Ruize Xu

Homepage: https://rick-xu315.github.io

Email : ruize.xu.gr@dartmouth.edu

Mobile : +1-917-407-9605

EDUCATION EXPERIENCE

- **Dartmouth College, NH** — Sep. 2025 - May 2030 (expected)
  PhD in Computer Science
  Advisor: Nikhil Singh

- **Columbia University, NY** — Jan. 2024 - May 2025
  MS in Computer Science-Thesis Track
  GPA: 4.08/4.0,   Advisor: Zhou Yu

- **Renmin University, China** (Ranked 20th globally on CSRankings 2025) — Sep. 2019 - Jun. 2023
  BS in Data Science

- **University of California, Davis, CA** — Jan. 2022 - Apr. 2022
  Exchange Student

RESEARCH INTEREST

Post-training Dynamics of Large Language Models:
  RL for Specialization, Policy Optimization, World Model Learning
Mechanistic and Representation-level Interpretability:
  Geometric Analysis of LLM Representations and Structure Emergence, Causal and Data-centric Interpretation

RESEARCH EXPERIENCE

- **Columbia University** — Feb 2024 - Present
  - **Mentor**: Junfeng Yang
  - **Project**: Adversarial analysis of diffusion-based image detectors via internal guidance (CVPR 2024 Workshop).
  - **Mentor**: Zhou Yu
  - **Project**: Built resume-job matching framework with hard-negative mining and synthetic data generation(ACL 2025).

- **Alibaba Security Group** — Jul 2023 - Dec 2023
  - **Mentor**: Pengda Qin
  - **Projects**: Built supervised fine-tuning data for 7B/14B LVLMs for multi-image QA (MME 1400+). Expanded LVLMs with MoE to improve unseen-task generalization (+50%) while maintaining prior performance (95%).

- **Baidu Vision** — Nov 2022 - Jun 2023
  - **Mentor**: Yumeng Zhang
  - **Projects**: Improved NeRF for dynamic unbounded autonomous-driving scenes via tri-plane decomposition on NuScenes.

- **Renmin University of China** (Undergraduate Research) — Jun 2021 - Jun 2024
  - **Advisor**: Di Hu
  - **Projects**: Studied modality imbalance in multi-modal learning; proposed modality-wise normalization to improve weak uni-modalities in fine-grained audio-visual tasks (ICASSP 2023).

PUBLICATION LIST

1. **Reinforcement World Model Learning for LLM-based Agents.**
   Xiao Yu, Baolin Peng, **Ruize Xu**, Yelong Shen, Pengcheng He, Suman Nath, Nikhil Singh, Jiangfeng Gao, Zhou Yu.
   Preprint.(Paper)

2. **Dyna-Think: Synergizing Reasoning, Acting, and World Model Simulation in AI Agents.**
   Xiao Yu, Baolin Peng, **Ruize Xu**, Michel Galley, Hao Cheng, Suman Nath, Jianfeng Gao, Zhou Yu.
   Preprint. (Paper)

3. **AI Agents for Web Testing: A Case Study in the Wild.**
   Naimeng Ye*, Xiao Yu*, **Ruize Xu***, Tianyi Peng, Zhou Yu.
   NeurIPS 2025, LAW Workshop. (Paper)

4. **ConFit v2: Improving Resume-Job Matching using Hypothetical Resume Embedding and Runner-Up Hard-Negative Mining.**
   Xiao Yu*, **Ruize Xu***, Chengyuan Xue*, Jinzhong Zhang, Xu Ma, Zhou Yu.
   Findings of ACL 2025. (Paper)

5. **From Detection to Deception: Are AI-Generated Image Detectors Adversarially Robust?**
Yun-yun Tsai, **Ruize Xu**, Chengzhi Mao, Junfeng Yang.
CVPR 2024, Responsible Generative AI Workshop. (Paper)

6. **Comclip: Training-free Compositional Image and Text Matching.**
Kenan Jiang, Xuehai He, **Ruize Xu**, Xin Eric Wang.
NAACL 2024. (Paper)

7. **MMcosine: Multi-modal Cosine Loss Towards Balanced Audio-Visual Fine-grained Learning.**
Ruize Xu, Ruoxuan Feng, Shi-Xiong Zhang, Di Hu.
ICASSP 2023. (Paper and Code)

Honors and Awards

- Tinker Research Grant **$5,000**, *Thinking Machines Lab.* 2026

- Renmin University of China (RUC): National 1st Prize, Chinese Mathematics Competitions; Mathematical Contest in Modeling / Interdisciplinary Contest in Modeling. Meritorious Winner (Top 7%) 1st Prize Scholarship for Academic Excellence (Top 5%); President Scholarship for Exchange Students; Undergraduate Research Fund; Dean's MingDe Data Science Talents Nomination (17/321); Academic Excellence Award (Top 3% GPA) 2020–2022

Service

- **Reviewer**: ICASSP 2024, Neurips LAW Workshop 2025,ICLR 2026 World Models Workshop, ICML 2026

- Table Tennis School Team: RUC&Columbia