



COLLEGE of ENGINEERING
AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

PhD Qualifying Examination

Mingjie Wang

Monday April 4, 2022 at 9am via Zoom

*Enhancing Learning Capabilities of Convolutional Neural Networks for
Fundamental Vision Problems*

Chair: Dr. Joe Sawada

Advisor: Dr. Minglun Gong

Co-Adviso: Dr. Simon Yang [SoE]

Non-Advisory: Dr. Fei Song

Non-Advisory: Dr. Graham Taylor [SoE]

Abstract:

During the past decade, Convolutional Neural Networks (CNNs) have been dominating the realm of computer vision and de facto standard for modern data-driven algorithms. They exhibited state-of-the-art performances in many vision tasks, thanks to their strong representative capacities. Hence, further diving into CNNs' learning capabilities has very broad impacts. This proposal reviews recent advances in promoting the boom of CNNs, and introduces several enhancements through novel mechanisms, structures, or paradigms, such as stochastic regularizations, layer-wise attention, multi-scale aggregation, tree structure, and auxiliary tasks. Since there are essentially two types of learning tasks, classification and regression, these two fundamental vision problems are chosen accordingly as large applications with demands of strengthening the models' capabilities.

Extensive experiments on widely-used benchmark datasets demonstrate the effectiveness of the proposed strategies in enhancing learning capabilities, thereby achieving superior performances in classification and counting accuracy. Ablation studies and visualization analysis are also performed to shed light on the impacts and behaviors of individual components. Although impressive progress in enhancing the learning capabilities of CNNs has been achieved, these proposed models mainly focus on exploiting novel mechanisms or aggregation strategies. In future work, I plan to make further efforts to intensify the capability of architectures from the perspectives of cutting-edge learning protocols, such as weakly-supervised learning, self-supervised proxy, transformer-based, or exclusively multi-layer perceptron (MLP)- based paradigms.