

# Zhaoyang Li

 6083205509

 zhl165@ucsd.edu

 LinkedIn

 Github

 Homepage

## Education

### MS University of California San Diego

Sep 2023 - Jun 2025

- Electrical and Computer Engineering (Intelligent Systems, Robotics & Control)

### BS University of Wisconsin-Madison,

Jan 2021 - May 2023

- **Double Major** in Computer Science & Mathematics

## Research Interests

My research focuses on creating agents that can perceive, reason, and act in the real world. My work integrates multimodal models with robot learning to turn language goals into actionable sub-tasks and train robust, sample-efficient policies. My goal is to build interpretable, safe robots that generalize beyond curated demonstrations.

## Publications & Preprints

\* indicates equal contributions.

- 1. **Zhaoyang Li\***, Zhan Ling\*, Yuchen Zahou, Litian Gong, Erdem Biyik, Hao Su, "ORIC: Benchmarking Object Recognition under Contextual Incongruity in Large Vision-Language Models", Accepted at ICCV 2025 Workshop on MMRAgI; extended version in submission. [\[Paper\]](#) [\[Code\]](#)
- 2. **Zhaoyang Li**, Sushaanth Srinivasan, Ninad Ekbote, Pengtao Xie, "A Multi-modal Large Language Model for Predicting Mechanisms of Drug Interactions", Under Review, 2025
- 3. Tongzhou Mu\*, **Zhaoyang Li\***, Stanisław Wiktor Strzelecki\*, Xiu Yuan, Yunchao Yao, Litian Liang, Aditya Gulati, Hao Su, "When Should We Prefer State-to-Visual DAGger Over Visual Reinforcement Learning?", published in Association for the Advancement of Artificial Intelligence (AAAI), 2025 [\[Paper\]](#) [\[Code\]](#)
- 4. Yifei Zhang, Yusen Jiao, Jiayi Chen, **Zhaoyang Li**, Huaxiu Yao, Jieyu Zhang, Frederic Sala, "Just Select Twice: Leveraging Low-Quality Data to Improve Data Selection," ATTRIB Workshop at NeurIPS 2024; extended version in submission. [\[Paper\]](#)

## Research Experience & Project

### Preference Guidance for Diffusion Policy via Energy-Based Model (Led project)

Advisor: Prof. Erdem Biyik, USC

Aug 2025 – present  
Los Angeles, CA

- Developed an Energy-Based Model to align action trajectories with human preferences and applied it as exact energy guidance for fine-tuning diffusion policies.
- Running large-scale experiments across RoboMimic, OpenAI Gym (HalfCheetah, Hopper, Walker2d), PushT, and real-robot evaluation on WidowX.
- Project is ongoing, with a paper in preparation targeting Robotics: Science and Systems (RSS).

### Vision-Language Model Benchmarking (Led project)

Advisor: Prof. Hao Su, UC-San Diego

Sep 2024 - July 2025  
San Diego, CA

- Developed ORIC, a framework that generates context-incongruous object-recognition data for training and evaluation, revealing uncertainty-driven LLM failures.
- Designed dual LLM- and CLIP-guided sampling to construct challenging ORIC-Bench and ORIC-style datasets.
- Improved Qwen3-VL-8B-Instruct via Visual-RFT on ORIC-style data, enhancing performance across benchmarks and aligning predictions with human reasoning.

### Multi-modal Language Model for Drug Interaction Prediction (Led project)

Advisor: Prof. Pengtao Xie; UC-San Diego

Dec 2024 - May 2025  
San Diego, CA

- Fine-tuned a multi-modal LLM with SMILES inputs to predict drug interaction status, degree, and mechanisms, integrating chemical informatics and NLP.
- Achieved strong performance: METEOR 0.42, BLEU-1 0.25, semantic similarity 0.57; outperforming GPT-4o (METEOR 0.16, BLEU-1 0.11, semantic similarity 0.30).

### **Empirical Analysis of State-to-Visual (S2V) Imitation vs. Visual RL (Co-leading project)**

Advisor: Prof. Hao Su, UC-San Diego

- Benchmarked State-to-Visual DAgger vs. visual RL across 16 tasks from ManiSkill, DM-Control, and Adroit.
- Analyzed performance trade-offs, efficiency, and computational costs.
- Built a standardized S2V pipeline and derived practical recommendations.

Feb 2024 - Sep 2024  
San Diego, CA

### **Modality Transfer for PET and MRI Images**

Advisor: Prof. Vikas Singh; UW-Madison Medical Science Center, Computer Vision Group

- Enhanced image translation with self-attention, MobileNetV2, and total variance loss in the pix2pix framework.
- Proposed U-TransGan model achieving PSNR 32, 0.98 correlation, and 0.92 SSIM.

May 2022 - Jun 2023  
Madison, WI

### **Wisconsin Science and Computing Emerging Research Stars (WISCERS)**

Advisor Prof. Chunming Zhang

- Implemented Robust Principal Component Analysis with Complex Noise, applying it to feature analysis of 2D images in computer vision. Explored dimension reduction and optimized PCA algorithms.

Jun 2022 - Aug 2022  
Madison, WI

### **Simulation of the Connected and Automated Driving Systems**

Advisor: Prof. Bin Ran; The Connected Automated Vehicle Highway System Group

- Simulated and optimized traffic systems in CARLA, enhancing traffic management models for improved efficiency.
- Refined object detection algorithms, including YOLO and Faster R-CNN, to improve vehicle detection and traffic control systems.

Sep 2021 - May 2022  
Madison, WI

## **Industry Experiences**

---

### **Computer Vision Algorithm Engineer | Mech-Mind Robotics Technologies Ltd.**

- Developed algorithms for structured light 3D cameras, improving image accuracy and optimizing point cloud reconstruction for laser systems.
- Led the refinement of internal camera distortion models, significantly enhancing imaging fidelity and calibration precision.

Jun 2023 - Sep 2023  
Beijing, China

### **Backend Engineer | Quanzhou YouGouZan Network Technology Co., Ltd.**

- Developed an online shopping mall on WeChat using SQL and Java, enabling functionalities like product search, browsing, recommendations, ordering, and payment.

Jun 2020 - Aug 2020  
Quanzhou, China

## **Teaching**

---

### **Teaching Assistant at UW-Madison**

- CS540: Introduction to Artificial Intelligence

Spring 2023

### **Peer Mentor at UW-Madison**

- CS537: Introduction to Operating System

Fall 2022

## **Professional Services**

---

### **Reviewer**

- AAAI Conference
- AAAI Workshop: Large Language Models and Generative AI for Health

## **Technical Skills**

---

- **Language:** English (Proficient, TOEFL:108 GRE:325+3.5), Chinese(Native)
- **Programming Language:** Assembly, C, C++, Java, C#, Python, R, Matlab, Latex, MySql, SQL
- **Frameworks:** Frameworks: OpenCV, Tensorflow, Pytorch, Sklearn, SPM12, SimpleITK, Torchio, Carla
- **Developer Tools:** IntelliJ IDEA, VS Code, Vim, Visual Studio, Jest, Xcode, UE, QT