

# PENG Zhenghao

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

Bilibili: 彭正皓

Github: [pengzhenghao](#)

Email: [pengzh@ie.cuhk.edu.hk](mailto:pengzh@ie.cuhk.edu.hk)

## EDUCATION

---

### The Chinese University of Hong Kong (CUHK)

August 2019 - Present

- Graduate student under the supervision of [Prof. Zhou Bolei](#) at Multimedia Lab, Department of Information Engineering.

### Shenzhen Institutes of Advanced Technology (SIAT), Chinese Academy of Sciences

June 2018 - Sept. 2018

- Research intern at Multimedia Research Center, under the supervision of [Prof. Qiao Yu](#).

### University of California, Berkeley

July 2017 - Aug. 2017

- Summer session.

### Shanghai Jiao Tong University (SJTU)

Sept. 2015 - July 2019

- B.E. in Naval Architecture and Ocean Engineering.
- Member of Zhiyuan Honors Program.
- Research assistant at Advanced Computer Architecture Laboratory with [Prof. Jiang Li](#).

## RESEARCH PAPERS

---

[1] Quanyi Li\*, **Zhenghao Peng\***, and Bolei Zhou. Efficient learning of safe driving policy via human-ai copilot optimization. In *International Conference on Learning Representations, 2021 (ICLR 2022)* [ [PDF](#), [Code](#), [Website](#) ]

[2] **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](#) ]

[3] **Zhenghao Peng**, 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](#) ]

[4] Hao Sun, Ziping Xu, Meng Fang, **Zhenghao Peng**, Jiadong Guo, Bo Dai, and Bolei Zhou. Safe exploration by solving early terminated mdp. *arXiv preprint arXiv:2107.04200*, 2021 [ [PDF](#) ]

[5] 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](#) ]

[6] **Zhenghao Peng**, Hao Sun, and Bolei Zhou. Non-local policy optimization via diversity-regularized collaborative exploration. *arXiv preprint arXiv:2006.07781*, 2020 [ [PDF](#) ]

[7] Hao Sun, **Zhenghao Peng**, Bo Dai, Jian Guo, Dahua Lin, and Bolei Zhou. Novel policy seeking with constrained optimization. *arXiv preprint arXiv:2005.10696*, 2020 [ [PDF](#) ]

[8] Hao Sun, Jiankai Sun, **Zhenghao Peng**, Dahua Lin, and Bolei Zhou. Learning with identity and uniqueness through social constraint. In *NeurIPS 2019 Deep RL Workshop*. IEEE, 2019

[9] 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](#) ]

[10] **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 [1]

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 [2]**

*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 [3]**

*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 [5]**

*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 [6]**

*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.

### **Approximate Random Dropout [9]**

*Apr. 2018 - Sept. 2018*

*Supervised by Prof. Jiang Li*

- Replaced the random dropout in neural network training with predefined dropout patterns, enabling skips of the redundant zero-multiplication and thus speedup. The method reduced the training time by 20%-77% (dropout rate is 0.3-0.7) on MLP and LSTM with low accuracy drop.
- I provided proof on the equivalence of this approach and the random dropout and polished the paper.

### **Neural Network Based Approximate Computing Framework [10]**

*Sept. 2017 - May 2018*

*Supervised by Prof. Jiang Li*

- Fused the conventional approximate computing architecture as an end-to-end structure. Applied the Hadamard product at the output of predictor and the activation of approximator to learn jointly.
- Reduced 74% training time; Improved the invocation by 50.6%, averagely. In JPEG benchmark, 60% parameters are reduced and  $32\times$  speedup are achieved.

### **Handwritten Arithmetic Expression Recognition**

*May 2017 - June 2017*

*Course Project*

- Expanded MNIST dataset with 10000+ operators images and built an End-to-End system providing recognition and calculation service of handwritten arithmetic expression via WeChat.

## AWARDS AND HONORS

---

Teaching Assistant Awards	<i>Term 1 &amp; Term 2, 2020 - 2021, CUHK</i>
Postgraduate Studentship	<i>2019 - 2022, CUHK</i>
Zhiyuan Honors Scholarship	<i>2015 - 2017, SJTU</i>

## PRESS COVERAGE

---

BastillePost, HK01, HK Commercial Daily, South China Morning Post, Sing Tao Daily, HK Economic Journal, Ta Kung Pao. (See my [homepage](#)) *December, 2021*

## 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 (原神), Badminton, Cycling, Billiards, Movie, Science Fiction