

PENG Zhenghao

Homepage: <https://pengzhenghao.github.io/>

295, Engineering VI, UCLA

Github: [pengzhenghao](#)

Email: pzh@cs.ucla.edu

RESEARCH INTERESTS

My research interests cover multi-agent RL (for potential usage in open-world games) and human-AI interaction (like playing with AI). Hope you enjoy my works!

EDUCATION

University of California, Los Angeles (UCLA) *September 2022 - Present*

- PhD student at the Department of Computer Science, supervised by [Prof. Zhou Bolei](#).

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

- MPhil student at the Department of Information Engineering, supervised by [Prof. Zhou Bolei](#).

University of California, Berkeley (UCB) *July 2017 - Aug. 2017*

- Summer session.

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

- Bachelor of Engineering and member of Zhiyuan Honors Program.
- Research assistant supervised by [Prof. Jiang Li](#).

AWARDS AND HONORS

The Outstanding Tutors Award 2021 of the Faculty of Engineering *2021, CUHK*

Teaching Assistant Awards *Term 1 & Term 2, 2020 - 2021, CUHK*

Postgraduate Studentship *2019 - 2022, CUHK*

Zhiyuan Honors Scholarship *2015 - 2017, SJTU*

RESEARCH PAPERS

[1] Lan Feng*, Quanyi Li*, **Zhenghao Peng***, Shuhan Tan, and Bolei Zhou. Trafficgen: Learning to generate diverse and realistic traffic scenarios. In *2023 International Conference on Robotics and Automation (ICRA)*. IEEE, 2023 (**ICRA 2023**) [[PDF](#), [Website](#)]

[2] Zhenghai Xue, **Zhenghao Peng**, Quanyi Li, Zhihan Liu, and Bolei Zhou. Guarded policy optimization with imperfect online demonstrations. In *International Conference on Learning Representations, 2023 (ICLR 2023)* [[OpenReview](#)]

[3] Quanyi Li, **Zhenghao Peng**, Haibin Wu, Lan Feng, and Bolei Zhou. Human-AI shared control via policy dissection. *Advances in Neural Information Processing Systems*, 2022 (**NeurIPS 2022**) [[PDF](#), [Code](#), [Website](#)]

[4] Hao Sun, Ziping Xu, Meng Fang, **Zhenghao Peng**, Jiadong Guo, Bo Dai, and Bolei Zhou. Mopa: a minimalist off-policy approach to safe-rl. 2022 (Deep RL Workshop NeurIPS 2022)

[5] Hao Sun, **Zhenghao Peng**, Bo Dai, Jian Guo, Dahua Lin, and Bolei Zhou. Novel policy seeking with constrained optimization. 2022 (Deep RL Workshop NeurIPS 2022)

[6] Qihang Zhang, **Zhenghao Peng**, and Bolei Zhou. Learning to drive by watching youtube videos: Action-conditioned contrastive policy pretraining. *European Conference on Computer Vision*, 2022 (**ECCV 2022**) [[PDF](#), [Website](#)]

[7] Quanyi Li*, **Zhenghao Peng***, Zhenghai Xue, Qihang Zhang, and Bolei Zhou. Metadrive: Composing diverse driving scenarios for generalizable reinforcement learning. *IEEE transaction on Pattern Analysis and Machine Intelligence*, 2021 (**TPAMI**) [[Paper](#), [Code](#), [Website](#)]

- [8] Boli Fang, **Zhenghao Peng**, Hao Sun, and Qin Zhang. Meta proximal policy optimization for cooperative multi-agent continuous control. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022
- [9] Mingxin Huang, Yuliang Liu, **Zhenghao Peng**, Chongyu Liu, Dahua Lin, Shenggao Zhu, Nicholas Yuan, Kai Ding, and Lianwen Jin. Swintextspotter: Scene text spotting via better synergy between text detection and text recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022 (CVPR 2022)*
- [10] Quanyi Li*, **Zhenghao Peng***, and Bolei Zhou. Efficient learning of safe driving policy via human-AI copilot optimization. In *International Conference on Learning Representations, 2022 (ICLR 2022)* [[PDF](#), [Code](#), [Website](#)]
- [11] **Zhenghao Peng***, Quanyi Li*, Chunxiao Liu, and Bolei Zhou. Safe driving via expert guided policy optimization. In *5th Annual Conference on Robot Learning, 2021 (CoRL 2021)* [[PDF](#), [Code](#), [Website](#), [Poster](#)]
- [12] **Zhenghao Peng**, Quanyi Li, Ka Ming Hui, Chunxiao Liu, Bolei Zhou, et al. Learning to simulate self-driven particles system with coordinated policy optimization. *Advances in Neural Information Processing Systems*, 34, 2021 (**NeurIPS 2021**) [[PDF](#), [Code](#), [Website](#), [Poster](#)]
- [13] Quanyi Li*, **Zhenghao Peng***, Qihang Zhang, Chunxiao Liu, and Bolei Zhou. Improving the generalization of end-to-end driving through procedural generation. *arXiv preprint arXiv:2012.13681*, 2020 [[PDF](#), [Repo](#), [Website](#)]
- [14] **Zhenghao Peng**, Hao Sun, and Bolei Zhou. Non-local policy optimization via diversity-regularized collaborative exploration. *arXiv preprint arXiv:2006.07781*, 2020 [[PDF](#)]
- [15] Zhuoran Song, Dongyu Ru, Ru Wang, Hongru Huang, **Zhenghao Peng**, Jing Ke, Xiaoyao Liang, and Li Jiang. Approximate random dropout. In *Design, Automation & Test in Europe Conference & Exhibition, 2019. DATE'19*. IEEE, 2019 [[PDF](#)]
- [16] **Zhenghao Peng**, Xuyang Chen, Chengwen Xu, Naifeng Jing, Xiaoyao Liang, Cewu Lu, and Li Jiang. Axnet: Approximate computing using an end-to-end trainable neural network. In *Proceedings of the 2018 International Conference on Computer-Aided Design. ICCAD'18*. IEEE/ACM, 2018 [[PDF](#)]

RESEARCH EXPERIENCES

Efficient Learning through Human-AI Copilot [10]

July 2021 - November 2021

Supervised by [Prof. Zhou Bolei](#)

- Proposed the Human-AI Copilot (HACO) algorithm for human-in-the-loop RL that trains agents from human interventions, partial demonstrations and free exploration, even without reward.
- HACO achieves high sample efficiency, high safety and low human cognitive cost.
- Please visit <https://decisionforce.github.io/HACO/>.

Safe Reinforcement Learning System via Expert in the Loop [11]

March 2021 - June 2021

Supervised by [Prof. Zhou Bolei](#)

- Proposed an Expert Guided Policy Optimization (EGPO) framework for safe RL, which incorporates the guardian mechanism in the interaction of agent and environment to ensure safe and efficient exploration.
- The experiments on safe driving shows EGPO can achieve training and test-time safety and better performance.
- Please visit <https://decisionforce.github.io/EGPO/>.

Simulating Realistic Traffic Flow via Multi-agent RL [12]

Feb. 2021 - May 2021

Supervised by [Prof. Zhou Bolei](#)

- Developed a novel MARL method called Coordinated Policy Optimization (CoPO) to incorporate social psychology principle to learn neural controller for a population of autonomous driving vehicles.
- The vehicles population learned by CoPO achieves superior performance and exhibits complex and socially compliant behaviors that improve the traffic efficiency and safety.
- Please visit: <https://decisionforce.github.io/CoPO/>

Autonomous Driving Simulator MetaDrive [7]

July 2020 - Present

Supervised by [Prof. Zhou Bolei](#)

- Developed the [MetaDrive](#), an open-ended and highly customizable driving simulator based on Panda3D and Bullet.
- Utilized procedural generation to generate infinite driving scenes with different road networks and traffic flows.

Efficient Asynchronous Reinforcement Learning [14]

Jan. 2020 - July 2020

Supervised by [Prof. Zhou Bolei](#)

- Proposed Ensemble Policy Optimization (EPO) framework that trains multiple heterogeneous policies simultaneously solving the same task while maintaining the diversity of the ensemble.
- EPO substantially improves sample efficiency in continuous locomotion tasks compared to the single-policy optimization counterparts.

TEACHING EXPERIENCES

- CS269 Seminar on Reinforcement Learning at UCLA, Fall, 2022-23
- IERG5350 Reinforcement Learning at CUHK, Term 1, 2021-22
- CSCI2100E Data Structures at CUHK, Term 2, 2020-21
- IERG5350 Reinforcement Learning at CUHK, Term 1, 2020-21
- IERG6130 Seminar on Reinforcement Learning at CUHK, Term 2, 2019-20

MISCELLANEOUS

Programming Languages: Python, Matlab, HTML, CSS, C++, etc.

ML Frameworks: Ray, RLlib, TensorFlow, PyTorch, Keras, etc.

Skills: Git, \LaTeX , PyCharm, Keynote, Photoshop, Final Cut, Cantonese, etc.

Hobbies: Genshin Impact, Badminton, Cycling, Hiking, Movie, Science Fiction