

# **Deep convolutional neural network models of the retinal response to natural scenes**

Lane McIntosh\*, Niru Maheswaranathan\*, Aran Nayebi, Surya Ganguli, Stephen Baccus

A central challenge in sensory neuroscience involves understanding neural computations and circuit mechanisms underlying responses to ethologically relevant, natural stimuli. However, the ubiquity of cascaded nonlinear processes like synaptic transmission and spiking dynamics in multilayered circuits has presented significant obstacles to the goal of learning accurate computational models of circuit responses to natural stimuli from neural recordings.

To address this, we employ deep convolutional neural networks (CNNs), which demonstrate success at many pattern recognition tasks (LeCun et al. 2015). These models cascade multiple layers of filtering and rectification — exactly the elementary computational building blocks thought to underlie complex functional responses of sensory circuits. Previous work utilized these models to understand properties in IT cortex (Yamins et al. 2013), but not in early sensory areas where knowledge of neural circuitry can provide important validation for such models.

We demonstrate that CNNs are considerably more accurate at capturing retinal responses to held-out natural scenes than linear-nonlinear (LN) models and related models, such as generalized linear models (GLMs). Furthermore, we find CNNs generalize significantly better across classes of stimuli (white noise vs. natural scenes) they were not trained on. Remarkably, analysis of these CNNs reveals internal units selective for visual features on the same small spatial scale as the main excitatory interneurons of the retina, bipolar cells. Moreover, probing the model with reversing gratings, paired flashes, and contrast steps reveals that the CNN learns nonlinear retinal response properties such as frequency doubling and adaptation, even though the CNNs were not trained on such stimuli.

Overall, this work demonstrates the power of CNNs to not only accurately capture sensory circuit responses to natural scenes, but also uncover the circuit's internal structure and function. Moreover, our methods can be readily generalized to other sensory modalities and stimulus ensembles.

[1] LeCun et al. Deep learning. *Nature* 521, 436–444 (2015).

[2] Yamins, et al. Hierarchical Modular Optimization of Convolutional Networks Achieves Representations Similar to Macaque IT and Human Ventral Stream. *Advances in Neural Information Processing Systems* 26 (2013).