CS Seminar

DETR and Beyond: Advancing Open-Set Object Detection with Language and Visual Prompts

Prof. Lei Zhang
Hong Kong University of Science and Technology (Guangzhou)

Date:
Dec 01, 2023
Friday
14:00-16:00

Venue:
Room 308
Chow Yei Ching Building
The University of Hong Kong

Abstract:

This talk presents the evolution of object detection from the traditional closed-set frameworks to the open-set domain, leveraging the DETR algorithm as our foundational technology. We first review the DETR model, setting the stage for our DETR-series research: DAB-DETR, DN-DETR, and DINO, which establish DETR-like models as a mainstream detection framework. We then discuss our steps towards open-set object detection, introducing Grounding DINO, which utilizes language prompts to extend detection boundaries, and T-Rex, which employs visual prompts for object detection in unseen domains. These works mark significant steps toward object detection systems that are not limited by predefined labels but are instead equipped to grow and adapt through interaction with prompts. Our talk will explore these advancements, demonstrating their potential to redefine the landscape of object detection.

About the Speaker:

Lei Zhang is the Chair Scientist of Computer Vision and Robotics at International Digital Economy Academy (IDEA) and an Adjunct Professor of Hong Kong University of Science and Technology (Guangzhou). Prior to this, he was a Principal Researcher and Research Manager at Microsoft, where he has worked since 2001 in MSRA (Beijing), MSR (Redmond), and other computer vision-related product teams. His research interests are mainly in computer vision and machine learning, with a particular focus on generic visual recognition at large scale. Since joining IDEA Research in 2021, he has conducted a series of research works in the direction of object detection, among which the DINO algorithm for the first time established DETR-like algorithms as the SOTA of object detection. He has published more than 150 papers in top conferences and journals and holds more than 60 US-granted patents. He was named as IEEE Fellow for his contribution in large-scale visual recognition and multimedia information retrieval.

All are welcome!
For enquiries, please call 2859 2180 or email
enquiry@cs.hku.hk
Department of Computer Science
The University of Hong Kong