The Effect of Pooling in a Deep Learning Model of Perceptual Learning

Modeling Sensitive Initial Conditions for Perceptual Learning

For correspondence: Rachel Lee (rachelstephlee@gmail.com)

Rachel Lee, Andrew Saxe
Stanford University

INTRODUCTION

- The neural mechanisms of visual perceptual learning remain unclear. Even focusing on orientation discrimination tasks alone, empirical studies have found widely different results, with some arguing that the site of plasticity is in early cortical layers such as V1 and V4 and others arguing it is in later decision layers[1-6].
- Here we consider two studies with nearly identical paradigms but widely divergent results: Schoups et al. (2001), and Ghose et al. (2002) both recorded changes in V1 after Macaque monkeys underwent extensive training in orientation discrimination[1, 3]. Both groups conducted follow-up studies in V4 using their respective experimental paradigms[2, 4] **(Fig 1).** The most significant difference is that one used fixed spatial frequency for training stimuli while the other used randomized spatial frequency for training stimuli (**Table 1**).
- Ghose et al. (2002) found substantial generalization to untrained positions, and no changes in V1; while Schoups et al. (2001) found the opposite. Explaining this sensitivity to task details is a key challenge for theory.

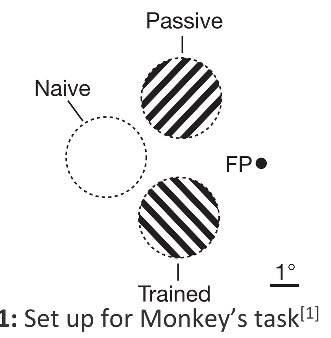


Fig 1: Set up for Monkey's task[1]

	Task parameters					Results		
	Location	Choices	Phase	Precision	Spatial frequency	V1	V4	Transfer
Schoups	Fixed	CW/CCW	Random	Higher	Fixed	Change	Change	No
Ghose	Fixed	Same/ Different	Random	Lower	Random	No Change	Change	Yes

Table 1: Summary of differences between the two experimental paradigms.

In summary, we aim to answer these questions:

1: How can we capture our visual cortex's marked sensitivity to the details of task paradigms?

2: Why do seemingly trivial task details lead to divergent results?

DEEP LEARNING MODEL OF PERCEPTUAL LEARNING

We develop a deep learning model of perceptual learning in orientation discrimination tasks. Our model has four main components **(Fig 2):**

a. Deep, chain-like structure
b. Pooling over phases
c. Initial orientation tuning
d. Gradient descent learning

We suggest that learning in a deep, layered structure can be highly nonlinear, amplifying the effects of small task differences.

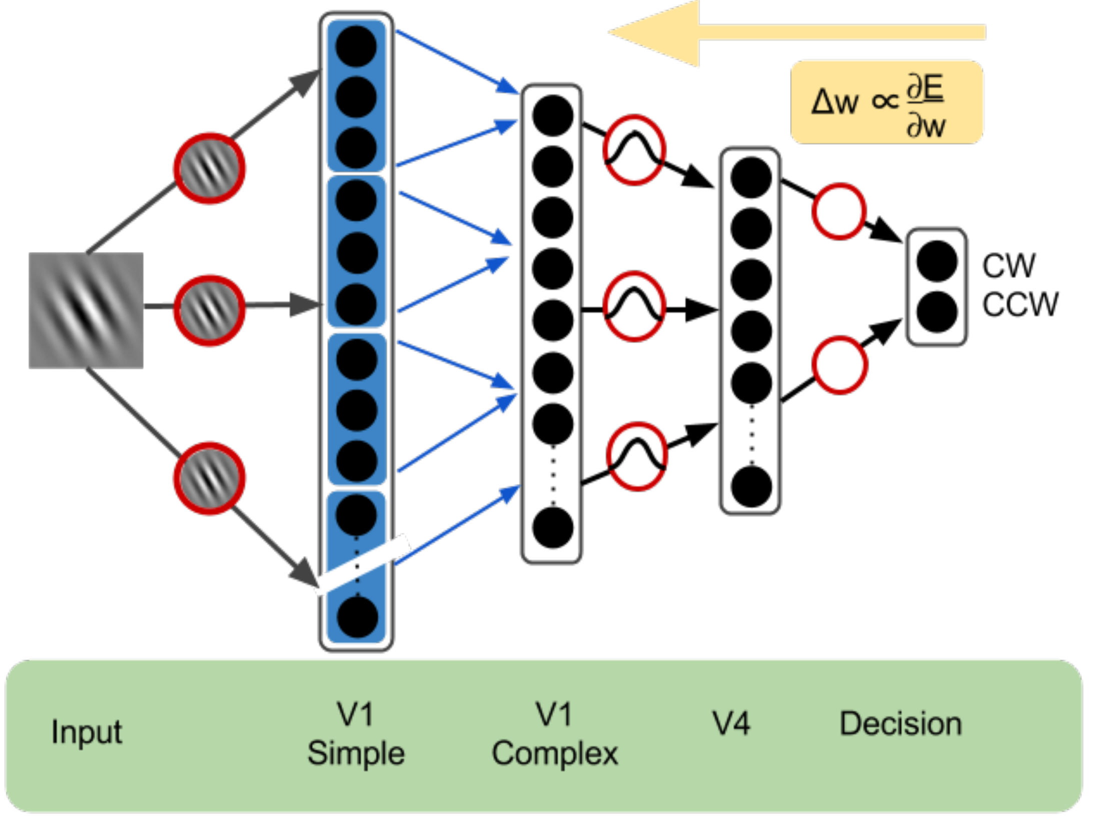


Fig 2: Deep Learning Model of Perceptual Learning with four main components color-coded. Activation function for phase pooling layer is sum of squares, for decision layer is softmax, and rectified firing otherwise. During learning, we allow changes in all layers but phase pooling layer. Decision has two choices: Clockwise (CW) or Counter-clockwise (CCW).

MODELING

We run our model with identical parameters, except for those that differ in the two different paradigms: Spatial frequency (fixed vs randomized) and task precision (lower for Ghose et al).

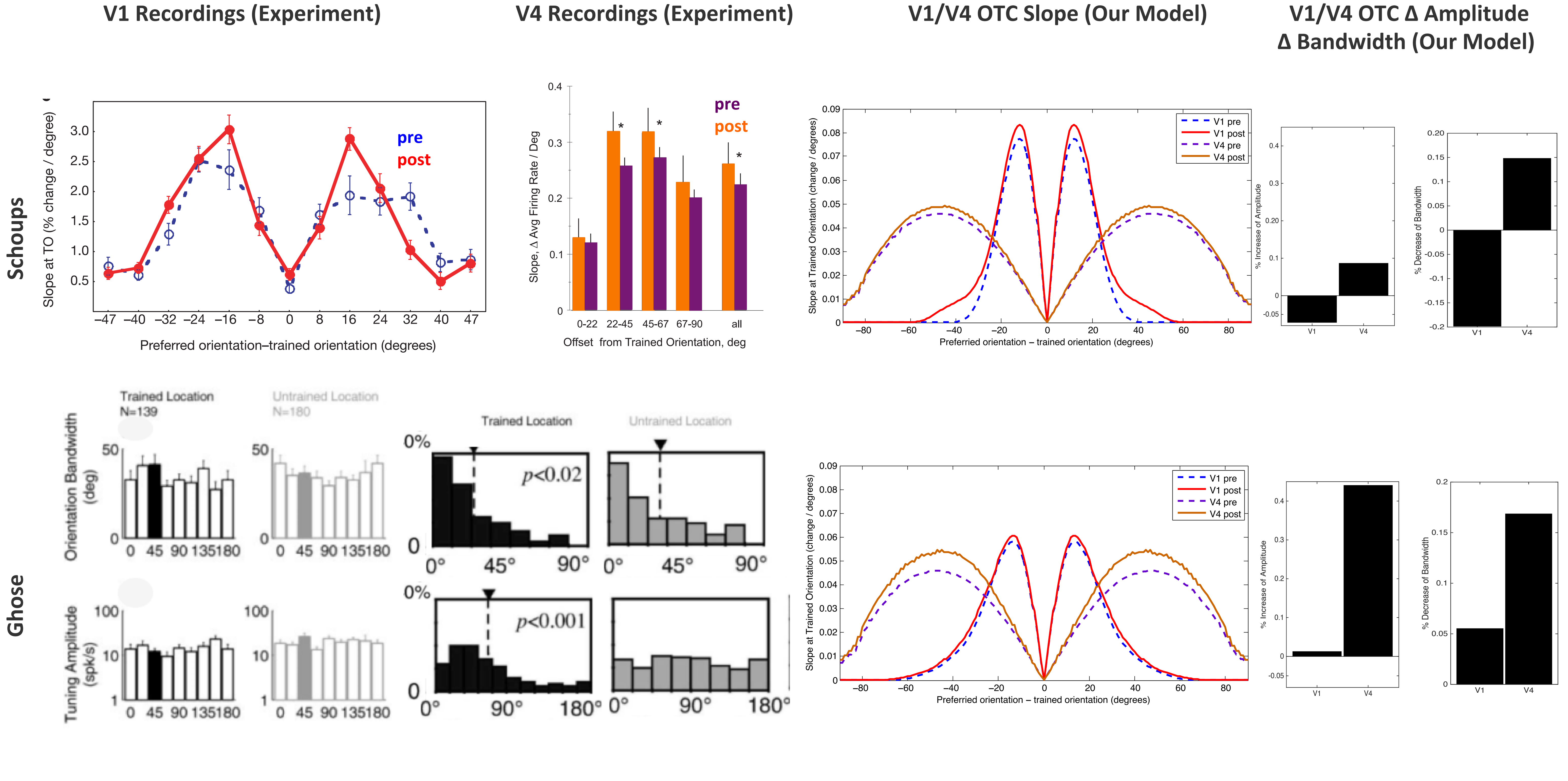


Fig 3: Modeling results. First two columns show empirical results from the four studies, latter two columns are results from our model.
Top Row (from left to right): Under Schoups paradigm: V1 OTC Slope[1], V4 OTC Slope[2], V1 and V4 OTC Slope, V1 and V4 OTC Δ Bandwidth and Δ Amplitude
Bottom Row (from left to right): Under Ghose paradigm: Δ Bandwidth and Δ Amplitude in V1 OTC[3], same for V4 OTC[4], V1 and V4 OTC Slope, V1 and V4 OTC Δ Bandwidth and Δ Amplitude

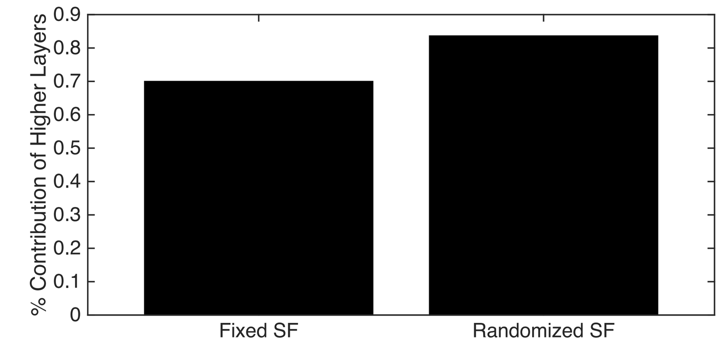
REFERENCES

[1] Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature, 412(6846), 549–53
[2] Raiguel, S., Vogels, R., Mysore, S. G., & Orban, G. a. (2006). Learning to see the difference specifically alters the most informative V4 neurons. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 26(24), 6589–602.
[3] Ghose, G. M., Yang, T., & Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. Journal of Neurophysiology, 87(4), 1867–88.
[4] Yang, T., & Maunsell, J. H. R. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 24(7), 1617–26.
[5] Dosher, B. a, & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. Proceedings of the National Academy of Sciences of the United States of America, 95(23), 13988–93.
[6] Petrov, A. a, Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: an incremental reweighting model. Psychological Review, 112(4), 715–43.
[7] Fahle, M. (1997). Specificity of learning curvature, orientation, and vernier discriminations. Vision Research, 37, 1885–1895.
[8] Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. Nature, 387(6631), 401–6. doi:10.1038/387401a0

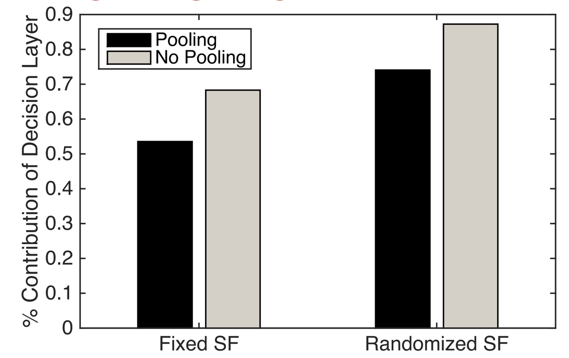
THEORY

We now consider *why* these differences in experimental paradigms lead to different results. In our model, we identify three factors:

1. **High versus low precision task:** Monkeys in the randomized spatial frequency task improve more slowly and hence see more low precision discriminations. Low precision discriminations change higher layers more.



2. **Larger changes downstream of pooling:** When spatial frequency is randomized, error can be decreased most quickly by changing connections to neurons that pool over SF. V1 does not pool over SF so weight changes target higher areas.



3. **Improved transfer:** Points 1 and 2 result in higher layers changing more. These higher layers contain large spatial receptive fields and hence transfer better.

DISCUSSION

Gradient descent in deep networks is highly nonlinear, such that even small task differences can cause widely different behavior.

Tuning changes that follow the gradient direction in a deep network model predict detailed changes in neural tuning in these tasks across multiple cortical areas.

Pooling alters learning dynamics to favor changes downstream of pooled representations.

Even in fine orientation discrimination tasks, "irrelevant" parameters such as spatial frequency can massively affect learning (in this instance, by forcing changes to higher levels that pool over SF.)

Our results highlight the intricate interplay between invariance, hierarchy, task, and learning-induced changes in neural tuning—factors that long have been the basis of experimental inferences about the locus of changes in the visual hierarchy[5,7,8].

Depth may be a key factor in other domains and learning phenomena.

R.L. was supported by Stanford Undergraduate Advising and Research, and COSYNE Undergraduate Travel Grant. A.S. was supported by a MBC Traineeship