BANQUET KEYNOTE TALK

Tuesday, October 4, 2022 at 7:00pm

The Future of Visual Computing via Foundation Models

Ce Liu
Azure Cognitive Services
Microsoft
USA

Abstract: Thanks to big data, computing power and modern network architecture, we are seeing a wave of continuous breakthroughs find their way into people's everyday lives. While modern AI has reached human parity on a few well-defined research benchmarks, a rapidly growing number of disjointed AI tasks are needed to mimic human intelligence in understanding the open and complex world. As each AI task is often defined by the statistics manifested from large amounts of task-specific data, we end up building expensive silos without a synergistic way of knowledge sharing and transferring among the different AI tasks.

In this keynote I will share the future of visual computing via large-scale image-language foundation models, such as CLIP and Florence (image to text) and Dall-E (text to image), as a new AI paradigm to integrate fragmented tasks. Empowered by a semantic layer learned from the latest transformers, these foundation models have demonstrated not only unprecedented capabilities in zero-shot and few-shot transfer learning for new tasks in the wild, but also fascinating potentials to unify common visual computing tasks such as recognition, detection, segmentation, captioning and image editing. I will also discuss how the research communities can develop disruptive and creative AI systems using foundation models of various modalities.

Speaker Bio-Sketch: Ce Liu received his B.E. and M.E. from Tsinghua University in 1999 and 2002, respectively. He received a PhD from MIT Department of Electrical Engineering and Computer Science in 2009. He worked at Microsoft Research Asia from 2002 to 2003, Microsoft Research New England from 2009 to 2014, and Google Research from 2014 to 2021. He was appointed as an adjunct assistant professor at Boston University to teach machine learning in 2013. He has published more than 70 papers in peer-reviewed conferences and journals. He received the best student paper award at NIPS 2006 and CVPR 2009, and the best paper award honorable mention at CVPR 2019. He is a recipient of TPAMI Young Research Award in 2016. His team won ECCV'2020 robust vision challenge on optical flow. He has been serving as area chairs for CVPR/ICCV/ECCV/NeurIPS/ICML/ICLR, an associate editor for TPAMI, and served as a Program co-Chair for CVPR 2020.