

陈芝漩 | Zhixuan Chen

多模态大模型 / 视觉语言模型 / 大模型后训练

zchenhi@connect.ust.hk | zhi-xuan-chen.github.io | github.com/zhi-xuan-chen | Google Scholar: 7AMNEKAAAAAJ

HKUST 计算机博士在读 | 多模态基础模型、文本提示视觉理解与区域级视觉语言生成 | 第一作者成果发表于 Nature Computational Science、IEEE TMI、MICCAI

教育经历

香港科技大学 (HKUST) | 计算机科学与工程 博士 2023.09 - 至今
导师：陈浩教授 | 研究方向：多模态大模型、视觉语言模型、计算机视觉

电子科技大学 (UESTC) | 电子科学与技术 工学学士 2019.09 - 2023.06
荣誉毕业 | 本科优秀毕业论文 | 四川省优秀大学毕业生

专业能力

多模态建模 视觉语言联合训练、跨模态对齐、CLIP / 对比学习
视觉理解与生成 文本提示分割、区域级 Grounding、三维视觉编码、长文本生成
模型训练与适配 参数高效微调、强化学习微调 (RL/RFT)、Prompt Learning

代表项目

PathSegmentor：文本提示驱动的可提示分割基础模型 Accepted
第一作者 | Nature Computational Science | JIF 20.3 (2025) | arXiv:2506.20988

- 提出首个面向病理图像的文本提示分割基础模型，以自然语言替代点、框等空间提示，实现语言指令到像素级掩码的端到端预测。
- 构建 PathSeg 数据集，整合 21 个公开来源、27.5 万组图像-掩码-标签样本和 160 个类别，支撑大规模视觉语言对齐训练。
- 较空间提示和文本提示基线，整体 Dice 分别提升 0.145 和 0.429，并验证外部数据泛化能力。

Reg2RG：区域级指代、定位与视觉语言生成 IEEE TMI 2025
第一作者 | github.com/zhi-xuan-chen/Reg2RG

- 提出区域引导的视觉语言生成框架，将三维视觉表征、区域定位与大语言模型生成统一建模。
- 增强局部目标的细粒度感知和跨模态对齐能力，实现从视觉区域到自然语言描述的可解释生成。

Dia-LLaMA：大语言模型驱动的三维视觉内容生成 MICCAI 2025
第一作者 | Early Accept | github.com/zhi-xuan-chen/Dia-LLaMA

- 将视觉编码、先验知识与语言推理结合，研究 LLM 驱动的长文本生成和三维视觉 token 高效编码。
- 围绕视觉语言联合训练完成模型设计、实验验证与开源实现。

论文成果

- [1] Zhixuan Chen, Junlin Hou, Liqi Lin, et al. Segment Anything in Pathology Images with Natural Language. Nature Computational Science, Accepted. 第一作者, JIF 20.3 (2025).
- [2] Zhixuan Chen, Yequan Bie, Haibo Jin, Hao Chen. Large Language Model with Region-guided Referring and Grounding for CT Report Generation. IEEE TMI, 2025. 第一作者。
- [3] Zhixuan Chen, Luyang Luo, Yequan Bie, Hao Chen. Dia-LLaMA: Towards Large Language Model-driven CT Report Generation. MICCAI, 2025. Early Accept, 第一作者。
- [4] Zhixuan Chen, Cheng Jin, Tianjing Zhang, Xiao Wu, Liangjian Deng. SpanConv: A New Convolution via Spanning Kernel Space for Lightweight Pansharpening. IJCAI, 2022. 第一作者, Long Oral (3.75%)。
- [5] Yequan Bie, Peng Xie, Zhixuan Chen, et al., Hao Chen. Visual Reinforcement Fine-Tuning via Bootstrapped Medical Reasoning. NeurIPS 2026 Submission.
- [6] Sunan He, Yuxiang Nie, et al., Zhixuan Chen, et al., Hao Chen. Towards Generalizable AI in Medicine via Generalist-Specialist Collaboration. Nature Biomedical Engineering, 2026.
- [7] Zhe Xu, Ziyi Liu, Junlin Hou, et al., Zhixuan Chen, et al., Hao Chen. A Versatile Pathology Co-pilot via Reasoning Enhanced Multimodal Large Language Model. arXiv, 2025.
- [8] Yuxiang Nie, Sunan He, Yequan Bie, Yihui Wang, Zhixuan Chen, Shu Yang, Hao Chen. ConceptCLIP: Towards Trustworthy Medical AI via Concept-Enhanced Contrastive Language-Image Pre-training. arXiv, 2025.
- [9] Yequan Bie, Andong Tan, Zhixuan Chen, Zhiyuan Cai, Luyang Luo, Hao Chen. Knowledge-Enhanced Explainable Prompting for Vision-Language Models. AAAI, 2026.
- [10] Yequan Bie, Luyang Luo, Zhixuan Chen, Hao Chen. XCoOp: Explainable Prompt Learning for Computer-Aided Diagnosis via Concept-guided Context Optimization. MICCAI, 2024.

代表性荣誉

国家奖学金 (连续三届, Top 2%)	2020 / 2021 / 2022
《人民日报》国家奖学金代表名录 (全国 100 人)	2023
黄昆奖学金 (全校 10 人)	2023
感恩中国近现代科学家奖学金 (Top 0.2%, 全校 12 人)	2023
四川省优秀大学毕业生 (Top 1%)	2023
美国大学生数学建模竞赛 Meritorious Winner (Top 7%)	2022
全国大学生光电设计竞赛国家一等奖 (Top 1.5%)	2021
四川省大学生数学建模竞赛一等奖	2021
全国大学生数学竞赛一等奖 (Top 8%)	2020
本科优秀毕业论文、校级优秀学生等荣誉	