Runyi Li (李 潤一)

lirunyi@g.ecc.u-tokyo.ac.jp · Runyi Li · Homepage

Education

Ph.D. Student, The University of Tokyo, Department of Information Science and Technology 2026.10 – 2029.09
Supervised by Prof. Yamasaki Toshihiko, research in multimodal LLMs, agentic AI and world models

M.S. Peking University, School of Electronic and Computer Engineering 2023.09 – 2026.06
Supervised by Prof. Jian Zhang, research in multimodal learning, generative models and trustworthy AI

B.S. Sichuan University, Department of Computer Science 2019.09 – 2023.06
Honor Graduation, supervised by Prof. Lei Zhang, research in medical image analysis

Selected Publications

Multimodal LLMs, Agentic AI & World Models

- IROHA: Long-Horizon Visual Reasoning with an Agentic Harness. **First** author. *ongoing project*.
- Dual Exam: Controlled Evaluation of Multimodal Agents and World Models. Sixth author. *Tech report at XGen Lab, 2026*.
- Bridging Video Understanding and Generation in a Unified Framework. **Co-first** author. *Tech report at TikTok, 2026*.
- Vision-Language-Action or World-Action? Just Action Models. Third author. *ICLR'27 under review*.
- FakeShield: Explainable Image Forgery Detection and Localization via Multi-modal Large Language Models. Third author. *ICLR'25*.
- EditHero: A Benchmark for Long-Horizon Part-Level 3D Editing and Vibe Modeling. Fourth author. *Tech report at Alaya Studio, 2026*.
- WorldRover: A Scalable Synthetic Video Data Engine for World Exploration with Rich Annotations. Third author. *Tech report at Alaya Studio, 2026*.
- Agentic World Modeling: Foundations, Capabilities, Laws, and Beyond. Contributor. *Survey, TMLR under review*.
- PoseWeaver: LLM-Driven Agentic Control of 2D Human Pose in Text-Image Generation. Fourth author. *AAAI27' under review*.

Generative Modeling & Trustworthy AI

- OmniSSR: Zero-shot Omnidirectional Image Super-Resolution using Stable Diffusion Model. **Co-first** author. *ECCV'24* ***Oral***.
- Mirai: Autoregressive Visual Generation Needs Foresight. Fourth author. *CVPR'26*.
- Video-Mirai: Autoregressive Video Diffusion Models Need Foresight. Third author. *NeurIPS'26*.
- LAFR: Efficient Diffusion-based Blind Face Restoration via Latent Codebook Alignment Adapter. **First** author. *BMVC'26*.
- EditGuard: Versatile Image Watermarking for Tamper Localization and Copyright Protection. Second author. *CVPR'24*.
- OmniGuard: Hybrid Manipulation Localization via Augmented Versatile Deep Image Watermarking. Fourth author. *CVPR'25*.
- Rooting Adaptive Watermarks for 3D Gaussian Generation Model. **First** author. *Springer Machine Intelligence Research*.

Experience

Alaya Studio *Research Intern* 2026.07 – Present
Mentored by Kaipeng Zhang and Zhixiang Wang; agentic AI, world models, 3D vision

XGen Lab *Research Intern* 2026.06
Mentored by Meng Chu; World models and agentic AI

ByteDance TikTok Group *Research Intern* 2025.06 – 2025.12
Mentored by Yuqi Wang; Unified video understanding and generation, MLLMs, VLMs

Universität Würzburg *Visiting Student* 2024.10 – 2026.04
Supervised by Prof. Radu Timofte; Diffusion-based generative AI, real-world image restoration and super-resolution

RabbitPre Intelligent Technology *Research Intern* 2024.04 – 2025.03
Trustworthy generative AI, privacy protection and personalized generation

Selected Honors and Awards

The UTokyo Fellowship Scholarship (Top 1%) 2026
Outstanding Graduate of Peking University (Top 3%) 2026
Outstanding Graduate of Sichuan Province (Top 1%) 2023
National Scholarship of China (Top 1%) 2022