CMAI: Interface of Math and Artificial Intelligence Seminar
The Chinese University of Hong Kong

This CMAI Interface of Math and AI Seminar is organized by Centre for Mathematical Artificial Intelligence (CMAI), under Department of Mathematics at CUHK.

Date: May 30, 2024 (Thursday)
Time: 10:00 am-11:30 am (Hong Kong Time)
Zoom Meeting: 905 330 9693

Multimodal Machine Learning and Its Applications

Speaker: Prof. Xingping Dong
Wuhan University

Abstract: With the vigorous development of deep learning, many artificial intelligence algorithms have achieved tremendous success in machine learning tasks, such as speech recognition, machine translation, and facial recognition. However, most of these tasks are based on single-modal data, while real-world data is often multimodal, as seen in movie data that includes video, audio, and text subtitles. The development of multimodal learning not only enables the effective utilization of the growing multimodal data but also provides a more comprehensive understanding of real-world data, enhancing the efficiency and capabilities of models. It can also handle more complex tasks, such as the recently highlighted text-to-image generation task.
This presentation will briefly introduce the basic concepts and development of multimodal learning, along with its applications in the visual and language domains. Through the analysis of specific tasks, including visual language navigation, language-specified target segmentation and tracking, we will explore the challenges and solutions that multimodal learning faces in practical applications, providing initial insights for the practical implementation of multimodal learning algorithms.

Bio: Xingping Dong is a professor in the School of Computer Science, Wuhan University. He was a research scientist in Inception Institute of Artificial Intelligence (IIAI), 2019-2023. In 2019, he received his Ph.D. degree in Beijing Lab of Intelligent Information Technology at School of Computer Science, Beijing Institute of Technology (BIT). He worked as a joint Ph.D. student in College of Engineering \& Computer Science, Australian National University (ANU). He has published 20+ conference and journal papers such as IEEE CVPR, ECCV, IEEE TPAMI, IEEE TIP, IEEE TMM, IEEE TNNLS, IEEE TCYB and IEEE TCSVT. His Google scholar citations are about 3186 times with H-index 23 . Four papers are selected as the ESI Highly Cited Paper including one ESI Hot Paper. He has also obtained many flagship scholarships such as National Natural Science Fund for Excellent Young scientists Fund Program (Overseas), and Excellent CAAI PhD thesis award. He is the regular reviewer for many famous conferences, such as IEEE CVPR, IEEE ICCV, ECCV, NeurIPS, ICML, ICLR, AAAI, and ACCV, and journals, such as IEEE TPAMI, IEEE TIP, IEEE TMM, IEEE TNNLS, TITS, PR, CVIU, NEUCOM, TVCJ, and SIVP. His current research interests include (1) Visual object tracking, (2) Deep reinforcement learning, (3) Image segmentation, (4) Few shot learning, (5) Neural rendering.