

Ruiqi Xian

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

Currently looking for summer interns, available from **May 20, 2024** to **August 26, 2024**

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

SKILLS

Language: Python, Pytorch, Tensorflow, Matlab, C/C++, Java, JavaScript, SQL

Devices: Qualcomm Robotics RB5 platform, ESP8266, MCS-51, Arduino, STM32, FPGA, Raspberry Pi

PROJECTS

- Learning Visual Features from Aerial Imagery Without Supervision** August 2023 – now
- Collected and curated data from public aerial datasets and web sources via unsupervised data retrieval.
 - Innovated a Self-Supervised Learning (SSL) algorithm, optimized for UAV-centric applications. This algorithm incorporates a unique self-distillation process with box-level contrastive views, enhancing the networks' ability to concentrate on human subjects within aerial imagery.
 - Advanced the self-distillation loss function by incorporating masked video modeling, boosting the student network's proficiency in learning visual features from video streams.
- Advanced Diagnostic Benchmark for Language Hallucination and Visual Illusions in Large Vision-Language Models** October 2023 – November 2023
- Developed HALLUSIONBENCH, an innovative diagnostic suite aimed at dissecting and understanding the complexities and failure modes in large visual-language models (LVLMs).
 - Conducted in-depth analyses on the latest iterations of LVLMs, such as GPT-4V and LLaVA-1.5, elucidating their limitations in challenges such as language hallucination and visual illusion.
- Prompt Learning for Action Recognition** May 2023 – August 2023
- Utilized the advantages of prompt learning to direct the models' attention toward action-related descriptions and instructions within input videos.
 - Integrated Large Language Models (LLMs), such as SAM (Segment Anything Model), into the training pipeline as prompt experts to enhance the optimization of the recognition model.
 - Introduced a learnable prompt methodology to guide the model's predictions while simultaneously acquiring input-invariant knowledge from prompt experts and input-specific prompt knowledge.
- Patch Similarity Guided Frame Selection for Aerial Action Recognition** January 2023 – May 2023
- Pioneered a resilient frame selection strategy, addressing aerial data challenges such as resolution variance, multi-scale content, and camera movement.
 - Introduced Patch Mutual Information (PMI) score to assess motion information between frames, enabling adaptive frame selection for comprehensive video representation.

Mutual Information-based Temporal Feature Alignment and Sampling for Aerial Action Recognition

October 2022 – January 2023

- Introduced a robust feature alignment technique leveraging mutual information, which concurrently acquires and aligns human-related features between frames. This enables recognition models to prioritize human actions while mitigating the impact of dynamic background and UAV ego-motions.
- Presented a frame sampling method rooted in information theory, enabling the selective sampling of discriminative frames related to human actions from UAV videos to improve recognition accuracy.

Aerial Action Recognition using Autozoom and Temporal Reasoning

March 2022 – October 2022

- Innovated an auto zoom algorithm for efficient autofocus, cropping, and scaling, augmenting the extraction of human action details against dominant backgrounds in UAV video streams.
- Introduced a temporal reasoning module with convolutions and attention mechanisms, optimizing recognition performance across low-power UAV platforms and high-end desktop systems.
- Deployed TensorFlow Lite methods on edge devices like Qualcomm Robotics RB5 Platform, integrating with ROS2 for hardware acceleration on Snapdragon CPUs and Adreno GPUs.
- Conducted comprehensive evaluations of machine learning APIs including Qualcomm's Snapdragon Neural Processing Engine (SNPE), Neural Network API (NNAPI), and Alibaba's Mobile Neural Network (MNN) on edge devices.

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

University of Maryland, College Park, Maryland, United States

January 2022 - May 2022

Graduate Teaching Assistant – Teaching ENPM 690 Robot learning

The Ohio State University, Columbus, Ohio, United States

January 2020 - May 2020

Research Assistant – Developed Honda Inline Measurement System, supervised by Dr. Lisa Fiorentini

Publications

HALLUSIONBENCH: An Advanced Diagnostic Suite for Entangled Language Hallucination & Visual Illusion in Large Vision-Language Models (Submission to CVPR2024)

Tianrui Guan, Fuxiao Liu*, Xiyang Wu, Ruiqi Xian, Zongxia Li, Xiaoyu Liu, Xijun Wang, Lichang Chen, Furong Huang, Yaser Yacoub, Dinesh Manocha, Tianyi Zhou*

PLAR: Prompt Learning for Action Recognition

(Submission to ICRA2024 <https://arxiv.org/abs/2305.12437>)

Ruiqi Xian, Xijun Wang*, Tianrui Guan, Dinesh Manocha*

PMI Sampler: Patch Similarity Guided Frame Selection for Aerial Recognition

(WACV2024 <https://arxiv.org/abs/2304.06866>)

Ruiqi Xian, Xijun Wang, Divya Kothandaraman, Dinesh Manocha

MITFAS: Mutual Information based Temporal Feature Alignment and Sampling for Aerial Action Recognition

(WACV2024 <https://arxiv.org/abs/2303.02575>)

Ruiqi Xian, Xijun Wang*, Dinesh Manocha*

AZTR: Aerial Video Action Recognition with Auto Zoom and Temporal Reasoning

(ICRA2023 <https://arxiv.org/abs/2303.01589>)

Ruiqi Xian, Xijun Wang*, Dinesh Manocha*