



李童

1997.06 中共党员 回族

浙江工业大学 计算机科学与技术学院

智能感知与系统教育部工程研究中心

导师：孙国道（国家优青）、梁荣华（国家万人）

研究方向：大数据可视化与可视分析、机器学习、跨模态理解

个人主页: <https://tongli97.github.io/>

邮箱: litong@zjut.edu.cn

联系方式: 19858140096

个人履历

2019.09 – 至今	浙江工业大学 计算机科学与技术学院 计算机科学与技术专业	硕博连读
◇ 博士课题：面向层次化模态理解的可视分析研究		
◇ 研究生期间在可视化领域发表论文 10 余篇 （一作 5 篇），含 5 篇 SCI 1/2 区 期刊论文和 4 篇 领域内高认可度会议，发表期刊总影响因子为 55.6		

代表性学术成果

- [1] **Li Tong**, Sun G, Zheng X, et al. CompoVis: Is Cross-modal Semantic Alignment of CLIP Optimal? A Visual Analysis Attempt [J], IEEE Transactions on Multimedea, 2025 (**SCI 1 区**, CCF-B, 一作);
- [2] **Li Tong**, Sun G, Chang B, et al. LANDER: Visual Analysis of Activity and Uncertainty in Surveillance Video [J], IEEE Transactions on Human-Machine Systems, 2024 (**SCI 2 区**, CCF-B, 一作);
- [3] **Li Tong**, Sun G, Liang R. SurVizor: Visualizing and Understanding the Key Content of Surveillance Videos [C], The China Visualization and Visual Analytics Conference, 2021 (国内本领域最高级别学术会议, 一作);
- [4] **Li Tong**, Sun G, Wang S, et al. ChartMind: Benchmark and Reasoning Insights of Multimodal Chart Question Answering [C], Reviewing, ACM Conference on Human Factors in Computing Systems, 2026 (领域**Top**会议, CCF-A, 一作);
- [5] Zheng X, **Li Tong**, Zhou Liu et al. MegaCQA: A chart question answering dataset for general visual understanding [J], Reviewing, Scientific Data, 2025 (**Nature** 子刊, 二作);
- [6] **Li Tong**, Sun G, Zheng X, et al. Probing the Compositional Understanding in VLMs with Visualization Representation and Insights [C], Poster, IEEE Pacific Visualization Conference, 2025 (领域三大会议之一, CCF-C, 一作);
- [7] **Li Tong**, Sun G, Zheng X, et al. Beyond Words: Unlocking Compositionality of Vision-Language Models with Visualization Insights [C], Poster, The China Visualization and Visual Analytics Conference, 2024 (国内本领域最高级别学术会议, 一作);
- [8] Chang B, Sun G, **Li Tong**, et al. MUSE: Visual Analysis of Musical Semantic Sequence [J], IEEE Transactions on Visualization and Computer Graphics, 2022 (领域 **Top** 期刊, CCF-A, 学生二作);
- [9] Xia W, Sun G, **Li Tong**, et al. Video Visualization and Visual Analytics: A Task-Based and Application-Driven Investigation [J], IEEE Transactions on Circuits and Systems for Video Technology, 2024 (**SCI 1 区**, CCF-B, 学生二作);
- [10] Jiang Q, Sun G, **Li Tong**, et al. AutoMA: Automated Generation of Multi-level Annotations for Time Series Visualization [C], IEEE Pacific Visualization Conference, 2025 (领域三大会议之一, CCF-C, 学生二作);
- [11] Jiang Q, Sun G, **Li Tong**, et al. Qutaber: Task-based Exploratory Data Analysis with Enriched Context Awareness [C], The China Visualization and Visual Analytics Conference, 2021 (国内本领域最高级别学术会议, 学生二作);
- [12] Wang Y, Sun G, Zhu Z, **Li Tong**, et al. E2Storyline: Visualizing the Relationship with Triplet Entities and Event Discovery [J], ACM Transactions on Intelligent Systems and Technology, 2023 (**JCR Q1**, 4/6);
- [13] 李童, 孙国道, 王海霞, 高海东, 谭谔, 梁荣华. 面向视觉语言模型组合性理解的可视分析方法[J]. 计算机辅助设计与图形学学报, (CCF 计算领域高质量科技期刊分区 **T1**, 一作);

学术服务

- ◇ 国内外著名会议/期刊审稿人: IEEE VIS (CCF-A), UbiComp/ISWC (CCF-A), IEEE VR (CCF-A), Euro Vis (CCF-B), IEEE TMM (CCF B), IJCHI (CCF B), npj Heritage Science (Nature 子刊) 等
- ◇ 学术会议/组织学生志愿者: 中国可视化与可视分析大会、中国图象图形学会

荣誉表彰

2023.01 第八届中国研究生智慧城市技术与创意大赛，国家三等奖（第一完成人）
2020.08 第十一届中国大学生服务外包创新创业大赛，国家三等奖（第一完成人）
2021.10 “运河杯”大学生课外学术科技基金立项（项目负责人）

中国科协
中国教育部

参与科研项目

以科研骨干参与课题组以下纵/横向科研项目：

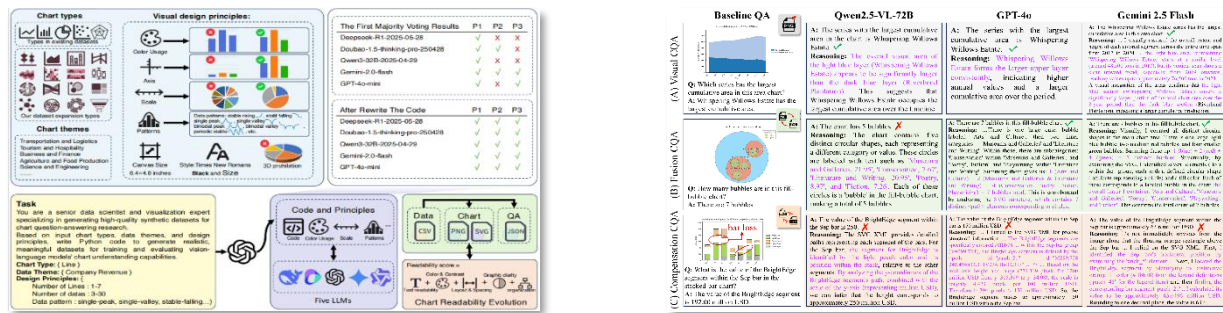
- [1] 制造企业数据空间设计理论与方法，国家重点研发计划项目
- [2] 面向跨设备多场景视频理解的可视分析技术研究，国家自然科学基金项目
- [3] 面向多模态数据的交互式对比学习和可视分析，国家自然科学基金项目
- [4] 多模态数据智能可视分析，国家自然科学基金项目优秀青年项目

科研经历概述

一、面向多模态大语言模型的推理可解释研究

(由国家自然科学基金重点项目支持, 2024.09-2025.10)

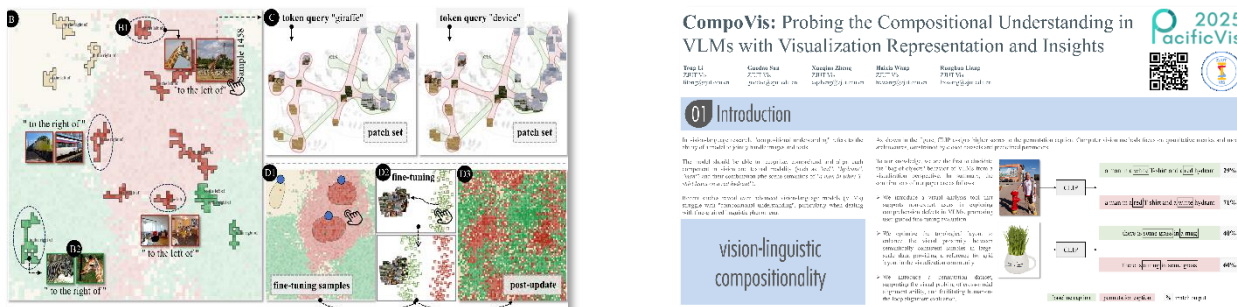
面向多模态大语言模型推理解析不足的挑战，设计数据与可视化双重约束机制，构建 230 万+图表问答基准，突破现有模型知识边界；系统评测 12 个代表性多模态大语言模型，通过视觉、融合、退化等对比实验分析推理能力边界；建立推理解构框架，量化分析多模态调用机制和思维链差异。数据集于 2025 年 9 月公开后下载量超 90 次，且研究成果已支持 3 余项后续研究。



二、面向跨模态语义对齐的可视分析研究

(由国家自然科学基金重点项目支持, 2022.10-2025.03)

面向视觉语言模型的跨模态对齐缺陷问题，提出了可交互的诊断分析方法。提出的拓扑优化算法使布局性能提升 37.5%，支持大规模数据的对齐模式识别；开发的原型系统支持像素级诊断反馈，突破了传统离线微调的局限；揭示了主流视觉语言模型的关键缺陷，为改进跨模态对齐提供见解支撑。数据集于 2025 年 1 月公开后下载量超 300 次，且研究成果已被工业缺陷检测团队借鉴。



三、面向监控视频语义理解的可视分析研究

(由国家自然科学基金重点项目支持, 2019.09-2022.09)

面向监控视频检索效率低与跨时序理解难等挑战，提出跨时序语义模式建模框架，实现时序视频流到结构化表征的有效转换；引入不确定性量化方法解构神经网络输出误差，融合呈现不确定性信息与视觉活动序列；开发的可视分析系统支持小时级视频交互式探索，实现视频-事件-帧多粒度分析。研究成果已在多个真实监控场景上验证有效，并与杭州市公安局开展技术交流。

