Jia Li

Ph.D. Candidate — Computer Vision & Multimodal Learning

Jia.Li@utdallas.edu
jiali-home.github.io

Education

- **The University of Texas at Dallas** — Aug 2023 – May 2028 (expected)
  Ph.D. in Computer Science (Advisor: Dr. Yapeng Tian) — *GPA: 4.0/4.0*
  - **Research Focus**: Computer Vision, Multimodal Learning, Out-of-Distribution (OOD) Detection, Egocentric Perception
- **Sichuan University** — Sep 2018 – Jun 2022
  B.S. in Mathematics — *GPA: 3.59/4.0*

Research Experience

- **SceneFocus: Enhancing Objects of Interest in Augmented Reality to Facilitate Complex Environment Perception for People with Low Vision [Under review]**:
  - Designed an **AR** system that recognizes multiple objects of interest and applies **importance-aware** visual augmentation to help users quickly grasp scene context while reducing visual overload in complex environments.
  - Ran a user evaluation with 15 participants with low vision across **kitchen** and **outdoor navigation** scenarios.
- **Do Audio-Visual Segmentation Models Truly Segment Sounding Objects? [Under review]**:
  - Created AVSBench-Robust (positive/negative audio-visual pairs) to test robustness to silent, ambient, and off-screen sound.
  - Developed a PyTorch pipeline with classifier-guided learning and negative sampling, which reduced erroneous segmentations in invalid audio cases by over 80% while maintaining SOTA performance.
- **Segment Every Out-of-Distribution Object [CVPR 2024]**:
  - Proposed S2M (Score-to-Mask), a framework to convert anomaly scores into OoD masks without manual thresholds.
  - Integrated dynamic prompt generation and outlier exposure to boost robustness under distribution shift, improving IoU by 15% and mean F1 by 25% over prior SOTA.

Publications & Preprints

- **Li, J.**, Zhao, W., Huang, Z., et al. (2025). Do audio-visual segmentation models truly segment sounding objects? *Manuscript under review.* [Paper] • [Code]
- Ou, Y., Wang, T., Zhang, Y., **Li, J.**, et al. (2025). Local uncertainty quantification for enhancing hallucination detection in code generation. *Manuscript under review.*
- **Li, J.**, & Tian, Y. (2025). From waveforms to pixels: A survey on audio-visual segmentation. *arXiv preprint.* [Paper]
- Zhao, W., **Li, J.**, et al. (2024). Segment every out-of-distribution object. *CVPR 2024.* [Paper] • [Demo]
- Huang, Z., **Li, J.**, Zhao, W., & Tian, Y. (2024). AV-Mamba: Cross-modality selective state space models for AVQA. *CVPR 2024 Sight and Sound Workshop.* [Paper]
- Wang, K., Abdoul Moktar, S., **Li, J.**, et al. (2024). Aleatoric and epistemic uncertainty in LLMs on ID and OOD QA tasks. *UDM-KDD Workshop 2024.*

Technical Skills

- **Programming**: Python, C++, CUDA, Bash, Git, LaTeX
- **Research Areas**: Computer Vision, Multimodal Learning, Out-of-Distribution (OOD) Detection, Audio-Visual Segmentation, Egocentric Perception, Transformers, Generative AI
- **Systems & Infra**: Distributed Training, High-Performance Computing, Training Pipelines, Docker, Slurm
- **Math & Foundations**: Probability, Statistics, Linear Algebra, Optimization, Uncertainty Quantification

Academic Experience

- **Reviewer**: AAAI'26, NeurIPS'25, ICML'25, AISTATS'25, NeurIPS'24, ICLR'25, ICML'25, AISTATS'25, TMLR
- **Research Assistant**: Prof. Feng Chen (May 2024 – Aug 2024); Prof. Yapeng Tian (Jan 2025 – Present)
- **Teaching Assistant**: Discrete Mathematics (Aug 2023 – Dec 2023); Operating Systems (Jan 2024 – May 2024); Advanced Computer Networks (Aug 2024 – Dec 2024)