

张通

✉ tongz27@uci.edu

☎ +86 18871162559

🌐 [the-star-sea](#)

🔗 [个人主页](#)

教育背景

复旦大学	上海
计算机科学与技术, 博士	2025年9月- 至今
加州大学尔湾分校	尔湾
计算机工程, 硕士	2023年6月- 2024年9月
南方科技大学	深圳
计算机科学与工程, 学士	2019年9月- 2023年6月

研究方向

- **Coding Agent**: 面向复杂软件环境的智能体构建、工具使用与经验积累
- 多模态推理 (**Multimodal Reasoning**): 结合视觉、语言与结构化表示的多步推理与可控生成

科研经历

基于视觉语言模型的SVG生成 PDF	伊利诺伊大学香槟分校
<i>Assistant Prof. Haohan Wang</i>	2023年9月- 2024年5月
PyTorch	
<ul style="list-style-type: none">• 提出S²VG², 首个结合视觉语言模型的SVG生成框架• 引入专用数据集SVG-SHAPE, 用于评估SVG生成方法与LLM的推理能力• 在LLM的SVG推理与视觉指标上取得了SOTA表现	
单样本可控头部三维重建 PDF	加州大学尔湾分校
<i>Prof. Xiaohui Xie</i>	2023年4月- 2023年6月
PyTorch	
<ul style="list-style-type: none">• 提出CVTHead, 一种单样本可控的头部三维重建框架• 在跨身份重建任务中与其他方法进行对比, 表现出更好的重建效果与稳定性• 在VoxCeleb1和VoxCeleb2数据集上取得了SOTA表现	
轨迹预测与驾驶视频描述生成 PDF	清华大学AIR实验室
<i>Assistant Prof. Hao Zhao</i>	2022年5月- 2022年9月
NumPy, PyTorch	
<ul style="list-style-type: none">• 改进了端到端Transformer模型, 实现驾驶视频的描述与解释生成• 在驾驶视频描述任务上取得了SOTA表现	

工作经历

研究实习	腾讯
<i>Dr. Dezhi Ran</i>	2025年9月- 2025年12月
PyTorch, Python	
<ul style="list-style-type: none">• 开发基于多模态大模型 (VLM) 的微信小程序智能体, 设计并实现Agent Memory和经验提取机制	
研究实习	vivo
<i>Dr. Bo Li</i>	2025年5月- 至今
PyTorch, Python	
<ul style="list-style-type: none">• 提出基于能量函数的注意力热图平滑算法, 提升图像编辑任务中的语义保真性与空间连贯性• 探索基于VLM与链式推理的框架, 构建图像模态上的“生成-反思-再推理”闭环流程, 实现多轮可控图像优化	
研究实习 PDF	香港中文大学 (深圳)

Prof. Haizhou Li

2024年11月 - 2025年4月

PyTorch

- 分析了11种防御机制在6个大语言模型上的表现，评估其对模型性能、过度拒答率与计算开销的影响
- 提出了9种元防御策略，系统化地分析防御机制的副作用，揭示安全性与性能间的权衡
- 实现并评估了音频语言模型的Token压缩策略

轻量化OCR模型开发 [Report](#)

OpenCV

Associate Prof. Shiqi Yu

2022年5月 - 2022年9月

PyTorch, ONNX, C++

- 在OpenCV Zoo中实现了PP-OCRv2模型的检测部分，并基于ONNX完成模型转换
- 实现了用于支持PP-OCRv2推理的多种C++算子
- 实现了文本检测的评估指标 (AP、Recall、Precision、Hmean)

Java课程教学助理

南方科技大学

Associate Prof. Yu Zhang

2022年3月 - 2022年6月

English, Java

- 设计并批改了大量课程作业与项目
- 管理了在线评测平台，用于学生代码提交的自动化评测

学术论文

Sinin Zhang, Yunfei Xie, Yuxuan Cheng, Haoyu Zhang and **Tong Zhang**, "PhysNote: Self-Knowledge Notes for Evolvable Physical Reasoning in Vision-Language Model", *ES-Reasoning @ ICLR*, 2026

Tong Zhang, Bangjie Yin Fengxiang Yang and Bo Li, "Energy-Driven Attention Mask Refinement for Semantic Image Editing", *Under Review*, 2025

Tong Zhang, Yiming Chen, Simin Chen, Zexin Li, Xianghu Yue, Cong Liu, Chenyu You and Haizhou Li, "Unintended Side Effects of Defense Mechanisms in Large Language Models: A Comprehensive Study", *Under Review*, 2025

Tong Zhang, Haoyang Liu, Peiyan Zhang, Yuxuan Cheng and Haohan Wang, "Beyond Pixels: Exploring Human-Readable SVG Generation for Simple Images with Vision Language Models", *Preprint*, 2024

Haoyu Ma, **Tong Zhang**, Shanlin Sun, Xiangyi Yan, Kun Han and Xiaohui Xie, "CVTHead: One-shot Controllable Head Avatar with Vertex-feature Transformer", *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024

Bu Jin, Xinyu Liu, Yupeng Zheng, Pengfei Li, Hao Zhao, **Tong Zhang**, Yuhang Zheng, Guyue Zhou and Jingjing Liu, "ADAPT: Action-aware Driving Caption Transformer", *IEEE International Conference on Robotics and Automation (ICRA)*, 2023