

Computational Neuroscience Seminar

Engineering a Large Scale Vision System by Leveraging Semantic Knowledge



Jonathon Shlens
Google Research

Monday, October 28, 2013

4:00-5:00pm

**Fung Auditorium, Powell-Focht Bioengineering Building
University of California, San Diego**

Abstract. Computer-based vision systems are increasingly indispensable in our modern world. Modern visual recognition systems have been limited though in their ability to identify large numbers of object categories. This limitation is due in part to the increasing difficulty of acquiring sufficient training data in the form of labeled images as the number of object categories grows unbounded. One remedy is to leverage data from other sources – such as text data – both to train visual models and constrain their predictions. In this talk I will present our recent efforts at Google to build a novel architecture that employs a deep neural network to identify visual objects employing both labeled image data as well as semantic information gleaned from unannotated text. I will demonstrate that this model matches state-of-the-art performance on academic benchmarks while making semantically more reasonable errors. Most importantly, I will discuss how semantic information can be exploited to make predictions about image labels not observed during training. Semantic knowledge substantially improves "zero-shot" predictions achieving state-of-the-art performance on predicting tens of thousands of object categories never previously seen by the visual model.