Recurrent CNN Modeling of Temporal Neural Response Recovers Connectivity Patterns of Early Visual Cortex

Harold Rockwell[1,2,4], Tai-Sing Lee[2,3,4]

[1]Department of Biological Sciences, [2]Center for the Neural Basis of Cognition, [3]School of Computer Science, [4]Neuroscience Institute; Carnegie Mellon University

Motivations and Summary

- Neural network models are often considered as uninterpretable “black boxes,” but under the right circumstances, they can provide inferences about the real neural systems they model.
- CNNs are well established as the best tools for predicting visual neurons’ response to complex natural image stimuli, but this body of work considers only the mean firing rate across the whole presentation time, ignoring the physiological and computational evidence for the importance of recurrence.
- We find that adding simple recurrence to a standard feedforward method of neural prediction allows us to successfully predict a discretized form of the PSTH with state-of-the-art performance.
- Furthermore, we find that the recurrent connectivity weights learned by our model have patterns consistent with functional connectivity studies in V1 (association fields and cross-orientation inhibition).

Methods

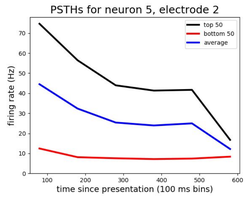
- Our data is composed of 34 consistent units from a multi-electrode array recording, in response to 2,250 natural images each shown for ten 500-ms presentations, collected by Gaya Mohankumar and Stephen Tsou.
- We use latent-space representations of the images, taken from the intermediate layers of a feedforward Imagenet-trained DenseNet, as the input to a single learned recurrent convolutional layer and standard factorized readout (fixed throughout time steps) to predict the spike count within 100 ms time bins.
- The recurrent update equation works as follows:

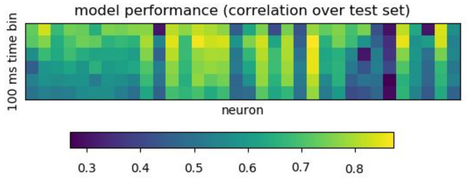
$$H_t = \phi(W_f X + W_r H_{t-1} + b)$$

(where H_t is the layer’s activation at time bin t, H_0=0, ϕ is a Softplus nonlinearity, X the input, and W_f and W_r the feedforward and recurrent convolutional weight matrices)

Results

Predictive performance remains high for individual time bins

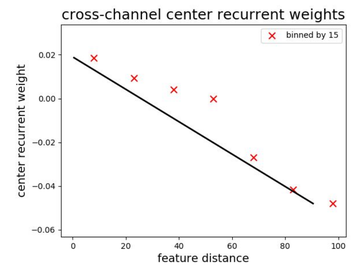




Left: some example discretized PSTHs, and a plot of predictive performance, measured by correlation with the true response over the test set, for each time bin and neuron.

A feedforward model trained on the full 600 ms response reaches an average correlation of 0.74, while the average correlation in each time bin of our model is 0.63. However, when the response of our model is summed across time bins and compared with the full 600 ms response, its average correlation is 0.75; comparable to the model trained directly on that response.

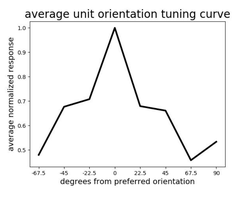
Convolutional layer units with less similar feature tuning have more negative connections

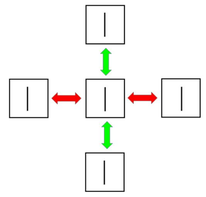


Left: binned values of the center recurrent weight between channels of the model’s recurrent convolutional layer, plotted against the normalized distance of their test-set tuning curves. The relationship is weak but significant, with a correlation of -0.20 (considering all values; not binning).

The less similar the feature tuning of two channels, the more negatively they interact when in the same location. This competitive interaction is expected in cortical connectivity from the theory of cross-orientation inhibition.

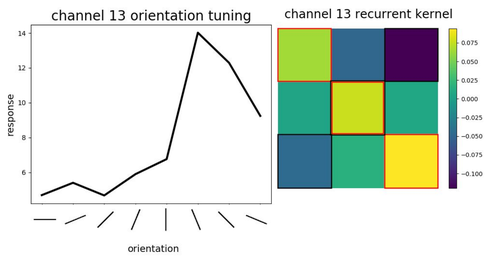
Convolutional layer units’ orientation tuning aligns with their recurrent kernels

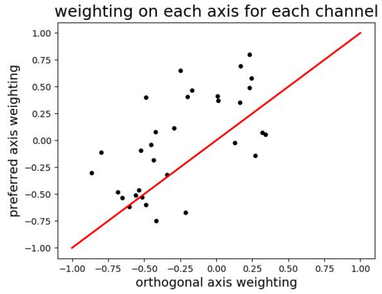




Left: Average feedforward orientation tuning of our model’s convolutional layer, and an illustration of the association field. Bottom left: an example of how the alignment is computed for a single channel (bottom). Below: the kernels’ weighting on the preferred axis versus the orthogonal one for each channel.

The average difference between the recurrent kernels’ weighting on the preferred orientation and the orthogonal is 0.23 (±0.06 SE), out of a maximum possible 0.66.





References:

Blakemore et. al (1970). “Lateral inhibition between orientation detectors in the human visual system.” https://doi.org/10.1038/228037a0

Cadena et. al (2019). “Deep convolutional models improve predictions of macaque V1 responses to natural images.” https://doi.org/10.1371/journal.pcbi.1006897

Kapadia et. al (1999). “Dynamics of spatial summation in primary visual cortex of alert monkeys.” https://doi.org/10.1073/pnas.96.21.12073