

Ruiqi Xian

<https://ricky-xian.github.io>

rxian@umd.edu | 9122 Bridgewater Street, College Park, MD | (614) 648-7971

Currently looking for summer internships from **May 22, 2025** to **Sep 1, 2025**

EDUCATION

- University of Maryland**, College Park, Maryland, United States August 2021 - now
Doctor of Philosophy in Electrical and Computer Engineering (with interests in Robotics and Computer Vision)
Advisor: Dr. Dinesh Manocha
- The Ohio State University**, Columbus, Ohio, United States August 2018 - May 2020
Master of Science in Electrical and Computer Engineering
Advisor: Dr. Lisa Fiorentini
- China University of Mining and Technology**, Xuzhou, Jiangsu, China September 2013 - June 2017
Bachelor of Engineering in Information Engineering

RESEARCH

Self-supervised Learning

Video Pretraining

- Developed a self-supervised learning method for video pre-training, leveraging object-level information to reduce memory requirements, accelerate pre-training speed for video foundation models via masked autoencoding, and improve model performance in downstream video recognition tasks.

Multi-Camera Multi-Object Tracking

- Developed a self-supervised multi-camera multi-object tracking framework leveraging single-view distillation and cross-view reconstruction to disentangle view-agnostic/specific features, enabling robust over-time and cross-view tracking without the need for camera calibration or manual annotations.

Cross-Modal Global Navigation

- Developed robot localization methods using LiDAR and satellite maps, addressing scale variations and representation gaps between modalities for global navigation in complex outdoor environments.

Large Vision-Language Models (LVLMs)

Hallucination

- Conducted comprehensive analysis of state-of-the-art LVLMs (e.g., GPT-4V, LLaVA-1.5) to identify and understand limitations in areas like language hallucination and visual illusions and developed a novel diagnostic suite for dissecting the complexities and failure modes of LVLMs.

LVLMs with Robotics

- Investigated the safety and security vulnerabilities of integrating LVLMs into robotic systems and designed adversarial attacks to expose potential weaknesses in current state-of-the-art LVLM-based robotic systems.

LVLMs for Video Understanding

- Integrated Large Language Models (LLMs) as "prompt experts" within the training pipeline for video recognition models and utilized learnable prompts to effectively guide the model's attention toward action-related descriptions within video inputs, significantly improving recognition accuracy.

UAV Video Understanding

Information Theory Guided UAV Action Recognition

- Developed information theory-based methods for robust feature alignment and adaptive frame selection in UAV videos to mitigate the impact of dynamic backgrounds and UAV motion while efficiently capturing the most informative frames for action recognition.

Edge Computing

- Developed real-time action recognition models specifically tailored for resource-constrained UAV

platforms. The algorithms were integrated with ROS2 for hardware acceleration, enabling real-time deployment on low-power UAV systems with Qualcomm CPU and GPUs.

EXPERIENCE

- University of Maryland**, College Park, Maryland, United States May 2022 - now
Graduate Research Assistant – Focus on UAV perception algorithms, supervised by Dr. Dinesh Manocha
- NEC Laboratory, America**, Princeton, New Jersey, United States May 2024 – August 2024
Research Intern – Focus on Multi-view multi-object tracking, supervised by Deep Patel, Iain Melvin
- University of Maryland**, College Park, Maryland, United States January 2022 - May 2022
Graduate Teaching Assistant – Teaching ENPM 690 Robot learning
- Ohio State University**, Columbus, Ohio, United States January 2020 - May 2020
Research Assistant – Developed Honda Inline Measurement System, supervised by Dr. Lisa Fiorentini

Selected Publications

- [10] **Ruiqi Xian**, Iain Melvin, Deep Patel, Sanjoy Kundu, Martin Renqiang Min, Dinesh Manocha. “CALIBFREE: Self-Supervised Feature Disentanglement for Calibration-Free Multi-Camera Multi-Object Tracking”. (Submitted to CVPR 2025)
- [9] **Ruiqi Xian**, Xiyang Wu, Tianrui Guan, Xijun Wang, Boqing Gong, Dinesh Manocha. “SOAR: Self-supervision Optimized UAV Action Recognition with Efficient Object-Aware Pretraining”. (Submitted to ICRA 2025)
- [8] **Ruiqi Xian***, Tianrui Guan*, Xijun Wang, Xiyang Wu, Mohamed Elnoor, Daeun Song, Dinesh Manocha. “AGL-NET: Aerial-Ground Cross-Modal Global Localization with Varying Scales”. (IROS2024) [Paper](#)
- [7] Xiyang Wu, **Ruiqi Xian**, Tianrui Guan, Jing Liang, Souradip Chakraborty, Fuxiao Liu, Brian Sadler, Dinesh Manocha, Amrit Singh Bedi. “On the Safety Concerns of Deploying LLMs/VLMs in Robotics: Highlighting the Risks and Vulnerabilities”. (CVPR 2024 VLADR workshop) [Project Page](#)
- [6] **Ruiqi Xian***, Xijun Wang*, Tianrui Guan, Dinesh Manocha. “SCP: Soft Conditional Prompt Learning for Aerial Video Action Recognition”. (IROS2024) [Paper](#)
- [5] Tianrui Guan*, Fuxiao Liu*, Xiyang Wu, **Ruiqi Xian**, Zongxia Li, Xiaoyu Liu, Xijun Wang, Lichang Chen, Furong Huang, Yaser Yacoob, Dinesh Manocha, Tianyi Zhou. “HALLUSIONBENCH: An Advanced Diagnostic Suite for Entangled Language Hallucination & Visual Illusion in Large Vision-Language Models”. (CVPR 2024) [Project Page](#)
- [4] **Ruiqi Xian**, Bryan I Vogel, Celso M De Melo, Andre V Harrison, Dinesh Manocha. “Real-time human action recognition from aerial videos using autozoom and synthetic data” (DCS 2024)
- [3] **Ruiqi Xian**, Xijun Wang, Divya Kothandaraman, Dinesh Manocha. “PMI Sampler: Patch Similarity Guided Frame Selection for Aerial Recognition”. (WACV 2024) [Paper](#)
- [2] **Ruiqi Xian***, Xijun Wang*, Dinesh Manocha. “MITFAS: Mutual Information based Temporal Feature Alignment and Sampling for Aerial Action Recognition”. (WACV 2024) [Paper](#)
- [1] **Ruiqi Xian***, Xijun Wang*, Dinesh Manocha. “AZTR: Aerial Video Action Recognition with Auto Zoom and Temporal Reasoning”. (ICRA 2023) [Paper](#)