

# 彭正皓

主页: <https://pengzhenghao.github.io/>

邮箱: [pzh@cs.ucla.edu](mailto:pzh@cs.ucla.edu)  
GitHub: [pengzhenghao](https://github.com/pengzhenghao)  
LinkedIn [Google Scholar](https://www.linkedin.com/in/pengzhenghao/)

## 研究兴趣

强化学习、跨模态大模型 (VLA/VLM)、人在回路中学习 (Human-in-the-loop Learning)。

## 教育经历

### 加州大学洛杉矶分校 (UCLA)

2022 年 9 月—2026 年 4 月 (预计)

- 计算机科学博士生 (Ph.D.), 导师: [周博磊教授](#)。

### 香港中文大学 (CUHK)

2019 年 8 月—2022 年 7 月

- 信息工程哲学硕士 (M.Phil.), 多媒体实验室 (MMLab), 导师: [周博磊教授](#)。

### 上海交通大学 (SJTU)

2015 年 9 月—2019 年 7 月

- 船舶与海洋工程工学学士 (B.Eng.), 致远荣誉计划成员。科研助理, 导师: [蒋力教授](#)。

## 实习与工作经历

### NVIDIA, 圣克拉拉, 加利福尼亚

2025 年 6 月—2025 年 12 月

*Autonomous Vehicle Group @ NVIDIA Research* 研究实习生

经理: [Boris Ivanovic](#)

- 设计 *Counterfactual VLA (CF-VLA)*: 一种“自我反思”的 VLA 框架, 在执行动作前先批判并修正自身计划。
- 搭建 rollout-filter-label 数据引擎: 从模型 rollouts 中挖掘失败案例、筛选场景, 并使用大规模 VLM 作为教师模型自动标注反事实推理轨迹 (reasoning traces)。
- 将 Qwen2.5-VL 融入自动驾驶 VLA: 设计领域专用视频编码器/解码器, 实现多节点 FSDP 训练与批量 rollouts, 并自动化数据集整理与扩充 (10M+ 视频片段)。
- 在大规模自动驾驶数据集上获得一致提升: 轨迹误差最高降低 17%, 碰撞率最高降低 21%, 并呈现自适应推理 (在高风险/困难场景中思考更多)。

### Waymo, 山景城, 加利福尼亚

2023 年 6 月—2023 年 9 月

*Waymo Research* 研究实习生

导师: [Justin Fu](#), [Rowan McAllister](#)

- 使用强化学习微调多智能体行为模型; 采用不依赖价值网络的 REINFORCE 训练大规模 Transformer (思路与后来的 GRPO 相近, 但更早实现)。
- 在 Waymo Open Sim Agents Challenge (WOSAC) 中降低碰撞与出界率等安全指标, 同时保持生成场景的真实感, 综合得分提升 22%。
- 构建新的评测框架: 在模型生成的真实仿真场景中更好地排名与测试自动驾驶规划器。

## 奖励与荣誉

- UCLA 学位论文年奖 (Dissertation Year Award), 2025–2026
- Amazon Fellowship, 2024–2025, UCLA
- University Fellowship, 2023–2024、2024–2025, UCLA
- 工程学院优秀导师奖 (Outstanding Tutors Award), 2021, CUHK
- Teaching Assistant Award, 2020–2021 学年第一、第二学期, CUHK
- Postgraduate Studentship, 2019–2022, CUHK
- 致远荣誉奖学金 (Zhiyuan Honors Scholarship), 2015–2017, SJTU

## 代表性论文

完整论文列表: [Google Scholar](#)。以下列出部分代表性论文:

13. **Alpamayo-R1: Bridging Reasoning and Action Prediction for Generalizable Autonomous Driving in the Long Tail.**  
Yan Wang, ..., Zhenghao Peng, ..., Marco Pavone (预印本 / Preprint)
12. **Counterfactual VLA: Self-Reflective Vision-Language-Action Model with Adaptive Reasoning.**  
Zhenghao Peng, Wenhao Ding, ..., Marco Pavone (投稿中 / In Submission)
11. **Predictive Preference Learning from Human Interventions.**  
Haoyuan Cai, Zhenghao Peng, and Bolei Zhou (NeurIPS 2025 Spotlight) [PDF, Code, Webpage]
10. **Robot-Gated Interactive Imitation Learning with Adaptive Intervention Mechanism.**  
Haoyuan Cai, Zhenghao Peng, and Bolei Zhou (ICML 2025) [PDF, Code, Webpage]
9. **Data-Efficient Learning from Human Interventions for Mobile Robots.**  
Zhenghao Peng, Zhizheng Liu, and Bolei Zhou (ICRA 2025) [Webpage, PDF]
8. **Improving Agent Behaviors with RL Fine-tuning for Autonomous Driving.**  
Zhenghao Peng, Wenjie Luo, Yiren Lu, T. Shen, C. Gulino, A. Seff, and Justin Fu (ECCV 2024) [PDF]
7. **Learning from Active Human Involvement through Proxy Value Propagation.**  
Zhenghao Peng, W. Mo, C. Duan, Q. Li, and Bolei Zhou (NeurIPS 2023 Spotlight) [PDF, Webpage]
6. **ScenarioNet: Open-Source Platform for Large-Scale Traffic Scenario Simulation and Modeling.**  
Quanyi Li\*, Zhenghao Peng\*, Lan Feng, Zhizheng Liu, Chenda Duan, Wenjie Mo, and Bolei Zhou (NeurIPS 2023) [PDF, Code, Webpage]
5. **MetaDrive: Composing Diverse Driving Scenarios for Generalizable Reinforcement Learning.**  
Quanyi Li\*, Zhenghao Peng\*, Z. Xue, Q. Zhang, and Bolei Zhou (TPAMI 2021) [PDF, Code, Webpage]
4. **Efficient Learning of Safe Driving Policy via Human-AI Copilot Optimization.**  
Quanyi Li\*, Zhenghao Peng\*, and Bolei Zhou (ICLR 2022) [PDF, Code, Webpage]
3. **Safe Driving via Expert Guided Policy Optimization.**  
Zhenghao Peng\*, Quanyi Li\*, C. Liu, and Bolei Zhou (CoRL 2021) [PDF, Code, Webpage, Poster]
2. **Learning to Simulate Self-Driven Particles System with Coordinated Policy Optimization.**  
Zhenghao Peng, Q. Li, K. Hui, C. Liu, and Bolei Zhou (NeurIPS 2021) [PDF, Code, Webpage, Poster]
1. **Non-local Policy Optimization via Diversity-regularized Collaborative Exploration.**  
Zhenghao Peng, Hao Sun, and Bolei Zhou (arXiv 2020) [PDF]

## 开源项目

---

[MetaDrive](#) 是强化学习与自动驾驶领域被广泛使用的开源驾驶模拟器，获得约 ~1000 GitHub stars 与 350+ 学术引用。更多项目请见 [GitHub](#)。

## 学术报告

---

- Human-in-the-loop Agent Learning, 密歇根大学 EECS 598: Action and Perception 课程客座讲座 (受邀: Stella Yu), 2024 年 5 月。

## 教学经历

---

- UCLA: CS260R 强化学习 (2025 冬、2023 秋、2022 秋)
- CUHK: IERG5350 强化学习 (2019–2020 第二学期、2020–2021 第一/第二学期、2021–2022 第一学期)

## 其他

---

- 审稿服务: NeurIPS、ICML、CVPR、ECCV/ICCV、ICLR、CoRL、RSS、IROS、ICRA、AAAI、TNNLS、IJCV、RA-L 等。
- 技能: PyTorch、JAX、Ray、RLLib、ROS2、Docker、Transformers、FSDP 等。