LEI ZHANG

Email: lez023@ucsd.edu | Phone: (323)824-7872 | Homepage | Google Scholar

Education

University of California, San Diego — San Diego, USA
Ph.D. in Computer Science — *Sep. 2024 – Present*
Advisor: Prof. Julian McAuley

Zhejiang University — Hangzhou, China
M.S. in Computer Science — *Sep. 2021 – Jun. 2024*

South China University of Technology — Canton, China
B.E. in Software Engineering — *Sep. 2017 – Jun. 2021*

Publications

- **L. Zhang**, J. Zhang, B. Lei, S. Mukherjee, X. Pan, B. Zhao, C. Ding, Y. Li, D. Xu, "Accelerating Dataset Distillation via Model Augmentation," *CVPR*, 2023 **Highlight**. PDF
- **L. Zhang**, Z. Wang, X. Dong, Y. Feng, X. Pang, Z. Zhang, K. Ren, "Towards Fairness-aware Adversarial Network Pruning," *ICCV*, 2023. PDF
- J. Zhang*, **L. Zhang***, G. Li, C. Wu, "Adversarial Examples for Good: Adversarial Examples Guided Imbalanced Learning," *ICIP*, 2022. (* equal contribution) PDF
- J. Zhang, B. Li, J. Xu, S. Wu, S. Ding, **L. Zhang**, C. Wu, "Towards Efficient Data Free Black-box Adversarial Attack," *CVPR*, 2022. PDF
- **L. Zhang**, F. Shu, S. Ren, B. Zhao, H. Jiang, C. Xie, "Filter & Align: Leveraging Human Knowledge to Curate Image-Text Data," *In submission to TMLR*, 2024. PDF
- F. Shu*, Y. Liao*, L. Zhuo*, C. Xu*, **L. Zhang***, G. Zhang*, H. Shi*, L. Chen, T. Zhong, W. He, S. Fu, H. Li, B. Li, Z. Yu, S. Liu, H. Li, H. Jiang, "LLaVA-MOD: Making LlaVA Tiny Via MOE-Knowledge Distillation" *In submission to ICLR 2025.* (* core members) PDF
- F. Shu*, **L. Zhang***, H. Jiang, C. Xie, "Audio-visual LLM for Video Understanding," *Arxiv Preprint*, 2024. (* equal contribution) PDF
- S. Ren, X. Li, H. Tu, F. Wang, F. Shu, **L. Zhang**, J. Mei, L. Yang, P. Wang, H. Wang, A. Yuille, C. Xie, "Autoregressive Pretraining with Mamba in Vision," *Arxiv Preprint*, 2024. PDF
- W. He, S. Fu, M. Liu, X. Wang, W. Xiao, F. Shu, Y. Wang, **L. Zhang**, Z. Yu, H. Li, Z. Huang, L. Gan, H. Jiang, "MARS: Mixture of Auto-Regressive Models for Fine-grained Text-to-image Synthesis," *Arxiv Preprint*, 2024. PDF
- W. Zhang, T. Lin, J. Liu, F. Shu, H. Li, **L. Zhang**, W. He, H. Zhou, Z. Lv, H. Jiang, J. Li, S. Tang, Y. Zhuang, "HyperLLaVA: Dynamic Visual and Language Expert Tuning for Multimodal Large Language Models," *Arxiv Preprint*, 2024. PDF

Experience

Research Assistant — Sept. 2024 – Present
University of California, San Diego. — *San Diego, USA*
- Supervisior: Prof. Julian McAuley and Prof. Zhijian Liu
- Research focus: Multimodal Large Language Model

Research Intern — Apr. 2023 – Jun. 2024
University of California, Santa Cruz. — *Remote*
- Supervisor: Prof. Cihang Xie
- Research focus: Vision-Language Learning

Research Intern — May. 2023 – Jun. 2024
Alibaba Group. — *Beijing, China*
- Research focus: Multimodal Understanding

Reserach Intern Dec. 2022 – Feb. 2023
Microsoft Research Asia. *Remote*

- Supervisor: Dr. Xun Guo
- Research focus: Text-to-Image with Diffusion Model

Projects

Efficient LLaVA via Mixture of Experts and Knowledge Distillation *Mar. 2024 – Sept. 2024*

- Integrate a sparse MoE architecture into the language model to strike a balance between computational efficiency and model expressiveness.
- Propose progressive knowledge transfer strategy (a) Mimic Distillation on dense and sparse architecture separately to facilitate hierarchical knowledge transfer. (b) Preference Distillation to adjust probability distribution on preference data.
- Ours-2B exceeds Qwen-VL-Chat-7B by an average of 8.8% on comprehension benchmarks with 0.3% training data and 23% training parameters.
- Ours-2B matches RLHF-based methods with 7B and 13B parameters on hallucination benchmarks.

Audi-Visual Large Language Model *Jun. 2023 – Oct. 2023*

- Propose modality-augmented training, integration of modality-specific tokens, for joint end-to-end training on different modalities.
- Curate a high-quality instruction dataset ranging from multi-turn conversations and audio-visual narratives to complex reasoning tasks.
- Our method outperforms non-LLM-based InterVideo by 6.6% and LLM-based Valley by 4.4% on video understanding benchmarks.

Technical Skills

Languages: Python, C/C++, LaTeX
Frameworks: Pytorch
Languages: TOEFL 105, GRE 325+3.0