Mathematical Analysis of Deep Convolutional Neural Networks

Ding-Xuan Zhou

City University of Hong Kong
Email: mazhou@cityu.edu.hk

Deep learning has been widely applied and brought breakthroughs in speech recognition, computer vision, and many other domains. The involved deep neural network architectures and computational issues have been well studied in machine learning. But there lacks a theoretical foundation for understanding the approximation or generalization ability of deep learning methods such as deep convolutional neural networks. This talk describes a mathematical theory of deep convolutional neural networks (CNNs). In particular, we discuss the universality of a deep CNN, meaning that it can be used to approximate any continuous function to an arbitrary accuracy when the depth of the neural network is large enough. Our quantitative estimate, given tightly in terms of the number of free parameters to be computed, verifies the efficiency of deep CNNs in dealing with large dimensional data. Some related distributed learning algorithms will also be discussed.