

# YUNFEI XIE

Email: xieyunfei01@gmail.com · Tel: +1-831-295-9285 · Homepage: <https://yunfeixie233.github.io/>

## RESEARCH INTERESTS

---

Reinforcement Learning, Vision-Language Models

## EDUCATION

---

### Rice University

09/2025-

Ph.D. in Computer Science, Supervisor: Prof. Chen Wei

### Huazhong University of Science and Technology

09/2021-06/2025

Bachelor of Engineering in Artificial Intelligence

GPA: 87.1/100

Scholarship & Honors: Science and Technology Scholarship (top 2% in school)

## PUBLICATION LIST (GOOGLE SCHOLAR)

---

### Reinforcement Learning

1. Preprint [link]: **Yunfei Xie**, Yinsong Ma, Shiyi Lan, Alan Yuille, Junfei Xiao, Chen Wei. “Play to Generalize: Learning to Reason Through Game Play”

### Vision Language Models / Large Language Models

1. ICLR 2025 [link]: **Yunfei Xie**, Ce Zhou, Lang Gao, Juncheng Wu, Xianhang Li, Hong-Yu Zhou, Sheng Liu, Lei Xing, James Zou, Cihang Xie, Yuyin Zhou. “MedTrinity-25M: A Large-scale Multimodal Dataset with Multigranular Annotations for Medicine” (received 350+ stars on GitHub)

2. Preprint [link]: **Yunfei Xie**, Kevin Wang, Bobby Cheng, Jianzhu Yao, Zhizhou Sha, Alexander Duffy, Yihan Xi, Chen Wei, Pramod Viswanath, Zhangyang Wang. “COPER: Agentic Context Significantly Improves and Stabilizes LLM in Multi-Player Game”

3. Preprint [link]: **Yunfei Xie**, Juncheng Wu, Haoqin Tu, Siwei Yang, Bingchen Zhao, Yongshuo Zong, Qiao Jin, Cihang Xie, Yuyin Zhou. “A Preliminary Study of o1 in Medicine: Are We Closer to an AI Doctor?”

### Computer Vision

1. ECCV 2024 [link]: **Yunfei Xie**, Cihang Xie, Alan Yuille, and Jieru Mei. “From Pixels to Objects: A Hierarchical Approach for Part and Object Segmentation Using Local and Global Aggregation”

2. Preprint [link]: **Yunfei Xie**, Alan Yuille, Cihang Xie, Yuyin Zhou, Jieru Mei. “Few-Shot Medical Image Segmentation via Supervoxel Transformer”

3. IEEE ISBI 2024 [link]: **Yunfei Xie**, Ce Zhou, Jieru Mei, Xianhang Li, Cihang Xie, and Yuyin Zhou. “Brain Tumor Segmentation Through Supervoxel Transformer”

4. Preprint [link]: Jiawei Mao, Xiaoke Huang, **Yunfei Xie**, Yuanqi Chang, Mude Hui, Yuyin Zhou. “Story-Adapter: A Training-free Iterative Framework for Long Story Visualization” (received 900+ stars on GitHub)

5. Preprint [link]: **Yunfei Xie**, Yuxuan Cheng, Juncheng Wu, Haoyu Zhang, Yuyin Zhou, Shoudong Han. “SCING: Towards More Efficient and Robust Person Re-Identification through Selective Cross-modal Prompt Tuning”

## ONGOING PROJECTS

---

### Reinforcement Learning in ARC-AGI-3 Benchmark [link]

- Exploring the development of an online reinforcement learning infrastructure and algorithm for ARC-AGI-3, a challenging long-horizon decision-making task designed to evaluate general intelligence
- Officially funded by the ARC-AGI organization with a \$10,000 research grant

## RESEARCH EXPERIENCE

---

### Research Intern, Rice University

02/2025-06/2025

Advised by Prof. Chen Wei

- Developed a novel approach to generalization in reinforcement learning with verifiable reward (RLVR): proposed a post-training paradigm, Visual Game Learning, where MLLMs developed out-of-domain general-

ization capabilities for multimodal reasoning through playing arcade-like games

## **Research Intern, VLAA lab at University of California, Santa Cruz**

02/2024-02/2025

Advised by Prof. Yuyin Zhou and Prof. Cihang Xie

- Developed an automated data synthesis pipeline for vision language models: created multi-granular visual and textual annotations for medical images, leveraging expert grounding models for ROI annotation, Retrieval-Augmented Generation (RAG) for linking medical knowledge to specific images, and fine-tuned medical-specific MLLMs for caption generation
- Explored consistency in image generation: proposed a training-free and efficient framework for generating high-quality, fine-grained long stories using an iterative paradigm that progressively optimizes image generation by repeatedly incorporating text and global constraints
- Constructed a comprehensive medical benchmark for holistic evaluation of o1's medical capabilities

## **Research Intern, CCVL lab at Johns Hopkins University**

08/2023-02/2024

Advised by Prof. Alan Yuille

- Investigated representation learning in vision transformers: integrated hierarchical representations of superpixels and group tokens into vision transformers to capture both local detail for parts and global context for objects in fine-grained segmentation
- Extended representation learning to 3D medical imaging: adapted superpixels to 3D supervoxels and integrated them with 3D vision transformers for medical segmentation

## **ACADEMIC SERVICE**

---

Reviewer:

2026: CVPR, ICLR

2025: ICCV, ICLR, CVPR, ICML, TMM, NeurIPS, TPAMI

2024: CVPR, ICML, IEEE ISBI

## **PROFESSIONAL SKILLS**

---

Programming Languages: Python, Shell,  $\LaTeX$ , HTML

Programming Tools: Git, PyTorch, Docker, Slurm, Linux Ops, Vim