

# LUO YU

roythuly.github.io | github.com/Roythuly | Google Scholar

*"Scaling Intelligence through Reinforcement Learning – Bridging the gap between RL and Foundation Models."*

## RESEARCH INTERESTS

My research focuses on **Scaling Intelligence** by Reinforcement Learning, developing architectures, data recipes, and algorithms to endow models with complex reasoning and persistent memory.

**Scalable RL Algorithms:** Efficient and stable algorithms for decision-making in non-stationary environments.

**LLM Post-training:** Unlocking reasoning and memory capabilities in LLMs via Agentic RL post-training.

**Embodied AI:** Developing Vision-Language-Action models for robust robotic generalization and sim-to-real transfer.

## EXPERIENCE

### Senior Researcher

Jan. 2025 – Present

Foundation Model Group, 2012 Labs, Huawei Technologies Co., Ltd.

Beijing, China

- Lead RL post-training for openPangu models (1B–718B), building end-to-end pipelines and proposing novel algorithms ( $R^2$ VPO) that significantly improved Math/Code reasoning benchmarks (e.g., +10 on AIME).

### Research Intern

2022 – 2023

Institute for AI Industry Research (AIR), Tsinghua University

Beijing, China

- Conducted research on hierarchical RL for robotic manipulation, enhancing skill decomposition and sim-to-real efficiency.

## EDUCATION

### Ph.D. in Computer Science and Technology

Jun. 2019 – Dec. 2024

Tsinghua University – Advisor: Prof. Fuchun Sun

Beijing, China

- Thesis: Efficient Reinforcement Learning under Non-Stationary Data Distribution

### M.S. in Control Science and Engineering

Sep. 2016 – Jun. 2019

Beijing Institute of Technology – Advisor: Prof. Yuanqing Xia

Beijing, China

- Thesis: Aperiodic-Triggered Robust Model Predictive Control (Outstanding Master Thesis)

### B.S. in Engineering Mechanics

Sep. 2012 – Jun. 2016

Tsinghua University

Beijing, China

## SELECTED PUBLICATIONS

- [1] **Yu Luo**, Shuo Han, Yihan Hu, Dong Li, Jianye Hao. *Ratio-Variance Regularized Policy Optimization for Efficient LLM Post-training*. arXiv 2026.
- [2] Pinzheng Wang, Shuli Xu, Juntao Li, **Yu Luo**, Dong Li, Jianye Hao, Min Zhang. *Re<sup>2</sup>: Unlocking LLM Reasoning via Reinforcement Learning with Re-solving*. ICLR 2025.
- [3] Lei Lv, Yunfei Li, **Yu Luo**, Fuchun Sun, Tao Kong, Jiafeng Xu, Xiao Ma. *Flow-Based Policy for Online Reinforcement Learning*. NeurIPS 2025.
- [4] OpenPangu Team (**Yu Luo**). *Pangu Embedded: An Efficient Dual-system LLM Reasoner with Metacognition*. arXiv. 2025
- [5] OpenPangu Team (**Yu Luo**). *Pangu Ultra MoE: How to Train Your Big MoE on Ascend NPUs*. arXiv. 2025
- [6] **Yu Luo**, Tianying Ji, Fuchun Sun, Jianwei Zhang, Huazhe Xu, Xianyuan Zhan. *OMPO: A Unified Framework for Reinforcement Learning under Policy and Dynamic Shifts*. ICML 2024 (**Oral**).
- [7] **Yu Luo**, Tianying Ji, Fuchun Sun, Jianwei Zhang, Huazhe Xu, Xianyuan Zhan. *Offline-Boosted Actor-Critic: Adaptively Blending Optimal Historical Behaviors in Deep Off-Policy RL*. ICML 2024.
- [8] **Yu Luo**, Tianying Ji, Fuchun Sun, Huaping Liu, Jianwei Zhang, Mingxuan Jing, Wenbing Huang. *Goal-Conditioned Hierarchical Reinforcement Learning with High-Level Model Approximation*. IEEE TNNLS 2024 (**SCI Q1, Top**).
- [9] Tianying Ji, Yongyuan Liang, Yan Zeng, **Yu Luo**, et al. *ACE: Off-Policy Actor-Critic with Causality-Aware Entropy Regularization*. ICML 2024 (**Oral**).
- [10] Guowei Xu, Ruijie Zheng, Yongyuan Liang, Xiyao Wang, Zhecheng Yuan, Tianying Ji, **Yu Luo**, et al. *DRM: Mastering Visual Reinforcement Learning through Dormant Ratio Minimization*. ICLR 2024 (**Spotlight**).
- [11] Tianying Ji, **Yu Luo**, Fuchun Sun, Mingxuan Jing, Fengxiang He, Wenbing Huang. *When to Update Your Model: Constrained Model-based Reinforcement Learning*. NeurIPS 2022 (**Spotlight**).