Zhenghao "Mark" PENG

Homepage: https://pengzhenghao.github.io/
Email: pzh@cs.ucla.edu
Github: pengzhenghao
LinkedIn Google Scholar

RESEARCH INTERESTS

Reinforcement Learning, Multimodal LLM (VLA, VLM) and Human-in-the-loop Learning.

EDUCATION

University of California, Los Angeles (UCLA) *September 2022 - April 2026 (Expected)*

- Ph.D. student at the Department of Computer Science, supervised by Prof. Bolei Zhou.

The Chinese University of Hong Kong (CUHK) *August 2019 - July 2022*

- M.Phil. in Information Engineering, supervised by Prof. Bolei Zhou at Multimedia Lab (MMLab).

Shanghai Jiao Tong University (SJTU) *Sept. 2015 - July 2019*

- B.Eng. in Naval Architecture and Ocean Engineering. Member of Zhiyuan Honor Program.
- Research assistant supervised by Prof. Li Jiang.

EXPERIENCE

NVIDIA, Santa Clara, CA *June 2025 – December 2025*
Research Intern in Autonomous Vehicle Group at NVIDIA Research *Manager: Boris Ivanovic*

- Designed *Counterfactual VLA (CF-VLA)*, a self-reflective VLA framework that critiques and corrects its own actions before execution.
- Built a rollout–filter–label data engine that mines failure cases from model rollouts, filters scenes, and uses large VLM as a teacher to auto-label counterfactual reasoning traces.
- Integrated Qwen2.5-VL into the driving VLA with domain-specific video encoders/decoders, implemented multi-node FSDP training and batched rollouts, and automated dataset curation for 10M+ video clips.
- Experimented on large-scale driving datasets and showed consistent gains in all metrics, demonstrating up to 17% lower trajectory error, and 21% lower collisions than non-reasoning baselines. The model demonstrates ***adaptive reasoning***: it thinks more in difficult, high-risk scenarios.

Waymo, Mountain View, CA *June 2023 - September 2023*
Research Intern in Waymo Research *Mentored by Justin Fu and Rowan McAllister*

- Fine-tuned multi-agent behavior models with reinforcement learning. Our method uses REINFORCE to train large transformer without the value net, similar to GRPO though much earlier than it.
- Reduced collision and off-road rates and improved safety-critical metrics on the Waymo Open Sim Agents Challenge (WOSAC), raising composite performance by 22%.
- Built a new evaluation framework that better ranks and tests autonomous driving planners in realistic scenarios generated by our model.

AWARDS AND HONORS

- Dissertation Year Award 2025-2026, UCLA
- Amazon Fellowship, 2024-2025, UCLA
- University Fellowship, 2023-2024, 2024-2025, UCLA
- The Outstanding Tutors Award 2021 of the Faculty of Engineering , 2021, CUHK
- Teaching Assistant Award, Term 1 2020 - 2021, Term 2 2020 - 2021, CUHK
- Postgraduate Studentship, 2019 - 2022, CUHK
- Zhiyuan Honors Scholarship, 2015 - 2017, SJTU

SELECTED PAPERS

For the complete list of publications, please visit my Google Scholar. Selected papers:

[13] **Alpamayo-R1: Bridging Reasoning and Action Prediction for Generalizable Autonomous Driving in the Long Tail**.
NVIDIA: Yan Wang, ..., Zhenghao Peng, ..., Marco Pavone **(Preprint)**

[12] **Counterfactual VLA: Self-Reflective Vision-Language-Action Model with Adaptive Reasoning**.
Zhenghao Peng, Wenhao Ding, ..., Marco Pavone **(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]

OPEN-SOURCE PROJECTS

MetaDrive is a widely used open-source driving simulator for reinforcement learning and autonomous driving, which received ~1000 GitHub stars and 350+ citations. For all my open-source projects, see GitHub.

TALKS

- Human-in-the-loop Agent Learning, EECS 598: Action and Perception Guest Lecture, invited by: Stella Yu, May 2024

TEACHING EXPERIENCES

- CS260R Reinforcement Learning, UCLA, 2025 Winter & 2023 Fall & 2022 Fall
- IERG5350 Reinforcement Learning, CUHK, Term 1, 2021-22 & Term 1 and 2, 2020-21 & Term 2, 2019-20

MISCELLANEOUS

Review Service: NeurIPS, ICML, CVPR, ECCV/ICCV, ICLR, CoRL, RSS, IROS, ICRA, AAAI, TNNLS, IJCV, RA-L, *etc.* **Skills:** PyTorch, Jax, Ray, RLLib, ROS2, Docker, Transformers, FSDP, *etc.*