

蒋铖杰简历



✉ jiangcj25@mails.tsinghua.edu.cn • 📄 Google Scholar • ☎ +86-186-0158-0580

研究方向: 可控视频生成 (世界模型) • 多模态大模型 • 图像融合 • 遥感

🎓 教育背景

本科: 中央财经大学 (计算机科学与技术) 绩点: 90.17/100 排名: 1/30 2021.9 ~ 2025.7

硕士: 清华大学 (计算机技术) 绩点: 3.8/4.0 导师: 袁春教授 2025.9 ~ 至今

△ 科研成果 (一作/共一 × 4, 通讯 × 1, 其他 × 3)

图像生成

- **Towards Unified Semantic and Controllable Image Fusion: A Diffusion Transformer Approach**
Jiayang Li*, **Chengjie Jiang***, Junjun Jiang†, Pengwei Liang, Jiayi Ma, Liqiang Nie [TPAMI 2026 \(IF: 18.6\)](#) 🔄
- 提出统一的 DiT 图像融合框架, 支持多类融合任务, 并扩展到文本可控融合与多模态分割。
- **RIS-FUSION: Rethinking Text-Driven Infrared and Visible Image Fusion From The Perspective of Referring Image Segmentation**
Siju Ma, Changxiyu Gong, Xiaofeng Fan, Yong Ma, **Chengjie Jiang†** [ICASSP 2026 Oral](#) 🔄
- 从指代分割视角重构文本驱动的红外-可见光融合, 提升可解释性与下游一致性。
- **Two in One: Robust Fusion of Infrared and Visible Images in Rainy Condition**
Jing Li, Jiafeng Yan, **Chengjie Jiang**, Bin Yang† [JAS 2026 \(IF: 19.2\)](#)
- 将图像融合和去雨任务耦合, 实现任务间的互促进。
- **Where Fusion Meets Dehazing: A Coupled Framework for Robust Visible-Infrared Image Fusion in Haze**
Jing Li, Jiafeng Yan, **Chengjie Jiang**, Bin Yang, Yu Liu† [TIP 在投 \(IF:13.7\)](#)
- 将图像融合和去雾任务耦合, 实现任务间的互促进。

多模态理解

- **GRASP: Geospatial Pixel Reasoning via Structured Policy Learning**
Chengjie Jiang, Y. Zhou, J. Yan, J. Li†, J. Li, Y. Zhou, H. He, J. Li [\[Major Revision\] ISPRS 在投 \(IF:12.2\)](#)
- 提出基于强化学习的遥感指令分割框架, 在无需密集掩膜监督下提升语言到像素级分割的泛化能力。
- **Look Where It Matters: Training-Free Ultra-HR Remote Sensing VQA via Adaptive Zoom Search**
Yunqi Zhou*, **Chengjie Jiang***, Chun Yuan, Jing Li† [AAAI 2027 在投](#) 🏠
- 提出无需训练的超高分辨率遥感问答流水线, 降低 token 与显存开销的同时提升精度与推理效率。
- **FOVIS: Foveated Vision for Ultra-High-Resolution Remote Sensing Reasoning**
Y. Zhou*, **Chengjie Jiang***, H. Zheng, X. Wang, S. Xu, Z. Long, L. Shi, X. Fan, C. Yuan† [NeurIPS 2026 在投](#)
- 对模型的注意力机制进行改进, 实现模型能够基于问题意图动态选择关键区域进行精细注意力计算。
- **EmBRACE-3K: Embodied Reasoning and Action in Complex Environments**
M. Lin*, W. Huang*, Y. Li, **Chengjie Jiang**, K. Wu, F. Zhong, S. Qian†, X. Wang, X. Qi† [ECCV 2026 在投](#)
- 构建 EmBRACE-3K 具身多模态基准, 评测模型在探索、动态空间与多阶段目标执行上的能力。

🏢 实习经历

- 腾讯光子工作室群 基础研究组—计算机视觉 & 多模态大模型 导师: 钱溇钜 2024.07 ~ 2025.06
- 研究多模态令牌剪枝与并行解码方法, 降低视觉 token 冗余和推理开销, 提升多模态大模型推理效率。
- 围绕 MLLM 视觉思维链推理开展研究, 后训练出一个游戏理解垂域大模型。
- 参与 EmBRACE-3K 数据集构建, 覆盖语言指令驱动的多步交互、动态空间-语义推理与具身任务评测。
- 参与游戏场景实时生成项目, 负责自动化生成街景招牌游戏资产。

🏆 荣誉奖项

- 美国大学生数学建模竞赛 M 奖 (6%) 2024.05
- 中央财经大学华为奖学金 (全校除会计学院外共评五人) 2023.12
- 全面发展一等奖学金两次 (每专业仅一名) 2022.11、2023.11
- 卓越学术奖学金一次 (全院包括研究生共 12 名) 2023.12

🛠 专业技能

- 编程语言: C++, Python, L^AT_EX • 英语能力: CET-4(551), CET-6(512) • 其他证书: FRM 一级认证