

蔡沛霖

+86 18739390111 | peilinca@usc.edu | caipeilin.com

教育经历

南加州大学 (University of Southern California)
计算机科学理学硕士

武汉大学
计算机科学工学学士

洛杉矶, 加利福尼亚州
2024年8月 - 2026年5月

武汉, 湖北
2020年9月 - 2024年6月

出版物

Secure On-Device Video OOD Detection Without Backpropagation 2025年10月
International Conference on Computer Vision (ICCV) 2025 [PDF] — [arXiv:2503.06166]

Shawn Li; **Peilin Cai**; Yuxiao Zhou; Zhiyu Ni; Renjie Liang; You Qin; Yi Nian; Zhengzhong Tu; Xiyang Hu; Yue Zhao

A Personalized Conversational Benchmark: Towards Simulating Personalized Conversations 2025年10月

NeurIPS Workshop MTI-LLM Spotlight (Top 5%) 2025 [PDF] — [arXiv:2505.14106]

Li Li; **Peilin Cai**; Ryan A. Rossi; Franck Dernoncourt; ...; Philip S. Yu; Xiyang Hu; Yue Zhao

研究经历

The Earth Simulator: 构建可扩展的街景世界模型 洛杉矶, 加利福尼亚州
研究助理. 南加州大学 Viterbi 学院 & GVL Lab. 指导教师: **Yue Wang** 教授 2025年6月 - 至今

- 构建街景世界模型: 将少量网络街景图像转换为沿用户定义轨迹、由动作调控的长时程探索视频。
- 采用完全自监督方式在互联网图像与视频上训练, 且无需相机位姿或外部标注, 并可借助易得数据扩展到多样化城市。
- 将基于图像的渲染与学习到的生成先验相结合, 以恢复场景结构、在必要时补全被遮挡内容, 并保持多视角与时序一致性。
- 生成符合几何约束的帧, 保留室外高频细节并实现稳定的长时滚动; 优于仅重建或仅生成的基线, 支持交互式探索、规划前瞻以及大规模合成经验生成。

LAM: 语言驱动的艺术化对象建模器 (Language Articulated Object Modelers) 洛杉矶, 加利福尼亚州

共同作者. 南加州大学 Viterbi 学院 & iLab. 指导教师: **Laurent Itti** 教授 2025年2月 - 2025年9月
已提交 *ICLR 2026*

- 提出将艺术化对象生成表述为“统一代码生成任务”的流程, 使几何与关节从零开始协同设计。
- 构建具身式代理流水线, 协调 *Link Designer* (层级结构推理)、*Geometry & Articulation Coders* (代码编写、编译、调试) 与 *Geometry & Articulation Checkers* (自我校验)。
- 以代码作为可解释的桥梁来选择关节类型并精确计算其位置关系, 确保部件间关系正确。
- 可依据文本提示自动生成复杂、过程化定义的艺术化对象, 展现出可靠性与可扩展性。

个性化会话基准: 迈向个性化会话的模拟 (PersonaConvBench) 洛杉矶, 加利福尼亚州
研究助理. 南加州大学 Viterbi 学院 & FORTIS 实验室. 指导教师: **Yue Zhao** 教授 2025年2月 - 2025年5月
已提交 *ICLR 2026*

- 共同创建 **PersonaConvBench**: 用于多轮会话中的个性化推理与生成的大规模基准。本人负责完整技术实现、数据获取与组织、全套实验的设计与执行以及多轮方法迭代。
- 定义三类核心任务——句子分类、影响力回归与以用户为中心的生成——覆盖10个基于Reddit的多样领域。
- 构建统一的提示与评测体系; 在使用个性化历史时, 相较最佳非会话基线, 情感分类取得+198%的相对提升。
- 公开基准、评测与代码, 促进面向长时程、用户自适应的LLM研究 (arXiv:2505.14106)。

SecDOOD: 无需反向传播的安全终端侧视频OOD 检测

研究助理·南加州大学 Viterbi 学院 & FORTIS 实验室·指导教师: Yue Zhao 教授
ICCV 2025 (海报)

洛杉矶, 加利福尼亚州
2024 年 11 月 - 2025 年 3 月

- 共同提出 SecDOOD: 一种安全的云—设备协作框架, 使终端侧在无需反向传播的情况下完成 OOD 检测; 本人负责完整技术实现、模型构建与长期方案迭代。
- 实现基于超网络 (HyperNetwork) 的个性化参数生成器, 使中心训练的模型能适配设备特定分布而无需本地微调。
- 设计动态特征采样与选择性加密, 仅加密最具信息量的通道, 在保持检测质量的前提下降低开销。
- 在多数据集/多 OOD 场景下的大量实验表明, 其性能可与完全微调的基线相当 (arXiv:2503.06166)。

PLayTR: 基于几何布局驱动的室内结构化平面重建

研究助理·武汉大学·指导教师: 夏桂松教授

武汉, 湖北
2023 年 6 月 - 2024 年 6 月

- 参考已有平面重建方法搭建原型; 通过将 2D 原始图像中编码的 3D 布局信息与“曼哈顿世界”线索引入交叉注意力 Transformer, 提高平面重建与分割精度。
- 将原方法对平面深度的“自顶向下全局预测”改进为“逐平面预测后, 结合掩码对深度图进行加权求和”。
- 设计数据生成模块, 从数据集中精确抽取满足条件的 3D 布局三线段集合, 并筛除无效数据。
- 重组评测流程, 引入新的深度监督策略, 用更高效的损失函数替换旧方法, 并为新的三线段组与平面组设计匈牙利匹配。
- 在不更换主干网络的情况下, 分割结果较原模型提升 0.7 个百分点; 重建结果显示我们解决了平面间接触不良这一突出问题。

技术技能

编程语言: Python, C/C++, JavaScript, HTML/CSS

开发工具: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm