

Wei Li

Austin, TX | weili@utexas.edu | 6197862930 | <https://weili1024.github.io/>

Education

Ph.D. in Computer Science , University of Electronic Science and Technology of China	2018 – 2022
M.S. in Computer Science , University of Science and Technology Beijing	2016 – 2018
B.S. in Computer Science , University of Science and Technology Beijing	2013 – 2016

Professional Experience

Postdoctoral Research Fellow , The University of Texas at Austin	2025 – Present
<ul style="list-style-type: none">Conducting research at the intersection of computer systems and risk-aware machine learning for large language model deployment.	
Researcher , Shanghai Artificial Intelligence Laboratory	2023 – 2024
<ul style="list-style-type: none">Developed InternLM-Xcomposer2, a free-form text-image composition and comprehension model achieving state-of-the-art on multimodal benchmarks.Contributed to InternLM2 technical report: architecture optimizations for robust vision-language understanding.Co-designed OmniCorpus, a 10 billion-image multimodal dataset interleaved with aligned text.Participated in Wanjuan-cc dataset release: safe, high-quality open-source English web-text for LLM training.	

Awards and Scholarships

- Postdoctoral Fellowship, The University of Texas at Austin, 2025
- Outstanding Ph.D. Dissertation Award, University of Electronic Science and Technology of China, 2022
- First Prize, ACM Student Research Competition (SRC), Asia Regional, 2021
- Runner-up, Best Student Paper Award, AAAI Conference on Artificial Intelligence, 2020
- Graduate Fellowship, University of Science and Technology Beijing, 2016–2018
- First Prize, National Undergraduate Mathematical Contest in Modeling (MCM/ICM), China, 2015

Selected Publications

- Dong X., Zhang P., Zang Y., Cao Y., Wang B., Ouyang L., Zhang S., Duan H., Zhang W., Li Y., Yan H., Gao Y., Chen Z., Zhang X., Li W., Li J., Wang W., Chen K., He C., Zhang X., Dai J., Qiao Y., and Lin D. InternLM-XComposer2-4KHD: A pioneering large vision-language model handling resolutions from 336 pixels to 4K HD. *Advances in Neural Information Processing Systems*, 37:42566–42592, 2024.
- Chen Z., Wang W., Tian H., Ye S., Gao Z., Cui E., Tong W., Hu K., Luo J., Ma Z., Ma J., Wang J., Dong X., Yan H., Guo H., He C., Shi B., Jin Z., Xu C., Wang B., Wei X., Li W., Zhang W., Zhang B., Cai P., Wen L., Yan X., Dou M., Lu L., Zhu X., Lu T., Lin D., Qiao Y., and Dai J. How far are we to GPT-4V? Closing the gap to commercial multimodal models with open-source suites. *Science China Information Sciences*, 67(12):220101, 2024.
- Li Q., Chen Z., Wang W., Wang W., Ye S., Jin Z., Chen G., He Y., Gao Z., Cui E., Yu J., Tian H., Zhou J., Xu C., Wang B., Wei X., Li W., Zhang W., Zhang B., Cai P., Wen L., Yan X., Li Z., Chu P., Wang Y., Dou M., Tian C., Zhu X., Lu L., Chen Y., He J., Tu Z., Lu T., Wang Y., Wang L., Wang L., Lin D., Qiao Y., Shi B., He C., and Dai J. OmniCorpus: A unified multimodal corpus of 10 billion-level images interleaved with text. *The Thirteenth International Conference on Learning Representations.*, 2024.
- Cai Z., Cao M., Chen H., Chen K., Chen X., Chen X., Chen Z., Chen Z., Chu P., Dong X., Duan H., Fan Q., Fei Z., Gao Y., Ge J., Gu C., Gu Y., Gui T., Guo A., et al. InternLM2 Technical Report. *CoRR.*, 2024.
- Qiu J., Lv H., Jin Z., Wang R., Ning W., Yu J., Zhang C., Li Z., Chu P., Qu Y., Shi J., Lu L., Peng R., Zeng Z., Tang H., Lei Z., Hong J., Chen K., Fei Z., Xu R., Li W., Tu Z., Yan H., and He C. Wanjuan-CC: A safe and high-quality open-sourced English web-text dataset. *arXiv preprint arXiv:2402.19282*, 2024.
- Dong X., Zhang P., Zang Y., Cao Y., Wang B., Ouyang L., Wei X., Zhang S., Duan H., Cao M., Zhang W., Li Y., Yan

- H., Gao Y., Zhang X., Li W., Li J., Chen K., He C., Zhang X., Qiao Y., Lin D., and Wang J. InternLM-XComposer2: Mastering free-form text-image composition and comprehension in vision-language large models. *arXiv preprint arXiv:2401.16420*, 2024.
- Ren B., Li Y., Mehta N., Timofte R., Yu H., Wan C., Hong Y., Han B., and Wu Z. The Ninth NTIRE 2024 Efficient Super-Resolution Challenge Report. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.
 - Zhang P., Dong X., Wang B., Cao Y., Xu C., Ouyang L., Zhao Z., and Duan H. InternLM-XComposer: A vision-language large model for advanced text-image comprehension and composition. *arXiv preprint arXiv:2309.15112*, 2023.
 - Ershov E., Savchik A., Shepelev D., Banic N., Brown M. S., and Timofte R. NTIRE 2022 Challenge on Night Photography Rendering. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
 - Du D., Zhu P., Wen L., Bian X., Lin H., Hu Q., Peng T., Zheng J., Bai X., Wang Y., Shi J., Gao Z., Zhang Y., Li W., Han K., Zou Y., Wang G., Peng Z., Sun C., and Zhang L. VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.
 - Zhu P., Wen L., Du D., Bian X., Ling H., Hu Q., Nie Q., Cheng H., Liu C., and Wang X. VisDrone-DET2018: The Vision Meets Drone Object Detection in Image Challenge Results. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2018.

Invited Talks

- (2025) “Risk-Controlled Inference for Large Language Models” at NeurIPS 2025 Workshop on Responsible AI Systems.
- (2024) “Scalable Backend Architectures for Vision-Language Models” at CVPR 2024 Systems and Applications Workshop.
- (2023) “InternLM-XComposer2: High-Resolution Vision-Language Modeling” at NeurIPS 2023 Workshop on Multimodal Learning.
- (2022) “Risk Assessment in Distributed Machine Learning Pipelines” at KDD 2022 Workshop on ML Systems and Reliability.
- (2021) “Efficient Deployment of Large Language Models in Cloud-Native Environments” at ICML 2021 Systems ML Workshop.
- (2020) “High-Performance Computing for Machine Learning Workloads” at SC 2020 Workshop on HPC AI Convergence.
- (2019) “Integrating Risk Control into Reinforcement Learning Frameworks” at IJCAI 2019 Workshop on Safe and Explainable AI.

Skills & Tools

Languages & Frameworks: Python, PyTorch, TensorFlow, C++

ML Topics: Vision-Language Models, Multimodal Pretraining, Causal Abstraction, Reinforcement Learning

Systems: Distributed Training, Microservices, Anomaly Detection, Uncertainty Quantification, Low-Latency Inference Pipelines

Tools: Git, Docker, Kubernetes, Linux

Services and Outreach

- **Journal Reviewer:**
 - IEEE Robotics and Automation Letters (RA-L)
 - IEEE Transactions on Automation Science and Engineering (T-ASE)
 - IEEE Transactions on Intelligent Transportation Systems (T-ITS)
 - IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)
 - Transportation Research: Part C (TR-C)
- **Conference Reviewer:**
 - AAAI Conference on Artificial Intelligence (AAAI)

- American Control Conference (ACC)
- Conference on Computer Vision and Pattern Recognition (CVPR)
- Conference on Neural Information Processing Systems (NeurIPS)
- Conference on Robot Learning (CoRL)
- IEEE International Conference on Control, Automation, Robotics, and Vision (ICARCV)
- IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
- IEEE International Conference on Robotics and Automation (ICRA)
- IEEE Intelligent Transportation Systems Conference (ITSC)
- IEEE Intelligent Vehicles Symposium (IV)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)
- International Joint Conference on Artificial Intelligence (IJCAI)
- Learning for Dynamics and Control (L4DC)
- Modeling, Estimation, and Control Conference (MECC)
- North American Manufacturing Research Conference (NAMRC)
- Robotics: Science and Systems (RSS)
- World Congress of the International Federation of Automatic Control (IFAC)

- **Program Committee:**

- Associate Editor, IEEE Intelligent Transportation Systems Conference (ITSC), 2021–2024
- Lead Organizer, IEEE ICRA Workshop on Human-Centered Robot Learning in the Era of Big Data and Large Models, 2025
- Lead Organizer, NeurIPS Workshop on Progress and Challenges in Trustworthy Embodied AI, 2022
- Co-organizer, IROS Workshop on Multi-Agent Interaction and Relational Reasoning, 2021
- Co-organizer, IEEE IV Workshop on Development of Socially-Compliant Driving Behaviour for Automated Vehicles, 2023
- Program Committee, NeurIPS Workshop on ML for Autonomous Driving (ML4AD), 2022
- Program Committee, NeurIPS Workshop on Robot Learning, 2023
- Program Committee, Symposium on ML for Autonomous Driving, 2023
- Program Committee, ICCV Workshop on Visual Perception for Navigation in Human Environments: The JackRabbit Human Motion Forecasting, 2023
- Program Committee, RSS Pioneers, 2024
- UT Austin Computer Science Ph.D. Admission Committee

- **Outreaching:**

- Volunteering mentor in the STEM Muse Mentorship Program