

# HAI ZHANG

✉ zhanghai12138@tongji.edu.cn · 🌐 betray12138 (the same as zhihu)

## 👤 PERSONAL STATEMENT

A dedicated, practicing, and capable researcher in reinforcement learning (RL), specifically in model-based RL, offline RL and context-based meta RL, with absorbing more than 150 papers in related fields (see the details in zhihu). A confident and cohesive member in collaborative work, brilliant at communication and organization. A pragmatic and skillful back-end development engineer before diving into RL, with the mastering of C++, PYTHON, Golang, and some common techniques.

## 🎓 EDUCATION

**TONGJI UNIVERSITY, SHANGHAI, CHINA** 2022.09 – up to now  
*MASTER(without examination), COMPUTER SCIENCE AND TECHNOLOGY*  
with GPA of 85.19/100

**TONGJI UNIVERSITY, SHANGHAI, CHINA** 2018.09 – 2022.06  
*BACHELOR, COMPUTER SCIENCE AND TECHNOLOGY*  
with GPA of 4.76/5.0, general ranking of 8/155

## 👥 RESEARCH EXPERIENCE

\* means co-first author, † means corresponding author

**QIZHI INSTITUTE(Tsinghua University IIIS) SHANGHAI, CHINA** 2024.06 – up to now  
*RESEARCH INTERN* supervised by Prof. Yang Gao

- **OpenVLA-InternVL2.** We substitute the backbone of OpenVLA from Llama2 7B to InternVL2 1B to achieve computational and communication reduction. (FSDP → DDP)

**ZHEJIANG LAB HANGZHOU, ZHEJIANG, CHINA** 2023.07 – 2024.05  
*RESEARCH INTERN* supervised by Prof. Lanqing Li & Prof. Pheng-Ann Heng

- **Scrutinize What We Ignore: Controlling Task Representation Shift In Context-Based Offline Meta Reinforcement Learning.** Hai Zhang, Boyuan Zheng, Tianying Ji, Jinhang Liu, Anqi Guo, Junqiao Zhao†, Lanqing Li†. *Under review in ICLR 2025*
- **Towards an Information Theoretic Framework of Context-Based Offline Meta-Reinforcement Learning.** Lanqing Li\*, Hai Zhang\*, Xinyu Zhang, Shatong Zhu, Yang Yu, Junqiao Zhao†, and Pheng-Ann Heng. *In NeurIPS 2024 Spotlight*

**TIEV LAB SHANGHAI, CHINA** 2022.09 – up to now  
supervised by Prof. Junqiao Zhao

- **How to Fine-tune the Model: Unified Model Shift and Model Bias Policy Optimization.** Hai Zhang, Hang Yu, Junqiao Zhao†, Di Zhang, Chang Huang, Hongtu Zhou, Xiao Zhang, and Chen Ye. *In NeurIPS 2023*
- **Safe Reinforcement Learning with Dead-Ends Avoidance and Recovery.** Xiao Zhang, Hai Zhang, Hongtu Zhou, Chang Huang, Di Zhang, Chen Ye†, Junqiao Zhao†. *In IEEE Robotics and Automation Letters 2023 (Present in ICRA 2024)*
- **Focus On What Matters: Separated Models For Visual-Based RL Generalization.** Di Zhang, Bowen Lv, Hai Zhang, Feifan Yang, Hongtu Zhou, Chang Huang, Hang Yu, Chen Ye, Junqiao Zhao†, and Changjun Jiang. *In NeurIPS 2024*

## ⚙️ PROJECTS

---

### **Distributed Complete Vehicle Cloudization**

2022.05 – 2023.01

- Invention Patent (Submitted, Patent Number: 202310899331.2)
- Affiliated with the National Key R&D Program
- Core Technique: distributed framework architecture, k8s deployment

**NIO, SHANGHAI, CHINA**

2021.10-2022.03

*BACKEND DEVELOPMENT ENGINEER*

- LIFELONG CYCLE MANAGE SYSTEM (MongoDB replica set, MongoDB atomic operation, Nginx reverse proxy, docker deployment, Redis distributed lock)
- EDGE-SIDE DATA UPLOAD AND DISPLAY (Protobuf, web-socket, TTL index, split-table storage)
- DATA SYNCHRONIZATION SERVICE (Kafka consumer group, full and increment synchronization)
- 5T MAGNITUDE DATA MIGRATION AND BACKUP (MongoDB shared set, shared key setting, index setting, MongoDB cluster construction)

### **Unknown Environment Exploration and Application Device Based on Deep Reinforcement Learning**

2020.05-2021.03

- Innovation and Entrepreneurship Program for SHANGHAI University Students.
- Responsible for accomplishing end-to-end vehicle driving on the CARLA simulator using representation learning combined with reinforcement learning algorithms such as SAC, TD3, and PPO.
- Responsible for training visual models to work with the ROS system to fulfill target detection.

## ⚓ COMPETITIONS

---

### **RLChina Intelligent Agent Challenge Nonin Spring Season Curling Challenge**

- Second Place in finals, Sixth Place in total scores
- Responsible for optimizing PPO and rule-based intelligence to complete curling strikes in a POMDP environment

### **WAIC: Meta-verse Lights Up Autonomous Driving, AI Simulation Driving Competition**

- Second Place (Unique), won 40 thousand RMB
- Responsible for perceptual information processing and code implementation of the decision state machine

### **Intel Cup National College Students Embedded System Invitational Competition**

- National Second Prize (Leader)
- All Sight Security Eye - Intelligent Suspect Detection System
- Responsible for the overall architecture of the project implementation as well as the full-view image stitching

## 📌 OTHERS

---

- Good at badminton, hiking and photography; enthusiastic in travelling
- Outstanding Undergraduate Thesis
- Scholarship for outstanding students for three consecutive years