Deep learning models reveal diverse computations in the retina under natural scenes

Niru Maheswaranathan[1*], Lane McIntosh[*], David B. Kastner[2], Luke Brezovec, Aran Nayebi, Surya Ganguli and Stephen A. Baccus

Stanford University. [1]Currently at: Google Brain. [2]Currently at: UCSF

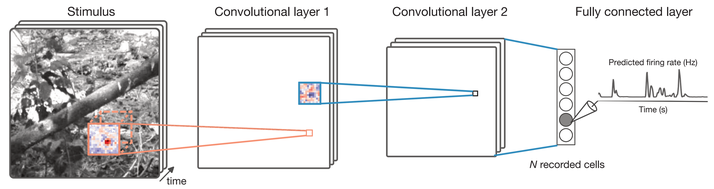
Introduction

Synthetic stimuli for probing sensory systems:

Advantages
- Control stimulus properties parametrically (e.g. orientation)
- Isolate particular mechanisms/effects

Disadvantages
- May push the circuit outside its natural operating regime
- Often requires clever insight and design

Goal: Understanding sensory responses to natural stimuli
- Are circuit computations discovered using structured stimuli relevant for natural vision?
- Do models trained on synthetic or natural stimuli exhibit different phenomenology?

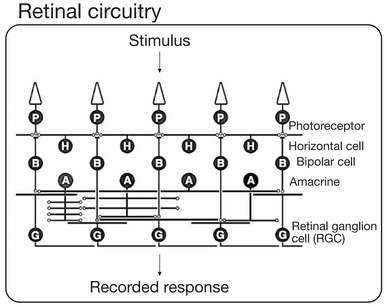


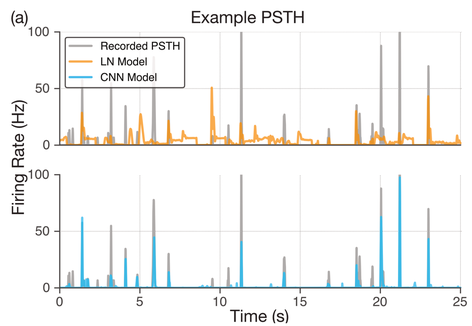
Our approach:
- Train a deep neural network to model retinal responses to natural stimuli
- See if the model exhibits known retinal phenomenology

CNNs accurately describe natural scene responses

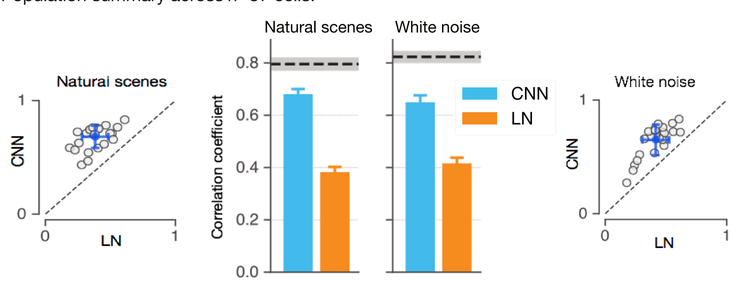
- Recorded salamander ganglion cell responses to white noise and natural scene stimuli.
- Fit CNN models to predict retinal spikes (blue), compared to LN models as a baseline (orange)

Retinal circuitry



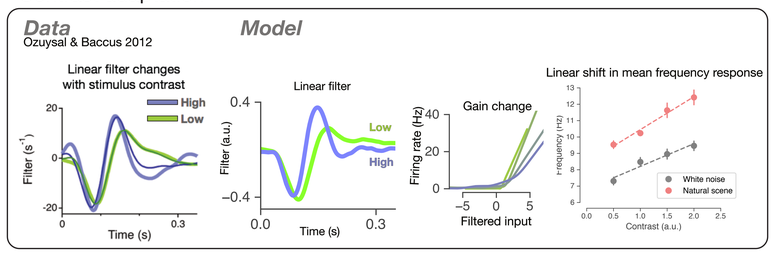


Population summary across *n*=37 cells:

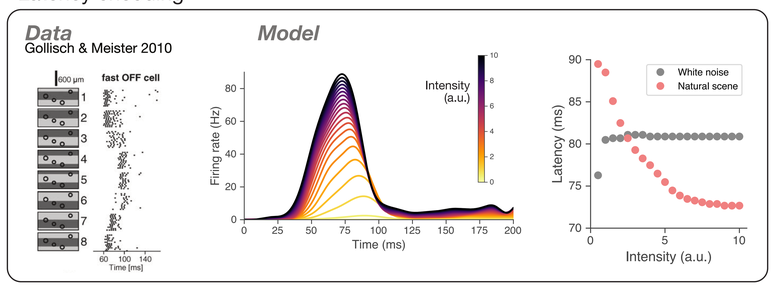


Generalization to diverse phenomena

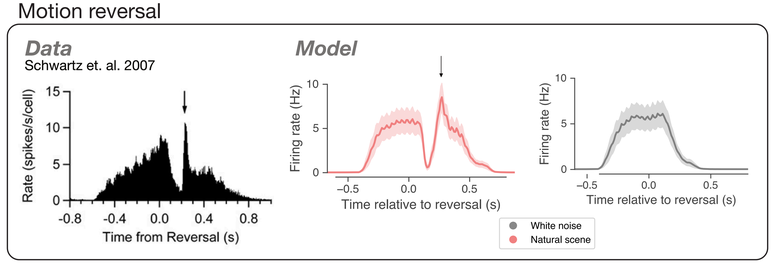
Contrast adaptation



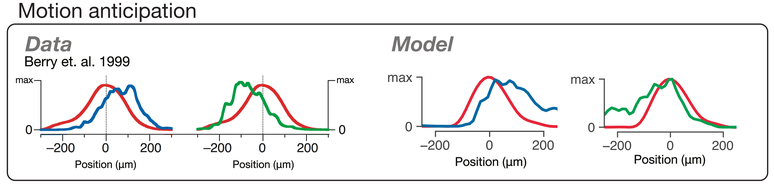
Latency encoding



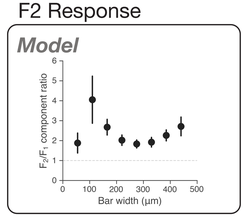
Motion reversal



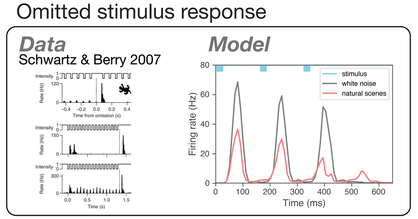
Motion anticipation



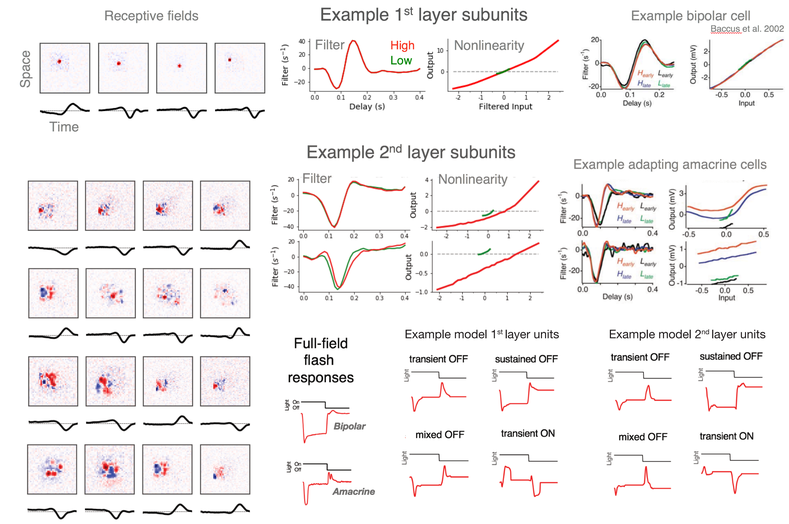
F2 Response

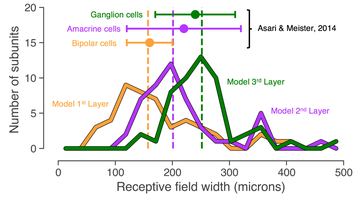


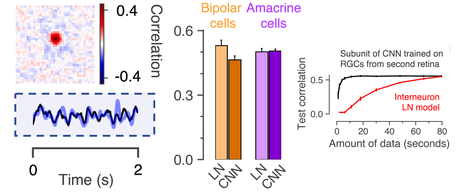
Omitted stimulus response



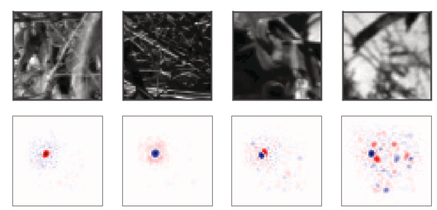
Comparing model units to retinal interneurons







Diversity in instantaneous receptive fields



Instantaneous receptive field (RF)
Gradient of model neuron wrt. stimulus
- Shows the most effective stimulus at each instant of time
- Visualizes stimulus-specific sensitivity

Conclusions

Implications for using deep networks to study neural systems:
- ✔ Remarkable match between retinal properties & internal model components
- ✔ Generalization beyond training distribution

Implications for retinal encoding of natural scenes:
- ✔ Diverse known nonlinear computations are engaged by natural stimuli
- ✔ Fitting models to natural stimuli is sufficient for capturing this phenomenology