Haoran He

haoran.he@connect.ust.hk · (+86) 135-2494-6726

Education

Hong Kong University of Science and Technology, Hong Kong 2024.2 – 2027.6 (Expected)

Ph.D. student in Electronic and Computer Engineering

Shanghai Jiao Tong University (SJTU), Shanghai, China 2019.9 – 2023.6

Bachelor's Degree in Computer Science (CS)

Interests

I am a 3rd-year Ph.D. student at Hong Kong University of Science and Technology, advised by Prof. Ling Pan. I received my bachelor's Degree at Shanghai Jiao Tong University in June 2023, advised by Prof. Weinan Zhang and Prof. Yong Yu.

My research goal is to develop an intelligent decision-making system that possesses optimality, generalizability, interpretability, and robustness. To achieve this, I primarily focus on:

- Generalist Reinforcement Learning and its application in the real world.
- Multimodal foundation models (e.g., unified omni models) and their applications.
- Reinforcement learning for large foundation models (e.g., post-training).
- Large foundation models for reasoning and decision-making.

Publications

See full list in my Google Scholar profile

- GARDO: Reinforcing Diffusion Models without Reward Hacking
  Haoran He, Yuxiao Ye, Jie Liu, Jiajun Liang, Zhiyong Wang, Ziyang Yuan, Xintao Wang, Hangyu Mao, Pengfei Wan, Ling Pan .
  In submission, 2026

- Scaling Image and Video Generation via Test-Time Evolutionary Search
  Haoran He, Jiajun Liang, Xintao Wang, Pengfei Wan, Di Zhang, Kun Gai, Ling Pan.
  In submission, 2026

- Random Policy Valuation is Enough for LLM Reasoning with Verifiable Rewards
  Haoran He, Yuxiao Ye, Qingpeng Cai, Chen Hu, Binxing Jiao, Daxin Jiang, Ling Pan.
  ICLR, 2026

- Pre-Trained Video Generative Models as World Simulators
  Haoran He, Yang Zhang, Liang Lin, Zhongwen Xu, Ling Pan
  AAAI 2026, ICLR 2025 World Model Workshop

- Rectifying Reinforcement Learning for Reward Matching
  Haoran He, Emmanuel Bengio, Qingpeng Cai, Ling Pan.
  ICML 2025

- Task-Agnostic Pre-training and Task-Guided Fine-tuning for Versatile Diffusion Planner
  Chenyou Fan, Chenjia Bai, Zhao Shan, **Haoran He**, Yang Zhang, Zhen Wang.
  ICML 2025

- Looking Backward: Retrospective Backward Synthesis for Goal-Conditioned GFlowNets
  Haoran He, Can Chang, Huazhe Xu, Ling Pan
  ICLR 2025

- On the Value of Myopic Behavior in Policy Reuse
  Kang Xu, Chenjia Bai, Shuang Qiu, **Haoran He**, Bin Zhao, Zhen Wang, Wei Li, Xuelong Li.
  TPAMI 2025

- Learning an Actionable Discrete Diffusion Policy via Large-Scale Actionless Video Pre-Training
  Haoran He, Chenjia Bai, Ling Pan, Weinan Zhang, Bin Zhao, Xuelong Li.

NeurIPS 2024

- Regularized Conditional Diffusion Model for Multi-Task Preference Alignment
  Xudong Yu, Chenjia Bai, **Haoran He**, Changhong Wang, Xuelong Li.
  NeurIPS, 2024

- Bridging the Sim-to-Real Gap from the Information Bottleneck Perspective
  Haoran He, Peilin Wu, Chenjia Bai, Hang Lai, LingXiao Wang, Ling Pan, Xiaolin Hu, Weinan Zhang.
  CoRL, 2024, Oral

- SAM-E: Leveraging Visual Foundation Model with Sequence Imitation for Embodied Manipulation
  Junjie Zhang, Chenjia Bai, **Haoran He**, Wenke Xia, Zhigang Wang, Bin Zhao, Xiu Li, Xuelong Li.
  ICML 2024

- Robust Quadrupedal Locomotion via Risk-Averse Policy Learning
  Jiyuan Shi, Chenjia Bai, **Haoran He**, Lei Han, Dong Wang, Bin Zhao, Xiu Li, Xuelong Li.
  ICRA 2024

- Diffusion Model is an Effective Planner and Data Synthesizer for Multi-Task Reinforcement Learning
  Haoran He, Chenjia Bai, Kang Xu, Zhuoran Yang, Weinan Zhang, Dong Wang, Bin Zhao, Xuelong Li.
  NeurIPS 2023

- Diffusion Models for Reinforcement Learning: A Survey
  Zhengbang Zhu, Hanye Zhao, **Haoran He**, Yichao Zhong, Shenyu Zhang, Haoquan Guo, Tingting Chen, Weinan Zhang.
  TPAMI 2024 (under review)

Experience

Intern at Seed — Jan. 2026 – Present

Role: Research Intern

Brief introduction: Research on RL post-training for unified multimodal foundation models.

- Be responsible for the RL post-training for SeedReam, Seedance, and SeeWorld, including both unified RL algorithms and distillation.

Intern at kuaishou Kling Team — Jan. 2025 – Jan. 2026

Role: Research Intern

Brief introduction: Research on RL post-training for video foundation models. Advised by Dr. Xintao Wang.

- We are investigating the RL efficacy for fine-tuning video models to obtain better performance.
- We are trying to address problems that are challenging for simply scaling up model size and data size.

Intern at Tencent AI Lab — Sep. 2024 – Dec. 2024

Role: Research Intern

Brief introduction: Research on world models for policy generalization. Advised by Dr. Zhongwen Xu.

- We built a generalist video-based world model as an environmental simulator.
- We proposed DWS, an effective model-based RL algorithm to learn well-performed policies across different environments.

Intern at Shanghai AI Lab — Oct. 2022 – Feb. 2024

Role: Research Scientist Intern

Brief introduction: Research on generalist reinforcement learning and robotics. Advised by Dr. Chenjia Bai.

- Aims to master quadruped locomotion task via a novel RL method. Adversarial methods for domain adaptation(e.g. sim2real) and skill discovery for performance boosting are proven to be effective.
- Learn a single policy that can tackle multiple tasks from the collected offline data.
- Proposed a new model named MTDIFF, which is trained on large-scale datasets for multi-task decision-making.
- Proposed a new model named VPDD, which is a generalist video-based multi-task agent pre-trained on large-scale human videos.

Professional Service

Reviewers:

- *Conferences:* ICML 2024-2026, IROS 2024, NeurIPS 2024-2026, AAAI 2025-2026, ICLR 2025-2026, ICRA 2025
- *Journals:* IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

Honors and Awards

1st Prize in school robocup competition	2019
Merit Student in SJTU	2020
Cosco shipping scholarship	2020
Honorable Mention in Mathematical Contest In Modeling	2020
C-class Excellent Scholarship	2021
Merit Student in SJTU	2022
B-class Excellent Scholarship	2022
Hong Kong PostGraduate Scholarship	2024

Skills

- Programming Languages: C == C++ == Python > Rust
- Platform: Linux, Mac, and Windows
- Tools: Asynchronous training and inference infrastructure; torch and tensorflow.
- Development: Machine learning, Reinforcement Learning, Data Analysis

Miscellaneous

- personal page: `https://tinnerhrhe.github.io`
- GitHub: `https://github.com/tinnerhrhe`
- Languages: English - Fluent, Mandarin - Native speaker
- Hobbies: Running, playing ping pong, and reading.