

Yazhe Niu

Phone: 86-15652586602 niuyazhe314@outlook.com

GitHub: <https://github.com/PaParaZz1> Zhihu: <https://www.zhihu.com/people/niu-ya-zhe-9>

EDUCATION

The Chinese University of Hong Kong

Hong Kong, China

Division of Electronic Engineering

MPhil-PhD Electronic Eng

08/2022 - present

• Supervisor: Professor WANG Xiaogang

Co-supervisor: Professor Li Hongsheng

Beihang University

Beijing, China

Department of Computer Science

Bachelor of Engineering at Beihang University

09/2016 - 06/2020

• Overall GPA: 3.8/4.0

Ranking: 8/223

RESEARCH INTERESTS

- Deep Reinforcement Learning (Large-scale RL/MCTS + RL/Diffusion + RL)
- Post-training and Applications of Foundation Models
- Distributed Deep Learning System and Infrastructure

SELECTED PUBLICATIONS & SUBMISSIONS [\[Google Scholar\]](#)

- **Yazhe Niu**, Yuan Pu, Zhenjie Yang, Xueyan Li, Tong Zhou, Jiyuan Ren, Shuai Hu, Hongsheng Li, Yu Liu “**LightZero: A Unified Benchmark for Monte Carlo Tree Search in General Sequential Decision Scenarios**”, accepted by NeurIPS 2023 (spotlight). [\[code\]](#)
- Yuan Pu*, **Yazhe Niu***, Jiyuan Ren, Zhenjie Yang, Hongsheng Li, Yu Liu “**UniZero: Generalized and Efficient Planning with Scalable Latent World Models**”, accepted by TMLR. [\[code\]](#)
- Yuan Pu*, **Yazhe Niu***, Jia Tang, Junyu Xiong, Shuai Hu, Hongsheng Li “**One Model for All Tasks: Leveraging Efficient World Models in Multi-Task Planning**”, Submitted to ICLR 2026. [\[code\]](#)
- “**SafeWork-R1: Coevolving Safety and Intelligence under the AI-45 Law**”, technical report (core leads) [\[web\]](#)
- Haolin Wang, Xueyan Li, **Yazhe Niu**[†], Shuai Hu, Hongsheng Li “**Empowering LLMs in Decision Games through Algorithmic Data Synthesis**”, accepted by ICLR 2025 SynthData Workshop (spotlight). [\[code\]](#)
- Rongkun Xue, Jinouwen Zhang, **Yazhe Niu**[†], Dazhong Shen, Bingqi Ma, Yu Liu, Jing Yang “**Pretrained Reversible Generation as Unsupervised Visual Representation Learning**”, accepted to ICCV 2025. [\[code\]](#)
- Jinouwen Zhang, Rongkun Xue, **Yazhe Niu**[†], Yun Chen, Hongsheng Li, Yu Liu “**Revisiting Generative Policies: A Simpler Reinforcement Learning Algorithmic Perspective**”, ICLR 2025 (review scores 8555). [\[code\]](#)

RESEARCH EXPERIENCES

Shanghai AI Lab

Shanghai, China

Part-time Researcher; Advisor: Yu Qiao & Chaochao Lu

07/2021 - present

- **Project: OpenDILab: Open-source Decision Intelligence (DI) Platform** [\[code\]](#) [\[zhihu\]](#)

➤ Overall responsibility for managing the research and operation, achieving a total of 28k+ GitHub stars

- Design and develop of a general DRL distributed system with stream and event mechanism, integrating over 90 environments and 120 algorithms for various RL research and application scenarios. [\[code\]](#)
- Develop essential algorithms and infrastructure (RFT/MCTS/Multi-agent) for foundation models.

SenseTime

Shanghai, China

Part-time Researcher; Advisor: Executive Director Yu Liu

06/2023 - 11/2024

- **Project: SenseMirage (Miaohua): Advanced AI Image Generation Platform** [\[web\]](#)
 - Accelerated a ~30B Diffusion Transformer foundation model training, achieving a ~40% enhancement through optimizations in data loading, operator speed, and computation/communication overlap.
 - Constructed a suite of toolkit to monitor the cluster with thousands of GPUs, ensuring training stability.

SenseTime

Shenzhen, China

Undergraduate Researcher; Advisor: PhD. Yu Liu

12/2019 - 06/2021

- **Project: SenseStar (DI-star): Large-Scale Deep Reinforcement Learning System for StarCraftII** [\[code\]](#)
 - Designed hierarchical reinforcement learning PPO Agent to beat hard level bot in StarCraftII.
 - Implemented AlphaStar distributed training and evaluation system with thousands of CPU/GPUs.
 - Improved its training performance by bags of tricks of the neural network architecture and optimization.

SenseTime

Beijing, China

Undergraduate Research Assistant; Advisor: PhD. Hongwei Qin

07/2018 - 11/2019

- **Project: Image Signal Processing (ISP) Pipeline with Deep Learning**
 - Designed and optimized the first AI + ISP SDK pipeline by hybrid programming with C++, Halide and PPL.
 - Designed a high-performance burst image align algorithm based on pyramid patch-matching and homography.

PROJECTS

Beihang University

Beijing, China

Teaching Assistant and Developer; Advisor: Prof. Ji Wu

06/2018 - 02/2019

- **Project: Transformation of Mandatory Oriented-Object Programming Course**
 - Redesigned 4 series of OOP coursework (16 weeks) and the double-blind code review mechanism.
 - Designed and implemented online judge platform and course forum (for ~300 students).

SKILLS

- **Programming Languages:** Proficient in Python, JavaScript, C & C++, CUDA, GDScript.
- **Deep Learning Framework:** Proficient in deep learning framework such as PyTorch, Triton, vLLM, SGLang.

HONORS&AWARDS

- Dean's Innovation Award in Sensetime Research (3 of 18 teams) *2020 - 2021*
- Merit Student in Beihang University (Top 5 % 223) *2016 - 2017, 2017 - 2018*

OTHERS

- **Personal Hobbies:** Football, Go, Real-time Strategy Games.