Xiao(Harold) Liu

xioliu@ucdavis.edu · +1 5304077636 · haroldliuj.github.io

Research Interests

Multimodal Language Understanding

Multimodal Large Language Model, Video Understanding, Summarization

Agentic AI

Multimodal Reasoning, Tool-use, Planning

Education

University of California, Davis — Davis, USA
Ph.D. in Computer Science (Expected 03/2026) — 09/2021–present
– Advisor: Dr. Jiawei Zhang

Southeast University — Nanjing, China
Bachelor of Engineering in Software Engineering — 09/2016–06/2020

University of California, Berkeley — Berkeley, USA
Exchange Student in Computer Science and Psychology — 08/2019–01/2020

Work Experience

Amazon Web Services, Bedrock — New York, USA
Applied Scientist Intern *(Return Offer)* — 06/2025–10/2025

- Developed an **innovative parameter-free learning agentic framework for complex image reasoning**, enabling vision-language models to solve challenging image-based questions by learning to use complex tools.
- **Completed a research paper** based on this work and **received a full-time return offer**.

Applied Scientist Intern — 06/2024–09/2024

- Developed **a novel training pipeline** that integrates tool-use capabilities into the parametric knowledge of visual language models, enhancing their reasoning abilities without increasing inference latency. This pipeline significantly improved model performance while maintaining the same inference time.
- Designed and curated **a large-scale dataset** comprising image operation reasoning steps and corresponding intermediate images.
- **A research paper** is under review and **received a intern return offer**.

RoastPic Inc — Davis, USA
Technical Lead, Co-founder — 06/2022–09/2023

- Played a role as a **co-founder** in conceptualizing and developing an advanced image processing system tailored for coffee bean analysis, leading to total **$500K investment** from a leading industry partner. Successfully deployed to **500+ beta testers**, collecting real-world usage data to refine the product.
- **Led and motivated a dynamic team of 5 developers as PM**, fostering a collaborative environment to drive the successful development and deployment of the product.
- Designed and built the full software stack using AWS, React Native, and MongoDB Atlas, with structured DB schemas and a robust CI/CD pipeline via GitHub Actions. Created a shared wiki knowledge base to support team collaboration.
- Applied **advanced deep learning techniques** for the detection of defects in green coffee beans, enhancing the precision of quality assessments. The detection model could significantly reduce human effort in the coffee QC process in the future.

Intel Beijing, China
Research Scientist Intern 01/2021–06/2021

- Developed and implemented **3 innovative visual-linguistic modality interaction schemas** for a visual dialogue model, leading to **SOTA performance** enhancements.
- Executed comprehensive ablation studies to assess the impact and interoperability of the model's components, confirming the effectiveness of the proposed methodologies.
- Led the structural revision of the previously rejected research paper, enhancing coherence and logical flow, resulting in **acceptance at the ACL DialDoc2021 Workshop**.

KG Data Nanjing, China
Machine Learning Engineer 07/2020–01/2021

- Pioneered the adaptation of the BERT model for the company's private medical QA dataset during BERT's initial release period. Led the development of a comprehensive fine-tuning and deployment pipeline for the BERT model, culminating in a significant performance improvement over the existing online QA model. Shared insights and techniques through an internal workshop, fostering knowledge dissemination within the company.
- Developed an innovative text clustering algorithm using Gibbs Sampling for the Dirichlet Multinomial Mixture model. **One patent was submitted** and is currently under review.

Publications

Google Scholar

[1] **Liu, Xiao**, J. Zhang. "AIGVE-MACS: Unified Multi-Aspect Commenting and Scoring Model for AI-Generated Video Evaluation". *Under Review of CVPR 2026.*

[2] G.Sun, **Liu, Xiao**, T.Williams, R.Samaniego. "Learning to Regulate: A New Event-Level Dataset of Capital Control Measures". *In working paper of The Journal of Finance and Data Science.*

[3] **Liu, Xiao**, X. Xiang, Z.Li, Z.Liu, J. Zhang. "AIGVE-Tool: AI-Generated Video Evaluation Toolkit with Multifaceted Benchmark". *Under Review of CVPR 2026.*

[4] **Liu, Xiao**, X. Xiang, Z.Li, Y.Wang, Z.Li, Z.Liu, W.Zhang, W.Ye, J. Zhang. "A Survey of AI-Generated Video Evaluation". *Under Review of ACM Computing Survey.*

[5] **Liu, Xiao**, J. Zhang. "GPTA: Generative Prompt Tuning Assistant for Synergistic Downstream Neural Network Enhancement with LLMs". *Under Review of COLM 2025.*

[6] **Liu, Xiao**, J. Lin, and J. Zhang. "Beyond Text: Unveiling Multimodal Proficiency of Large Language Models with MultiAPI Benchmark". *In the Workshop of Towards Knowledgeable Language Models at The 62nd Annual Meeting of the Association for Computational Linguistics (ACL) 2024.*

[7] H. Zhang, **Liu, Xiao**, and J. Zhang. "SummIt: Iterative Text Summarization via ChatGPT". In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 2023.

[8] H. Zhang, **Liu, Xiao**, and J. Zhang. "Extractive Summarization via ChatGPT for Faithful Summary Generation". In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 2023.

[9] **Liu, Xiao**, H. Zhang and J. Zhang. "DiffuSum: Generation Enhanced Extractive Summarization with Diffusion". In *Proceedings of the Association for Computational Linguistics(ACL)* 2023.

[10] H. Zhang, **Liu, Xiao**, and J. Zhang."Contrastive Hierarchical Discourse Graph for Scientific Document Summarization". In *The 4th workshop on Computa-tional Approaches to Discourse at the 61st Annual Meeting of the Association for Computational Linguistics(ACL)* 2023.

[11] H. Zhang, **Liu, Xiao**, J. Zhang. "HEGEL: Hypergraph Transformer for Long Document Summarization". In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 2022.

[12] L. Yang, F. Meng, **Liu, Xiao**, M. Wu, V. Ying, and J. Xu. "SeqDialN: Sequential Visual Dialog Network in Joint Visual-Linguistic Representation Space". In *Proceedings of the 1st Workshop on Document-grounded Dialogue and Conversational Question Answering*. Association for Computational Linguistics (ACL DialDoc) 2021.

[13] N. Ma, **Liu, Xiao**, Y. Gao. "Entity Linking Based on Graph Model and Semantic Representation". In *International Conference on Knowledge Science, Engineering, and Management 2019*.

Applied Research Experience

Photometric Assessment of Coffee Color Naming Schema — Davis, USA
Volunteer Researcher at UC Davis Coffee Center and Specialty Coffee Association — 06/2022–10/2023

- Supervised by Prof.William Ristenpart. Worked on a novel approach developed for differentiating coffee colors through photometric analysis and regression on near-infrared metrics for Specialty Coffee Association (SCA). A new color naming schema based on the ISCC–NBS System was proposed. A poster was presented at SCA Coffee Expo 2023, currently under SCA review, potentially setting a new industry standard.

Certifications and Achievements

UC Davis Graduate Group in Computer Science (GGCS) Research Fellowship	2023
Coffee Quality Institute Q Arabica Grader	2022
Coursera deeplearning.ai Deep Learning Specialization	2019
Nvidia Image Recognition Course Training	2019
China National Innovation Project Fellowship	2018
Third Prize, Southeast University School History System Development Competition	2017

Teaching Experience

Teaching Assistant at UC Davis, ECS 189G Deep Learning	Winter 2025
Teaching Assistant at UC Davis, ECS 189G Deep Learning	Winter 2024
Teaching Assistant at UC Davis, ECS 189G Deep Learning	Winter 2023
Teaching Assistant at UC Davis, ECS 189G Deep Learning	Winter 2022
Teaching Assistant at UC Davis, ECS 289G Advanced Deep Learning	Spring 2022

Skills

Programming Languages: Python, JavaScript, Java
Deep Learning: PyTorch, Hugging Face, Transformers, OpenAI API, LLM/VLM Finetuning
Platforms: Git, Docker, Amazon AWS, GitHub Actions
Project Management: Lark, Notion, Miro, Figma
Languages: Mandarin (native), English (fluent)