

Jianyu Lai

Mobile: +86-159-6058-5656

Wechat: alexlai2860

Email: jlai218@connect.hkust-gz.edu.cn

Github: alexlai2860

EDUCATION

-
- The Hong Kong University of Science and Technology (Guangzhou)** Guangzhou, China
• **School:** College of Future Technology **Major:** Data Science and Analytics September 2024 -
Degree: Master of Philosophy in Data Science and Analytics; Supervised by Prof. Lei Zhu
Research Proposal: Image Restoration, Image Generation, MLLM.
 - South China University of Technology** Guangzhou, China
• **School:** Automation Science and Engineering **Major:** Automation September 2020 - June 2024
Degree: Bachelor of Engineering in Automation
Key Courses: Principles of Microcomputers (4/4), Motor and Drive Foundation (4/4), Probability and Mathematical Statistics (3.7/4), Signal Analysis and Processing (3.7/4), Computer Network and Communication (3.7/4). **GPA:** 3.47/4

RESEARCH INTEREST

-
- Advanced image generation algorithms, such as diffusion models and autoregressive generative models, aim to produce highly realistic images.
 - Developing robust and generalizable solutions for real-world image restoration, particularly in adverse conditions.
 - Exploring multimodal large language models (MLLM) for image processing.

PUBLICATIONS

-
- **In submission: ACD-Desmoke: Adaptive Constrained Diffusion for Generalizable Zero-Shot Surgical Smoke Removal in Laparoscopy:** [Jianyu Lai](#)[†], Sixiang Chen[†], Yunlong Lin, Tian Ye, Song Fei, Heming Li, Zhaohu Xing, Lei Zhu.
Summary: Laparoscopic surgery is often hindered by surgical smoke, reducing visibility and precision. Existing desmoking methods are limited by reliance on costly synthetic data or complex training architectures. This paper proposes ACD-Desmoke, a zero-shot, training-free framework leveraging pre-trained diffusion models with adaptive constrained diffusion. By incorporating adaptive weight allocation and timestep control, it achieves robust, high-quality smoke removal across diverse laparoscopic datasets, offering a practical and scalable solution.
 - **In submission: GenHaze: Pioneering Controllable One-Step Realistic Haze Generation for Real-World Dehazing:** Sixiang Chen, Tian Ye, Yunlong Lin, Yijun Yang, Yun Liu, Yeying Jin, Song Fei, [Jianyu Lai](#), Zhaohu Xing, Fugee Tsung, Lei Zhu.
Summary: Real-world image dehazing is vital for visual quality, but existing methods struggle with complexity and controllability. We propose GenHaze, a framework using a pre-trained latent diffusion model to generate high-quality, controllable hazy images in one step. GenHaze enhances existing models with minimal fine-tuning, significantly improving performance across seven metrics. This work demonstrates the potential of realistic degradation generation, offering an efficient alternative to traditional approaches.
 - **CVPR'25: SnowMaster: Comprehensive Real-world Image Desnowing via MLLM with Multi-Model Feedback Optimization:** [Jianyu Lai](#)[†], Sixiang Chen[†], Yunlong Lin, Tian Ye, Yun Liu, Song Fei, Zhaohu Xing, Hongtao Wu, Weiming Wang, Lei Zhu. [[PDF](#), [Code](#)]
Summary: Motivated by the limited generalization of current desnowing models due to their reliance on synthetic data and the lack of large-scale real-world snowfall datasets, this paper introduces RealSnow10K, a large-scale annotated high-quality real-world snowfall dataset, and a preference dataset with 36,000 expert-ranked pairs. Furthermore, we propose SnowMaster desnowing framework which leverages MMPO-enhanced Q-instruct for semi-supervised training to improve real-world snow removal performance.

PROJECTS

-
- **2023.9 - 2024.5: Gait Face Fusion Recognition System Based on Multi Perspective Collaboration:** This is a College Students Innovation and Entrepreneurship Training Program (National) Project which develops a multi person gait recognition system with a visual operating interface to achieve long-distance, contactless, and non-cooperative multi person identity recognition. This project overcomes challenges in long-distance blur and close-range incomplete body movement by combining facial and gait features. Specifically, I lead the design and implementation of the occlusion gait restoration module, which significantly improved recognition accuracy.
 - **2022.11 - 2023.6: Auto-Focus-Assistant:** This project is an independent R&D initiative that leverages depth cameras and advanced computer vision algorithms to deliver a highly efficient auto-focus system for various camera setups. The system processes both spatial depth and RGB data from the depth camera, deploying efficient object detection models (**YolofastestV2**) to compute target-to-camera distances with high accuracy. These distances are converted into control signals for a **follow-focus motor (Nucleus-N)**, enabling precise adjustments to the focus ring. Iteratively optimized, the system supports multiple focusing modes, including manual, fully automatic, and continuous tracking via **SCRFD face detection** and object tracking algorithms. It also integrates **eye-tracking focus** and **radar waveform-based focus**, making it one of the most intelligent focusing systems currently available. [[Code](#), [Video](#)]

- **2022.9 - 2023.1: Original Aspiration: RoboMaster Short Film Project:** This is a team-driven filming project in which I serve as the **team leader**, overseeing a comprehensive range of responsibilities including **screenwriting, directing, cinematography, and post-production editing and color grading**. The short film narrates the journey of preparing for a robotics competition, which garnered nearly 4,000 views on a video-sharing platform and received commendations and shares from the event organizers. [Video]
- **2021.9 - 2022.7: National College Student Robot Competition (RoboMaster University Championship):** This is a robotics competition in which different universities participate in teams. Each competing team is required to develop multiple robots with different functionalities to engage in the contest. I am a member of the **aerial robotics (unmanned aerial vehicle - UAV) team**, primarily responsible for **designing and developing the target detection system for the UAV**, and joint debugging and flight testing with the electronic control group. Additionally, I maintain the visual code repository for the robotics laboratory, known as the Scut Robot Vision Library (SRVL). Concurrently, I serve as one of the UAV pilots within my team. [Code, Video]

SKILLS SUMMARY

- **Languages:** Python, C++.
- **Frameworks:** Opencv, Pytorch, Numpy, Onnxruntime.
- **Soft Skills:** Leadership, Science Writing, Engineering Experience, Independent Thinking, Critical Thinking.

PATENTS AND AWARDS

- **2023.6: Patent, Inventor & Patentee:** Automatic focusing system based on depth vision and target detection, CN219268990U
- **2022.8: National First Prize, Technical team member:** National College Student Robot Competition (RoboMaster University Championship)