# Convolutional Recurrent Neural Network Models of Dynamics in Higher Visual Cortex

**VSS 2018**

*2018.05.21*

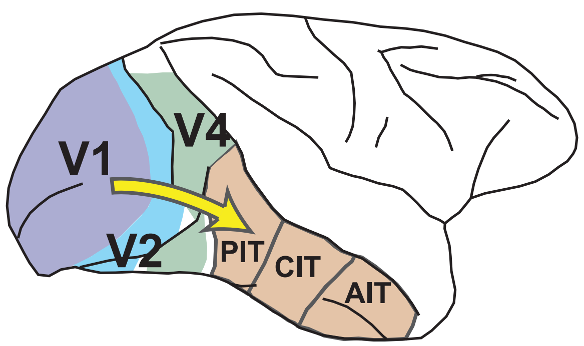

Aran Nayebi
Neurosciences PhD Program
Stanford University


Stanford
Neuro AI Lab
Stanford Neuroscience and Artificial Intelligence Lab

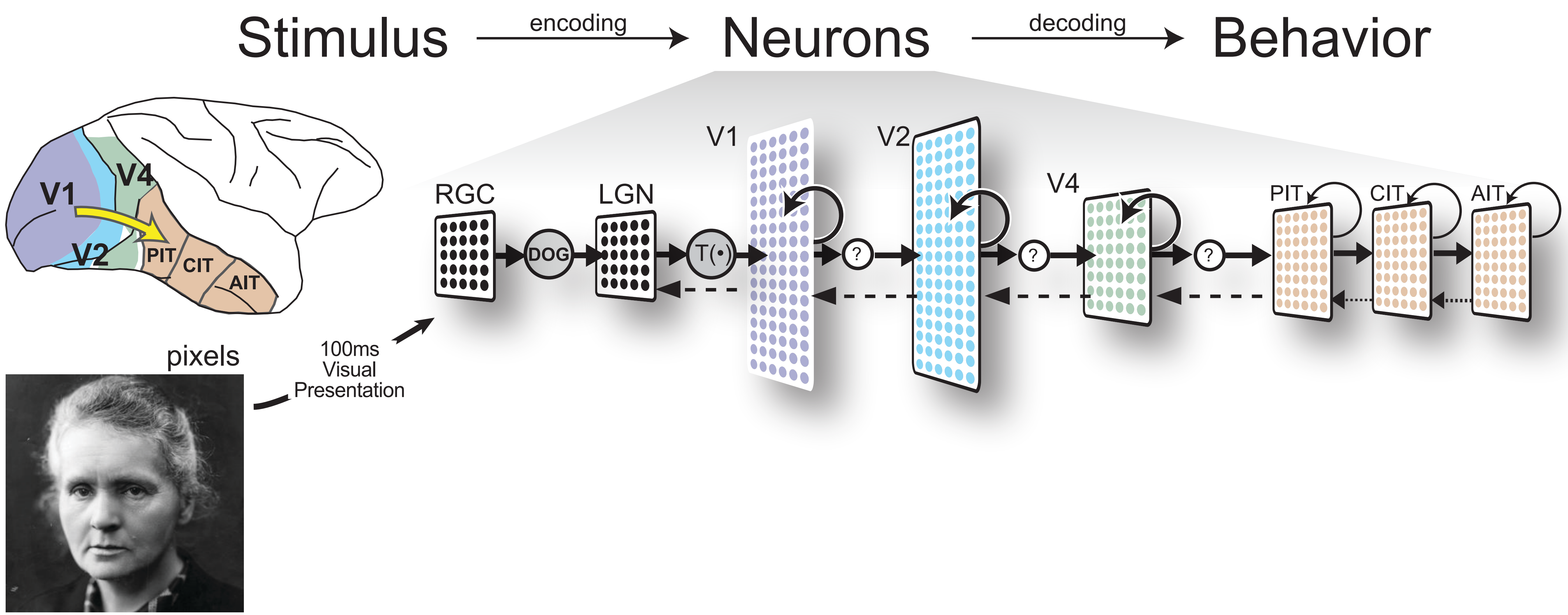
# Object Recognition is Hard (But Easy for Us)

# …thanks to the Ventral Stream

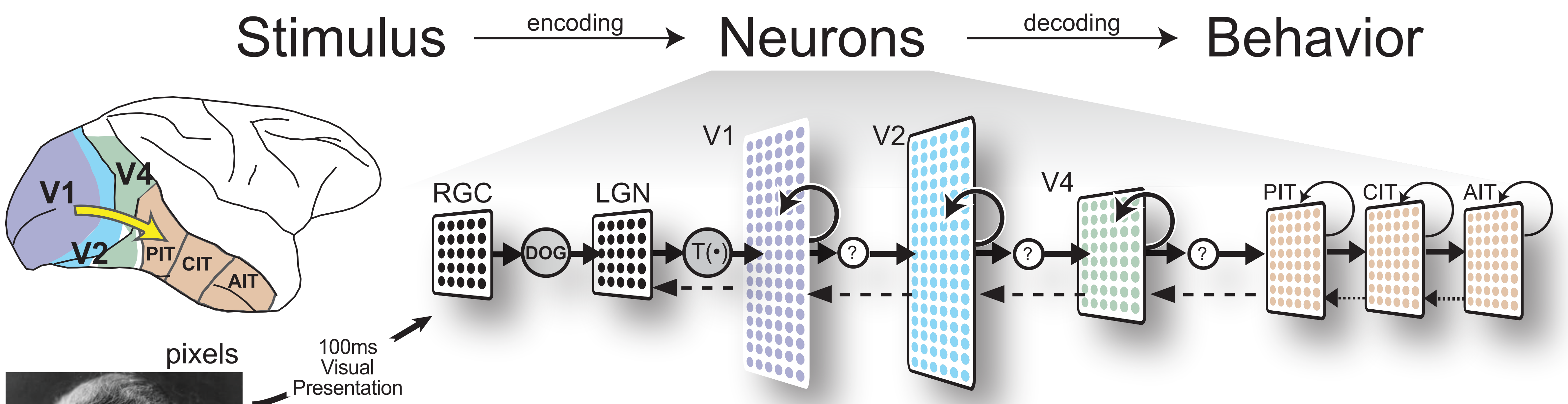
Stimulus $\xrightarrow{\text{encoding}}$ Neurons $\xrightarrow{\text{decoding}}$ Behavior

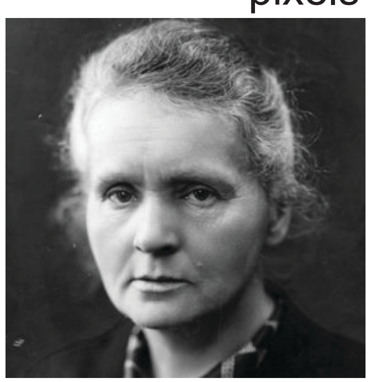


# …thanks to the Ventral Stream



# CNNs as Models of Object Recognition



Convolutional Neural Networks (CNNs)

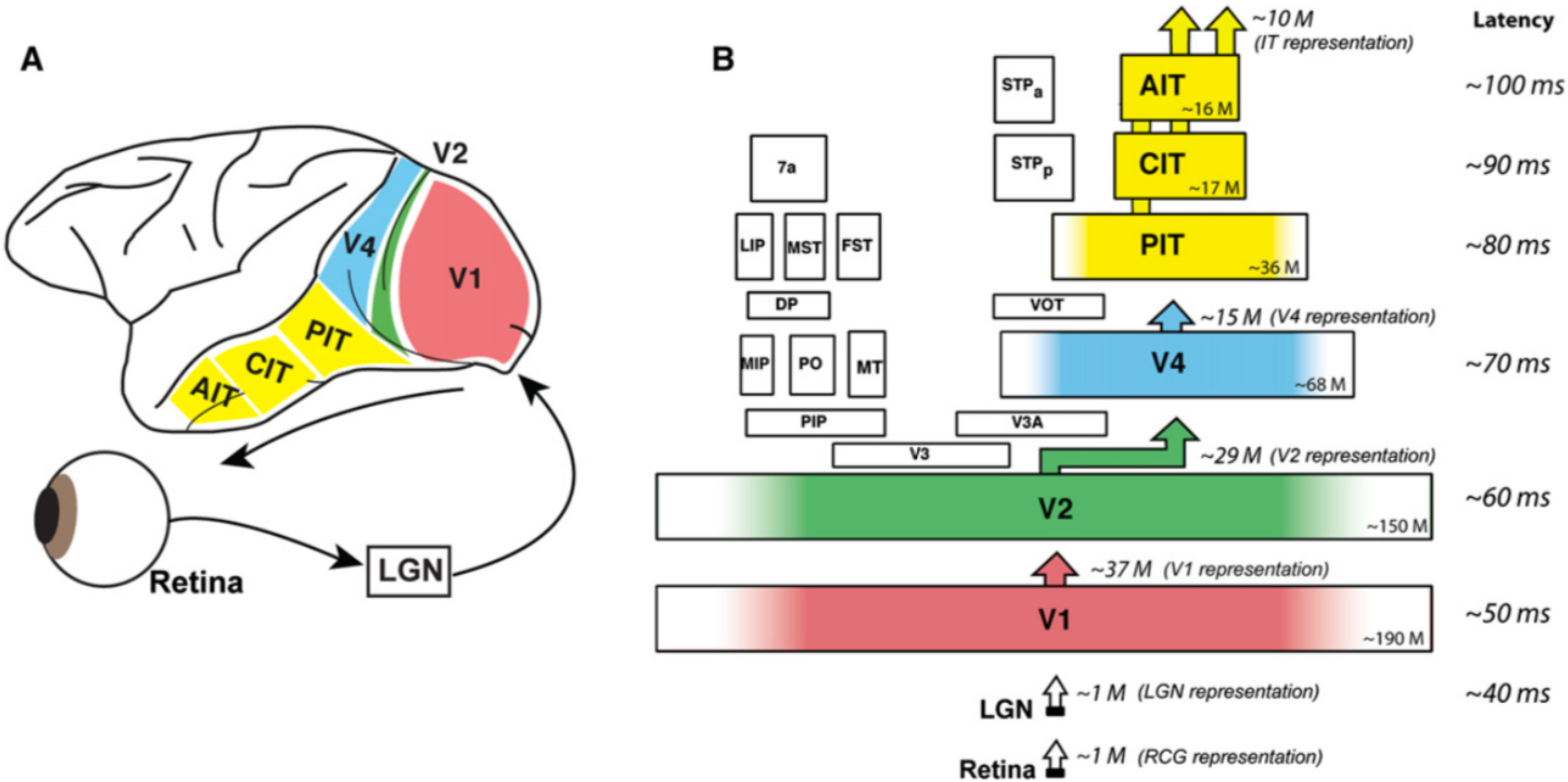
Fukushima, 1979; Lecun, 1995

CNNs are inspired by visual neuroscience:

1) **hierarchy**

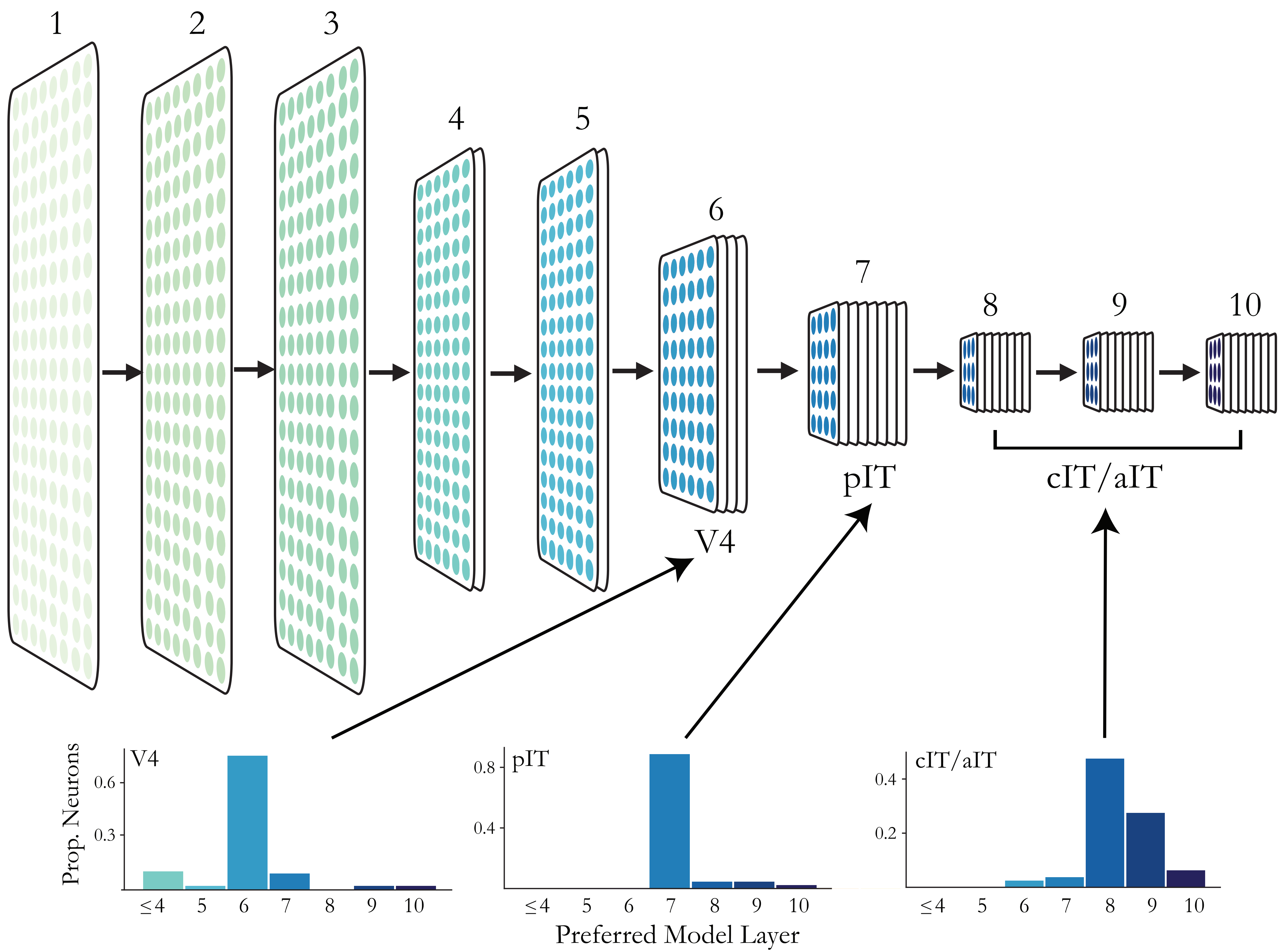
2) **retinotopy** (spatially tiled)

# ~10-12 "Layers" Plausible based on anatomy and timing

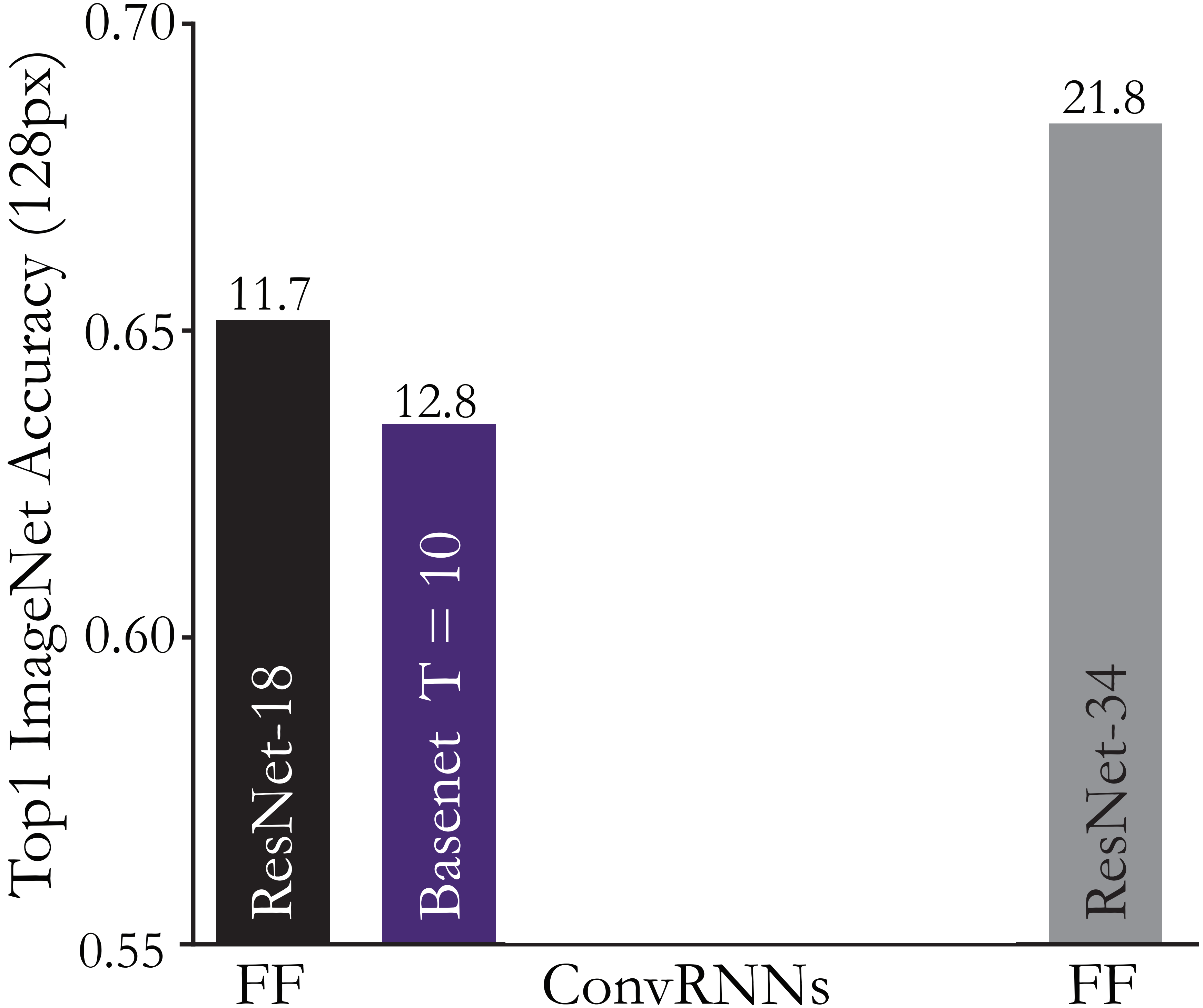



DiCarlo, Zoccolan, & Rust 2012

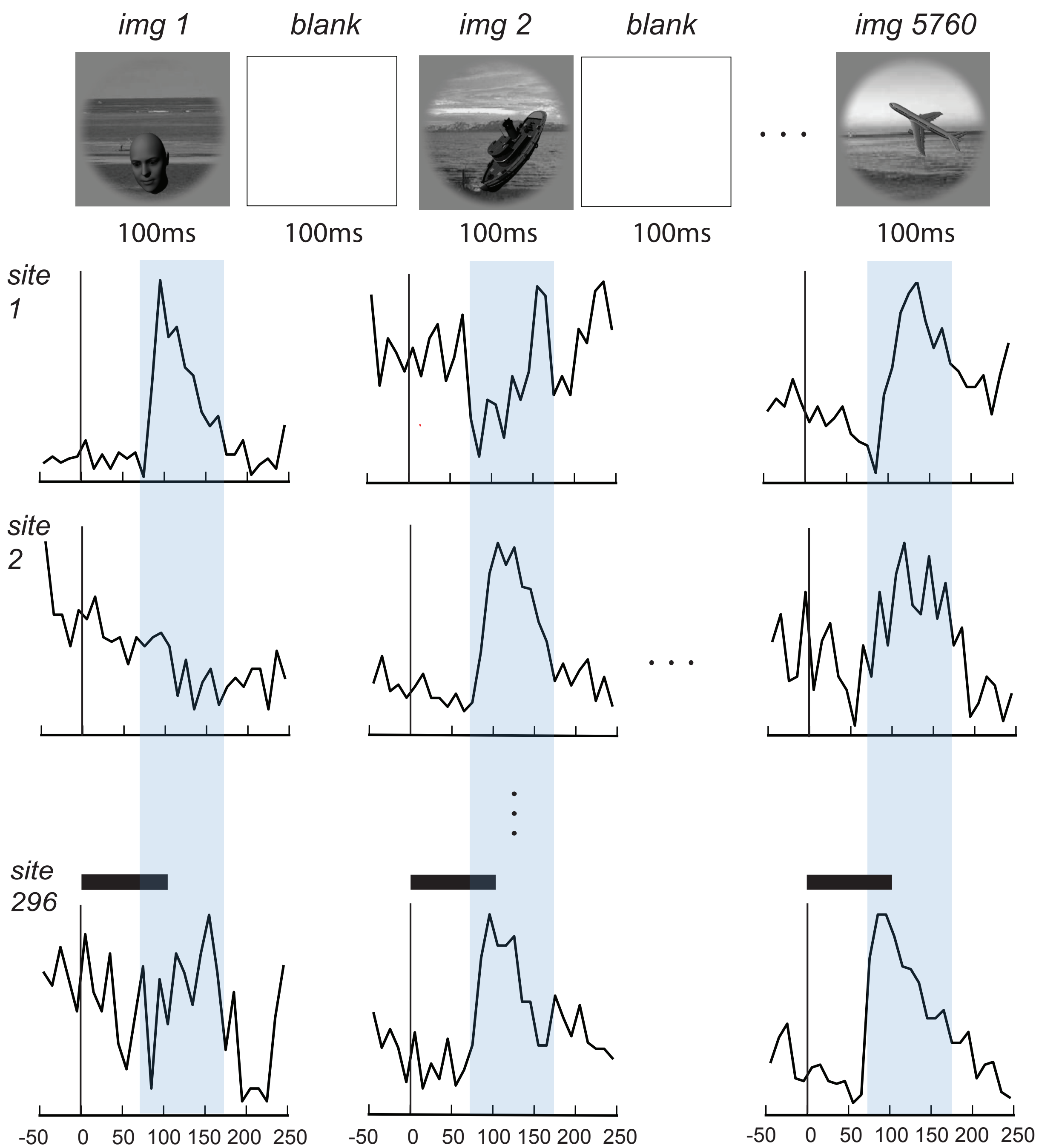
# ~10-12 "Layers" Plausible based on anatomy and timing



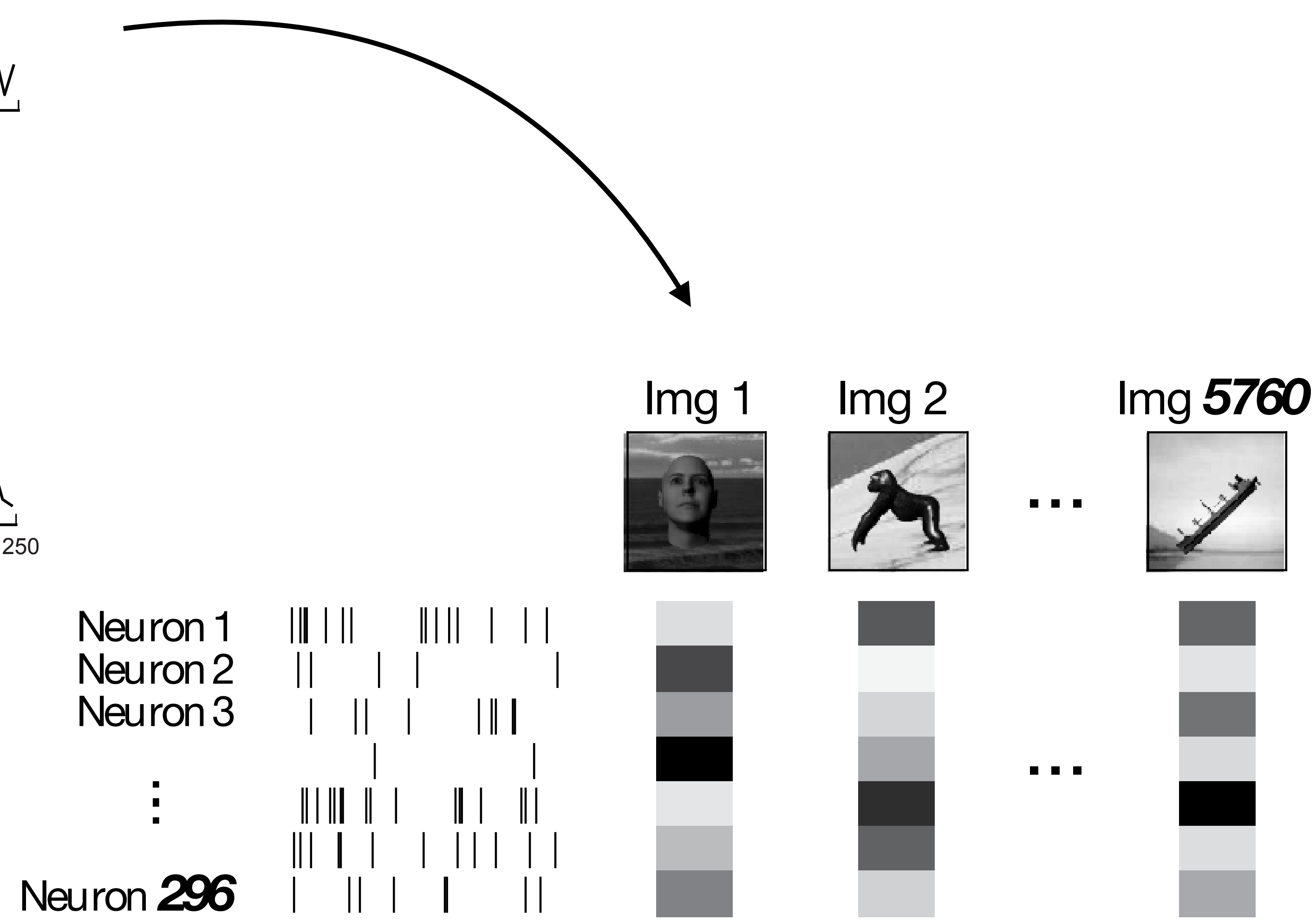
# …But, such Networks Are Far From Human Performance



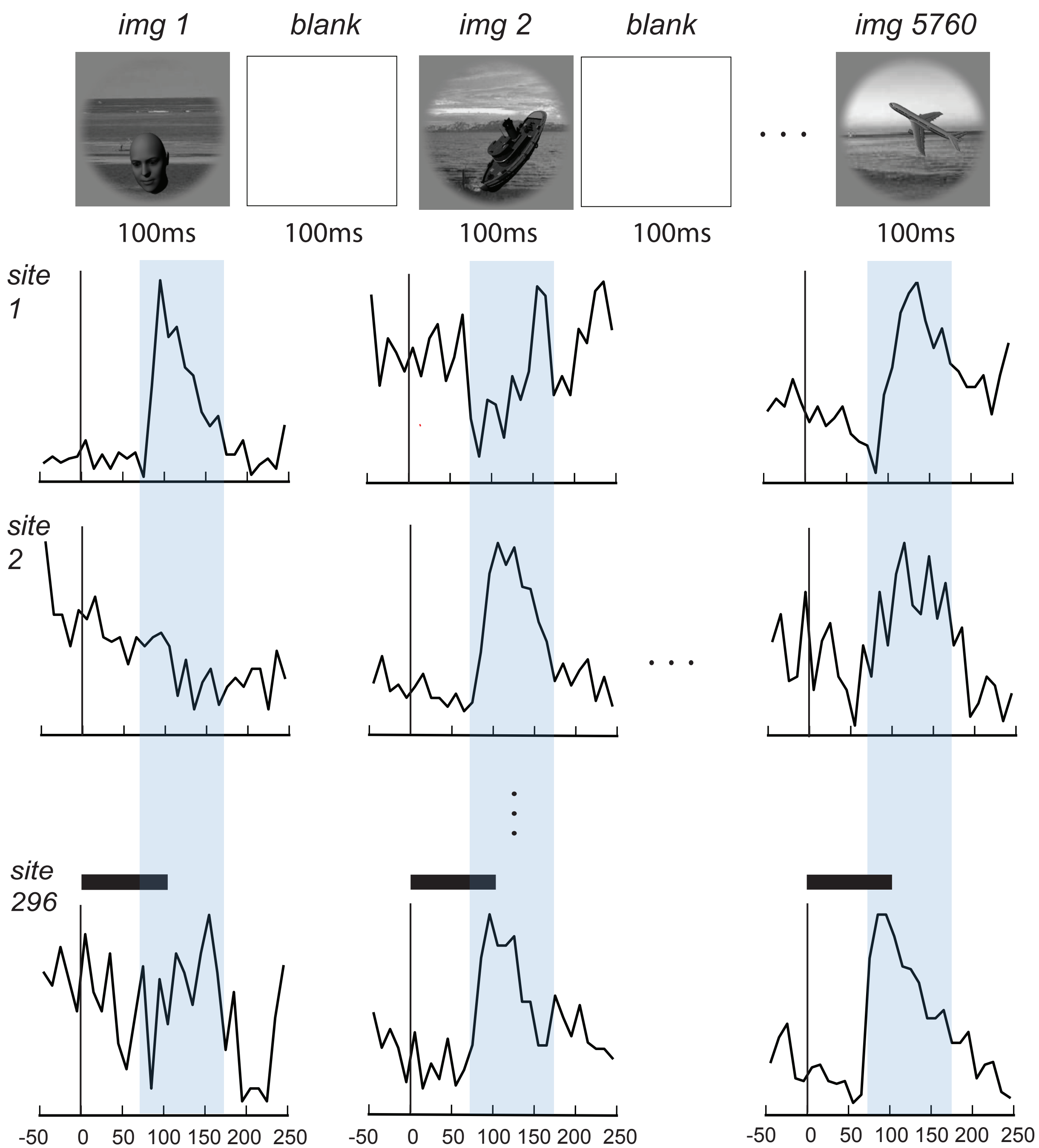
# So far, only explaining temporal average of responses




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

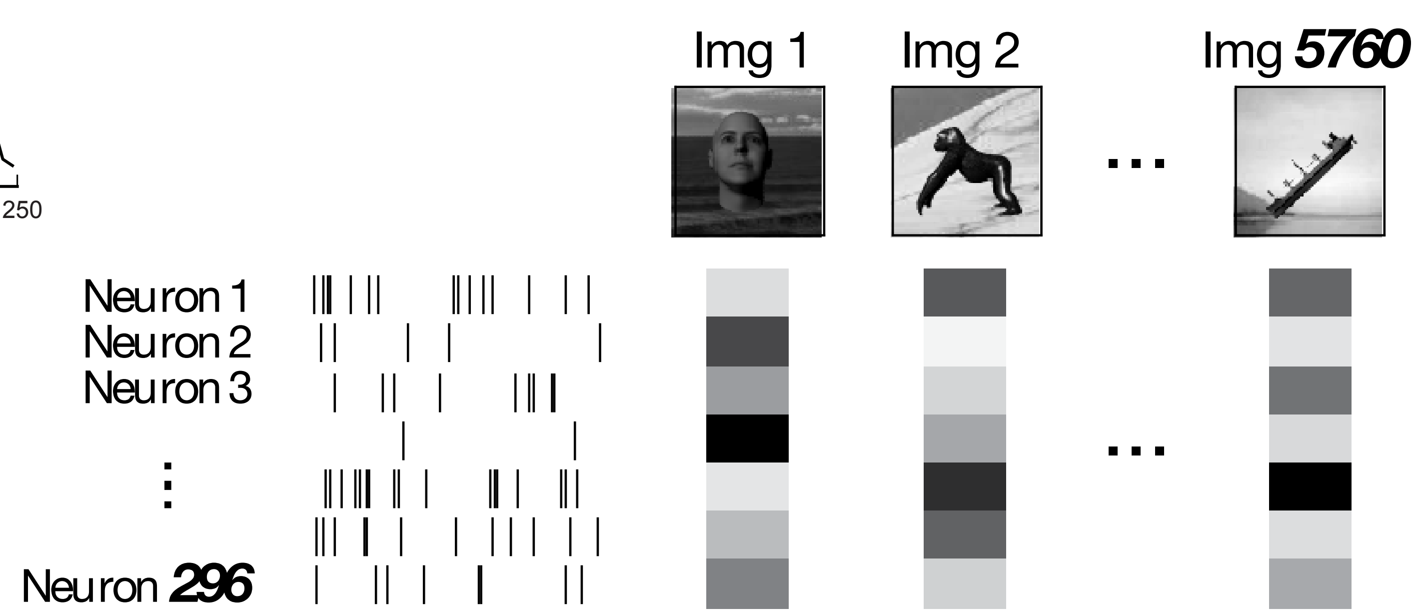


# So far, only explaining temporal average of responses



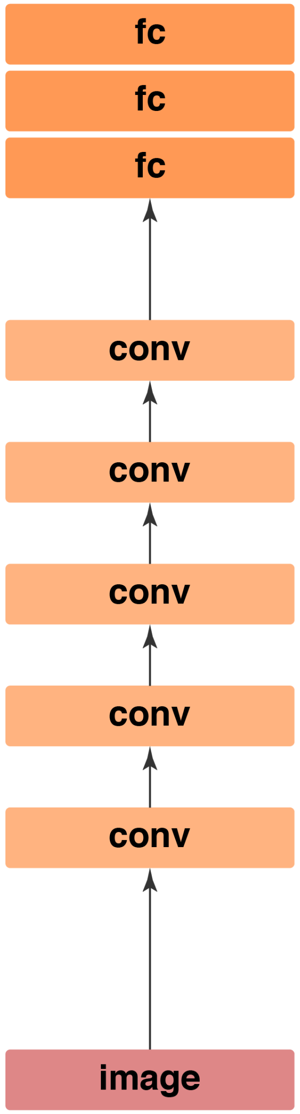

*e.g. Binned spike counts 70ms-170ms post stimulus presentation*

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

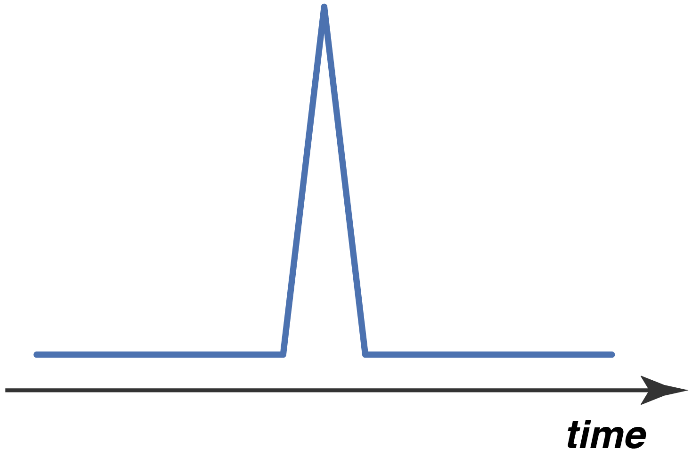


# Trajectory Possibilities

Simple feedforward networks simple dynamics:

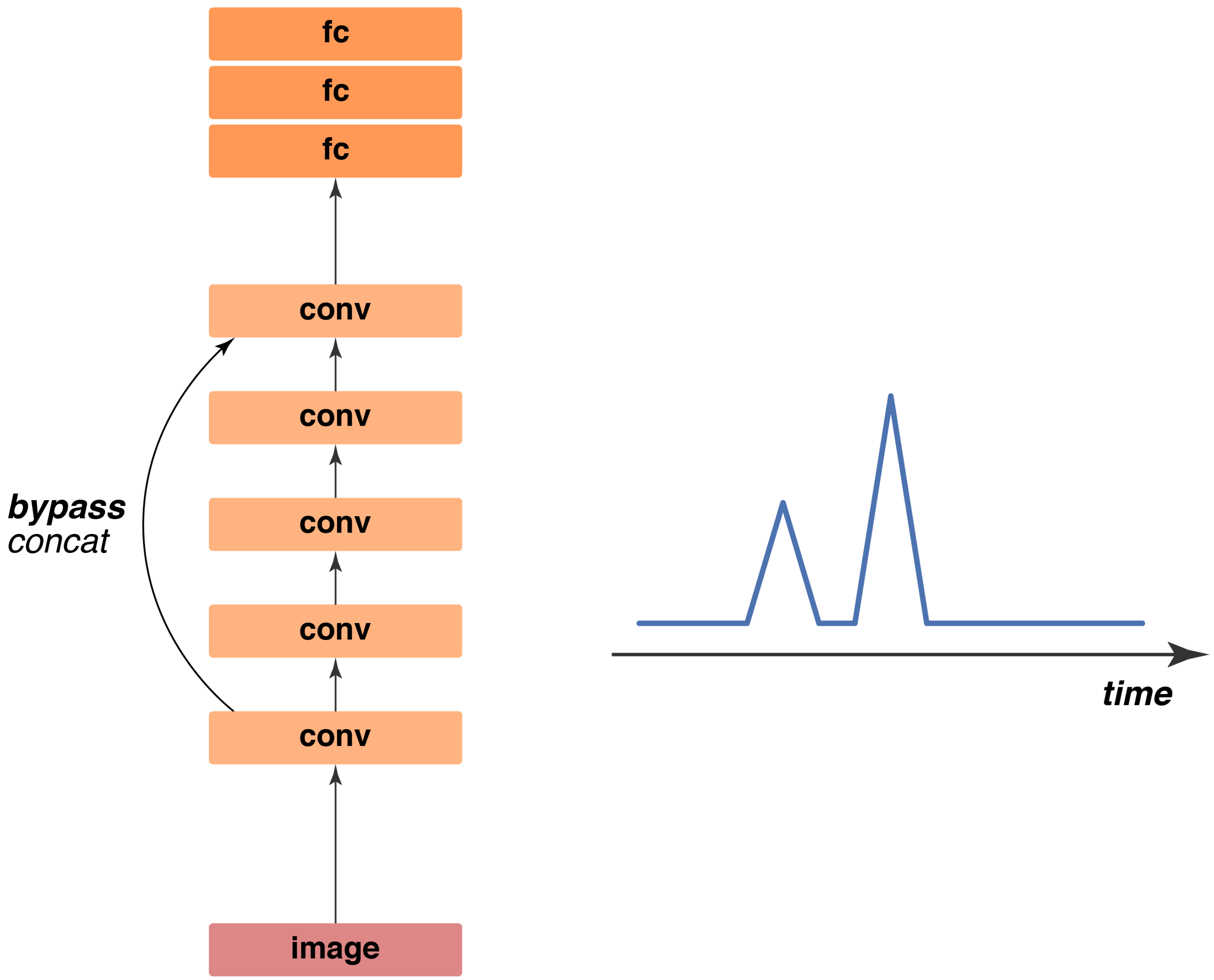



courtesy Jonas Kubilius



# Trajectory Possibilities

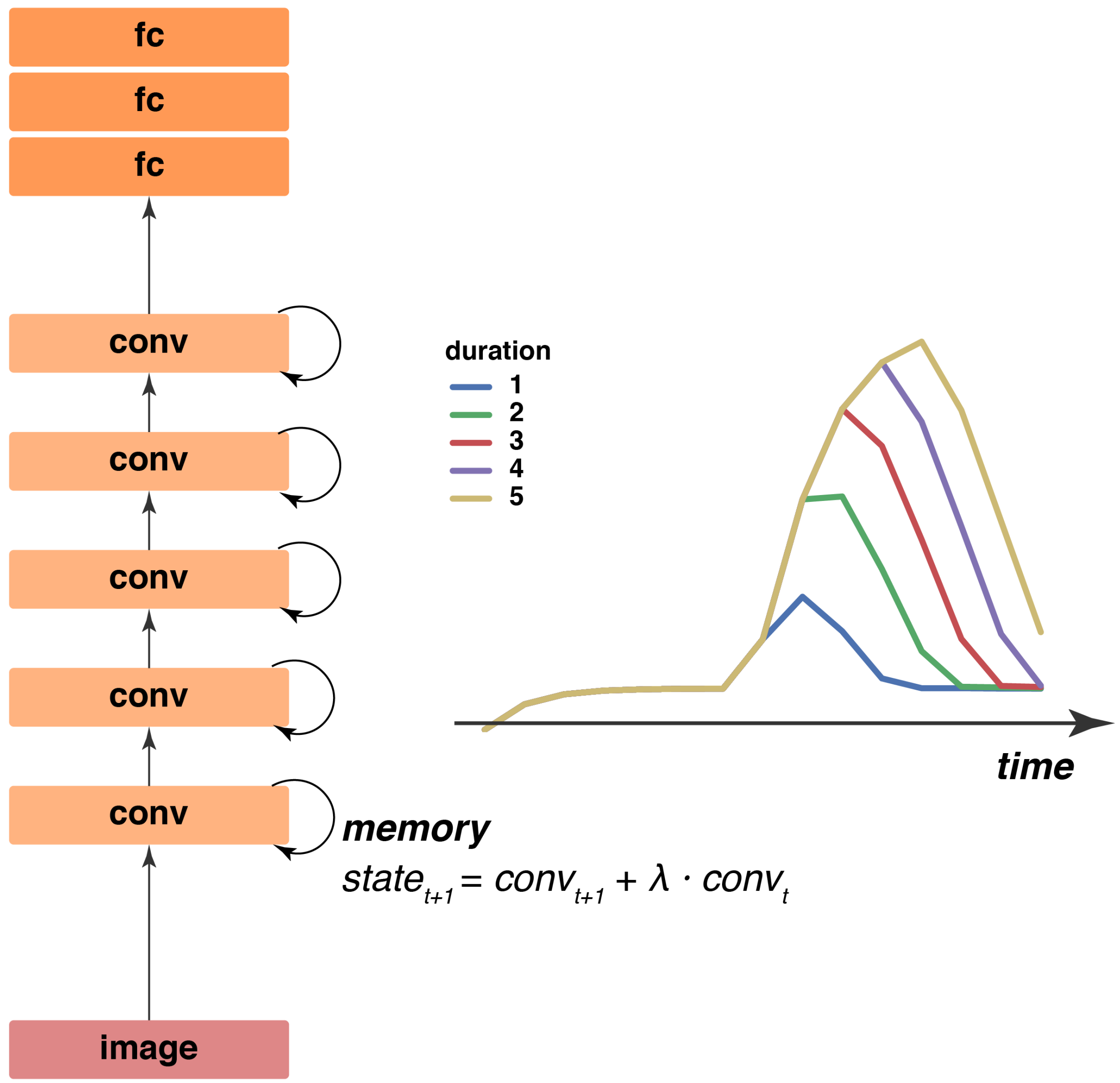
Dynamics more interesting with bypasses:




*courtesy Jonas Kubilius*

# Trajectory Possibilities

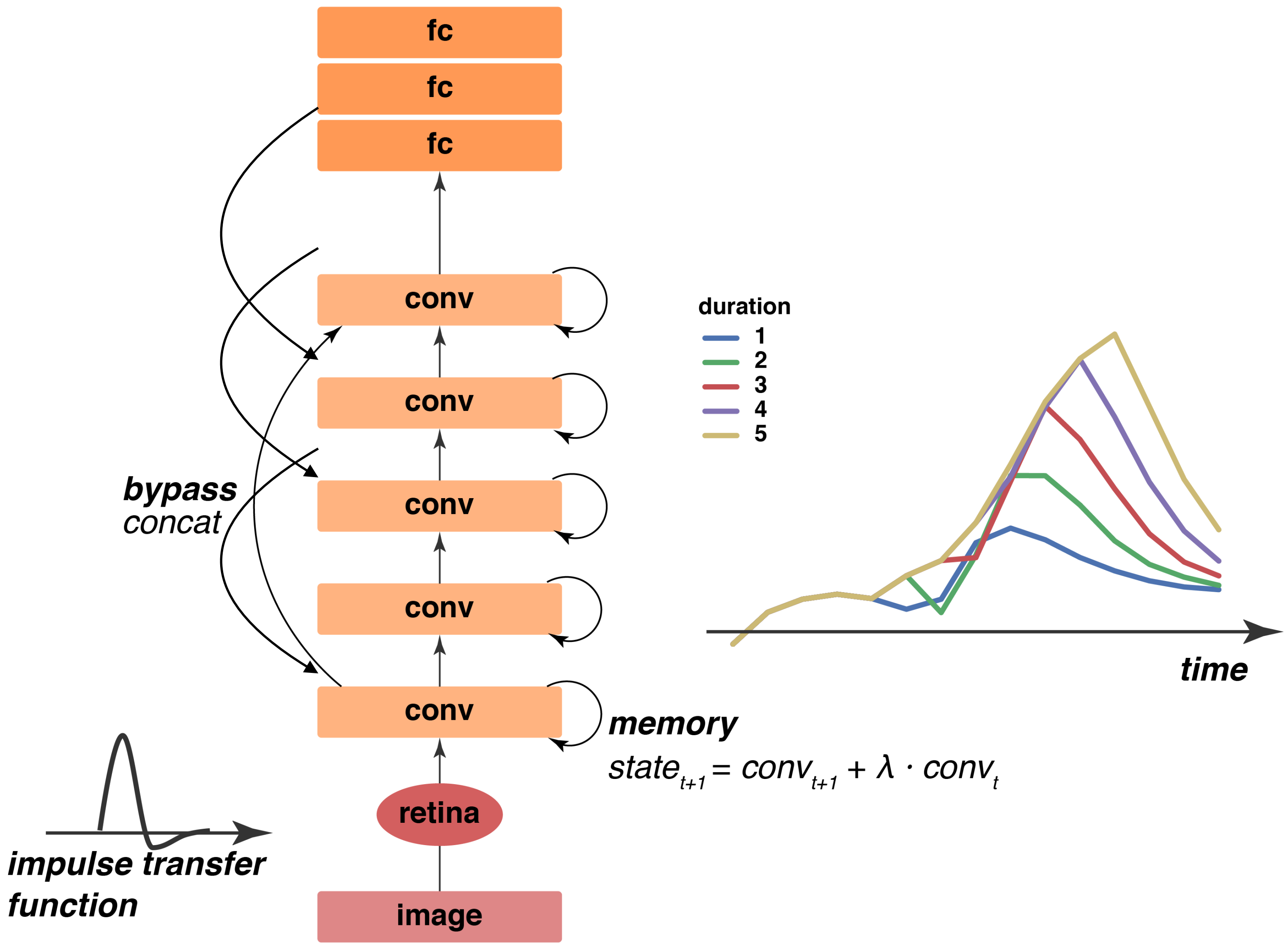
Dynamics more interesting with bypasses, local recurrence:




*courtesy Jonas Kubilius*

# Trajectory Possibilities

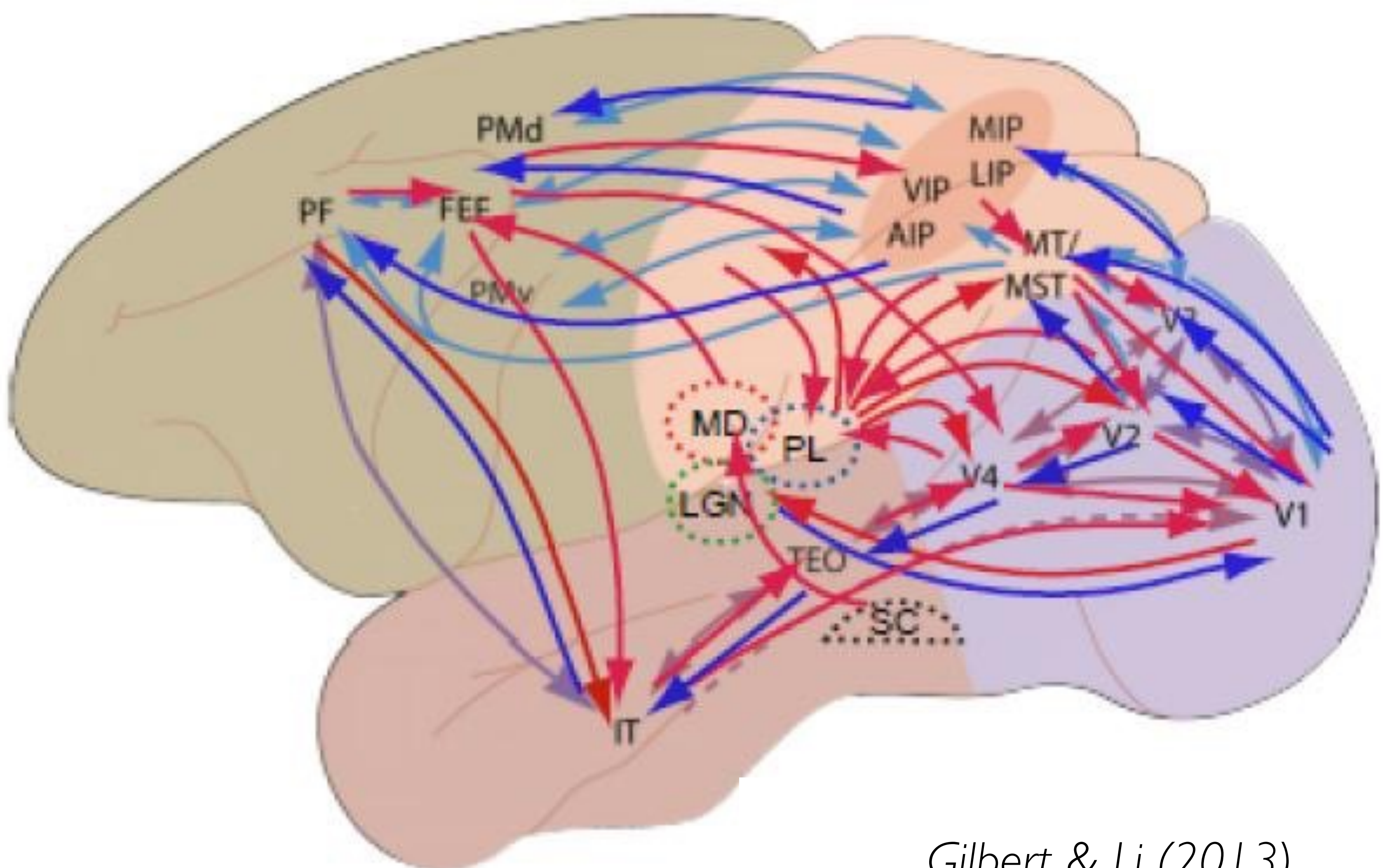
Dynamics more interesting with bypasses, local recurrence, long-range feedback:




*courtesy Jonas Kubilius*

# Dynamics result from recurrence

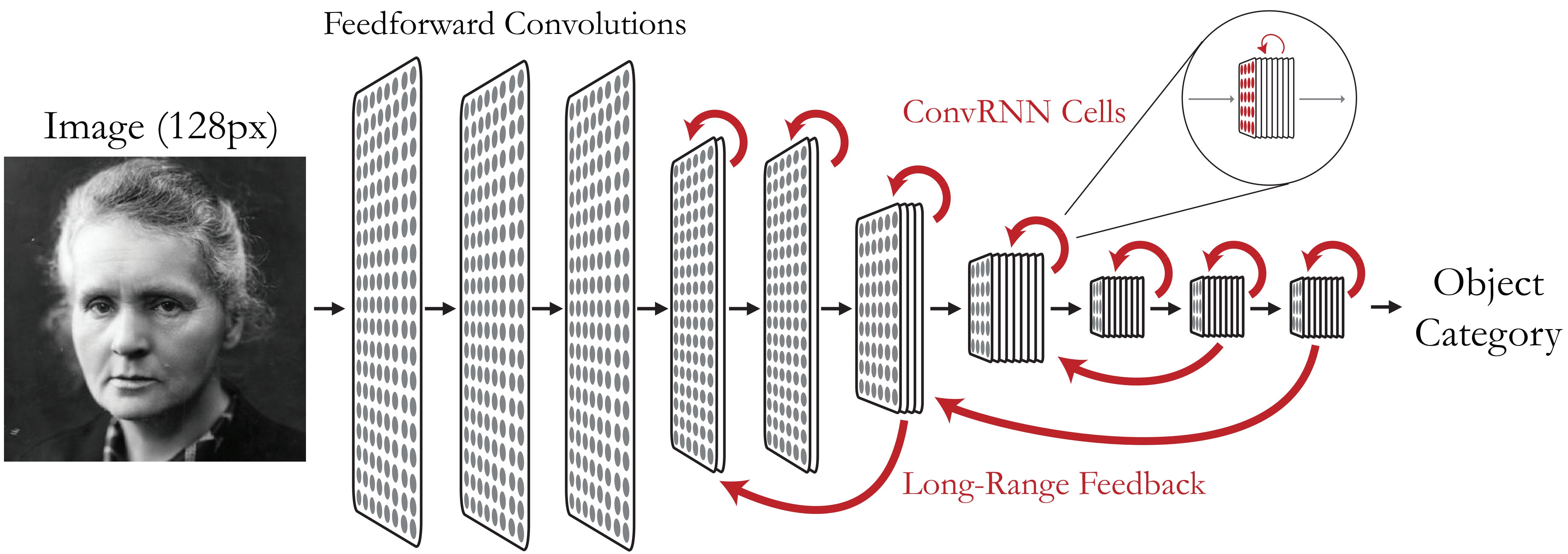
Feedbacks are everywhere anatomically:




*Gilbert & Li (2013)*

. . . but what are they for?

# Convolutional Recurrent Neural Networks (ConvRNNs)



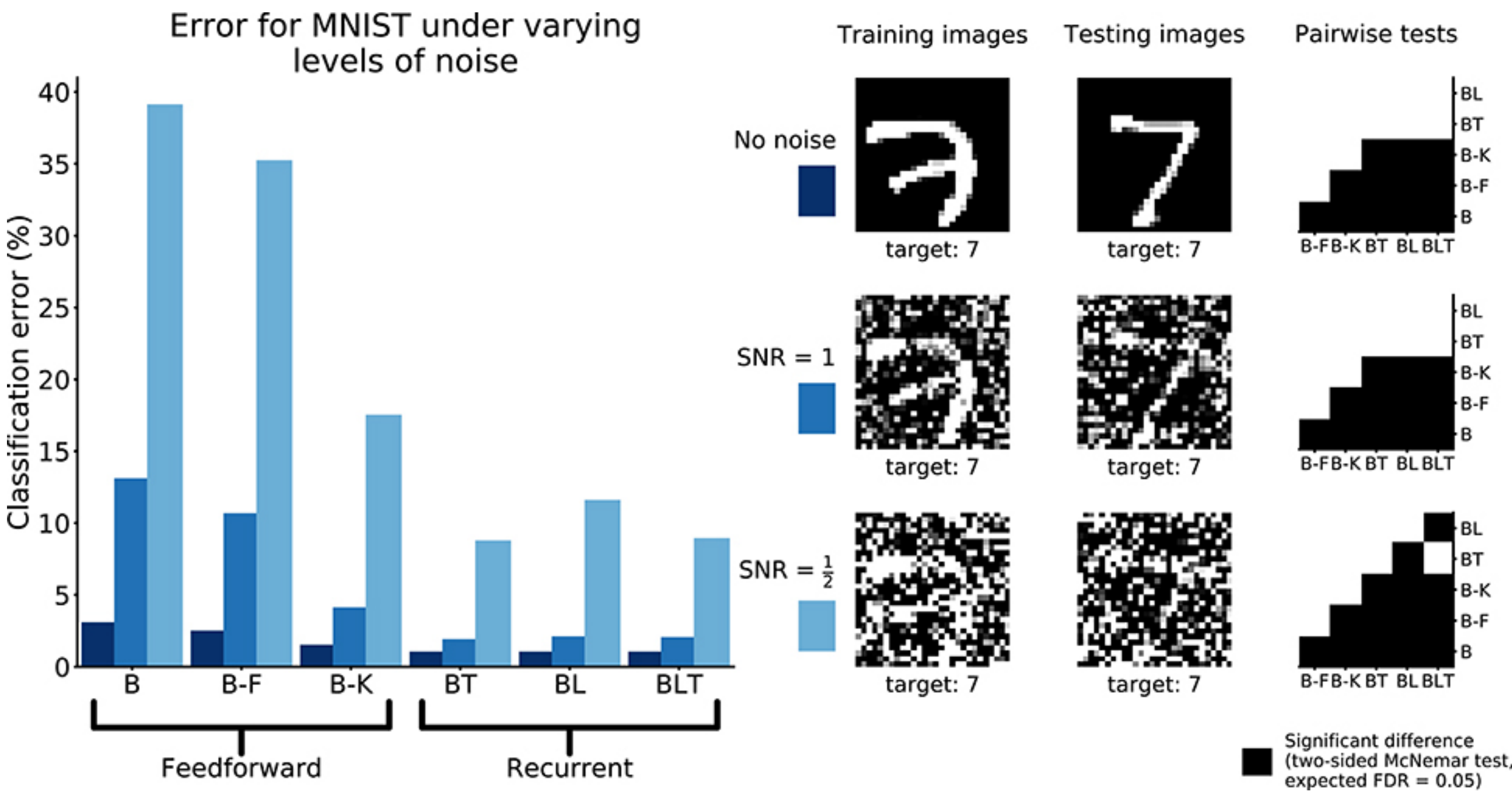
# Hypotheses for ConvRNNs

# Hypotheses for ConvRNNs - Occlusions

**Recurrent convolutional neural networks suppress occluders and enhance targets in occluded object recognition**

**Courtney J. Spoerer (courtney.spoerer@mrc-cbu.cam.ac.uk)**
Medical Research Council Cognition and Brain Sciences Unit,
15 Chaucer Road, Cambridge, CB2 7EF, UK

**Nikolaus Kriegeskorte (nikokriegeskorte@gmail.com)**
Medical Research Council Cognition and Brain Sciences Unit,
15 Chaucer Road, Cambridge, CB2 7EF, UK



# Hypotheses for ConvRNNs - Top Down Feature Attention

CBMM Memo No. 047 April 12, 2016

**Bridging the Gaps Between Residual Learning, Recurrent Neural Networks and Visual Cortex**

by

**Qianli Liao and Tomaso Poggio**
Center for Brains, Minds and Machines, McGovern Institute, MIT

Performance gains only on quite small datasets

**Feedback Networks**

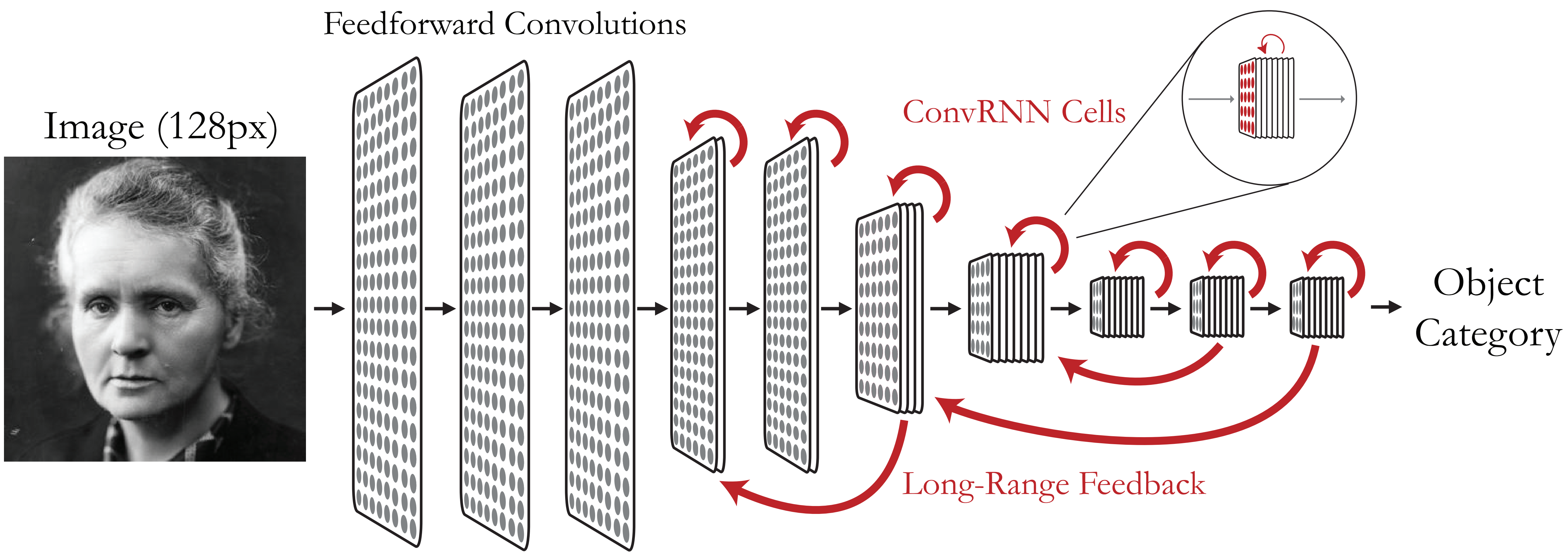
Amir R. Zamir[1,3*] Te-Lin Wu[1*] Lin Sun[1,2] William B. Shen[1] Bertram E. Shi[2]
Jitendra Malik[3] Silvio Savarese[1]
[1] Stanford University [2] HKUST [3] University of California, Berkeley
http://feedbacknet.stanford.edu/

Accepted as a workshop contribution at ICLR 2015

**ATTENTION FOR FINE-GRAINED CATEGORIZATION**

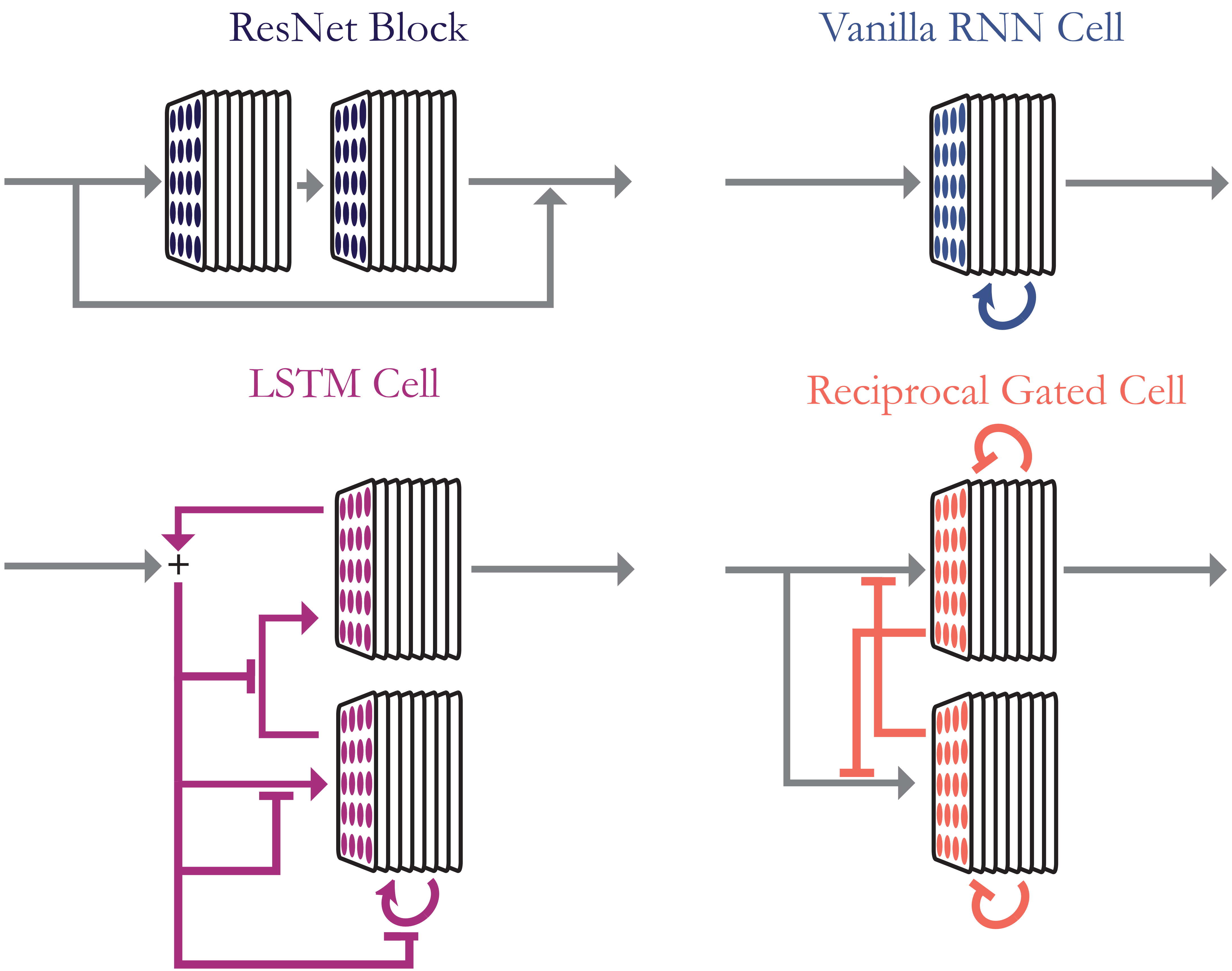
**Pierre Sermanet, Andrea Frome, Esteban Real**
Google, Inc.
{sermanet,afrome,ereal,}@google.com

# Improving ImageNet Performance with 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

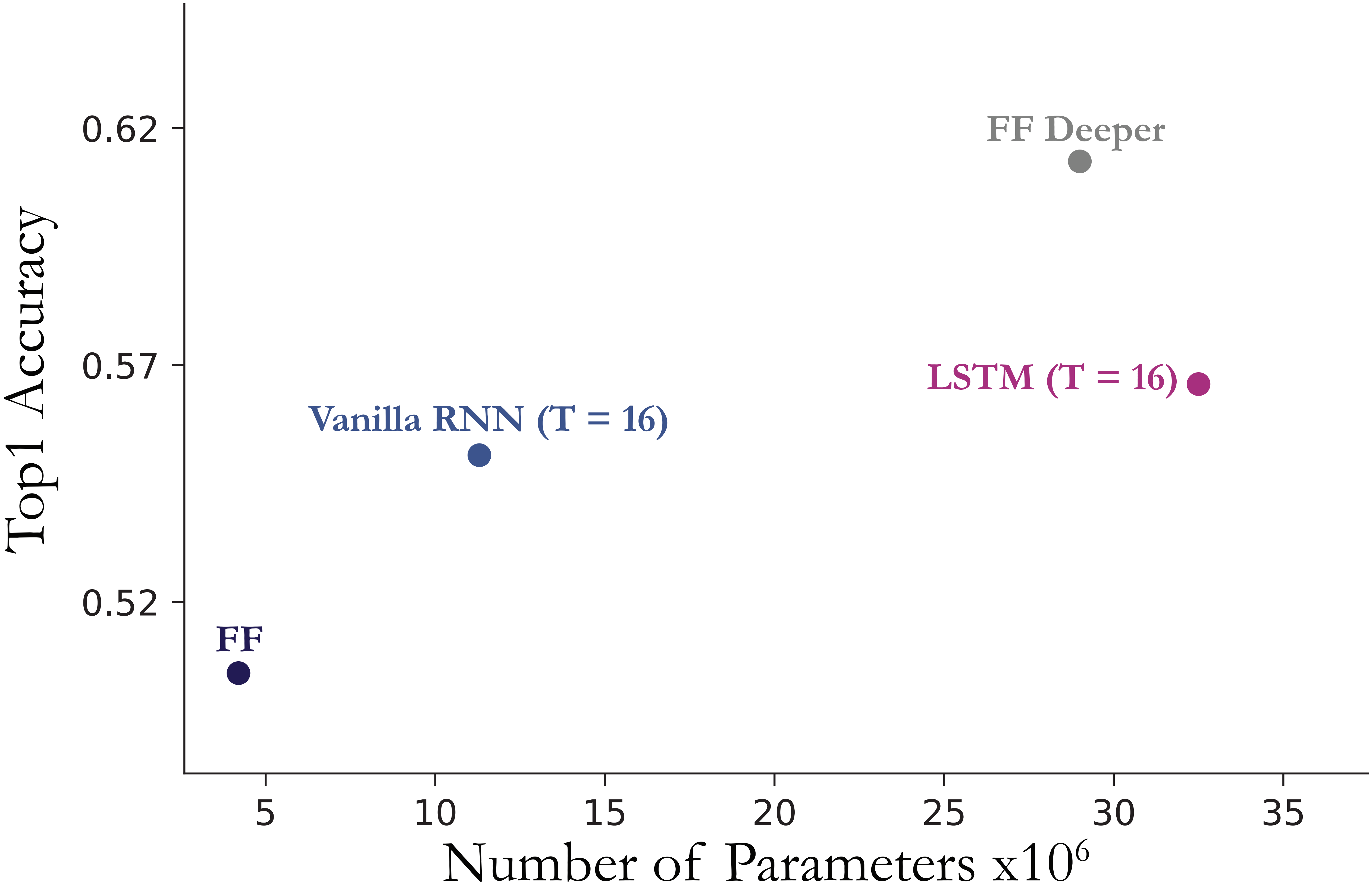
Two complementary principles:

**(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")

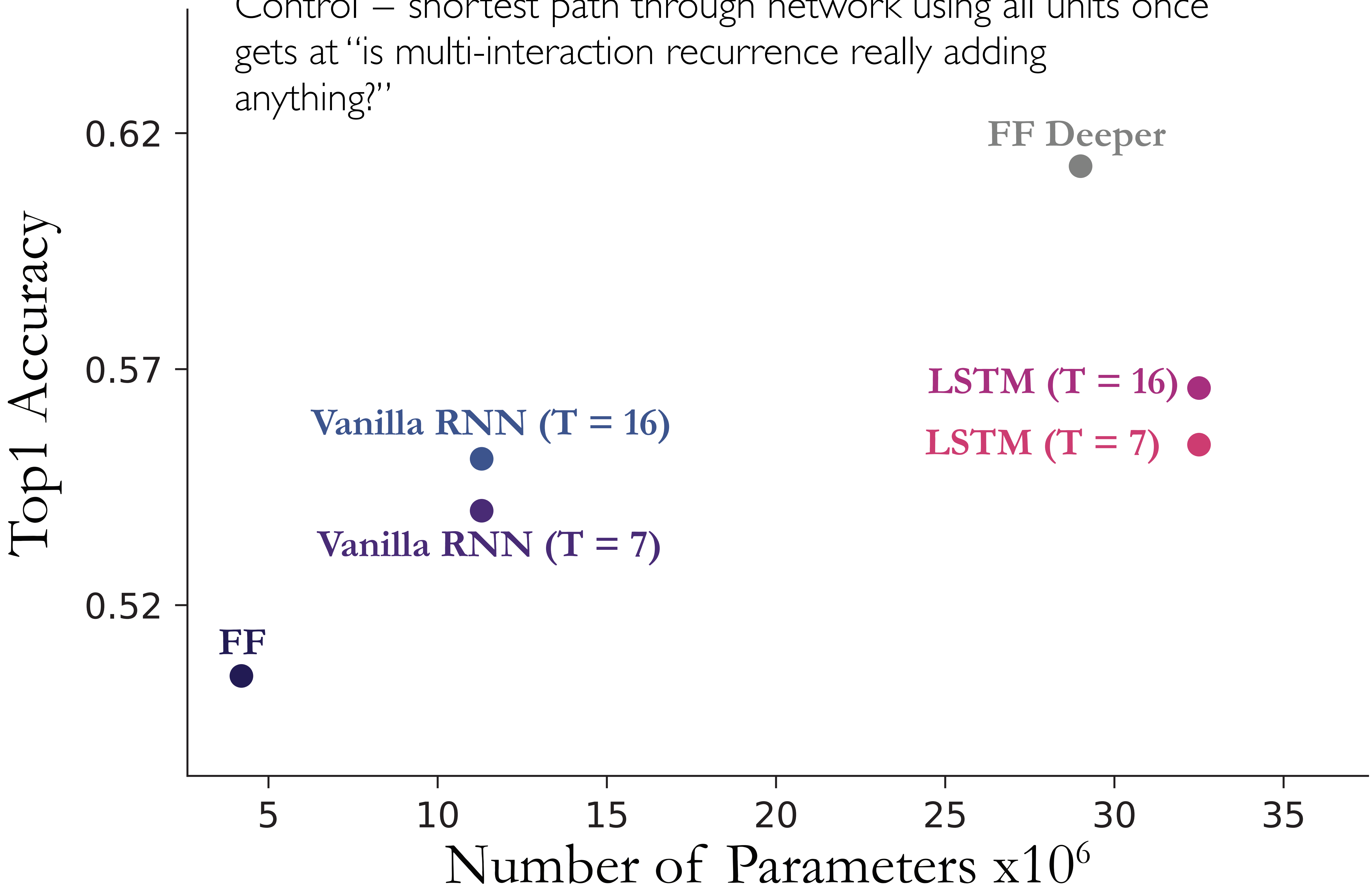
LSTM has **(1)** but not **(2)**; VanillaRNN has **(2)** but not **(1)**

# Not All Local Recurrence is Equal

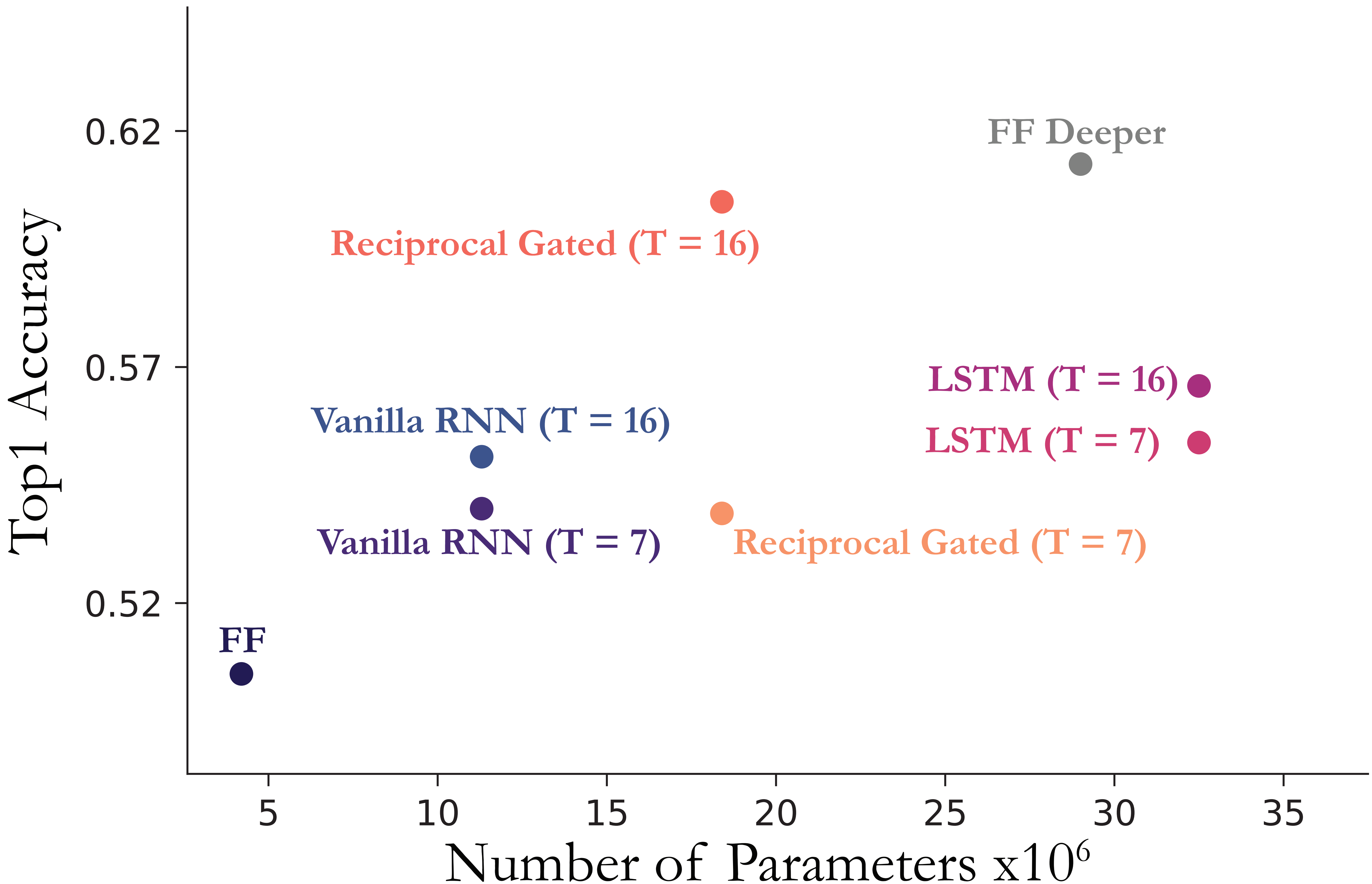


# Not All Local Recurrence is Equal

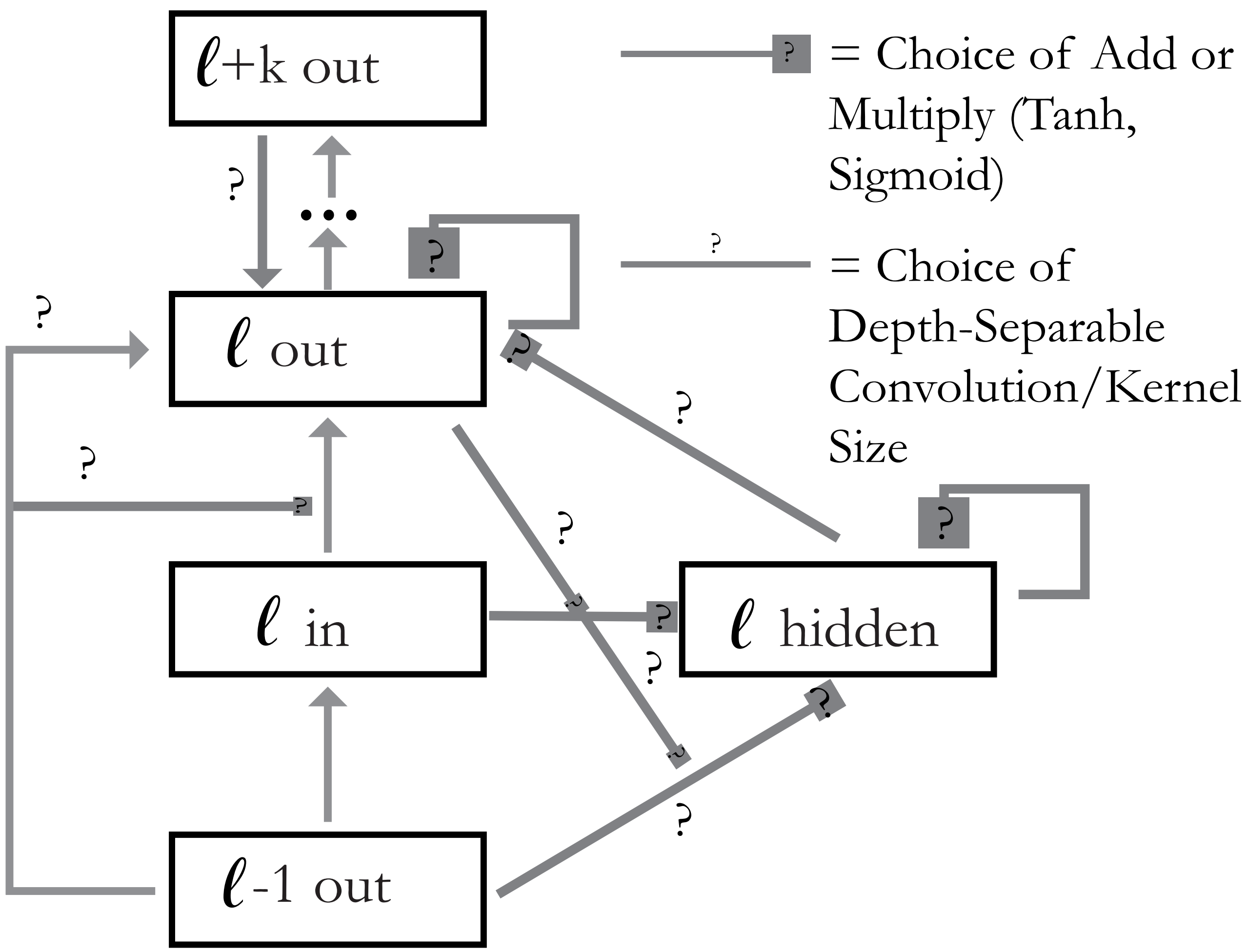
Control = shortest path through network using all units once gets at "is multi-interaction recurrence really adding anything?"



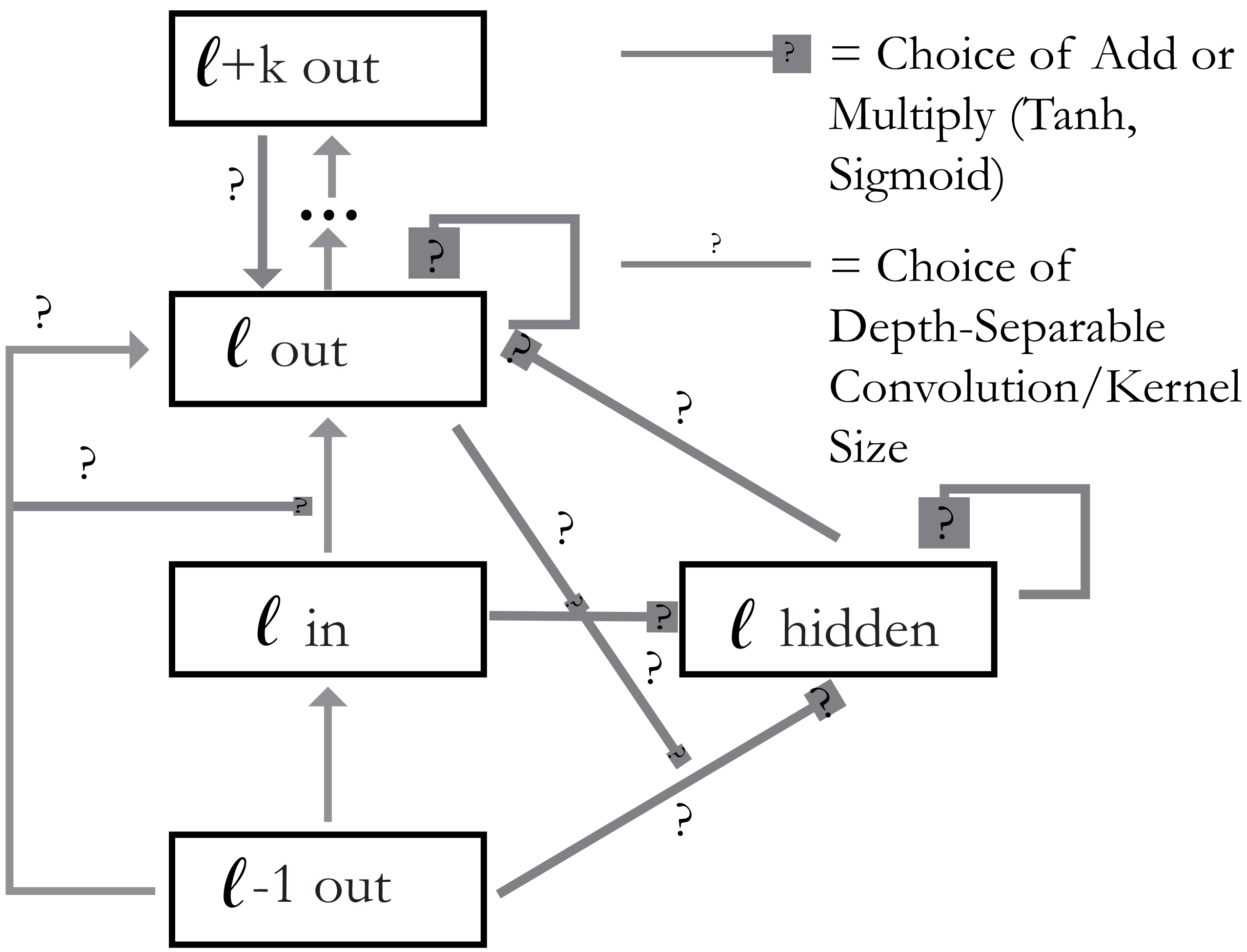
# Not All Local Recurrence is Equal

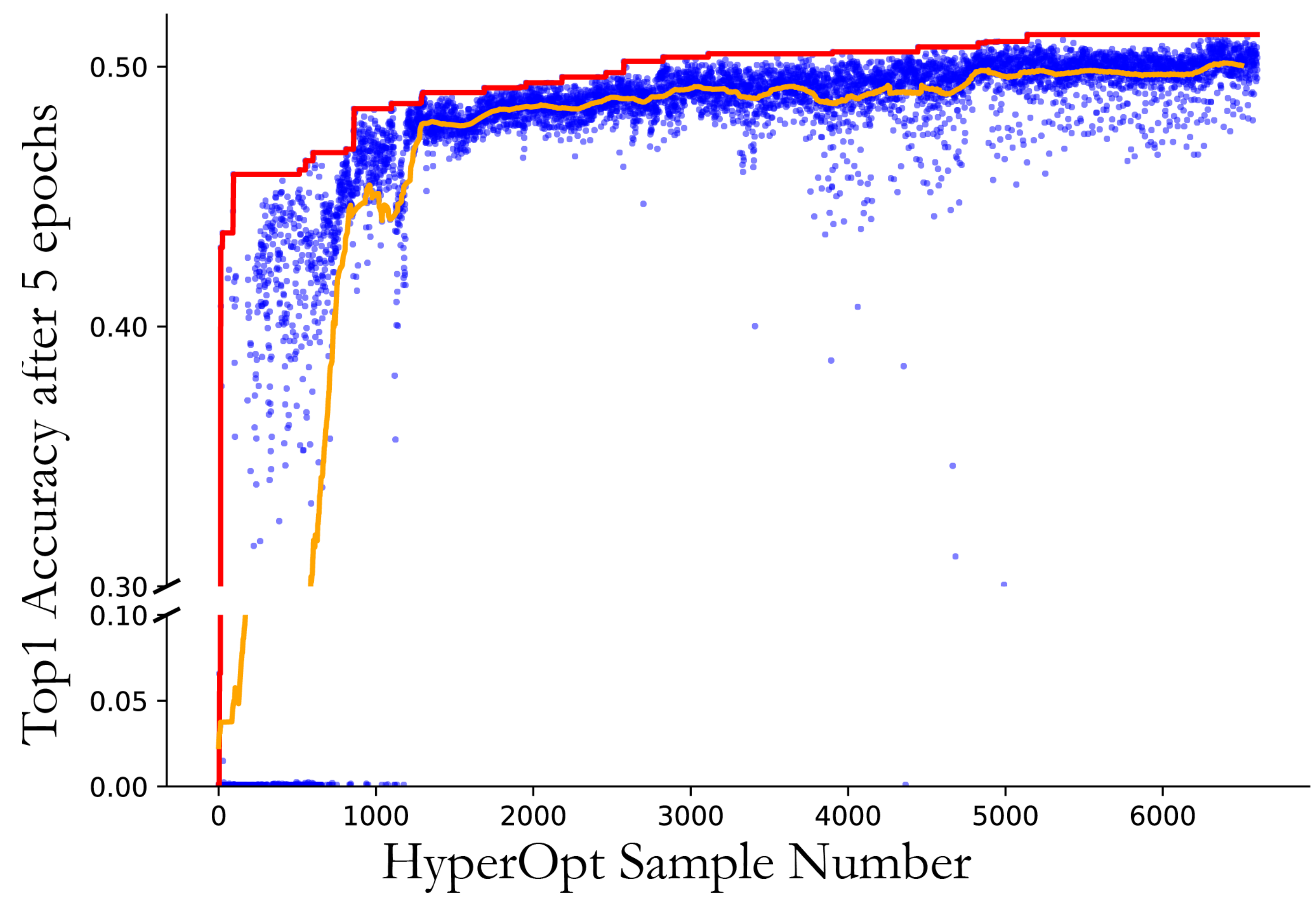


# Search Over Local and Global Recurrence

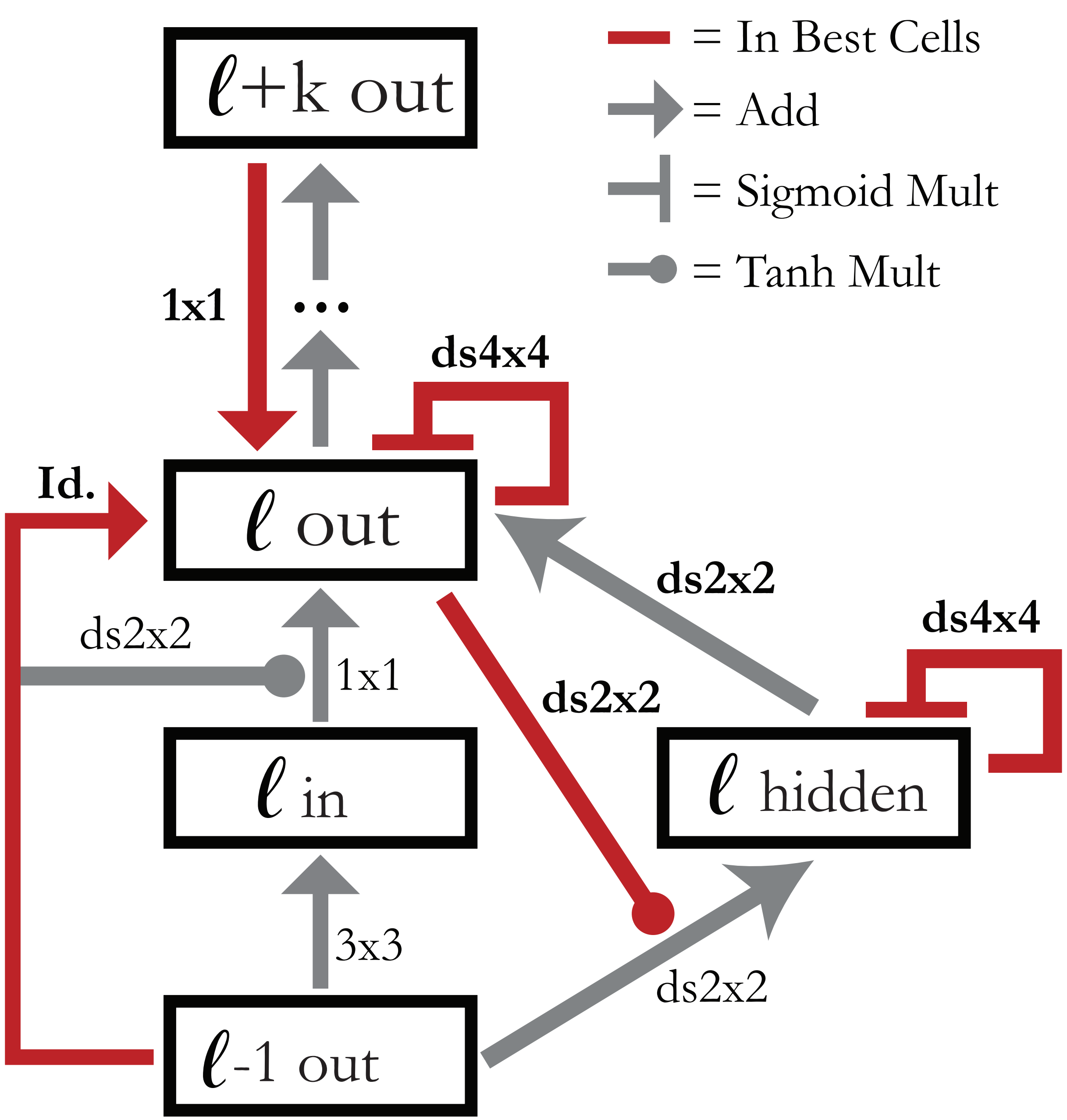


# Search Over Local and Global Recurrence

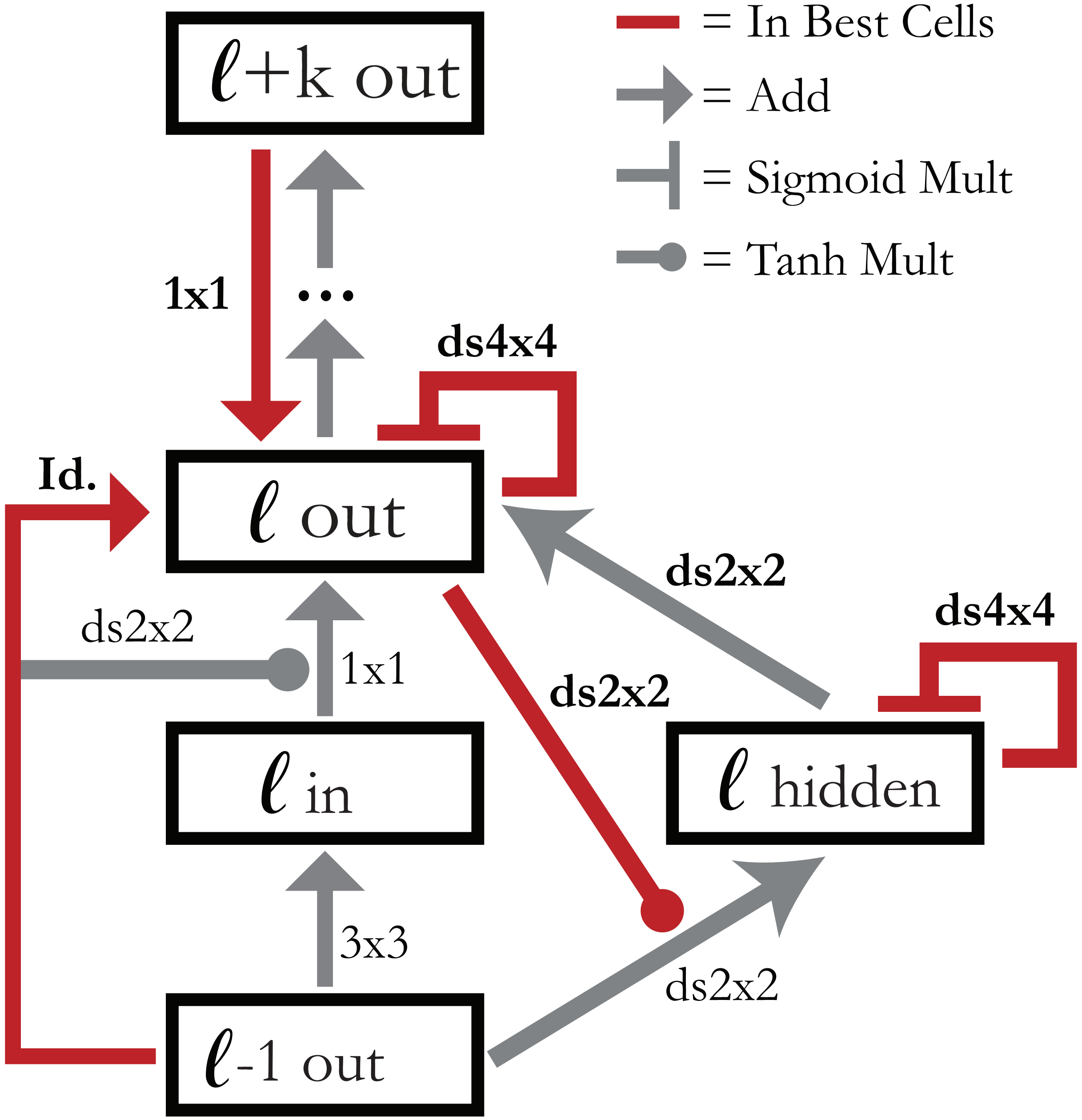


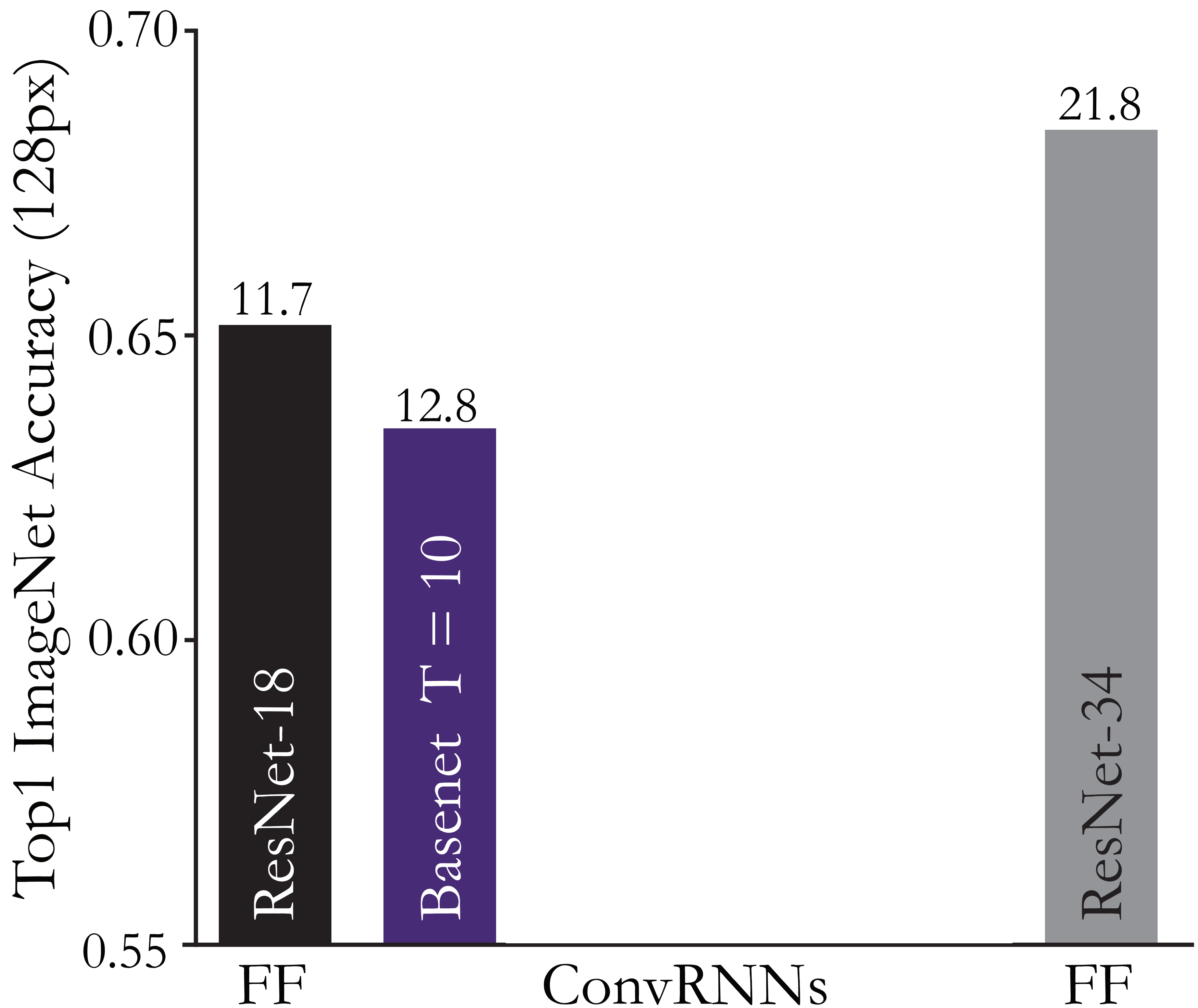
# Emergent Local and Global Connectivity Patterns



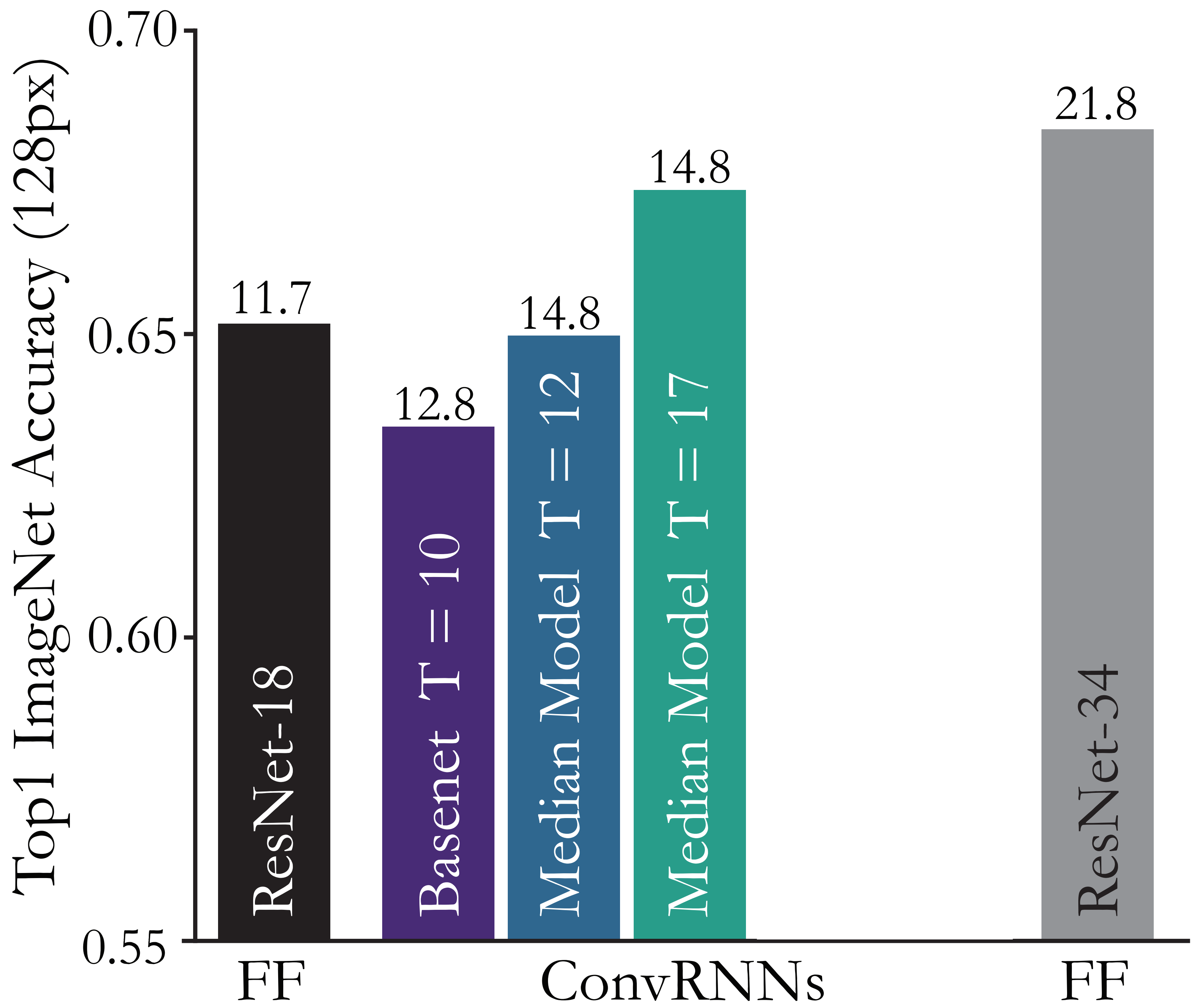
# Emergent Local and Global Connectivity Patterns



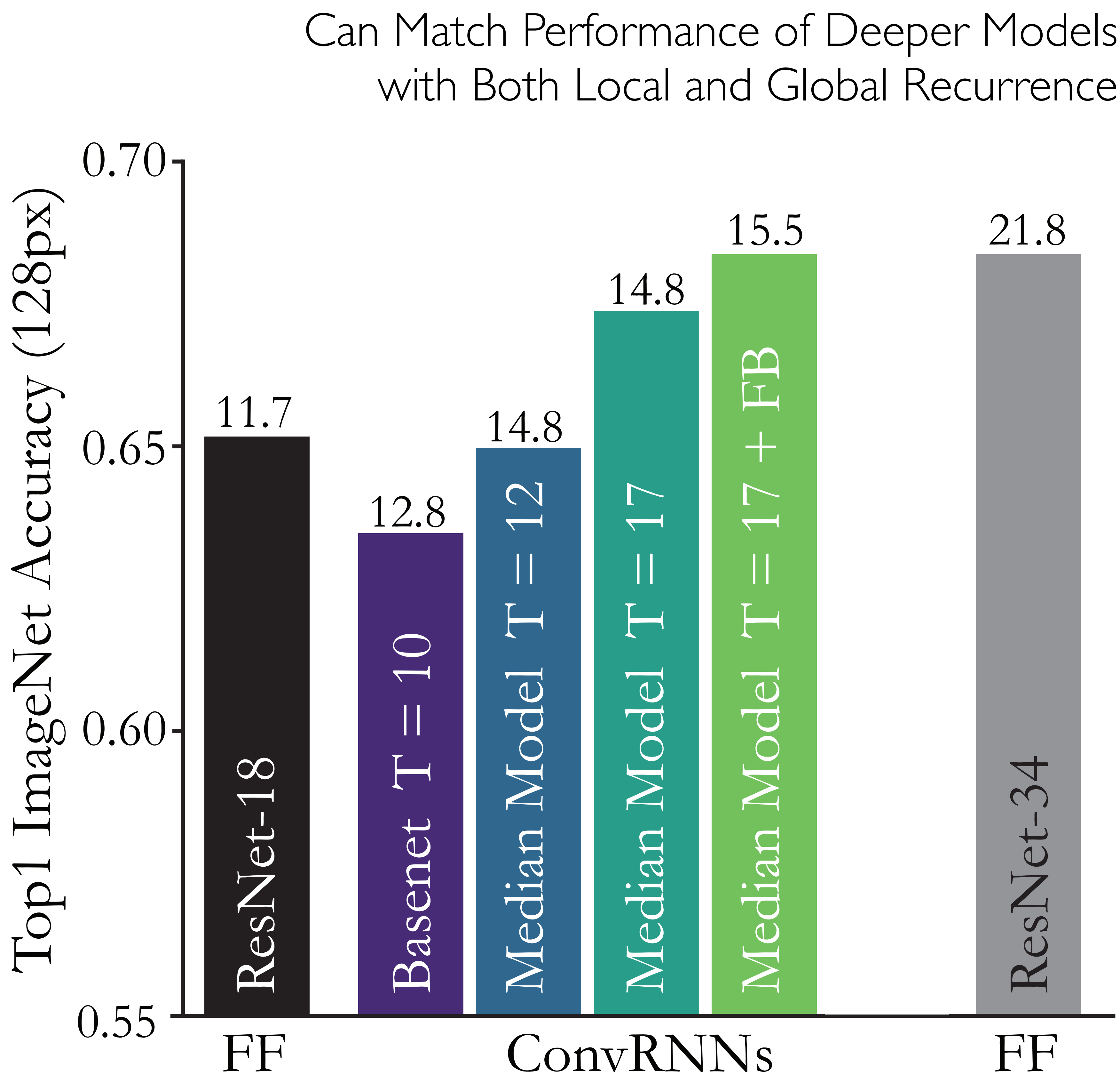

# Improving ImageNet Performance with ConvRNNs



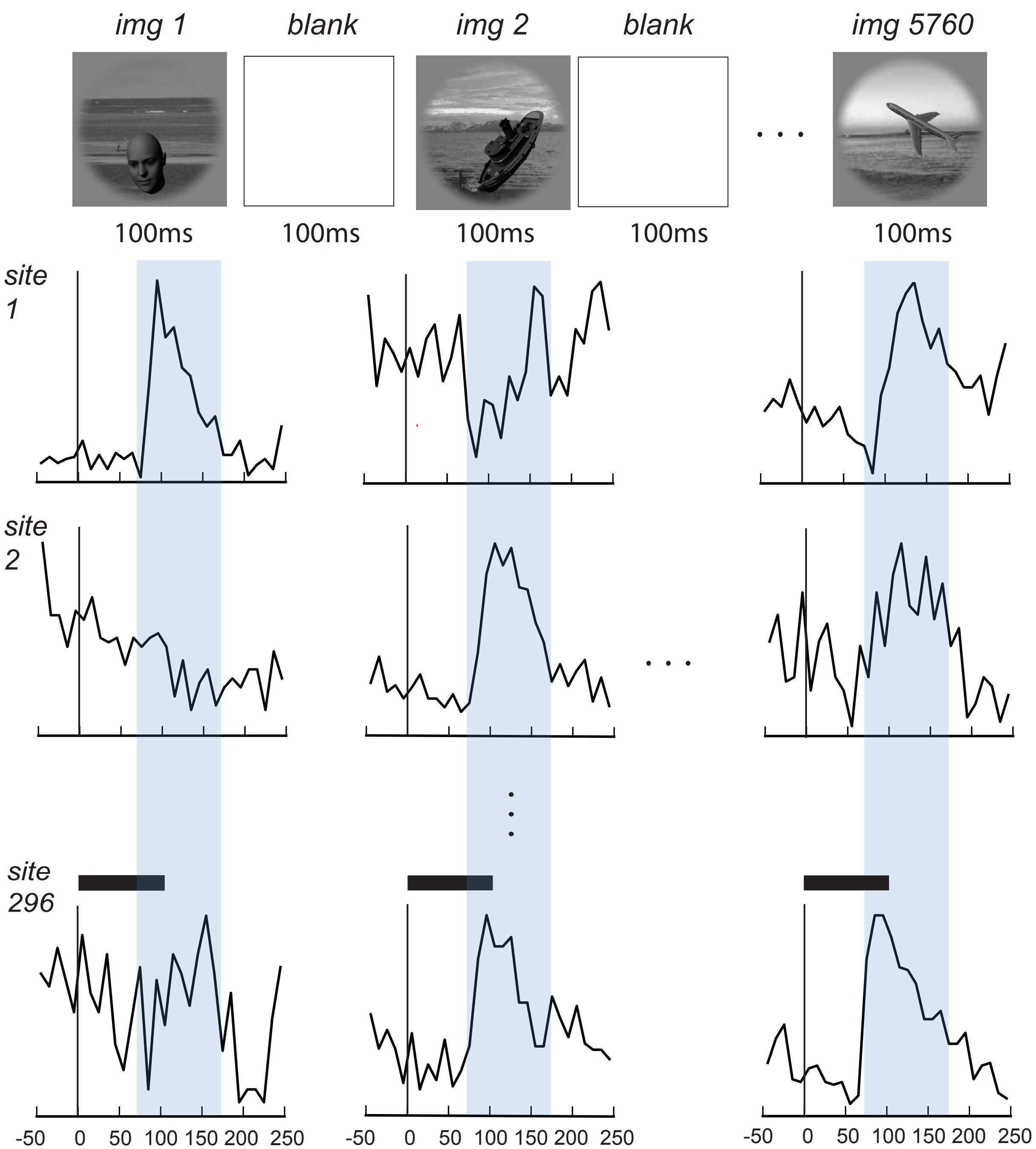
# Improving ImageNet Performance with ConvRNNs



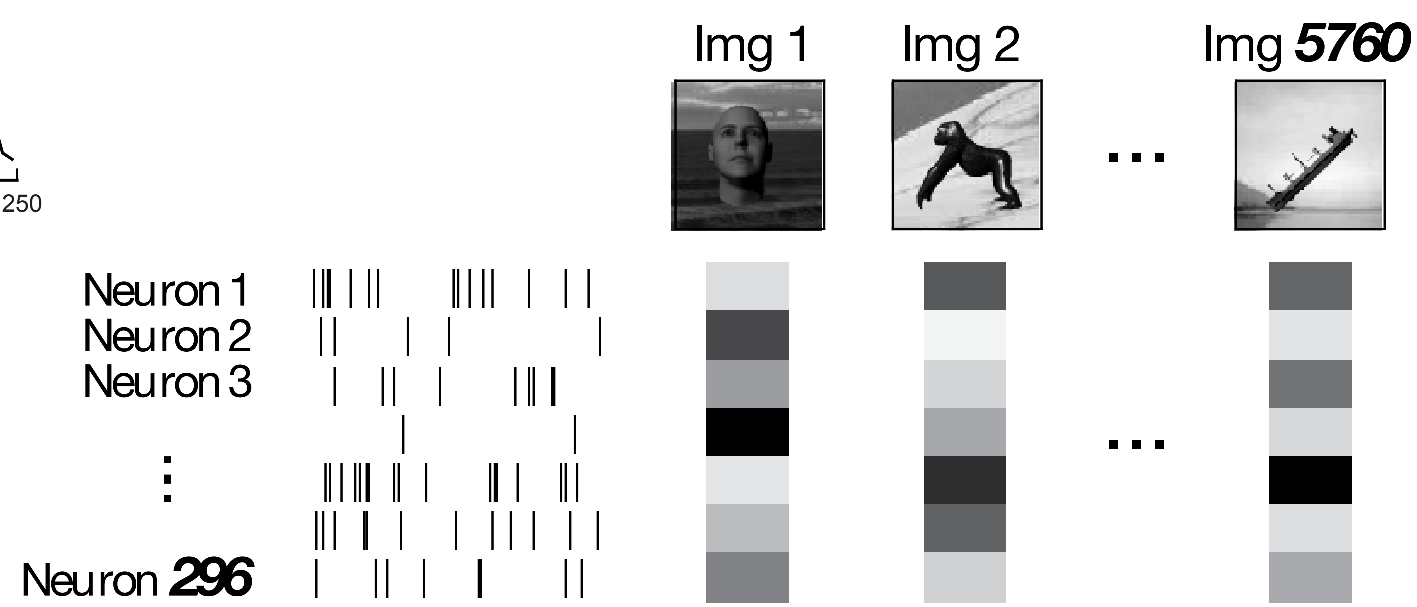
# Improving ImageNet Performance with ConvRNNs



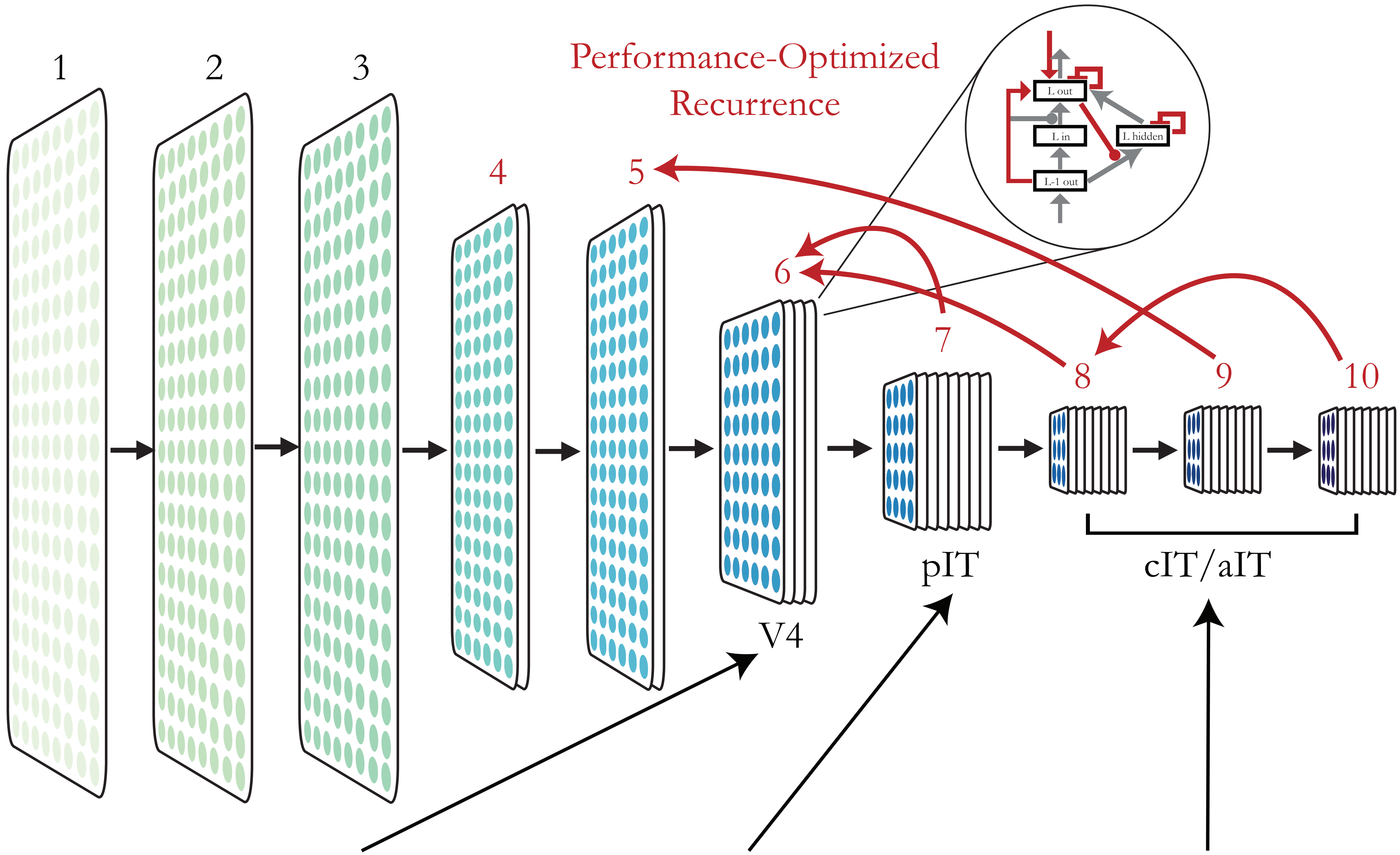
# Neural Predictivity with ConvRNNs



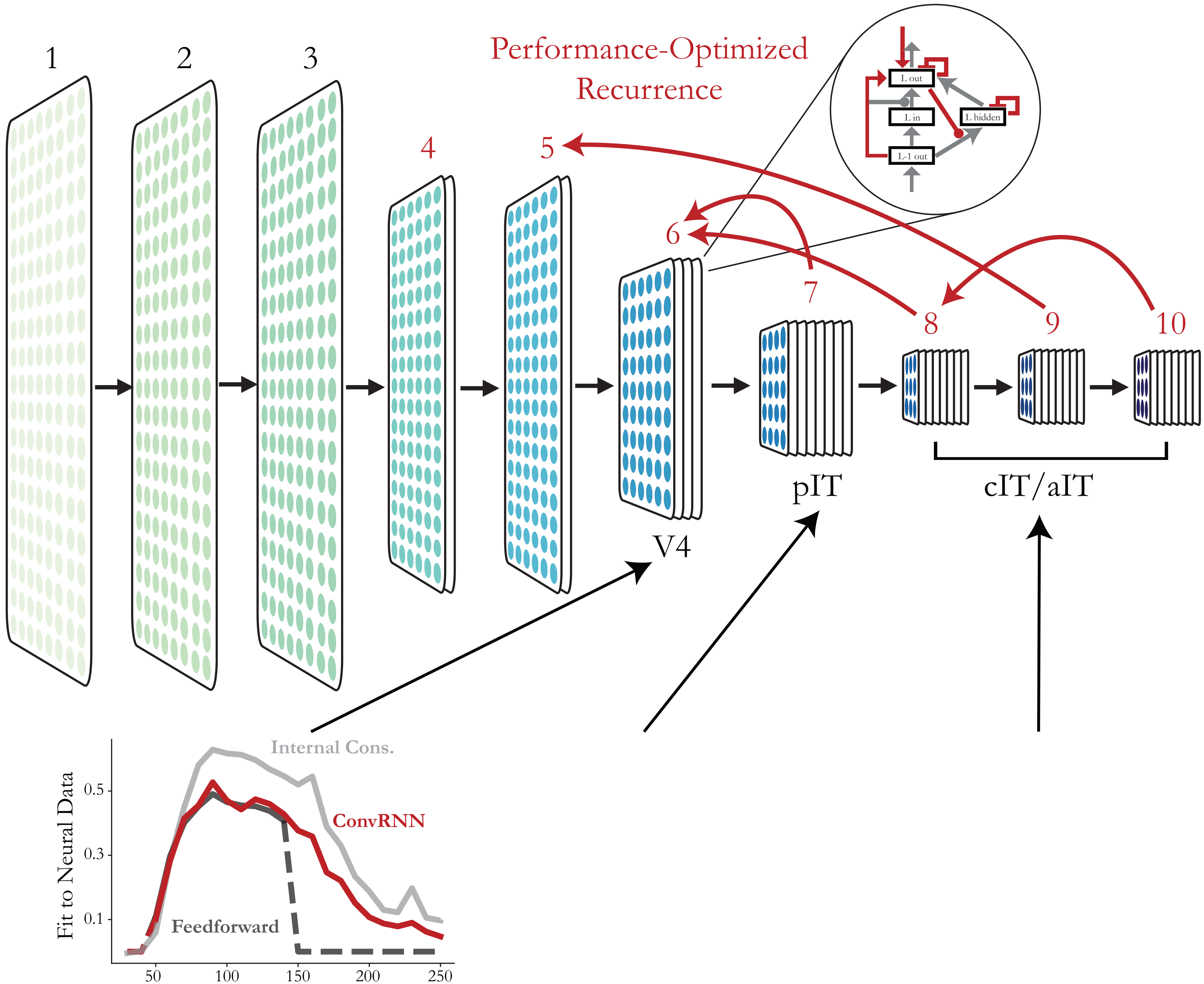

*Binned spike counts 70ms-170ms post stimulus presentation*



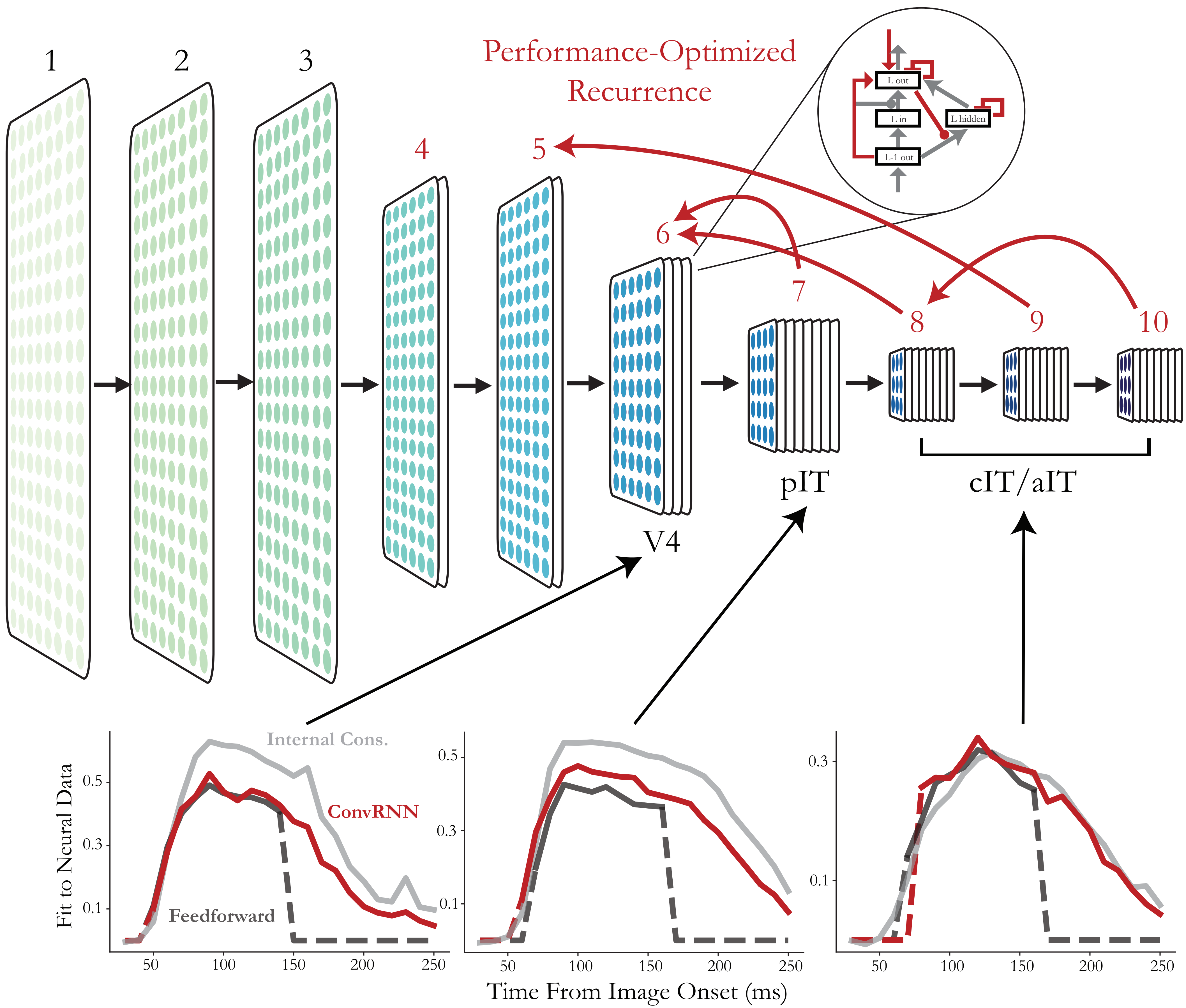
# Improved Neural Fit with ConvRNNs



# Improved Neural Fit with ConvRNNs



# Improved Neural Fit with ConvRNNs



# Conclusion

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

# Conclusion

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

- Exhaustive hyperparameter optimization results in strong selection for gating and bypassing in a local recurrent cell

# Conclusion

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

- Exhaustive hyperparameter optimization results in strong selection for gating and bypassing in a local recurrent cell

- With proper local recurrence in place, specific patterns of long-range feedback connections further improve performance

# Conclusion

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

- Exhaustive hyperparameter optimization results in strong selection for gating and bypassing in a local recurrent cell

- 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 dynamics in the primate ventral stream over feedforward models

# Conclusion

- At scale, one has to be careful with the choice of local recurrent architecture introduced into CNNs to improve performance
- Exhaustive hyperparameter optimization results in strong selection for gating and bypassing in a local recurrent cell
- 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 dynamics in the primate ventral stream over feedforward models
- Future work will explore the use of dynamic and self-supervised tasks for matching neural responses

# Collaborators

## Thanks!

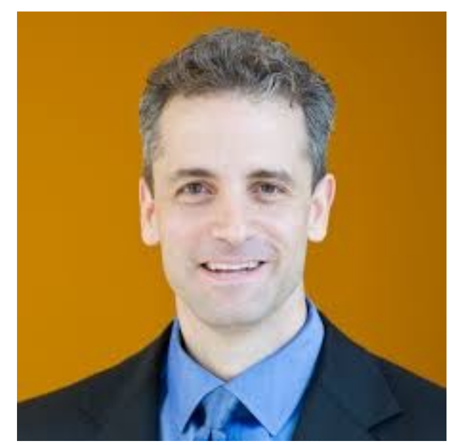
Dan Yamins

Dan Bear

Jonas Kubilius

Surya Ganguli

David Sussillo



Jim DiCarlo