Lixin XIU, PhD Student at the University of Tokyo

xiu-lixin106@g.ecc.u-tokyo.ac.jp riishin.github.io

Programming: Python, PyTorch
Research Interests: Multimodal Large Language Models (MLLMs), post-training, agent systems and information theory.

EDUCATION

The University of Tokyo, Tokyo, Japan — Oct 2024 — Sep 2027 (Expected)
PhD in Information Science and Technology
Supervisor: Prof. Hideki Nakayama

The University of Tokyo, Tokyo, Japan — Oct 2022 — Sep 2024
Master of Information Science and Technology, Department of Mathematical Informatics
Thesis: Enhancing neural network-based document retrieval with optimized indexing: An information-theoretic perspective

Xi'an Jiaotong University, Xi'an, China — Aug 2018 — Jul 2022
B.E. in Computer Science and Technology
Thesis: An automatic identification system for traffic signs based on deep neural betworks

PUBLICATIONS

Lixin Xiu, Xufang Luo, and Hideki Nakayama. A Comprehensive Information-Decomposition Analysis of Large Vision-Language Models. In *Proceedings of the 14th International Conference on Learning Representations* **(ICLR 2026)**.

Xin Du*, Lixin Xiu*, and Kumiko Tanaka-Ishii. Bottleneck-Minimal Indexing for Generative Document Retrieval. In *Proceedings of the 41st International Conference on Machine Learning*, PMLR 235:11888-11904, 2024. Selected as oral presentations **(ICML 2024 Oral)**.
***: Equal Contribution**.

RESEARCH EXPERIENCE

Intern — Dec 2025 — Mar 2026
Shanda AI Research Tokyo — Tokyo, Japan

- Built a post-training pipeline for role-play chatbots on open-source multilingual LLMs by generating, cleaning, and filtering high-quality synthetic dialogue data with frontier language model APIs.
- Fine-tuned smaller models via supervised fine-tuning (SFT) to distill persona consistency, adversarial robustness, and controllable conversational styles into locally deployable 8B/14B LLMs.

Research Assistant — Nov 2024 — Present
Beyond AI, the University of Tokyo and SoftBank — Tokyo, Japan

- Developed a generative retrieval-augmented generation (RAG) framework for a vanilla language model.
- Prepared a research manuscript and am currently revising it for submission to a top-tier venue.

SELECTED PROJECTS

Traffic Sign Recognition Contest, Summer Workshop 2021 — May 2021 — Jul 2021
School of Computing, National University of Singapore — Singapore

- Developed and optimized neural networks for traffic sign recognition, and delivered project presentations.
- Won third place and received an A grade.

HONORS & AWARDS

UTokyo SPRING GX (JPY 180,000/month stipend, JPY 360,000/year research grant) — Oct 2024 — Sep 2027
JASSO Scholarship — Oct 2022 — Mar 2023
Outstanding Graduate, Xi'an Jiaotong University — Jun 2022
Third Prize of the 12th Mathematics Competition of Chinese College Students — Dec 2020
Third-Class Scholarship for 2019-2020 Academic Year (University-level) — Sep 2020

SERVICE

Reviewer: ICLR 2026