Task-Driven Recurrent Models & Dissecting Neural Computations *In-Silico*

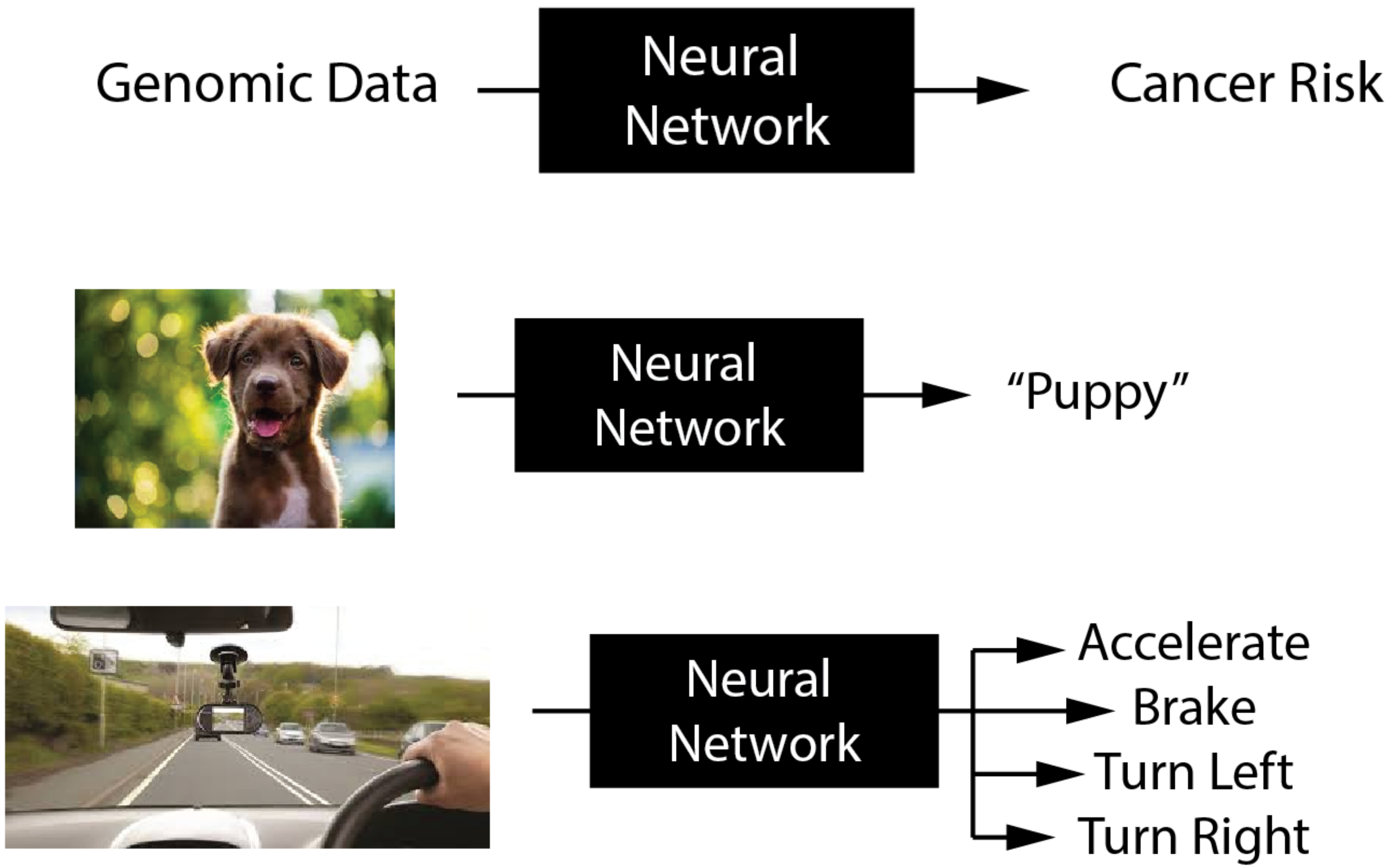
Aran Nayebi
Neurosciences PhD Program
Stanford University

Bernstein 2019
Brain against the Machine Workshop
2019.09.18

Why has deep learning become so popular?

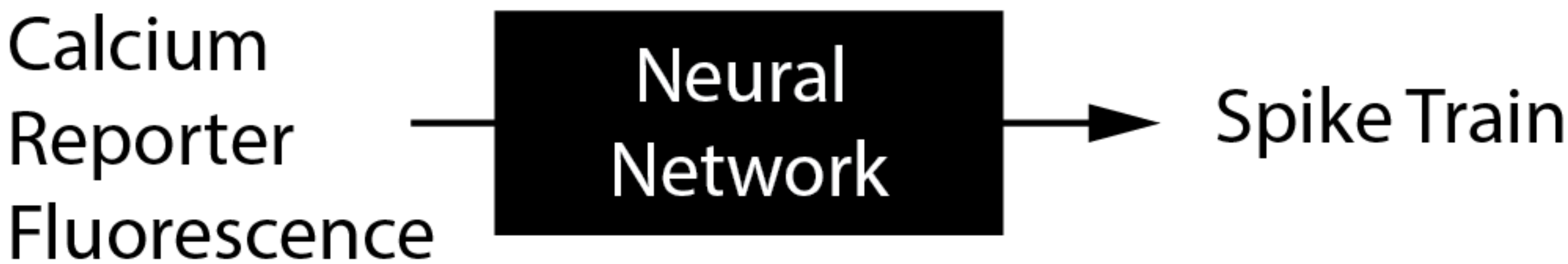
An effective *tool* for data inference problems.

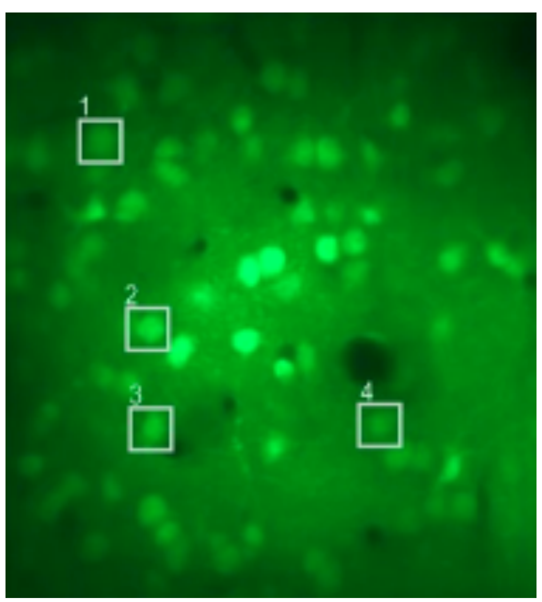
NNs are used as a tool in many disciplines

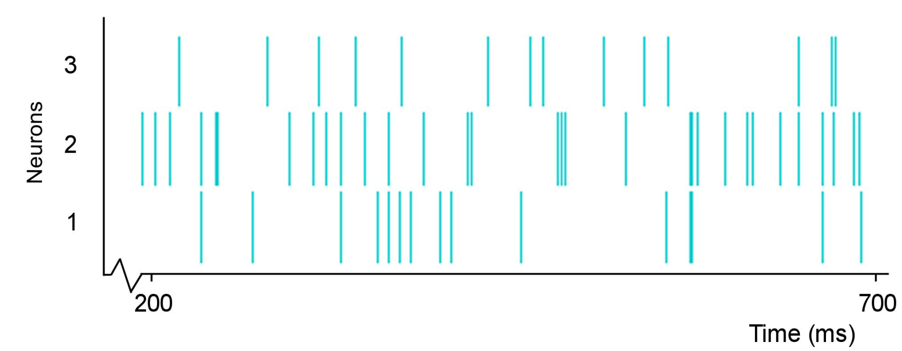


Zou, J., …, Telenti, A. *A primer on deep learning in genomics,* Nature Genetics, 2018
Krizhevsky, A., …, Hinton, G. *ImageNet Classification with Deep Convolutional Neural Networks,* NIPS, 2012
Bojarski, M., …, Zhao, J. *End to End Learning for Self Driving Cars*, arXiv, 2016

And they can be useful as a tool for neuroscience data



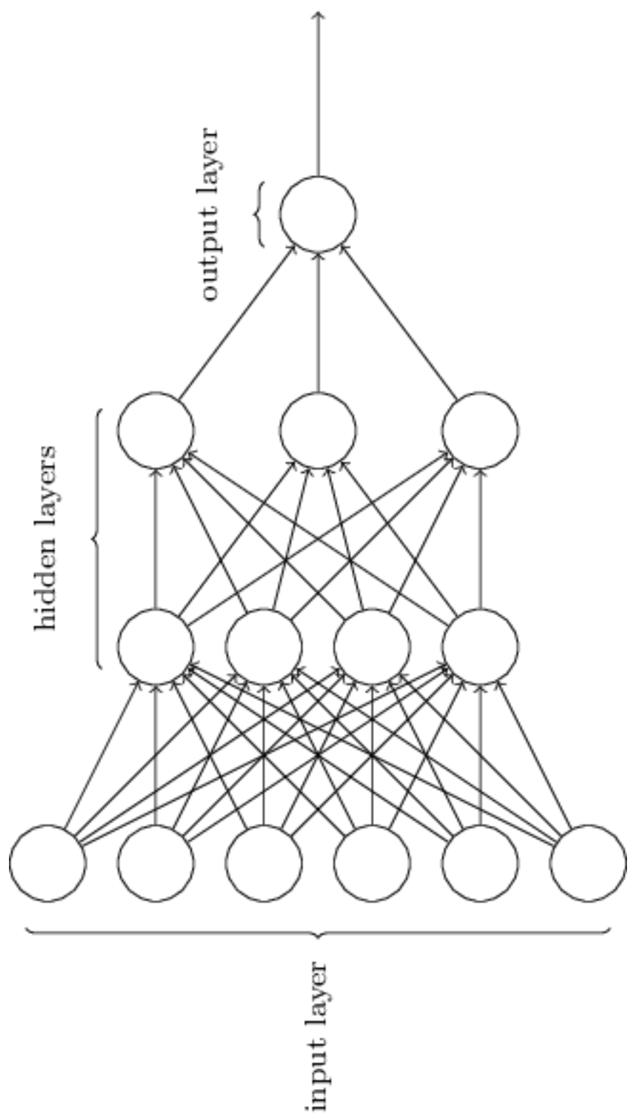




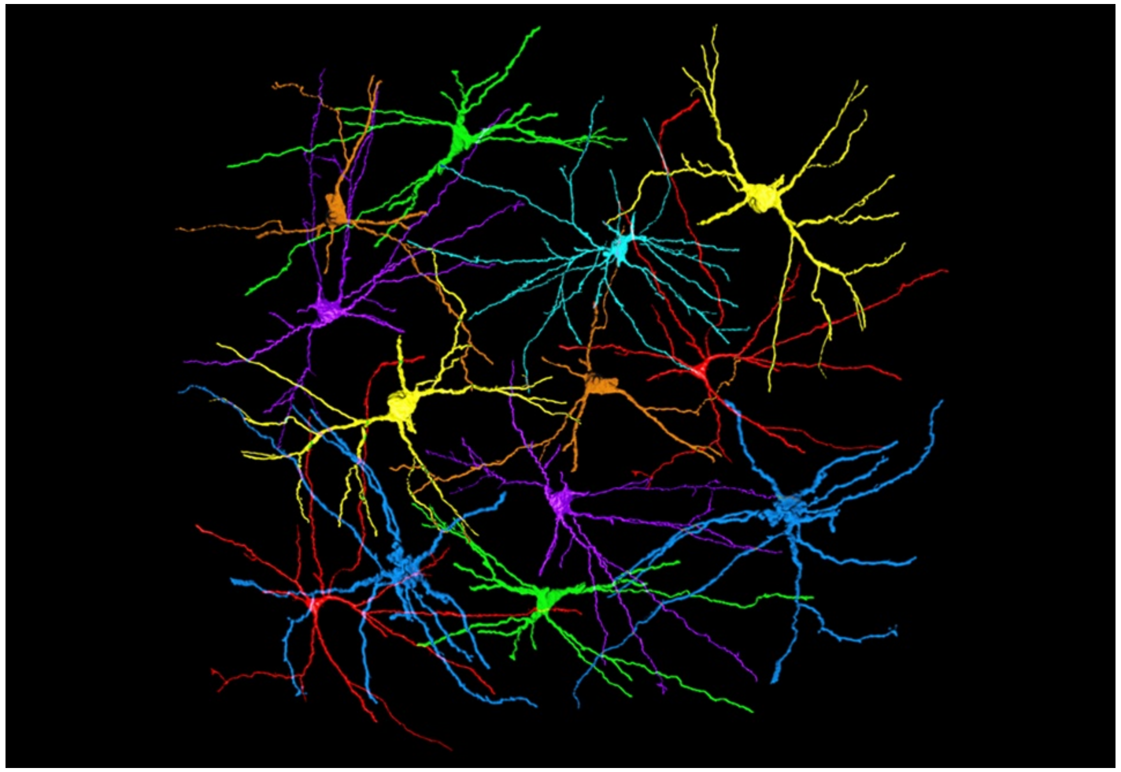
Not the kind of models we will discuss

Putting the *Neural* in Neural Network

Artificial NN



Biological NN



A Case Study: The Visual System

A Case Study: The Visual System

Guiding Questions:

A Case Study: The Visual System

Guiding Questions:

Can we use "biologically-inspired" neural network models to provide normative **(why?)** & mechanistic **(how?)** insights into a neural circuit?

A Case Study: The Visual System

Guiding Questions:

Can we use “biologically-inspired” neural network models to provide normative **(why?)** & mechanistic **(how?)** insights into a neural circuit?

Can we also use these “biologically-inspired” networks to improve an AI goal?

Current Approaches:

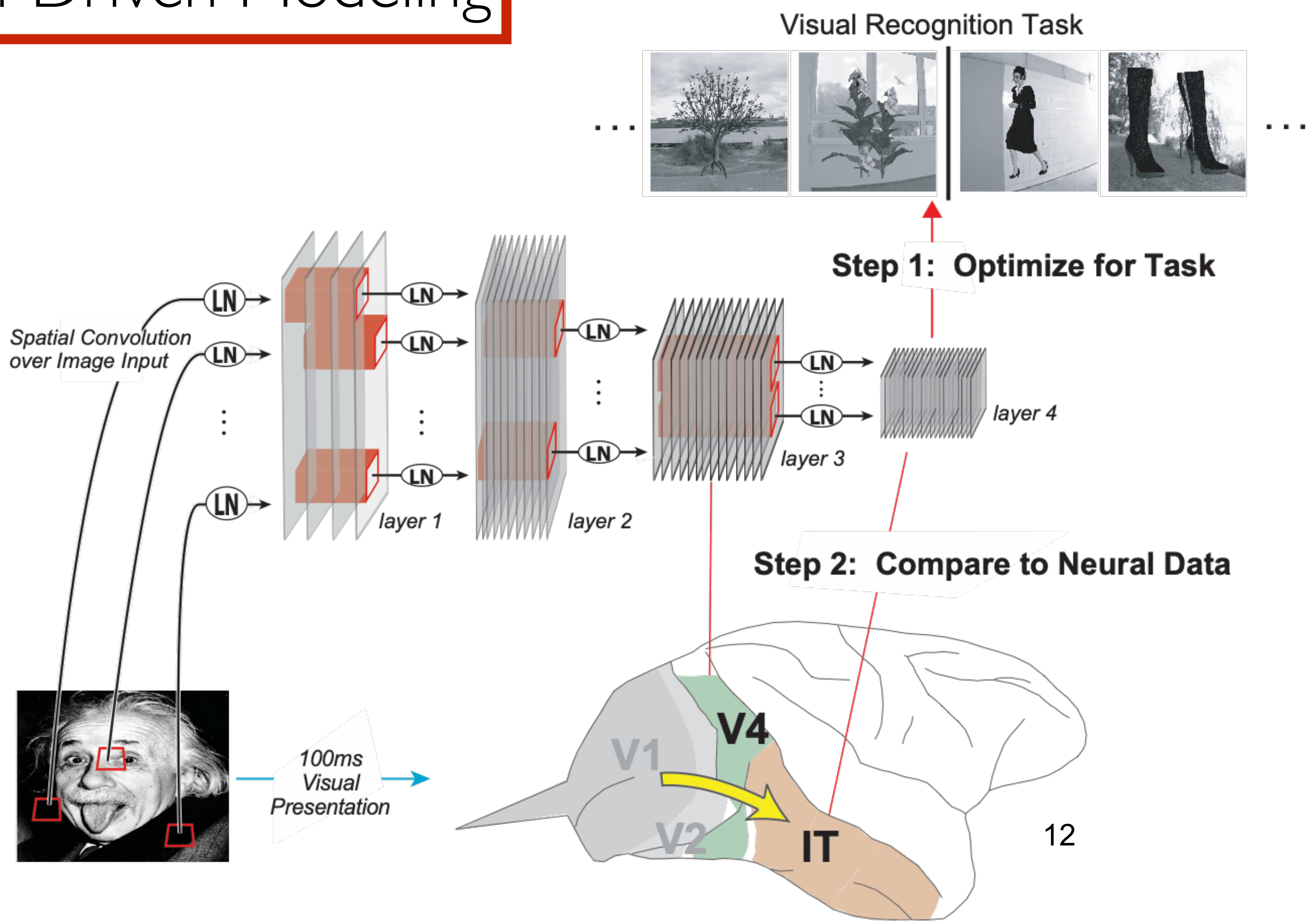
- Goal-Driven Modeling

- Direct Fits to Data

Roadmap

Current Approaches:

- Goal-Driven Modeling

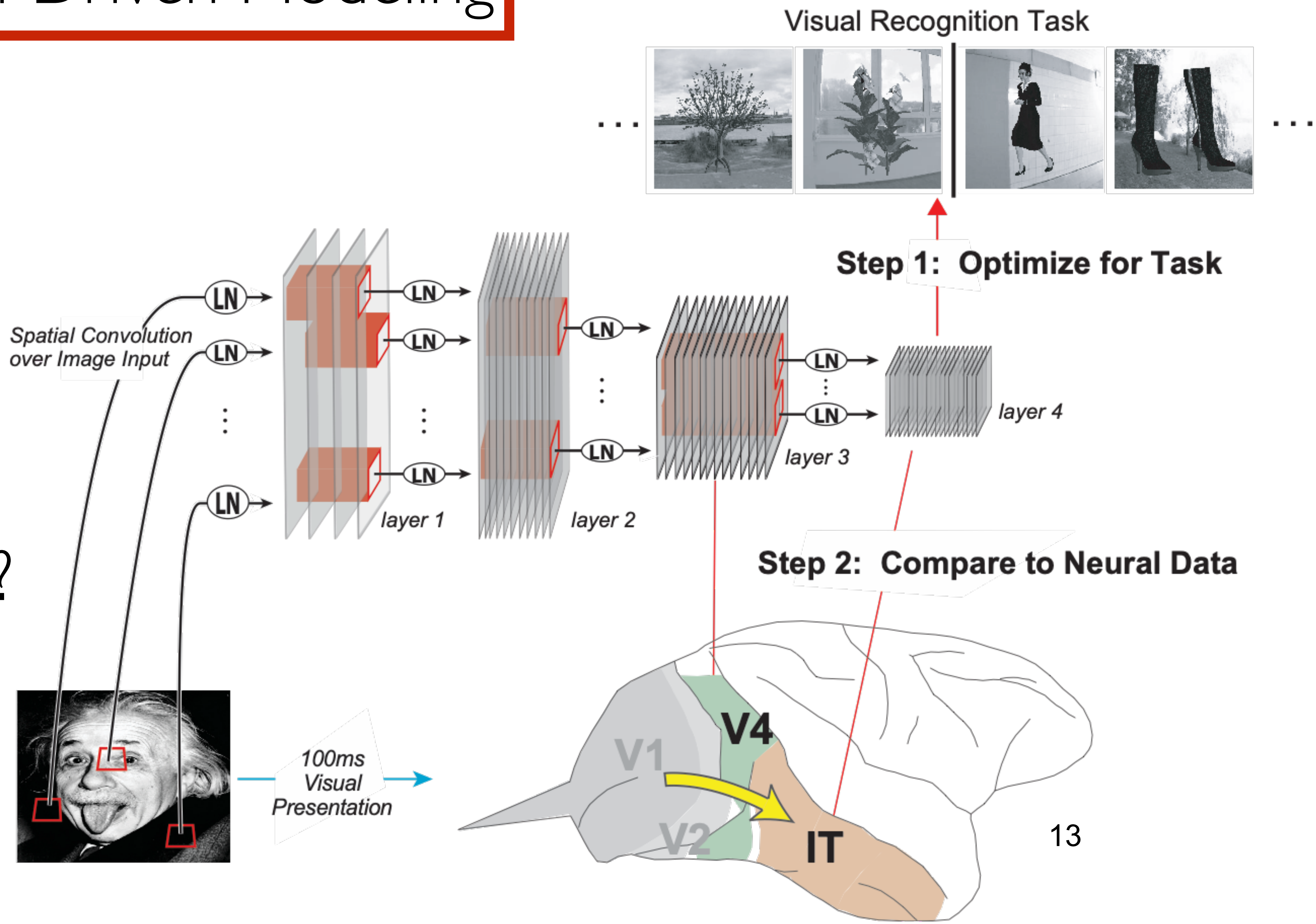


Roadmap

Current Approaches:

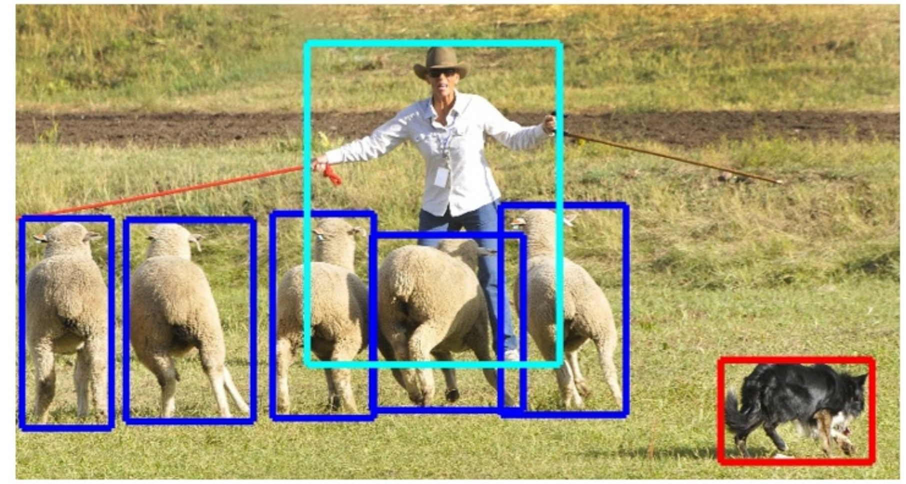
- Goal-Driven Modeling

For what behavioral "goal" has the system been evolved for?

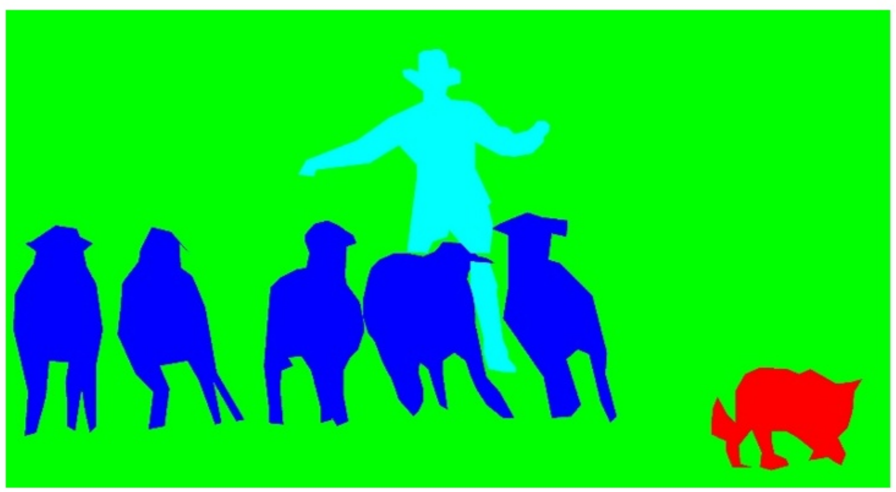


Classification, Segmentation, Localization, …

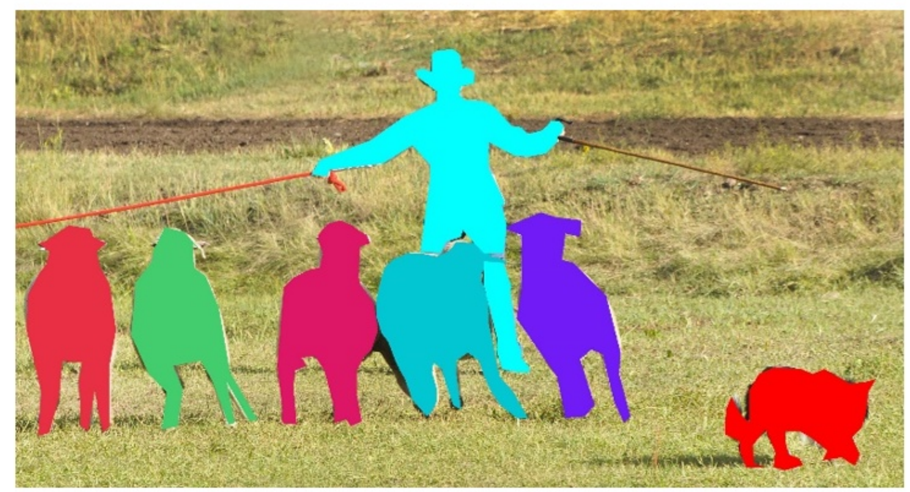
(a) **Image classification**



(b) **Object localization**



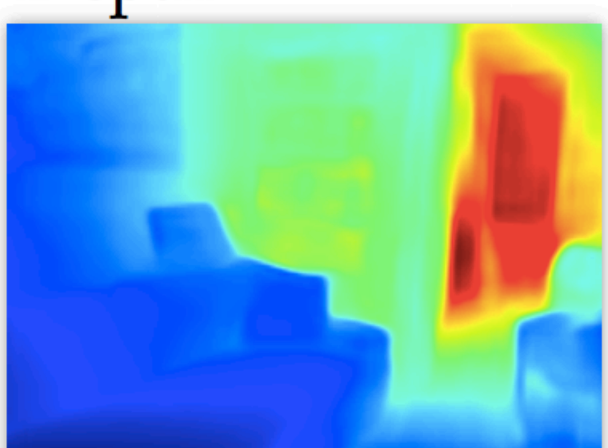
(c) **Semantic segmentation**



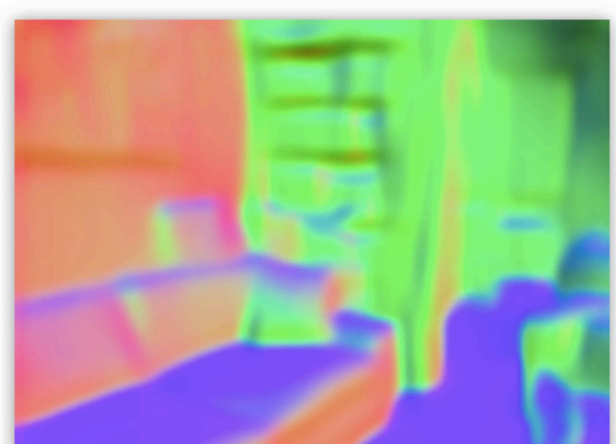
(d) **Instance segmentation**

Lin et al. 2014

Depth



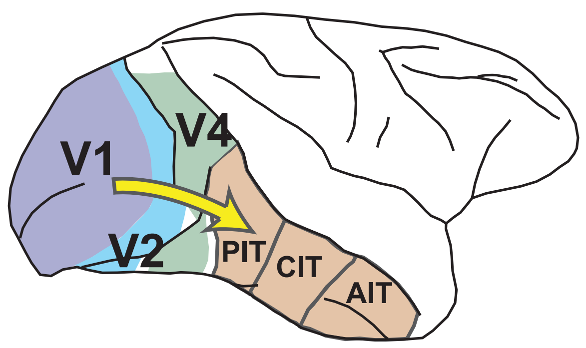
Normals



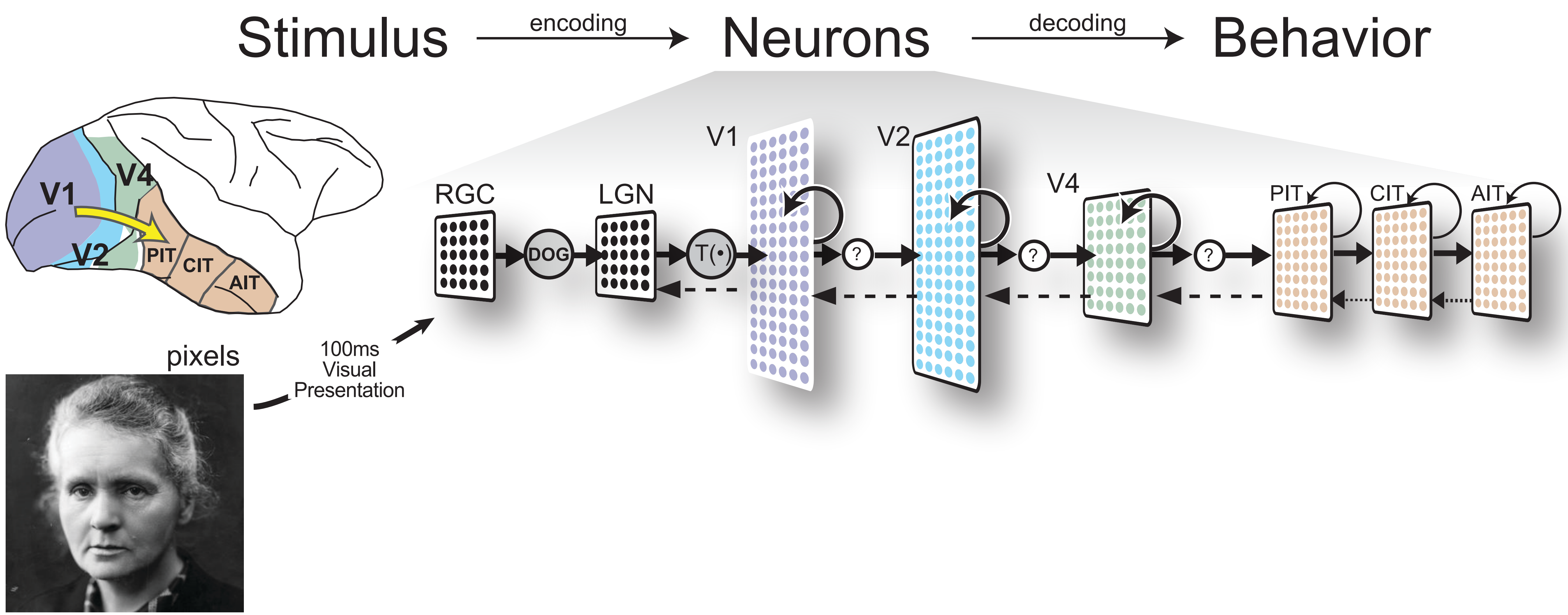
Eigen and Fergus 2015

…thanks to the Ventral Stream

Stimulus $\xrightarrow{\text{encoding}}$ Neurons $\xrightarrow{\text{decoding}}$ Behavior

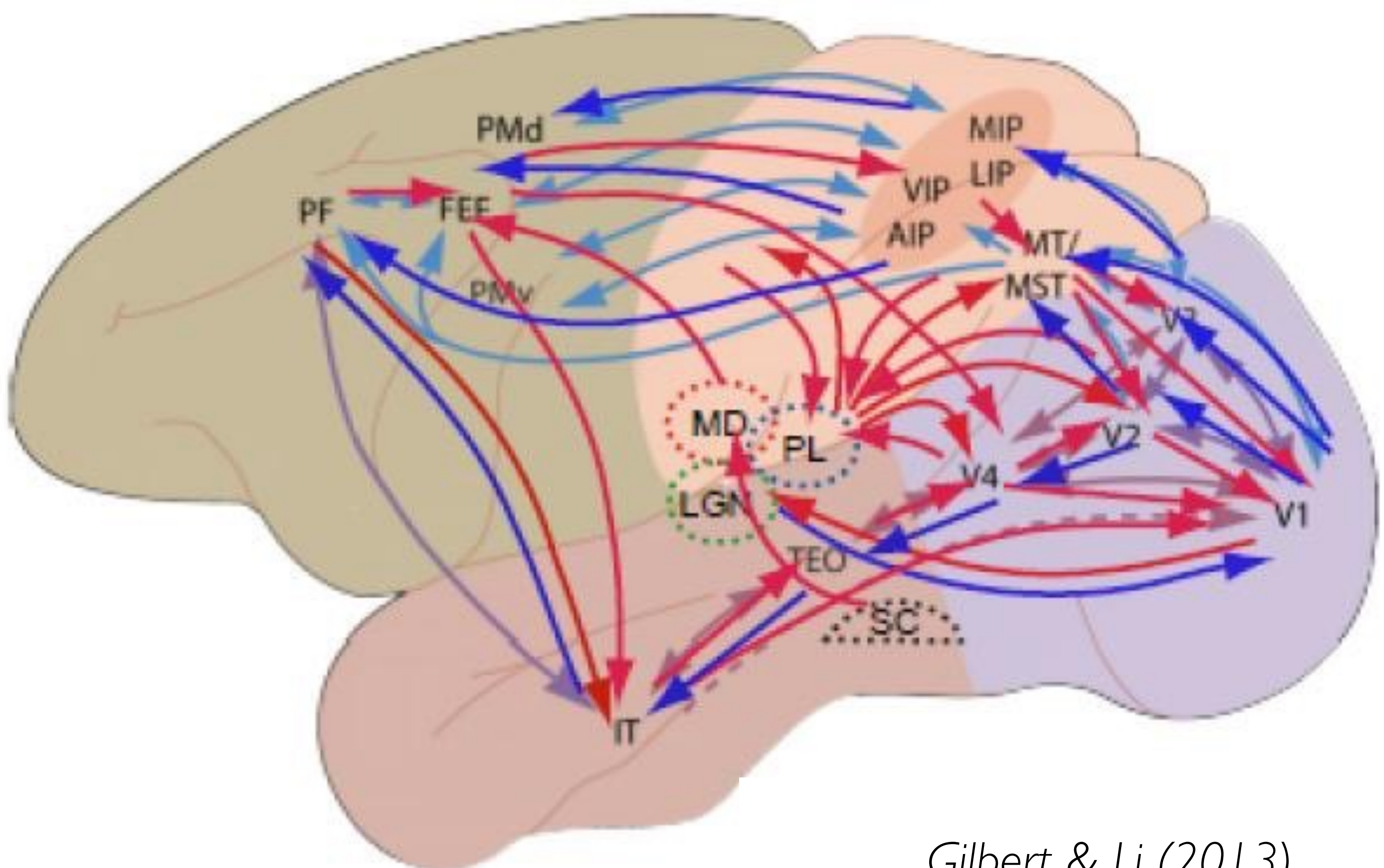


…which is somewhat feedforward



…but also not just feedforward

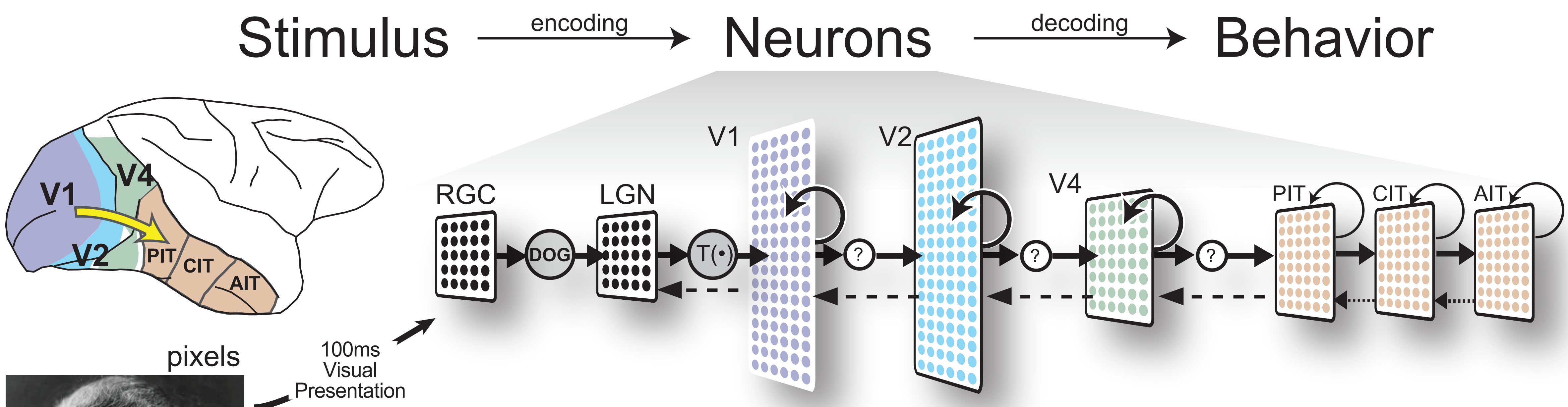
Feedbacks are everywhere anatomically:

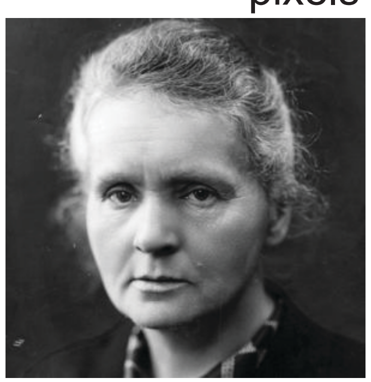


Gilbert & Li (2013)

…but what are they for?

CNNs as Models of Object Recognition





Convolutional Neural Networks (CNNs)

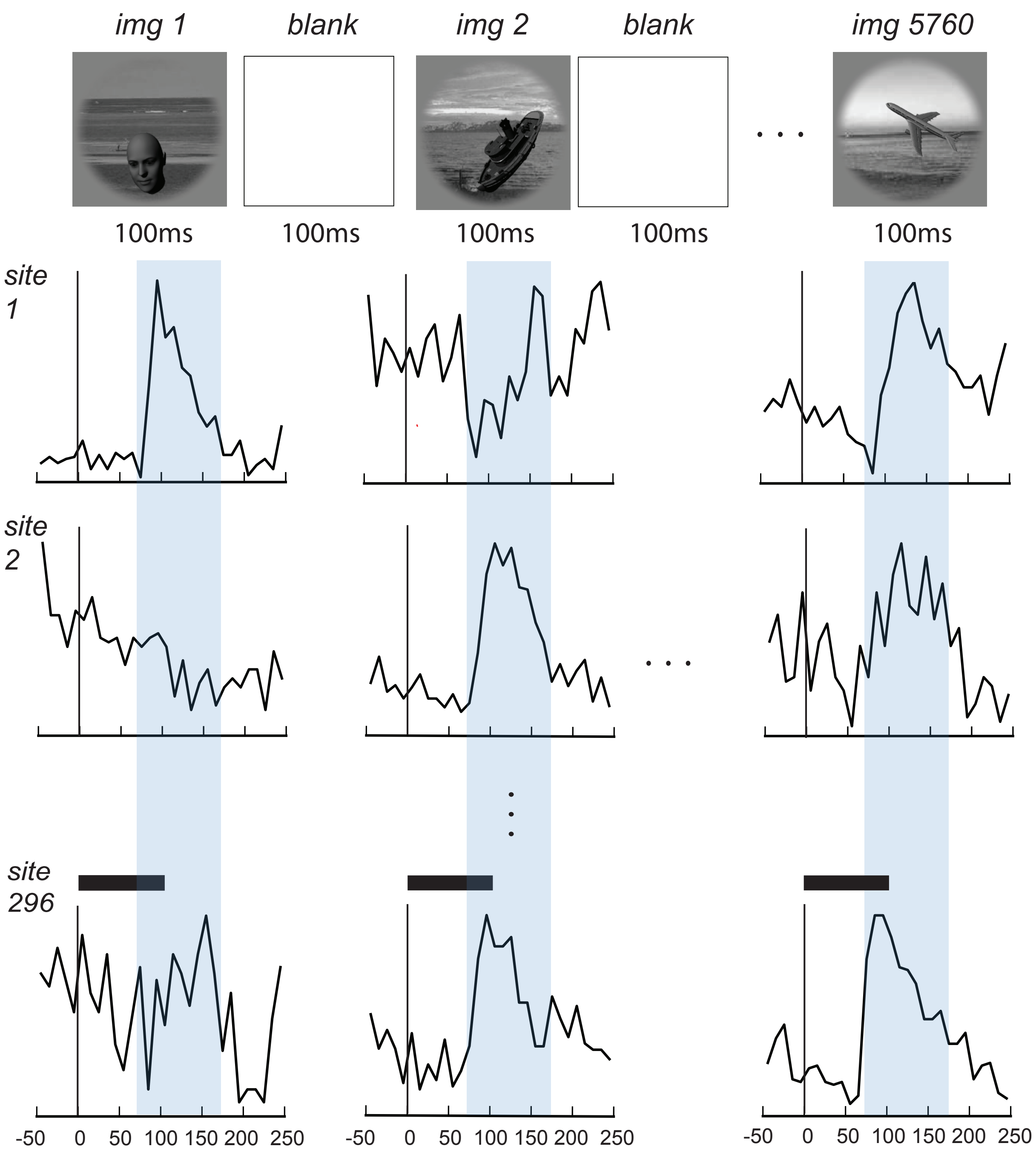
Fukushima, 1979; Lecun, 1995

CNNs are inspired by visual neuroscience:

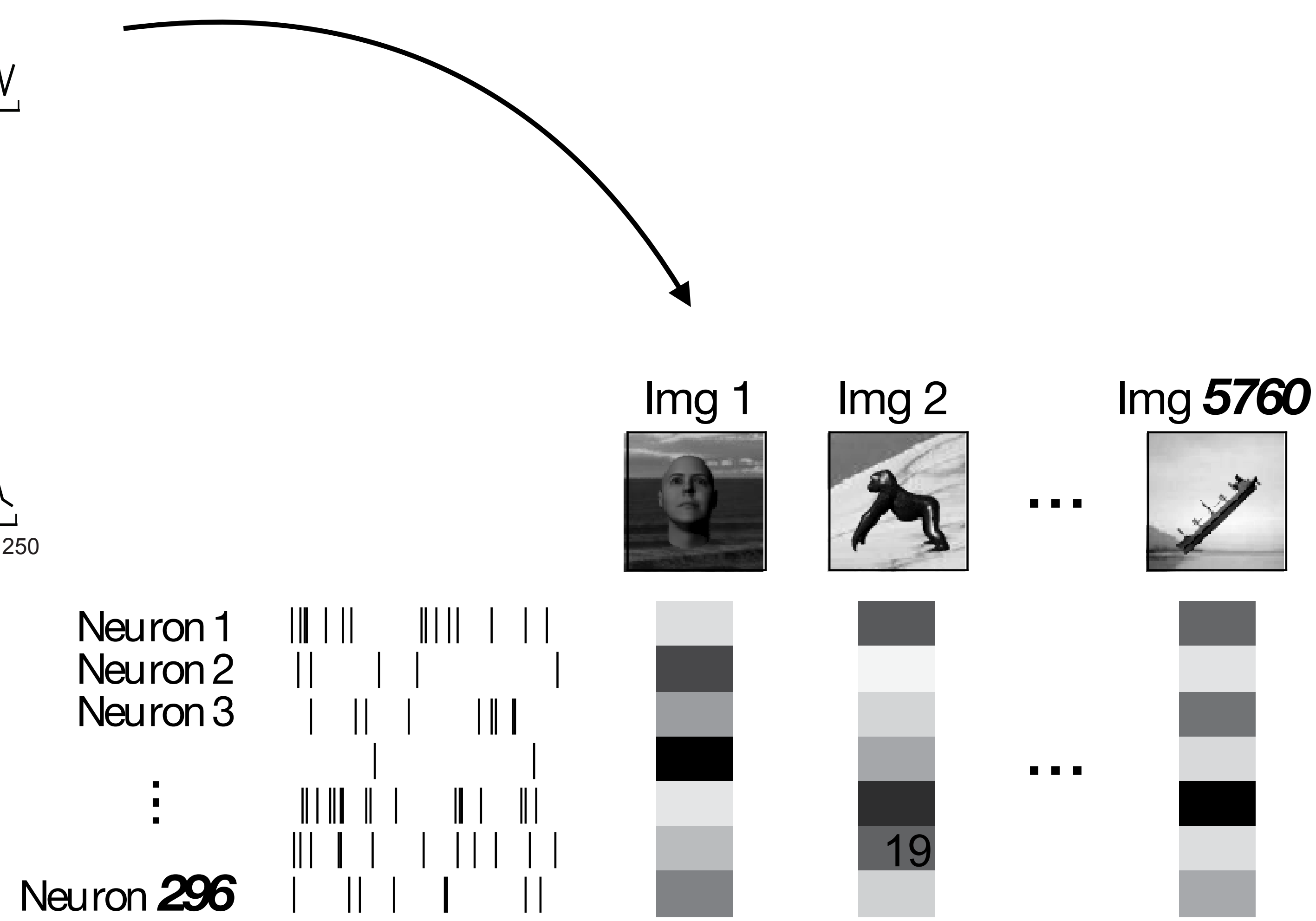
1) **hierarchy**

2) **retinotopy** (spatially tiled)

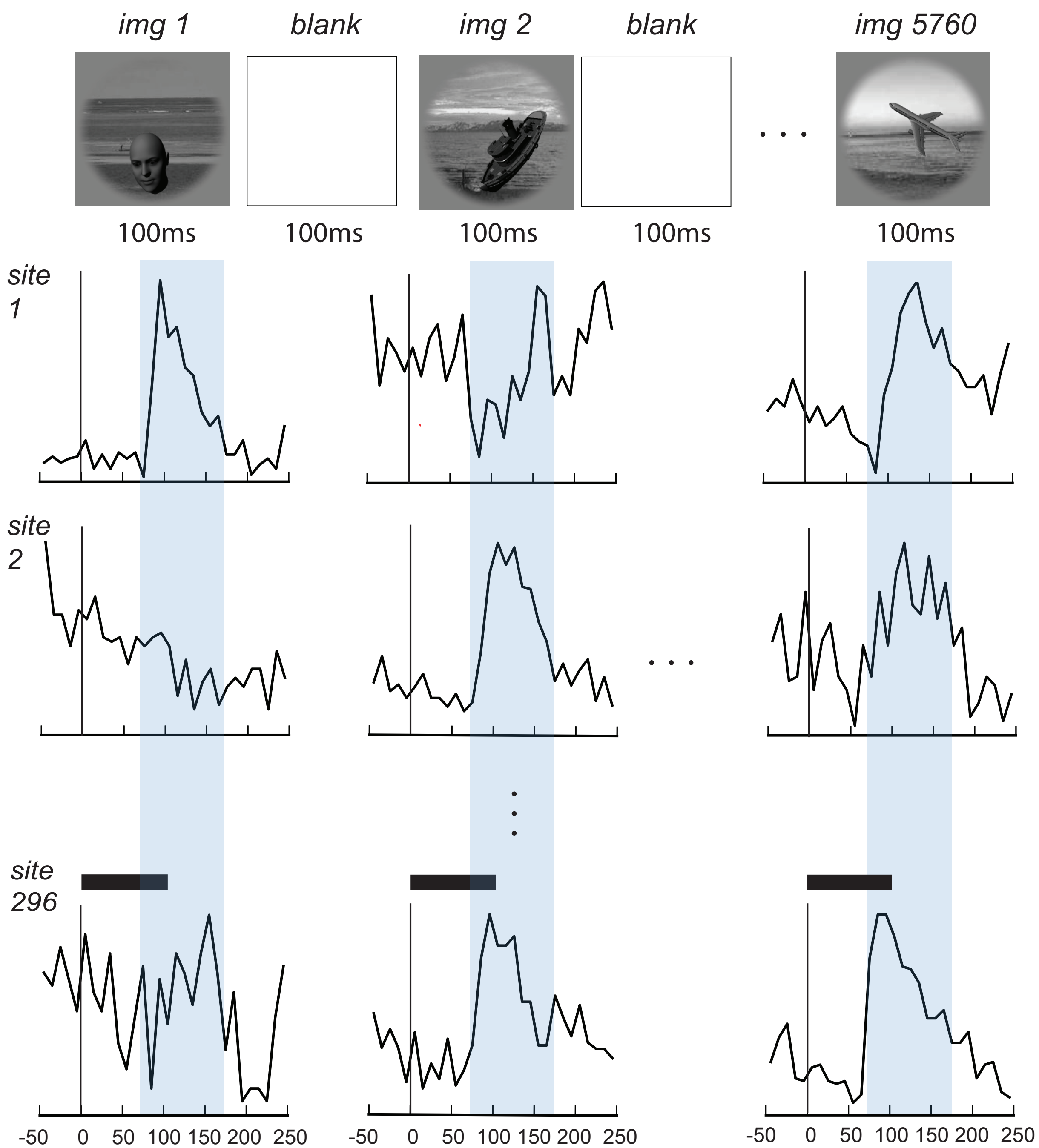
So far, only explaining temporal average of responses



e.g. Binned spike counts 70ms-170ms post stimulus presentation

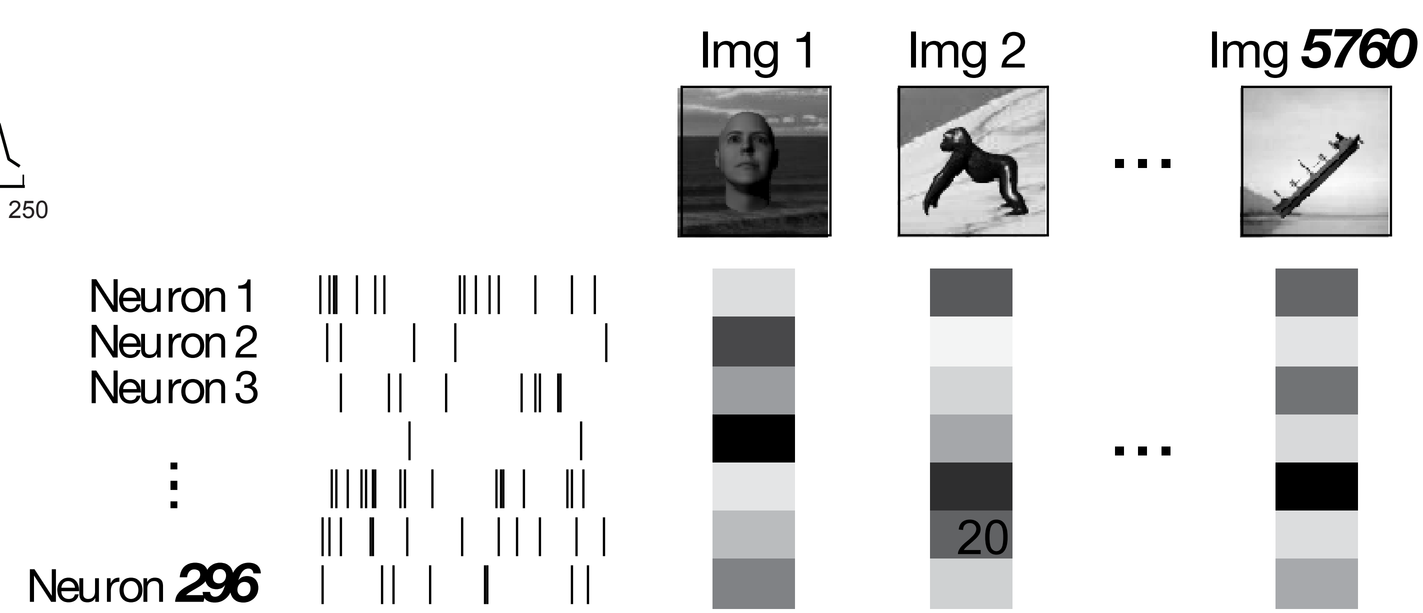


So far, only explaining temporal average of responses

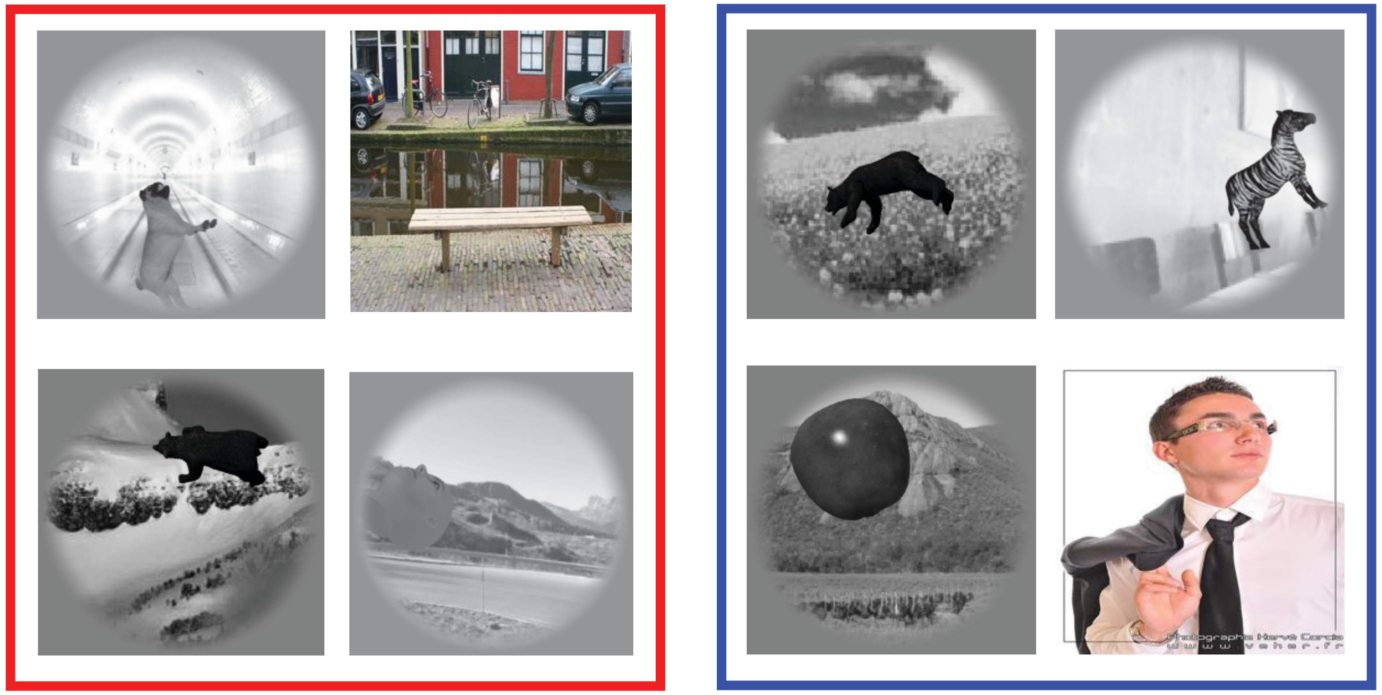


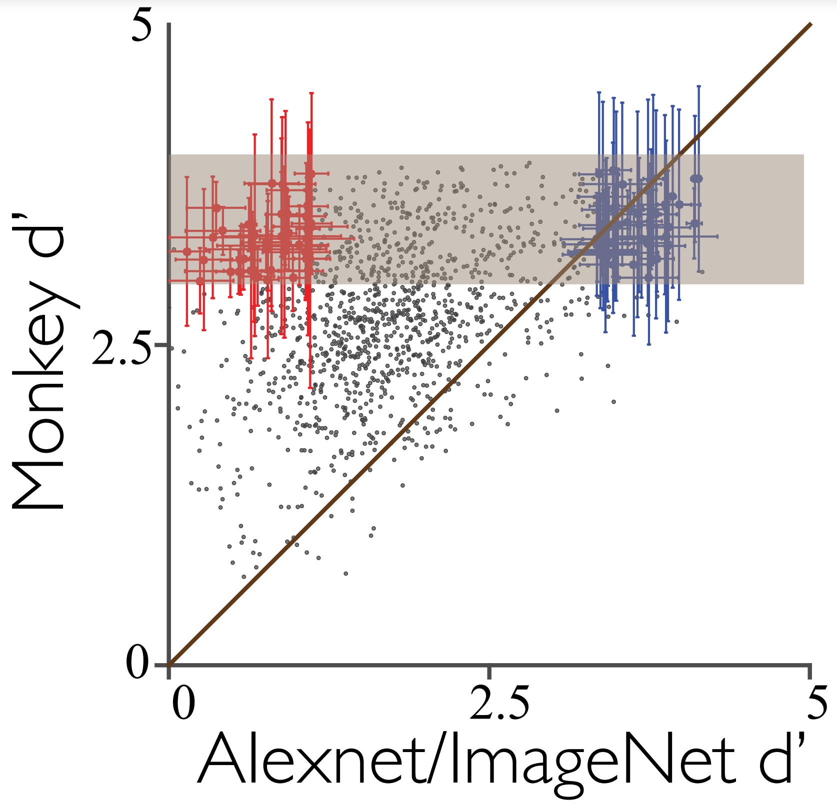
e.g. Binned spike counts 70ms-170ms post stimulus presentation

but actually the data is highly reliable at much finer grain —10ms bins



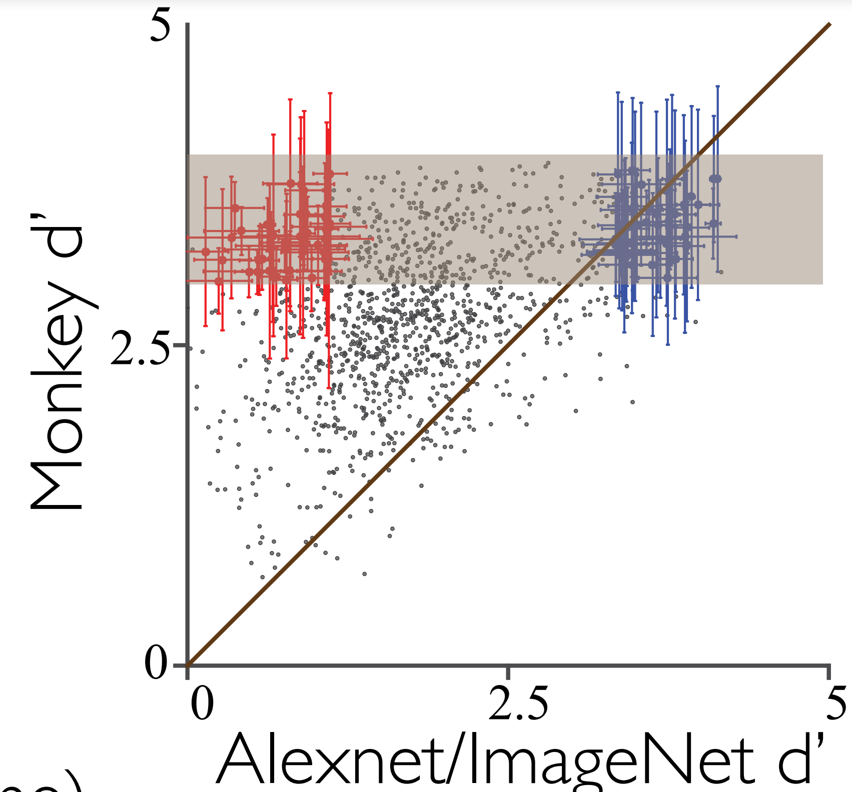
Evidence of Functional Relevance



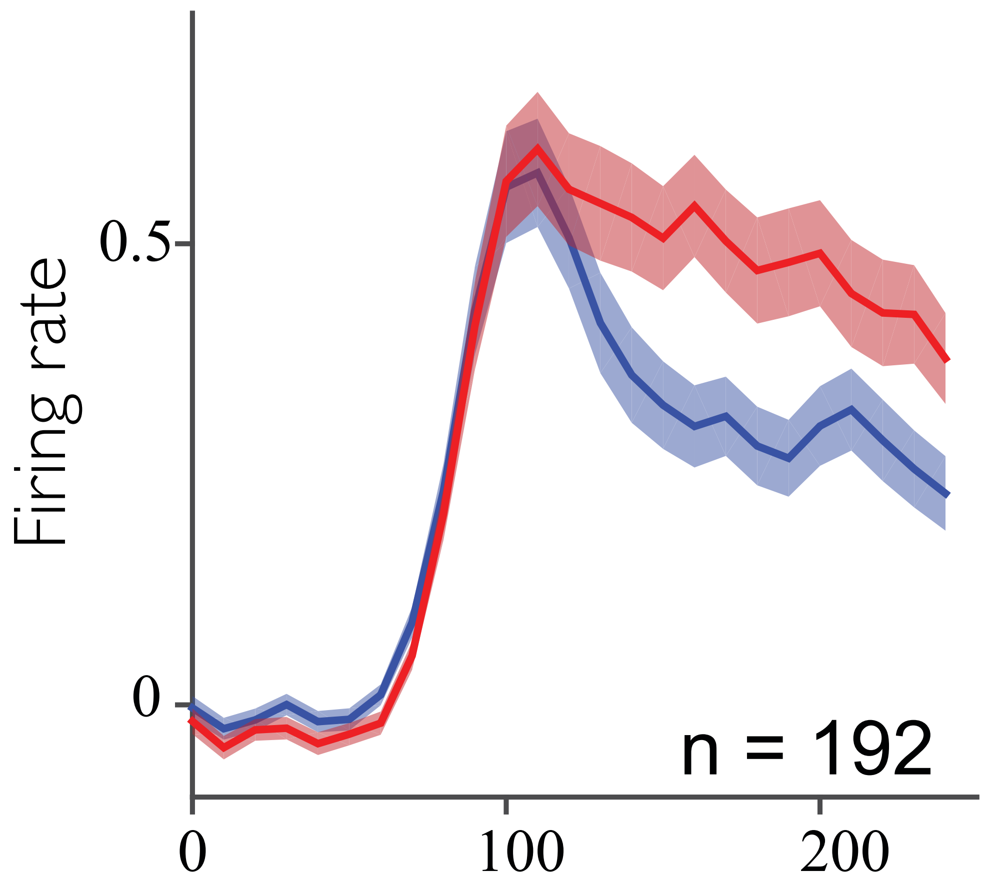


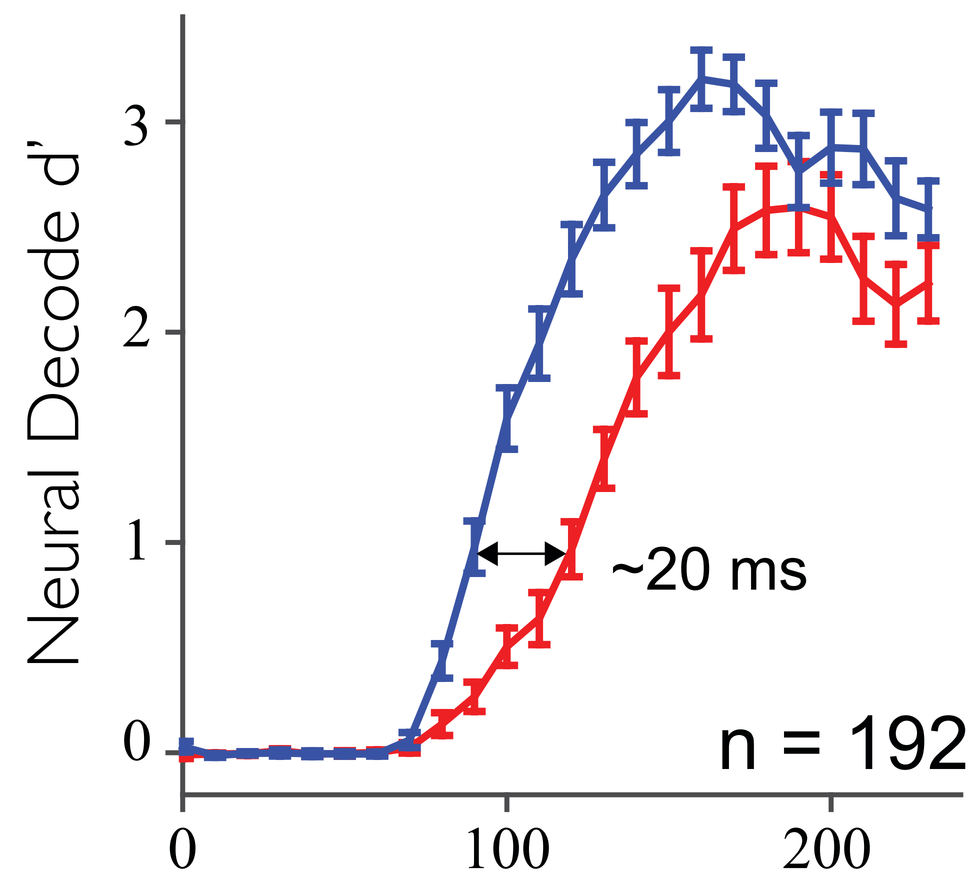
Kar et. al. (2018)

Evidence of Functional Relevance



… even in range before 250ms (rough saccade time)





Time post-image onset (ms)

Kar et. al. (2018)

Overview

Nayebi, A.*, Bear, D.*, Kubilius, J.*,…, Yamins, D.L.K. *Task-Driven Convolutional Recurrent Models of the Visual System*, NeurIPS 2018

- Expand architecture class (local and global recurrence)
- Parametrize local and global feedback motifs and optimize for performance on ImageNet
- Evaluate neural predictivity in V4 and IT temporal responses

Daniel Bear
(Stanford)

Jonas Kubilius
(MIT)

Overview

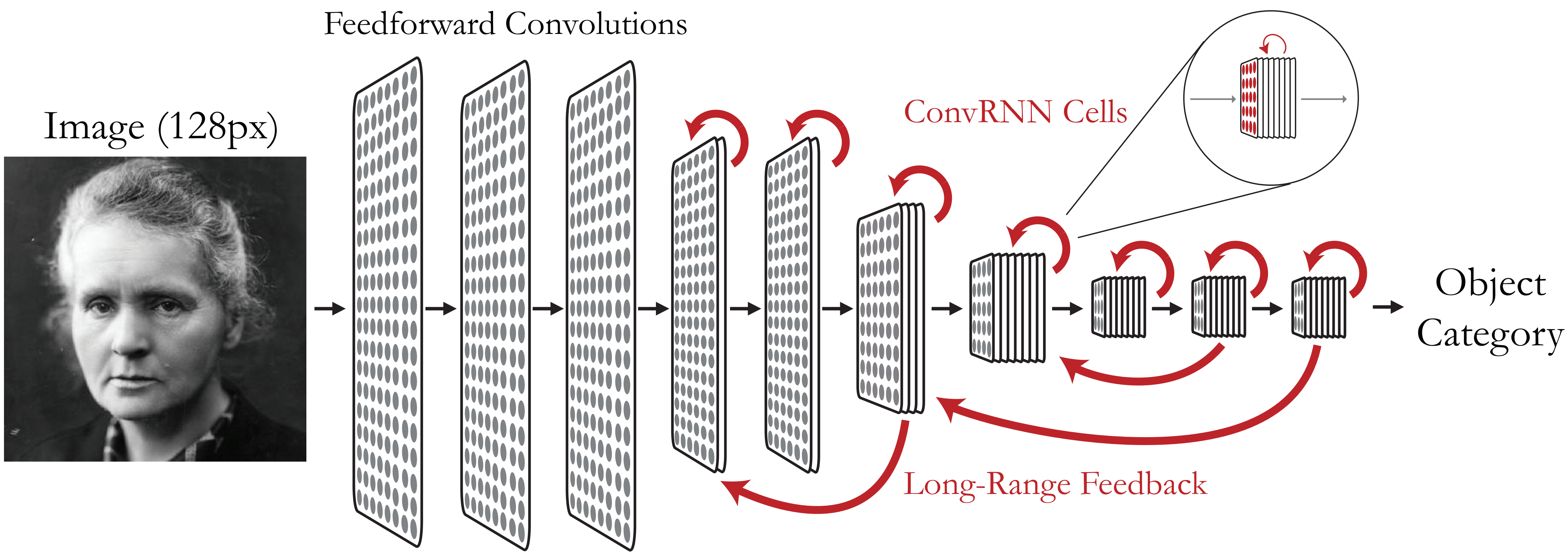
Nayebi, A.*, Bear, D.*, Kubilius, J.*,…, Yamins, D.L.K. *Task-Driven Convolutional Recurrent Models of the Visual System*, NeurIPS 2018

- Expand architecture class (local and global recurrence)
- Parametrize local and global feedback motifs and optimize for performance on ImageNet
- Evaluate neural predictivity in V4 and IT temporal responses

Daniel Bear
(Stanford)

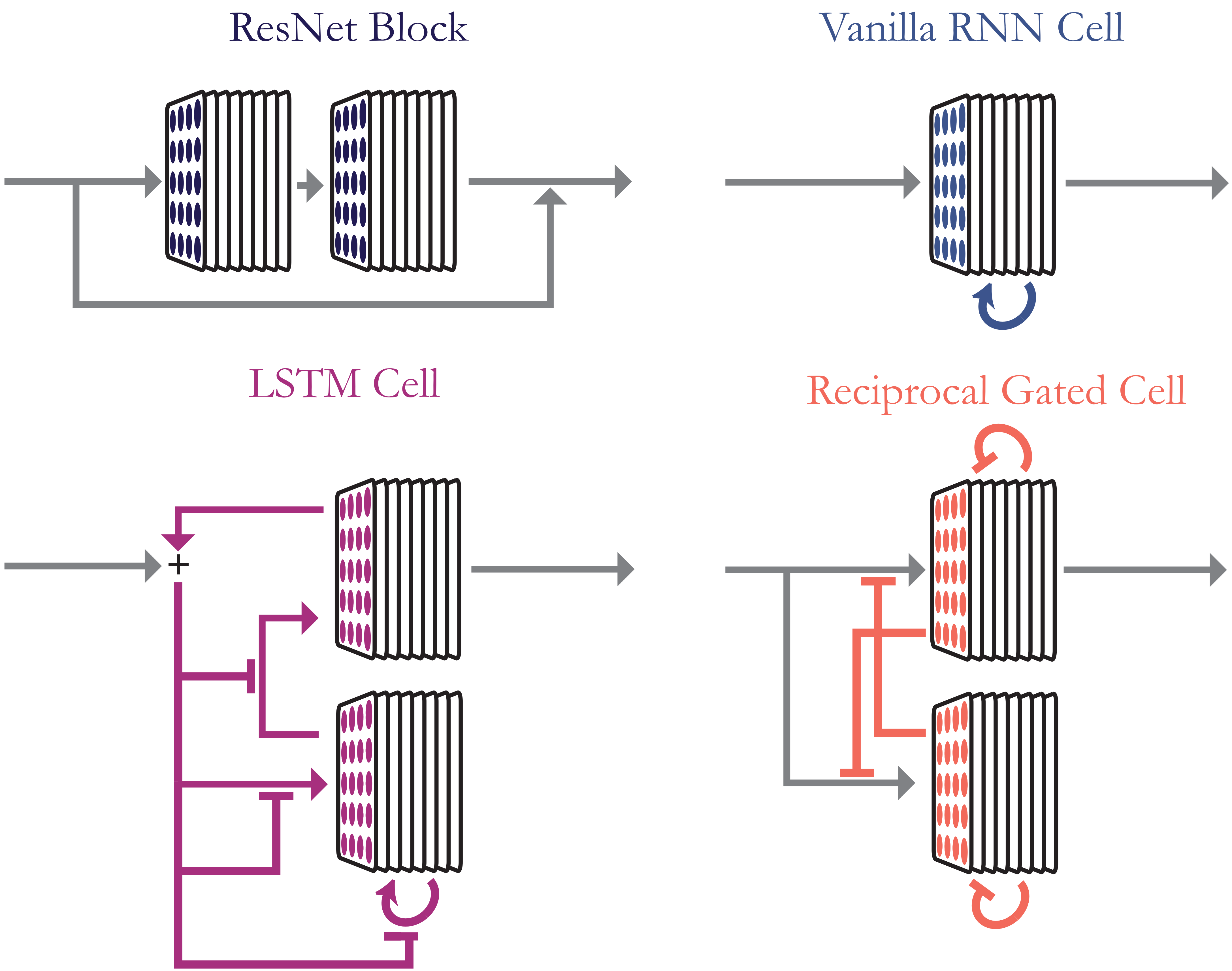
Jonas Kubilius
(MIT)

Convolutional Recurrent Neural Networks (ConvRNNs)



Each time-step (10 ms) is treated equally — including feedforward steps

Many Choices of Local Recurrence



Principles of Local Recurrence

Two complementary principles:

(1) gating = multiplication by input-dependent tensor w/ values in [0, 1]

(2) bypassing = when recurrent cell is in 0 state, input is unchanged (“performance preserving”)

Principles of Local Recurrence

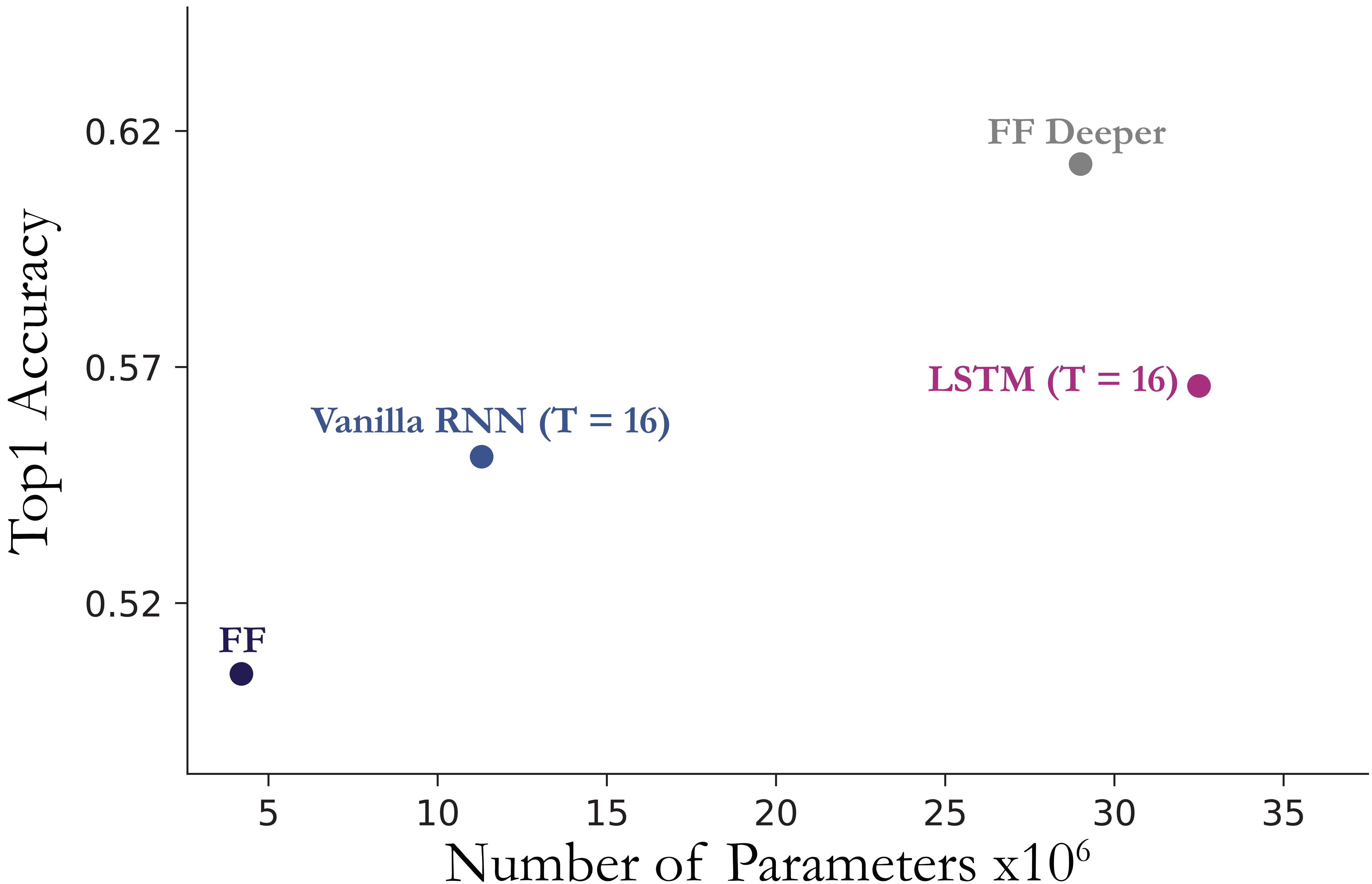
Two complementary principles:

(1) gating = multiplication by input-dependent tensor w/ values in [0, 1]

(2) bypassing = when recurrent cell is in 0 state, input is unchanged ("ResNet-like")

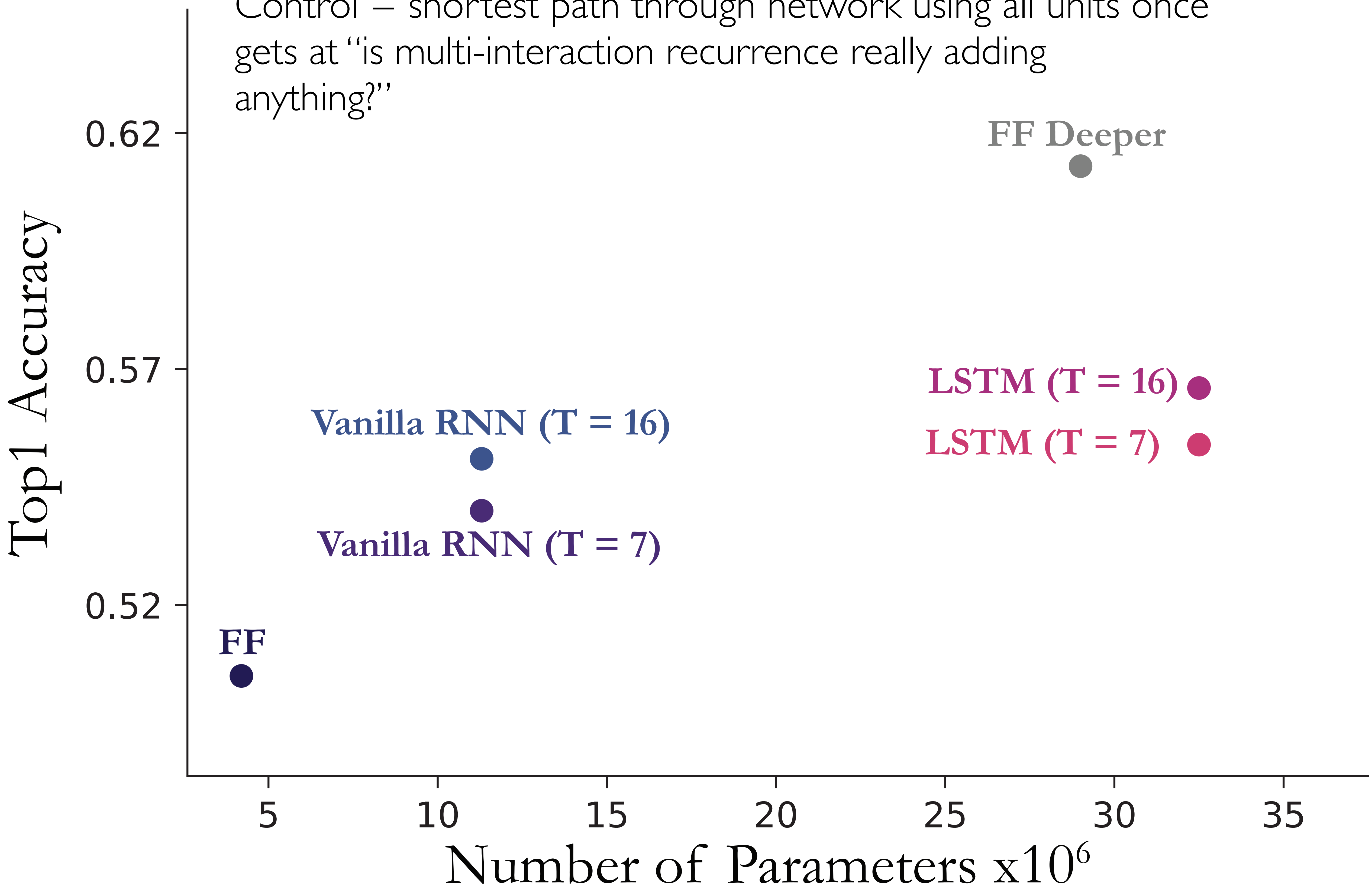
LSTM has **(1)** but not **(2)**; VanillaRNN has **(2)** but not **(1)**

Not All Local Recurrence is Equal

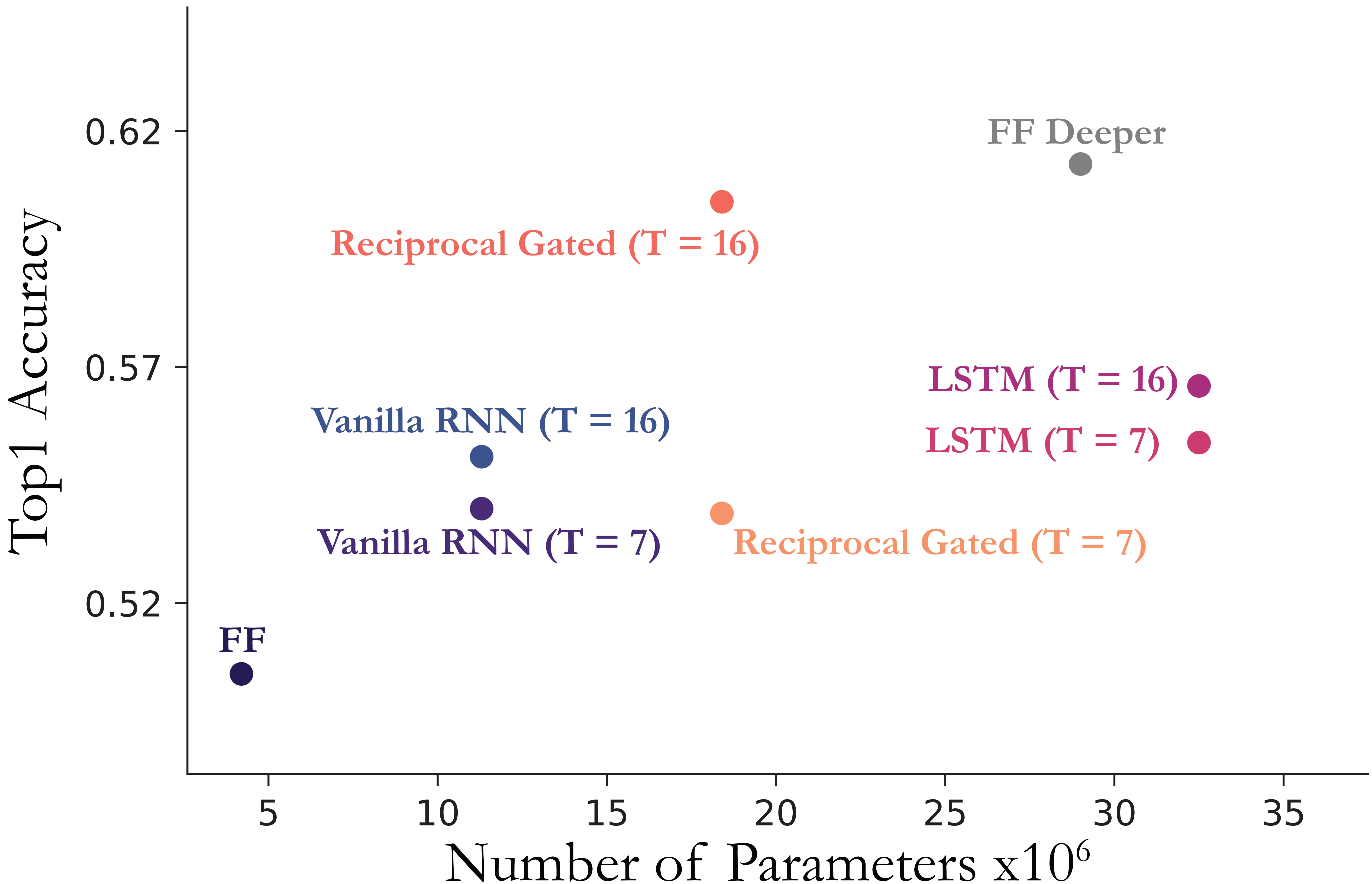


Not All Local Recurrence is Equal

Control = shortest path through network using all units once gets at "is multi-interaction recurrence really adding anything?"



Not All Local Recurrence is Equal



Large-Scale Search over Deep Recurrent Architectures

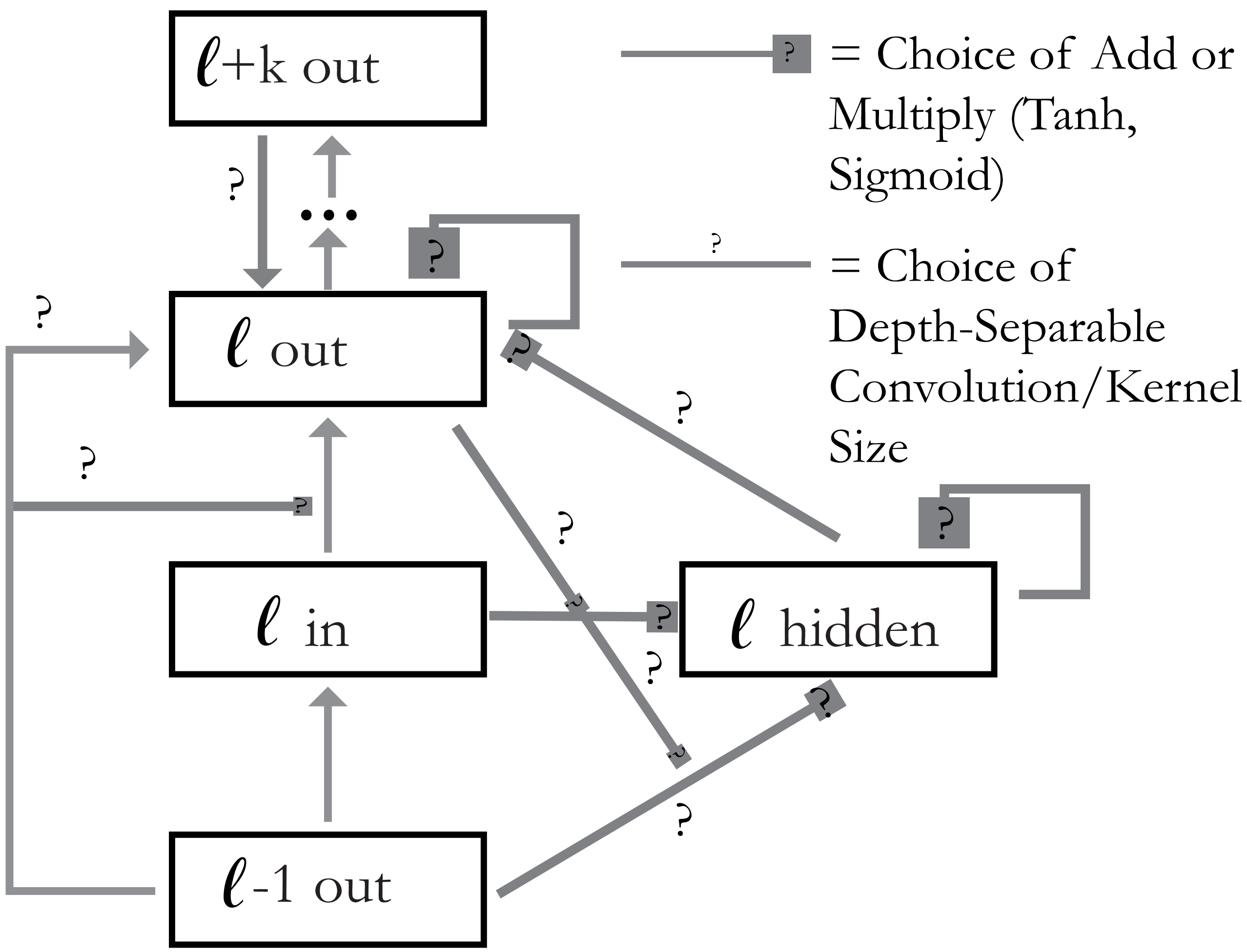
Nayebi, A.*, Bear, D.*, Kubilius, J.*,…, Yamins, D.L.K. *Task-Driven Convolutional Recurrent Models of the Visual System*, NeurIPS 2018

- Expand architecture class (local and global recurrence)
- Parametrize local and global feedback motifs and optimize for performance on ImageNet
- Evaluate neural predictivity in V4 and IT temporal responses

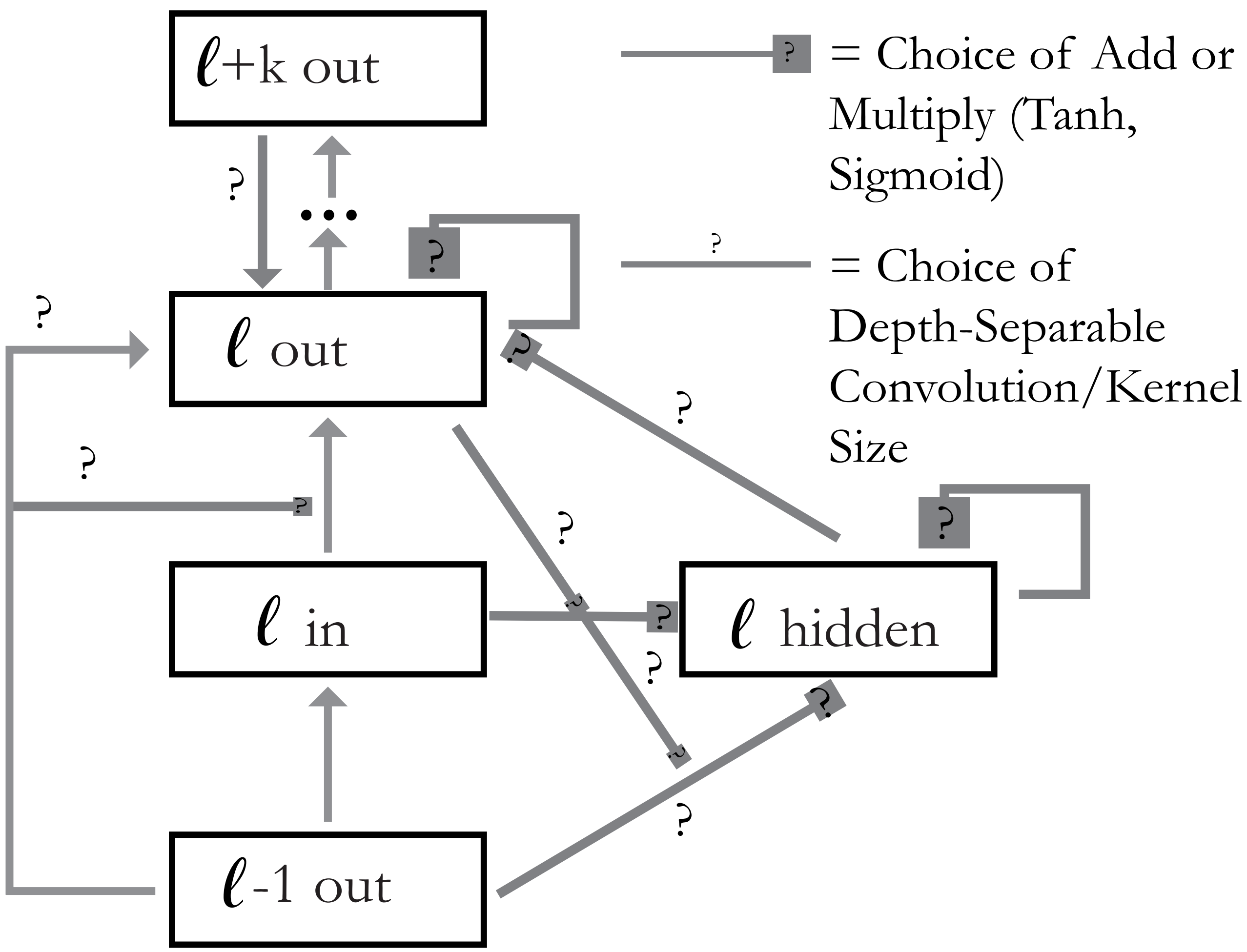
Daniel Bear
(Stanford)

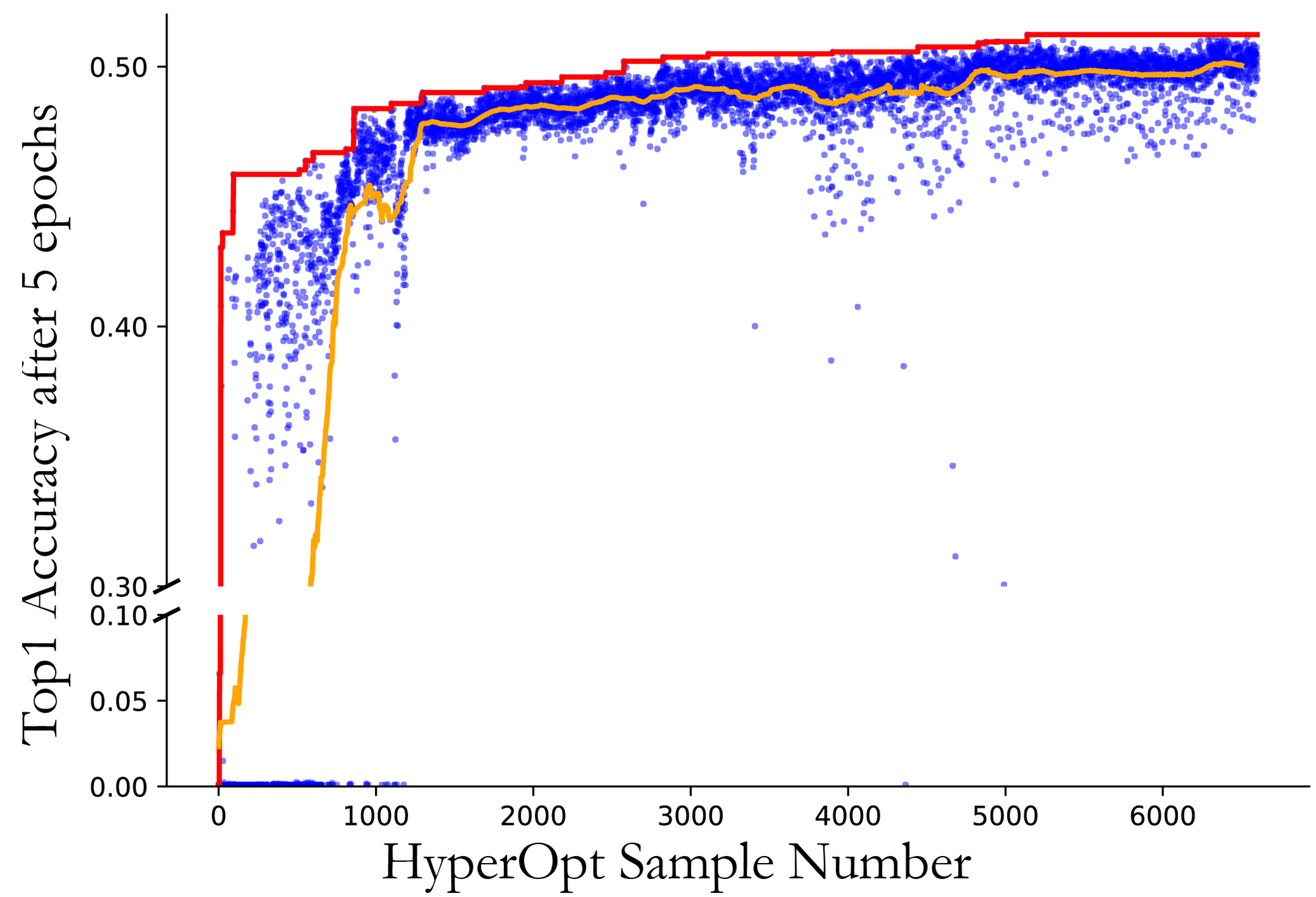
Jonas Kubilius
(MIT)

Search Over Local and Global Recurrence

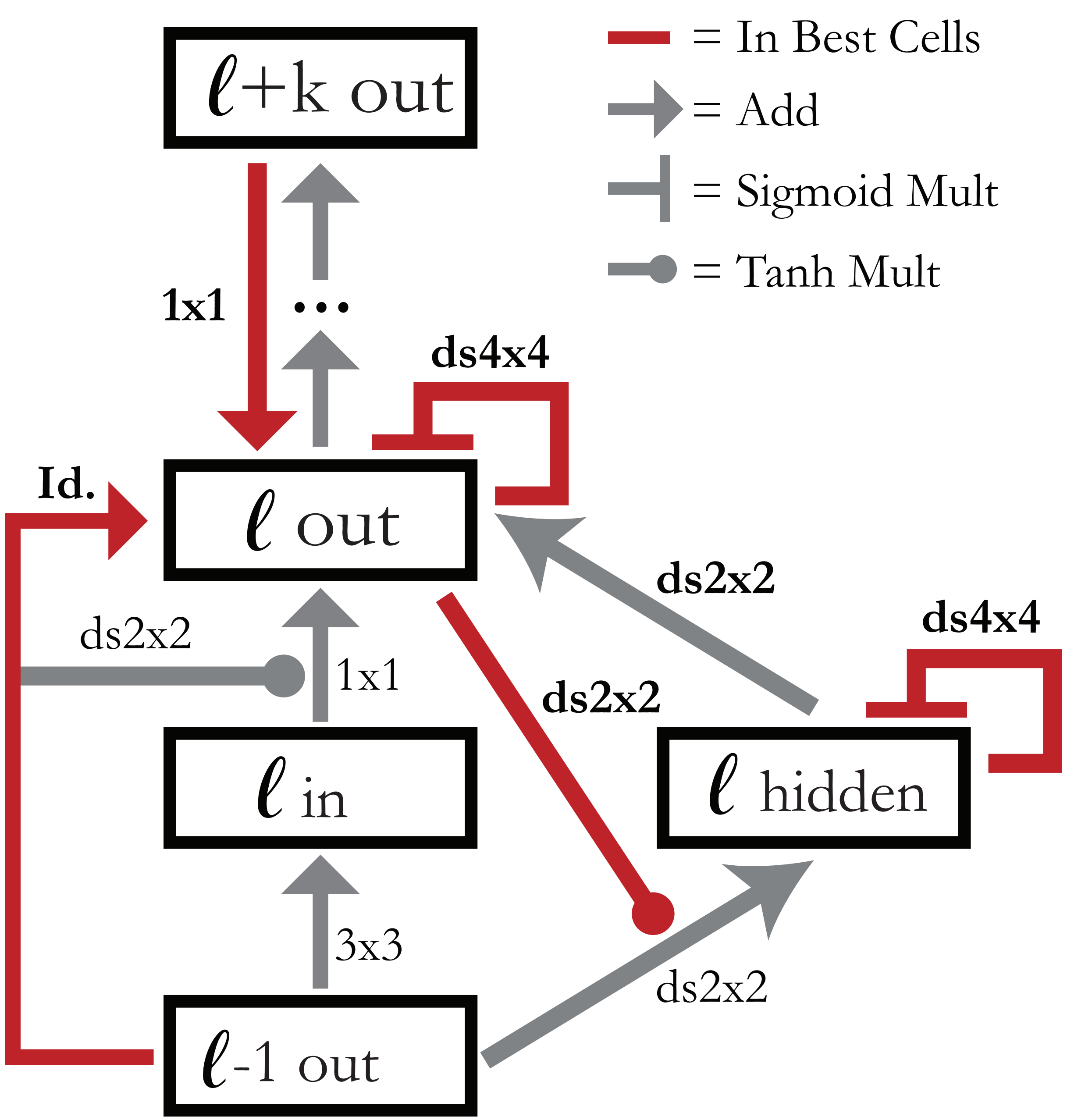


Search Over Local and Global Recurrence

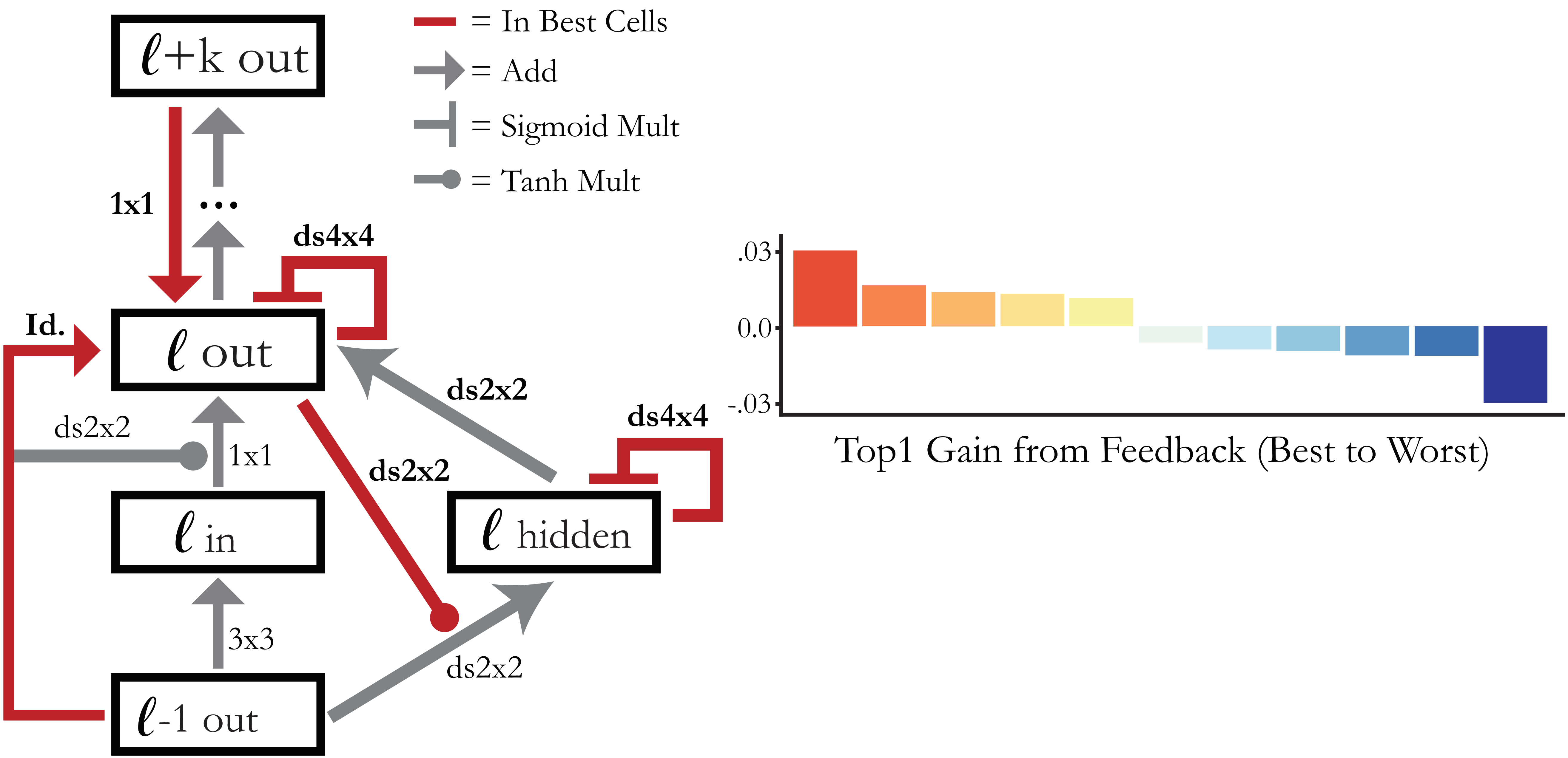




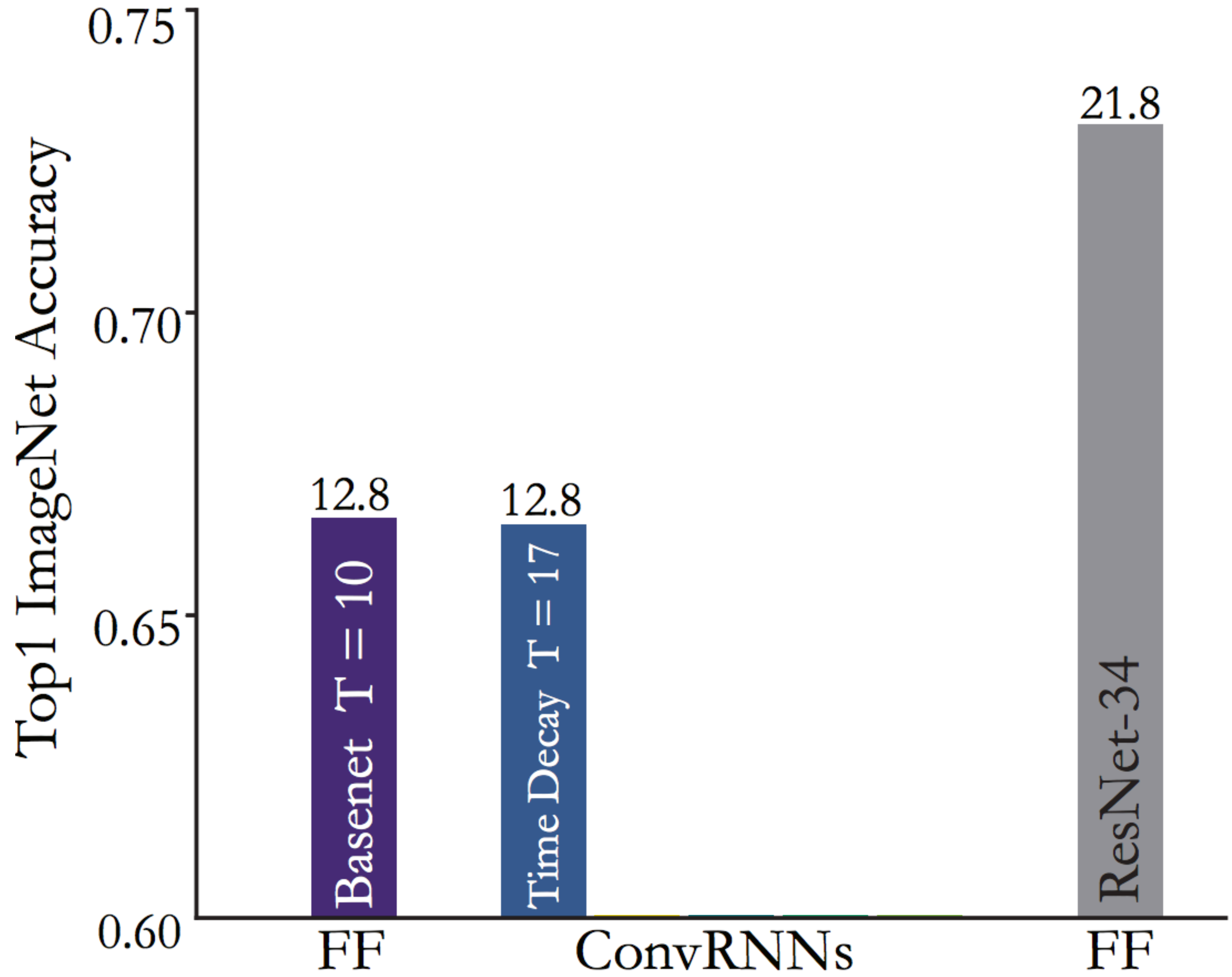
Emergent Local and Global Connectivity Patterns



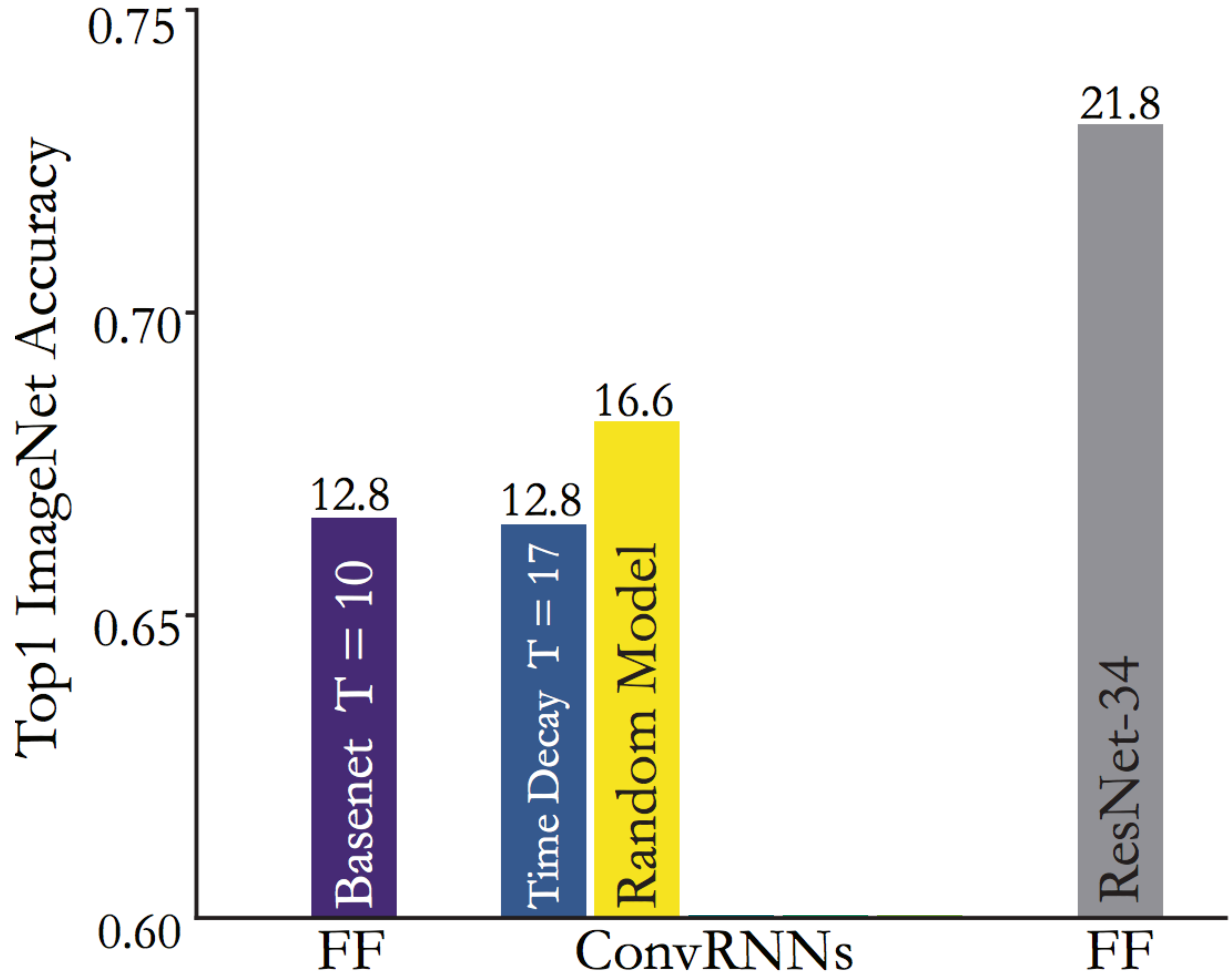
Emergent Local and Global Connectivity Patterns



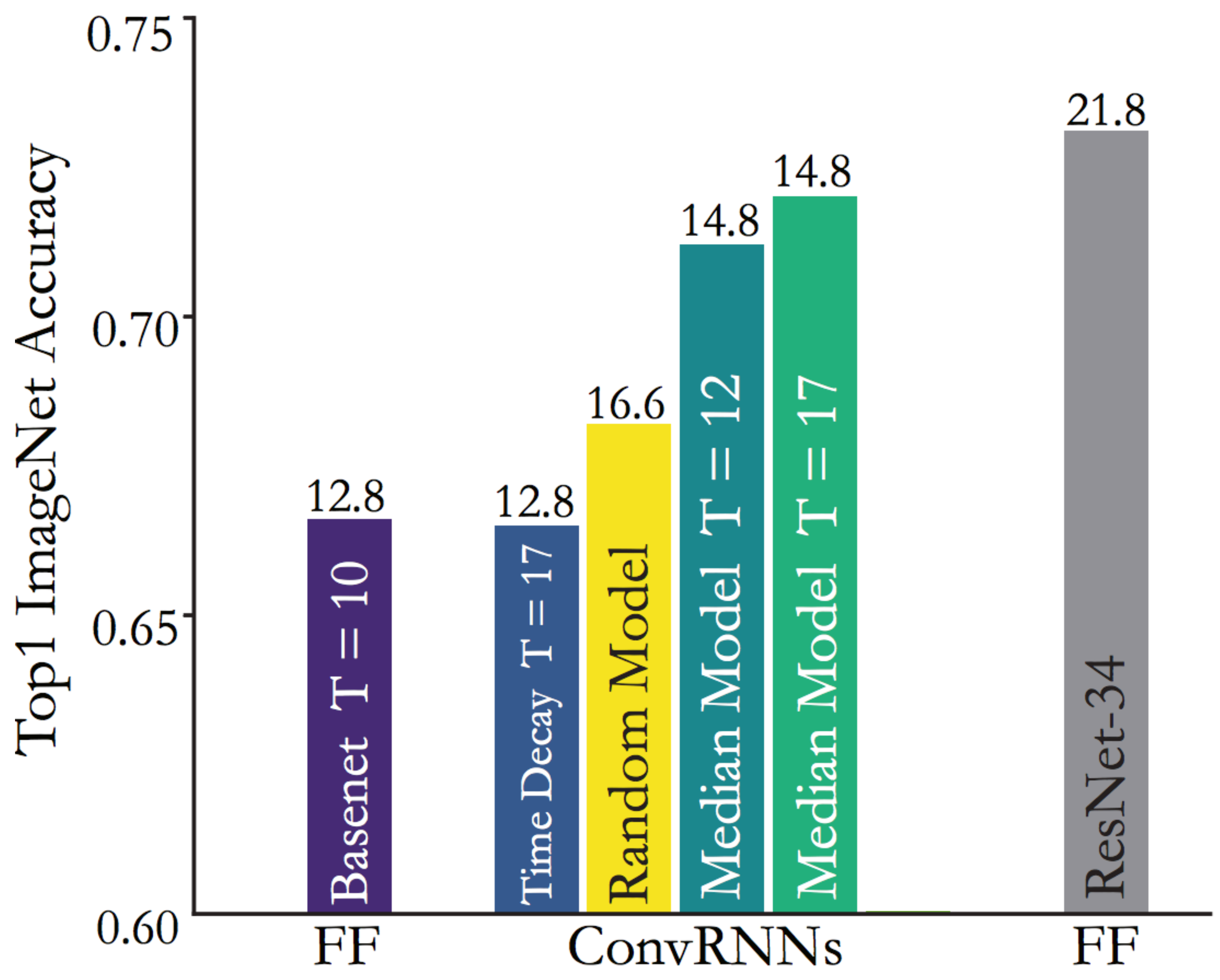
Improving ImageNet Performance with ConvRNNs



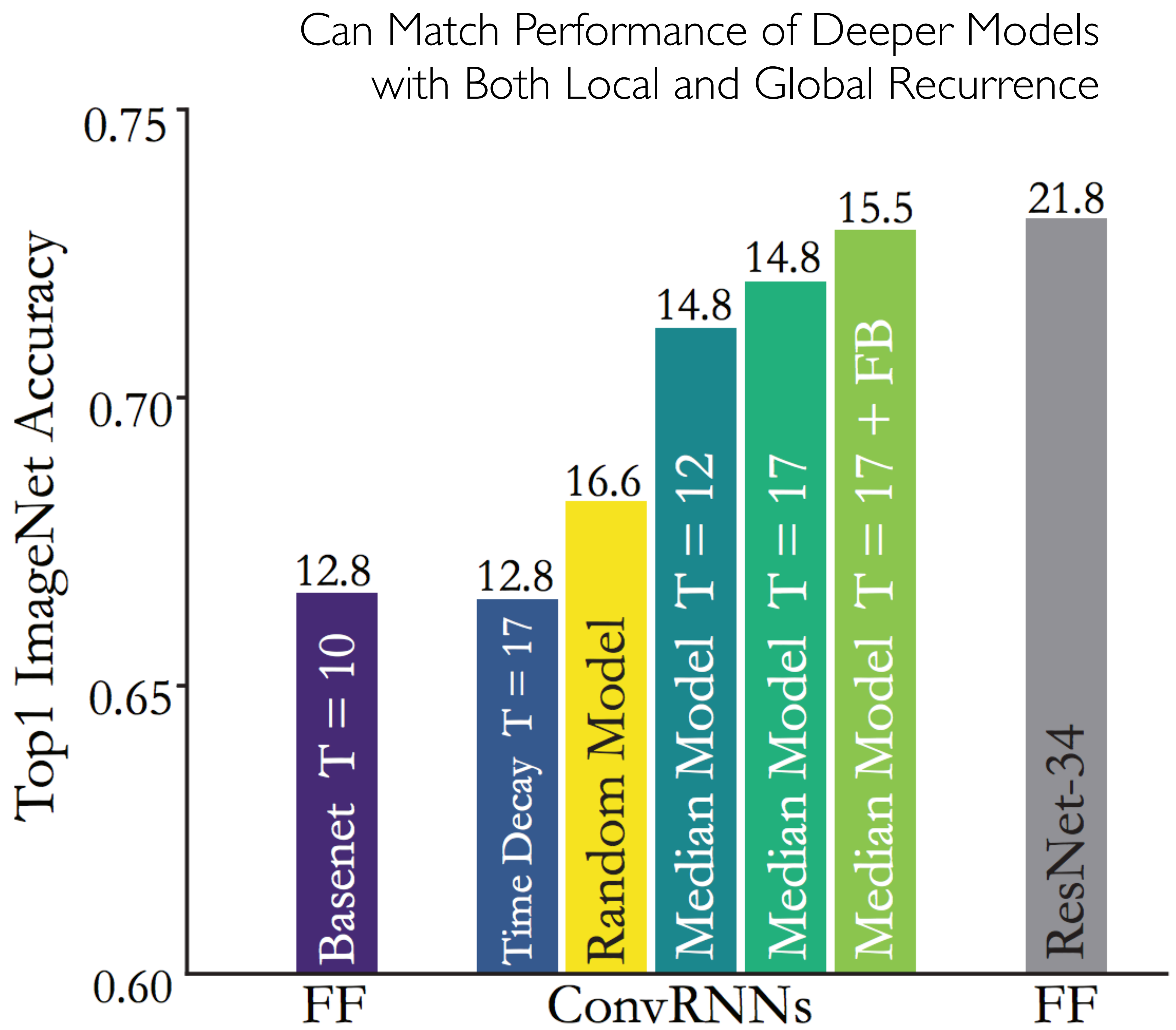
Improving ImageNet Performance with ConvRNNs



Improving ImageNet Performance with ConvRNNs



Improving ImageNet Performance with ConvRNNs



Role of Recurrence in Core Object Recognition

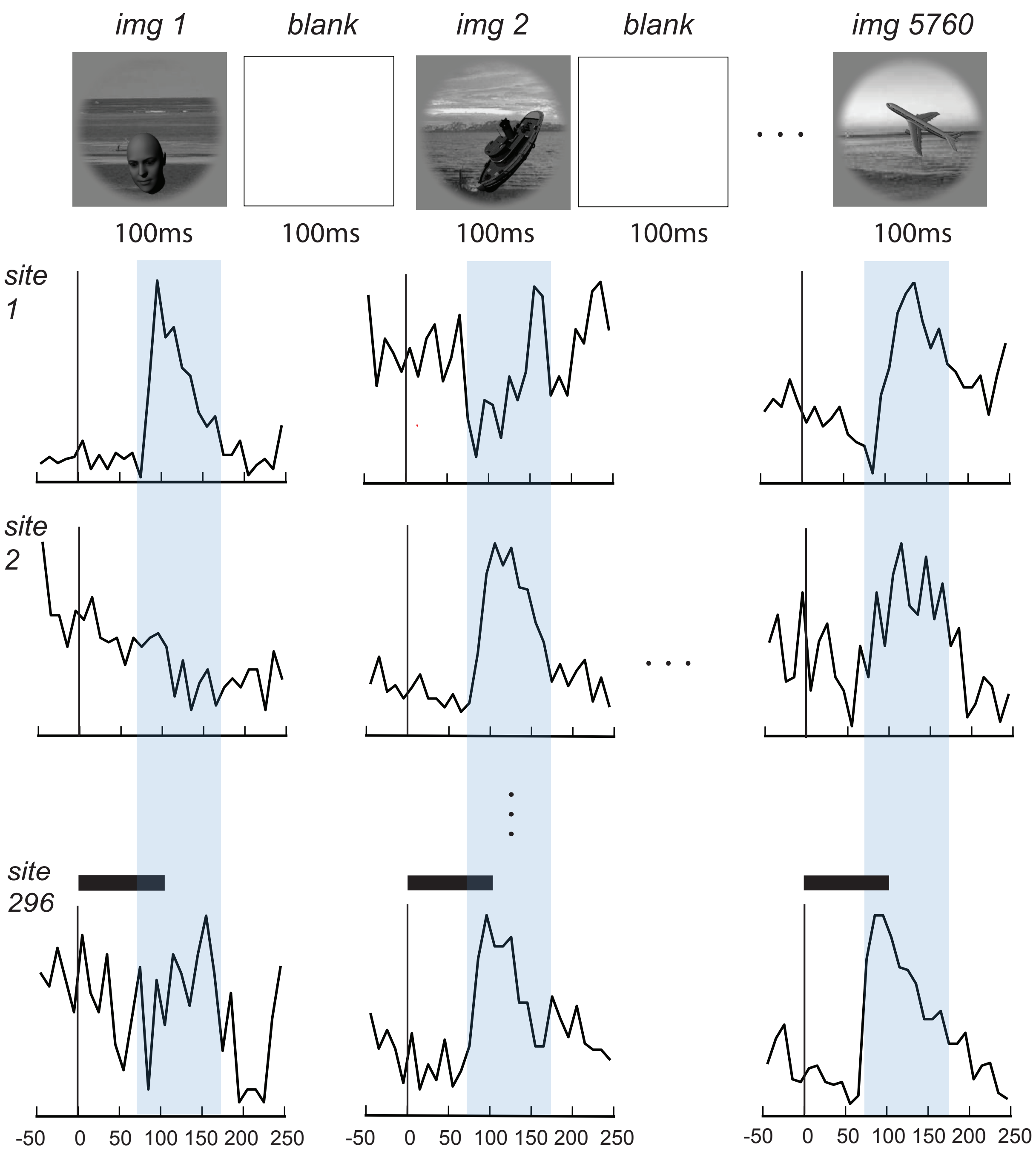
Nayebi, A.*, Bear, D.*, Kubilius, J.*,…, Yamins, D.L.K. *Task-Driven Convolutional Recurrent Models of the Visual System*, NeurIPS 2018

- Expand architecture class (local and global recurrence)
- Parametrize local and global feedback motifs and optimize for performance on ImageNet
- Evaluate neural predictivity in V4 and IT temporal responses

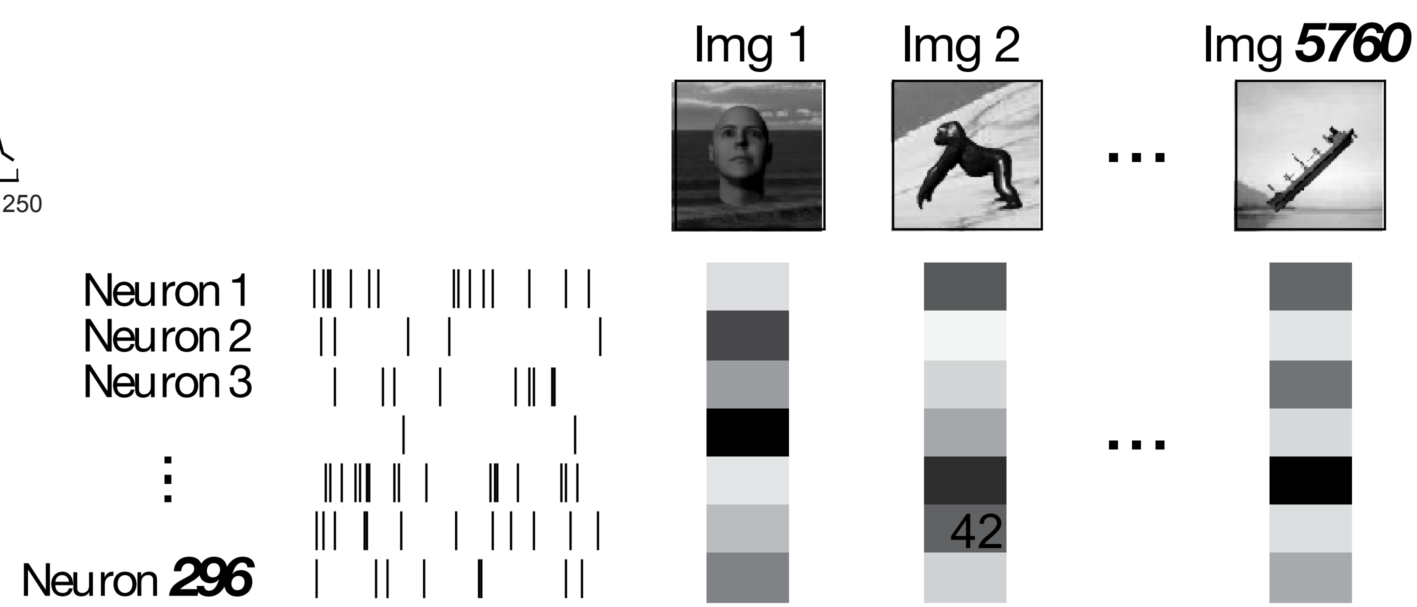
Daniel Bear
(Stanford)

Jonas Kubilius
(MIT)

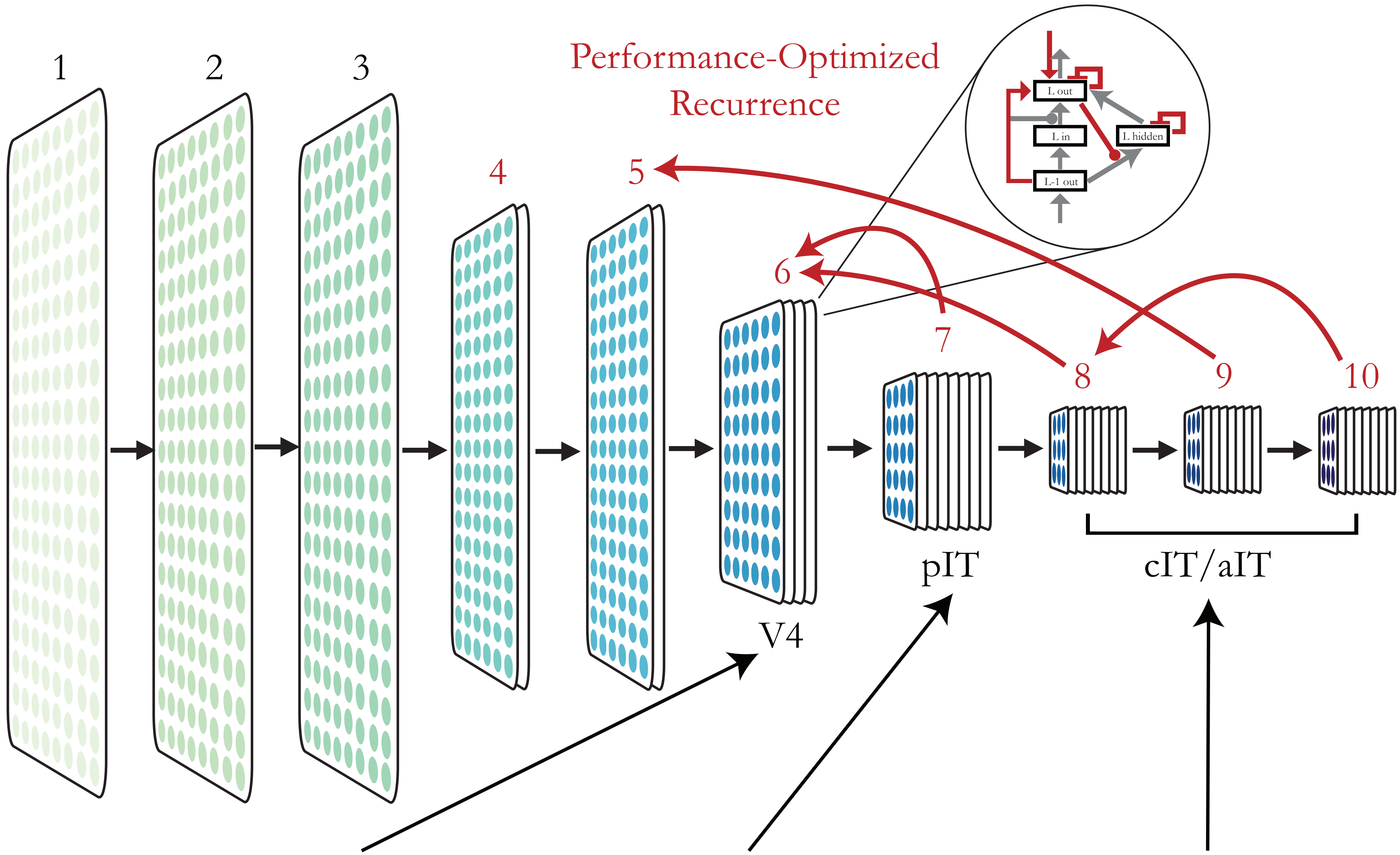
Neural Predictivity with ConvRNNs



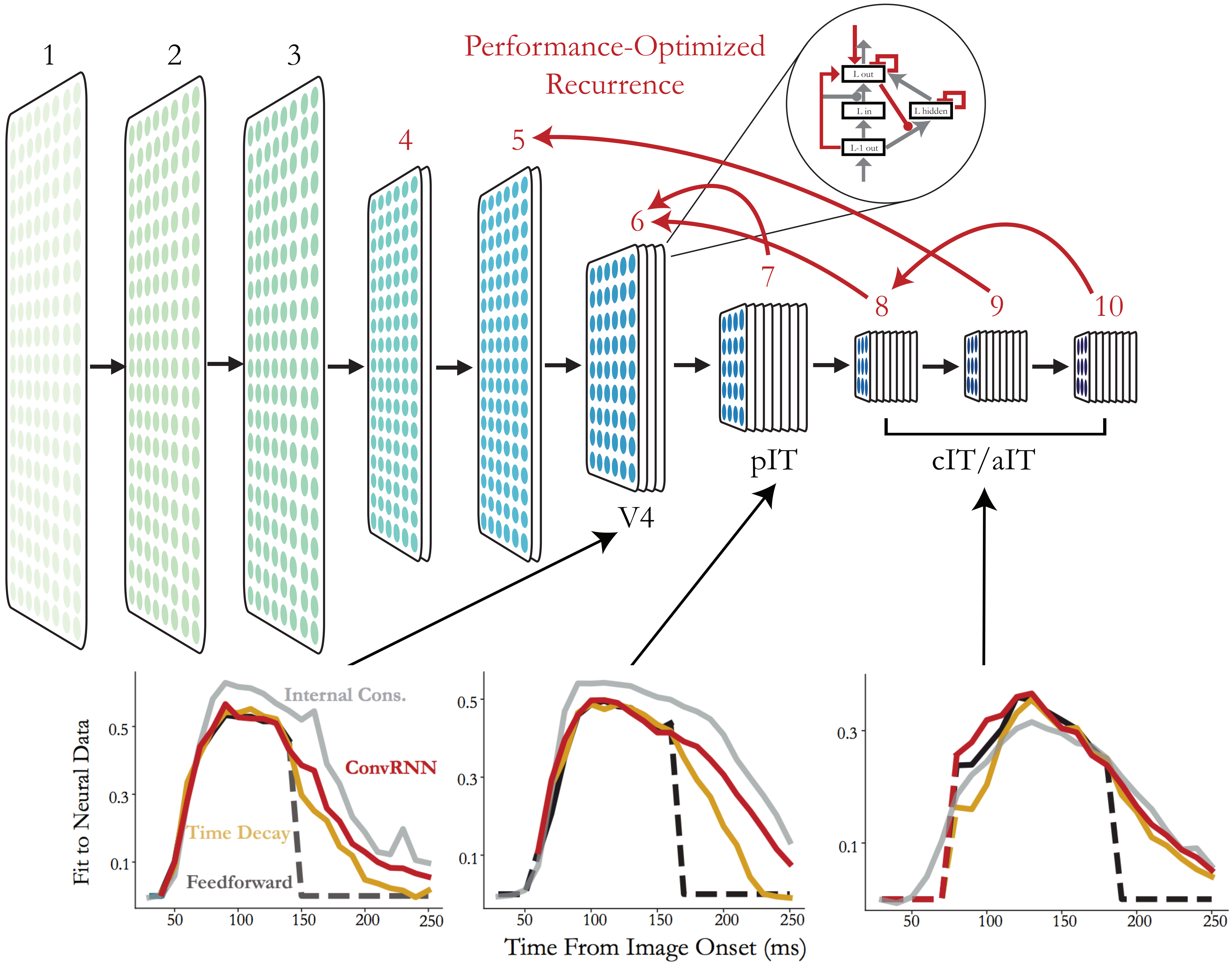
Binned spike counts 70ms-170ms post stimulus presentation



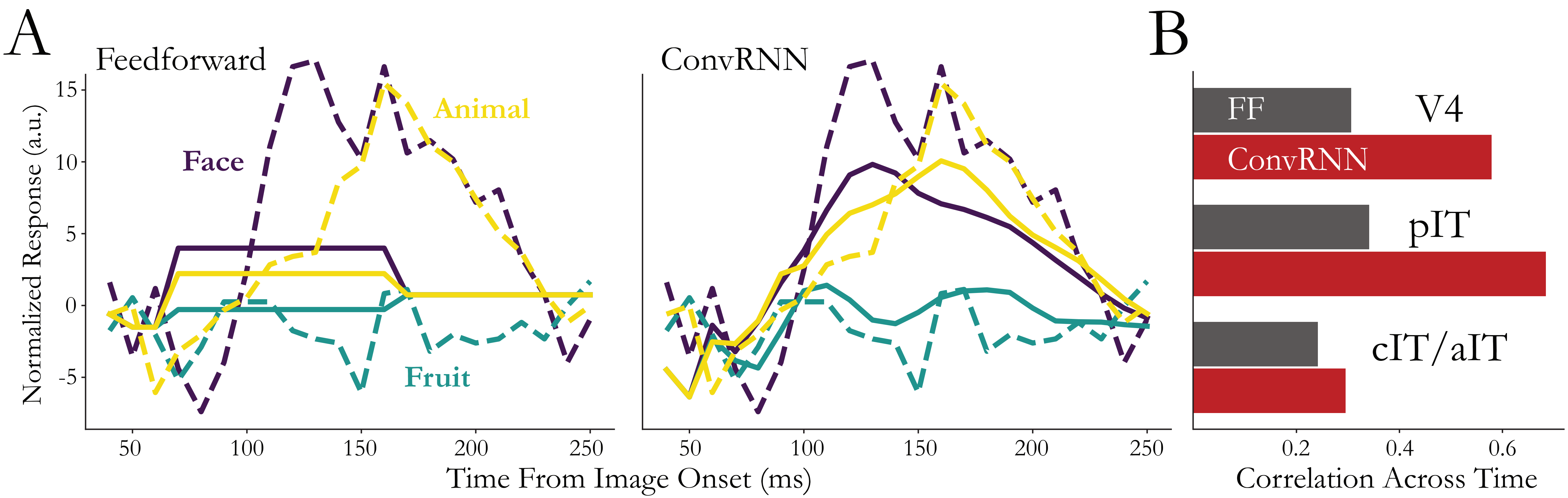
Temporal Neural Fit with ConvRNNs



Temporal Neural Fit with ConvRNNs



Improved Neural Fit with ConvRNNs



Takeaways

AI/Performance **Neuroscience**

Takeaways

AI/Performance **Neuroscience**

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance
  - *With proper local recurrence in place, specific patterns of long-range feedback connections further improve performance*

Takeaways

AI/Performance **Neuroscience**

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance
  - *With proper local recurrence in place, specific patterns of long-range feedback connections further improve performance*

- These performance-optimized dynamics provide strong estimates of neural and behavioral metrics in the primate ventral stream over feedforward & other recurrent models

Takeaways

AI/Performance **Neuroscience**

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance
  - *With proper local recurrence in place, specific patterns of long-range feedback connections further improve performance*

- These performance-optimized dynamics provide strong estimates of neural and behavioral metrics in the primate ventral stream over feedforward & other recurrent models

- These new convolutional recurrent architectures can be applied to many computer vision tasks (segmentation, movie prediction) without much modification

Takeaways **AI/Performance** **Neuroscience**

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance
  - *With proper local recurrence in place, specific patterns of long-range feedback connections further improve performance*

- These performance-optimized dynamics provide strong estimates of neural and behavioral metrics in the primate ventral stream over feedforward & other recurrent models

- These new convolutional recurrent architectures can be applied to many computer vision tasks (segmentation, movie prediction) without much modification

- We can use these models to explore a variety of normative questions across the entirety of the ventral stream (V1,V4, dorsal stream)

Current Approaches:

- Goal-Driven Modeling
- Direct Fits to Data

Current Approaches:

- Goal-Driven Modeling
- Direct Fits to Data

Current Approaches:

- Goal-Driven Modeling

- Direct Fits to Data

What can we learn when the behavioral goal is unclear but we are not data limited?

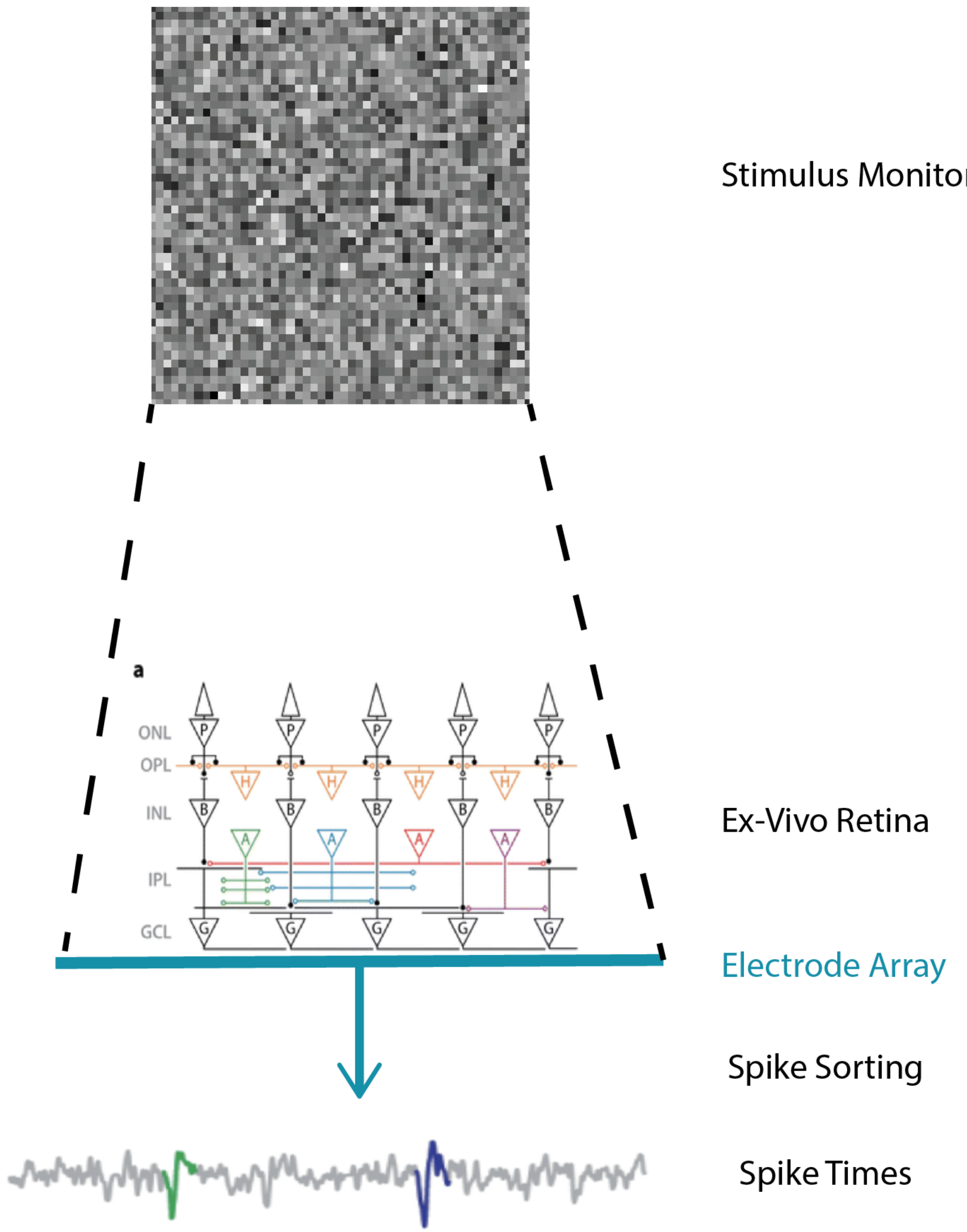
Roadmap

Current Approaches:

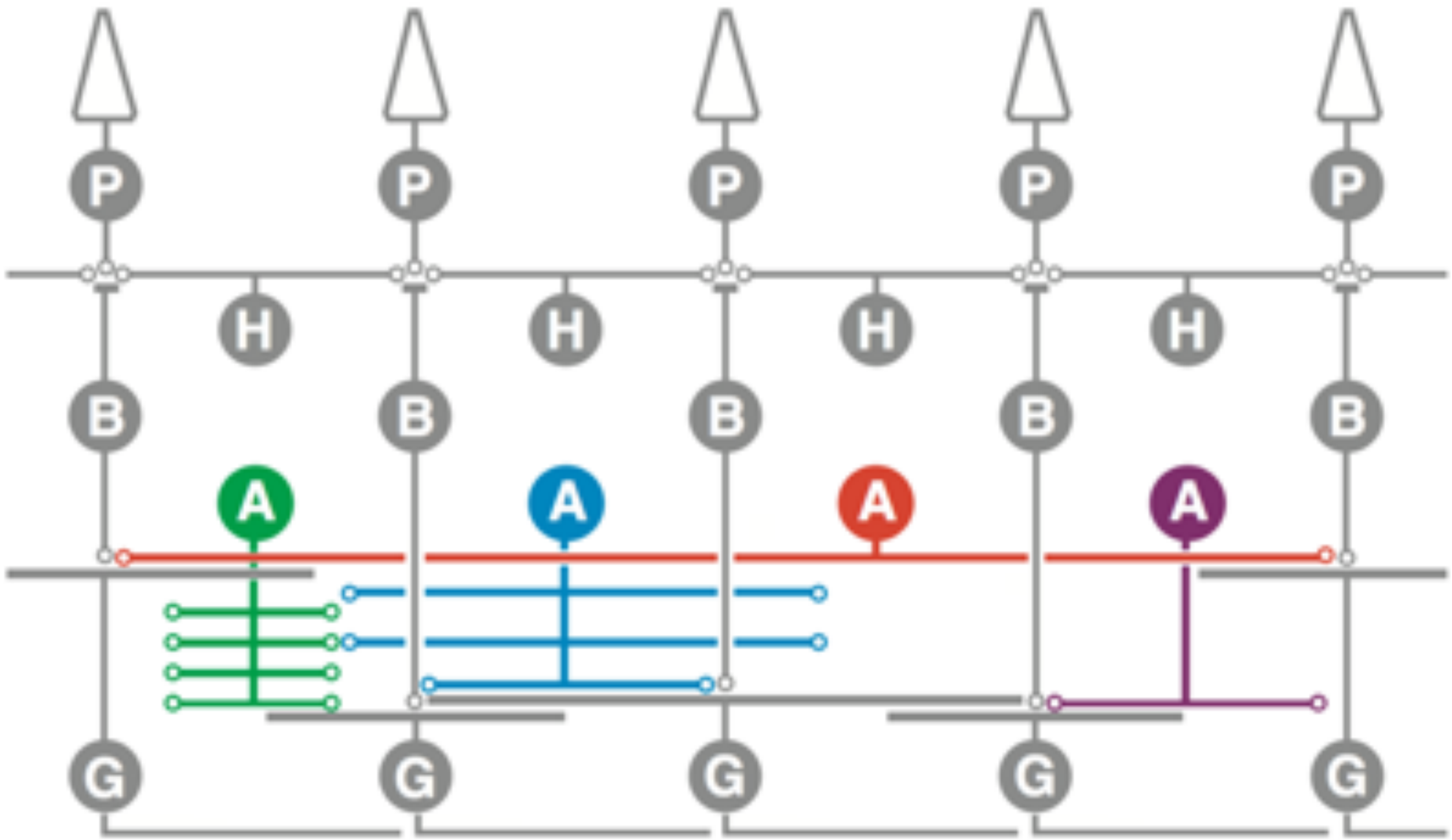
- Goal-Driven Modeling

- Direct Fits to Data

What can we learn when the behavioral goal is unclear but we are not data limited?

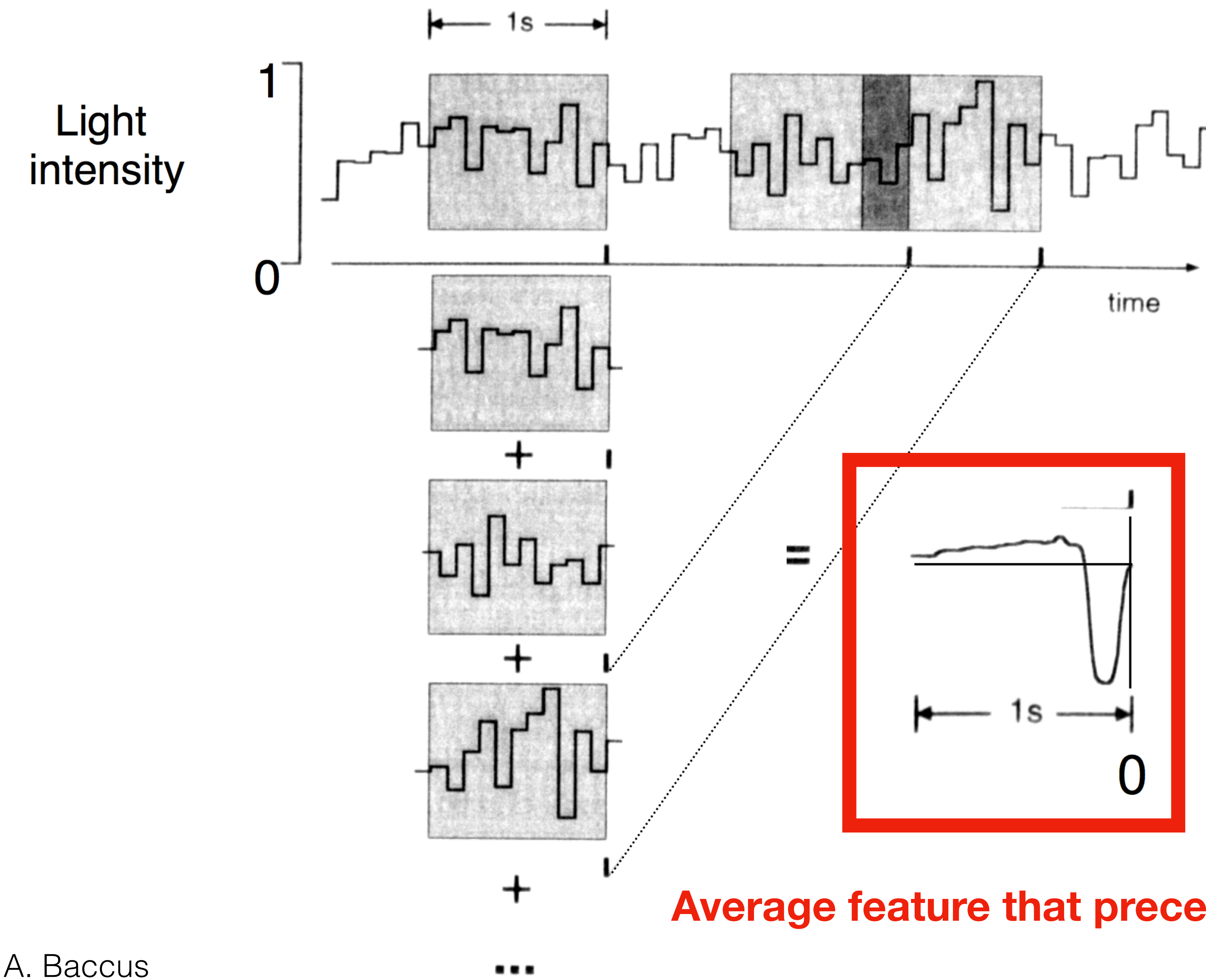


Retinal Circuitry



Credit: S. A. Baccus

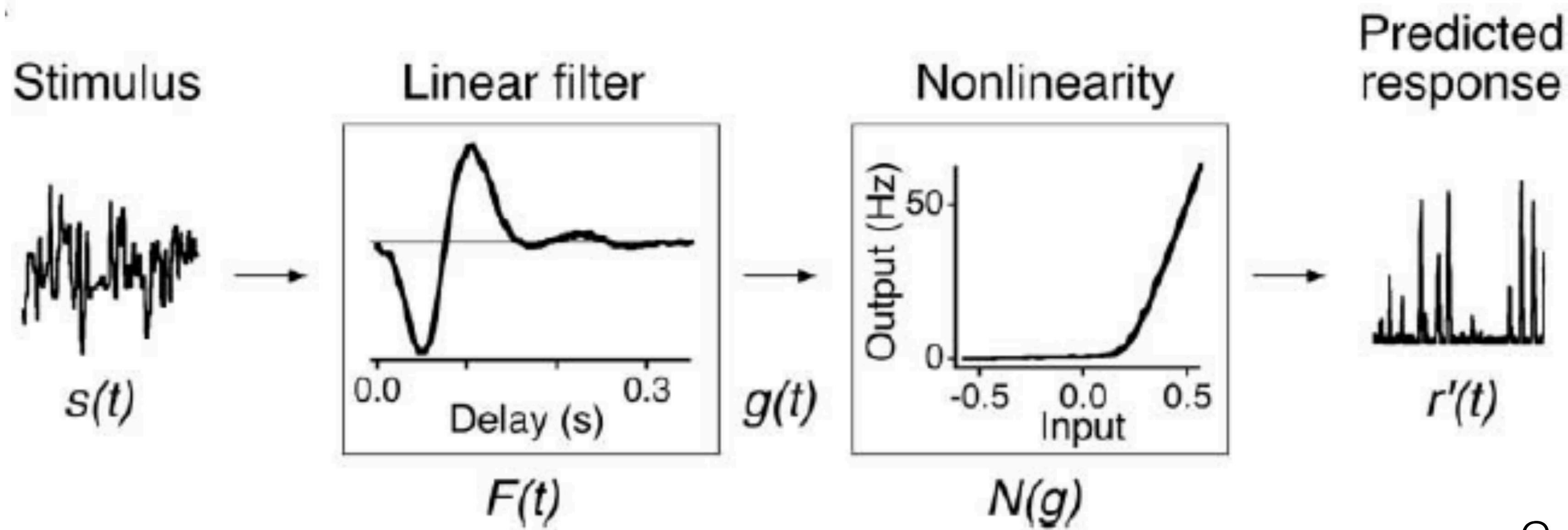
Traditional models of sensory encoding



Average feature that precedes a spike

Credit: S. A. Baccus

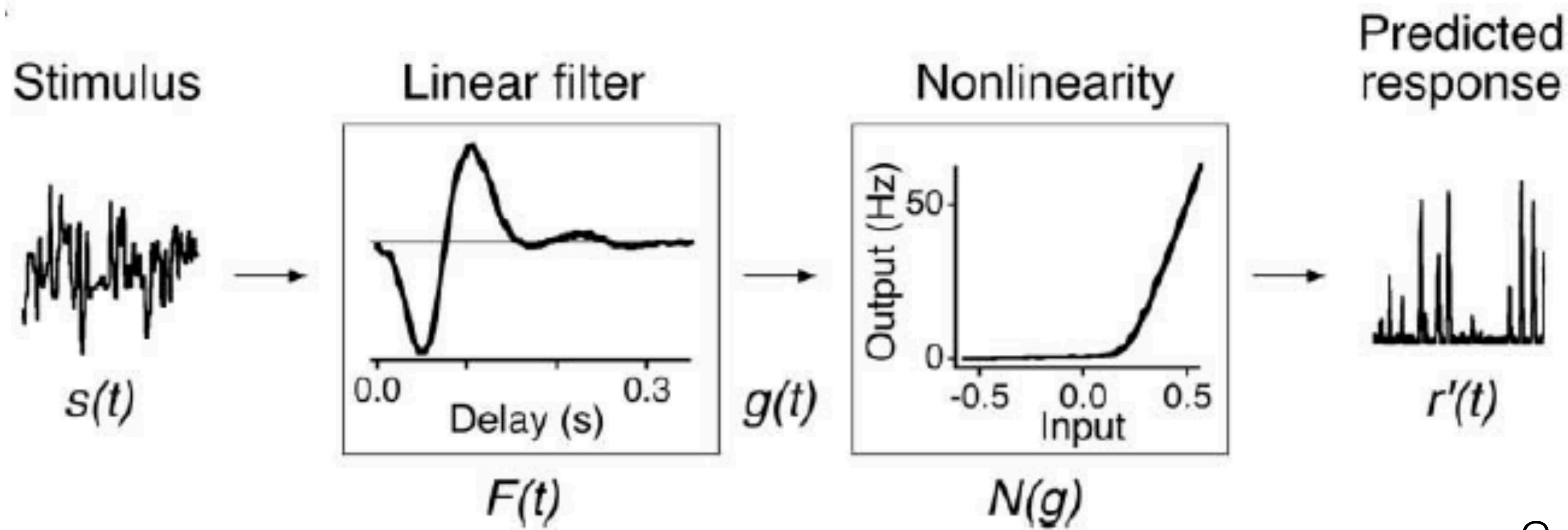
Drawbacks of LN Model



Credit: S. A. Baccus

Ignores multiple cell types in a single "layer"

Drawbacks of LN Model

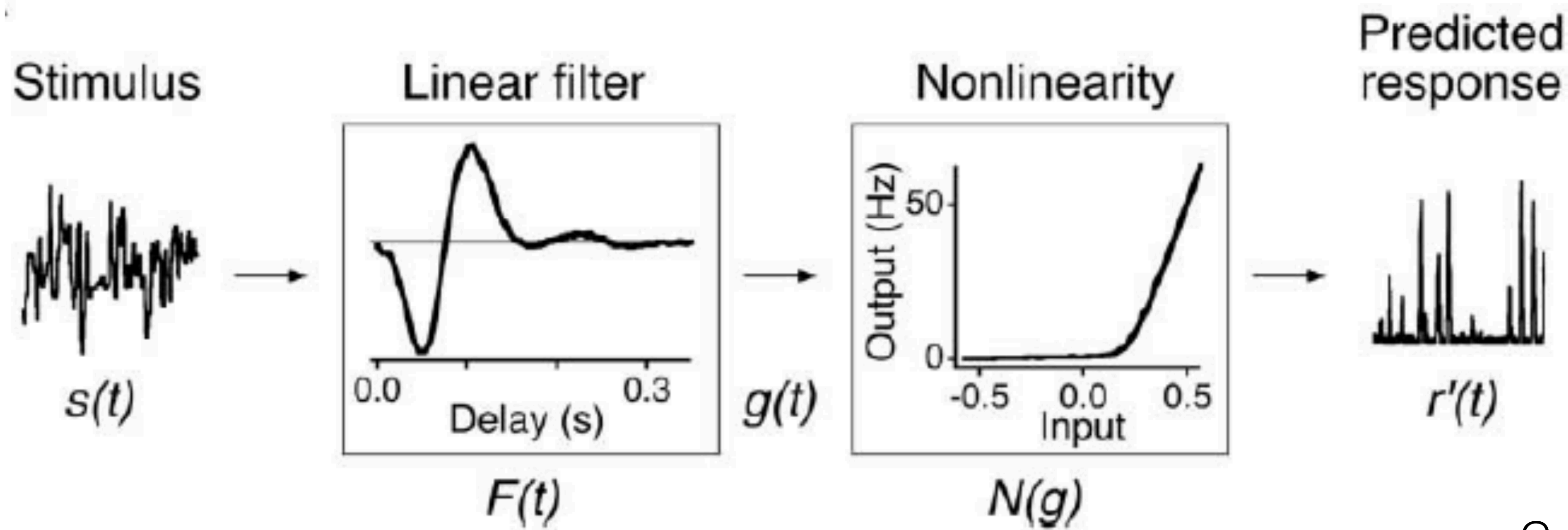


Credit: S. A. Baccus

Ignores multiple cell types in a single "layer"

Lack of multi-stage processing

Drawbacks of LN Model



Credit: S. A. Baccus

Ignores multiple cell types in a single "layer"

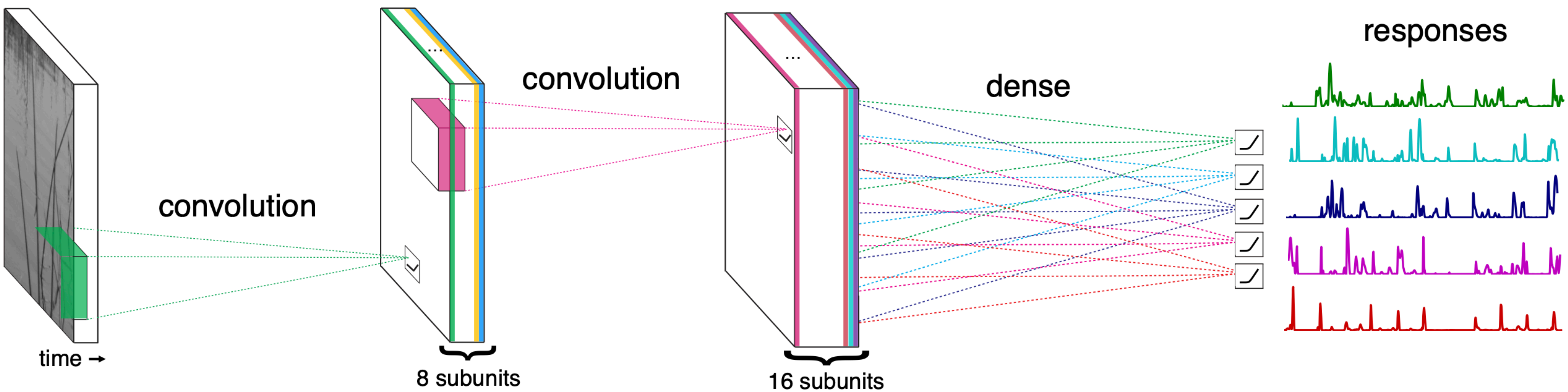
Lack of multi-stage processing

How much do these matter?

CNN Models of the Retina

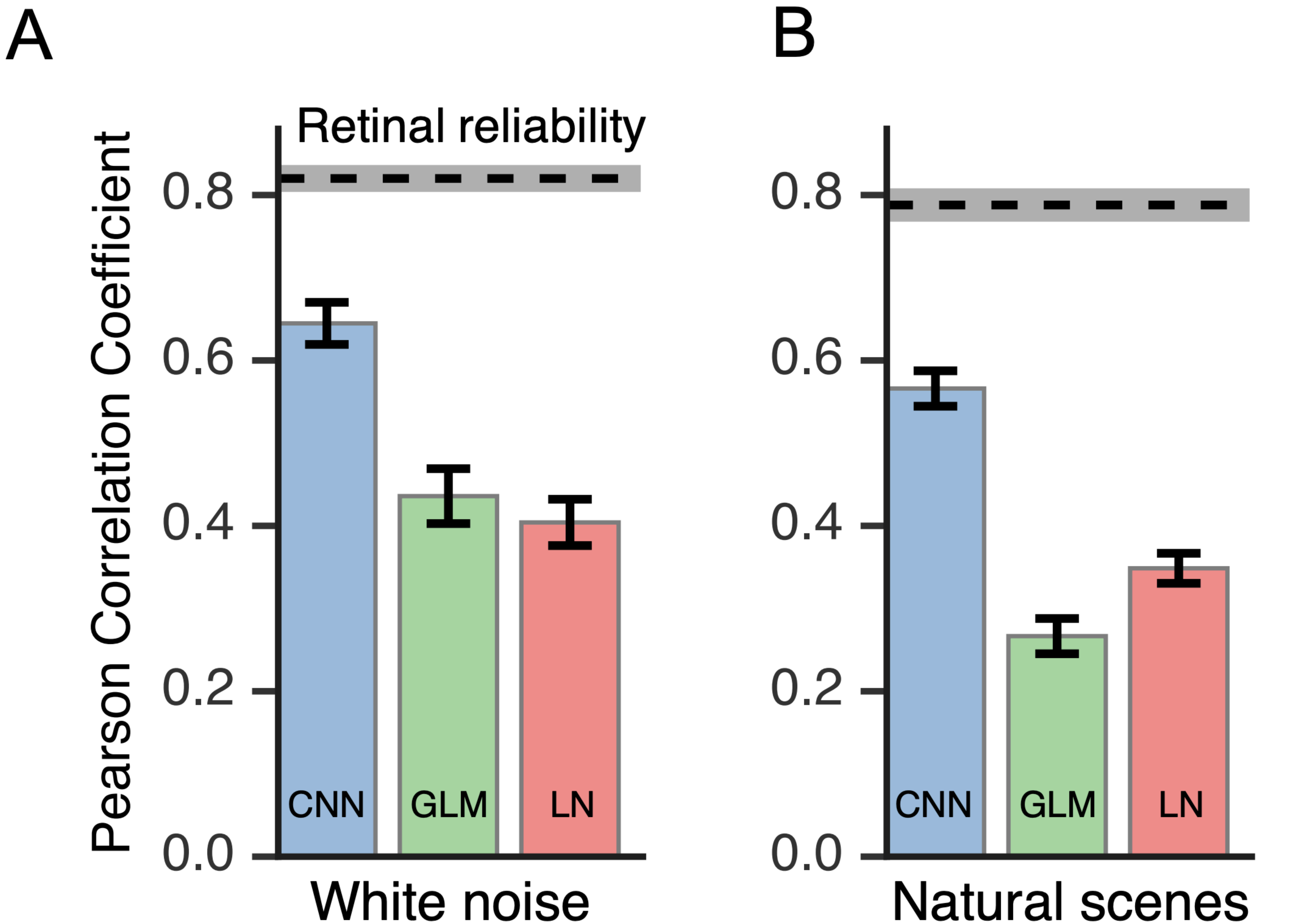
Lane McIntosh
(Stanford)

Niru Maheswaranathan
(Stanford)



McIntosh, L.*, Maheswaranathan, N.*, Nayebi, A., Baccus, S. *Deep Learning Models of the Retinal Response to Natural Scenes,* NIPS 2016

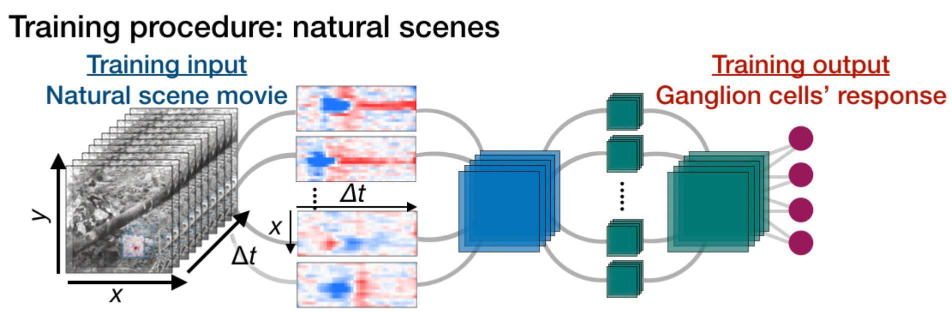
Held-Out Performance

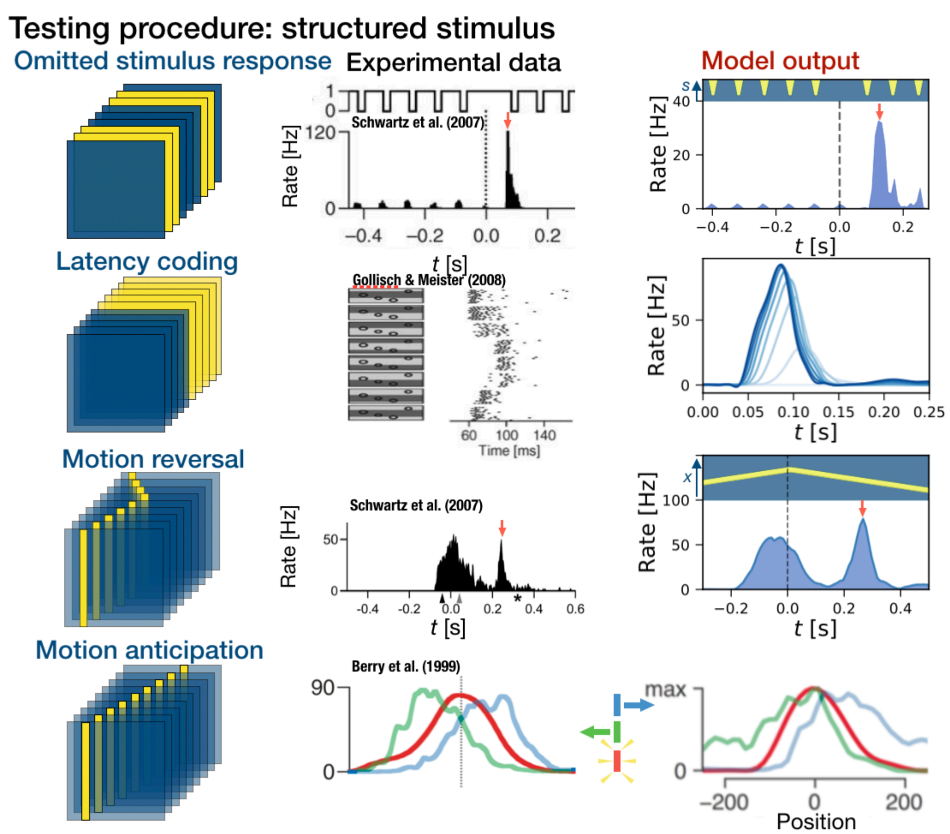


Are the deep network's "*computational mechanisms*" for generating neural responses the same as those in the brain?

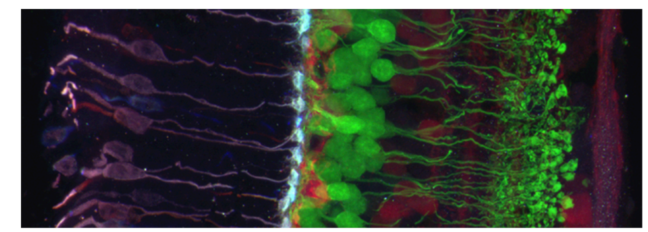
Can we bridge the Gap?

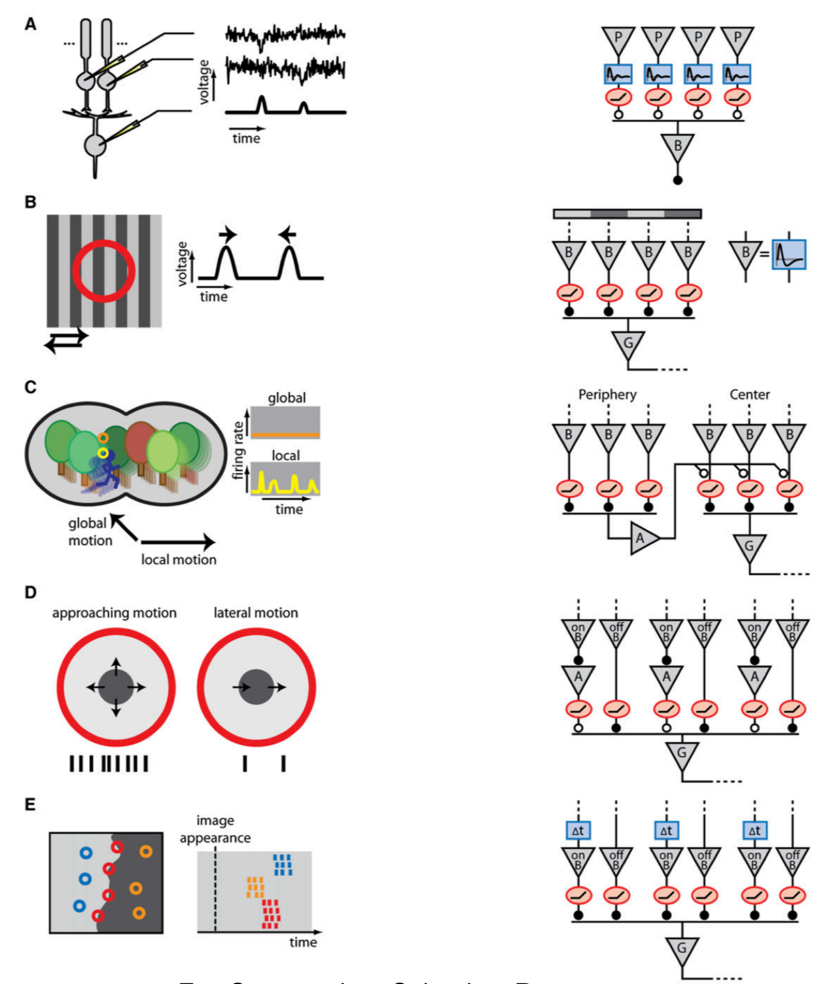
Deep learning model





Scientific hypotheses in neuroscience





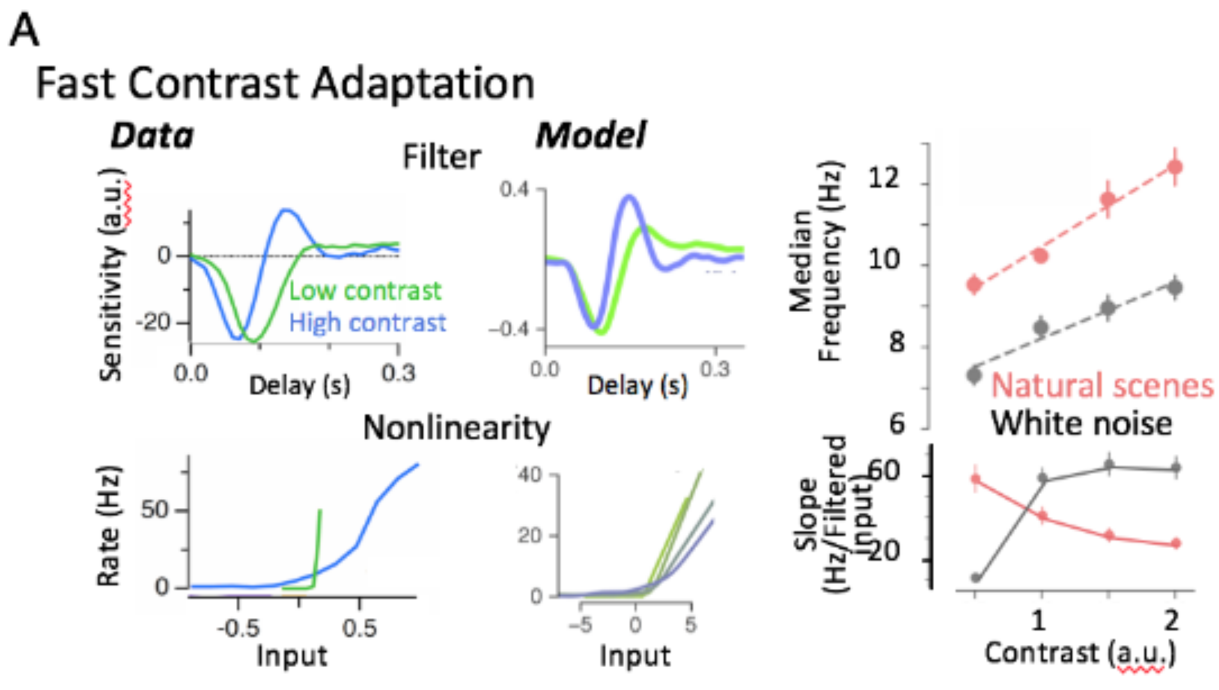
Eye Smarter than Scientists Believed.
Tim Gollisch, Markus Meister (2009)

Credit: H. Tanaka

Unified Mechanistic Model

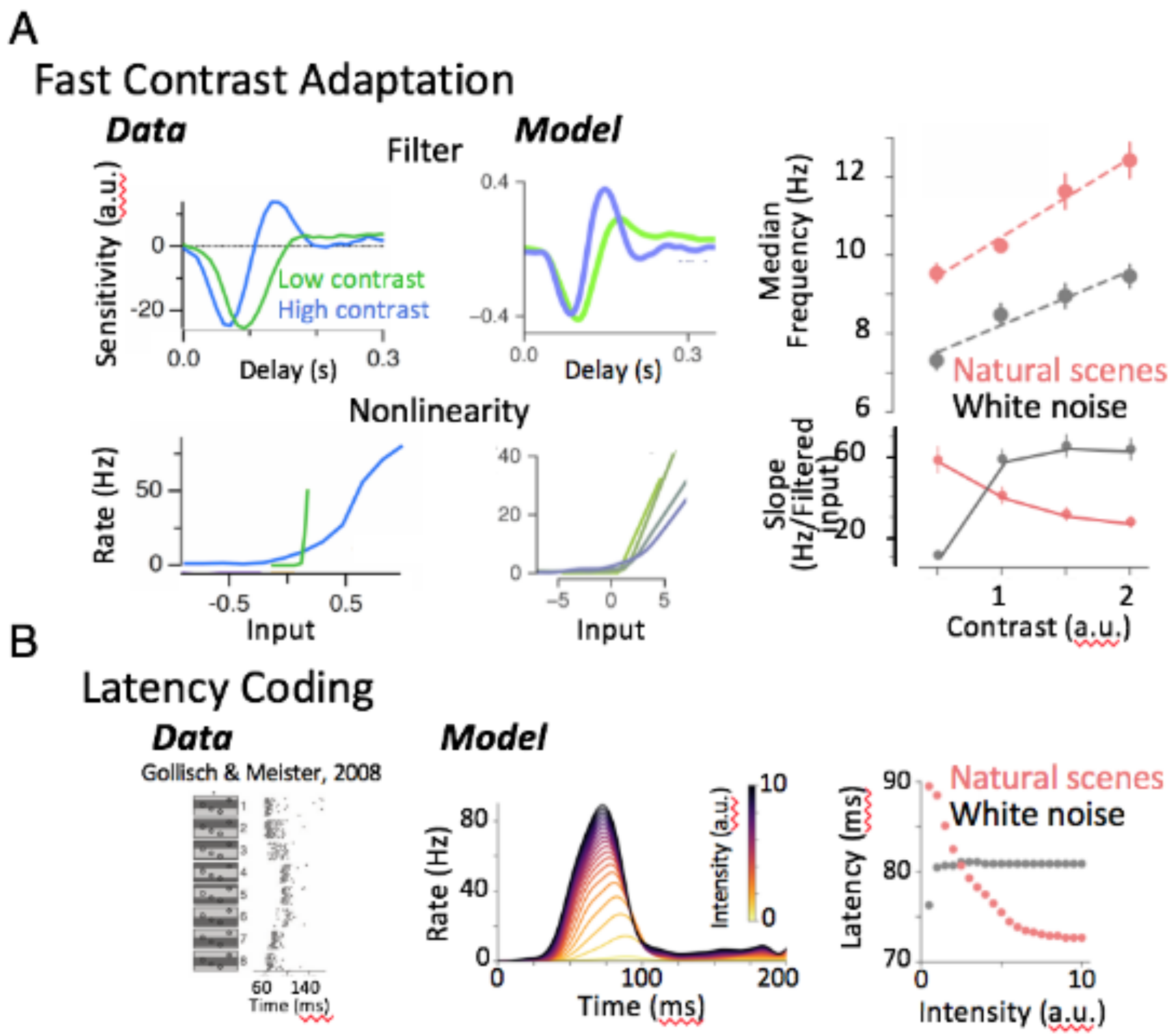
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



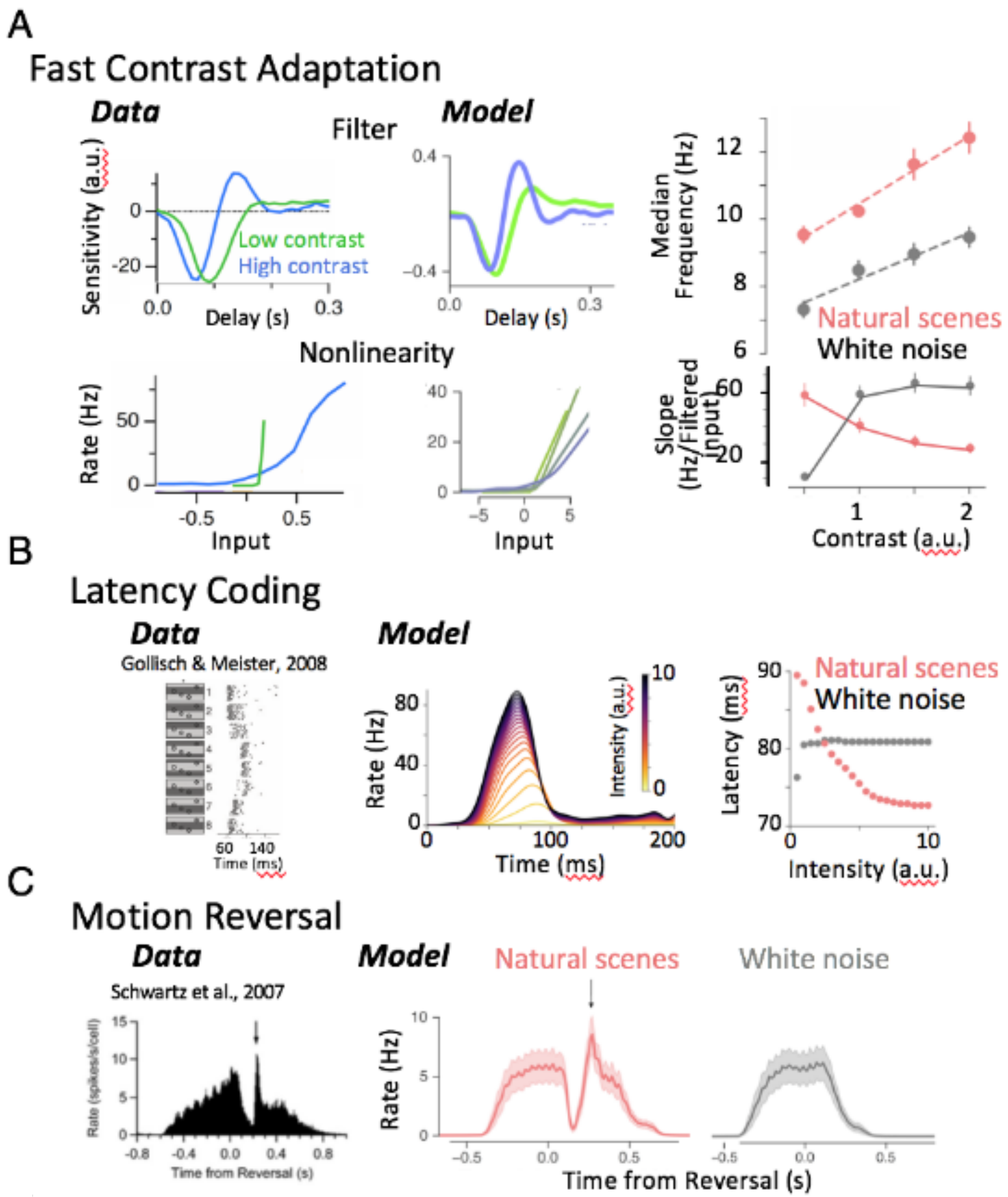
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



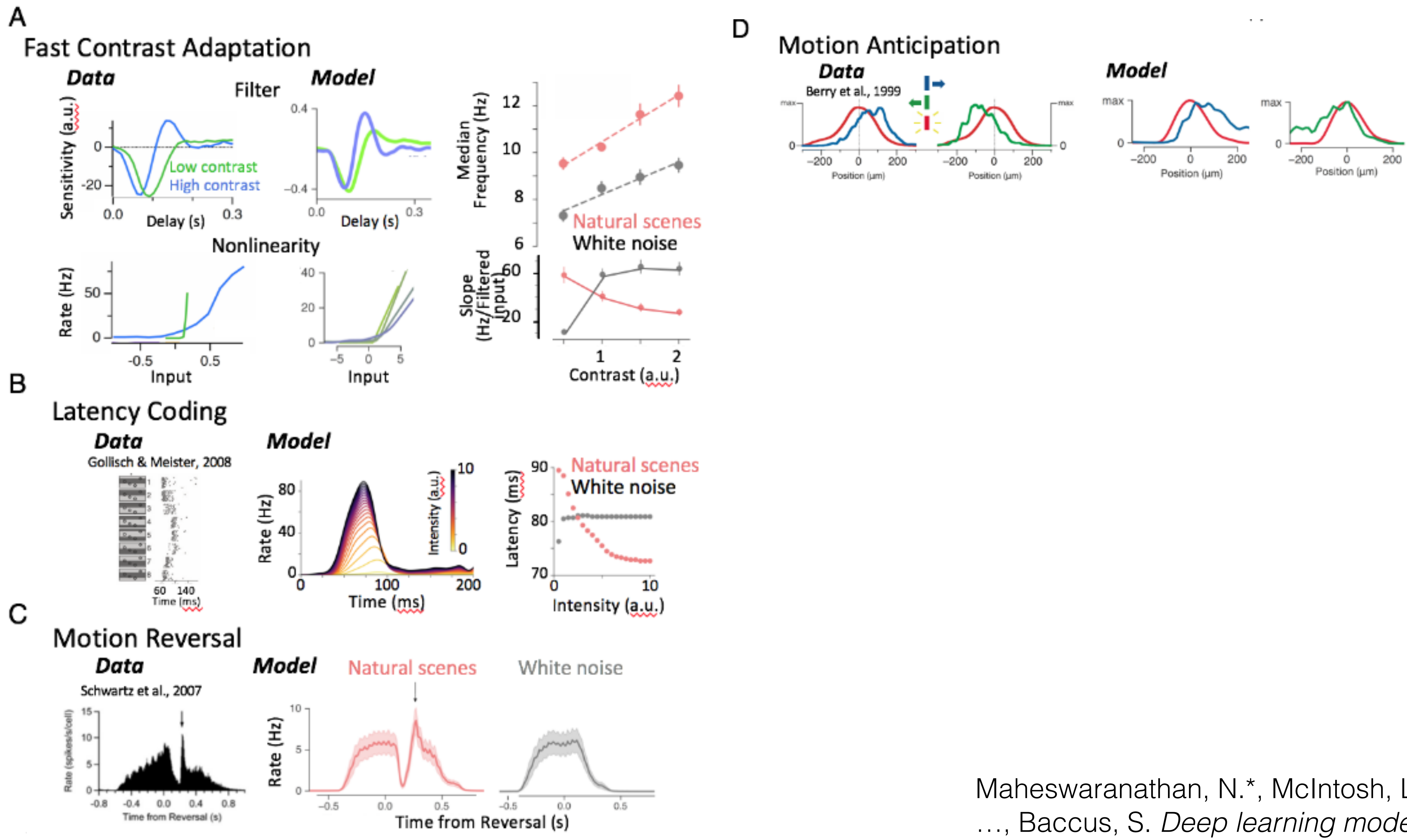
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



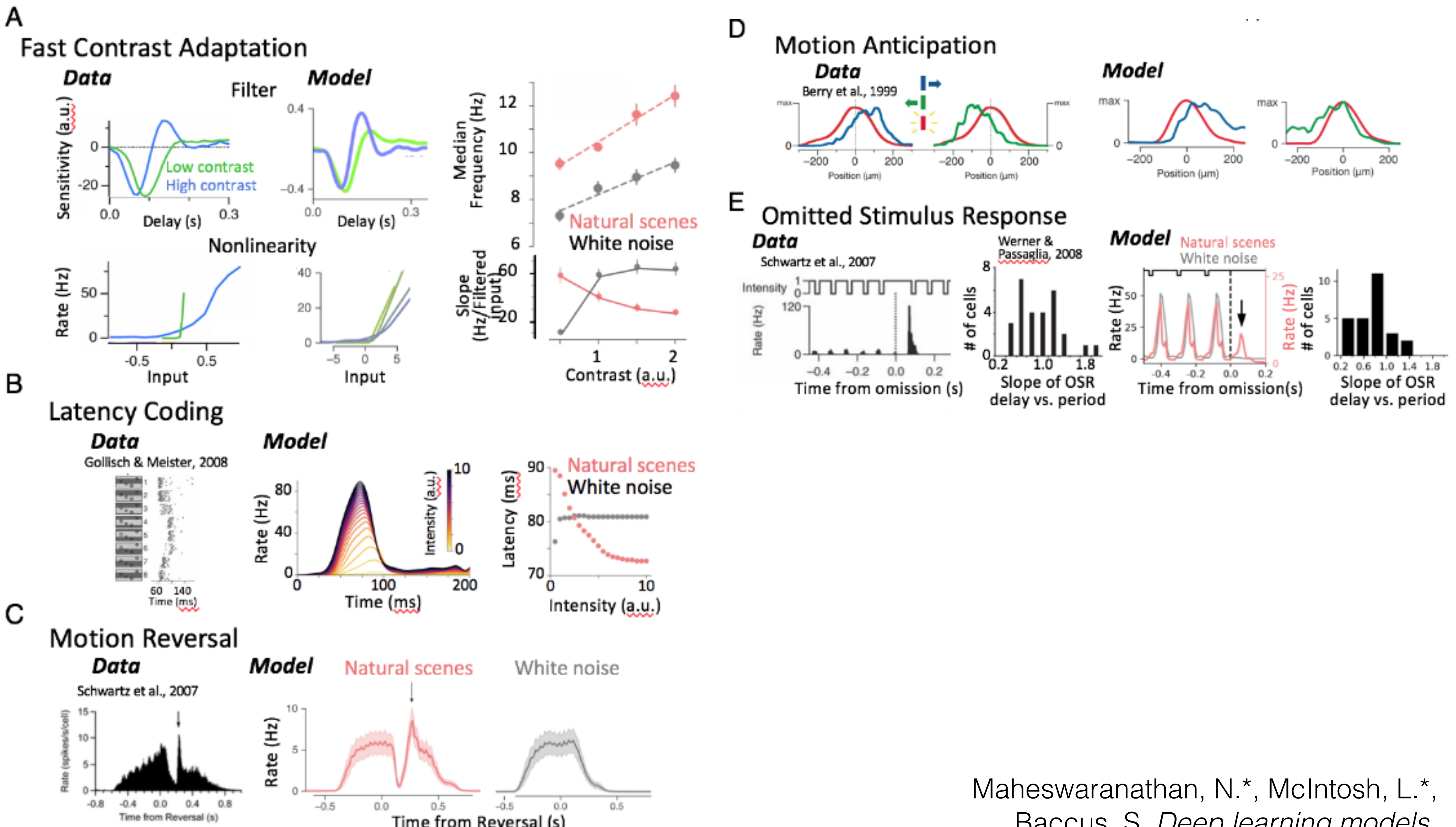
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



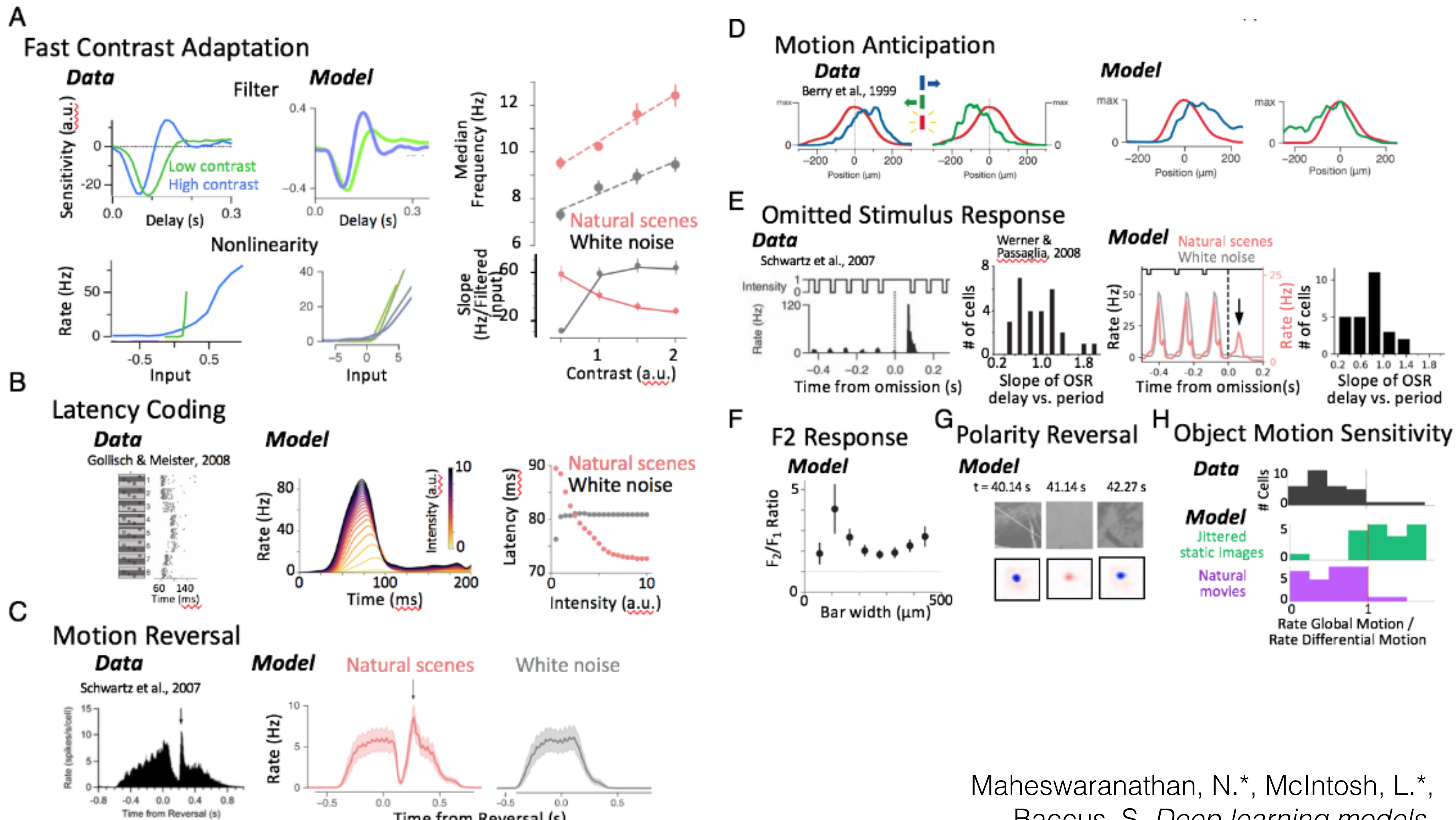
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



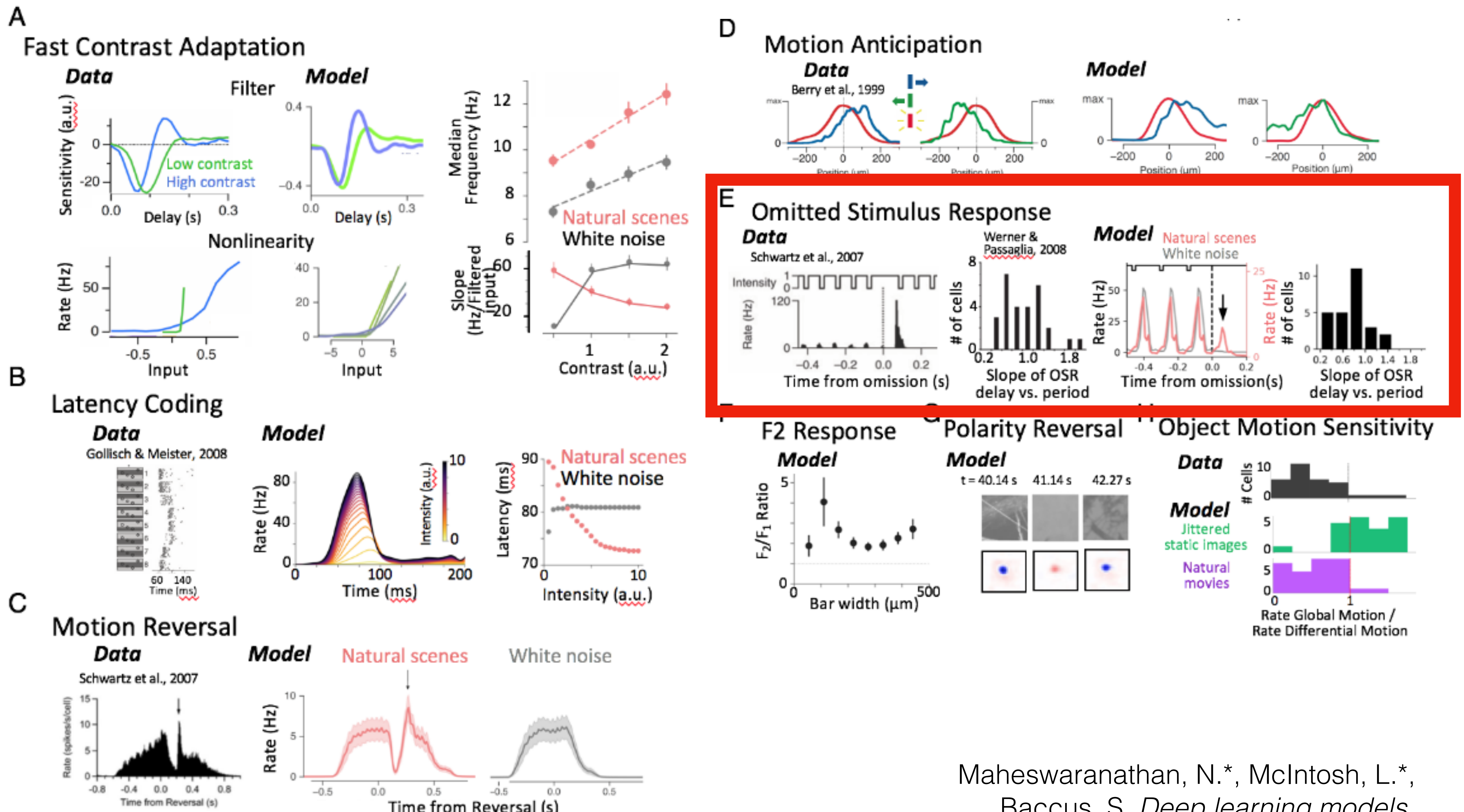
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Unified Mechanistic Model



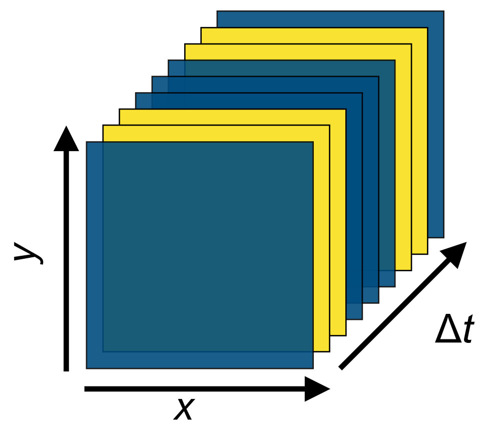
Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

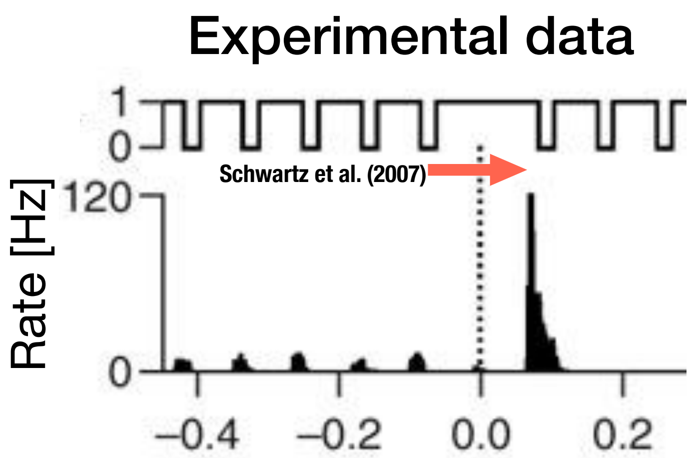
Unified Mechanistic Model

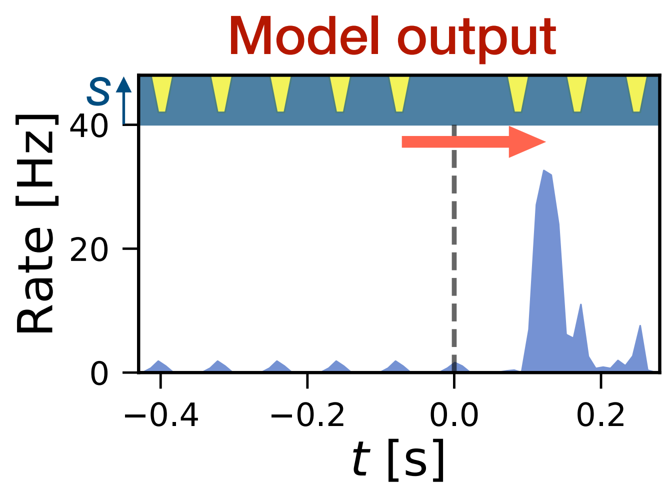


Maheswaranathan, N.*, McIntosh, L.*, …, Baccus, S. *Deep learning models reveal internal structure and diverse computations in the retina under natural scenes, bioArxiv 2018*

Mechanistic Predictions - Omitted Stimulus Response







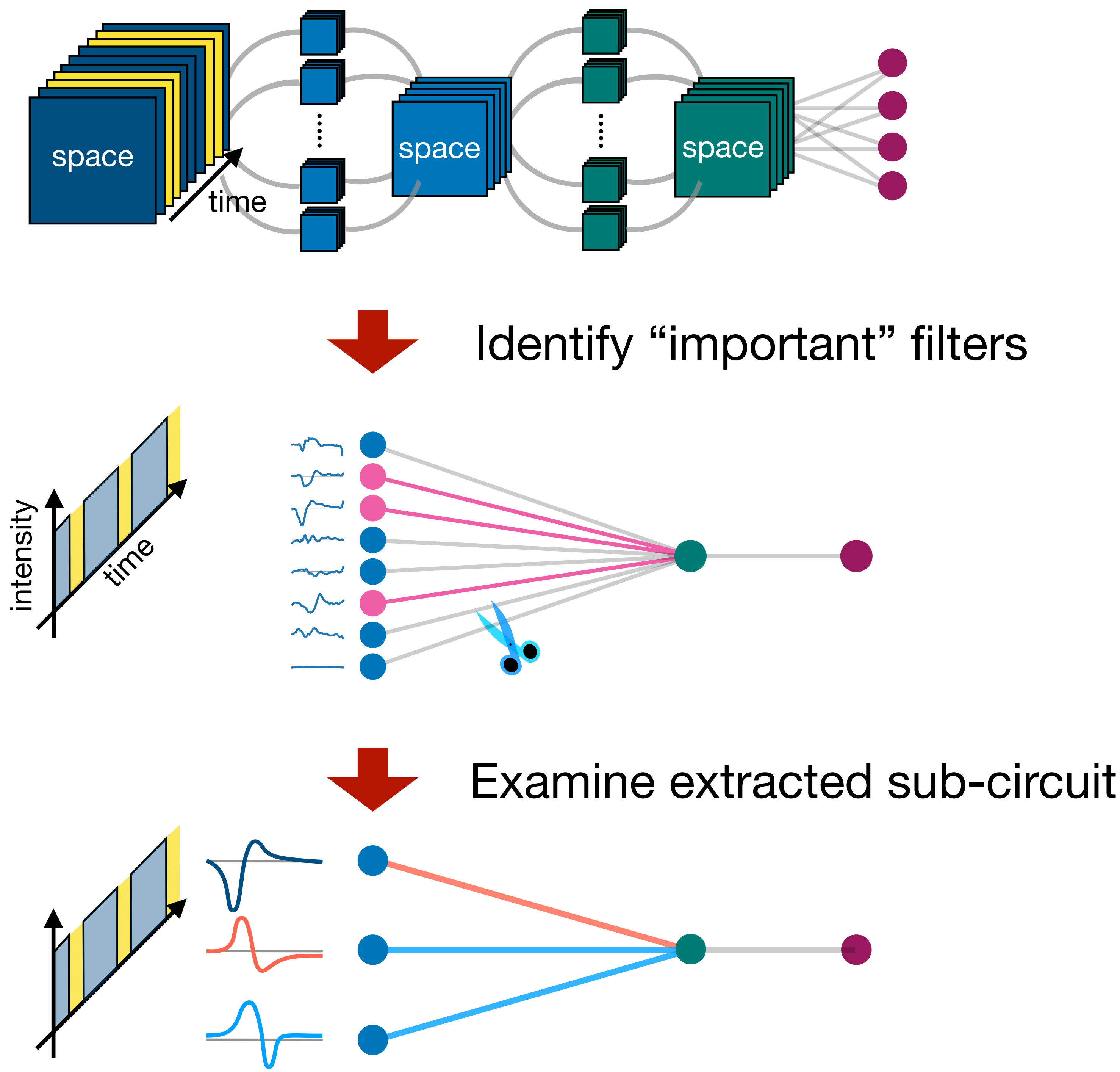
Hidenori Tanaka
(Stanford)

Q1. What computational mechanism causes the large amplitude burst?

Q2. How is the latency of the peak proportional to the period of the flashes?

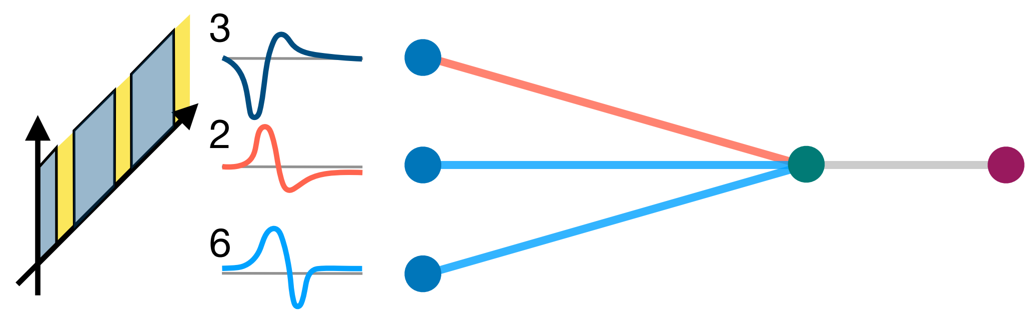
Tanaka, H., Nayebi, A. Baccus, S, Ganguli, S. *From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction,* NeurIPS 2019

Model Reduction Approach

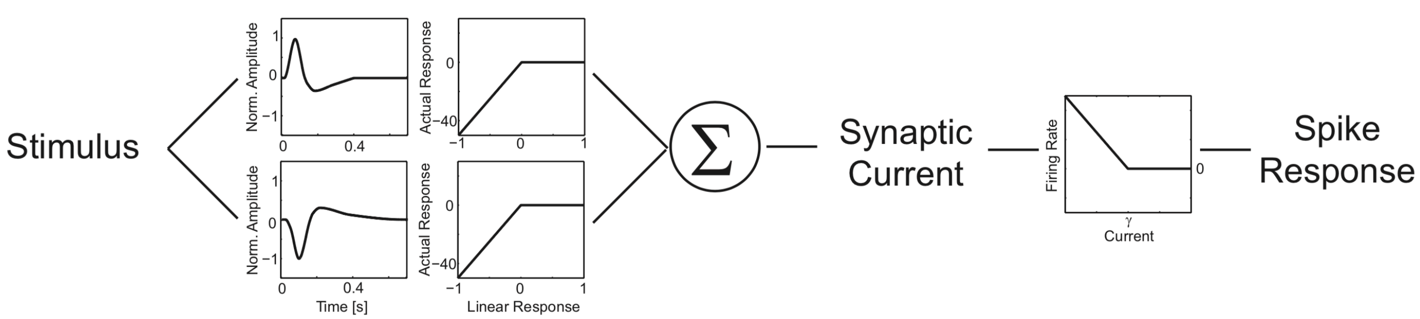


Tanaka, H., Nayebi, A. Baccus, S, Ganguli, S. *From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction,* NeurIPS 2019

OSR Mechanistic Prediction from Distilled Model

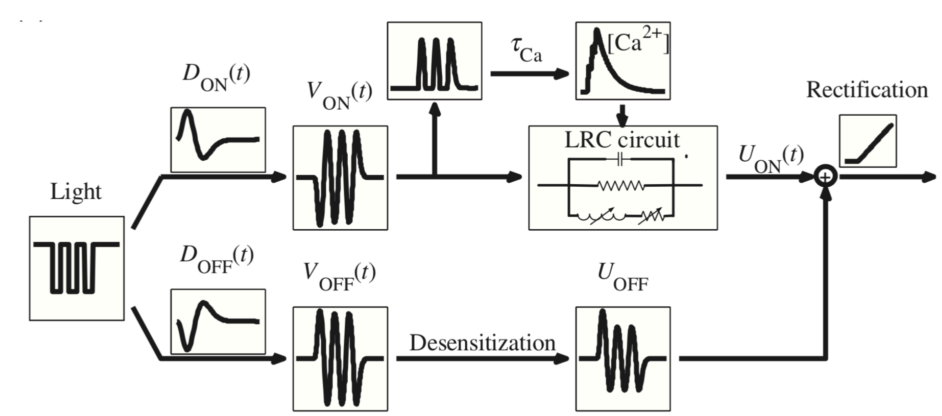


ON-OFF dual pathways model
➡ cannot reproduce predictive latency



B. Werner,…, L. Passaglia et al. (2008)

LRC circuit model
➡ no resonance found in bipolar cells



J. Gao,…, M.J. Berry et al. (2009)

Tanaka, H., Nayebi, A. Baccus, S, Ganguli, S. *From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction,* NeurIPS 2019

Takeaways

AI/Performance **Neuroscience**

Takeaways

AI/Performance **Neuroscience**

- In areas where a high-level behavioral goal is not immediately clear, if we can collect sufficient data then we can fit deep models directly

Takeaways

AI/Performance **Neuroscience**

- In areas where a high-level behavioral goal is not immediately clear, if we can collect sufficient data then we can fit deep models directly
    - In this data regime, models that more closely match the anatomical constraints of the system also generalize better on held-out data — can provide useful architectural insight into a retinal "front-end" which traditional CNN models lack

Takeaways

AI/Performance **Neuroscience**

- In areas where a high-level behavioral goal is not immediately clear, if we can collect sufficient data then we can fit deep models directly
  - In this data regime, models that more closely match the anatomical constraints of the system also generalize better on held-out data — can provide useful architectural insight into a retinal "front-end" which traditional CNN models lack

- Network models that are directly fit to ethologically-relevant stimuli generalize to other stimuli and recapitulate responses to simpler, structured stimuli

Takeaways

AI/Performance **Neuroscience**

- In areas where a high-level behavioral goal is not immediately clear, if we can collect sufficient data then we can fit deep models directly
  - In this data regime, models that more closely match the anatomical constraints of the system also generalize better on held-out data — can provide useful architectural insight into a retinal "front-end" which traditional CNN models lack

- Network models that are directly fit to ethologically-relevant stimuli generalize to other stimuli and recapitulate responses to simpler, structured stimuli
  - They provide a unified mechanistic model, which we can distill and probe to yield new mechanistic hypotheses ("cheap" *in-silico* experiments)

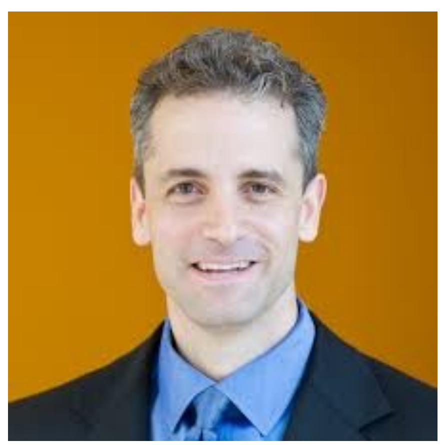
Acknowledgements

Surya Ganguli**

Daniel Yamins**

Stephen Baccus

Jim DiCarlo



Thanks!

Funding:

Neurosciences PhD Program

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

Jonas Kubilius

David Sussillo

Daniel Bear

Hidenori Tanaka

Lane McIntosh

Niru Maheswaranathan

Contact:
anayebi@stanford.edu