Deep networks: Behavior, Brain and Theory

Dan Kersten & Paul Schrater *February 5, 2019*

University of Minnesota, Spring Semester, 2019

Topics in Computational Vision Psy 8036*
Psy 5993 Section 034*

Instructors:

Dan Kersten: kersten@umn.edu

Paul Schrater: schrater@umn.edu

Meeting time: First meeting Tuesday, Jan 22th, 3:00 pm.

Place: Elliott N227

*Students can sign up for either Topics in Computational Vision Psy 8036 (Kersten) or Psy 5993 Section 034 (Schrater).

See Canvas page for updated weekly content.

Abstract

Recent rapid advances in deep learning networks have provided the means to produce "image computable models" of human vision–models that take natural images as input and produce accurate predictions of perceptual decisions. However, the current and future value of deep network research for understanding the brain's visual system faces both methodological and conceptual challenges. What are the best methods to compare deep networks to perceptual behavior and to the brain? And how can we achieve a conceptual understanding of the networks, to determine which elements are important and which are not? We will read and discuss empirical papers

that compare network models of object recognition to behavior and the brain. The seminar will also review work that is helping to understand what functions networks can compute, and the limitations on learning to generalize. Finally, we will discuss advances that will be needed to understand the human ability to interpret virtually any image–an ability that spans a wide range of visual tasks. The class format will include short introductory lectures by the instructors to provide historical context and weekly student presentations of current literature. Students will have the opportunity to collaborate on final programming projects. The course will also introduce and use Julia, a rapidly developing language for scientific programming, which is fast, flexible, and relatively easy to learn and use.

Background

There are good online resources for learning about artificial neural networks, and in particular deep convolutional neural networks. For video content, there is the Neural Networks for Machine Learning from Geoff Hinton's 2016 coursera lectures and Fei-Fei Li's Stanford 231n course, Convolutional Neural Networks for Visual Recognition. For books, see Ian Goodfellow's Deep Learning free online or for purchase, and Visual Cortex and Deep Networks: Learning Invariant Representations for purchase by Tomaso Poggio and Fabio Anselmi.

For an excellent basic background and review, watch the first 9 lectures of the Stanford 231n course.

> ⚡ **Danger**
>
> !!!DRAFT!!!

Introduction

Week 1 Deep architectures

The first class will cover background and overview of the problems of human image understanding and visual recognition and how these problems have been approached. We'll go over the goals of the seminar in the context of three questions: 1) Is computational vision close to producing biologically consistent, predictive, models of human visual recognition performance? 2) assuming that candidate models exist; how well do we understand them for example, to decide when two networks are equivalent? 3) What is missing from current theories, in terms of conceptual understanding, behavioral functionality and their neural bases? To address the first question above, we'll review empirical papers, from psychophysics and neuroscience aimed at understanding the basic-level or "core" function of rapid object identification with the goal of determining best ways of comparing network models to behavior and the brain. To address the second question, we will review theoretical work that seeks to understand what functions networks can compute, and how efficiently they can learn the parameters (e.g. "weights") of those functions. Finally, we will assess where research may need to go to understand human ability to interpret virtually any image–a challenge that will require advances in dynamic neural architectures that allow task flexibility.

Background:

For a short video Introduction to deep networks, see Lecture 1 from the MIT short course (6.S191): Introduction to Deep Learning.

To run Julia programs and Jupyter notebooks locally on your computer, first install Julia, and then use the anaconda distribution to install Jupyter. For video instructions see: installing Julia and Jupyter.

But the quickest and easiest way to start learning and using the Julia language is to sign in to JuliaBox where you'll immediately be able to create notebooks and access tutorials.

Week 2 Class cancelled due to weather.

Week 3 Deep networks and the brain & Intro to Julia programming

Background:

Video introduction to Julia programming.

Introduction to Julia for Data Science and Scientific Computing

Readings:

Kietzmann, T. C., McClure, P., & Kriegeskorte, N. (2018, June 5). Deep Neural Networks In Computational Neuroscience. bioRxiv. doi:10.1101/133504

Week 4 Deep networks and the brain reviews & Simulations with the Julia language

Readings:

Turner, M. H., Sanchez Giraldo, L. G., Schwartz, O., & Rieke, F. (2019, January). Stimulus- and goal-oriented frameworks for understanding natural vision. Nature Neuroscience, 22(1), 15–24. doi:10.1038/s41593-018-0284-0

Jacobs, R. A. & Bates, C. J. (2018, November 27). Comparing the Visual Representations and Performance of Humans and Deep Neural Networks. Current Directions in Psychological Science, 0963721418801342. doi: 10.1177/0963721418801342

Kay, K. N. (2018, October 15). Principles for models of neural information processing. NeuroImage. New Advances in Encoding and Decoding of Brain Signals, 180, 101–109. doi:10.1016/j.neuroimage.2017.08.016

Empirical tests

The most basic perceptual test is "look and see" which is what is often what is done by modelers. One can treat "adversarial examples" as behavioral tests to guide model development. However, one can be systematic and ask what parametric manipulations can be made of test images that we expect humans to generalize "for free". How best to design "adversarial tasks" for a model network? For example, based on typical experience, we expect little cost to testing on certain families of novel image variations. These manipulations can be based on 3D variables (e.g. object transformations such as 3D rotation, cast shadows, occlusion), or image variables (e.g. lower contrast, blur, noise). Manipulations can be based on highly artificial variations that we already know humans have generalization abilities, if limited, such as reverse contrast, non-linear histograms, morphs, and atypical occlusions. Given these manipulations, what are good quantitative measures that can be applied to both human and model observers? When does more model training solve the problem, and can we understand when to rule out a class of network architectures?

Most current perceptual tests have been with feedforward DCNNs, and that is where we start.

Week 5 Empirical tests of DCNNs: human object recognition

Methodological challenges and the design of adversarial tasks.

Readings:

Ullman, S., Assif, L., Fetaya, E., & Harari, D. (2016, March 8). Atoms of recognition in human and computer vision. Proceedings of the National Academy of Sciences, 113(10), 2744–2749. doi:10.1073/pnas.1513198113. pmid: 26884200

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric, 10.

Retrieved from http://openaccess.thecvf.com/content_cvpr_2018/CameraReady/0299.pdf

Elsayed, G. F., Shankar, S., Cheung, B., Papernot, N., Kurakin, A., Goodfellow, I., & Sohl-Dickstein, J. (2018, February 22). Adversarial Examples that Fool both Computer Vision and Time-Limited Humans. arXiv: 1802.08195 [cs, q-bio, stat]. Retrieved November 13, 2018, from http://arxiv.org/abs/1802.08195

Geirhos, R., Temme, C. R. M., Rauber, J., Sch�tt, H. H., Bethge, M., & Wichmann, F. A. (2018). Generalisation in humans and deep neural networks. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), Advances in Neural Information Processing Systems 31 (pp. 7549–7561). Curran Associates, Inc. Retrieved January 28, 2019, from http://papers.nips.cc/paper/7982-generalisation-in-humans-and-deep-neural-networks.pdf

Week 6 Empirical tests of DCNNs and methods of study

Including discussion of generative models, manifold discovery, posterior estimation – useful tools?

Readings:

Hill, M. Q., Parde, C. J., Castillo, C. D., Colon, Y. I., Ranjan, R., Chen, J.-C., ..., & O'Toole, A. J. (2018, December 28). Deep Convolutional Neural Networks in the Face of Caricature: Identity and Image Revealed. arXiv: 1812.10902 [cs]. Retrieved January 23, 2019, from http://arxiv.org/abs/1812.10902

Zhang, M., Feng, J., Ma, K. T., Lim, J. H., Zhao, Q., & Kreiman, G. (2018, December). Finding any Waldo with zero-shot invariant and efficient visual search. Nature Communications, 9(1). doi:10.1038/s41467-018-06217-x

Ricci, M., Kim, J., & Serre, T. (2018, February 9). Same-different problems strain convolutional neural networks. arXiv: 1802.03390 [cs, q-bio]. Retrieved December 11, 2018, from http://arxiv.org/abs/1802.03390

Luo, W., Li, Y., Urtasun, R., & Zemel, R. (n.d.). Understanding the Effective Receptive Field in Deep Convolutional Neural Networks, 9

Week 7 DCNNs and the brain

Zhou, B., Bau, D., Oliva, A., & Torralba, A. (2018). Interpreting Deep Visual Representations via Network Dissection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1–1. doi:10.1109/TPAMI.2018.2858759

Readings:

Rajalingham, R., Issa, E. B., Bashivan, P., Kar, K., Schmidt, K., & DiCarlo, J. J. (2018, February 12). Large-scale, high-resolution comparison of the core visual object recognition behavior of humans, monkeys, and state-of-the-art deep artificial neural networks. doi:10.1101/240614

Schrimpf, M., Kubilius, J., Hong, H., Majaj, N. J., Rajalingham, R., Issa, E. B., …, & DiCarlo, J. J. (2018, September 5). Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? doi:10.1101/407007

Breedlove, J. L., St-Yves, G., Olman, C. A., & Naselaris, T. (2018, November 9). Human brain activity during mental imagery exhibits signatures of inference in a hierarchical generative model. doi:10.1101/462226

Understanding deep networks

Week 8 Theory: shallow vs. deep

Universal approximators. Hierarchy as a solution to over-fitting.

Poggio, T., Mhaskar, H., Rosasco, L., Miranda, B., & Liao, Q. (2017, October 1). Why and when can deep-but not shallow-networks avoid the curse of dimensionality: A review. International Journal of Automation and Computing, 14(5), 503–519. doi:10.1007/s11633-017-1054-2

Garriga-Alonso, A., Aitchison, L., & Rasmussen, C. E. (2019). DEEP CONVOLUTIONAL NETWORKS AS SHALLOW GAUSSIAN PROCESSES, 16

Lin, H. & Jegelka, S. (2018, June 28). ResNet with one-neuron hidden layers is a Universal Approximator. arXiv: 1806.10909 [cs, stat]. Retrieved January 18, 2019, from http://arxiv.org/abs/1806.10909

Week 9 Theory: DCNNs and implicit generative models

Feedforward DCNNs, implicit generative models, texture, and maximum entropy.

Readings:

Xie, J., Zhu, S.-C., & Wu, Y. N. (2017). Synthesizing dynamic patterns by spatial-temporal generative convnet. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7093–7101). Retrieved from http://openaccess.thecvf.com/content_cvpr_2017/papers/Xie_Synthesizing_Dynamic_Patterns_CVPR_2017_paper.pdf

Zhang, Q., Wu, Y. N., & Zhu, S.-C. (2017, October 2). Interpretable Convolutional Neural Networks. arXiv: 1710.00935 [cs]. Retrieved February 22, 2018, from http://arxiv.org/abs/1710.00935

WU, Y. N., XIE, J., LU, Y., & ZHU, S.-C. (2018). Sparse and Deep Generalizations of the FRAME Model. Retrieved from https://www.intlpress.com/site/pub/files/_fulltext/journals/amsa/2018/0003/0001/AMSA-2018-0003-0001-a007.pdf

Week 10 Theory: invariance

Readings:

Tacchetti, A., Isik, L., & Poggio, T. A. (2018, September 15). Invariant Recognition Shapes Neural Representations of Visual Input. Annual Review of Vision Science, 4(1), 403–422. doi:10.1146/annurev-vision-091517-034103

Leibo, J. Z., Liao, Q., Anselmi, F., & Poggio, T. (2015, October 23). The Invariance Hypothesis Implies Domain-Specific Regions in Visual Cortex.

PLOS Computational Biology, 11(10), e1004390. doi:10.1371/journal.pcbi. 1004390

Azulay, A. & Weiss, Y. (2018, May 30). Why do deep convolutional networks generalize so poorly to small image transformations? arXiv: 1805.12177 [cs]. Retrieved October 4, 2018, from http://arxiv.org/abs/1805.12177

Week 11 Theory: normalization

The value and types of normalization: batch, spatial, temporal, and channel normalization.

Liao, Q., Kawaguchi, K., & Poggio, T. (2016, October 19). Streaming Normalization: Towards Simpler and More Biologically-plausible Normalizations for Online and Recurrent Learning. arXiv: 1610.06160 [cs]. Retrieved December 7, 2018, from http://arxiv.org/abs/1610.06160

What's missing?

Week 12 Compositionality and semantic accessibility

The problems of occlusion and articulation (e.g. body pose). Computing spatial relationships,...

Readings:

Burgess, C. P., Matthey, L., Watters, N., Kabra, R., Higgins, I., Botvinick, M., & Lerchner, A. (2019, January 22). MONet: Unsupervised Scene Decomposition and Representation. arXiv: 1901.11390 [cs, stat]. Retrieved February 5, 2019, from http://arxiv.org/abs/1901.11390

Tang, H., Schrimpf, M., Lotter, W., Moerman, C., Paredes, A., Ortega Caro, J., ..., & Kreiman, G. (2018, August 28). Recurrent computations for visual pattern completion. Proceedings of the National Academy of Sciences, 115(35), 8835–8840. doi:10.1073/pnas.1719397115

Zhang, Z., Xie, C., Wang, J.