THE CHINESE UNIVERSITY OF HONG KONG
Department of Electronic Engineering
Seminar

Towards General Multimodal Intelligence: Large Multimodal Models from Algorithm to Impact

Professor Hongsheng Li

Associate Professor
Department of Electronic Engineering

Date: September 15, 2025 (Monday)
Time: 9:30 a.m. – 11:30 a.m.
Venue: Room 222, 2/F, Ho Sin Hang Engineering Building, CUHK

Abstract

Advancing toward general multimodal intelligence requires models that can efficiently fuse heterogeneous inputs, build structured world models from unstructured data, and perform controllable generation and verifiable reasoning. In this talk, I will outline an end-to-end research agenda that addresses each pillar. Our LLaMA-Adapter, CLIP-Adapter, and SPHINX-X families offer compute-efficient pathways to high-performing LMMs, surpassing strong baselines on MME, SEED-Bench, and Video-Bench while enabling rapid adaptation in practical settings. In perception, our point-cloud detectors (PointRCNN, PV-RCNN, PartA2) introduced part-aware and point-voxel abstractions that now underpin industrial autonomy stacks at million-device scale. For generation, Lumina-T2X provides a unified, open framework capable of flexible-resolution synthesis across images, videos, 3D point clouds, and audio, with broad community adoption. For reasoning, our code-based self-verification approach significantly improves mathematical problem solving. I will discuss lessons learned on data curation, architecture scaling, and evaluation, and outline future directions in multimodal agents, safety, and energy-efficient training.

Biography

Hongsheng Li is currently an Associate Professor in the Department of Electronic Engineering at The Chinese University of Hong Kong, adjunct PhD advisor at Shanghai Jiao Tong University and University of Science and Technology of China. He was a Huashan Chair Professor at Xidian University and a Consultant Scientist at Shanghai AI Laboratory. He received the bachelor's degree from East China University of Science and Technology in 2006, and the doctorate degree in Computer Science from Lehigh University in 2012. He has over 64,000 citations on Google Scholar. He has published over 230 papers in premier conferences on AI, computer vision, and multimodal learning, including CVPR, ICCV, ECCV, NeurIPS, ICLR, and AAAI. He received the 2025 CUHK Research Excellence Award, 2020 IEEE Circuits and System Society Outstanding Young Author Award, 2022-2024 AI 2000 Most Influential Scholar Honorable Mention in Computer Vision, 2022-2024 Top 2% Scientists Worldwide, 2021 CUHK Young Researcher Award. He is/was an associate editor of Neurocomputing, IEEE Transactions on Circuits and Systems for Video Technology, and Transactions on Machine Learning Research. He serves as a Senior Area Chair of CVPR 2026 and ICCV 2025, and an Area Chair of NeurIPS 2021-2023, 2025, ICCV 2023-2025, ICML 2023-2025, CVPR 2023, ACM MM 2024-2025.

***** ALL ARE WELCOME *****

For enquiries, please contact Ms. Natalie Tsang, Department of Electronic Engineering, on 3943-8277.