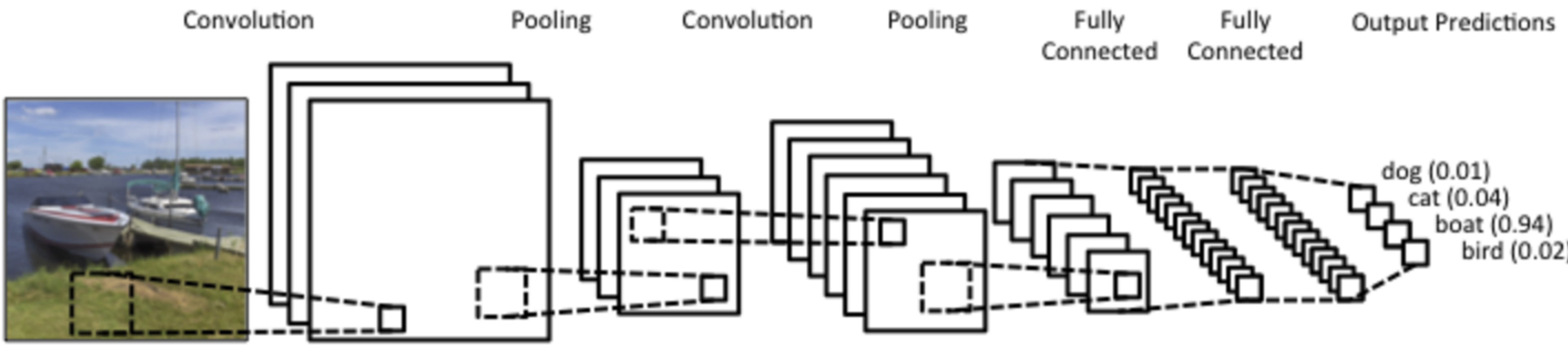
Attention and Modeling of Cognitive Processes

Grace Lindsay, PhD
Columbia University

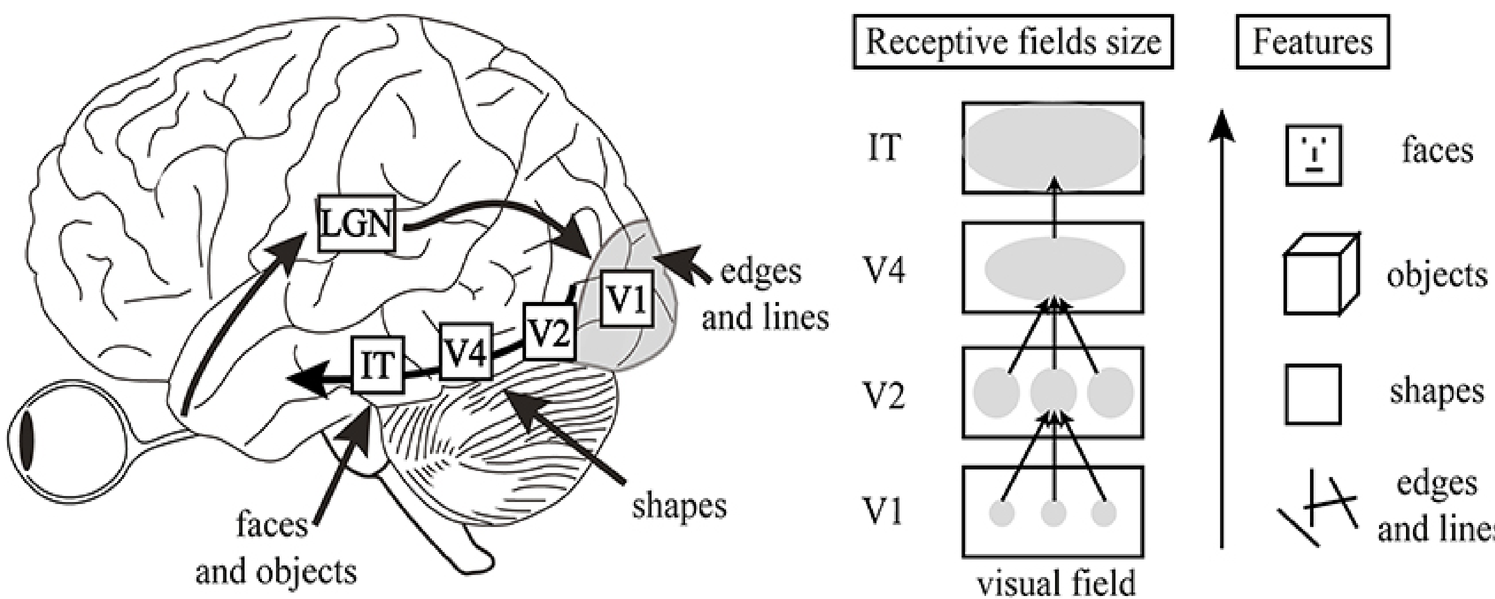
Deep Learning for Human Brain Mapping
Organization for Human Brain Mapping 2019

Convolutional neural networks were inspired by the brain's visual system.



wildml.com

Key components of the architecture: spatial layout and hierarchical feature extraction



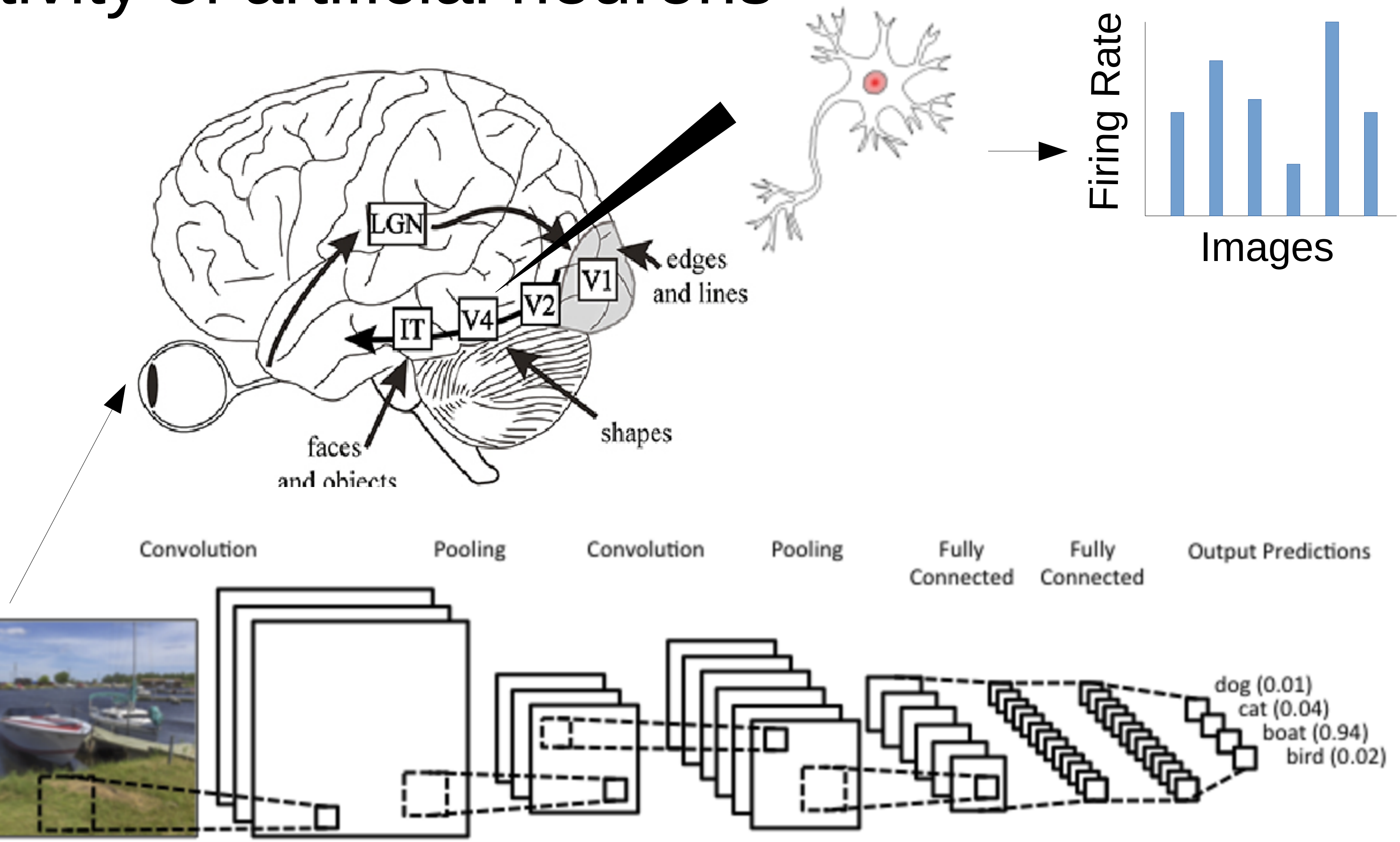
Herzog and Clarke, 2014

Treating CNNs as a model of the brain

- Comparing neural activity
- Comparing behavior

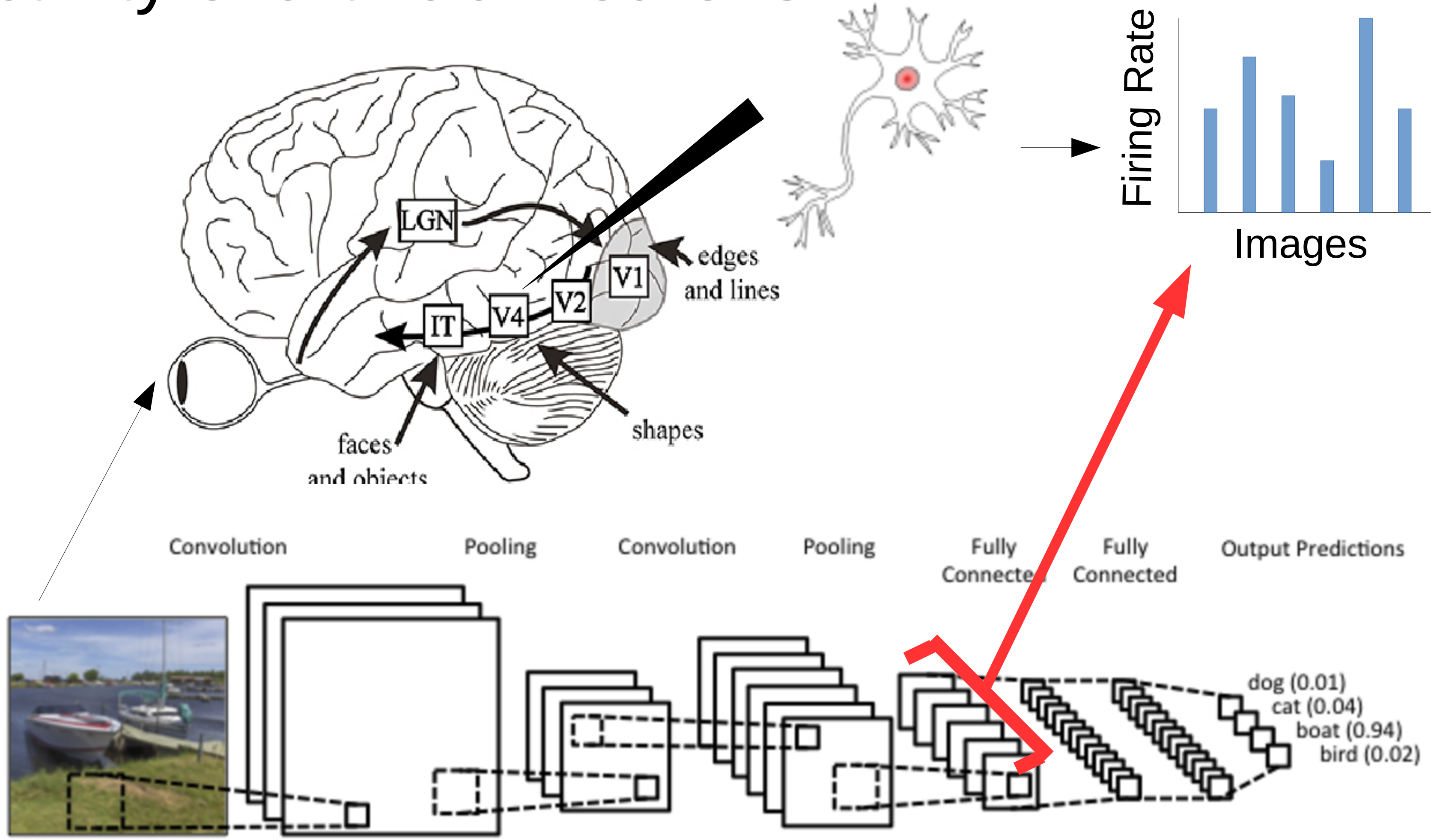
Comparing neural activity

1. Predict a real neuron’s firing rate from the activity of artificial neurons



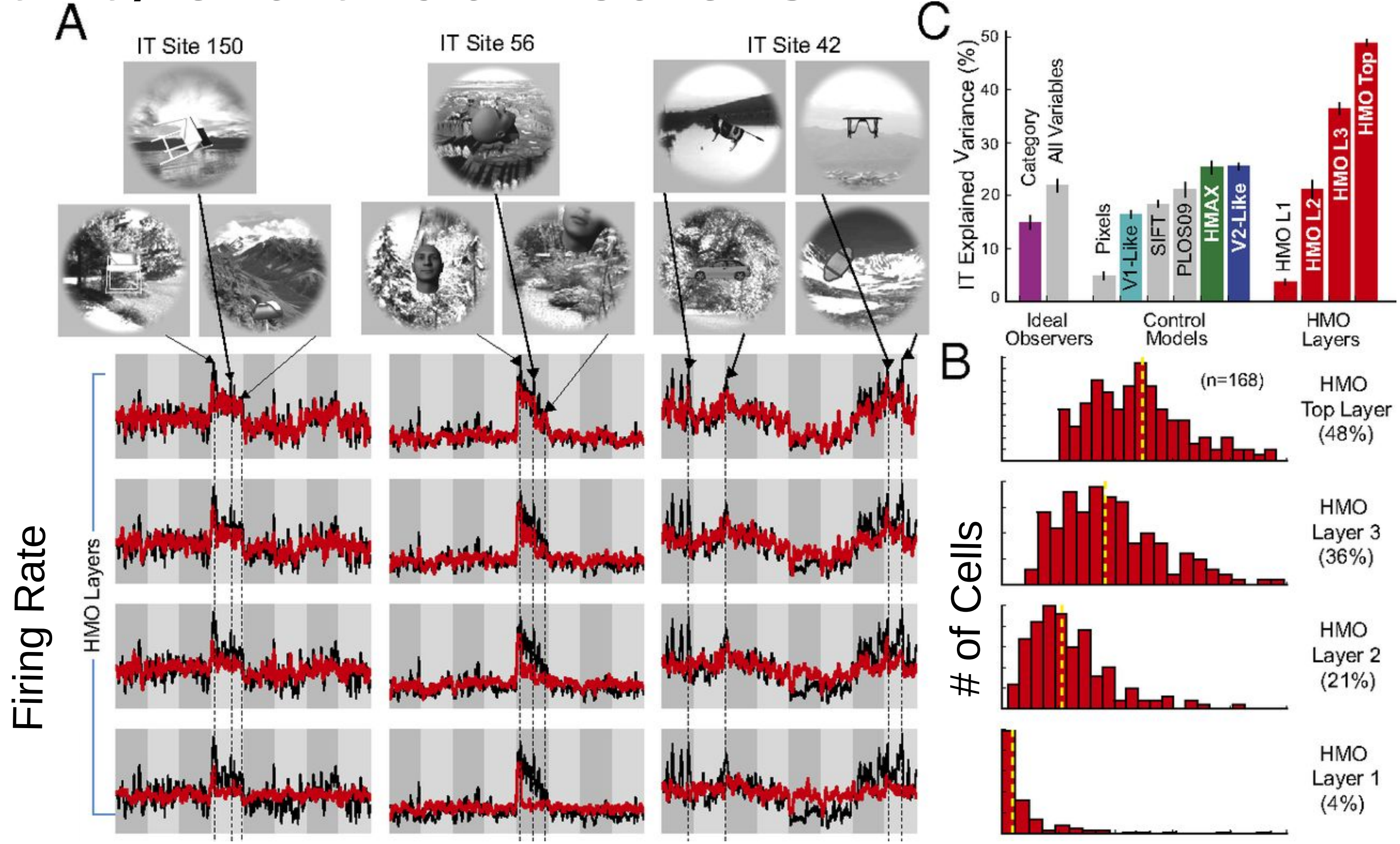
Comparing neural activity

1. Predict a real neuron’s firing rate from the activity of artificial neurons



Comparing neural activity

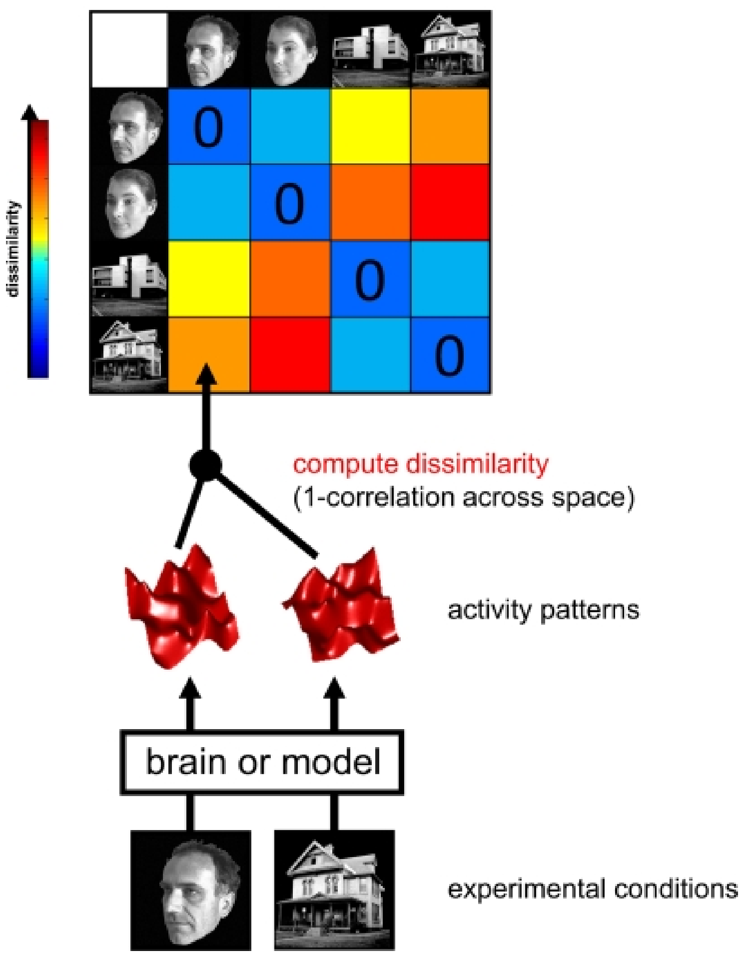
1. Predict a real neuron's firing rate from the activity of artificial neurons



Performance-optimized hierarchical models predict neural responses in higher visual cortex. Yamins et al., 2014

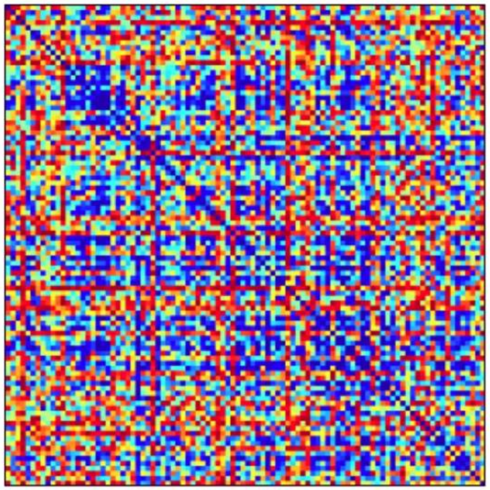
Comparing neural activity

1. Representational Similarity Analysis (RSA)

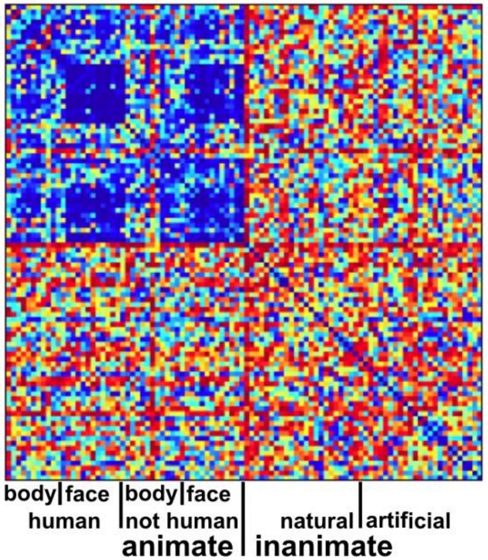


Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience. Kriegeskorte et al., 2008.

Early Visual Cortex

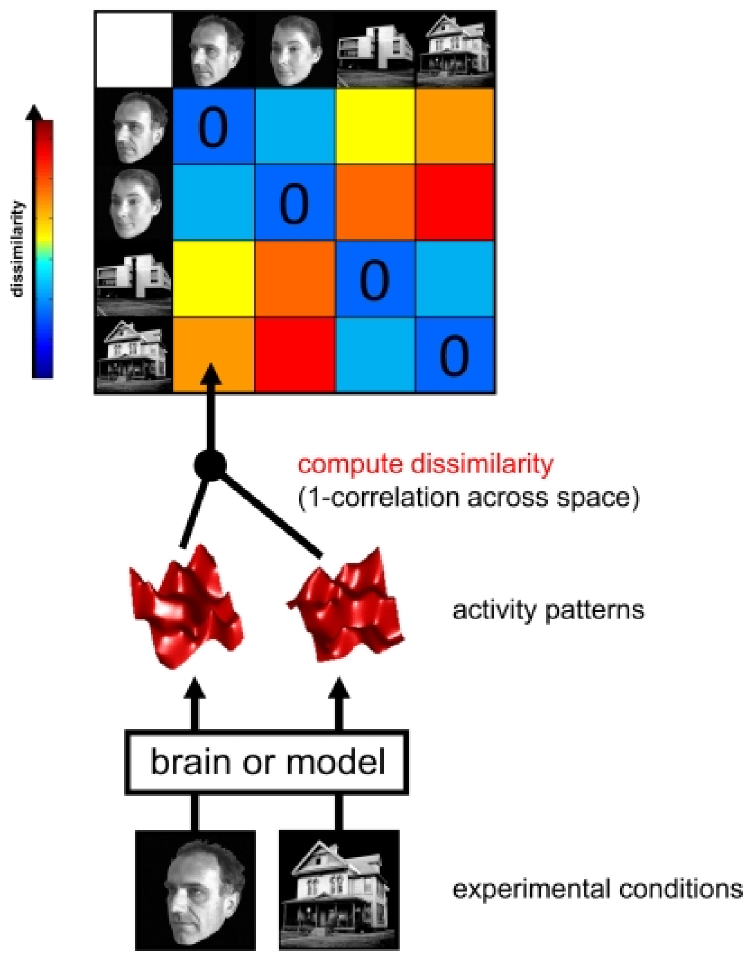


FFA

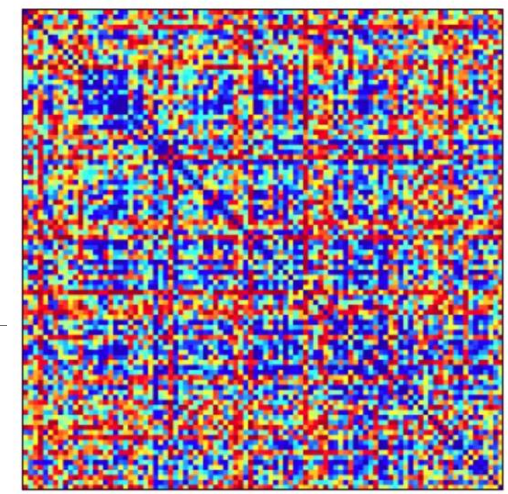


Comparing neural activity

1. Representational Similarity Analysis (RSA)

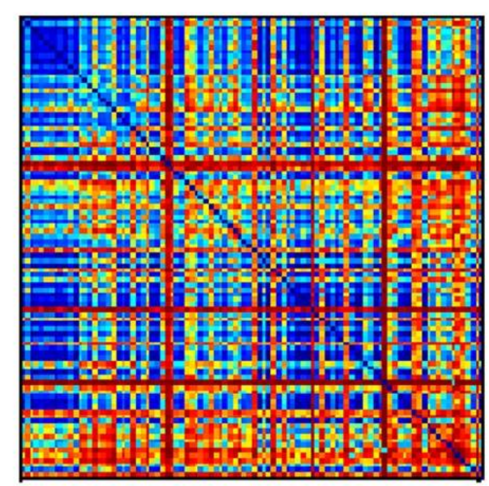


Early Visual Cortex



Measure Correlation

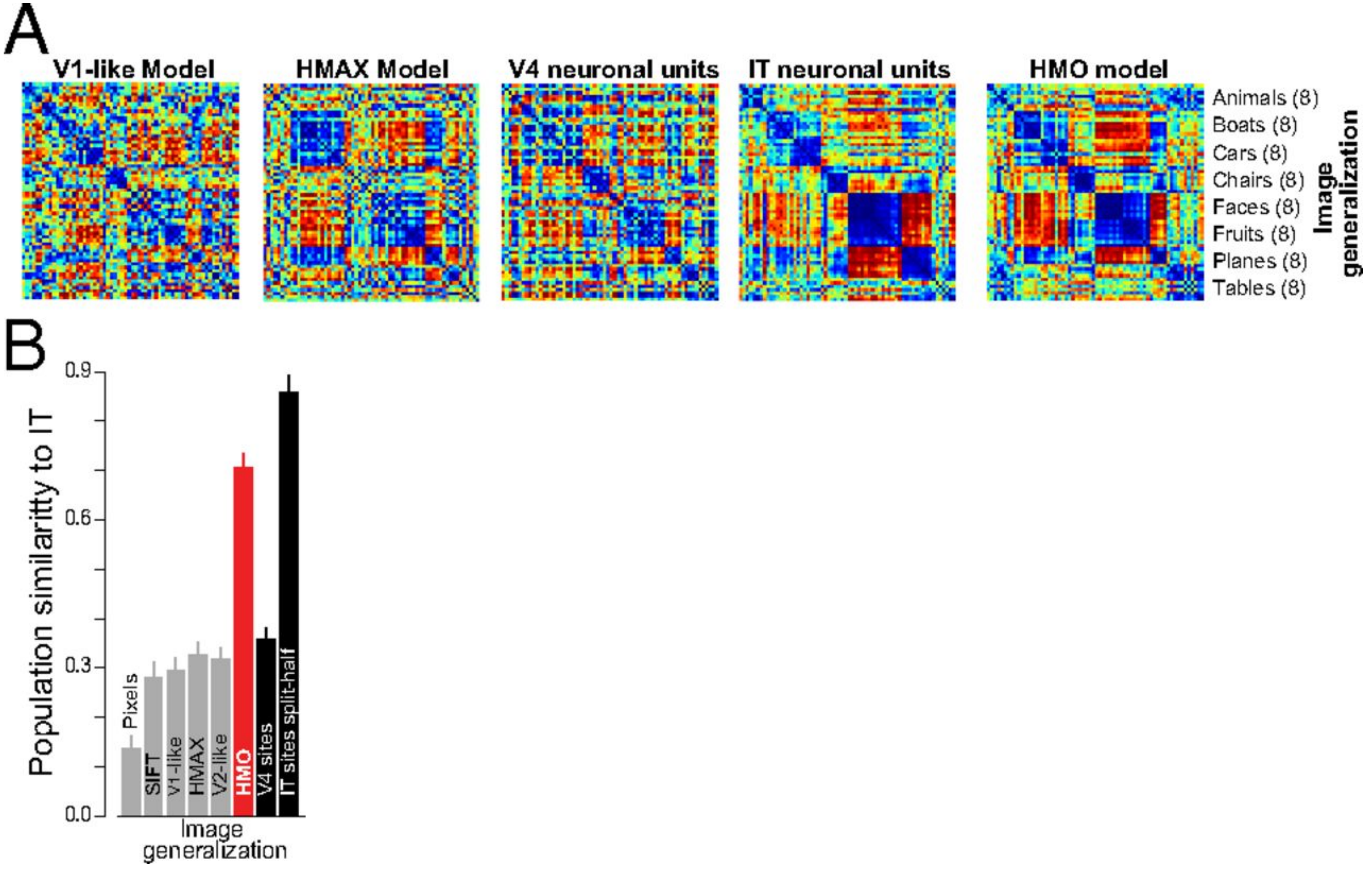
V1 model



Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience. Kriegeskorte et al., 2008.

Comparing neural activity

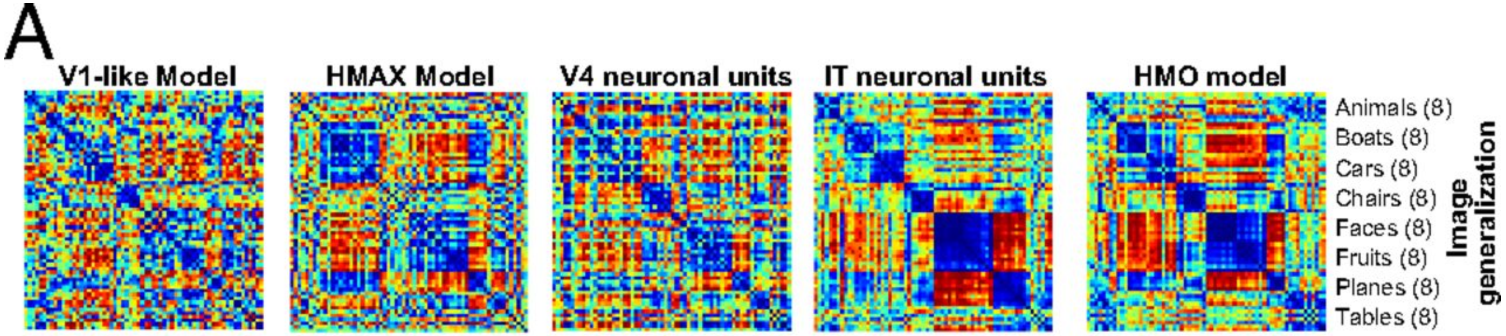
1. Representational Similarity Analysis (RSA)

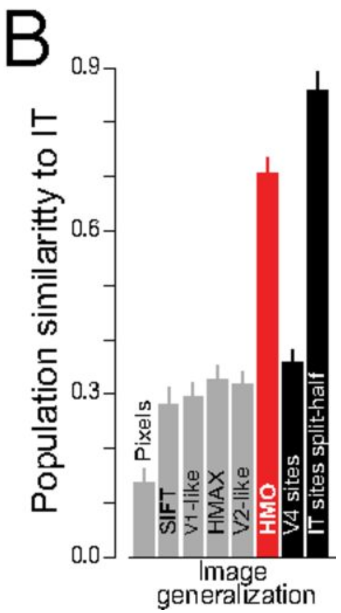


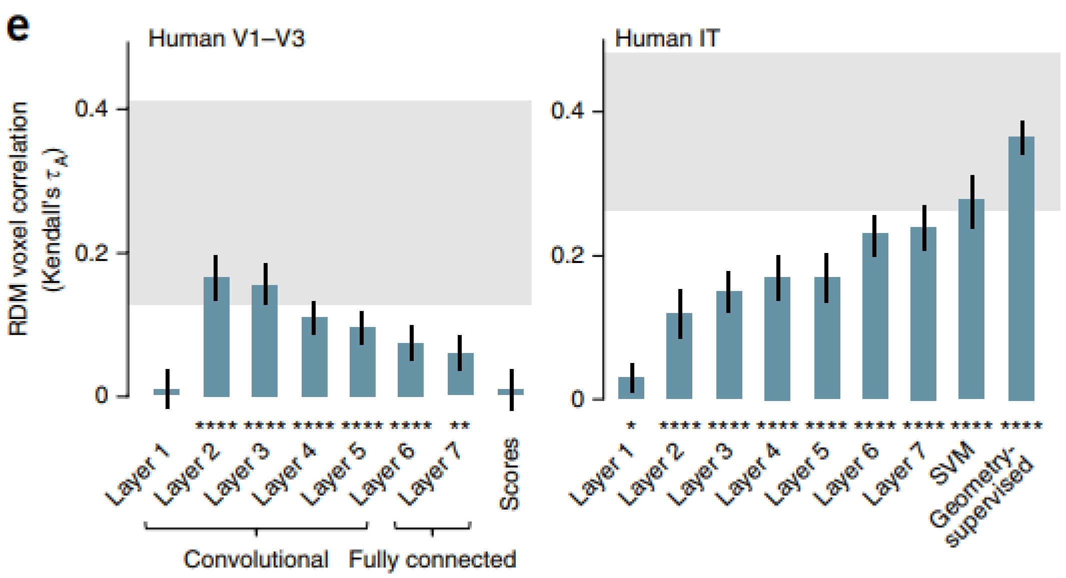
Performance-optimized hierarchical models predict neural responses in higher visual cortex. Yamins et al., 2014

Comparing neural activity

1. Representational Similarity Analysis (RSA)



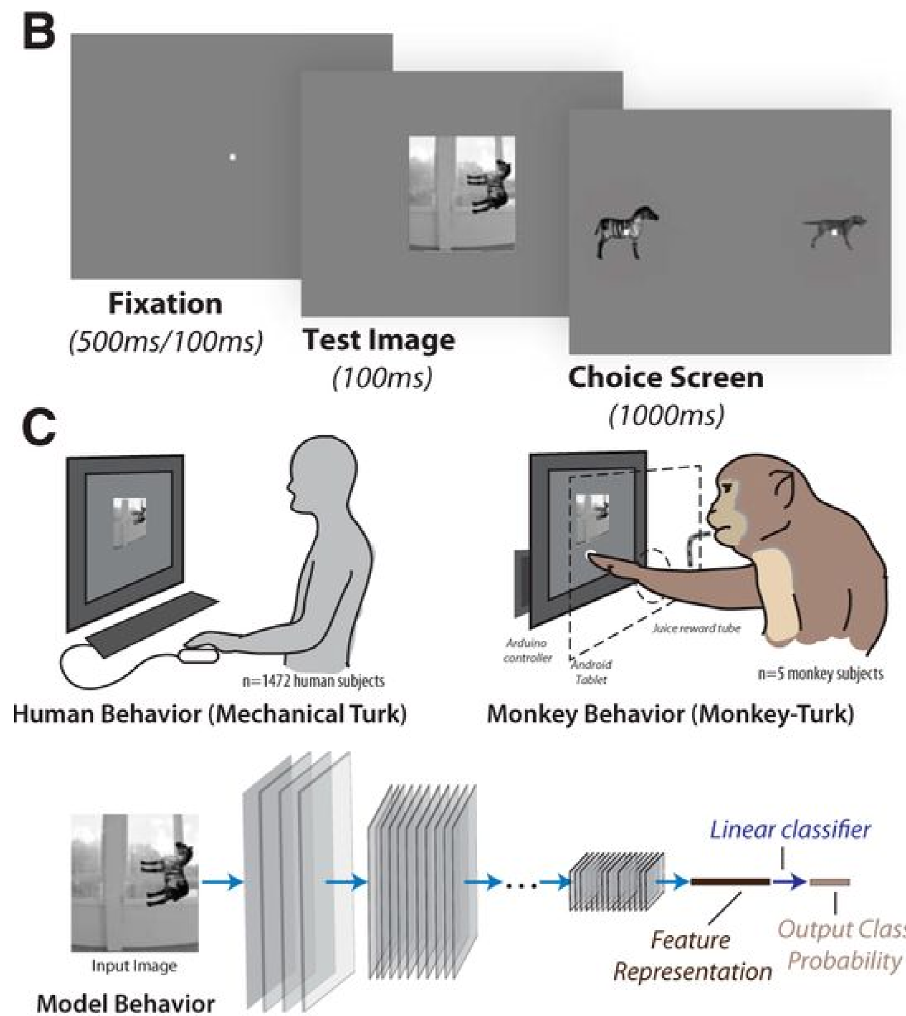




Performance-optimized hierarchical models predict neural responses in higher visual cortex. Yamins et al., 2014

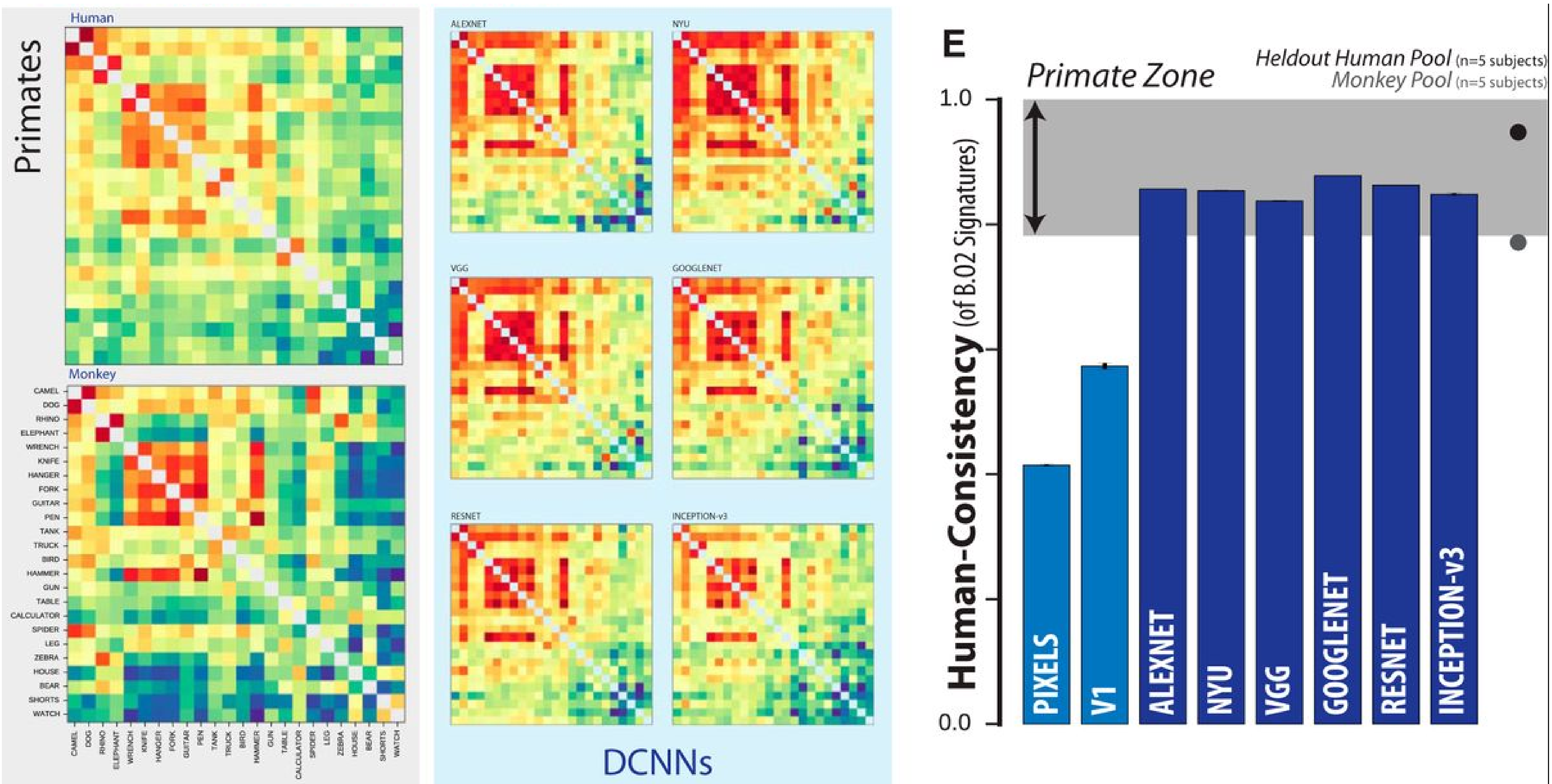
Deep supervised, but not unsupervised, models may explain IT cortical representation. Khaligh-Razavi & Kriegeskorte, 2014

Comparing Behavior



Large-Scale, High-Resolution Comparison of the Core Visual Object Recognition Behavior of Humans, Monkeys, and State-of-the-Art Deep Artificial Neural Networks. Rajalingham et al., 2018

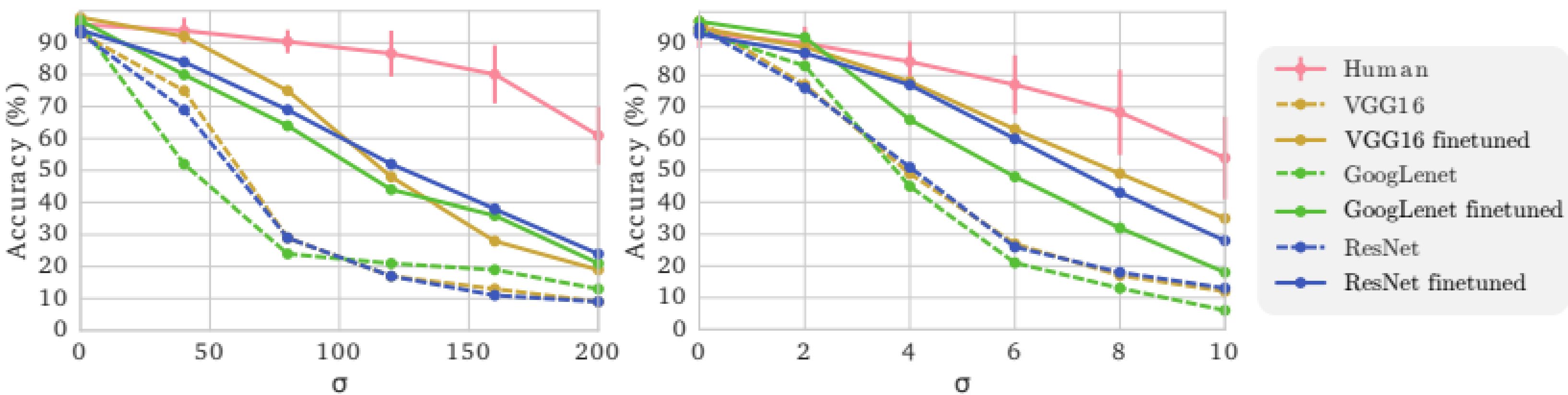
Comparing Behavior



Large-Scale, High-Resolution Comparison of the Core Visual Object Recognition Behavior of Humans, Monkeys, and State-of-the-Art Deep Artificial Neural Networks. Rajalingham et al., 2018

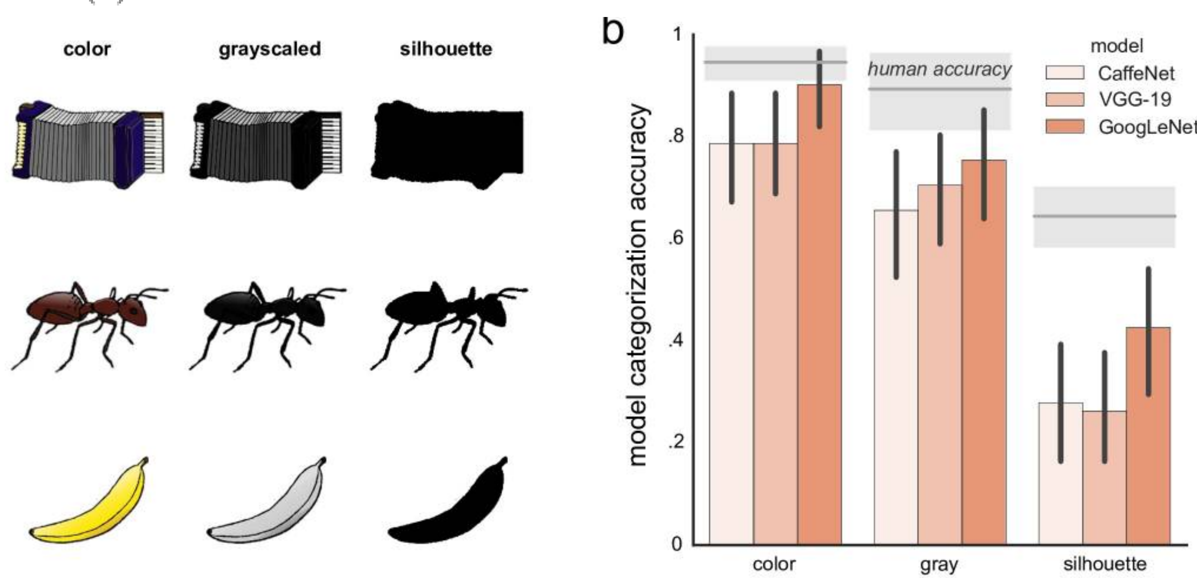
Comparing Behavior

Comparing how they fail:



(a) Noise (b) Blur

A Study and Comparison of Human and Deep Learning Recognition Performance Under Visual Distortions Dodge & Karam, 2017



Deep Neural Networks as a Computational Model for Human Shape Sensitivity. Kubilius et al., 2016

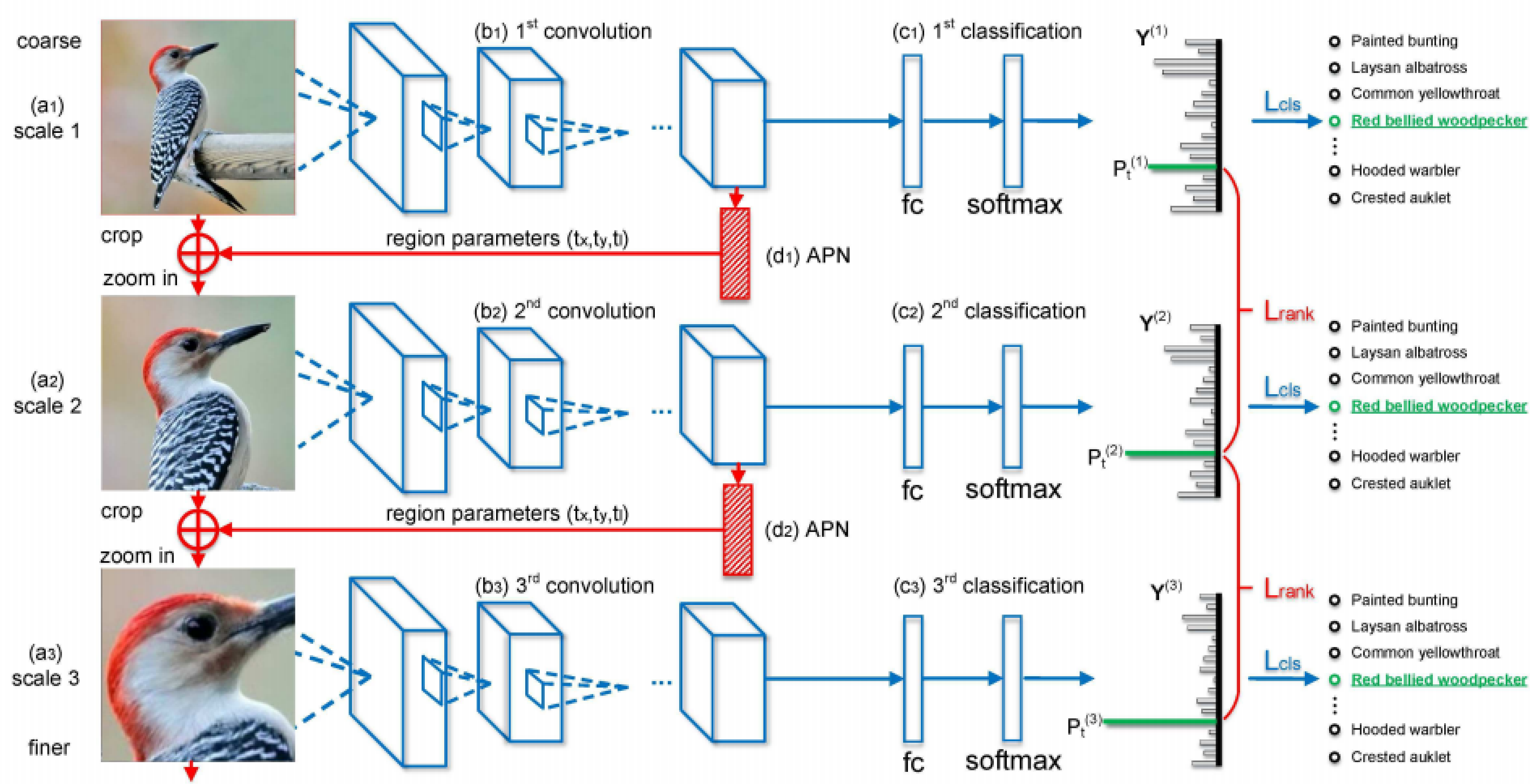
Attention

A word that has many meanings…

- Attention in machine learning
- Attention in neuroscience

Attention in Convolutional Neural Networks

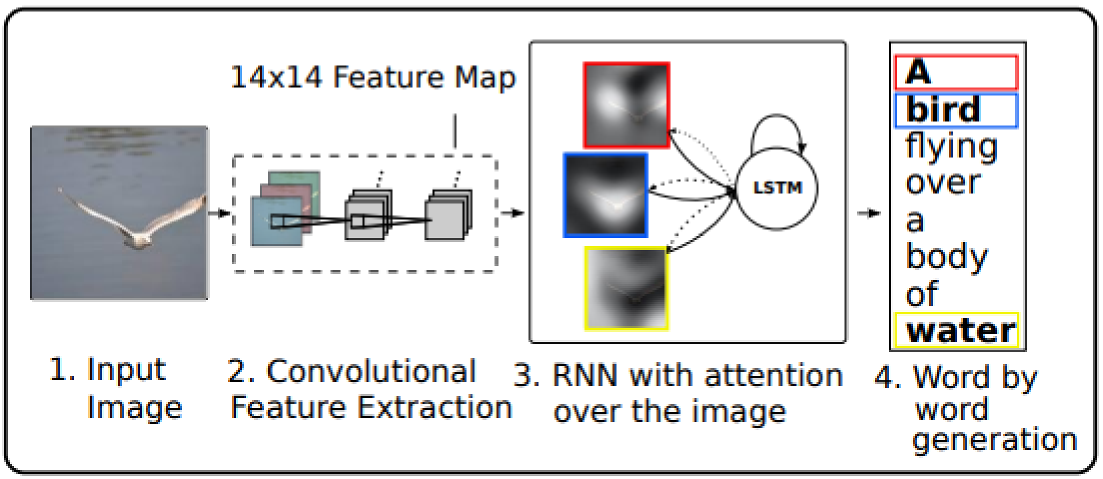
"Hard" spatial attention:



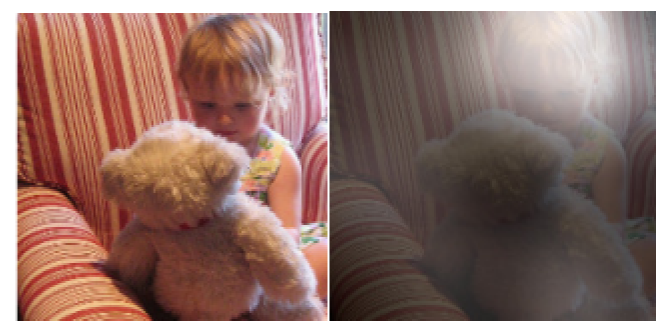
Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-grained Image Recognition. Fu et al., 2017.

Attention in Convolutional Neural Networks

"Soft" spatial attention:



A woman is throwing a <u>frisbee</u> in a park.

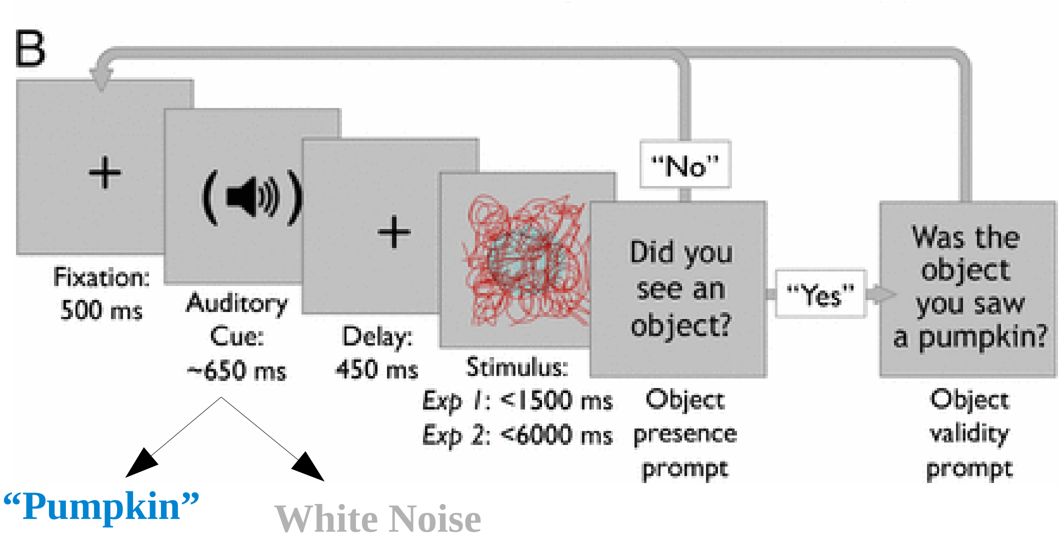


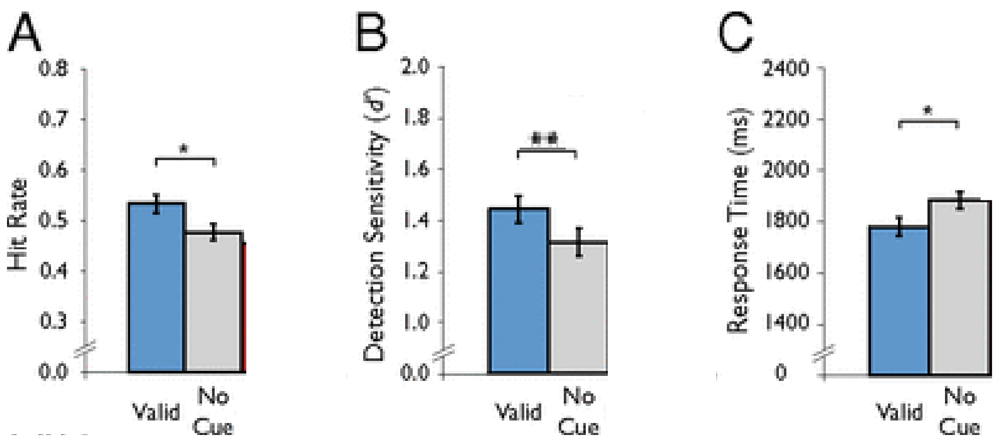
A little <u>girl</u> sitting on a bed with a teddy bear.

Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. Xu et al., 2015.

Attention in the Brain

Feature-based attention:



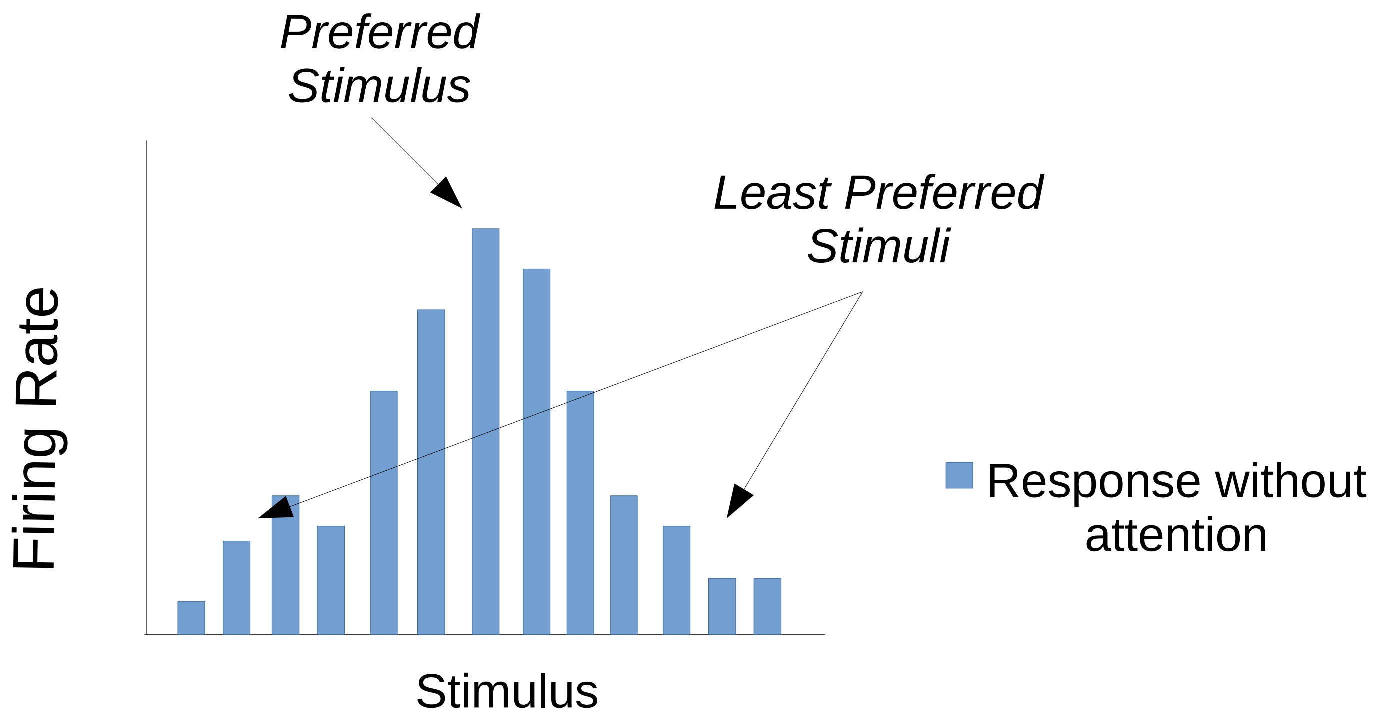


Lupyan and Ward, 2013

Attention in the Brain

- Attention modulates the firing of a neuron in a way that depends on its normal firing response.

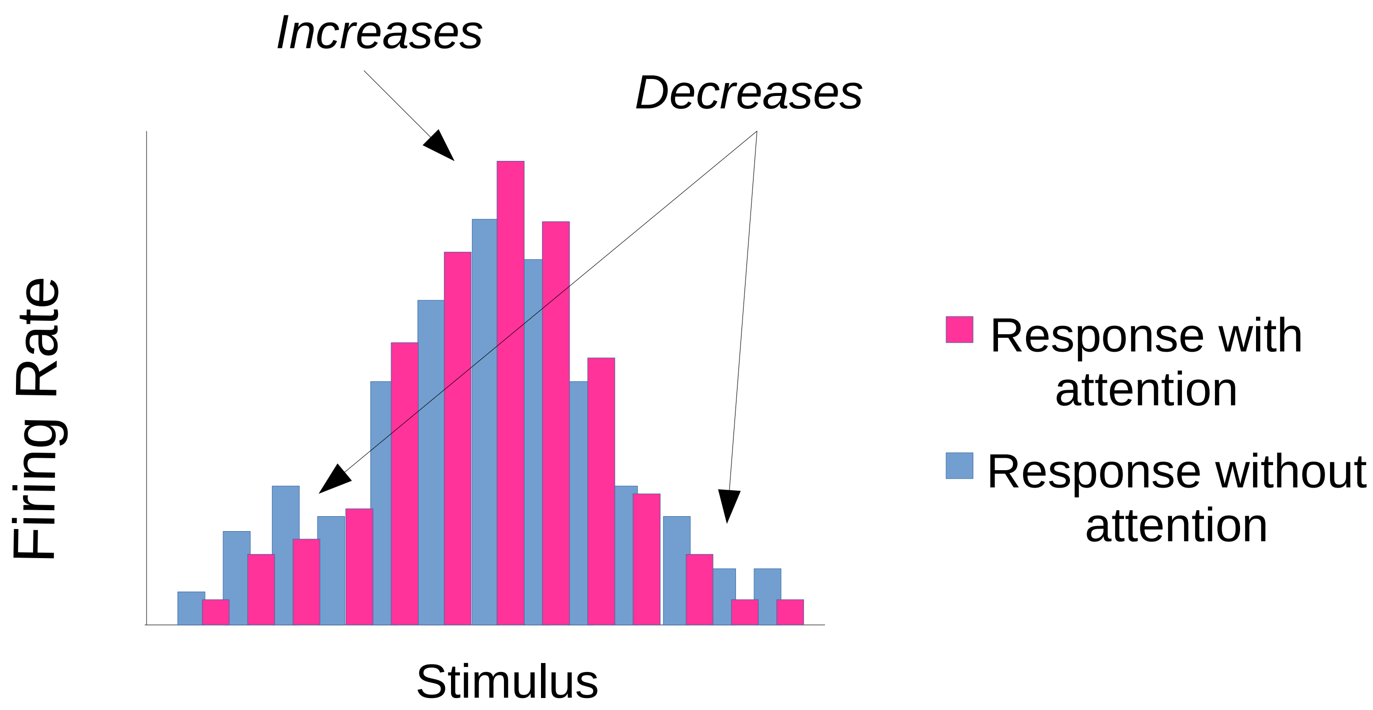
Tuning Curve:



Attention in the Brain

- "Feature Similarity Gain Model": Attention modulates the firing of a neuron in a way that depends on its normal firing response.

Tuning Curve:



Attention to preferred stimuli increases firing rates while attention to non-preferred decreases

How to model biological attention in convolutional neural networks

eLIFE
elifesciences.org

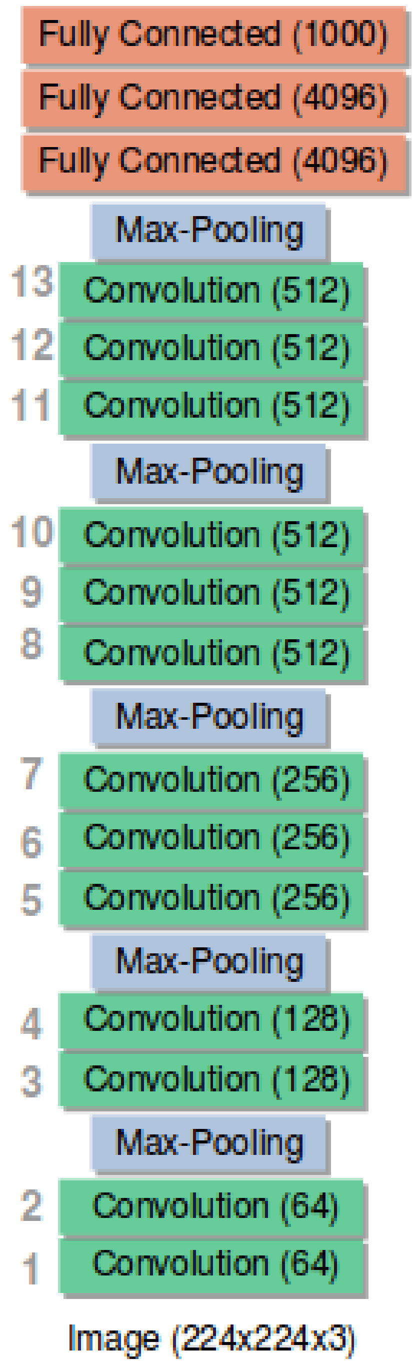
RESEARCH ARTICLE

How biological attention mechanisms improve task performance in a large-scale visual system model

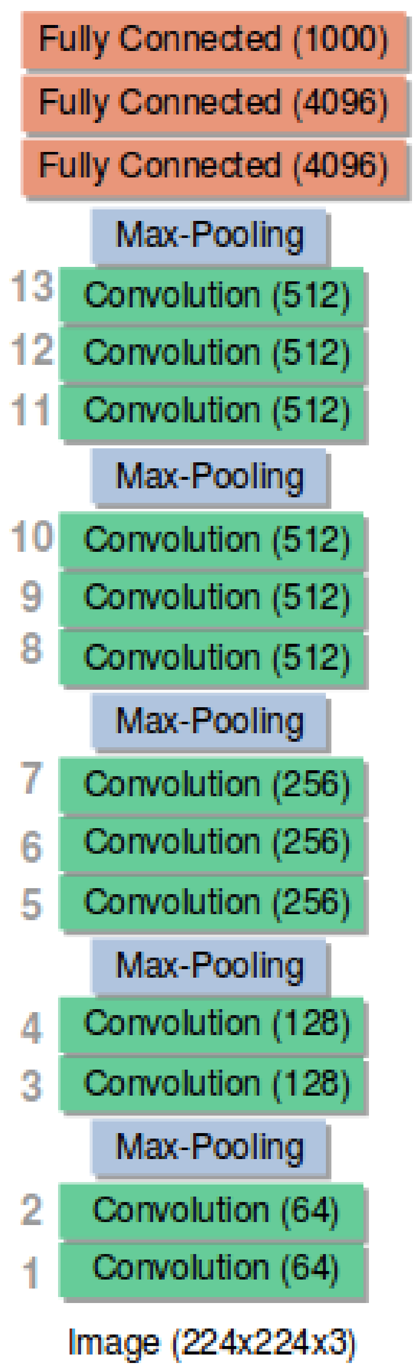
Grace W Lindsay[1,2]*, Kenneth D Miller[1,2,3,4]

[1]Center for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University, New York, United States; [2]Mortimer B. Zuckerman Mind Brain Behaviour Institute, Columbia University, New York, United States; [3]Swartz Program in Theoretical Neuroscience, Kavli Institute for Brain Science, New York, United States; [4]Department of Neuroscience, Columbia University, New York, United States

VGG-16

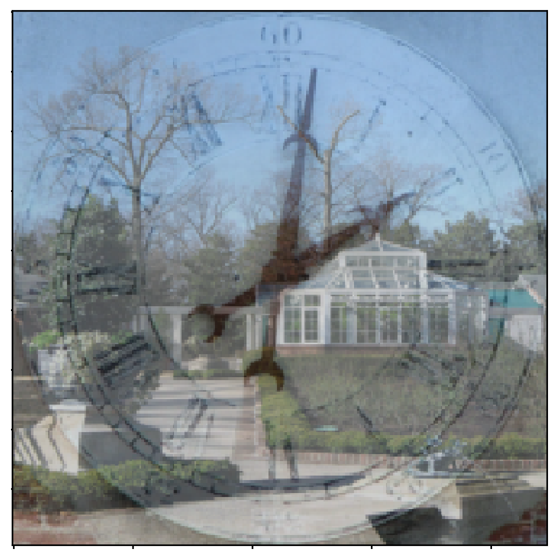

Fully Connected (1000)
Fully Connected (4096)
Fully Connected (4096)
Max-Pooling
13 Convolution (512)
12 Convolution (512)
11 Convolution (512)
Max-Pooling
10 Convolution (512)
9 Convolution (512)
8 Convolution (512)
Max-Pooling
7 Convolution (256)
6 Convolution (256)
5 Convolution (256)
Max-Pooling
4 Convolution (128)
3 Convolution (128)
Max-Pooling
2 Convolution (64)
1 Convolution (64)
Image (224x224x3)

VGG-16

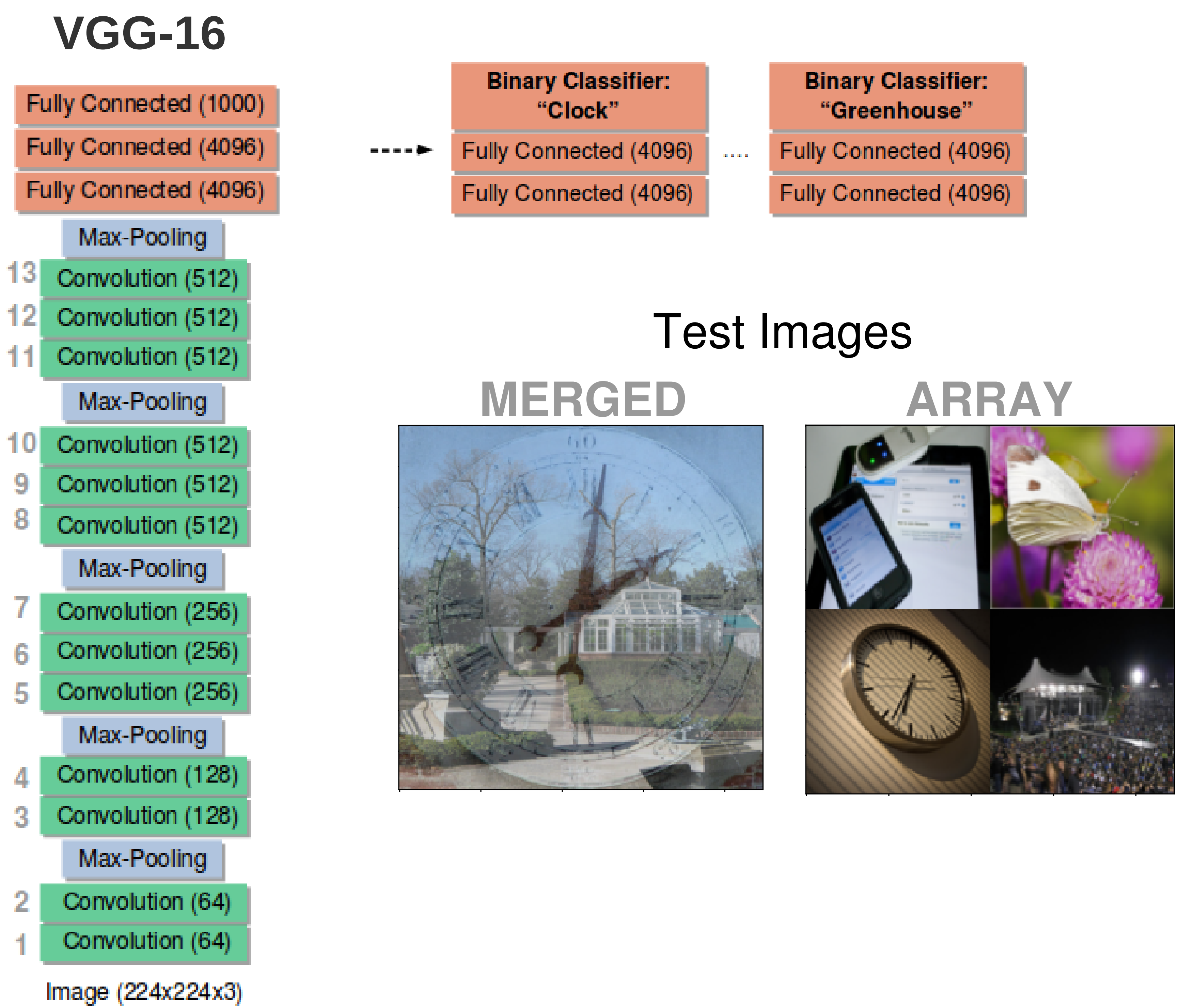


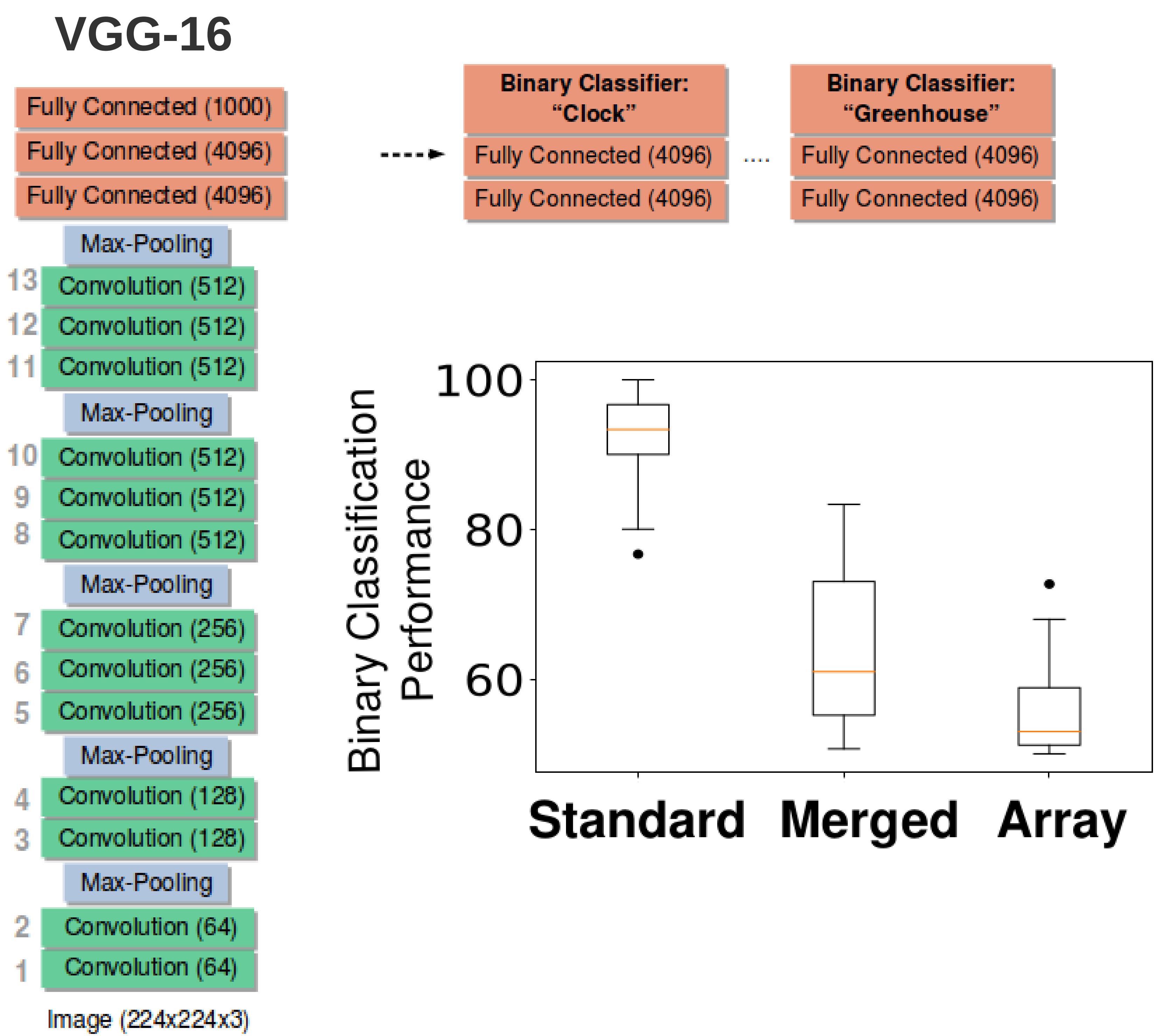
Test Images

MERGED



ARRAY

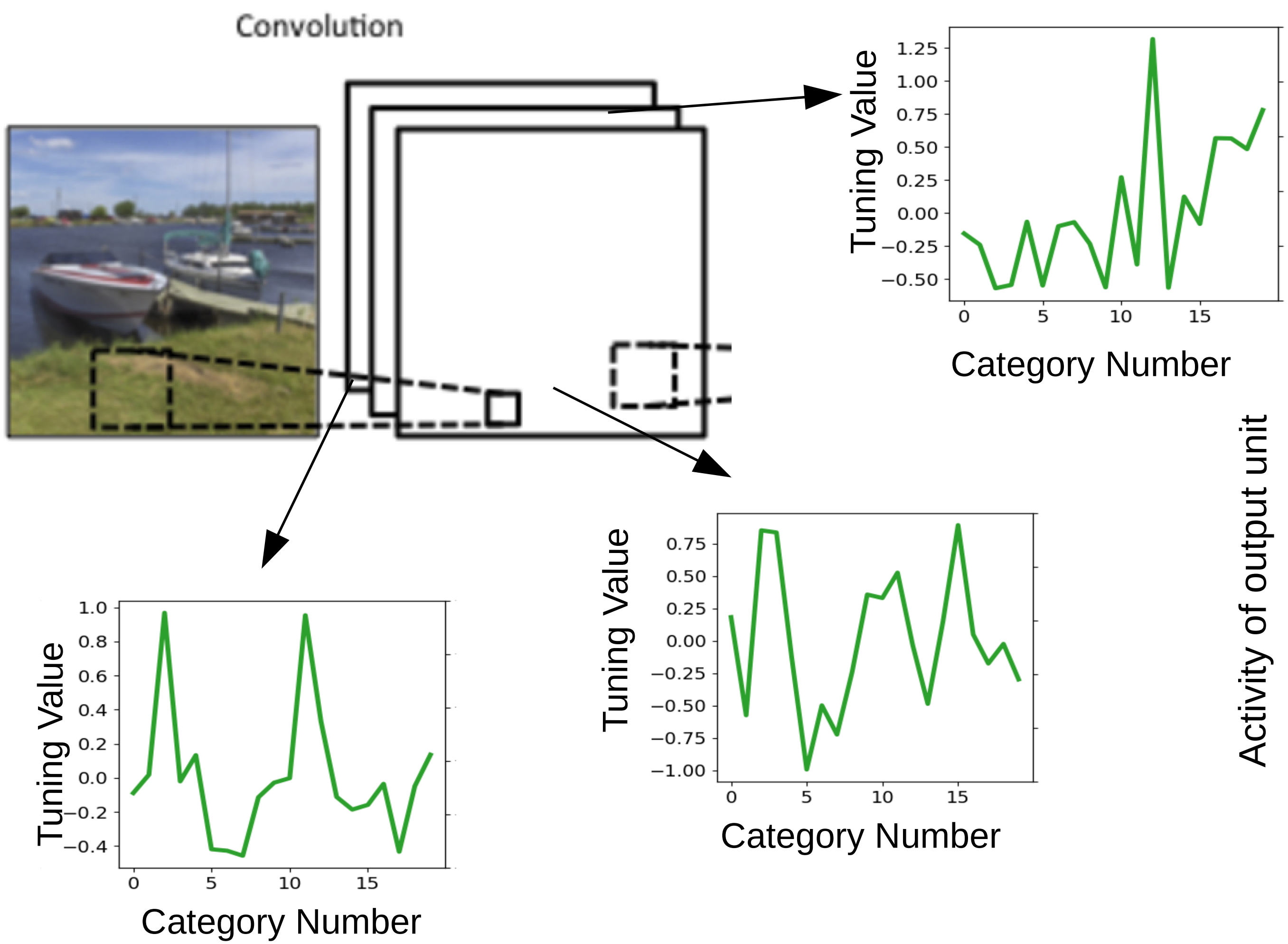

VGG-16
Fully Connected (1000)
Fully Connected (4096)
Fully Connected (4096)
Max-Pooling
13 Convolution (512)
12 Convolution (512)
11 Convolution (512)
Max-Pooling
10 Convolution (512)
9 Convolution (512)
8 Convolution (512)
Max-Pooling
7 Convolution (256)
6 Convolution (256)
5 Convolution (256)
Max-Pooling
4 Convolution (128)
3 Convolution (128)
Max-Pooling
2 Convolution (64)
1 Convolution (64)
Image (224x224x3)
Binary Classifier: "Clock"
Fully Connected (4096)
Fully Connected (4096)
Binary Classifier: "Greenhouse"
Fully Connected (4096)
Fully Connected (4096)
Test Images
MERGED
ARRAY



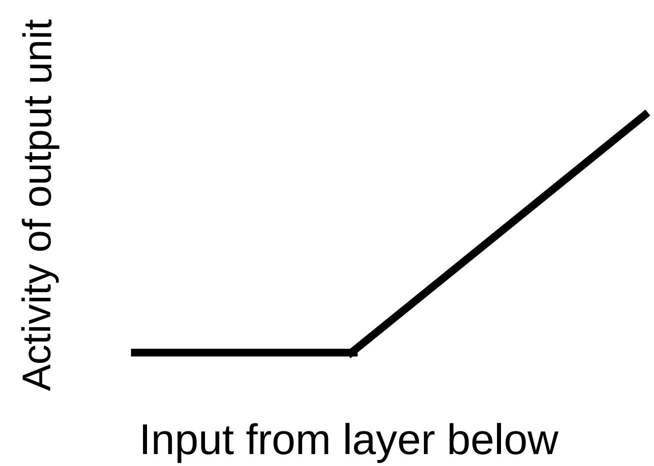
VGG-16
Fully Connected (1000)
Fully Connected (4096)
Fully Connected (4096)
Max-Pooling
13 Convolution (512)
12 Convolution (512)
11 Convolution (512)
Max-Pooling
10 Convolution (512)
9 Convolution (512)
8 Convolution (512)
Max-Pooling
7 Convolution (256)
6 Convolution (256)
5 Convolution (256)
Max-Pooling
4 Convolution (128)
3 Convolution (128)
Max-Pooling
2 Convolution (64)
1 Convolution (64)
Image (224x224x3)
Binary Classifier: "Clock"
Fully Connected (4096)
Fully Connected (4096)
Binary Classifier: "Greenhouse"
Fully Connected (4096)
Fully Connected (4096)
Binary Classification Performance
100
80
60
Standard
Merged
Array

Replicating the neural correlates of attention

1.) Make normalized category tuning curves for each feature map:

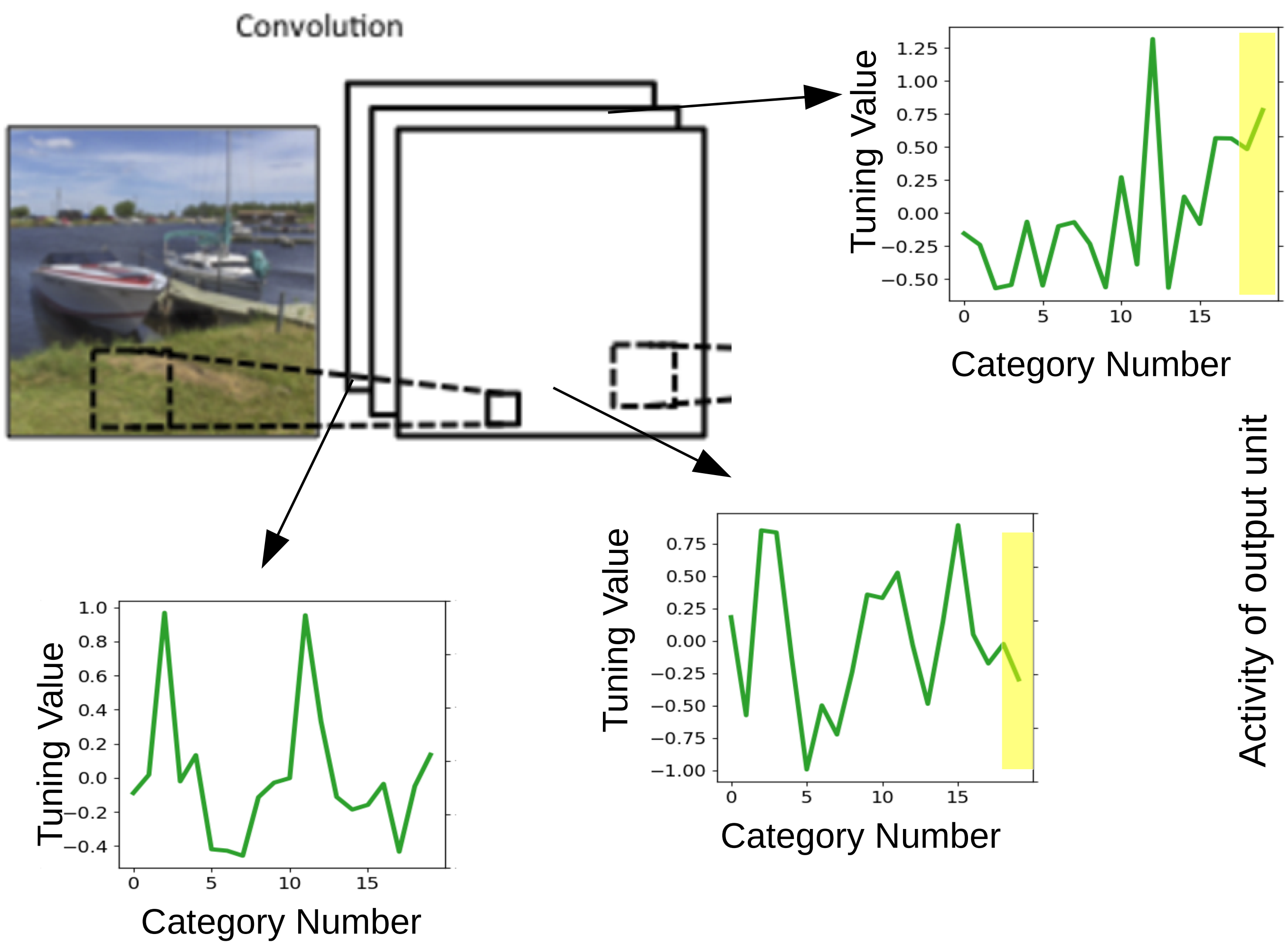


2.) When attention is applied to a category, the activity is scaled according to the tuning for that category:



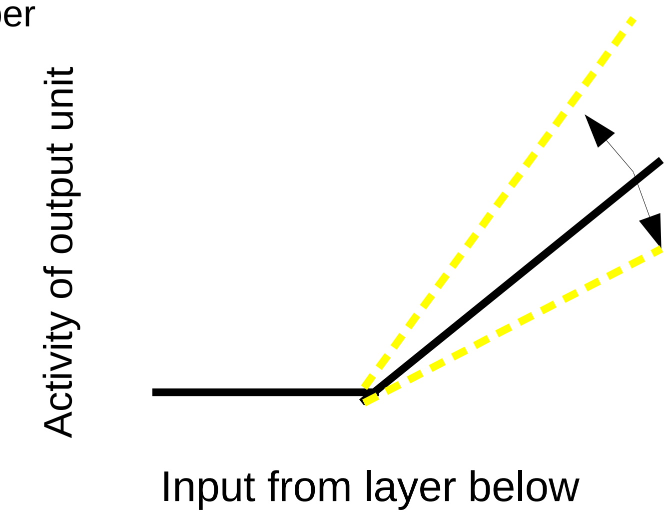
Replicating the neural correlates of attention

1.) Make normalized category tuning curves for each feature map:

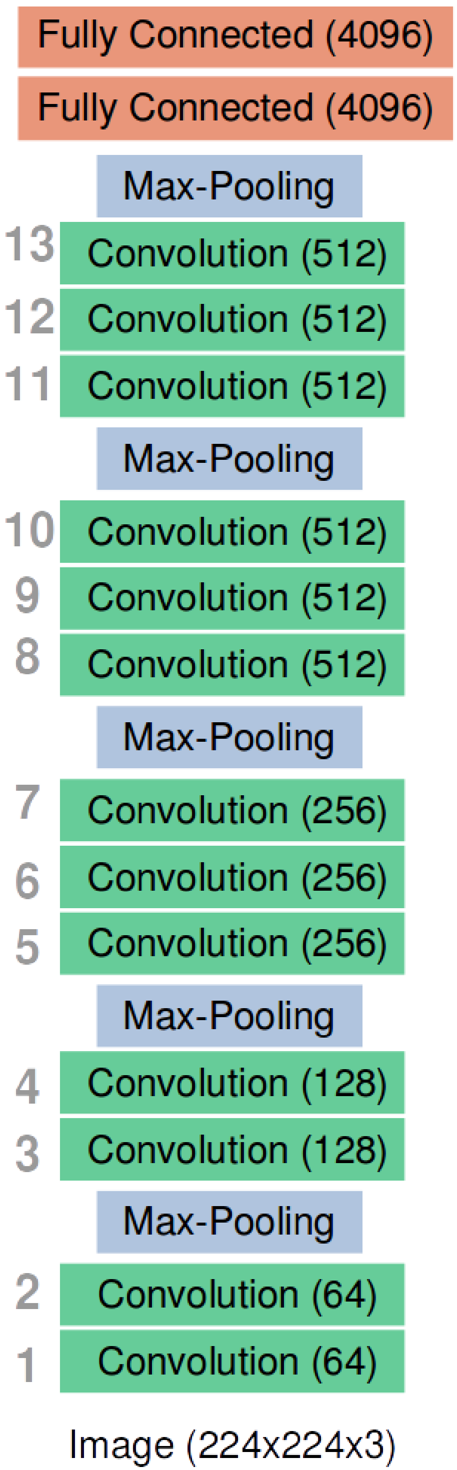


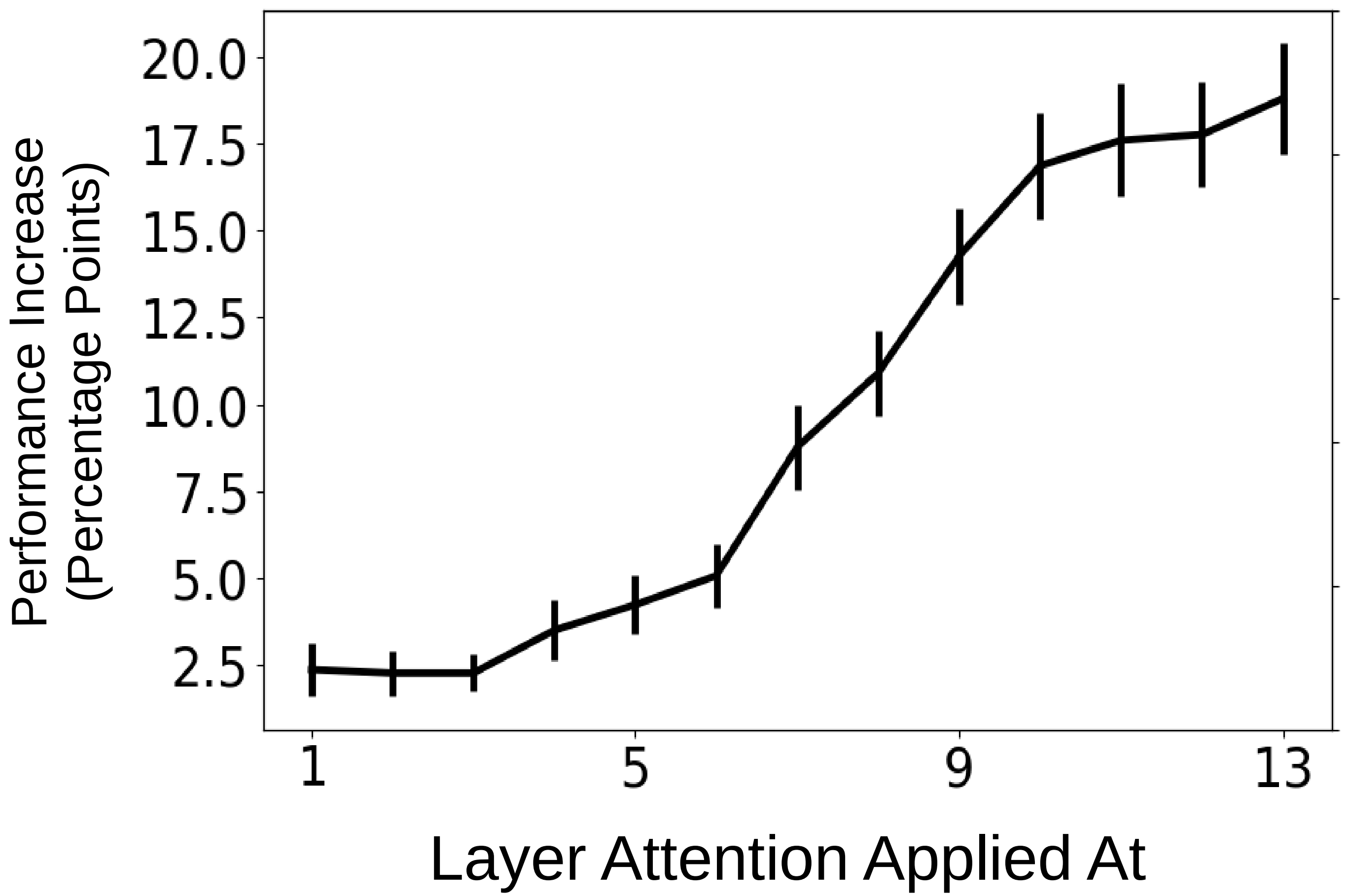
2.) When attention is applied to a category, the activity is scaled according to the tuning for that category:

$$x_{lk}^{ij} = (1 + \beta f_{lk}^{c})[I_{lk}^{ij}]_{+}$$

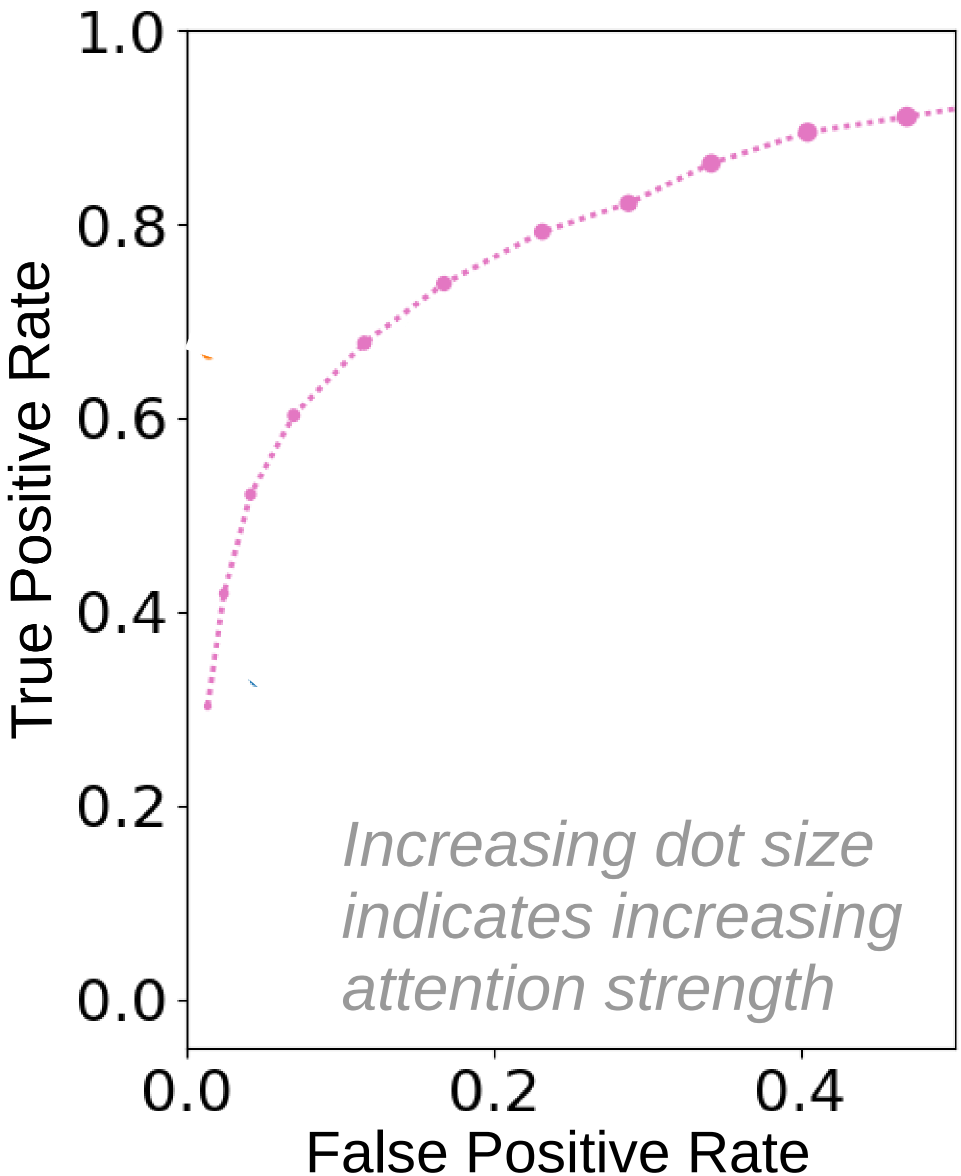


Implementing the neural correlates of attention lead to performance differences



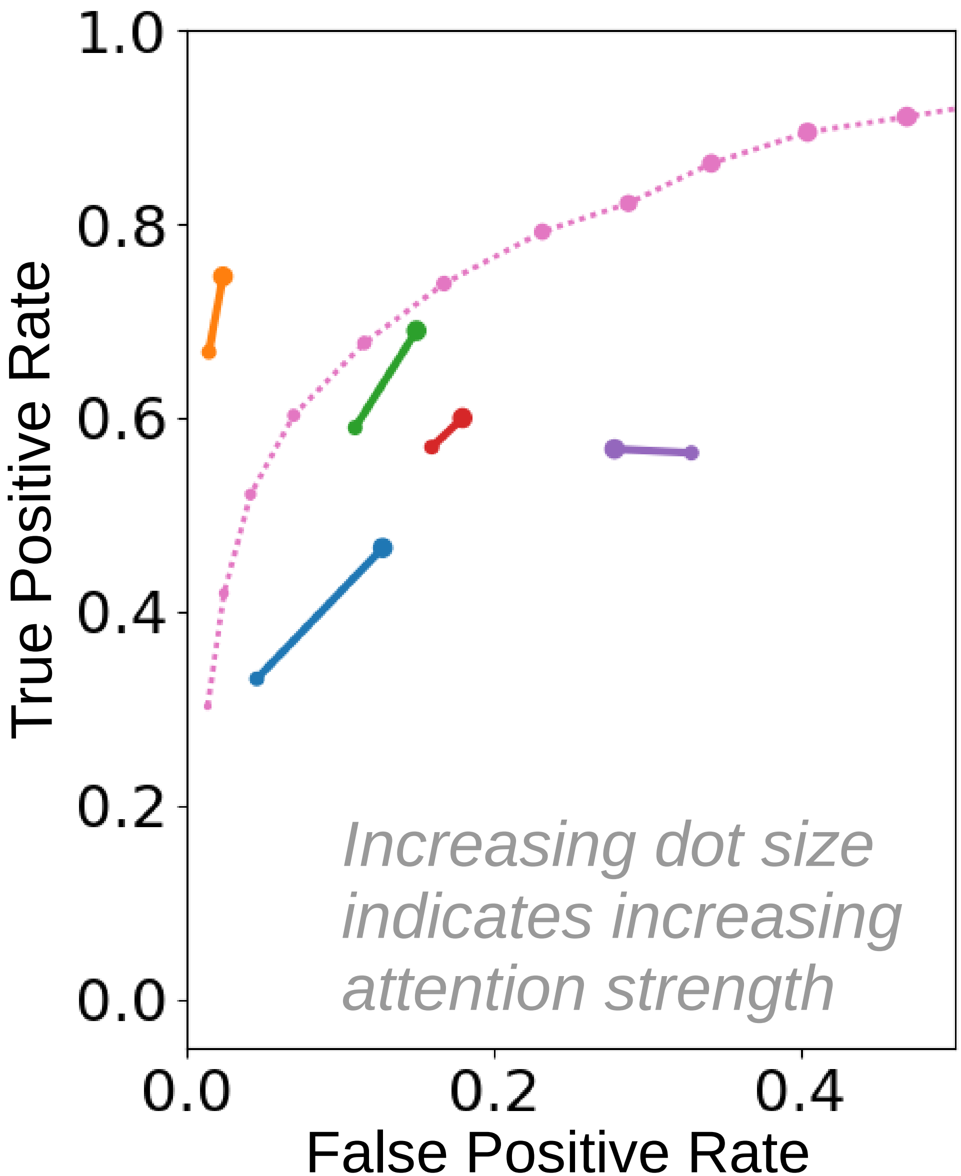


Relationship Between Attention Strength and Performance



- Increasing the overall strength of attention increases true ***and*** false positives

Relationship Between Attention Strength and Performance



- Increasing the overall strength of attention increases true ***and*** false positives
- This is true experimentally as well

Gary Lupyan and Michael J Spivey. PLoS One 2010

Gary Lupyan and Emily J Ward. PNAS 2013.

Mika Koivisto and Ella Kahila. Vision Research 2017.

On to the code!