THESIS DEFENSE

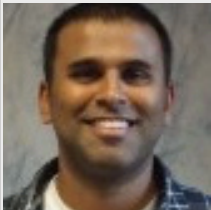## Learning to Learn for Small Sample Visual Recognition

Tuesday, May 8, 2018

3305 Newell Simon Hall

10:00 a.m.

Thesis Committee:
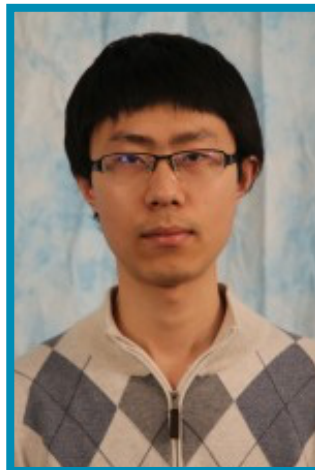
Martial Hebert
Chair

Deva Ramanan

Ruslan Salakhutdinov

Andrew Zisserman
University of Oxford

Yann LeCunn
Facebook AI Research
& New York University

Yuxiong Wang

Abstract

Understanding how humans and machines recognize novel visual concepts from few examples remains a fundamental challenge. Humans are remarkably able to grasp a new concept and make meaningful generalization from just few examples. By contrast, state-of-the-art machine learning techniques and visual recognition systems typically require thousands of training examples and often break down if the training sample set is too small.

This dissertation aims to endow visual recognition systems with low-shot learning ability, so that they learn consistently well on data of different sample sizes. Our key insight is that the visual world is well structured and highly predictable not only in data and feature spaces but also in task and model spaces. Such structures and regularities enable the systems to learn how to learn new recognition tasks rapidly by reusing previous experiences. This philosophy of learning to learn, or meta-learning, is one of the underlying tenets towards versatile agents that can continually learn a wide variety of tasks throughout their lifetimes. In this spirit, we address key technical challenges and explore complementary perspectives.

We begin by learning from extremely limited data (e.g., one-shot learning). We cast the problem as supervised knowledge distillation and explore structures within model pairs, i.e., models learned from few samples and models learned from large enough sample sets. To further decouple a recognition model from ties to a specific set of categories, we consider self-supervision using meta-data. We introduce an unsupervised meta-training phase and explore structures within a large collection of models. We then move on to learning from a medium sized number of examples and explore structures within a self-evolving model when learning from continuously changing data streams and tasks. Finally, we combine generative learning with meta-learning and explore joint structures in both data and task spaces.