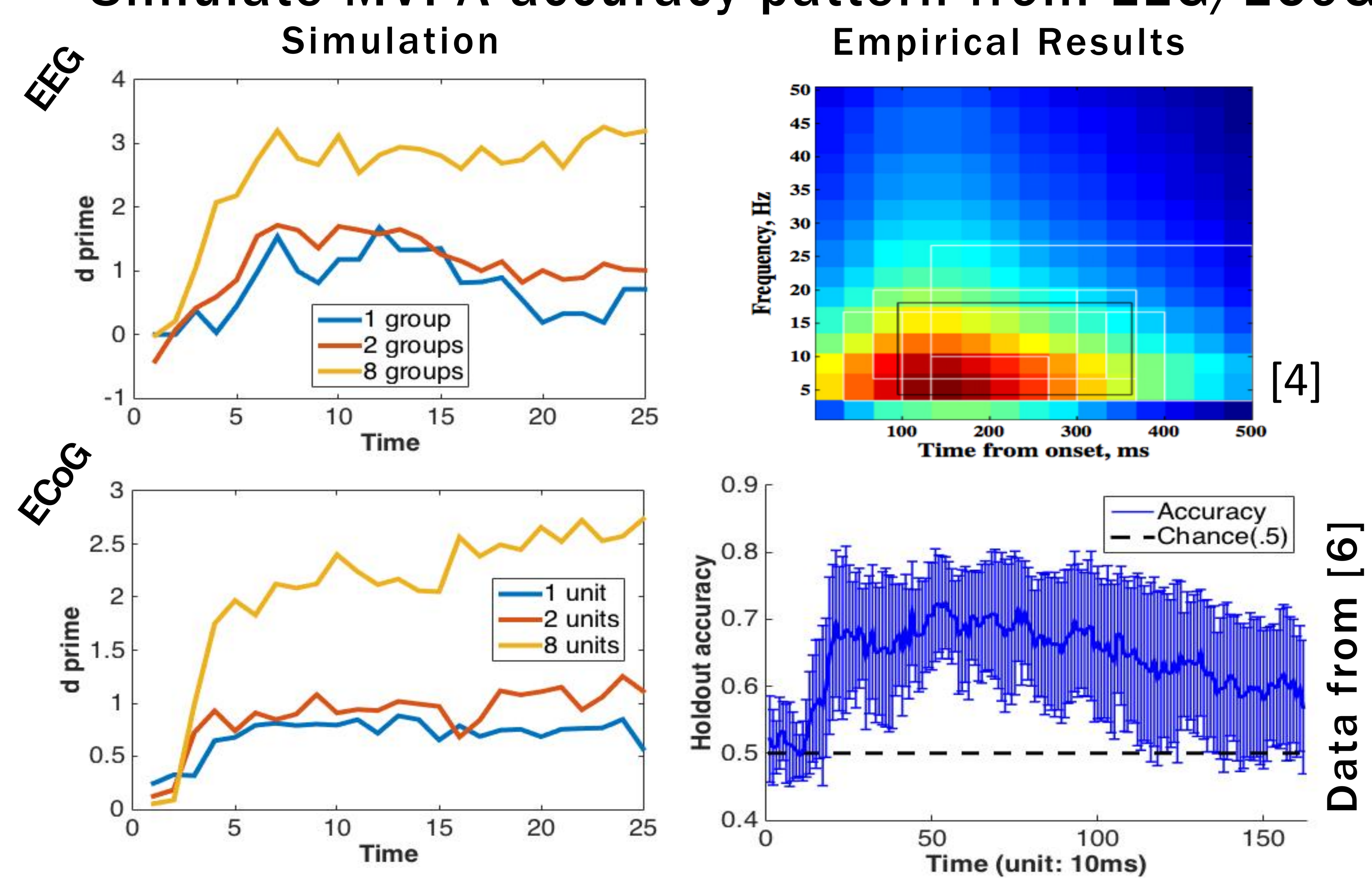
Different colors correspond to different sup. classes.

Simulate behavioral results for visual recognition



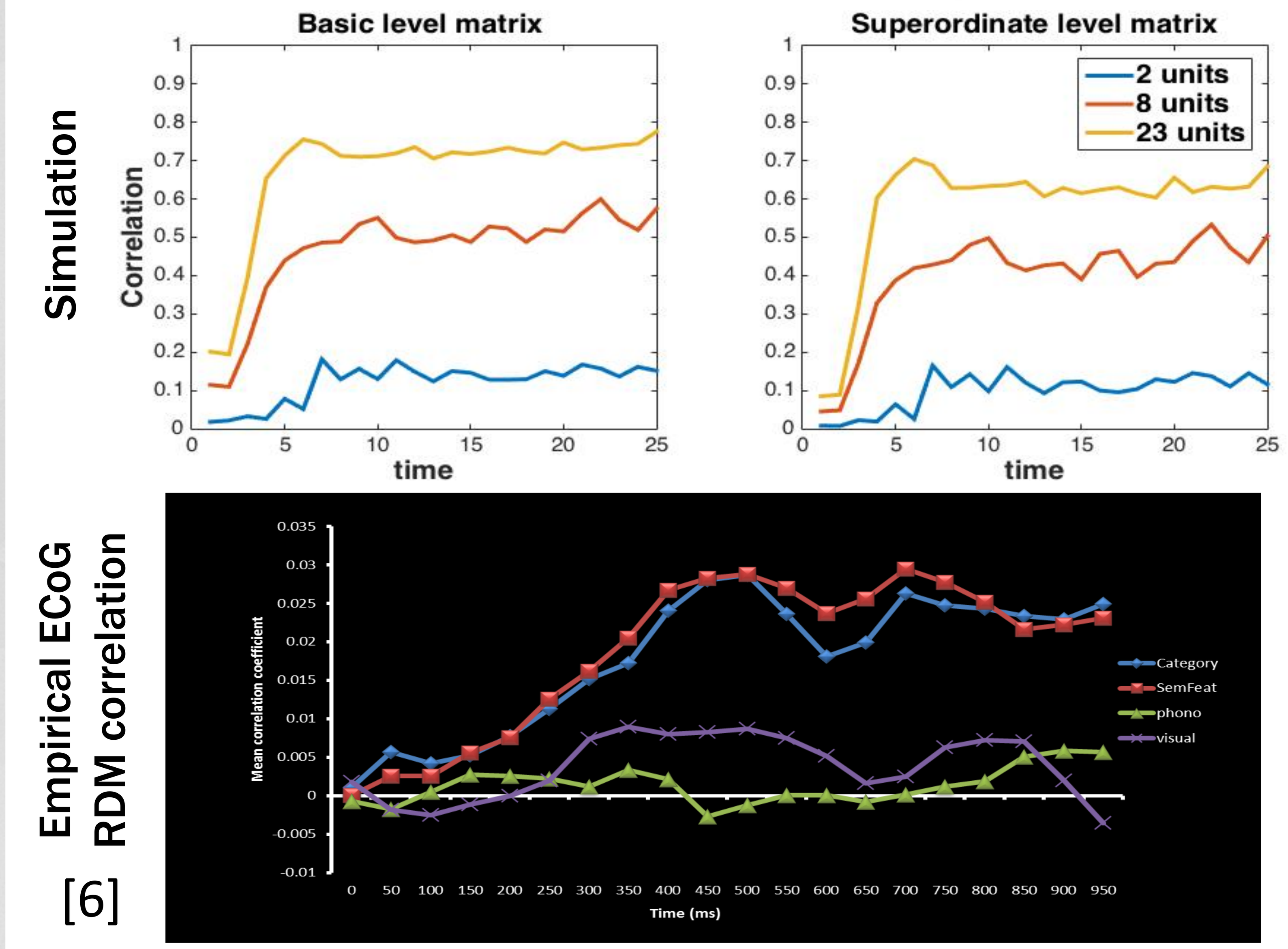
The model shows both a basic level advantage with slow presentation and ultra-rapid categorization pattern with brief presentation.

Simulate MVPA accuracy pattern from EEG/ECoG



- When signal is spatially smoothed (EEG) decoding peaks and drops
- When signal is sparsely sampled (ECoG) it can be decoded throughout

Simulate RDM temporal dynamics



Correlation of model RDMs [5] with idealized superordinate and basic-level RDMs is equivalent and stable over time[6].

Summary

Our recurrent neural network is consistent with:

i) Behavioral results:
- Basic level advantage in untimed setting
- Superordinate but not basic ultra-rapid classification

ii) Neuroimaging results
- Decoding over narrow window with EEG
- Decoding over broad window with ECoG
- RDM temporal correlation over broad window

Conclusion: The neural-cognitive mechanisms underlying ultra-rapid rapid & unconstrained visual object recognition can be interactive.

References

[1] Wu, C.-T., et al. (2015). *Journal of Cognitive Neuroscience*, *27*(1), 141–149.
[2] Serre, T., Oliva, A., & Poggio, T. (2007). *PNAS*, *104*(15), 6424–6429.
[3] Rogers, T. T., & Patterson, K. (2007). *J Exp Psychol Gen*, *136*(3), 451–469.
[4] Murphy, B., et al. (2011). *Brain and Language*, *117*(1), 12–22.
[5] Kriegeskorte, N. (2008). *Frontiers in Systems Neuroscience*. 2, 4.
[6] Chen, Y., et al. (2016). Cortex, 79, 1–13.

Simulation source code:
https://github.com/QihongL/categorization_PDP

Acknowledgement: This research was funded by the Hilldale Undergraduate/Faculty Research Fellowship. We thank Dr. Matthew Lambon Ralph et al. for sharing the ECoG data.

qihong.lu@wisc.edu