T7A & T7B: Scalable Deep Learning for Image Processing with Microsoft Cognitive Toolkit

Speakers: Cha Zhang (Microsoft Research, USA), Taifeng Wang (Microsoft Research, China)

Abstract

Deep learning has become the de facto standard method in most image processing problems. In the past few years, deep learning algorithms have met and exceeded human-level performance in image recognition. Nevertheless, training deep learning networks on a large data set remains very challenging. The sheer amount of computation needed to train a convolutional neural network can take months on large data sets. Combining that with the black art of hyper-parameter tuning, the community desperately needs tools to help train deep learning networks on multiple servers with multiple GPUs. This tutorial will introduce Microsoft's Cognitive Toolkit, also known as CNTK (https://github.com/Microsoft/CNTK), to the image processing community. Various algorithms supported by the toolkit will be presented. The benefits of CNTK in terms of speed and scalability relative to existing toolkits will also be described.

Speaker Bio:

Cha Zhang is a Principal Researcher in the Advanced Technology Group at Microsoft Research. He received the B.S. and M.S. degrees from Tsinghua University, Beijing, China in 1998 and 2000, respectively, both in Electronic Engineering, and the Ph.D. degree in Electrical and Computer Engineering from Carnegie Mellon University, in 2004. His current research focuses on applying various audio/image/video processing and machine learning techniques to multimedia applications. Dr. Zhang has published 3 books, more than 80 technical papers and 30+ U.S. patents. He won the best paper award at ICME 2007, the top 10% award at MMSP 2009, the best student paper award at ICME 2010, and the top 10% award at ICIP 2014. He was the Program Co-Chair for VCIP 2012, and the General Co-Chair for ICME 2016. He currently is an Associate Editor for IEEE Trans. on Circuits and Systems for Video Technology, and IEEE Trans. on Multimedia.

Taifeng Wang is a Lead Researcher in the Artificial Intelligence group, Microsoft Research Asia. He received his B.S and M.S. degrees in Electronic Engineering from University of Science and Technology of China in 2003 and 2006, respectively. His research interests include large-scale machine learning, distributed system, internet advertising, and search engine technique. His research focuses on modeling users' behavior in ads system to help the search engine to deliver better ads. Before working on ads, he had developed a large-scale graph learning platform which handles learning and mining tasks on graphs with billions of nodes. His latest research interest is to enable deep learning algorithms to scale well on distributed systems, and he actively contributes to the distributed learning algorithms in Microsoft's Cognitive Toolkit.