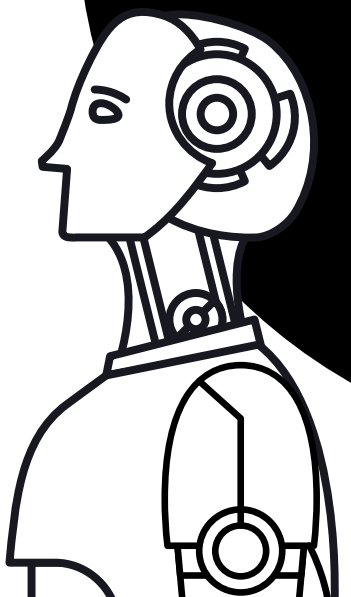
國立陽明交通大學電信工程研究所
系統組論文研討

Visual Prompt-Guided Vision-Language Models for Cardiac CT Diagnosis



主講人：國立臺北科技大學 電子工程系 曾柏軒 教授

日期：114 年 10 月 29 日（星期三）
時間：13:20~15:20
地點：工程四館219室

Abstract

Cardiovascular disease is a significant cause of mortality, and computed tomography (CT) plays a central role in the non-invasive assessment of coronary health. Recent advances in general-purpose medical vision–language models (VLMs) have demonstrated potential across diverse imaging tasks. However, their direct application to cardiac CT remains limited. Existing models face challenges in recognizing subtle plaques, consistently evaluating stenosis severity, and integrating complex three-dimensional cardiac anatomy, thereby restricting their clinical utility in tasks such as report generation and diagnostic question answering. To address these limitations, we integrate segmentation-aware prompting with SigLIP to enhance fine-grained anatomical feature alignment. Based on this visual prompt-augmented idea, we propose a multimodal VLM framework for cardiac CT analysis to improve medical report generation and visual question answering. An automated data synthesis pipeline expands high-quality cardiac CT–report pairs, alleviating the scarcity of large-scale annotated datasets. Evaluated on more than 3,000 cardiac CT stacks, the proposed framework achieves consistent improvements over existing VLM baselines, establishing a scalable foundation for clinically meaningful cardiac CT interpretation.

Biography

Po-Hsuan Tseng received the B.S. and Ph.D. degrees in communication engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 2005 and 2011, respectively. From Jan. 2010 to Oct. 2010, he was a Visiting Researcher with the University of California, Davis, Davis, CA, USA. He was an Associate Professor from 2012 to 2017 and an Assistant Professor from 2017 to 2022 with the Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan, where he has been a Full Professor since 2022. His current research interests include signal processing for networking and communications, including wireless localization and sensing, networking for the Internet of Things, software-defined networks, and medical image analysis.