A Large Dataset of Macaque V1 Responses to Natural Images Revealed Complexity in V1 Neural Codes

Shang Gao[1], Tianye Wang[2], Xie Jue[3], Daniel Wang[1], Tai Sing Lee[1†], Shiming Tang[2†]

[1]Carnegie Mellon University, Pittsburgh, USA [2]Peking University, [3]Chinese Institute for Brain Research, Beijing, China.

(†, corresponding author: taislee@andrew.cmu.edu, tangshm@pku.edu.cn)

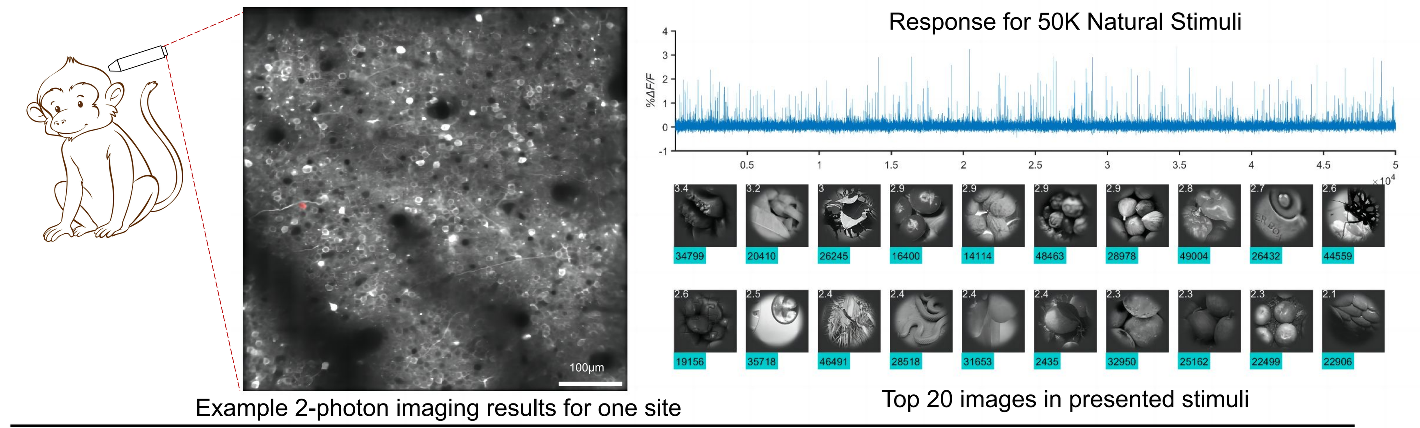
Virtual poster QR code

Motivations & Background

Research shows that Convolutional Neural Network is an excellent tool for modeling neural representation and computation in the ventral stream of macaque monkeys (Yamins and DiCarlo 2016) and for characterizing V1 receptive fields in monkeys (Zhang et al. 2019) and mice (Candena et al., 2019, Walker et al., 2019.). Our earlier studies showing that V1 neurons are tuned to more complex features beyond orientation-selective Gabor filters (Tang et al., 2018a., Zhang et al. 2019) were based on a parametric artificial pattern and might be biased. The extreme sparsity of macaque V1 responses to natural images in 2-P calcium imaging had prevented adequate CNN fitting of macaque V1 receptive fields (Tang et al. 2018b). This study provides a large dataset for better characterizations of the neural codes of macaque V1 neurons.

Experiments & Data

Using two-photon imaging with GCaMP5, we measured the responses of 1689 neurons from 6 sites of three awake behaving macaque monkeys to 30k-50k natural images. About 300 cells from each site were tracked across five days anatomically and based on responses to 200 fingerprint images. Monkeys performed fixation task. The images were presented in sequence with 1 second per image preceded by 1 second of gray screen. The 30k-50k images in the training set were presented once, 1000 images in the validation set were tested once with 10 repeats.

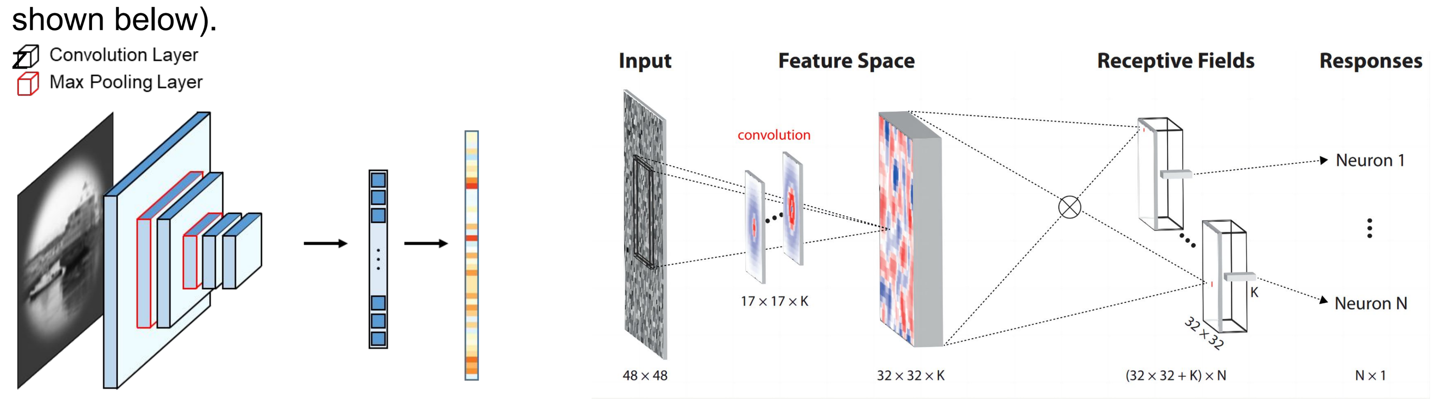


Example 2-photon imaging results for one site

Top 20 images in presented stimuli

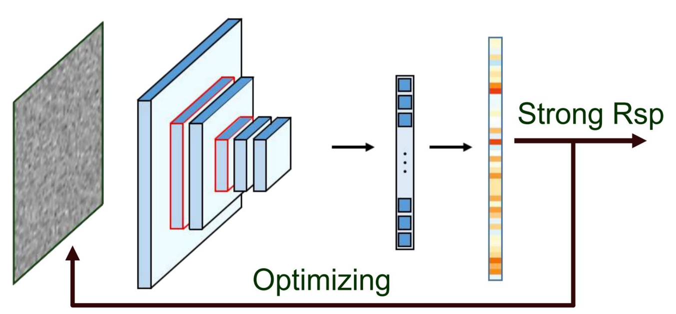
CNN modeling of V1 RFs

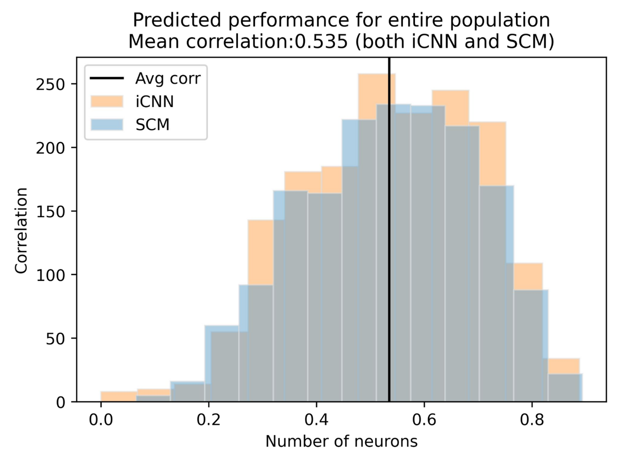
Individual CNN (iCNN) (Zhang et al. 2019) or shared core CNN (SCM) (Klindt et al. 2017), (with 4 conv layers) were fitted to the responses to 30k-50k training images to predict the responses to 1000 validation images. The metric used to evaluate the models was the Pearson correlation between neurons response and model-predicted responses. Predicted correlation for entire population of neurons is around 0.53. (Histogram of the performance distribution iCNN vs SCM shown below).

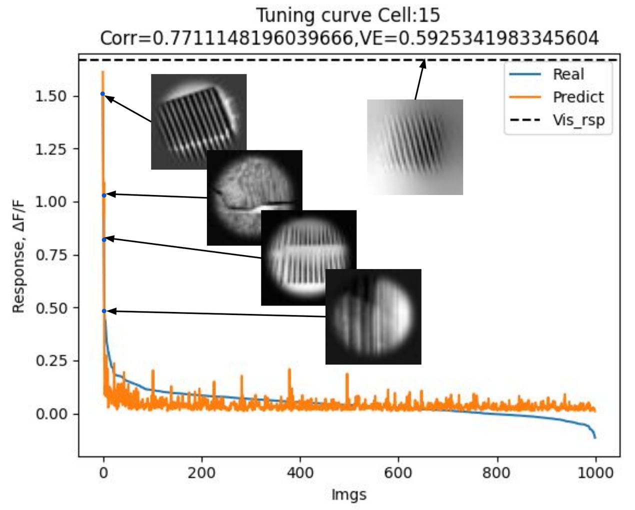


Individual-CNN, with 4 Convolutional layers and 2 Max Pooling layers

Shared-core model. Proposed by Klindt, et al. in 2017



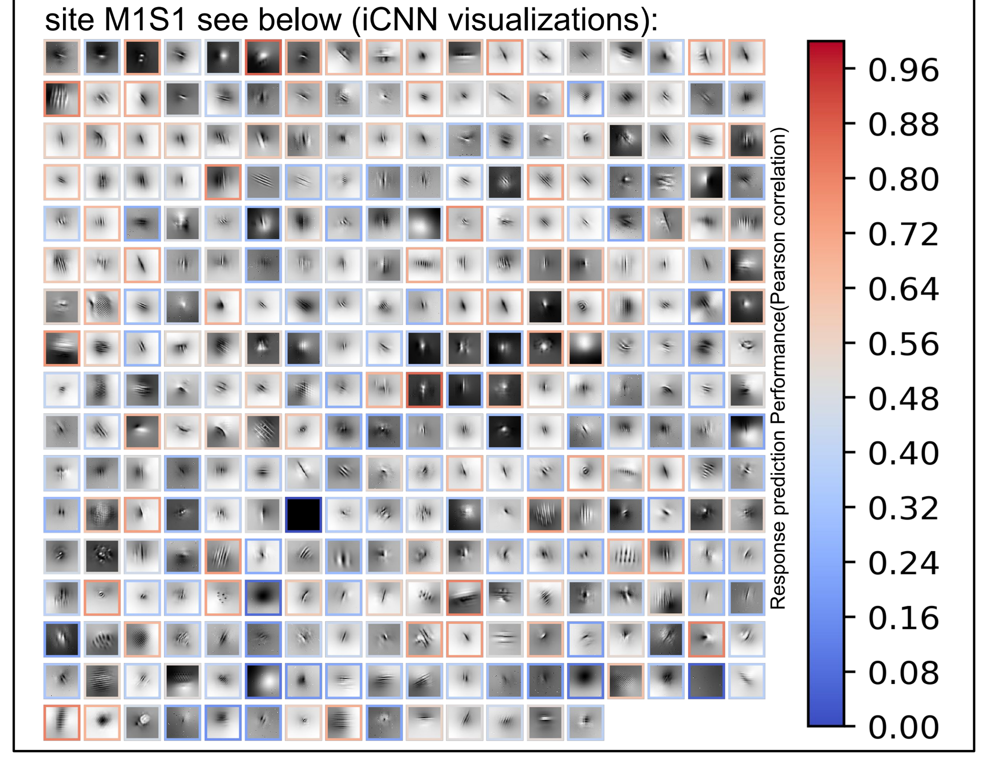




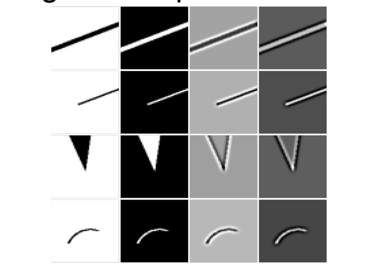
Ranked tuning curve (blue curve) for an example neuron against the prediction of the model (orange curve) with performance of 0.77 correlation. The neuron's preferred image is "visualized" by optimizing the input image via backpropagation to maximize the responses of the neurons. The dashed line shows the response of the model to the visualized image (see above). Validation set images with top responses were shown for comparison.

Diversity and Complexity of V1 tunings

Visualizations: By visualizing all neurons in the collected V1 data, we can observe complex tuning that differs from traditional Gabor-oriented filters. All neurons in a example site M1S1 see below (iCNN visualizations):

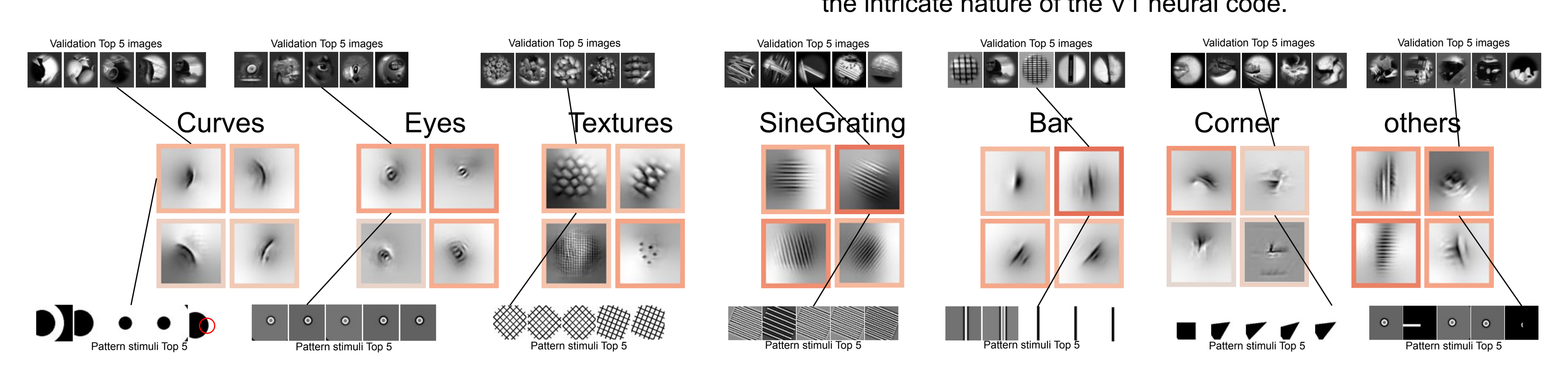


Augmented pattern stimuli



We tested the fitted CNN models with the Pattern stimuli used in Tang et al. 2018a, augmented with reverse contrast and Laplacian filtered versions to compare neurons' visualization to their classification based on responses to curves, junctions, bars and edges in the earlier study.

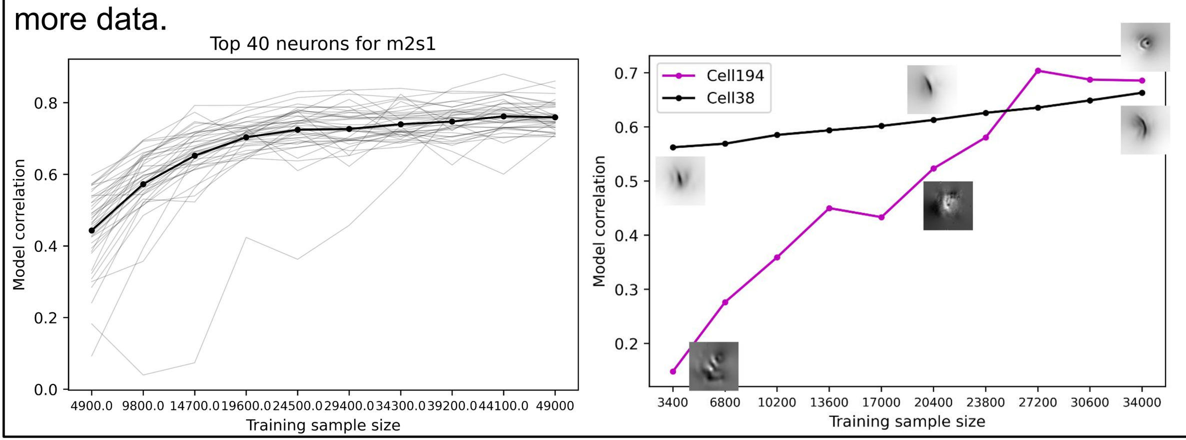
V1 neurons forms certain clusters: V1 neurons exhibit clustering behavior, as evidenced by the example neurons presented in the center panel below, each belonging to a labeled cluster. These neurons exhibit different patterns, such as curves, textures, and even eye-specific neurons. The top panel displays the top 5 validation images that elicited the largest responses from each neuron, while the bottom panel shows the top 5 Pattern stimuli. Notably, the preferred validation images and pattern stimuli of the neurons share similar shapes with the visualizations, indicating the intricate nature of the V1 neural code.



Data size matters

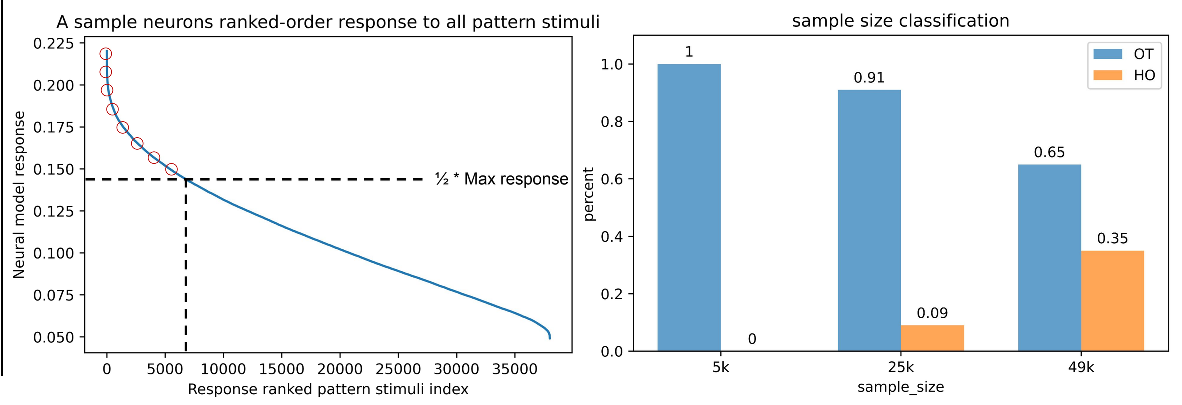
Left panel shows the response prediction performance of 40 top-performing neurons from site M2S1 as a function of the amount of data used to train the models. It shows the general trend, also evident for the entire population, response prediction performance improves with the number of training samples used.

Right panel shows the visualizations of two example neurons increase in complexity with more and more data.



One cell evolves from a traditional oriented edge to a curve while the other evolves from a grating to an eye.

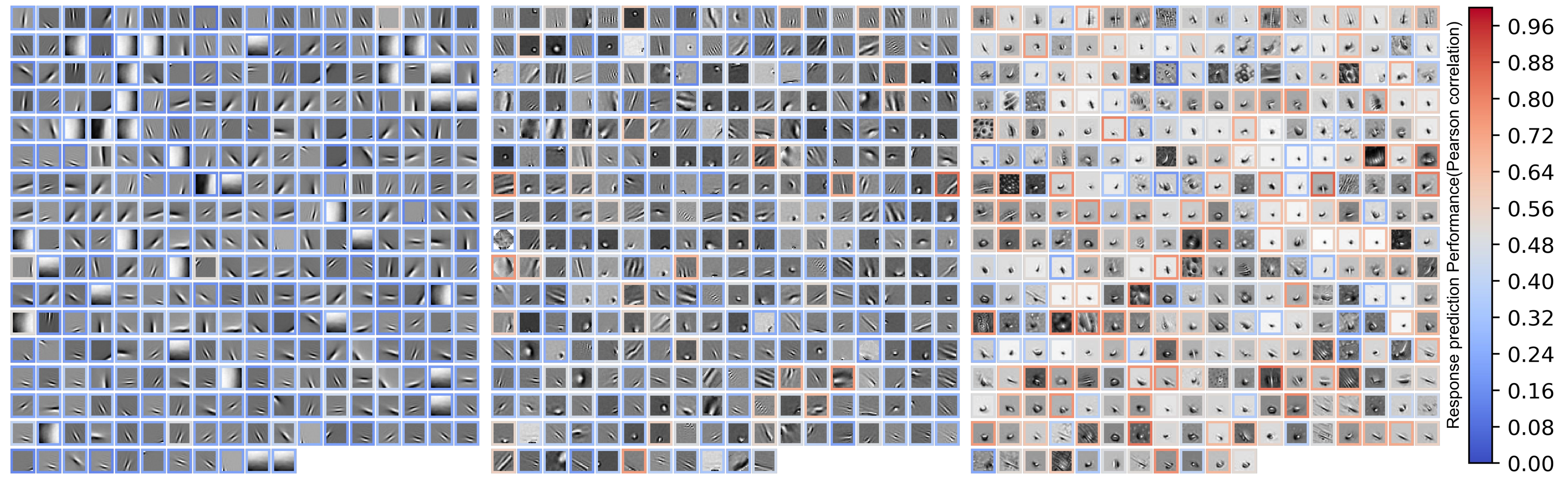
As the size of the dataset grows, we observed more neurons become classified as Higher-Order (HO). A neuron is classified as HO when all the pattern stimuli that elicit responses greater than 50% of its peak response belong to one or more higher order categories (curves, corners, crosses, rings). Neurons with oriented bars or edges above half-peak response are classified as OT (Oriented-Tuning).



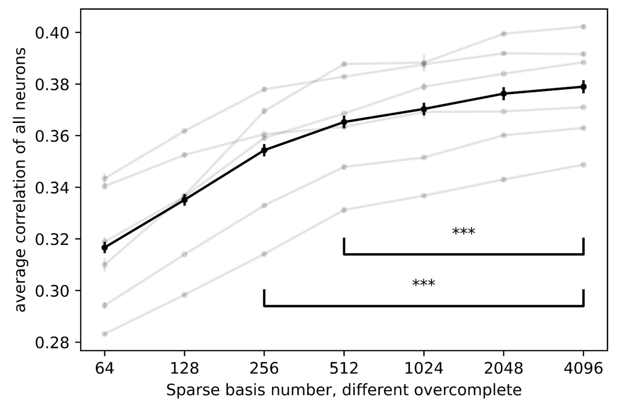
With limited data, neurons are predominantly classified as oriented tuned (OT), but models trained with more data revealed a higher proportion of neurons tuned to complex patterns (HO).

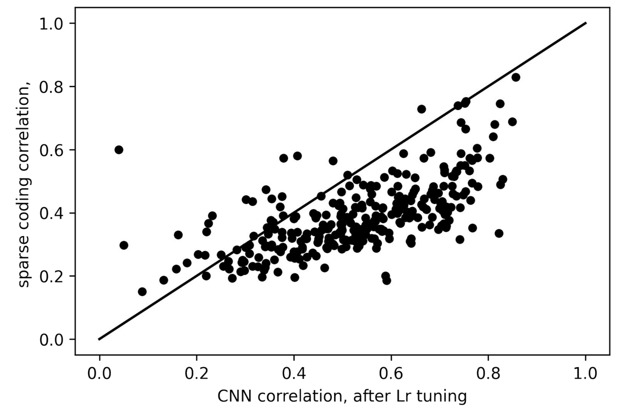
Receptive fields based on Over-complete sparse coding fit better than complex sparse coding

Overcomplete sparse coding (cite Olshausen, LeCun, Sommer) also yields more complex receptive fields than standard sparse coding (Olshausen and Field 1996). The best fitted "overcomplete" codes for neurons in site m2s1 (Middle Panel, 16X overcomplete) revealed curvature and corner neurons versus the standard Gabor filters in the standard sparse coding (Left Panel). CNN models' visualization (Right Panel) show a greater degree of diversity and complexity in neural codes.



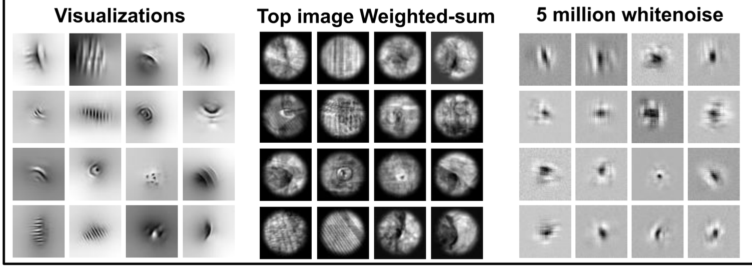
We evaluated the performance of linear-nonlinear (LN) models in predicting neural responses using linear filters learned through sparse coding from 16 x 16 natural image patches. Our results indicate that prediction performance improves with increasing overcompleteness of the representation, as shown in the right graph. To account for rotation and translation variations, we rotated each filter set by 18 orientations, shifted them by 25 positions, and reversed the contrast before identifying the best filter for each LN model to predict neuron responses.

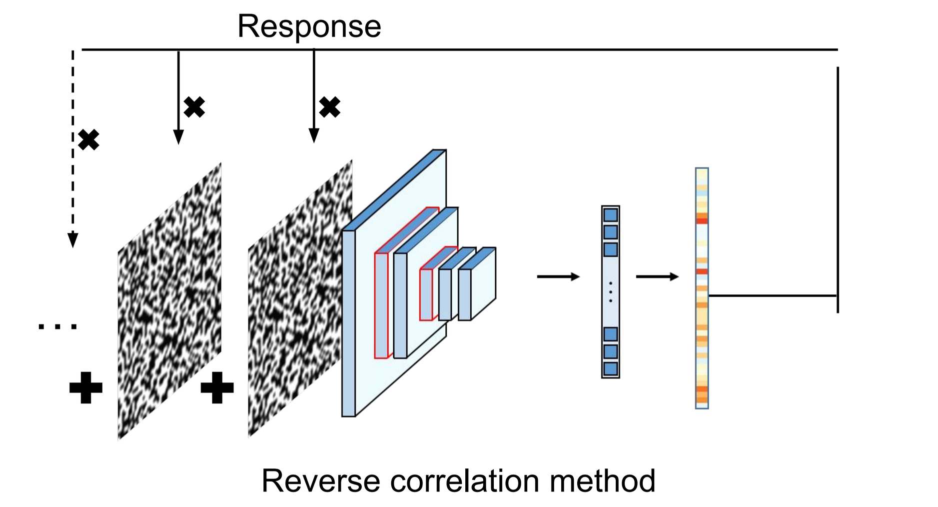




Natural images matters

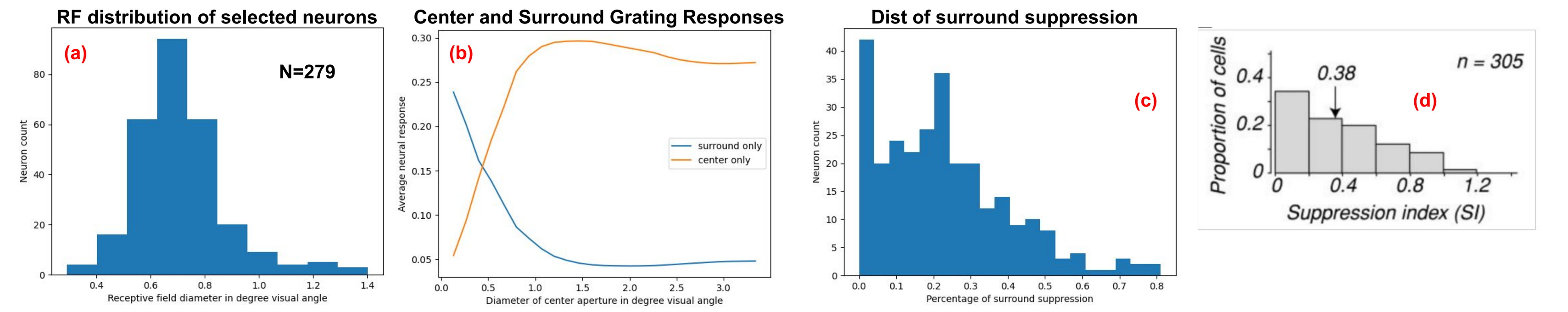
Natural images, with rich features, are crucial for recovering complex receptive fields, in addition to the amount of data. Testing CNN models with white noise image, we found that standard reverse-correlation techniques fail to recover complex pattern receptive fields even with 5 million white noise patterns. CNN visualization, top response weighted average stimuli, as well as the receptive fields recovered from white noises are shown for comparison.





Surround suppression

We tested the CNN models of 279 neurons (all sites combined) with good response prediction performance (> 0.7 in Pearson Correlation) with sine-wave grating with size, ranging from 1X to 7X receptive field size. Interestingly, these neurons trained with natural images exhibit the classical surround suppression effect automatically. (a) RF distribution of the neurons, as mapped by bars. (b) averaged responses of the selected CNN neurons to a sine-wave grating (averaged over 4 phases) of each cell's preferred orientation and spatial frequency inside (grating center-only) or outside (gray center, grating surround-only) apertures of different diameters. (c) distribution of the magnitude of the surround suppression index (MaxRsp - MinRsp) / MaxRsp, which is very similar to that reported in Cavanaugh et al. (2002) (d).



Summary and conclusion

- We collected extensive data on the response of 1689 Macaque V1 neurons to 30k-50k natural images. Using this data, we developed neural network models that more accurately predicts neural responses, and characterizes the receptive fields of the neurons.
- Our findings suggest that V1 neurons exhibit complexity beyond traditional oriented Gabor and tuned to curves, textures, eyes, and other higher order features.
- We demonstrate a large data set of natural images is important for revealing the complexity of receptive fields that white noise stimuli fail to recover.
- We also found complex receptive fields predicted by overcomplete sparse coding fit neural responses better than standard sparse coding, though still not as powerful as the CNN models.
- The CNN models automatically exhibit surround suppression, suggesting that models have captured neurons' sensitivity to context, and that these CNN models can potentially be used as neurons-in-silicon for carrying out "neurophysiological experiments".

Funding & References

Supported by the National Natural Science Foundation of China U1909205, and Center for Life Sciences, Peking University, China (Tang) and NSF grant CISE RI 1816568 and NIH R01 EY030226-01A1, USA (Lee).

- Cadena et al., 2019. PLOS Comp. Biol.
- Klindt et al., 2017. NeurIPS.
- Tang et al., 2018a. Curr. Biol.
- Tang et al. 2018b. e-Life
- Walker et al., 2019. Nat. Neurosci.
- Yamins and DiCarlo 2016 Nature Neuro.
- Zhang et al. 2019. J. Com Neuro.
- Cavanaugh, Bair and Movshon (2002) J. Neurophysiology.