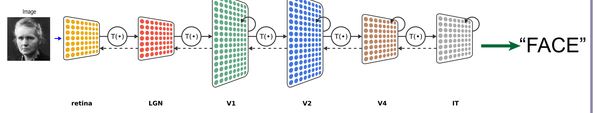
Convolutional recurrent neural network models of dynamics in higher visual cortex

III-83
COSYNE 2018

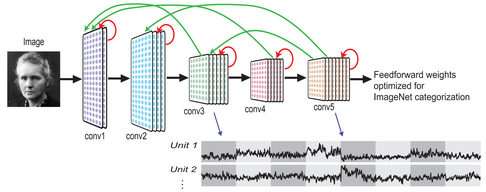
Aran Nayebi*[1], Jonas Kubilius*[2,3], Daniel Bear[1], Surya Ganguli[1], James J. DiCarlo[2], and Daniel L. K. Yamins[1]
[1]Stanford University [2]Massachusetts Institute of Technology [3]KU Leuven

Introduction & Approach

- The primate ventral visual stream for object recognition contains prominent corticocortical feed-forward and feedback connections.



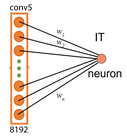
- For many images (we call them "control"), object category can be decoded from primate IT from feed-forward responses (Majaj, Hong, et al., 2015). But for some images (we call them "challenge"), decoding of object category in IT is delayed (Kar et al., 2017), suggesting the importance of feedback computations (currently not in use in the HCNN models).
- We propose convolutional recurrent networks (convRNNs) that have local and global feedback:



Recurrent weights optimized to match dynamics in V4 and IT

Predicting Neural Responses with convRNNs

Model features

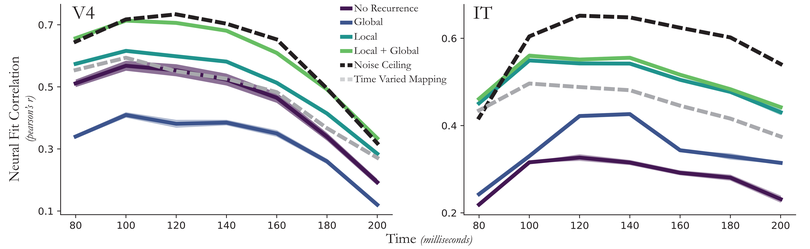


Neural Fitting Procedure

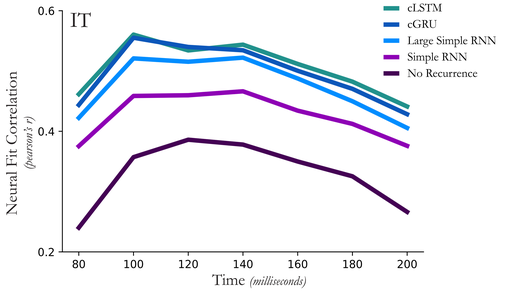
Randomly split the neural imageset into two groups (75% *train* / 25% *test*)

Use partial least square regression to estimate the mapping weights (**w**) (use *training* images and corresponding neural responses averaged across a specific time interval)

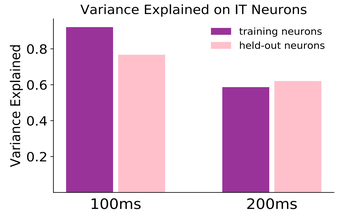
Given the weights, predict the neural responses of the *test* images.



- **Local recurrent** circuits substantially improve predictions of IT dynamics.
- **Long-range feedback** improves V4 predictions nearly to 100% of noise ceiling.
- **Local *gated* recurrence** is important for improved fit.

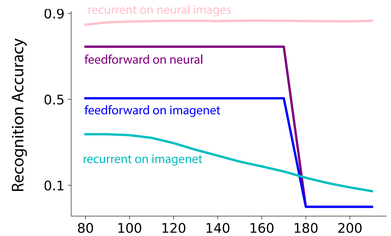


Generalization of convRNNs



Neural fits hold well even for cross neuron validation (as well as across held-out images and objects).

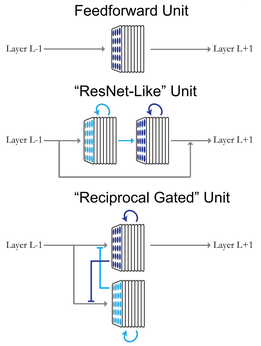
Performance on recognition task on neural images is improved
... but performance on ImageNet drastically worsens.

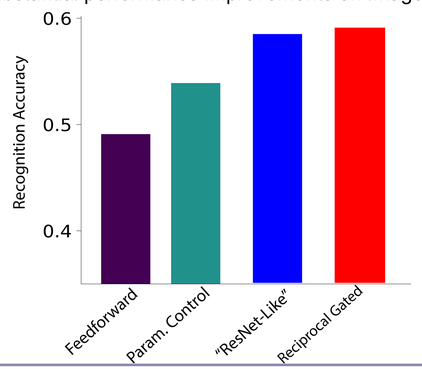


→ subtle "overfitting" to image-type.

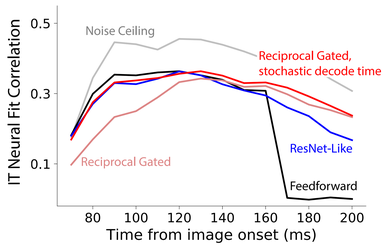
Improving ImageNet Performance with convRNNs

With better recurrent cells, substantial performance improvements on Imagenet:

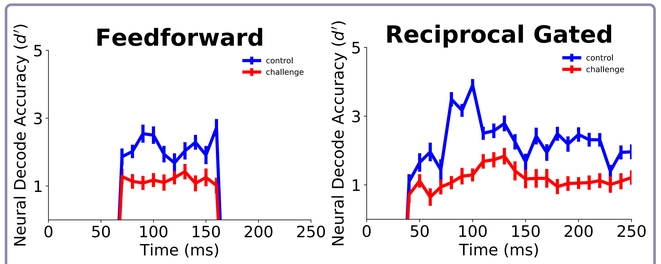




Good choice of decoder also leads to reasonable IT neural fits



Predicting Behavior (Preliminary)



- convRNNs seem to achieve higher neural decoding accuracy on both control and challenge images over a feedforward model.
- The solution for challenge images emerges later than the control images, exhibiting a longer temporal decoding latency.

Conclusions

- Recurrent convolutional models with feedforward task-optimized weights with global and local (gated) connections improve predictions of V4 and IT dynamics over feedforward models.
- Although these models generalize on held-out neurons and images, even demonstrating superior categorization performance on the neural images, their performance on ImageNet drastically worsens.
- With a good choice of recurrent cell and decoder, we can address this subtle overfitting, and obtain improved performance on ImageNet and predict neural responses at later timepoints over feedforward models and recurrent controls. Future work will address how the inclusion of such feedback connections improve performance on other visual tasks.

This research was supported by the European Union's Horizon 2020 research and innovation programme (No. 705498) to JK.