Tutorial 1: Architecture Design and Interpretability of Convolutional Neural Networks

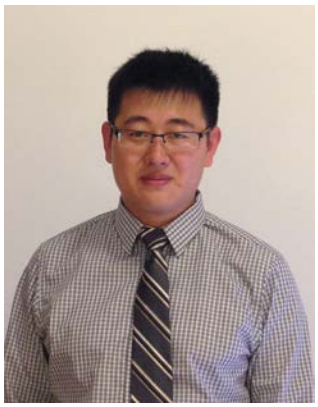
Deep learning techniques, in particular convolutional neural networks (CNNs), have been widely adopted for image processing in recent years. This tutorial will cover two important research topics on CNNs, namely, neural architecture design and interpretability, both in the context of image recognition or generation.

Firstly, architecture innovation of CNNs plays an important role in advancing the research on deep learning. From AlexNet and VGG, to ResNet and DenseNet, better architecture design has pushed the depth limit of deep models from 7 layers to hundreds of layers, leading to unprecedent performance on many real-world problems. This tutorial will review classical convolutional network architectures for image/video classification, segmentation and detection, discuss their underlying design principles, and analyze their strengths and weaknesses. We will also introduce the recent trend of developing highly efficient light-weighted deep models for mobile devices.

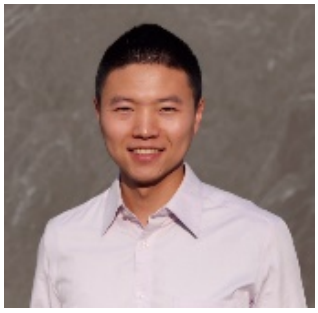
Secondly, interpreting deep models remains a very challenging task, which greatly limits its applicability in scenarios where people seek for a more transparent decision process. This tutorial will talk about recent progress on interpreting the deep visual representations in both the discriminative models and the generative models. The tutorial will show object detectors emerge in the network trained to classify scenes, while object synthesizers emerge in the network trained to draw scenes. Analytical tools are further developed to identify the emergent object concepts among the layers. The decision making of the networks is further visualized and explained.

Speakers:

Gao Huang
Tsinghua University, China



Jiashi Feng
National University of Singapore, Singapore



Bolei Zhou
The Chinese University of Hong Kong, China