Mouse visual cortex as a limited resource system that self-learns an ecologically-general representation

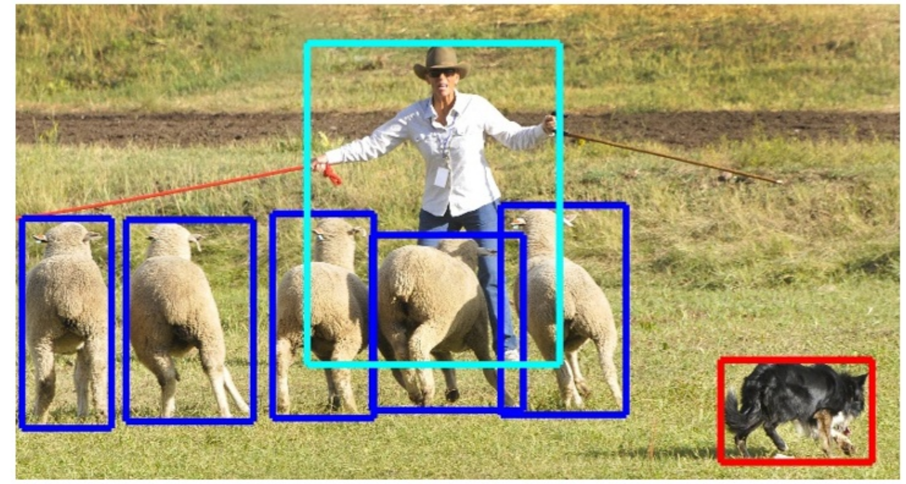
Aran Nayebi

WWNeuRise
2022.11.02

From Neurons to Behavior

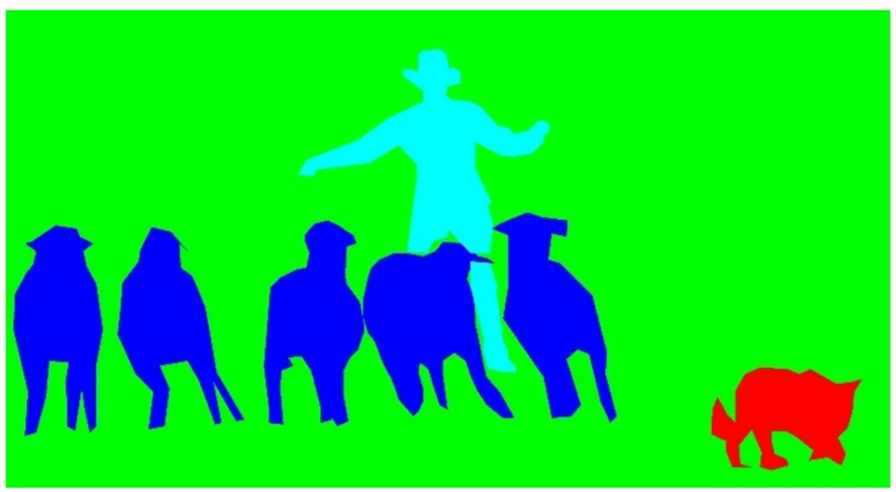
Classification, Segmentation, Localization, …

(a) **Image classification**

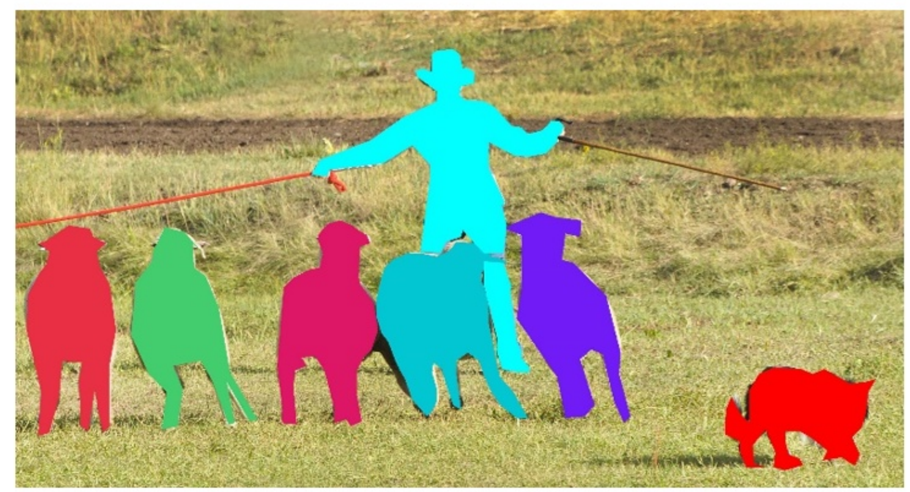


(b) **Object localization**

Lin et al. 2014

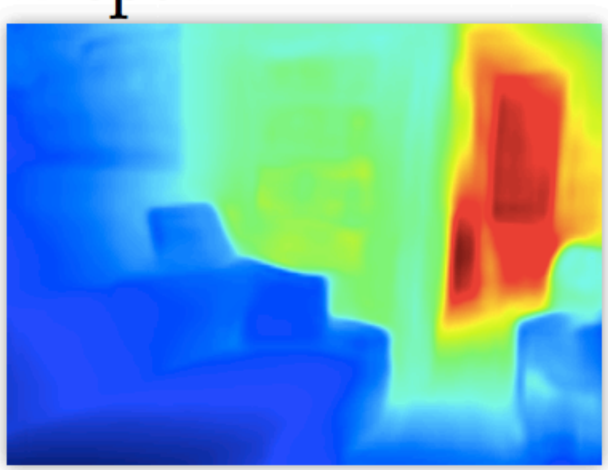


(c) **Semantic segmentation**

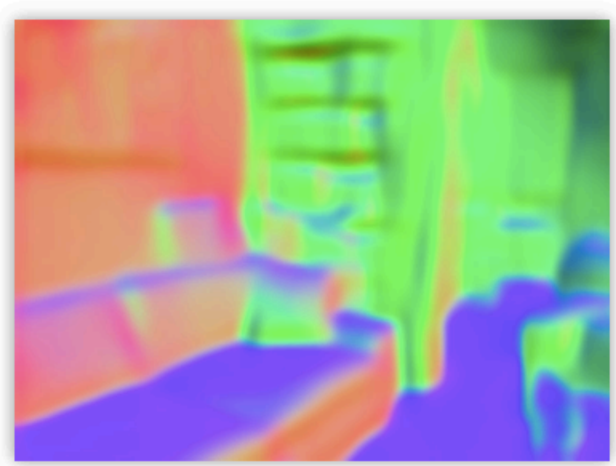


(d) **Instance segmentation**

Depth



Normals



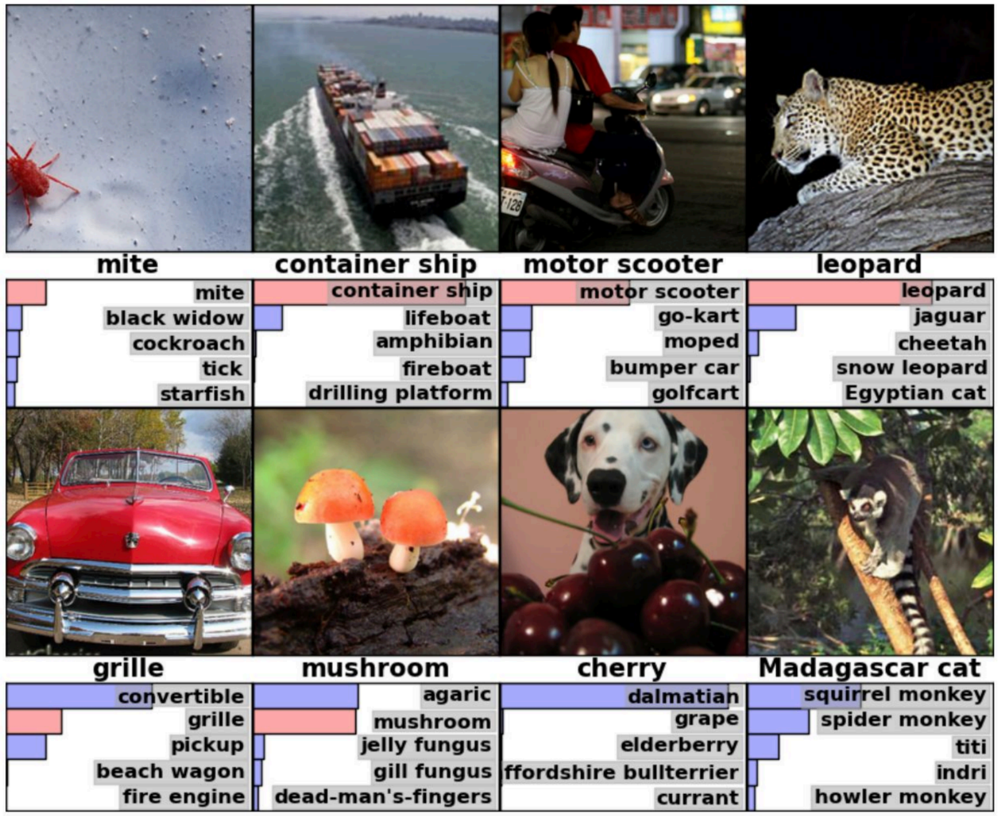
Eigen and Fergus 2015

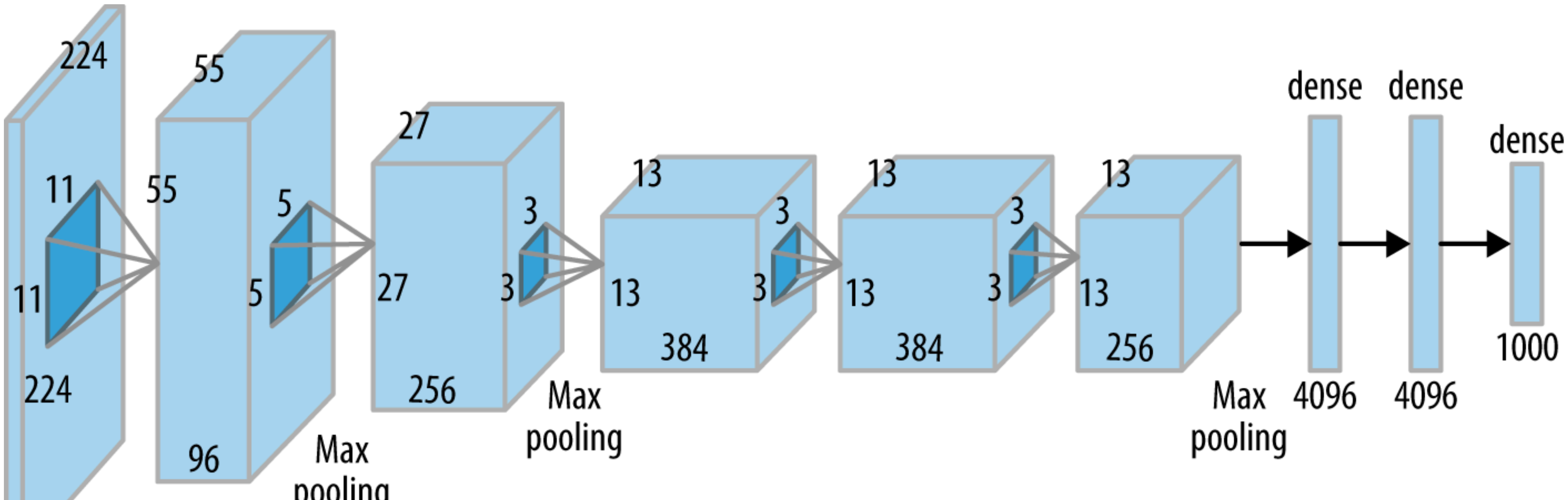
Convolutional Neural Networks (CNNs)

ImageNet Challenge

IMAGENET

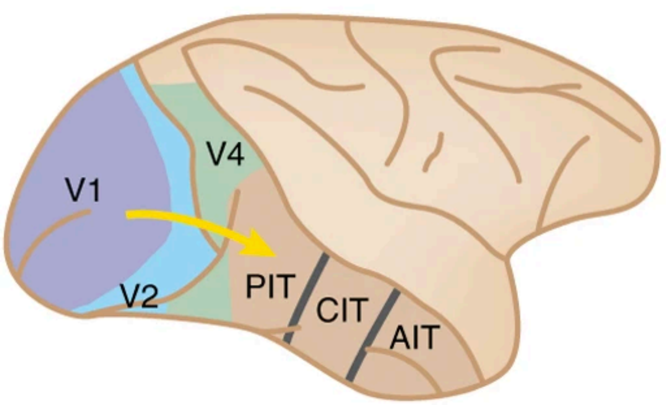
- 1,000 object classes (categories).
- Images:
  - 1.2 M train
  - 100k test.

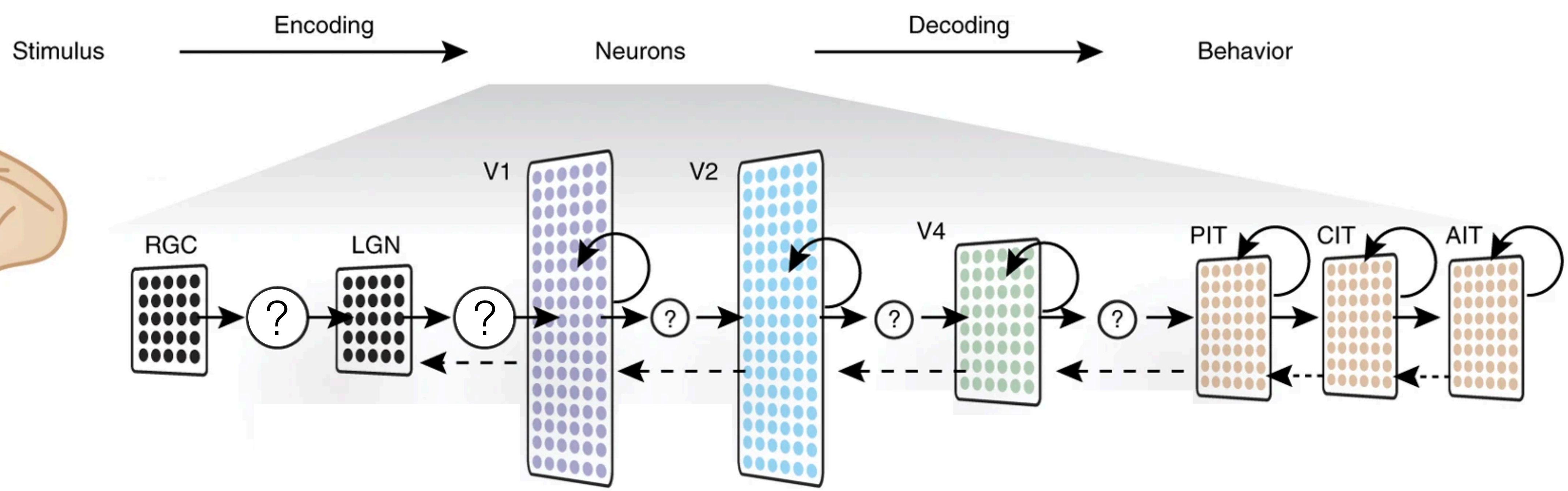




Krizhevsky et al. 2012

CNNs as Models of Object Recognition





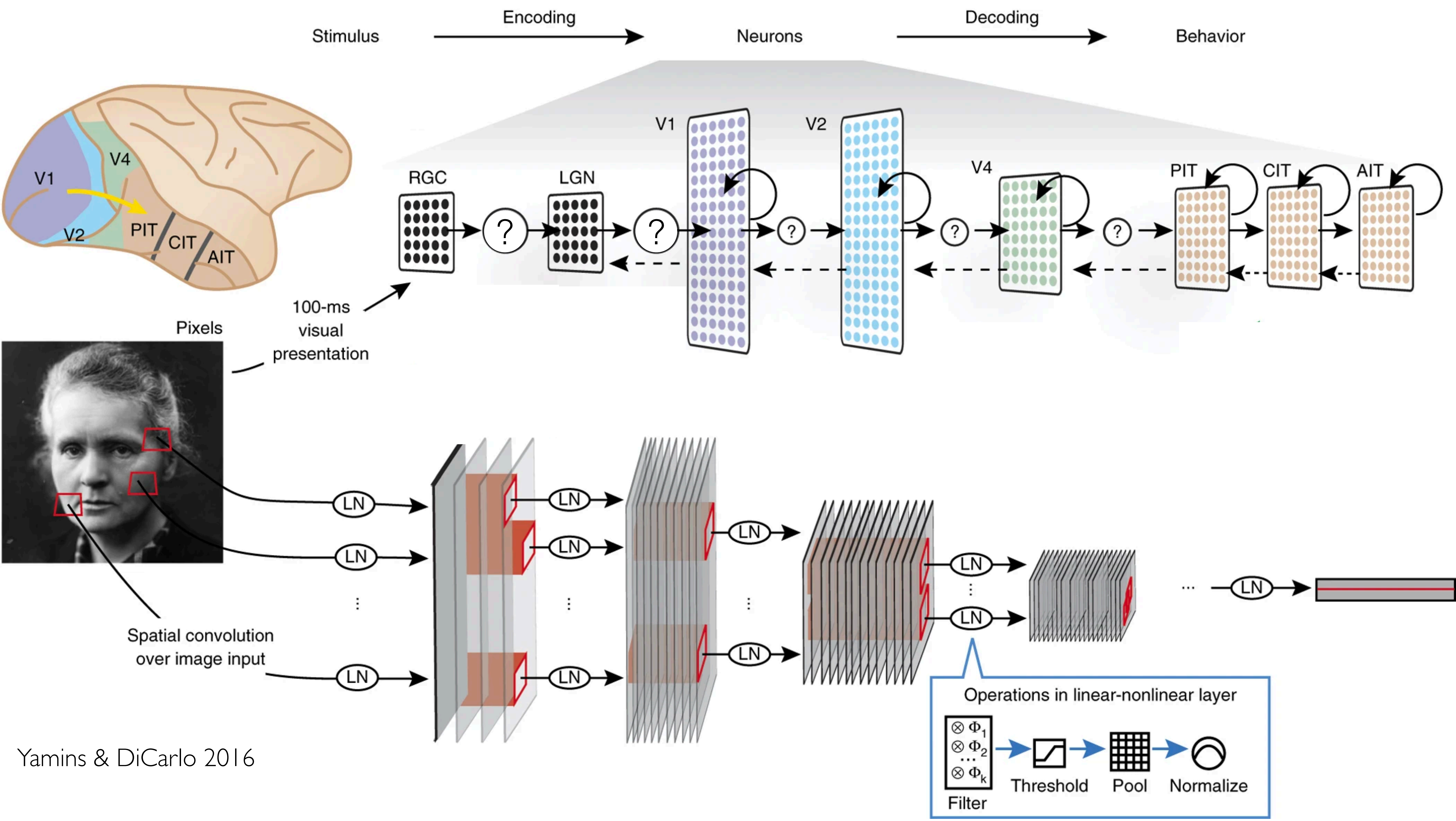
Convolutional Neural Networks (CNNs)

Fukushima, 1979; Lecun, 1995

CNNs are inspired by visual neuroscience:

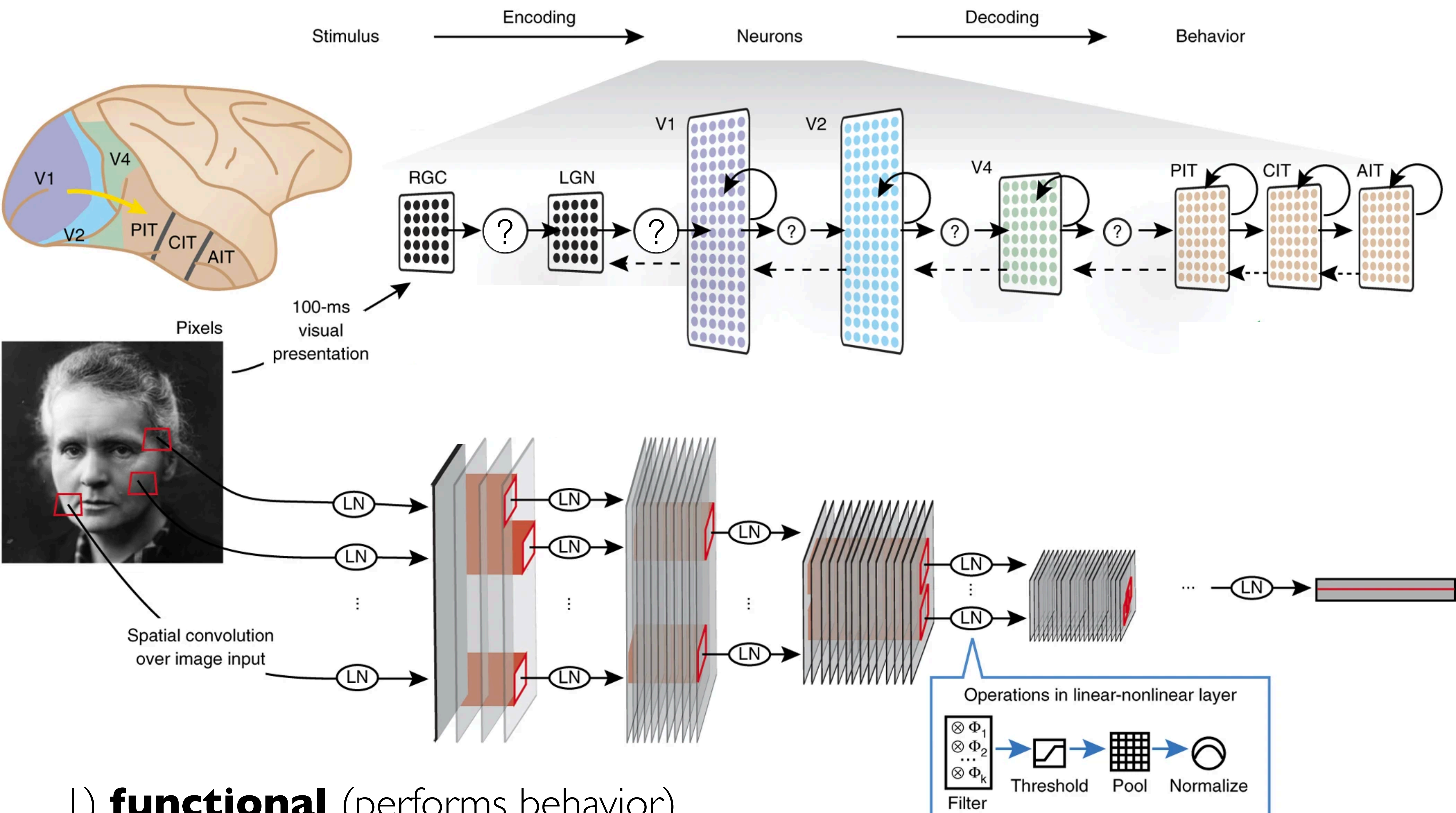
1) **hierarchy**
2) **retinotopy** (spatially tiled)

CNNs as Models of Object Recognition



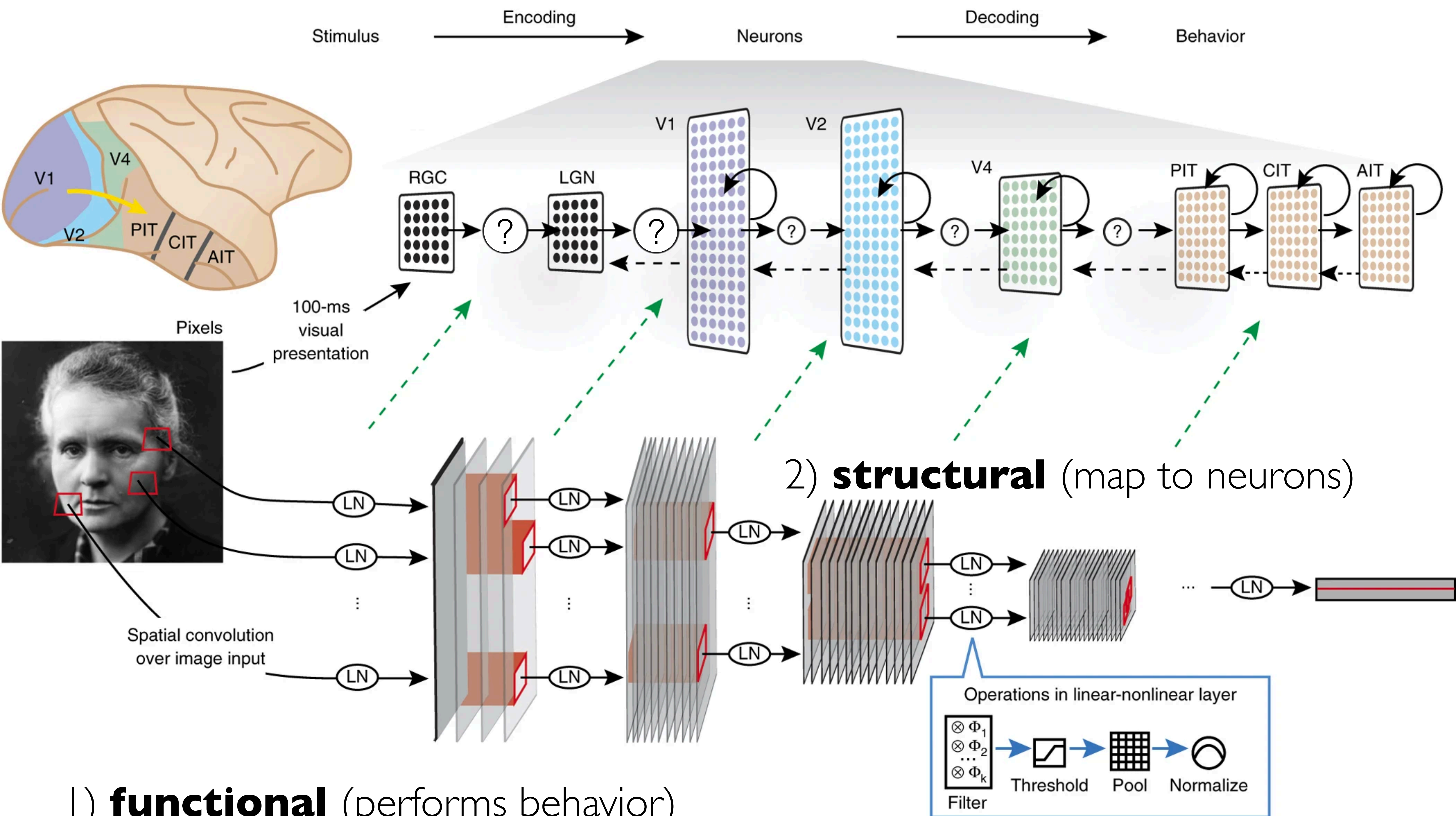
Yamins & DiCarlo 2016

CNNs as Models of Object Recognition



1) **functional** (performs behavior)

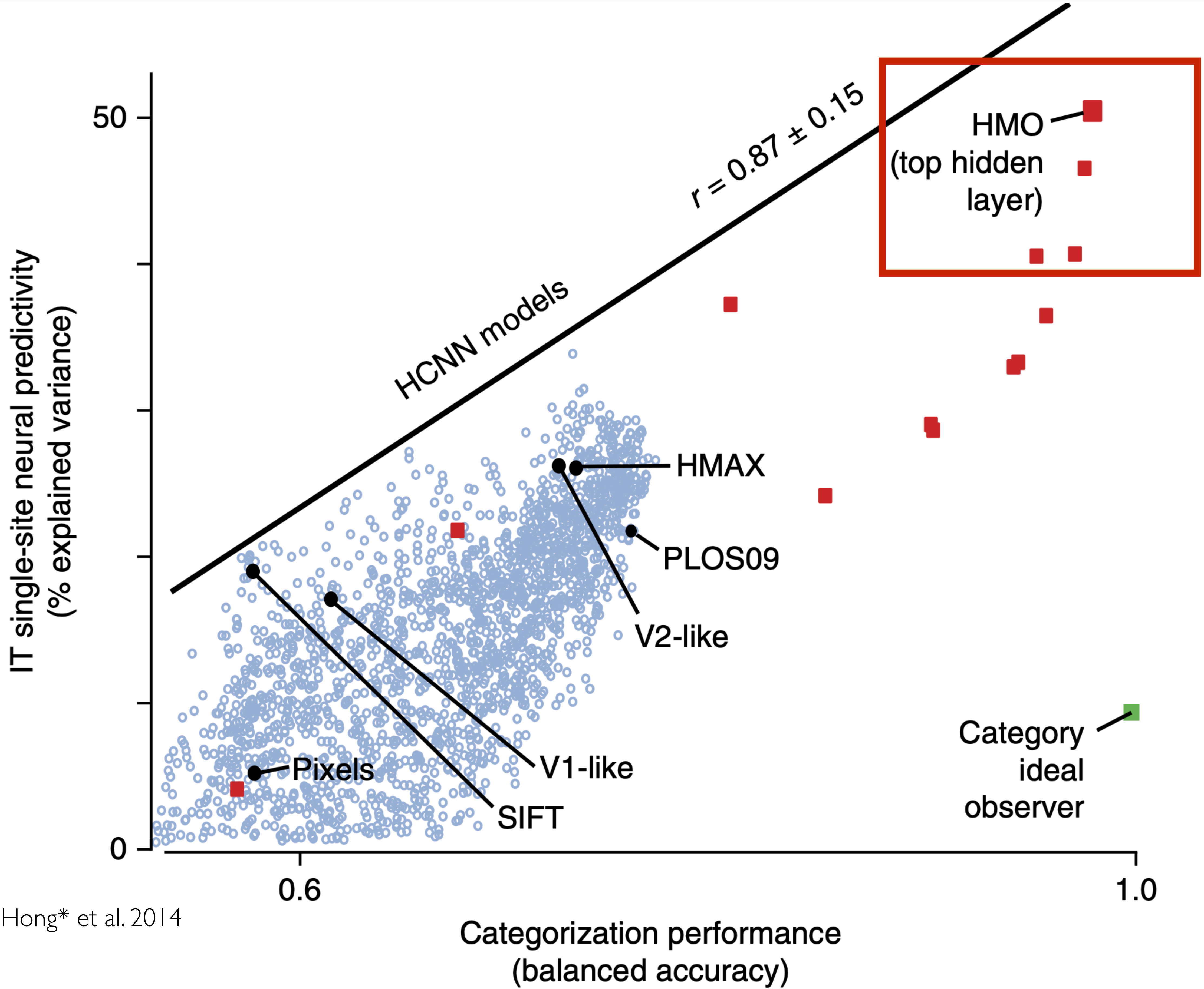
CNNs as Models of Object Recognition



2) **structural** (map to neurons)

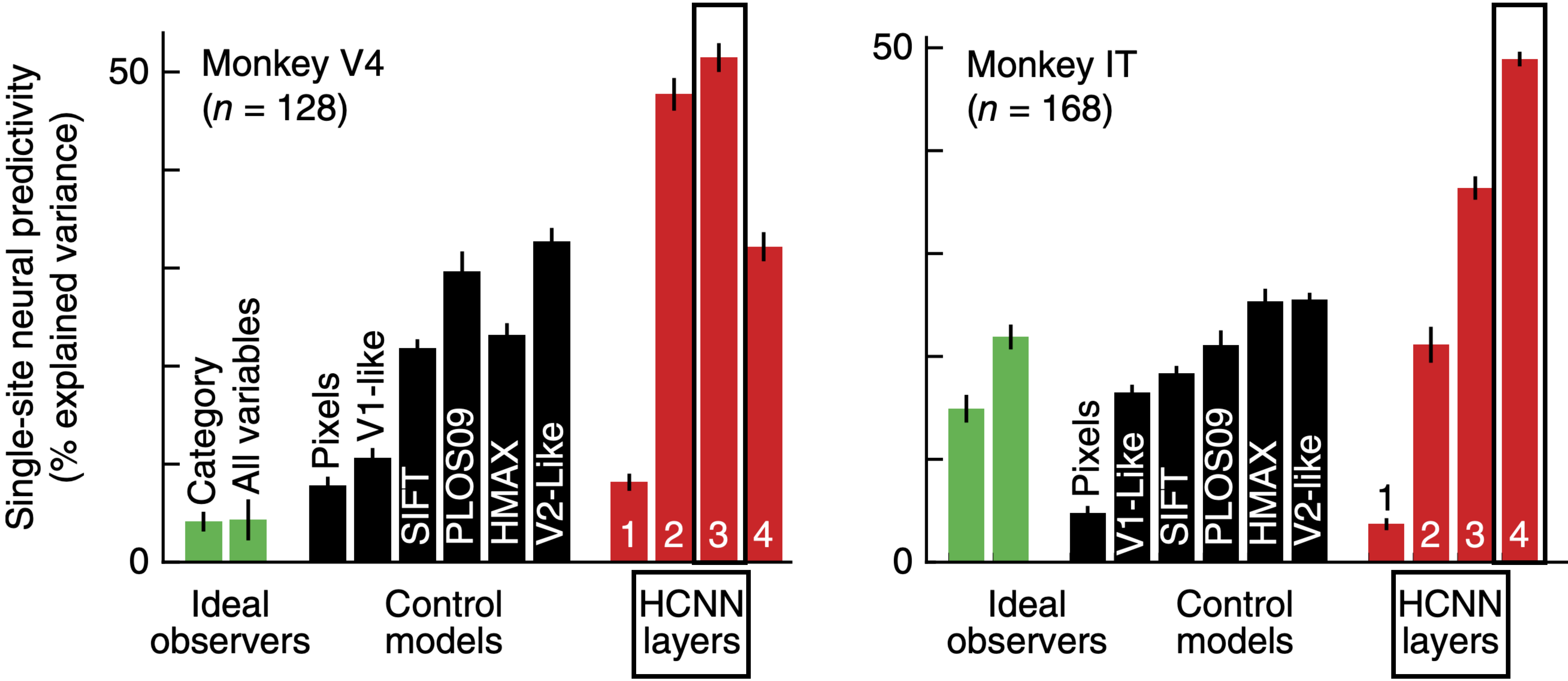
1) **functional** (performs behavior)

Categorization performance correlated with neural predictivity



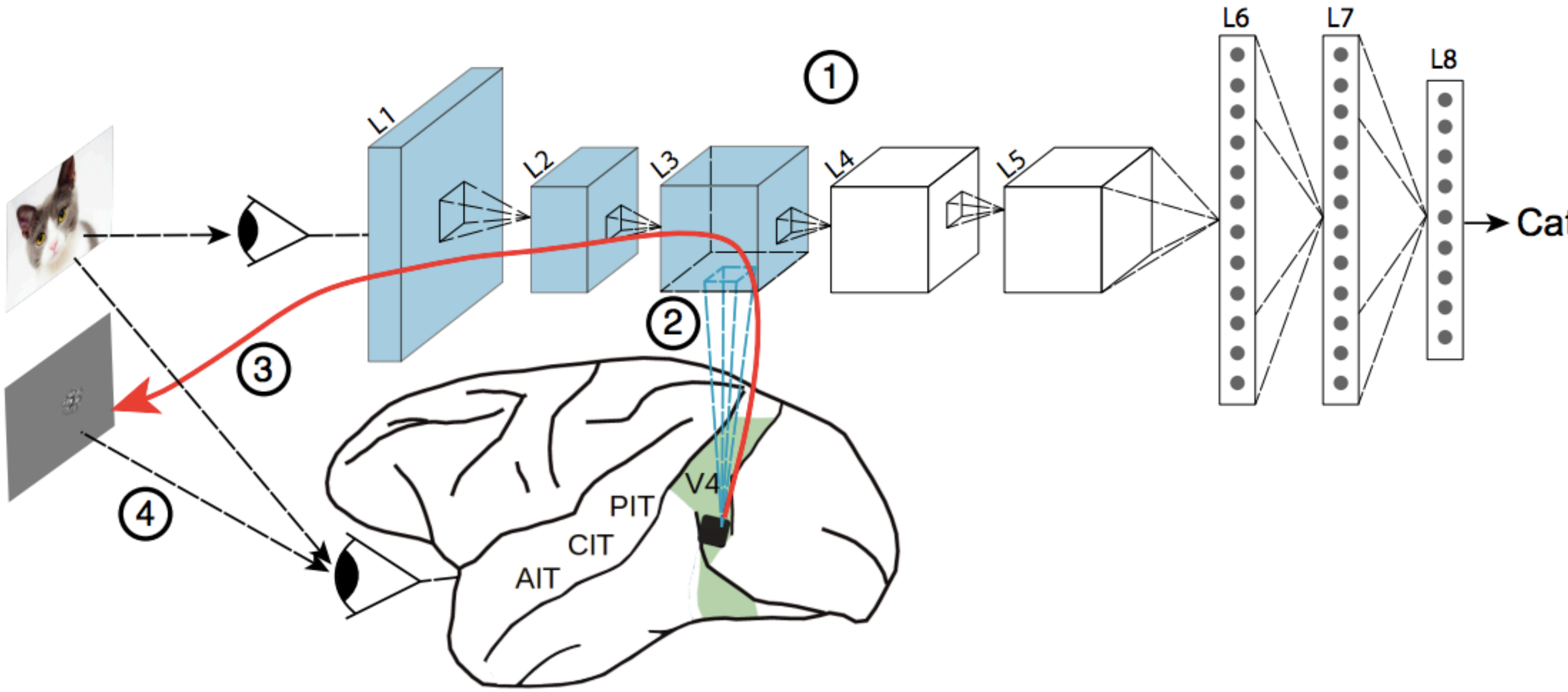
Yamins*, Hong* et al. 2014

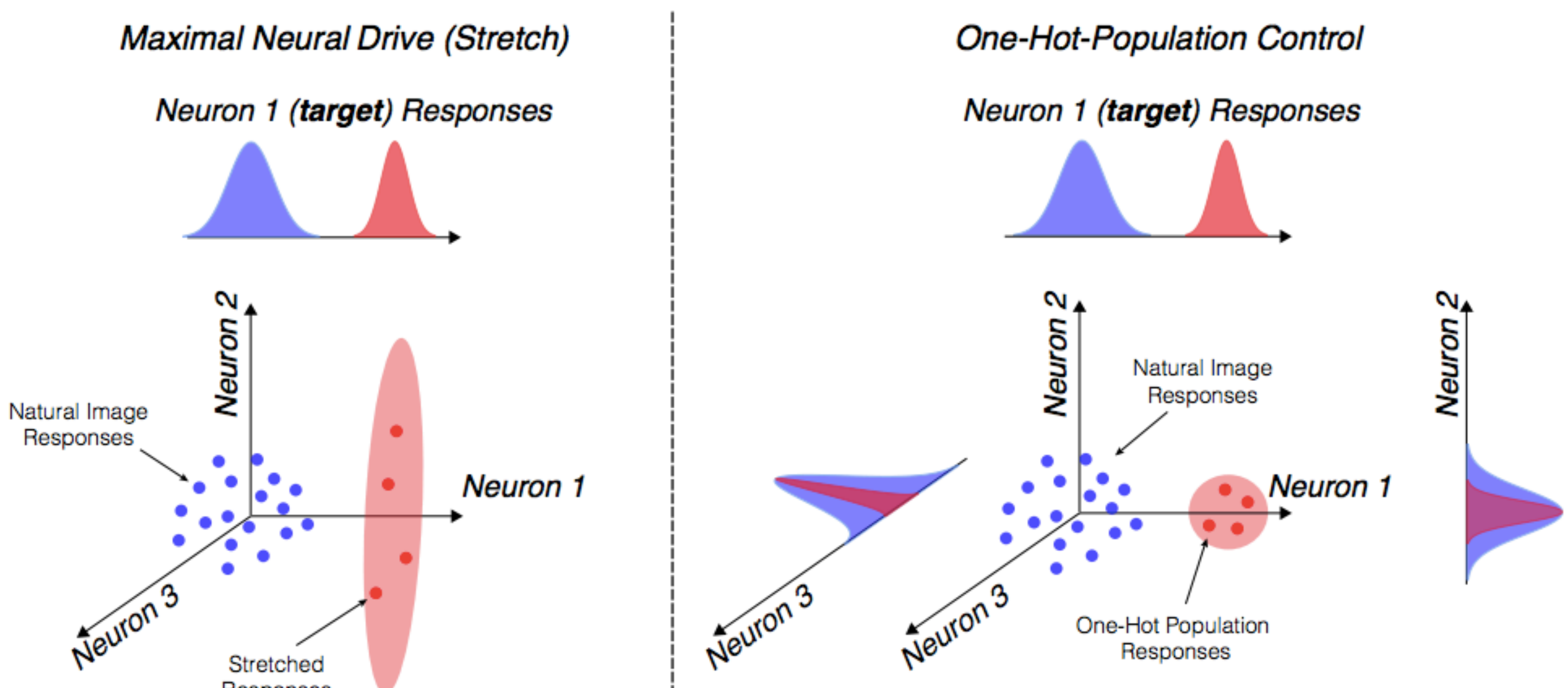
Hierarchy as a by-product of task optimization



Yamins*, Hong* et al. 2014

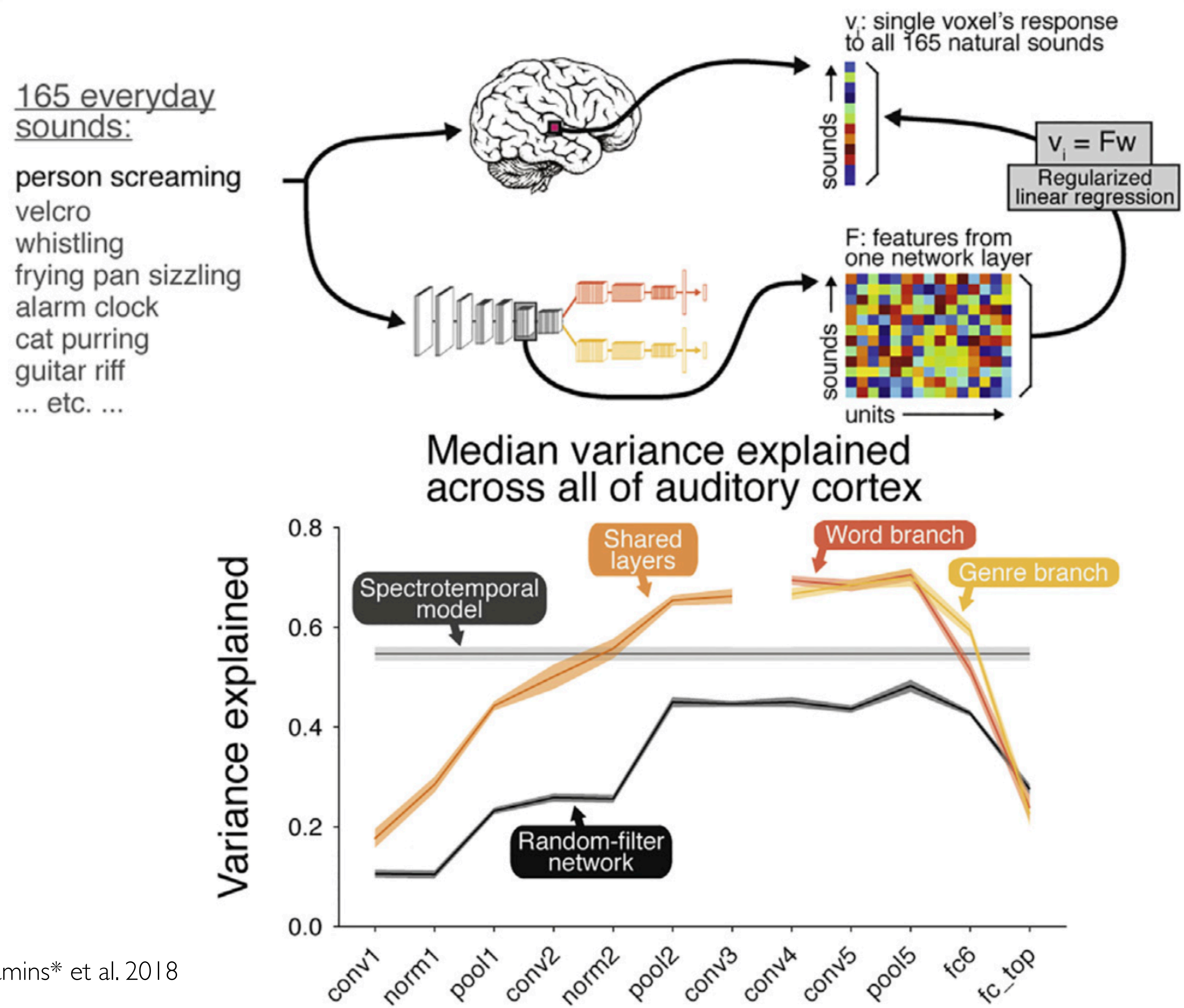
Neural population control of intermediate areas





Bashivan et al. 2019

Not just limited to visual cortex, but also auditory cortex

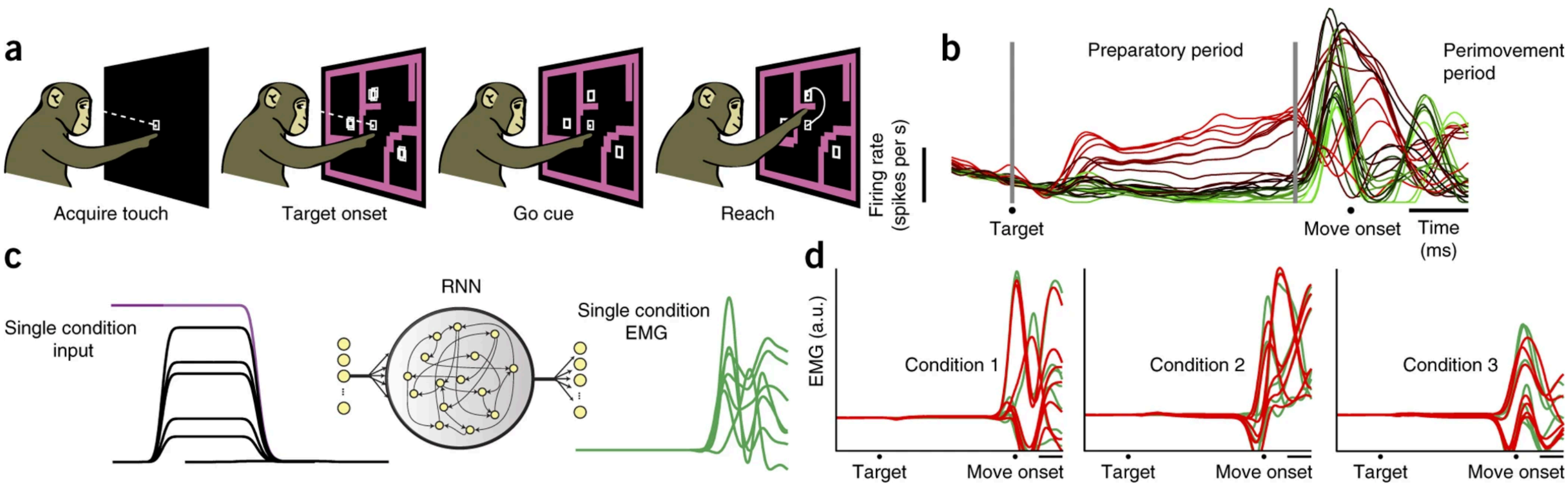


Kell*, Yamins* et al. 2018

Not just sensory, but applicable to motor areas

A neural network that finds a naturalistic solution for the production of muscle activity

David Sussillo, Mark M Churchland, Matthew T Kaufman & Krishna V Shenoy



RNN trained to mimic muscle activities (EMG) as a function of condition

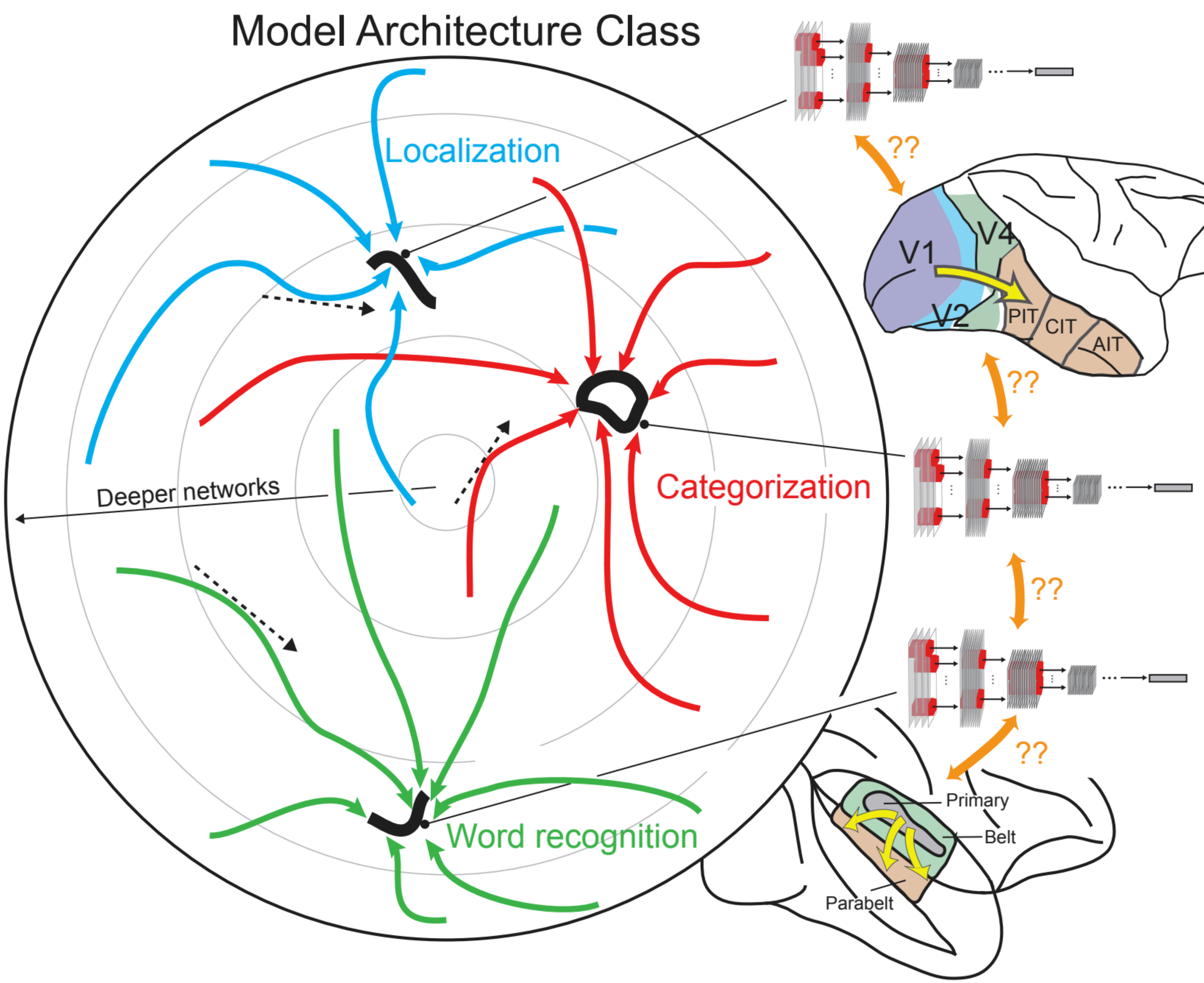
Goal-Driven Modeling (Sensory)

Sensory:

> Formulate comprehensive model class (**CNNs**)

> Choose challenging, ethologically-valid tasks (**categorization**)

> Implement generic learning rules (**gradient descent**)



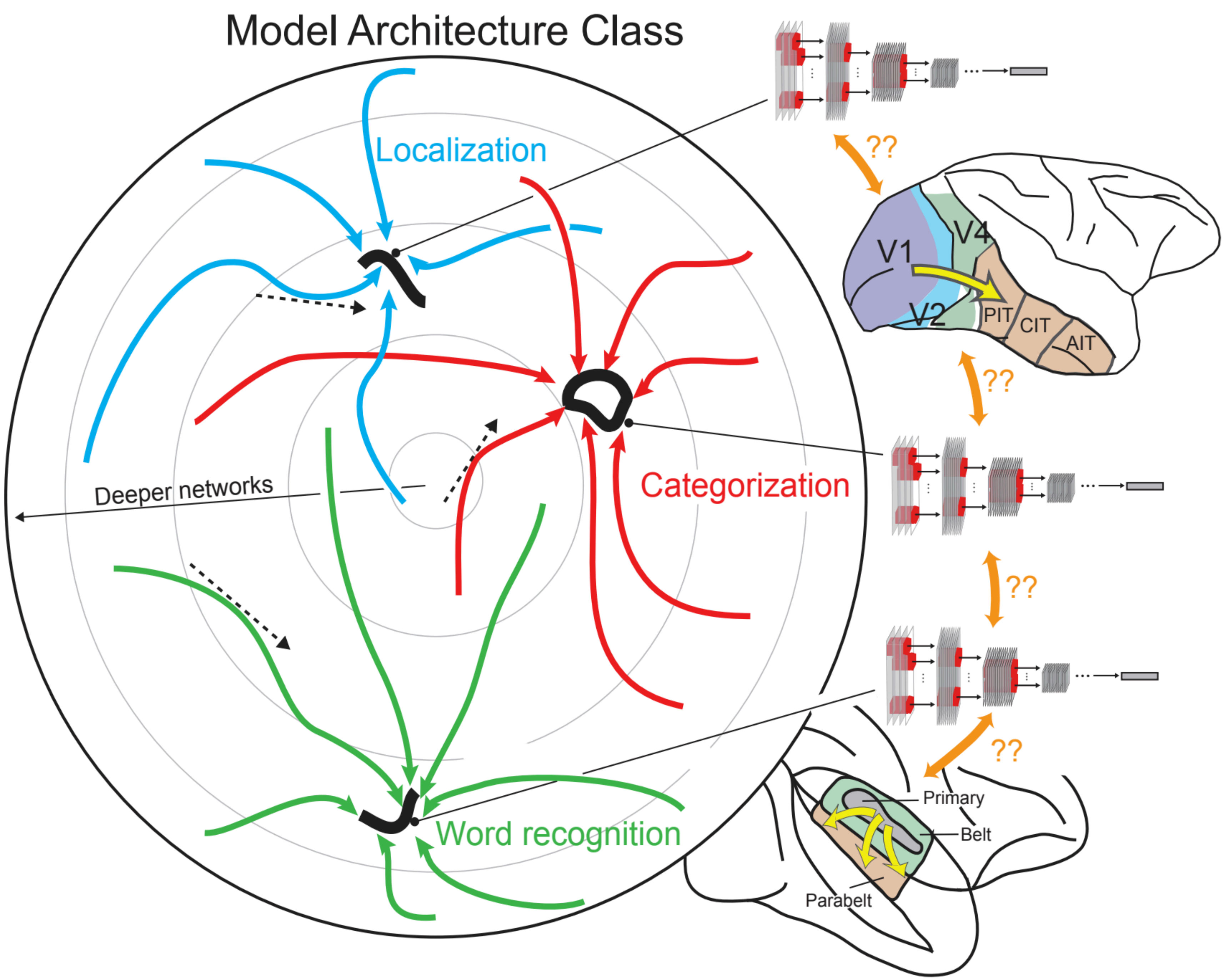
Goal-Driven Modeling (Motor)

Motor

> Formulate comprehensive model class (**RNNs**)

> Choose challenging, ethologically-valid tasks (**motion generation**)

> Implement generic learning rules (**gradient descent**)



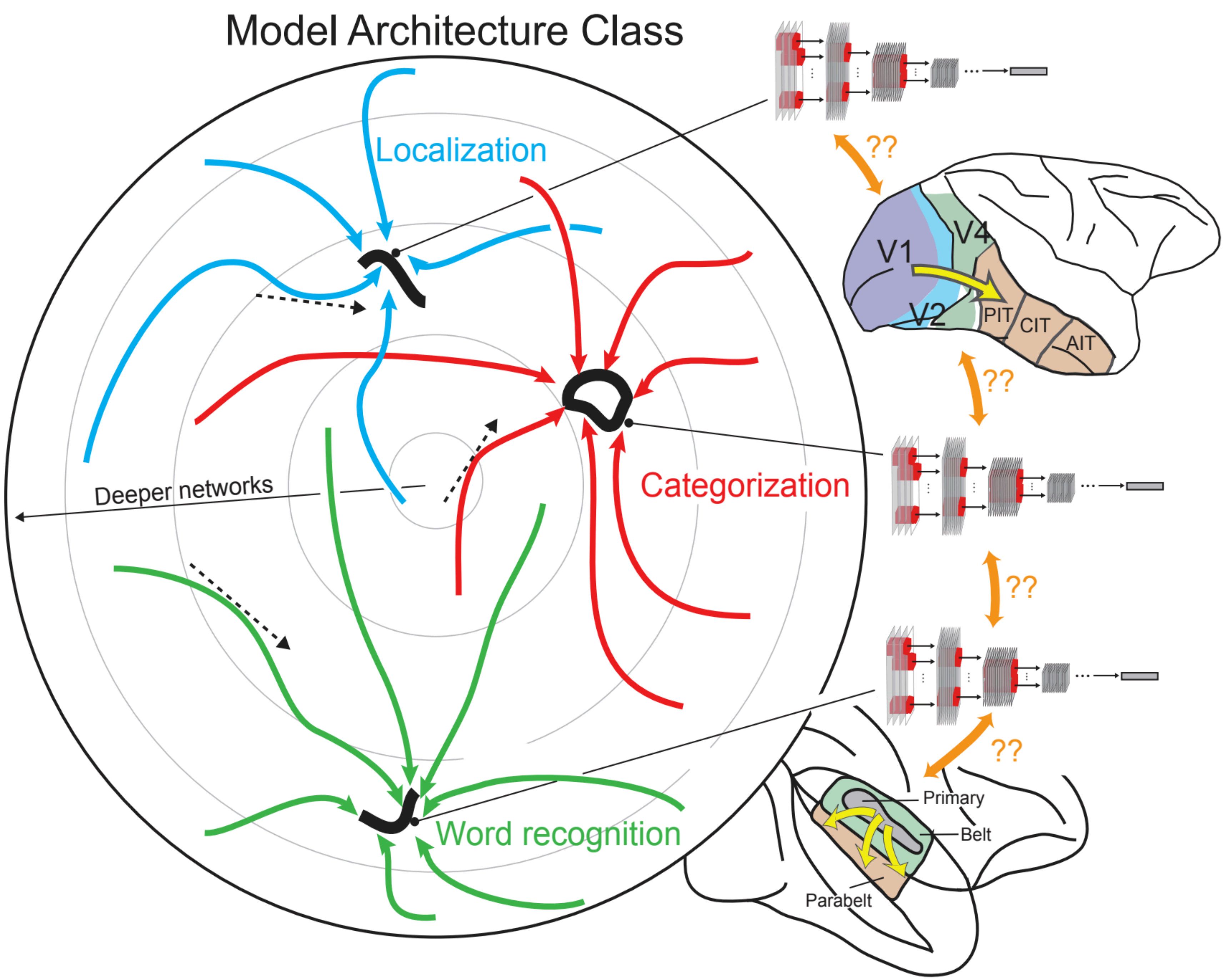
Goal-Driven Modeling

Motor

> Formulate comprehensive model class (**RNNs**)

> Choose challenging, ethologically-valid tasks (**motion generation**)

> Implement generic learning rules (**gradient descent**)



Similarity between sensory and motor:
Goal-driven optimization useful in both

Goal-Driven Modeling - Three Primary Components

"circuit"

1.

A = *architecture class*

"behavior"

2.

L = *loss function* ***D*** = *dataset*

"plasticity mechanism"

3.

Learning Rule

$$\underset{a \in \mathcal{A}}{\operatorname{argmin}}[L(p_a^*)]$$

genetic algorithms

where p^* is result of

backprop

$$\frac{dp_a}{dt} = -\lambda(t) \cdot \langle \nabla_{p_a} L(x) \rangle_{x \in \mathcal{D}}$$

Goal-Driven Modeling - Three Primary Components

"circuit"

1.

A *= architecture class*

"behavior"

2.

L *= loss function* ***D*** *= dataset*

A. Nayebi*, N. Kong* *et al.*
Mouse visual cortex as a limited resource system that self-learns an ecologically-general representation. *bioRxiv (2021)*

"plasticity mechanism"

3.

Learning Rule

$$\underset{a \in \mathcal{A}}{\operatorname{argmin}} [L(p_a^*)]$$

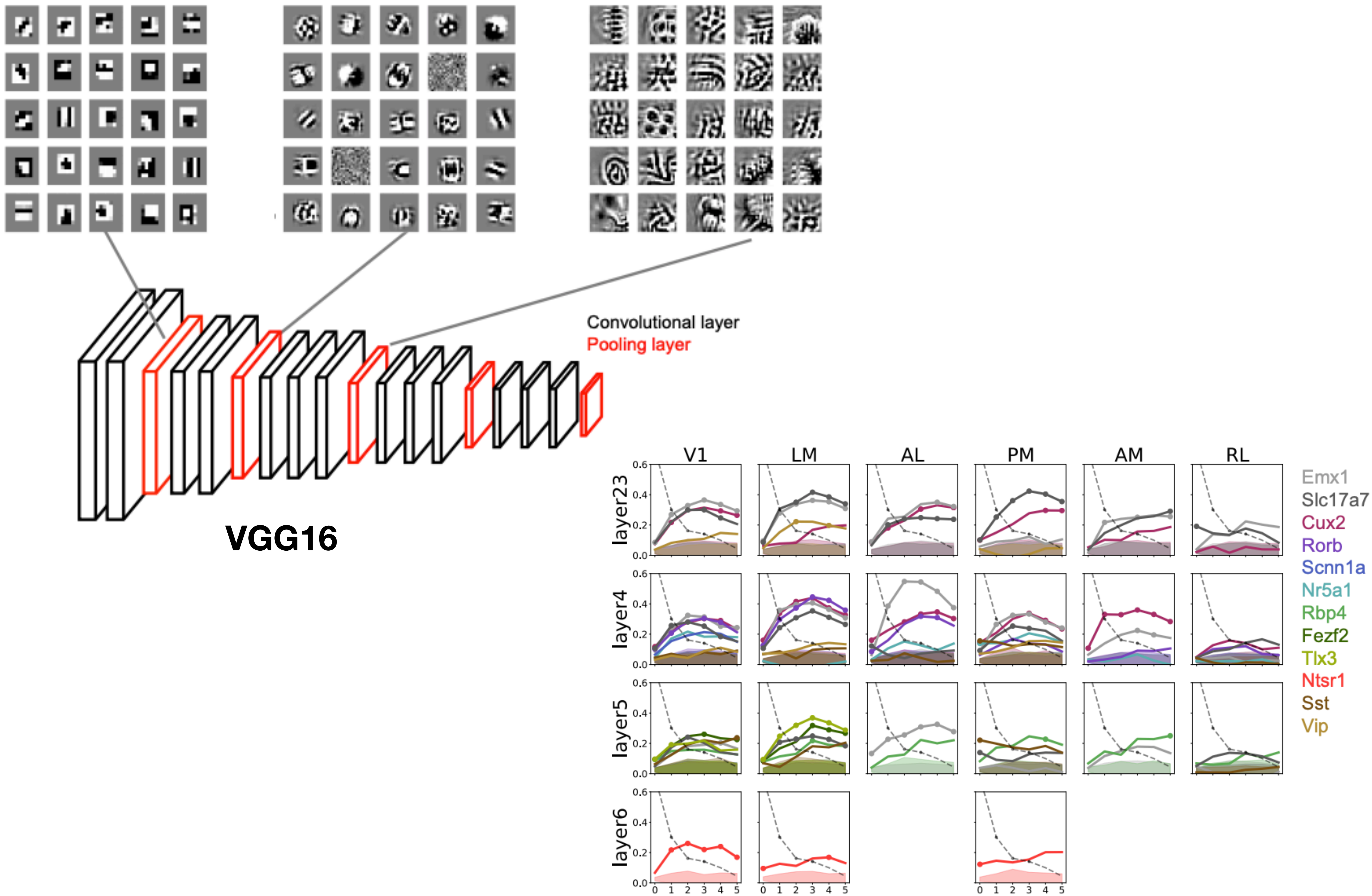
genetic algorithms

where p^* is result of

backprop

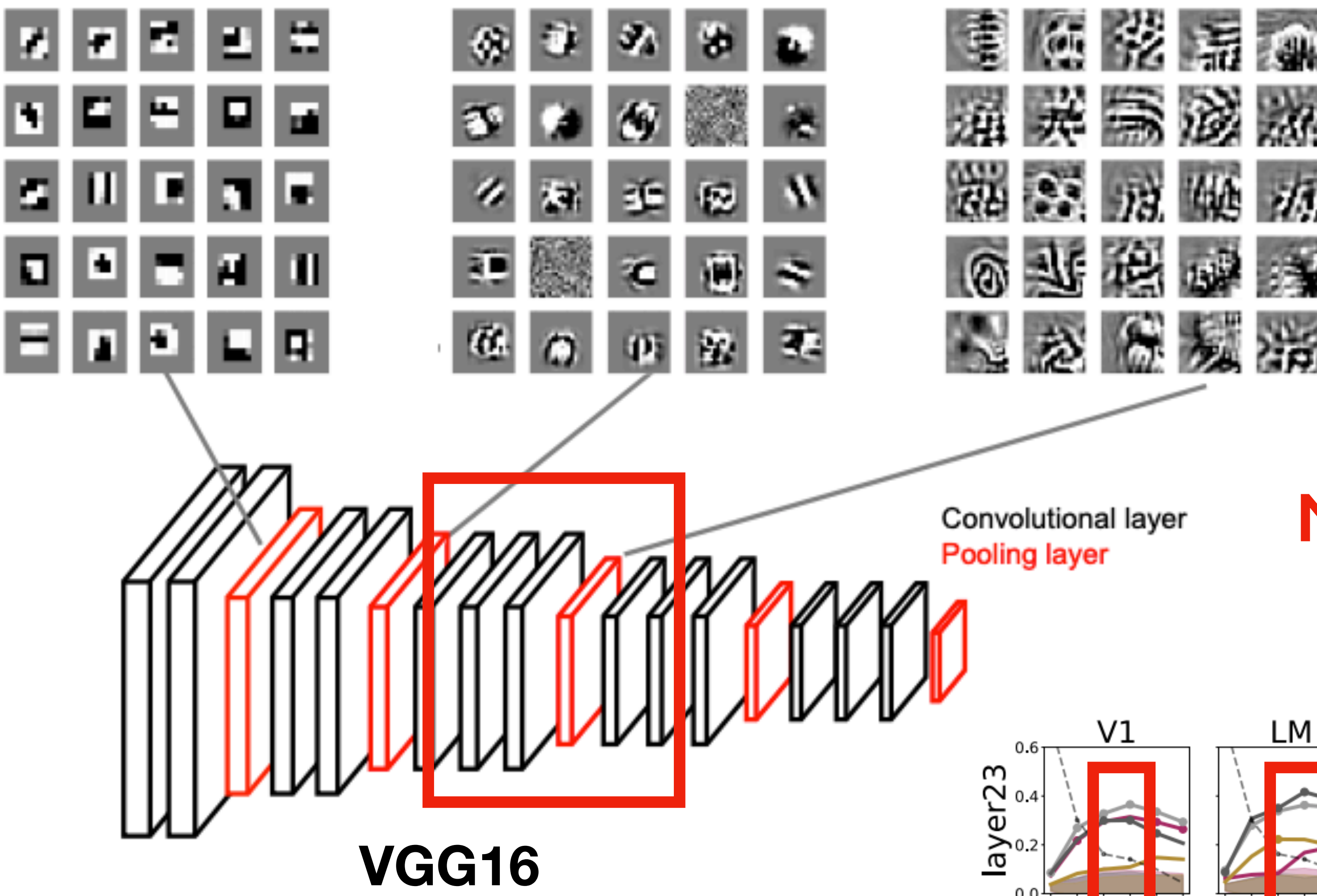
$$\frac{dp_a}{dt} = -\lambda(t) \cdot \langle \nabla_{p_a} L(x) \rangle_{x \in \mathcal{D}}$$

Deep models suggest mouse vision is representationally deep

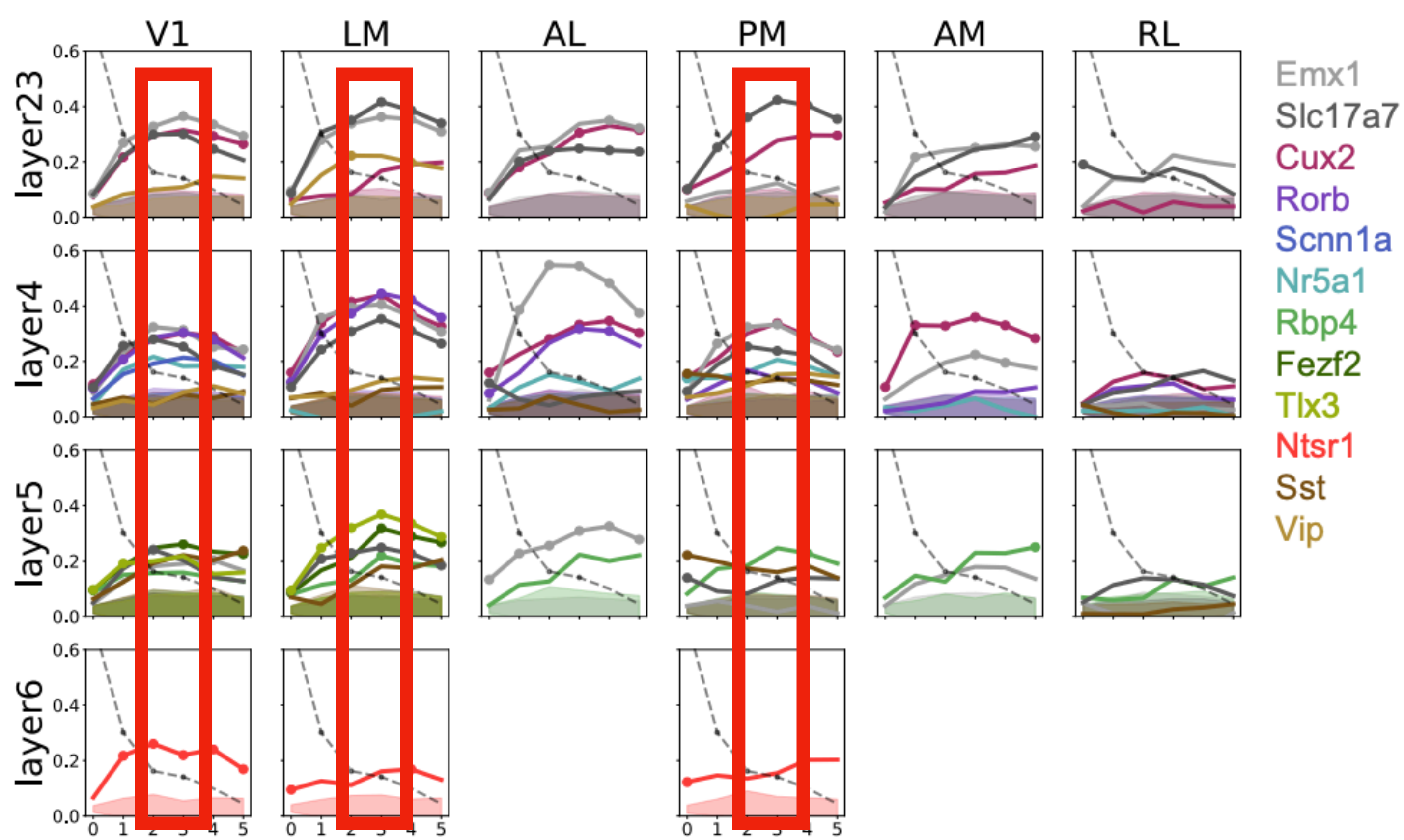


de Vries et al. 2020

Deep models suggest mouse vision is representationally deep

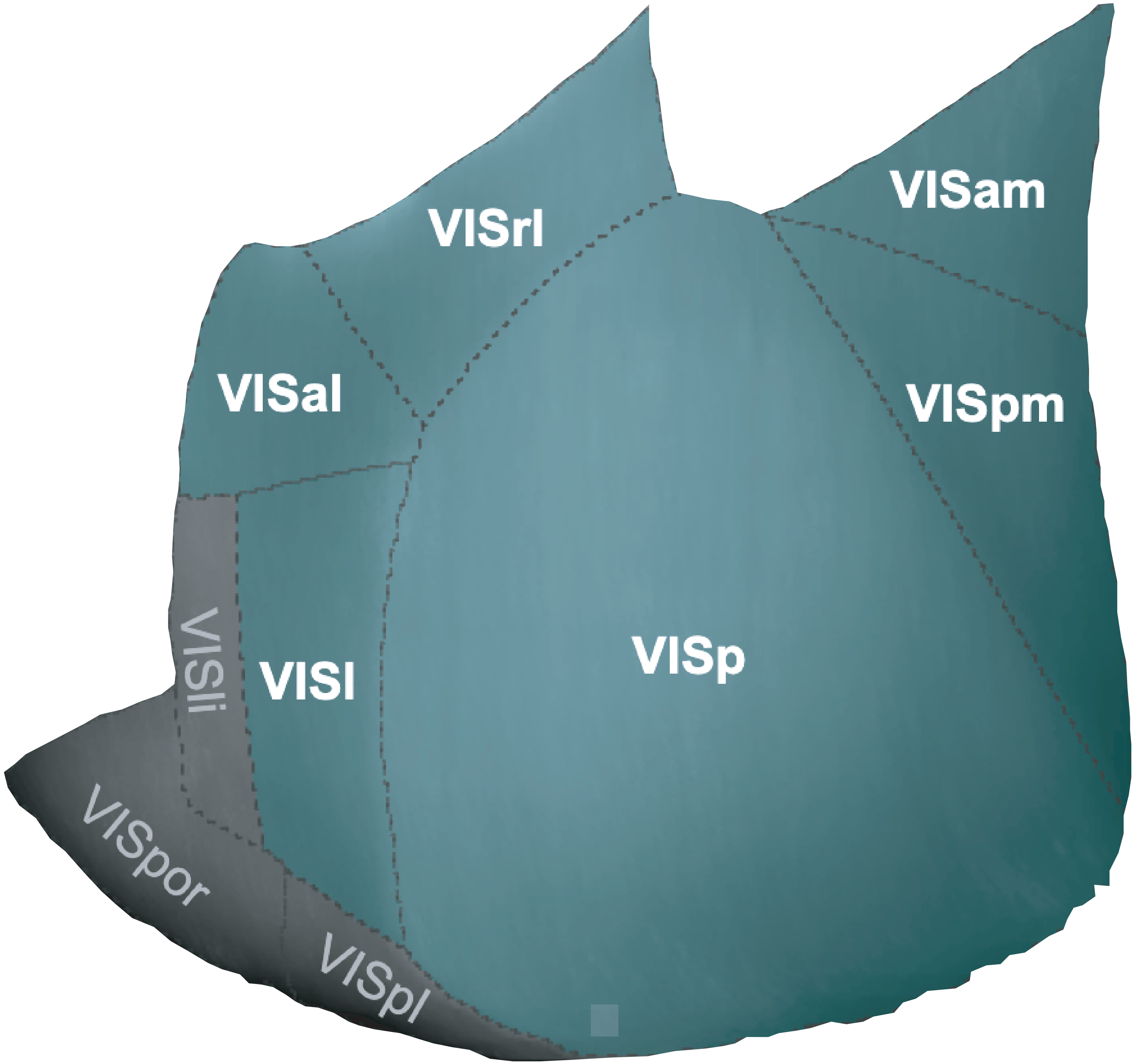


Mid-levels best predict visual areas



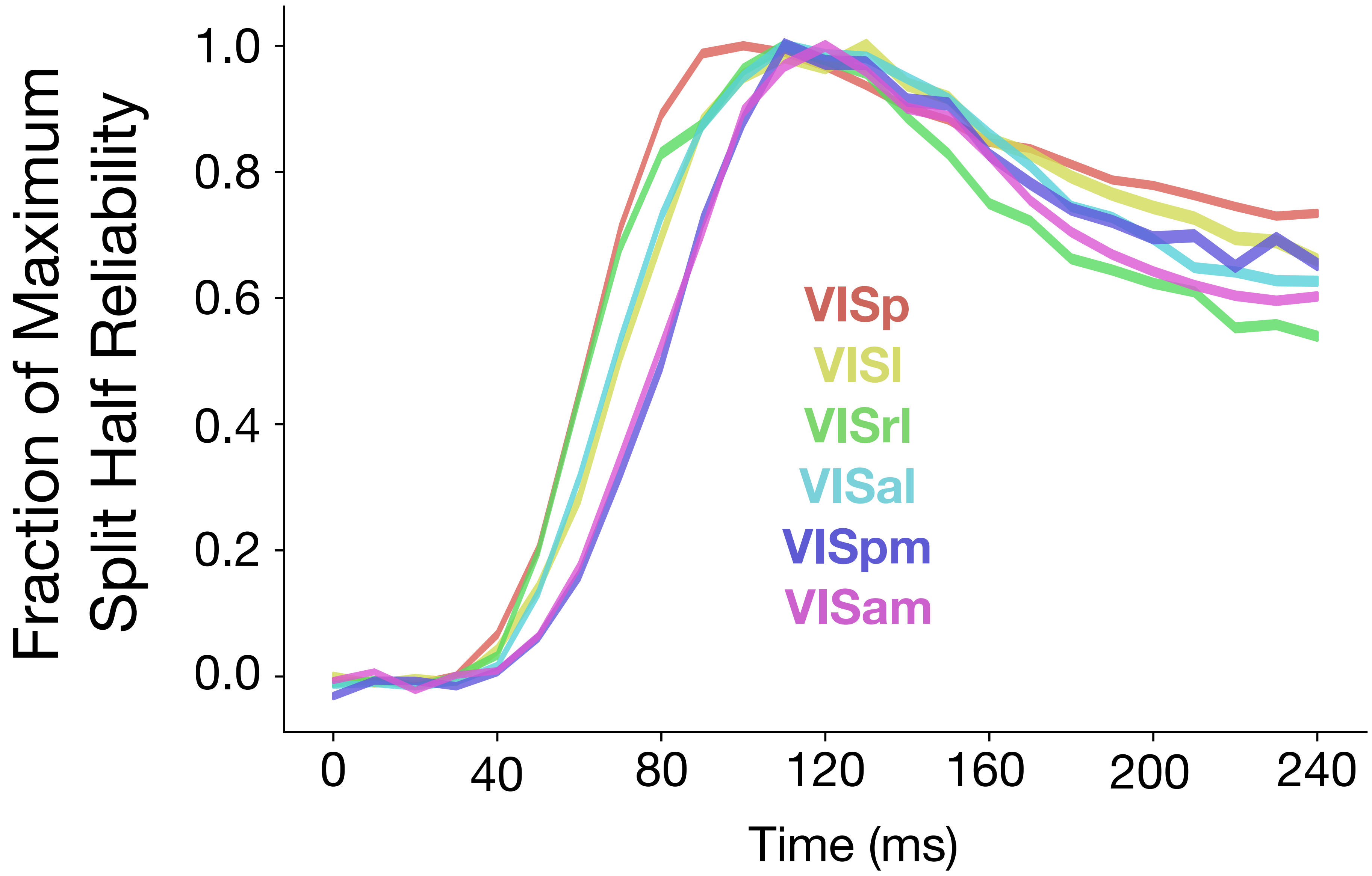
de Vries et al. 2020

But mouse visual cortex is anatomically shallow!

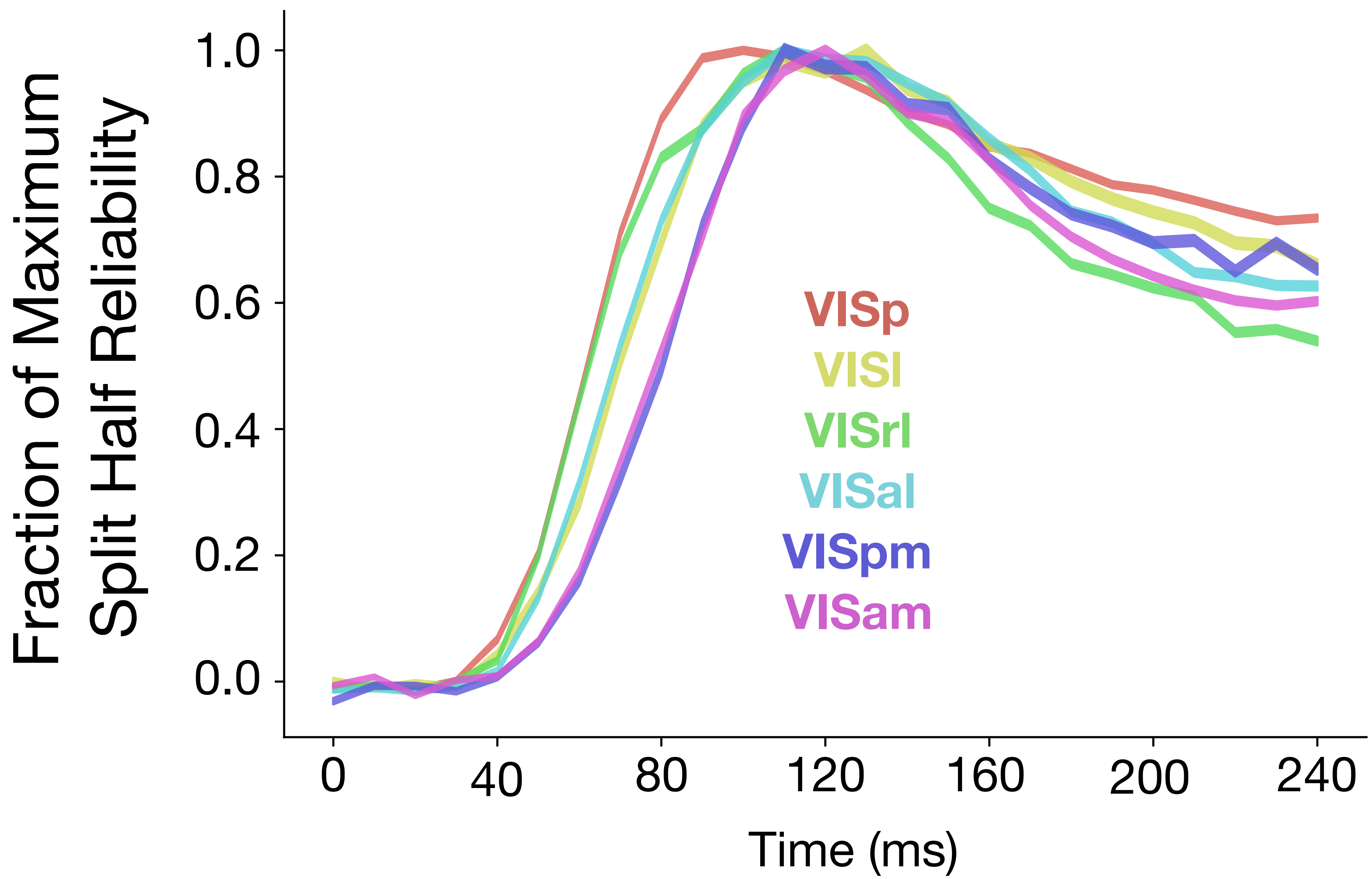


Allen Mouse Brain Atlas

But mouse visual cortex is anatomically shallow!

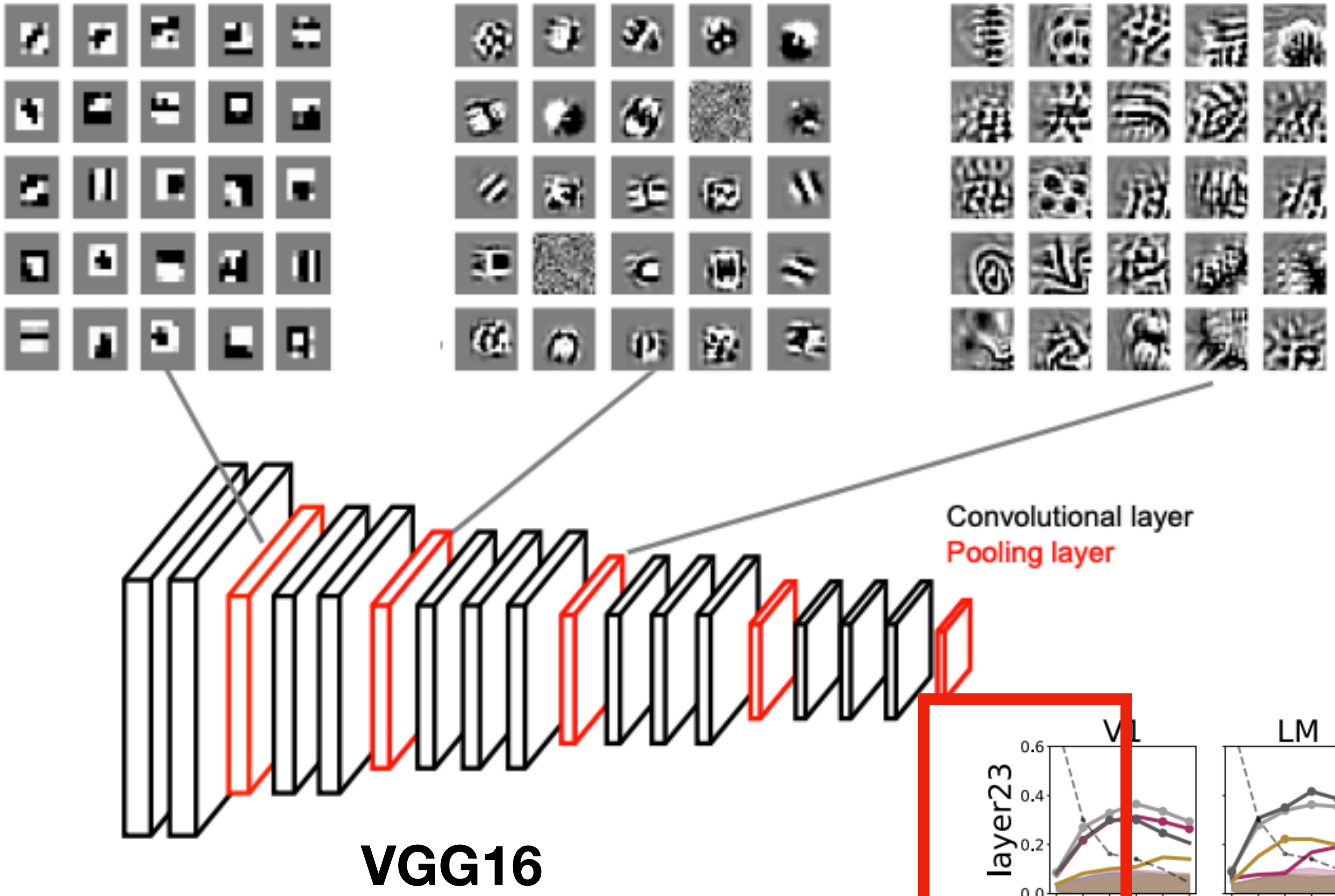


But mouse visual cortex is anatomically shallow!

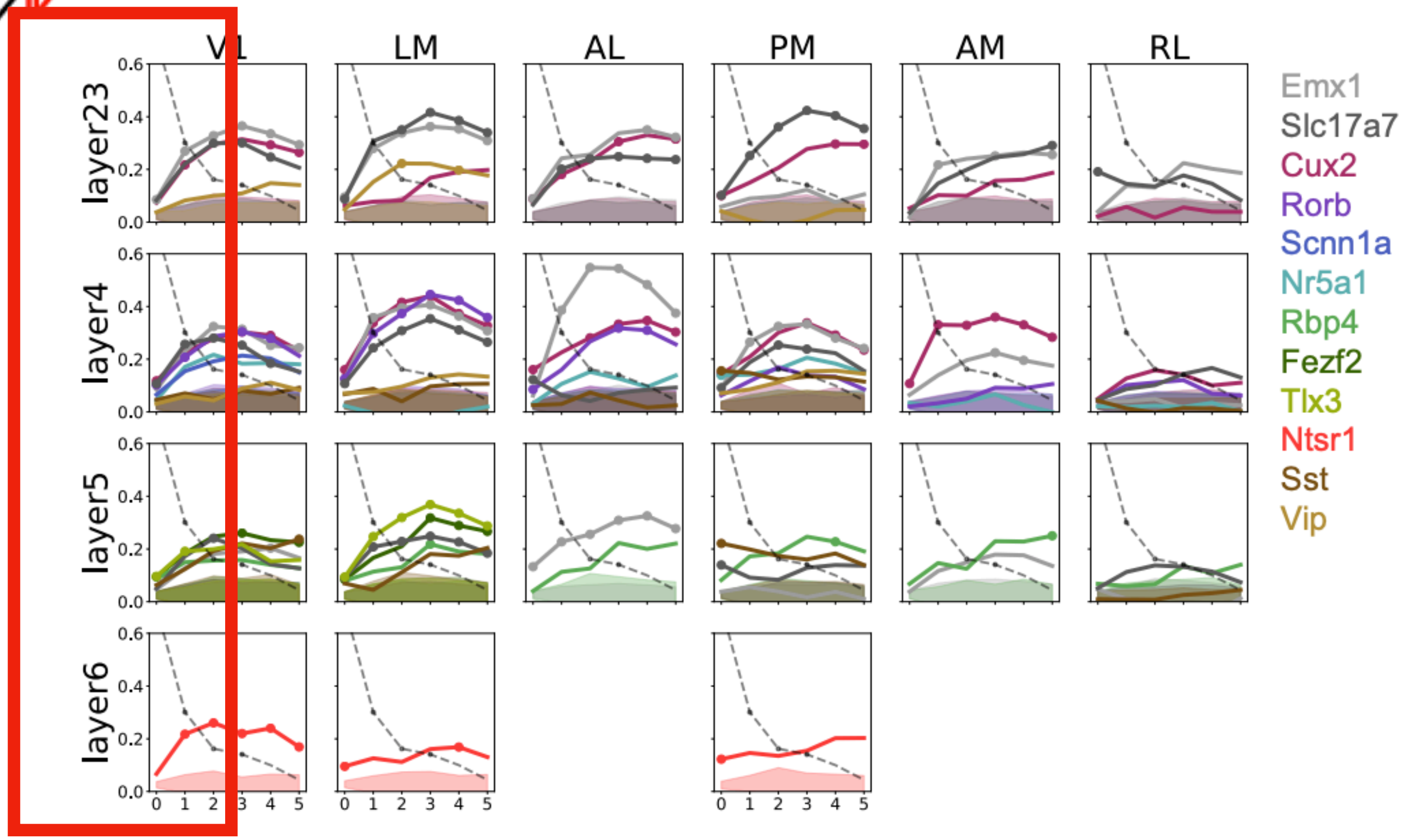


Suggests roughly 3-4 stages of processing based on reliability timing

Deep models are also a poor match to responses

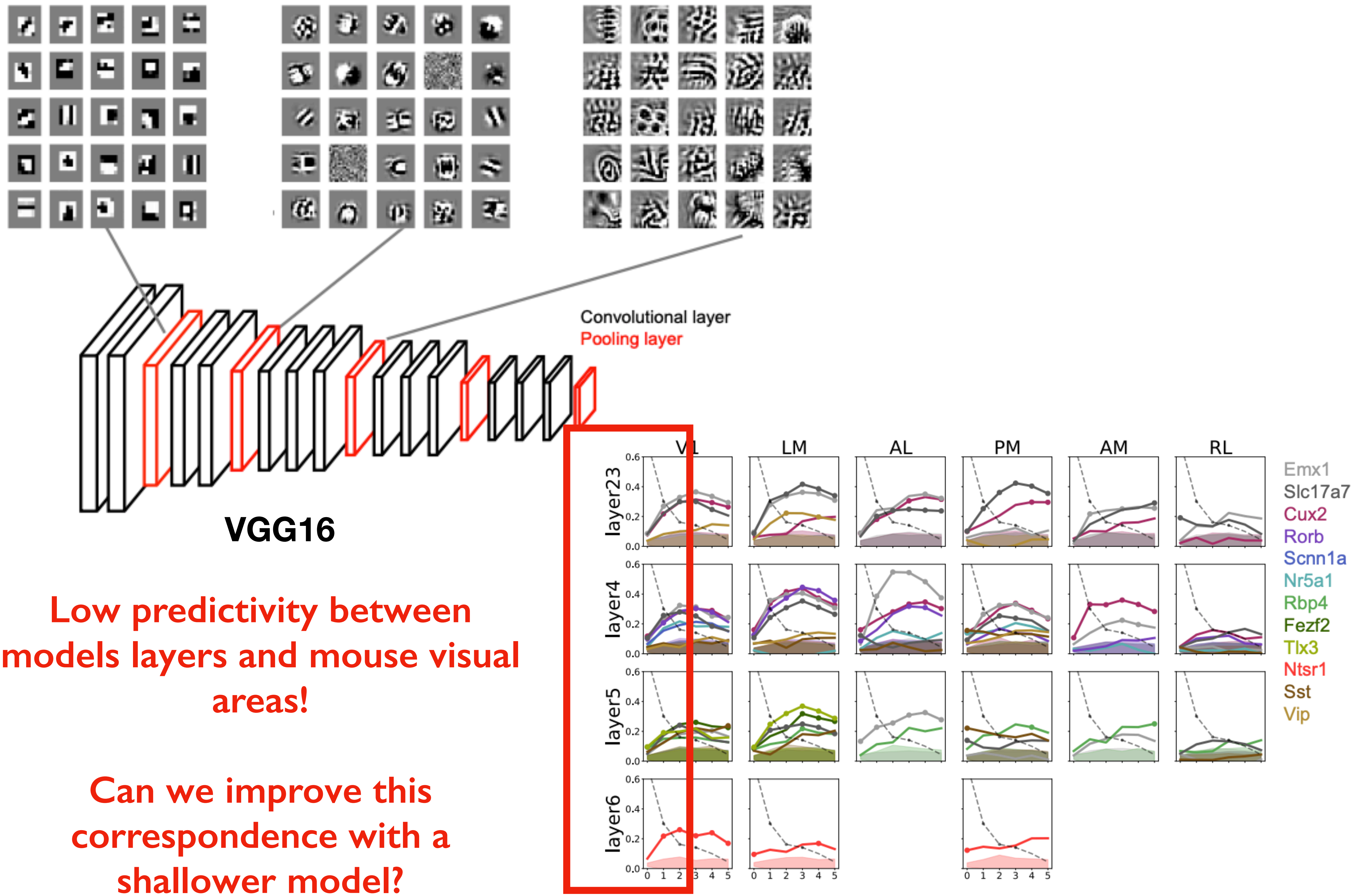


Low predictivity between models layers and mouse visual areas!



de Vries et al., 2020

Deep models are also a poor match to responses



Low predictivity between models layers and mouse visual areas!

Can we improve this correspondence with a shallower model?

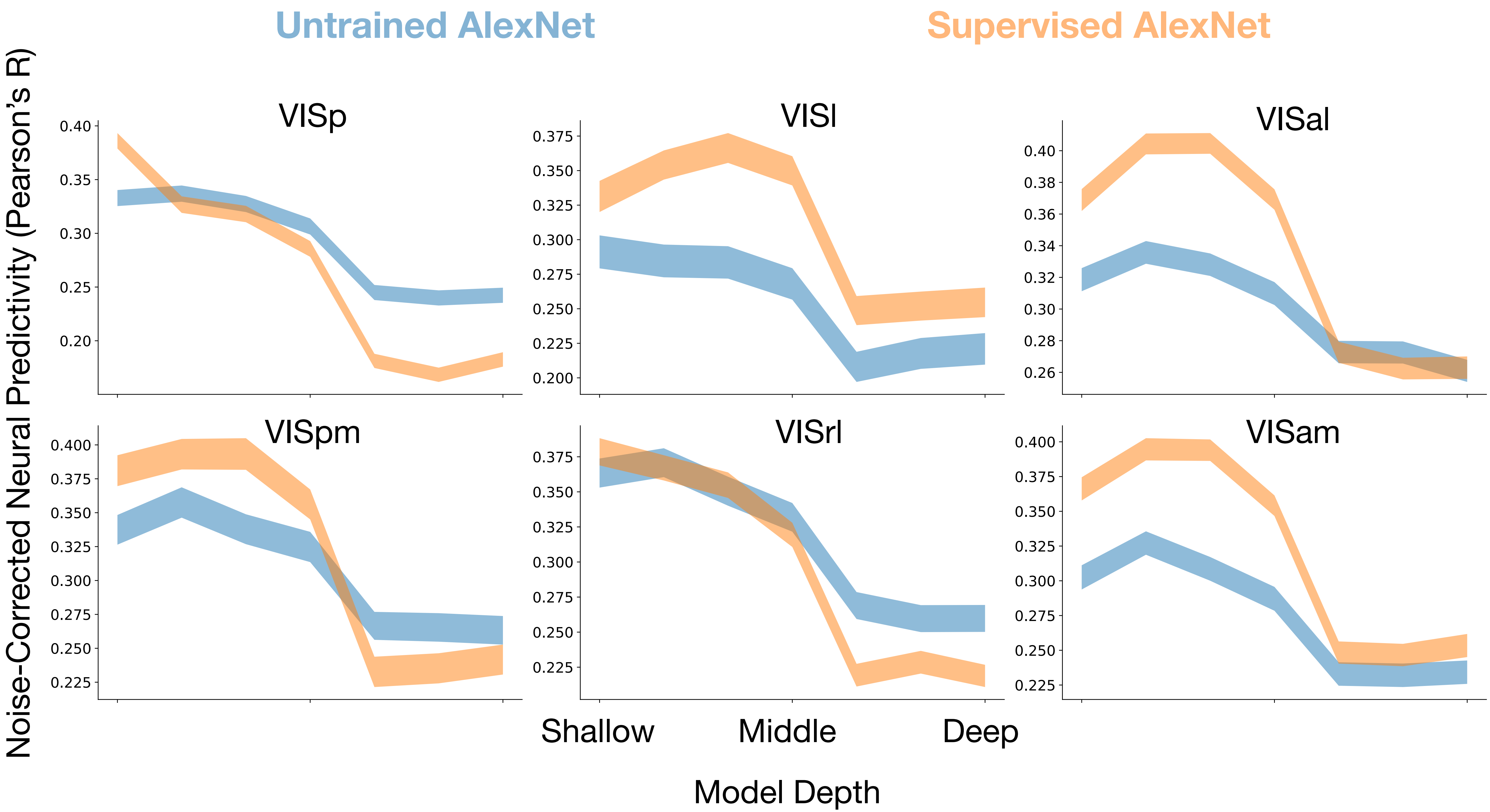
de Vries et al., 2020

Even shallower models?

Look at neural predictivity across model layer

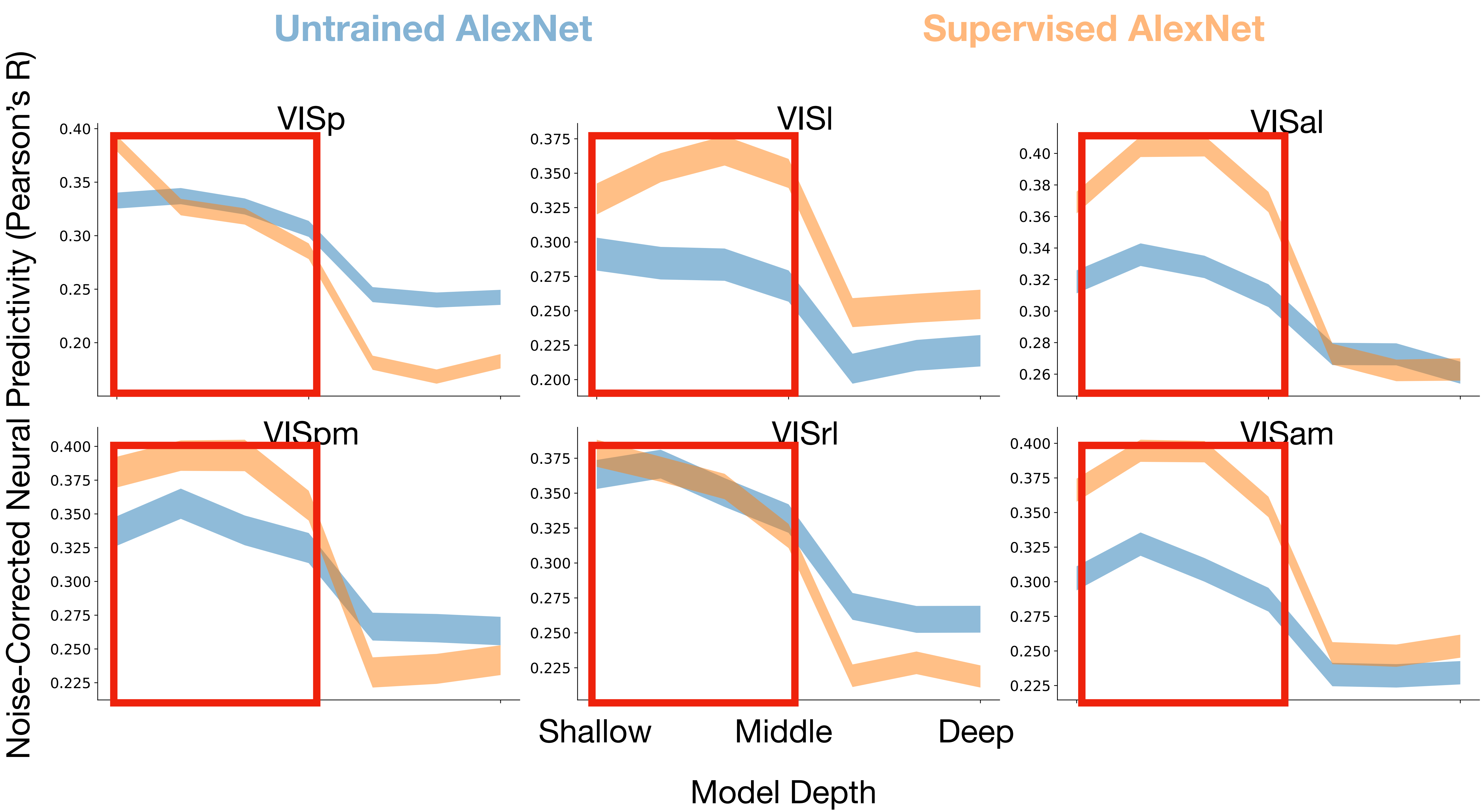
Even shallower models?

Look at neural predictivity across model layer

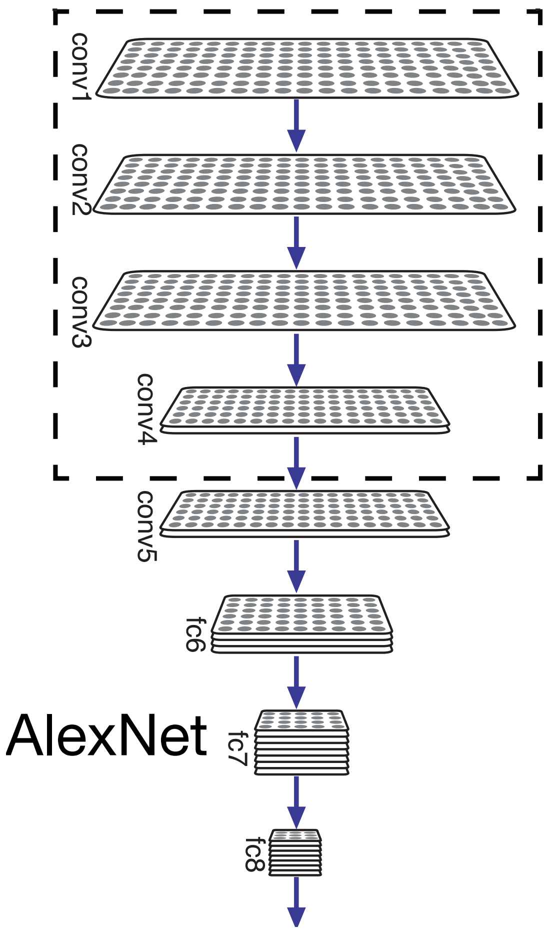


Even shallower models?

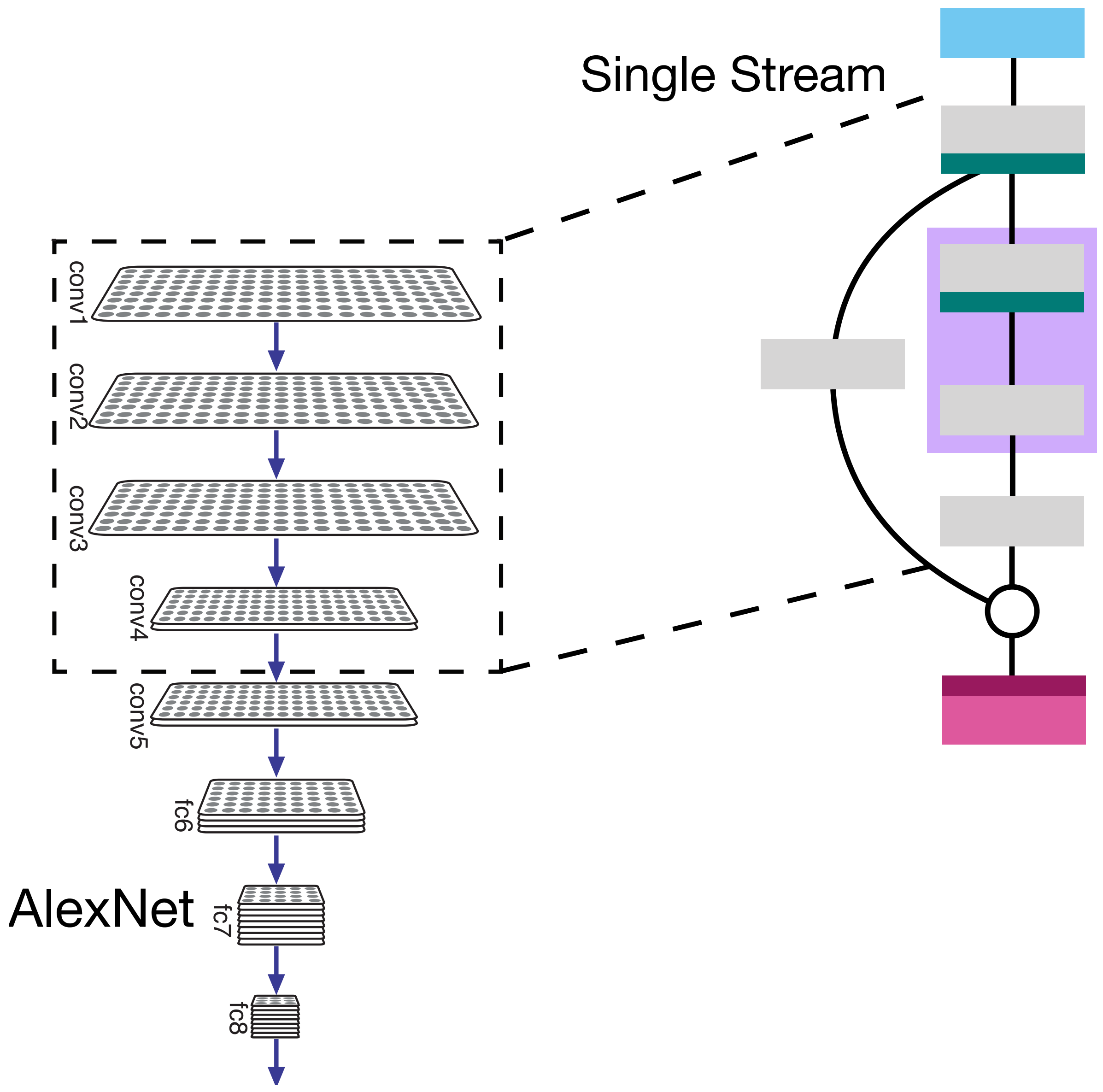
Look at neural predictivity across model layer



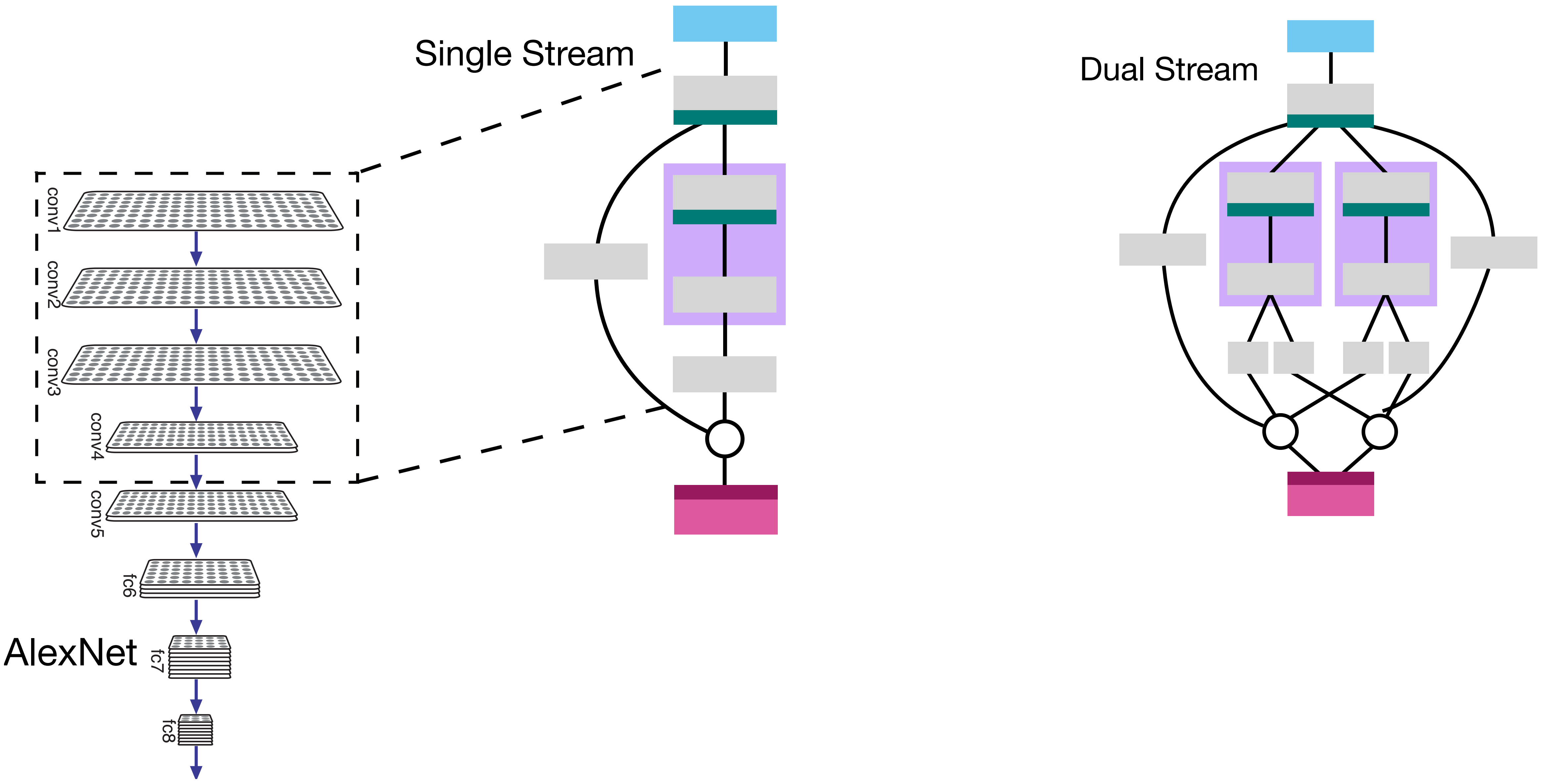
Even shallower models?



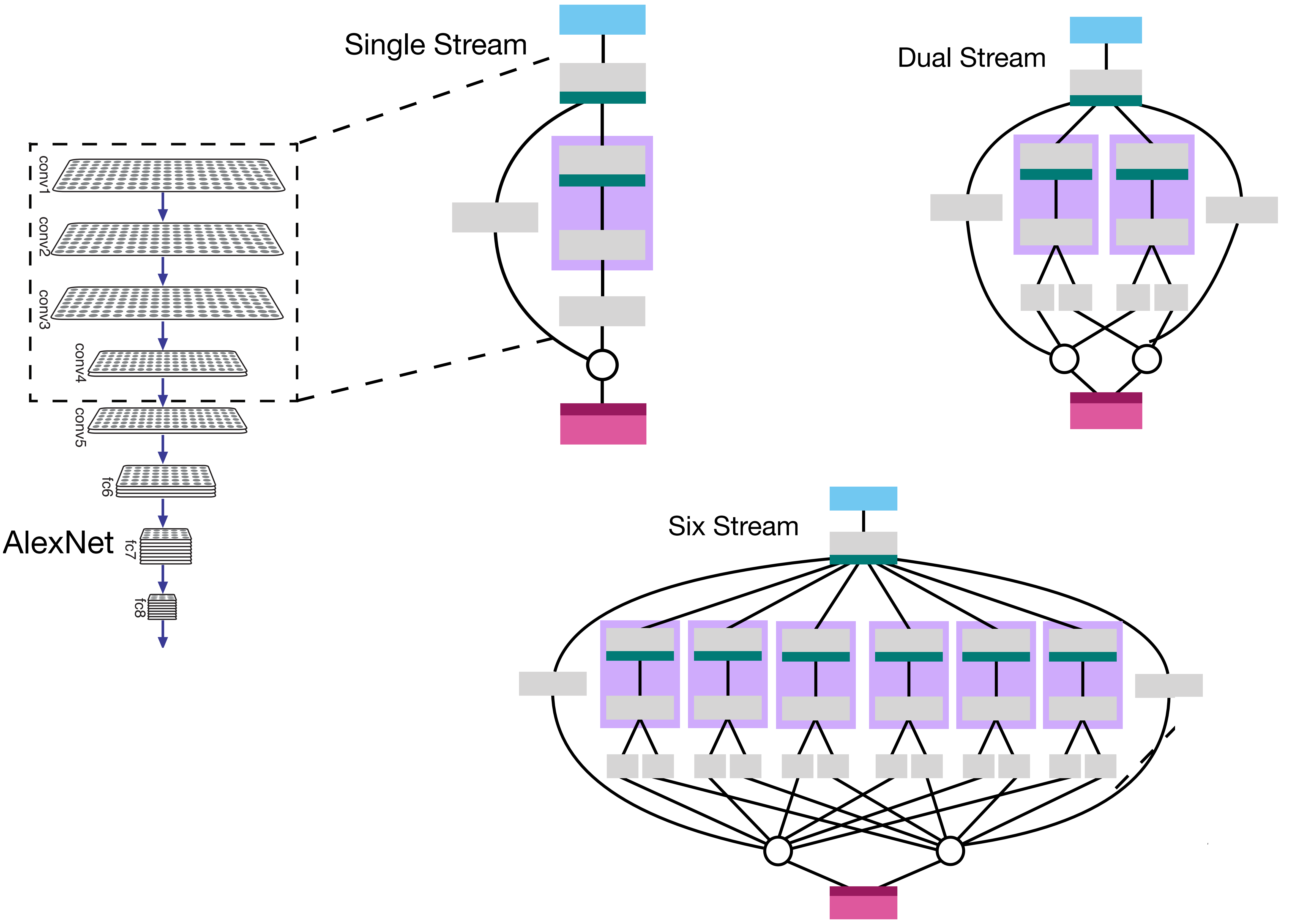
Even shallower models?



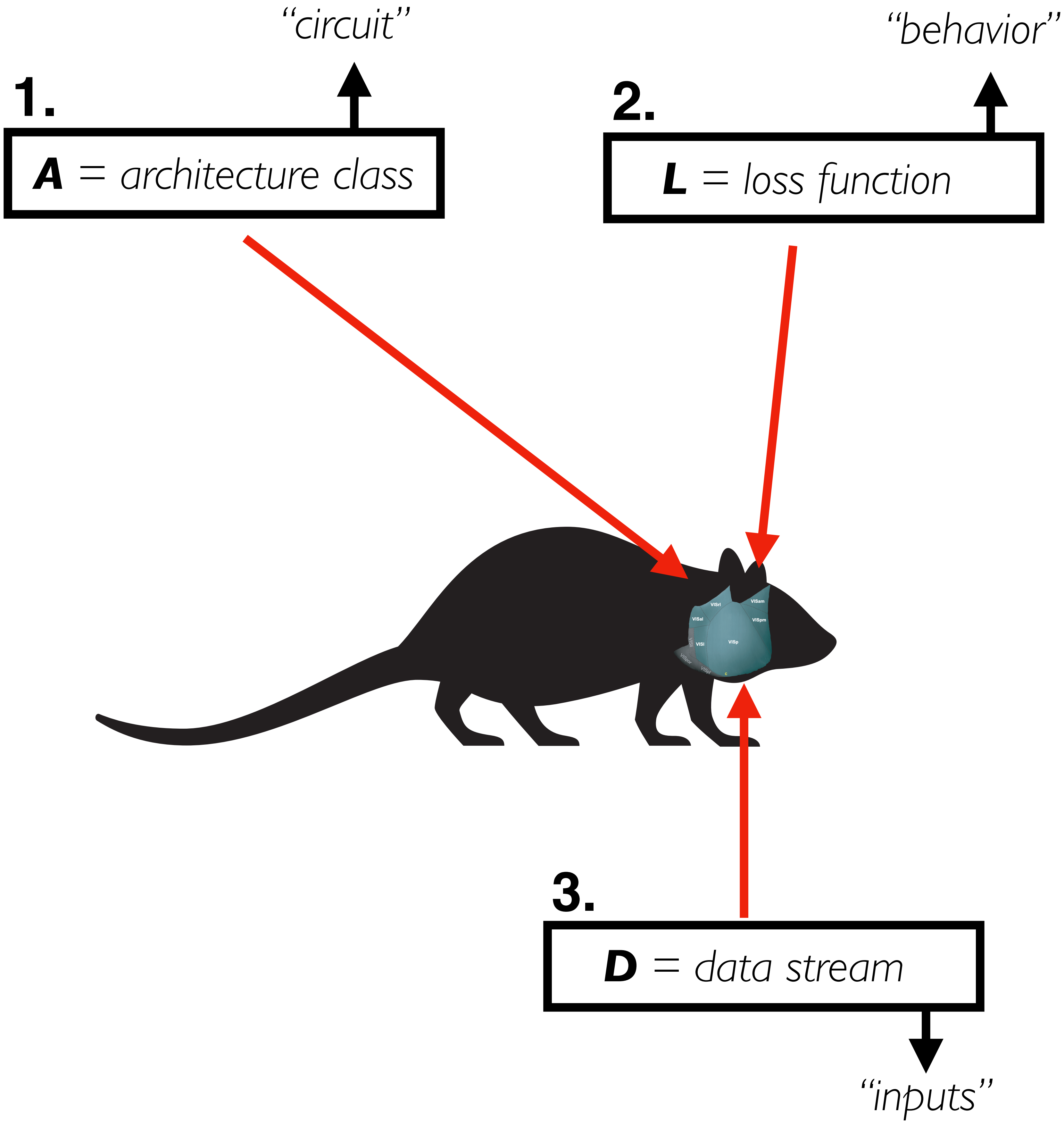
Even shallower models?



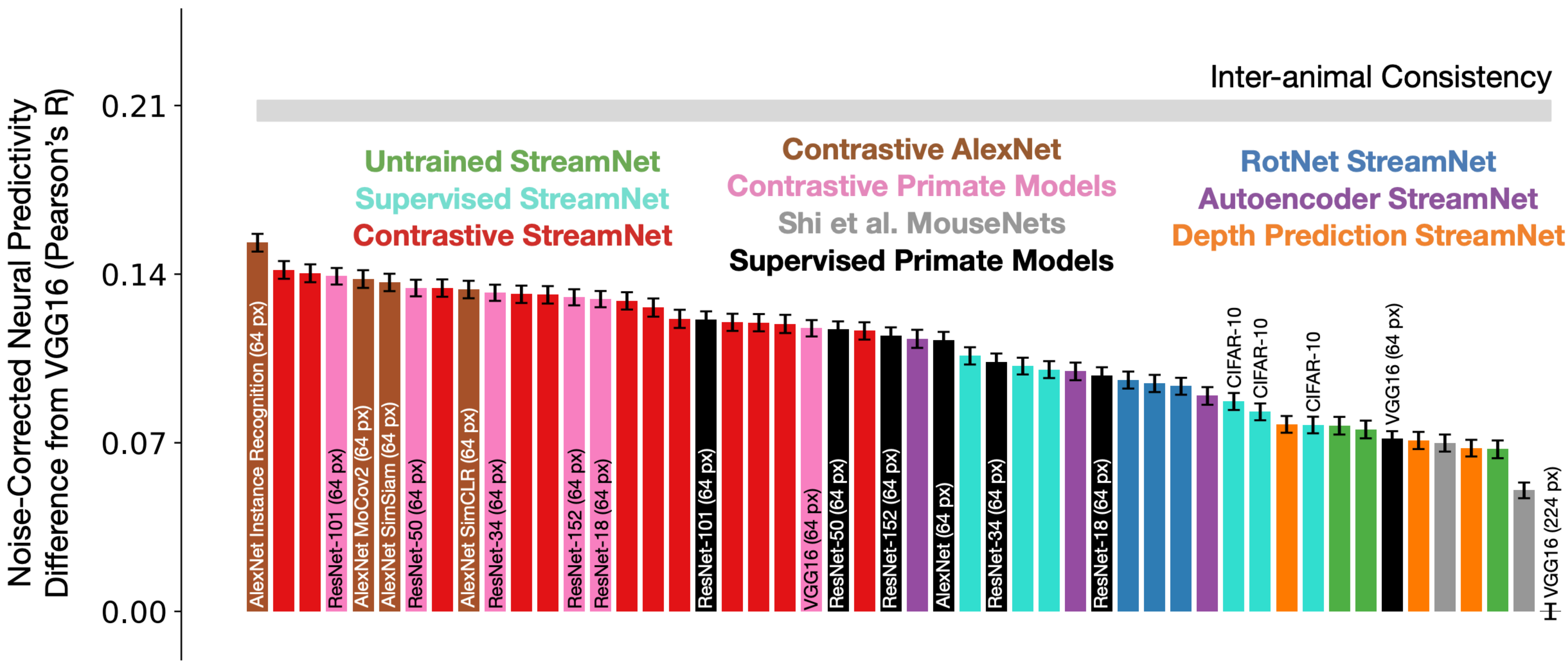
Even shallower models?



Putting it all together: Circuit, Behavior, Input

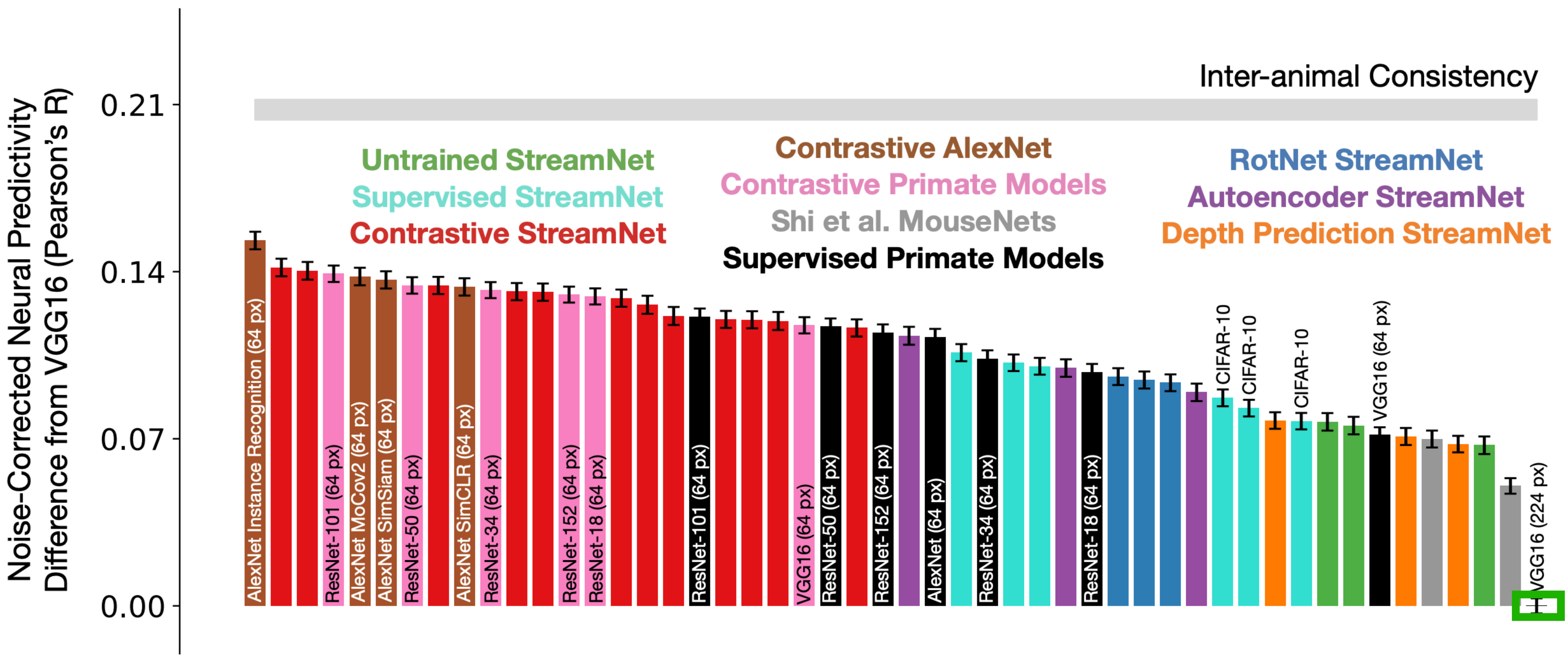


Substantially improving neural response predictivity of models of mouse visual cortex



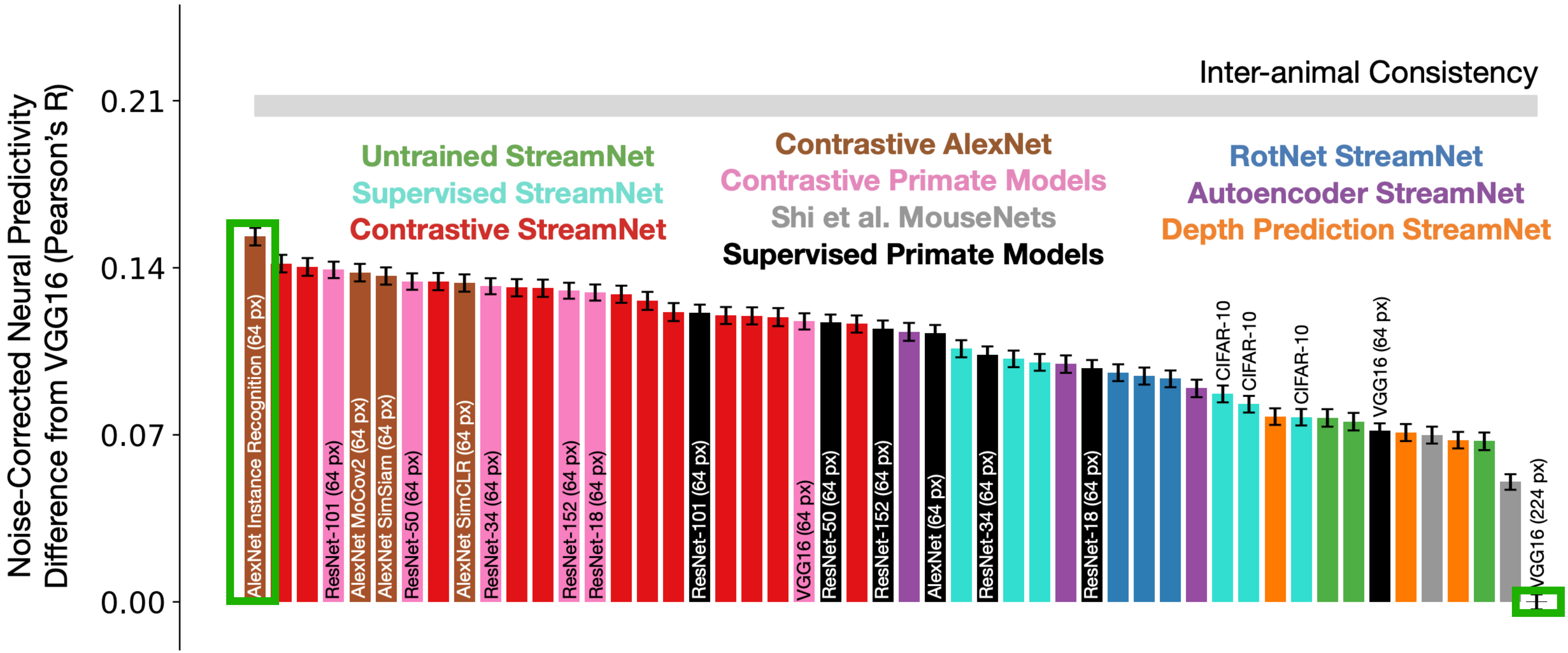
Substantially improving neural response predictivity of models of mouse visual cortex

Previous baseline (deep supervised primate model)

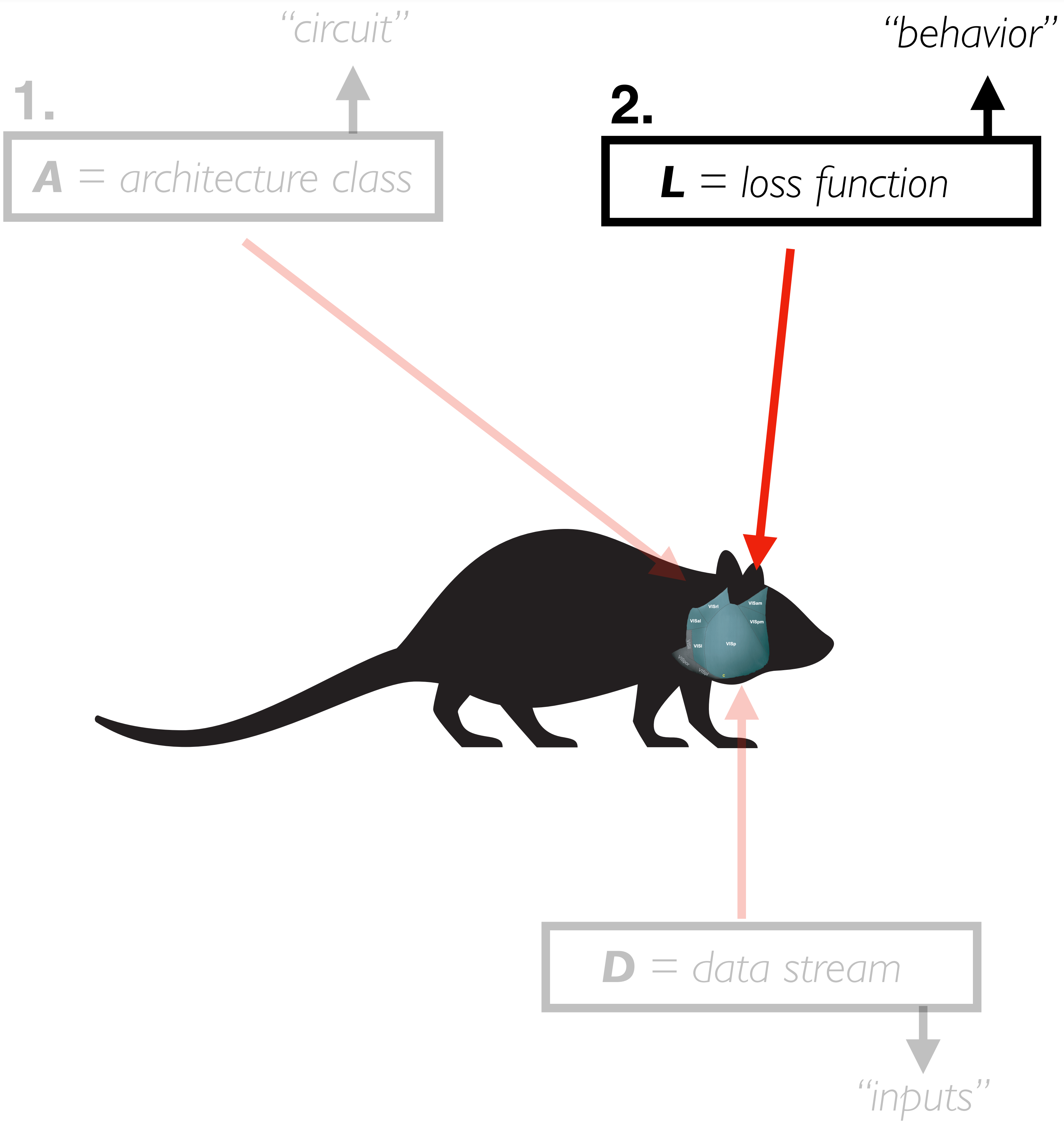


Substantially improving neural response predictivity of models of mouse visual cortex

Previous baseline (deep supervised primate model)
Shallow low resolution unsupervised model

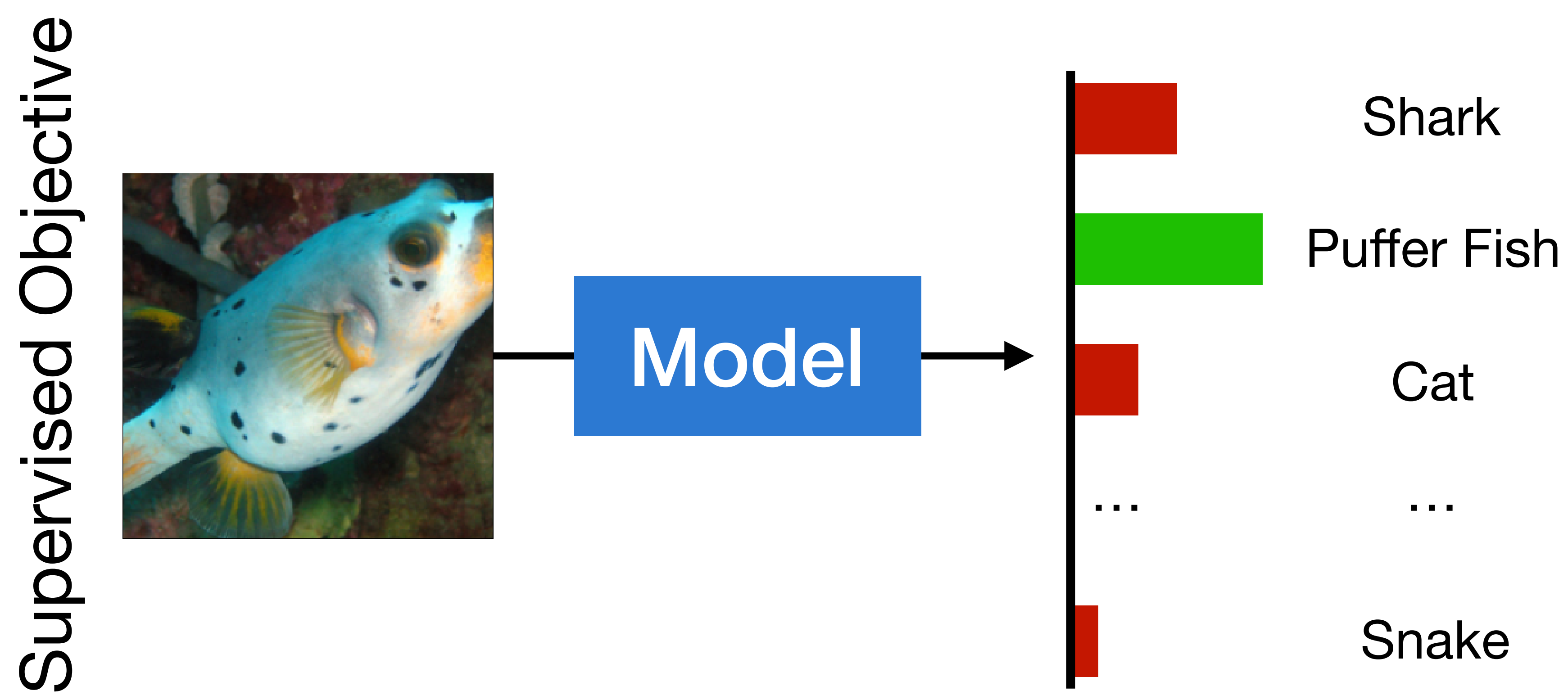


Distilling Constraints: Behavioral Goals



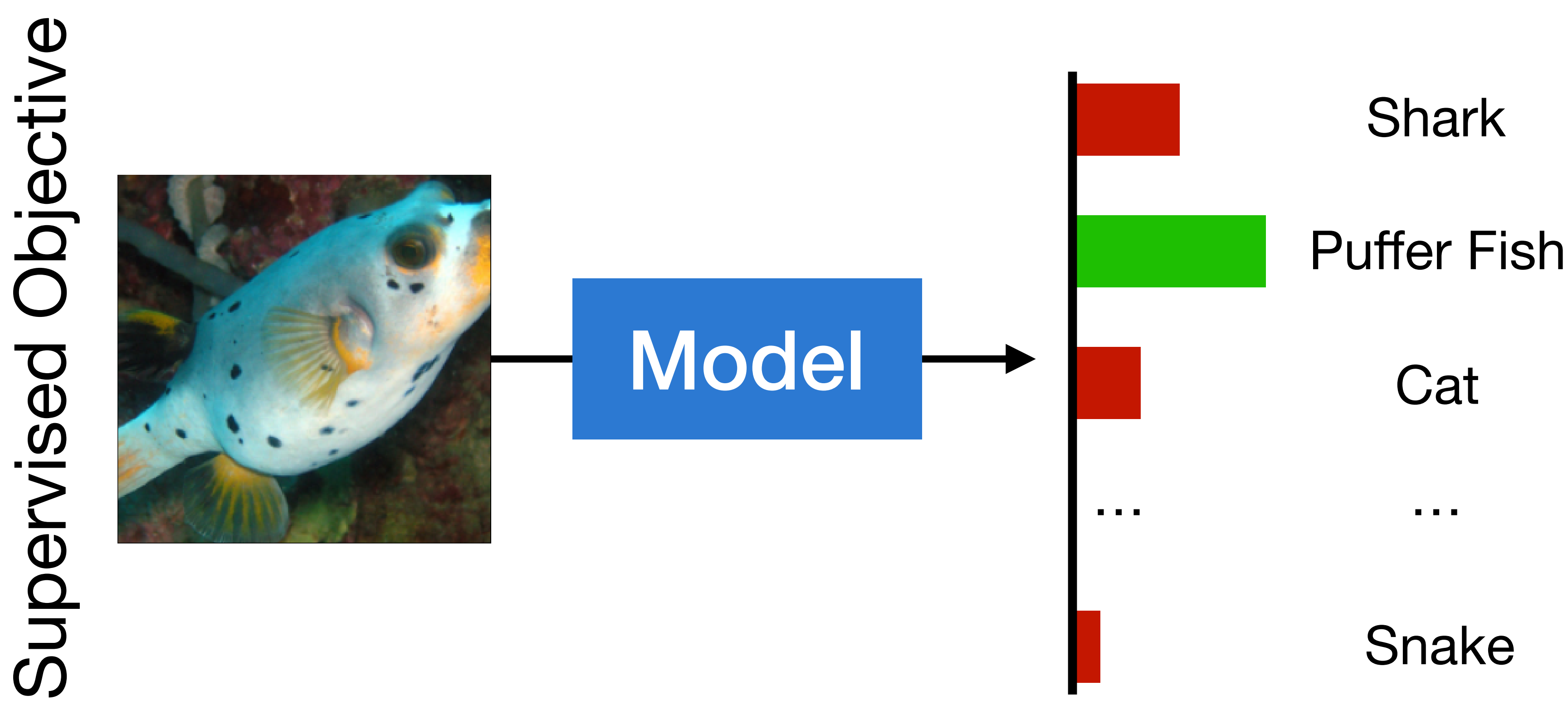
Distilling Constraints: Behavioral Goals

Typical setting: supervision with (1000) category labels



Distilling Constraints: Behavioral Goals

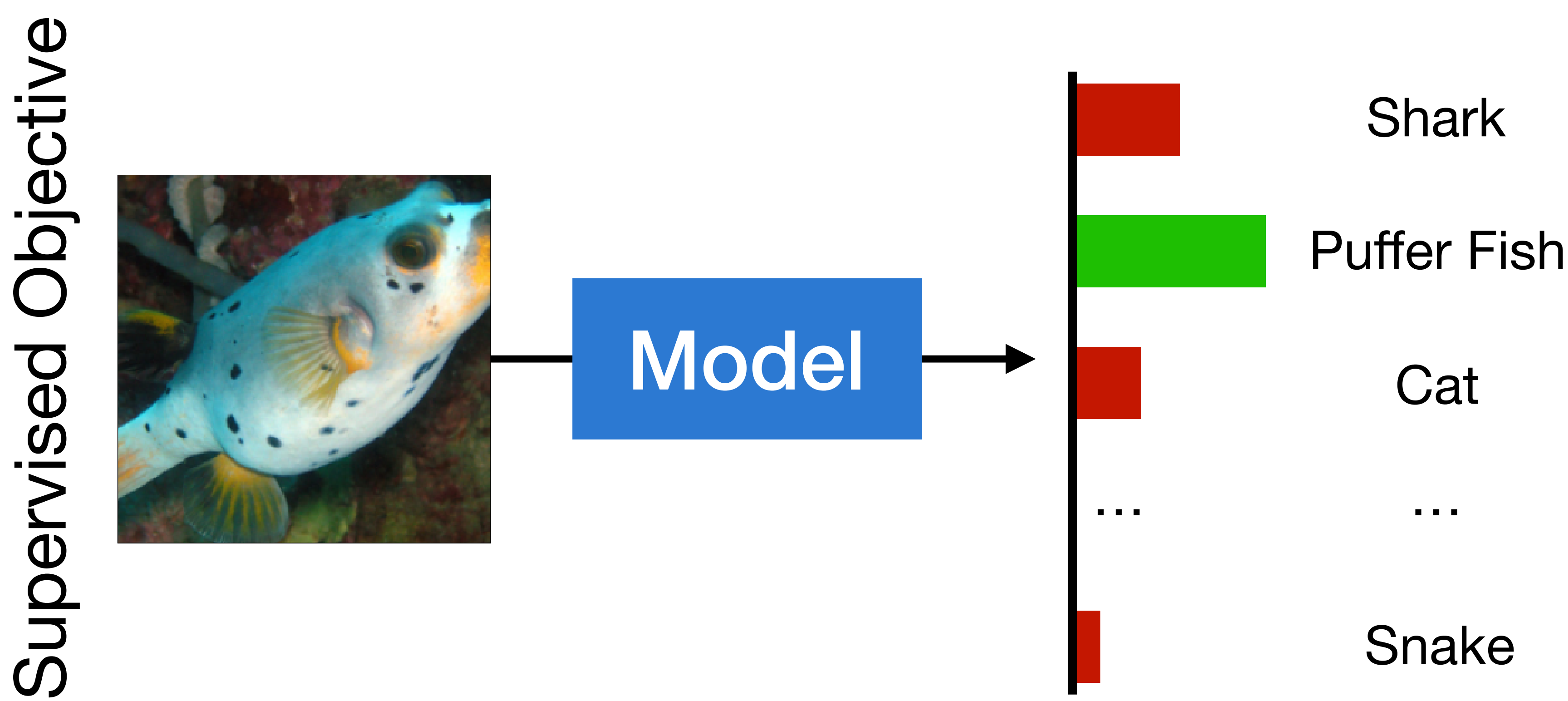
Typical setting: supervision with (1000) category labels
…but is very "unnatural" for mice!



Distilling Constraints: Behavioral Goals

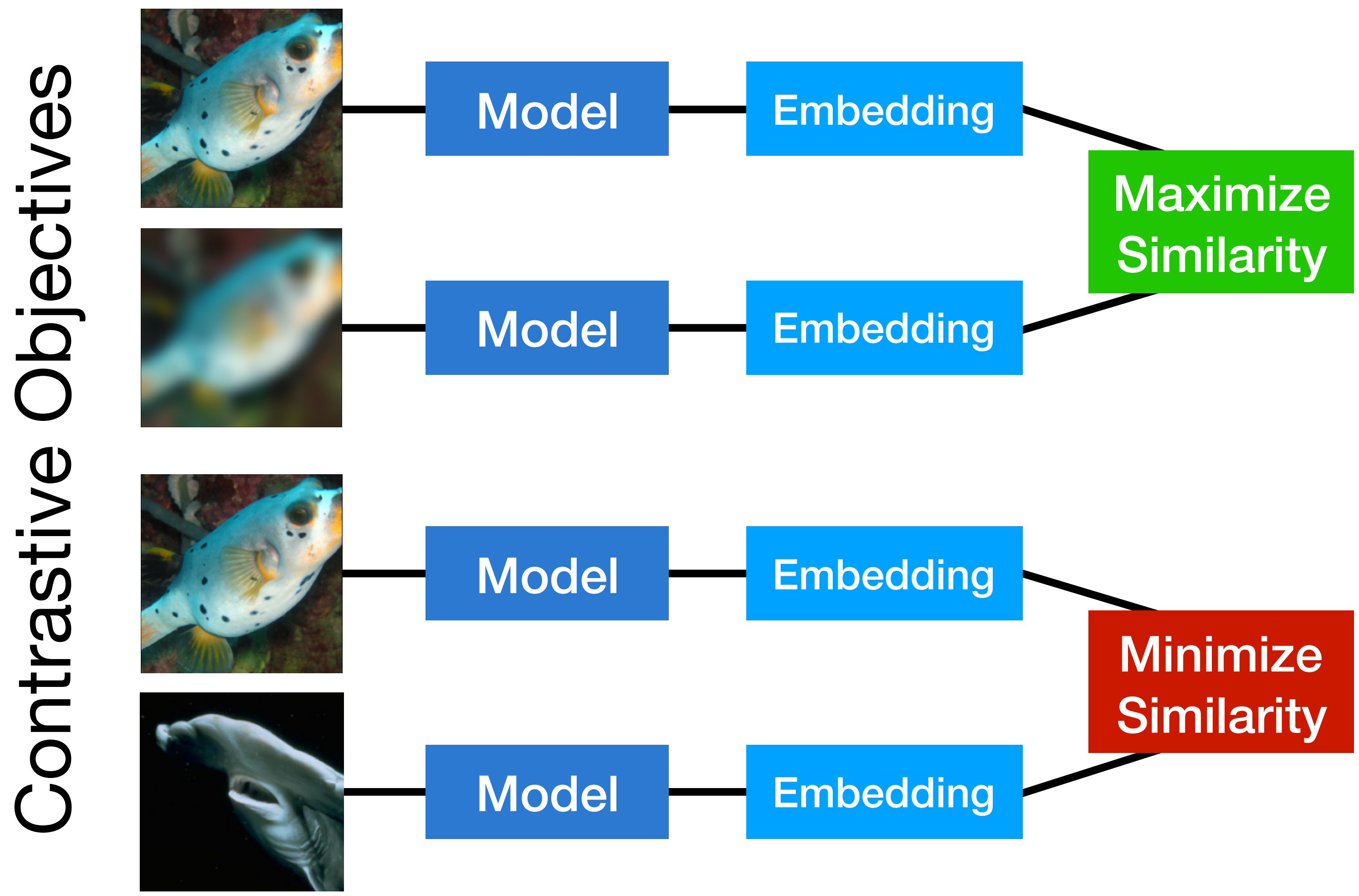
Typical setting: supervision with (1000) category labels
…but is very "unnatural" for mice!

Both the type and number of categories is unrealistic for mice

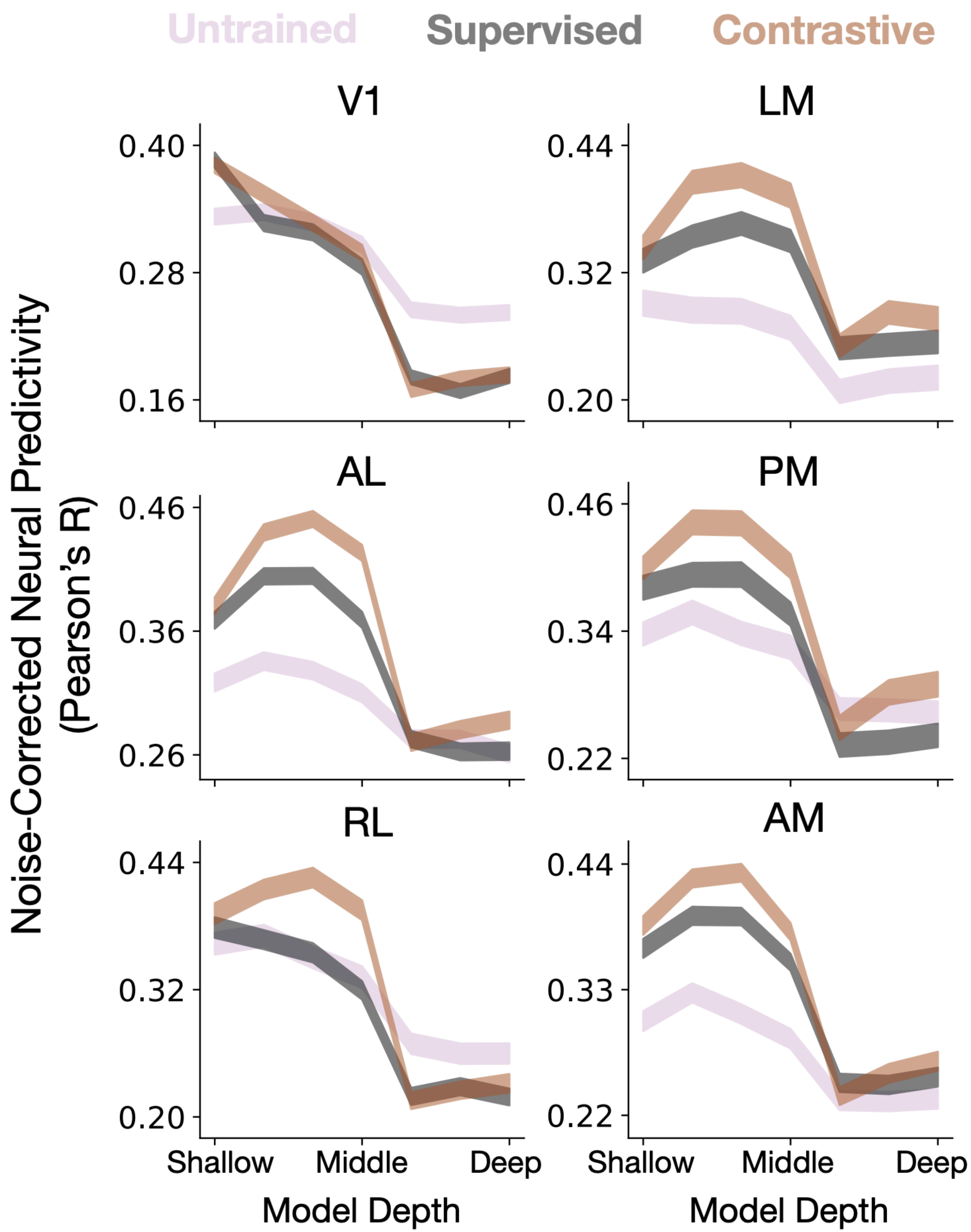


Distilling Constraints: Behavioral Goals

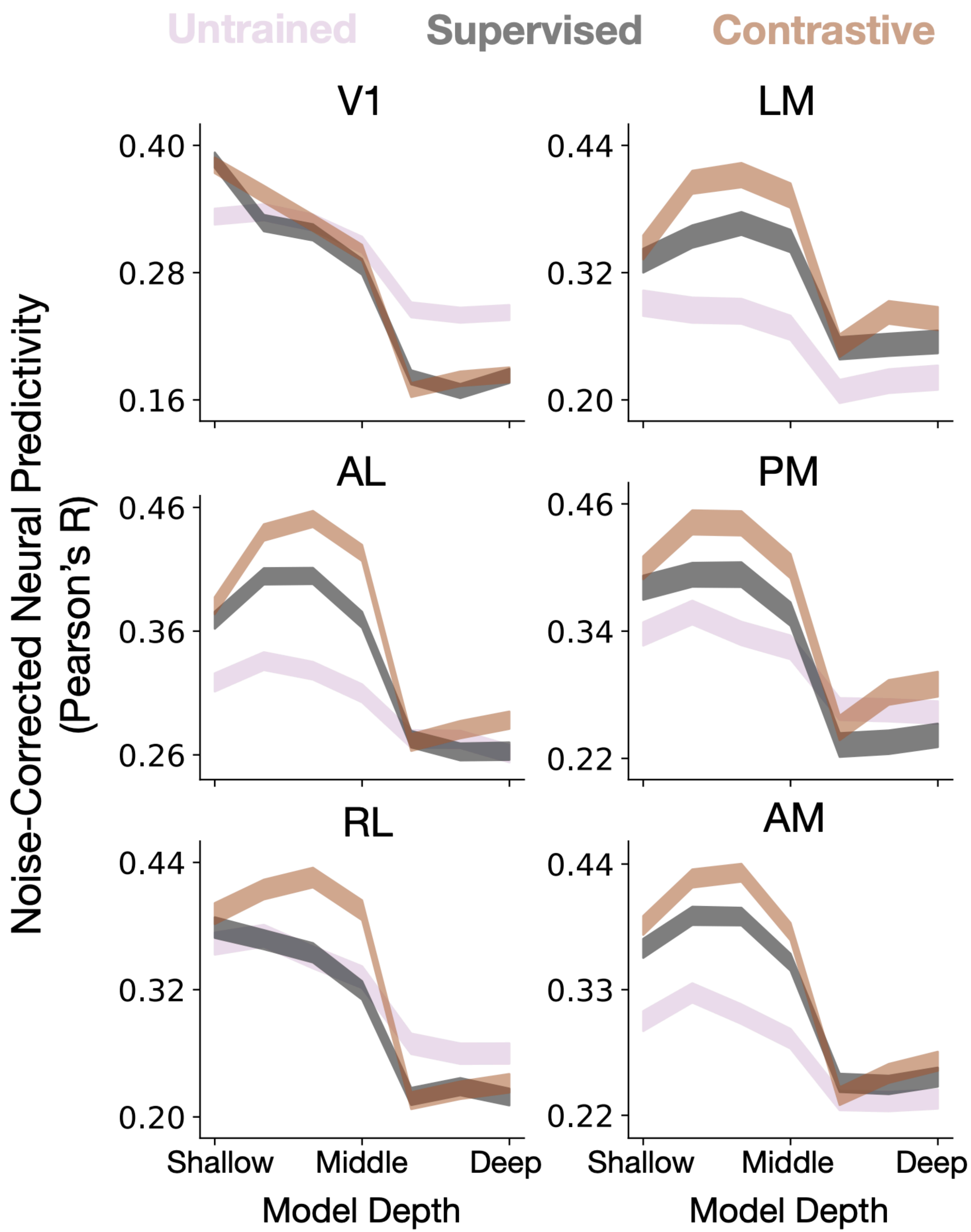
Consider unsupervised objectives, most notably "contrastive" objectives



Distilling Constraints: Behavioral Goals

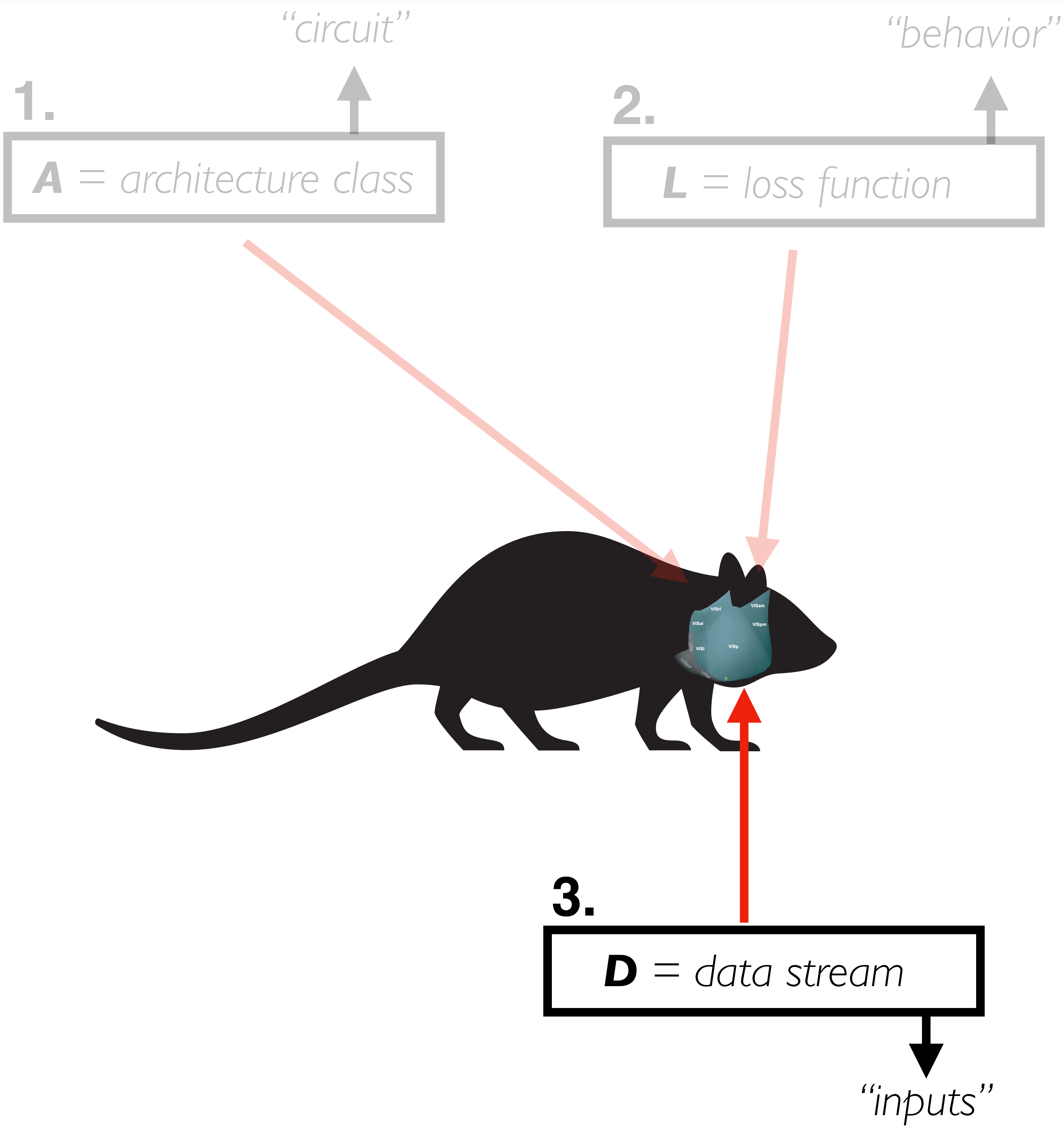


Distilling Constraints: Behavioral Goals



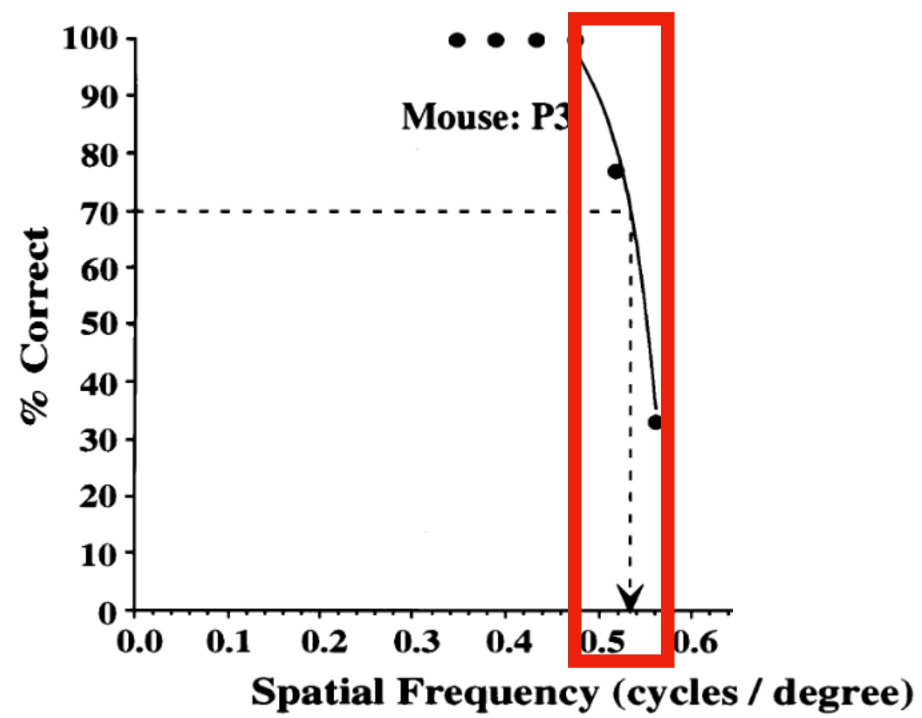
Using an unsupervised, contrastive objective function improves neural predictivity (architecture & data stream fixed)

Distilling Constraints: Inputs

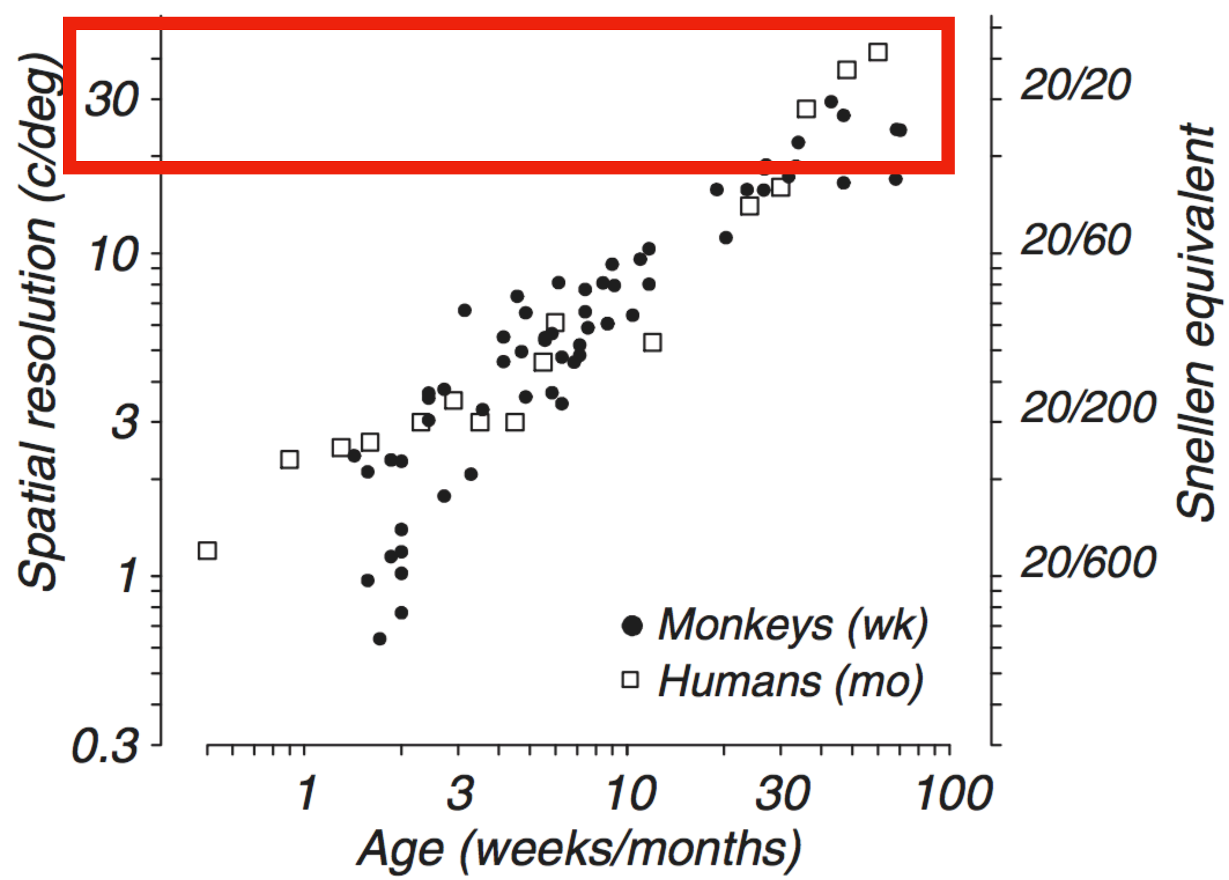


Distilling Constraints: Inputs

Mice

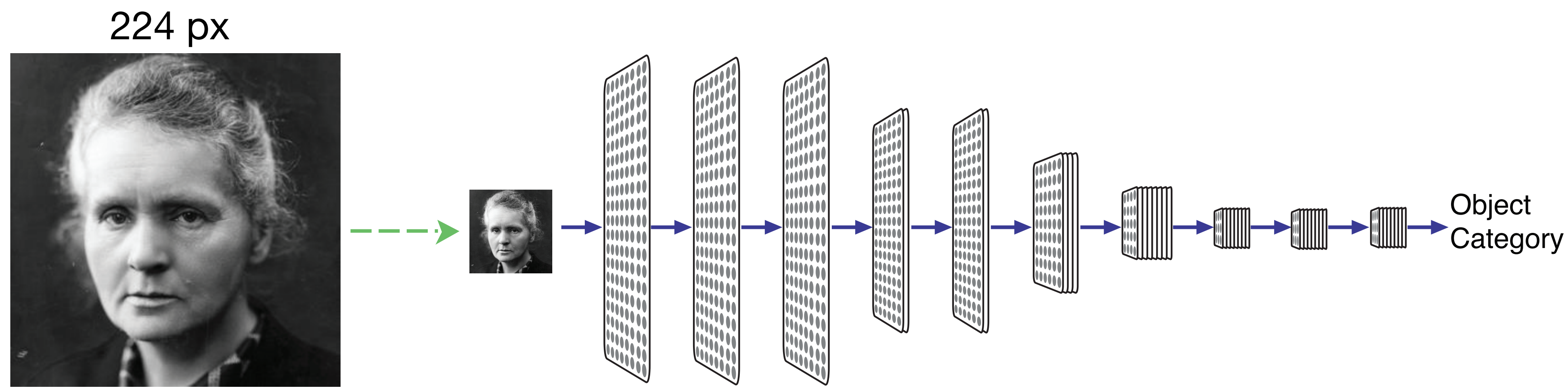


Primates

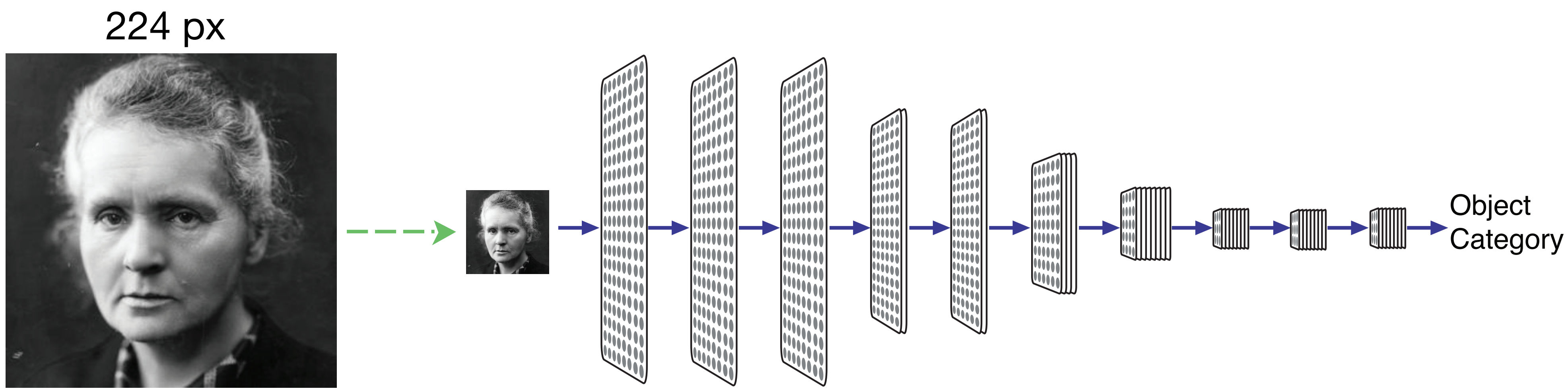


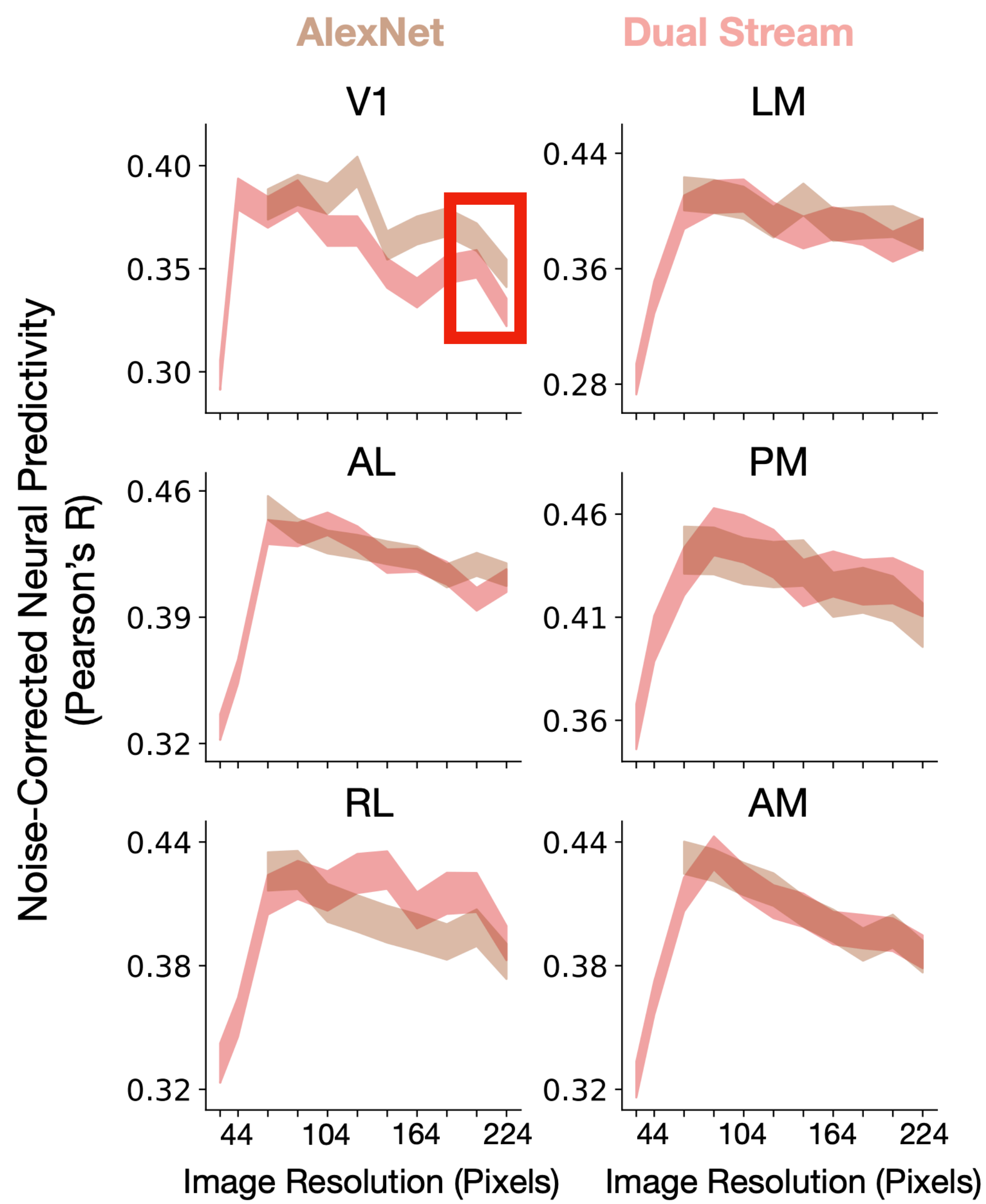
Prusky et al., 2000; Kiorpes, 2019

Distilling Constraints: Inputs

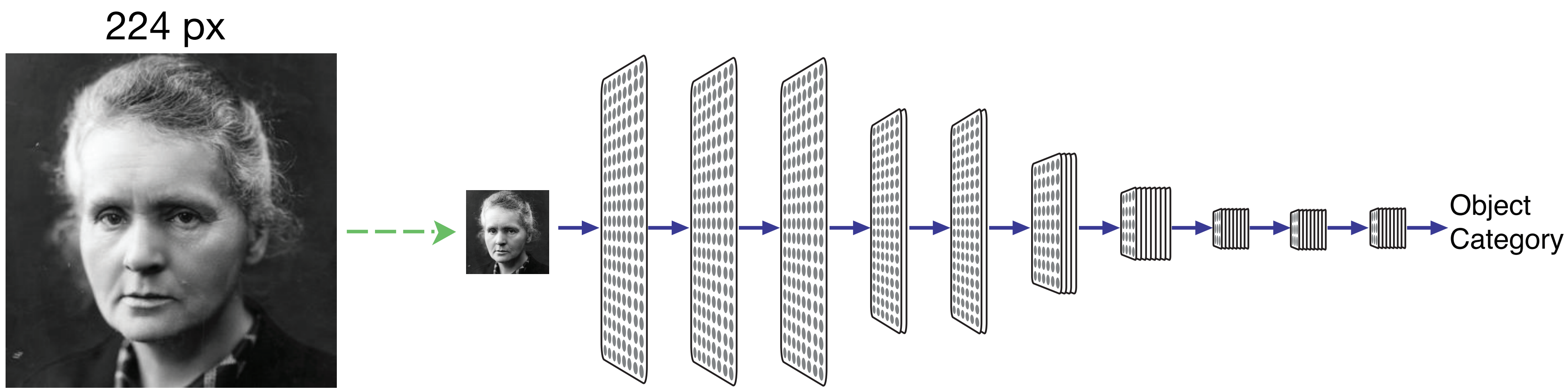


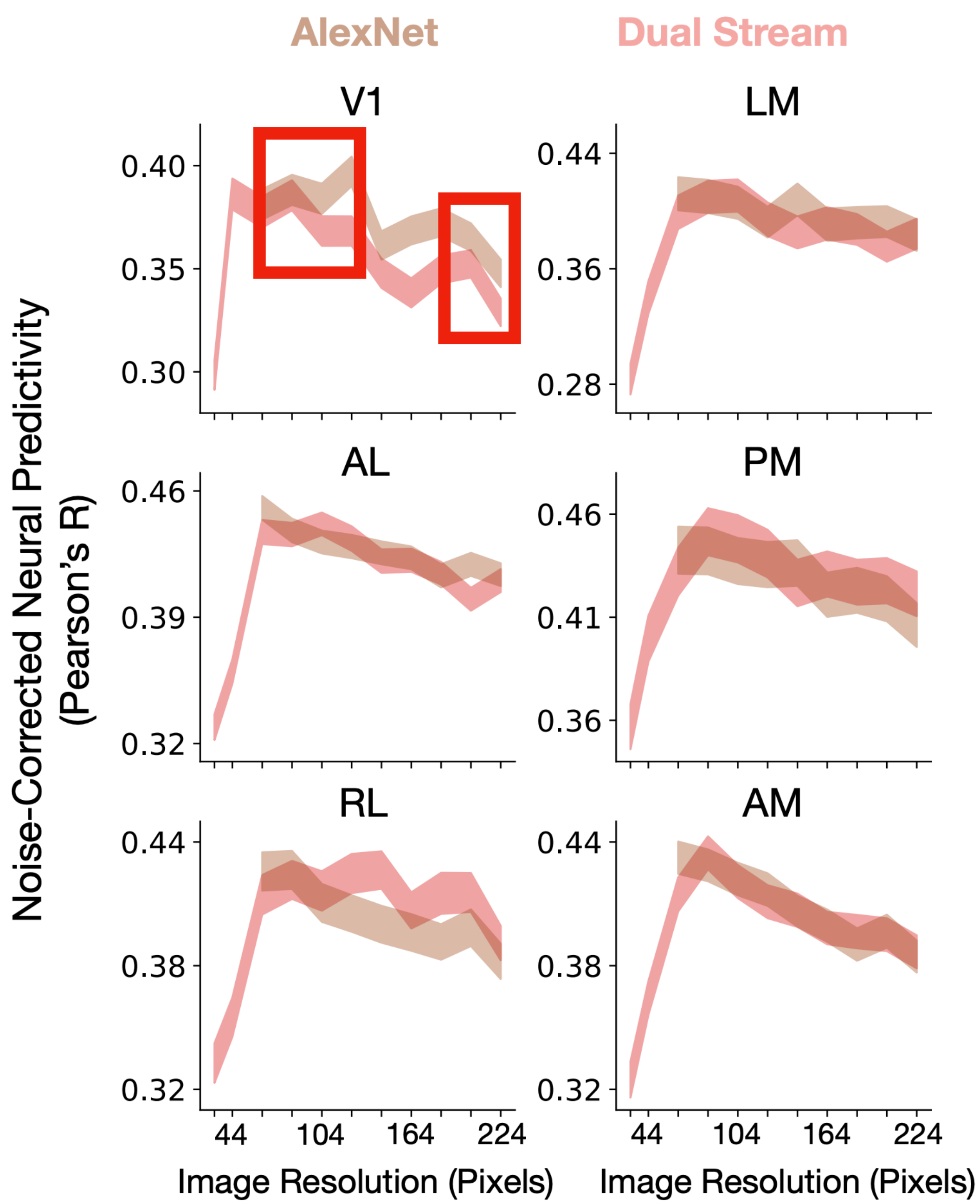
Distilling Constraints: Inputs



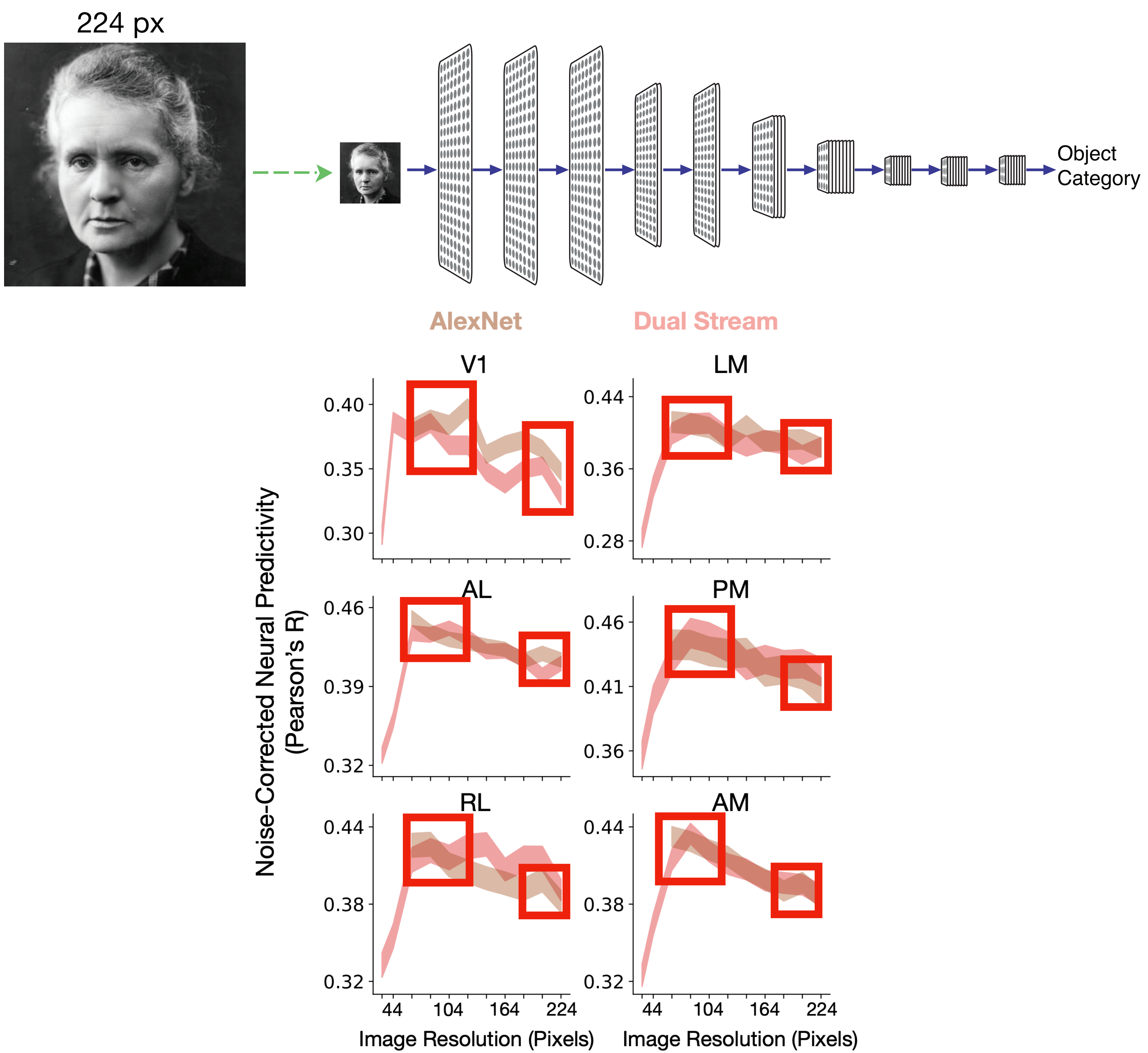


Distilling Constraints: Inputs

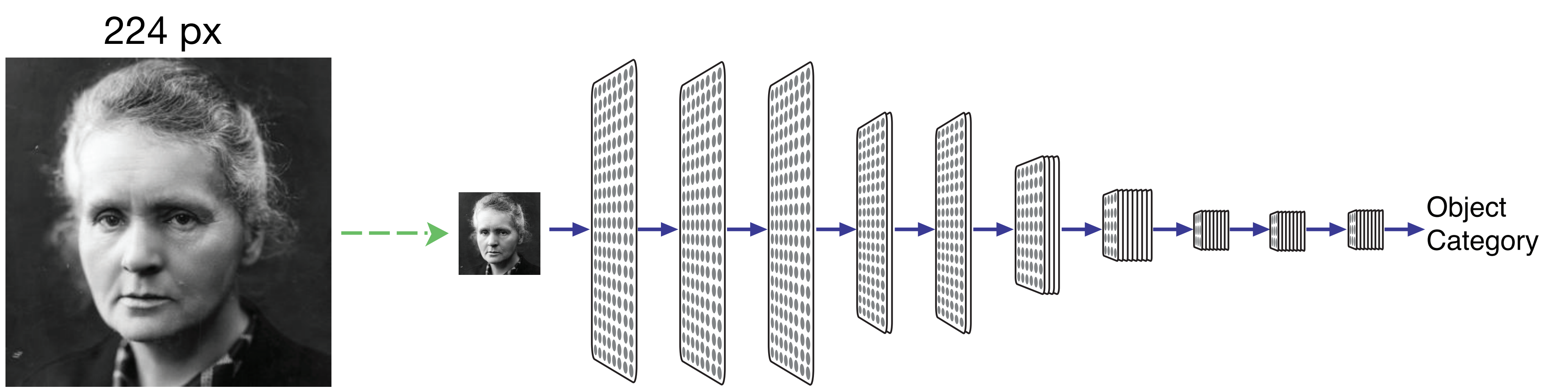


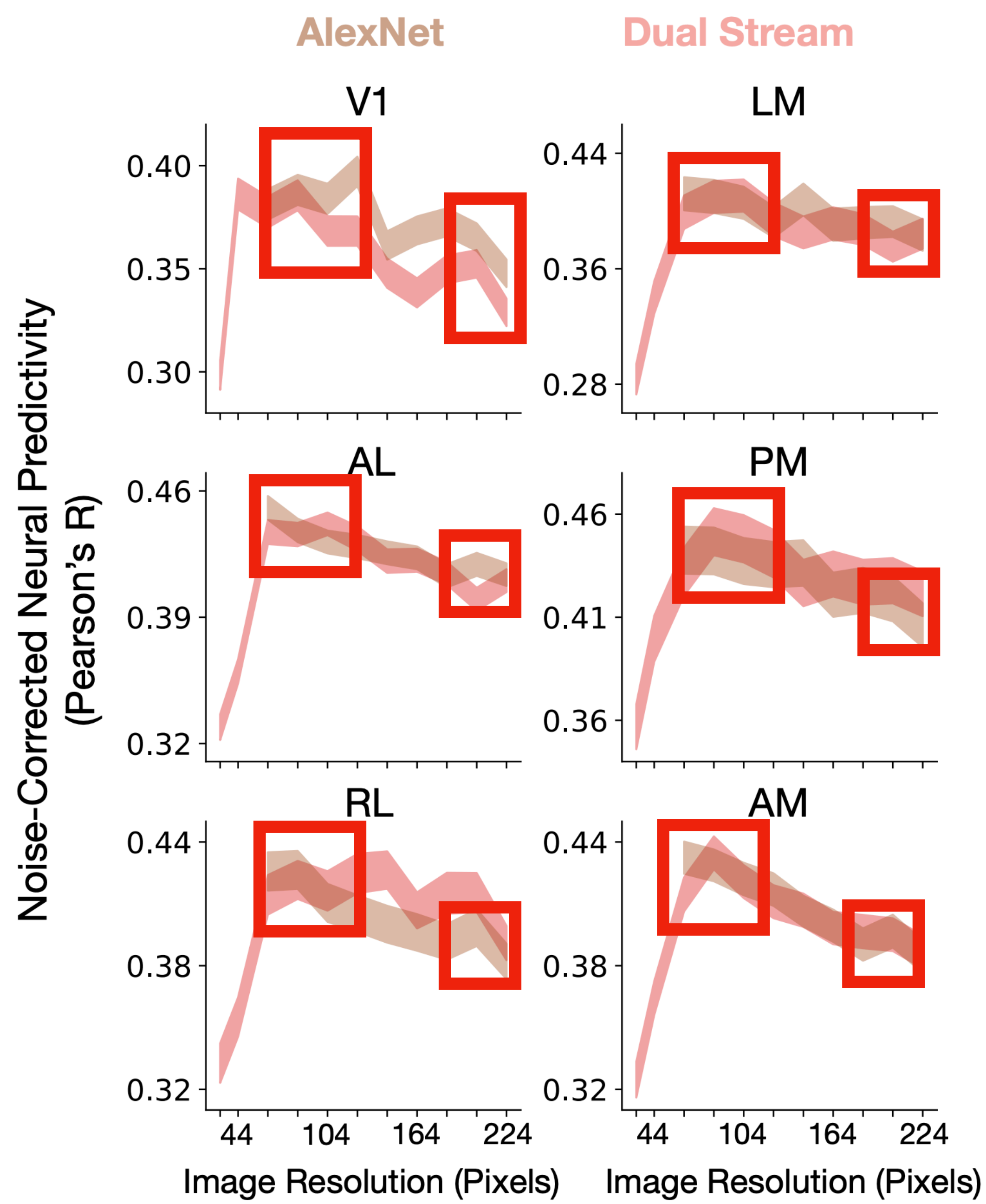


Distilling Constraints: Inputs



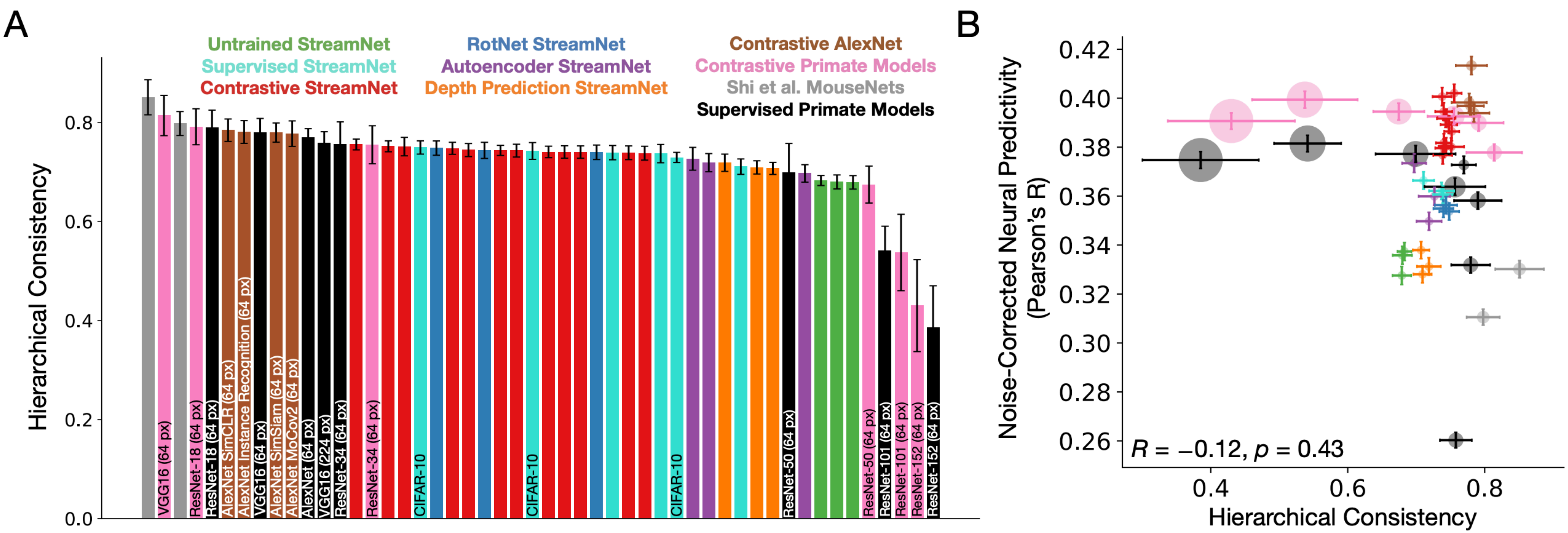
Distilling Constraints: Inputs





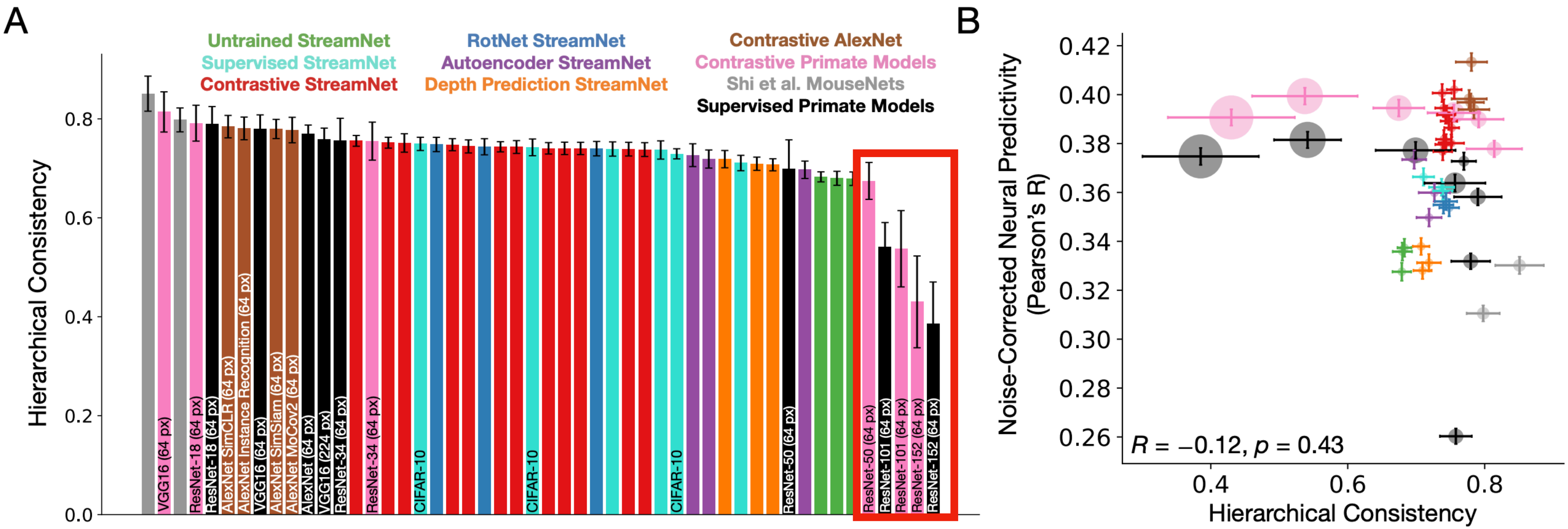
Reducing image resolution during training improves neural predictivity across visual areas

Quantifying hierarchical consistency of models with functional hierarchy of mouse visual cortex



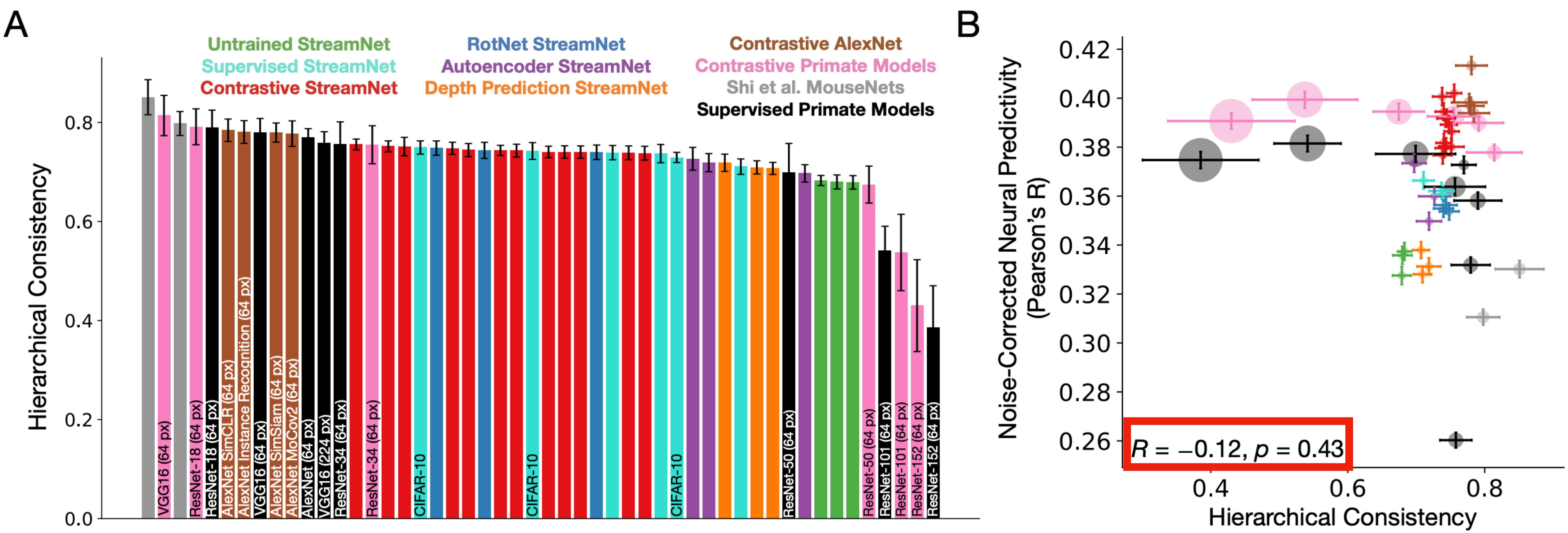
Quantifying hierarchical consistency of models with functional hierarchy of mouse visual cortex

Very deep models do not match this metric well (supervised or unsupervised)



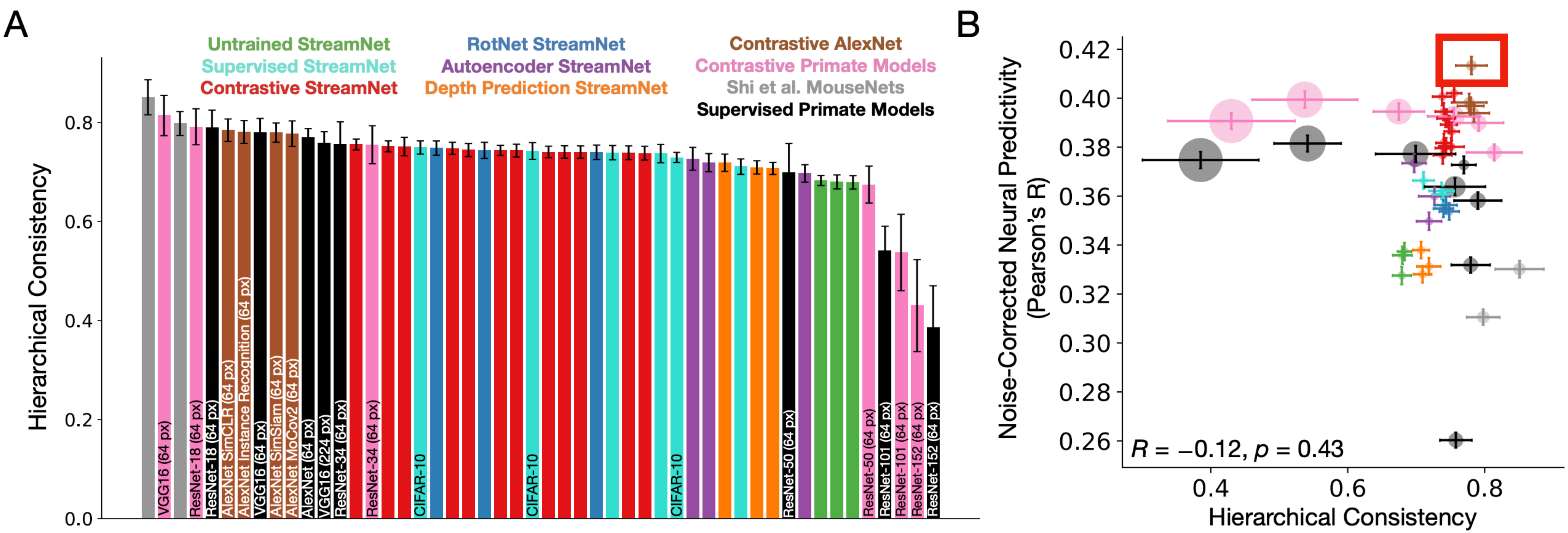
Quantifying hierarchical consistency of models with functional hierarchy of mouse visual cortex

The two metrics capture difference aspects of the variance



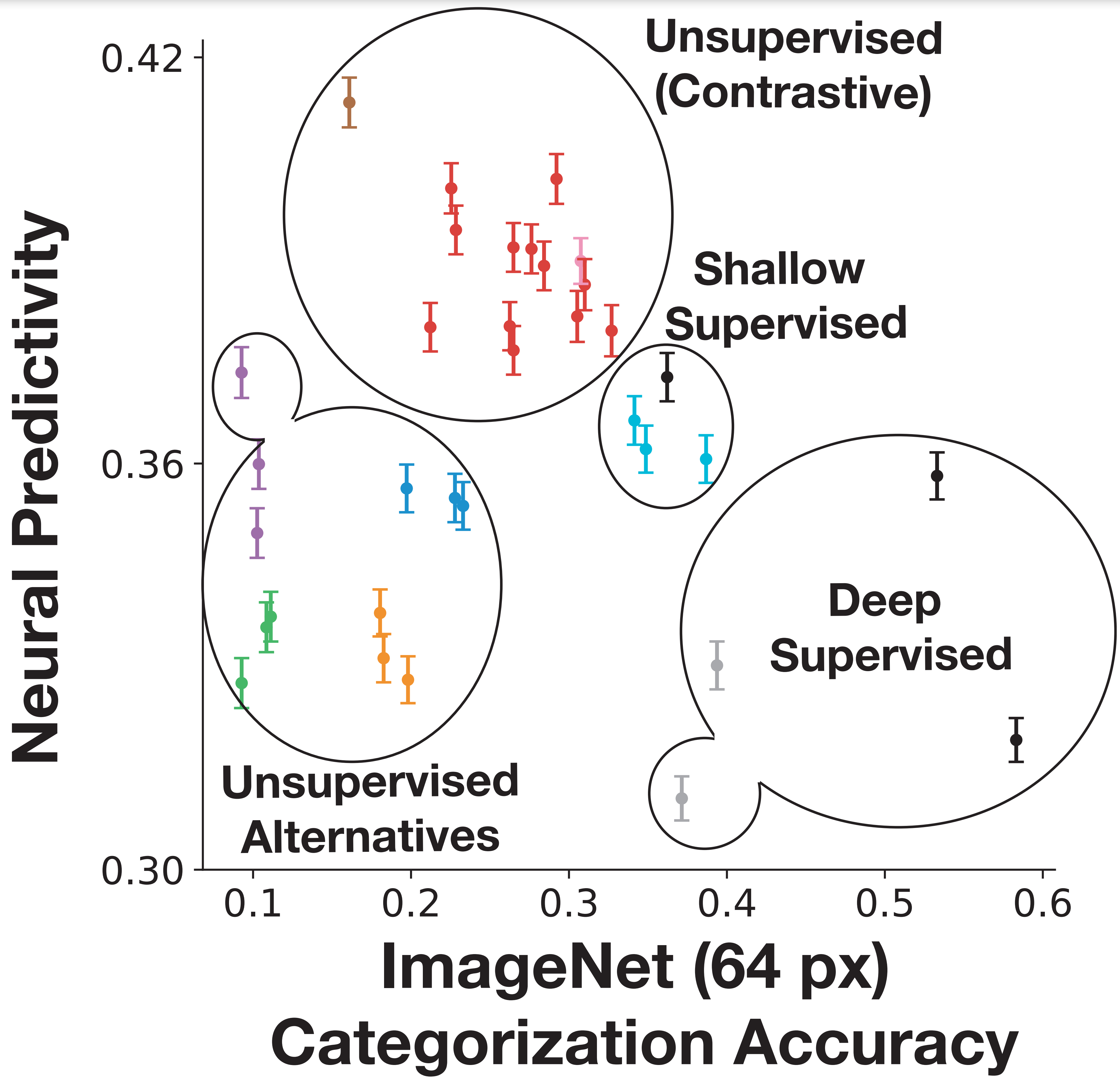
Quantifying hierarchical consistency of models with functional hierarchy of mouse visual cortex

Contrastive AlexNet appears to be roughly optimal for both

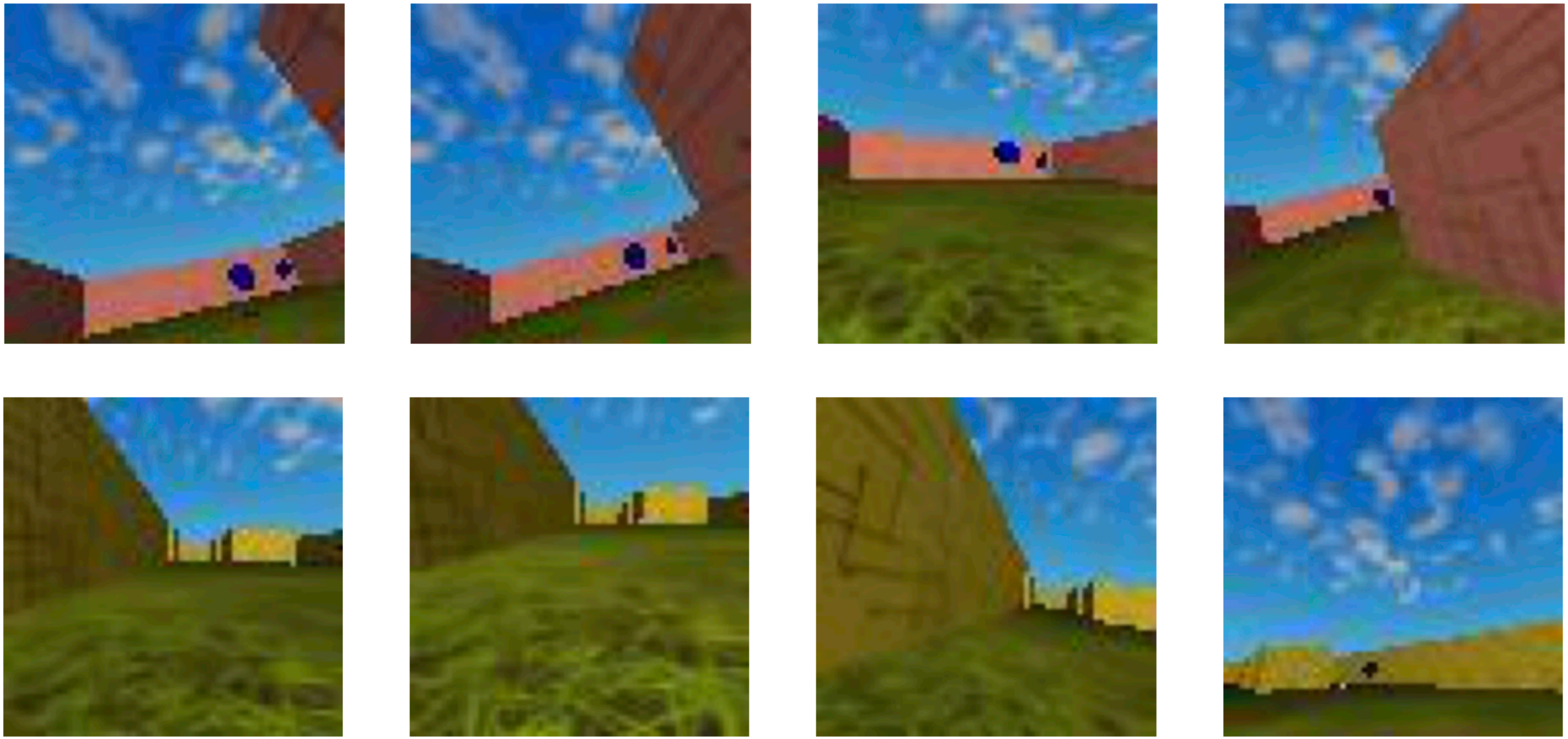


What is the ecological reason for unsupervised nets?

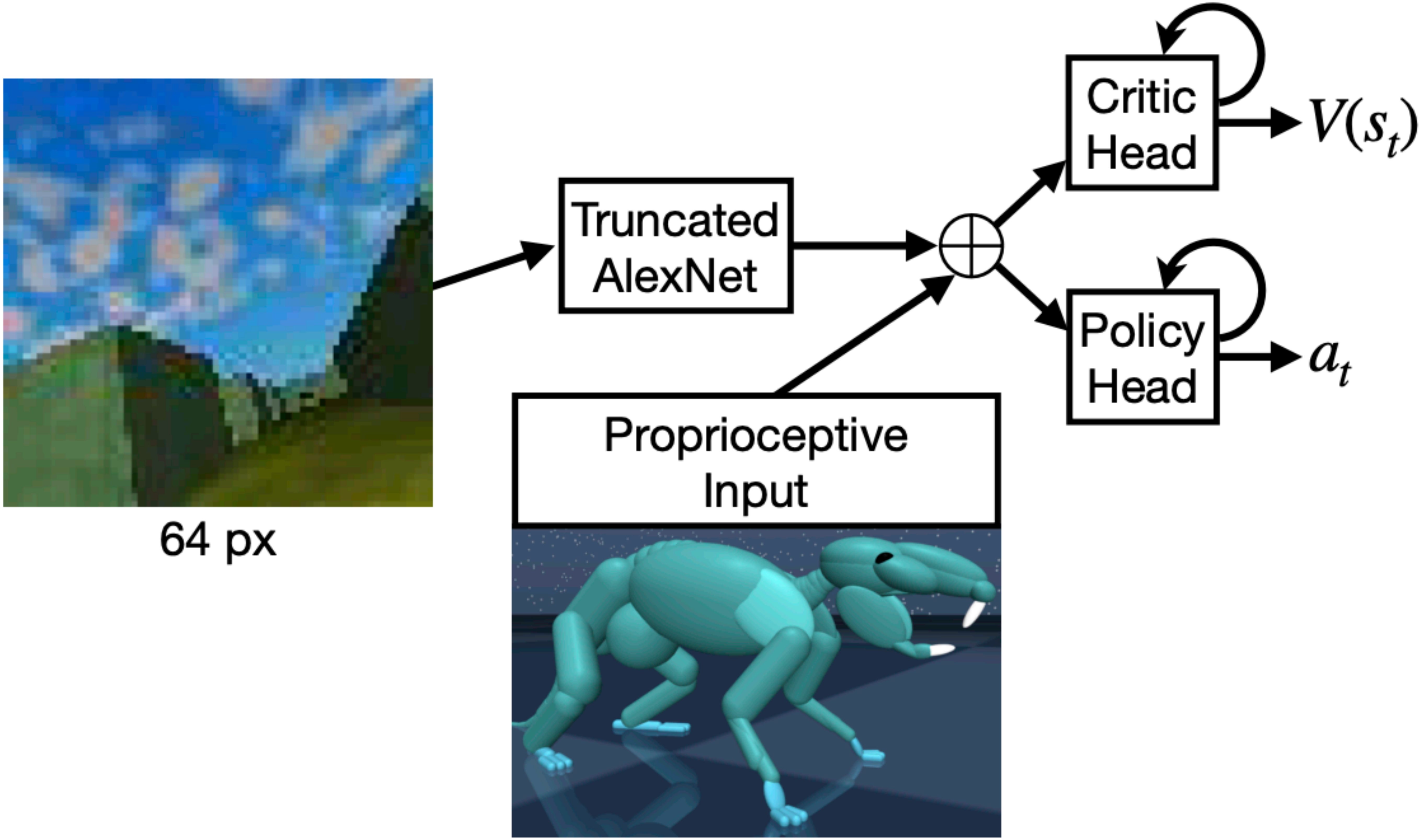
ImageNet categorization performance not correlated with neural predictivity



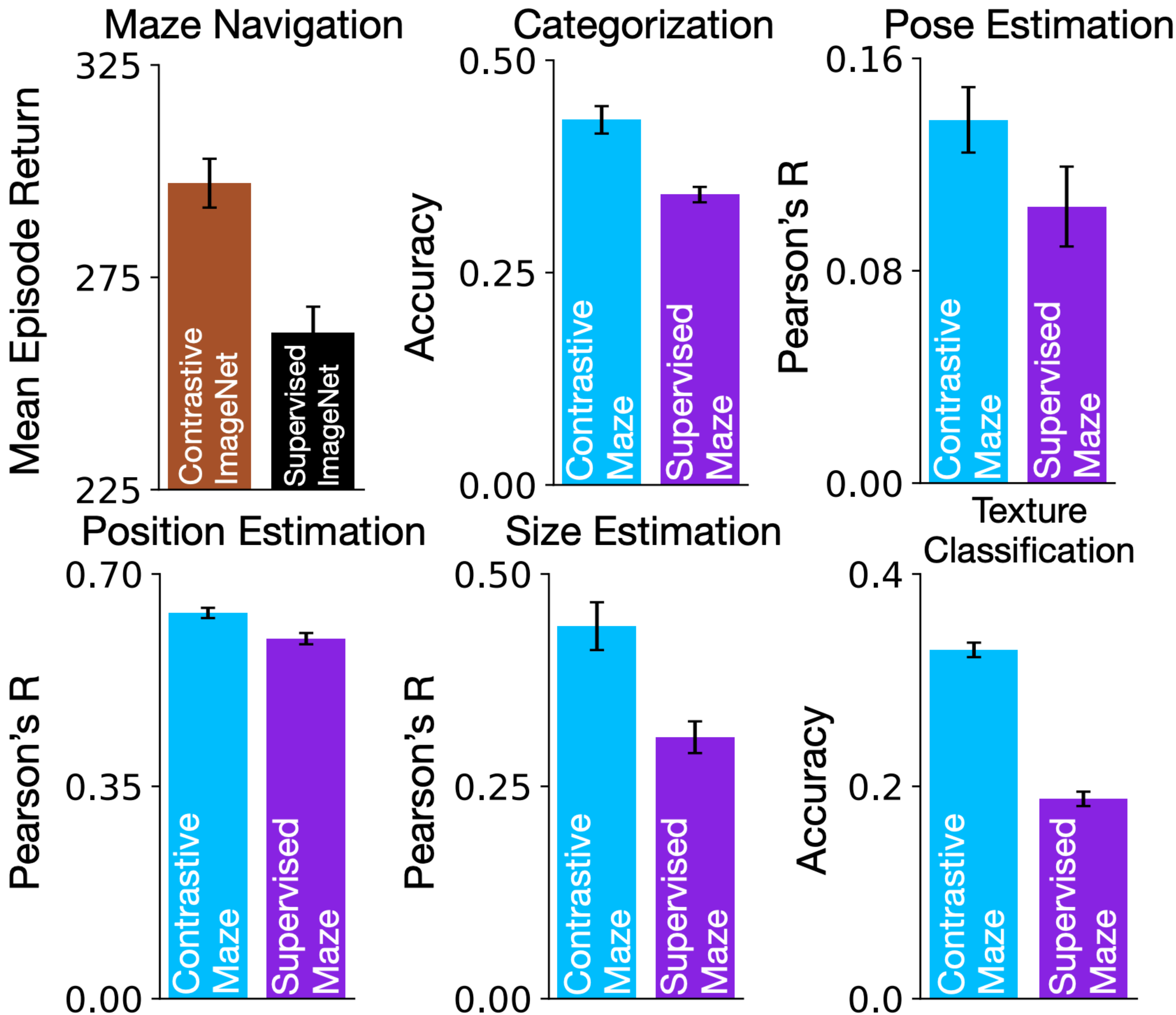
“Rodent Mazes” environment



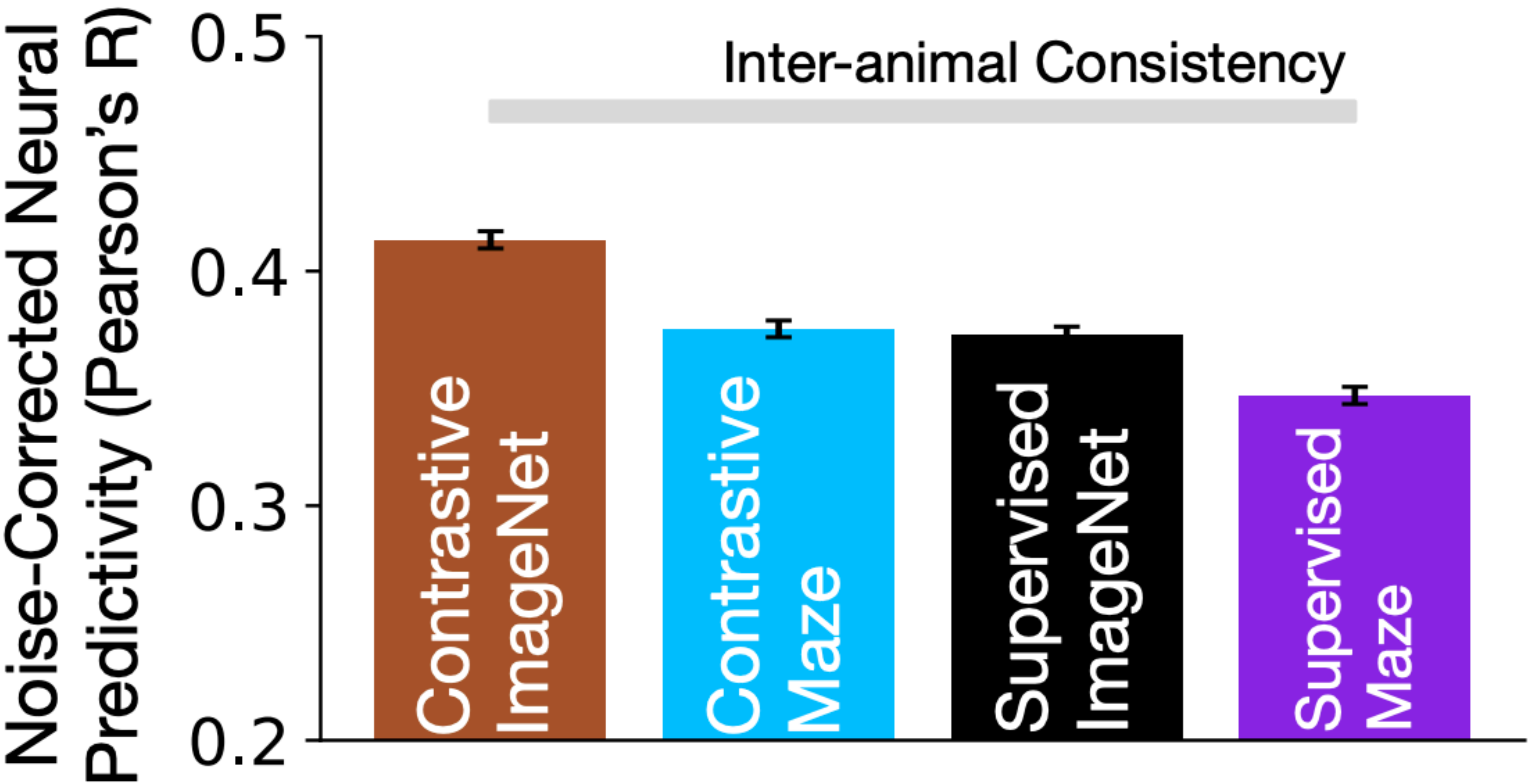
Schematic of RL Agent



Contrastive Models yield better transfer performance



Best models that match the neurons have the best transfer



Takeaways

- Mouse visual cortex is best captured (so far) by three ingredients:

Takeaways

- Mouse visual cortex is best captured (so far) by three ingredients:
  - **Circuit**: shallow architecture
  - **Behavior**: contrastive (unsupervised)
  - **Input**: low resolution

Takeaways

- Mouse visual cortex is best captured (so far) by three ingredients:
    - **Circuit**: shallow architecture
    - **Behavior**: contrastive (unsupervised)
    - **Input**: low resolution
- Mouse visual cortex is a general-purpose machine utilizing its limited resources to perform a variety of visual tasks

Takeaways

- Mouse visual cortex is best captured (so far) by three ingredients:
  - **Circuit**: shallow architecture
  - **Behavior**: contrastive (unsupervised)
  - **Input**: low resolution
- Mouse visual cortex is a general-purpose machine utilizing its limited resources to perform a variety of visual tasks
- This is all in contrast to the deep, high-resolution, and task-specific visual system of the primate

Takeaways

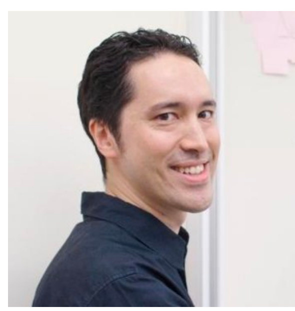
- Mouse visual cortex is best captured (so far) by three ingredients:
  - **Circuit**: shallow architecture
  - **Behavior**: contrastive (unsupervised)
  - **Input**: low resolution
- Mouse visual cortex is a general-purpose machine utilizing its limited resources to perform a variety of visual tasks
- This is all in contrast to the deep, high-resolution, and task-specific visual system of the primate
- Generic nature of the behavior could be used by other sensory systems

Acknowledgements

Thanks to my awesome collaborators!

Nathan C.L. Kong*

Chengxu Zhuang



Justin L. Gardner

Anthony M. Norcia

Daniel L.K. Yamins

Contact:
anayebi@mit.edu;
nclkong@stanford.edu

Funding:

Neurosciences PhD Program

Stanford Mind, Brain, Computation and Technology Training Program,
Wu Tsai Neurosciences Institute