Zhuoheng Li

zhuoheng.li@colorado.edu | linkedin.com/andy-zhli | andy-lzh.github.io | Google Scholar | Boulder, CO 80301

Research Interests

Computer Vision, Multimodal Learning, Hierarchical Segmentation, Accessibility

Education

University of Colorado Boulder — Boulder, CO
Ph.D. in Computer Science — *Aug 2024 - Present*

University Of California, Davis — Davis, CA
B.S in Computer Science — *Sep 2020 - June 2024*

Publications and Technical Report

- Chongyan Chen*, Yu-Yun Tseng*, **Zhuoheng Li**, Anush Kumar Venkatesh, Danna Gurari, "Accounting for Focus Ambiguity in Visual Questions". IEEE International Conference on Computer Vision (ICCV 2025)
- Zhuosheng Liu, **Zhuoheng Li**, Jiawei Zhang, C Titus Brown, Luxin Wang, "Pathogenic potential prediction of Vibrio parahaemolyticus by using pangenome data with high performance machine learning algorithms". bioRxiv 2025.04.08.647818
- Xiao Liu, Xinhao Xiang, Zizhong Li, Yongheng Wang, **Zhuoheng Li**, Zhuosheng Liu, Weidi Zhang, Weiqi Ye, and Jiawei Zhang, "A Survey of AI-Generated Video Evaluation". arXiv preprint arXiv:2410.19884.
- **Zhuoheng Li**, Zhuosheng Liu, Jiawei Zhang, "Parameter-Efficient Fine-Tuning for Vision-Language Models", Technical Report
- Zhengfeng Lai, **Zhuoheng Li**, Luca Cerny Oliveira, Joohi Chauhan, Brittany N. Dugger, Chen-Nee Chuah, "CLIPath: Fine-tune CLIP with Visual Feature Fusion for Pathology Image Analysis Towards Minimizing Data Collection Efforts", ICCV 2023 Workshop on Computer Vision for Automated Medical Diagnosis

Research Experience

IVC Group — Aug 2024 – Present
Graduate Research Assistant; Advised by Prof. Danna Gurari — *Boulder, CO*

- Curating hierarchical instance segmentation datasets and developing vision-language models that understand object compositions and support instance-level grounding to part and subpart level.
- Building models and datasets to detect focus ambiguity in visual questions and enable interactive disambiguation through ambiguity-aware vision-language reasoning.

UC Davis Coffee Center — Oct 2021 - Aug 2024
Research Assistant; Advised by Prof. William Ristenpart — *Davis, CA*

- Launched RoastPic, a coffee analytic application that uses computer vision and machine learning models to analyze size, color, and defects of coffee beans observed in an image; secured **$250K** in seed funding for this project.
- Presented this work, including through posters and demos at the Specialty Coffee Expo in 2022, 2023, and 2024 as well as a featured talk in Yunnan University's Coffee Forum in 2024.

IFM Lab — July 2023 - June 2024
Research Assistant; Advised by Prof. Jiawei Zhang — *Davis, CA*

- Led an empirical study on the impact of domain shifts when fine-tuning CLIP for downstream tasks; identified significant performance degradation on tasks involving complex spatial compositions.
- Designed models to predict the pathogenic potential of *Vibrio parahaemolyticus* using pangenome data, achieving AUC 0.98 and identifying key accessory genes associated with virulence through model feature analysis.
- Contributed to a survey on AI-generated video quality analysis by reviewing statistical video models used to assess spatial-temporal coherence.

RUbiNet Lab — Aug 2022 - Feb 2023
Research Assistant; Advised by Prof. Chen-Nee Chuah — *Davis, CA*

- Proposed a novel domain adaptation method for fine-tuning **CLIP** to downstream tasks using limited labeled data, achieving a **26.26%** accuracy improvement over Zero-Shot CLIP when fine-tuned with only **10%** of labeled data.

Industry Experience

RoastPic Inc. June 2023 - Aug 2024
Technical Co-Founder *Davis, CA*

- Led the development of image analysis services, internal tooling services, CI/CD pipelines, and infrastructure for a seed-round computer vision startup focused on analytics of coffee beans.
- Designed and implemented RESTful image analysis APIs using Django in Python to seamlessly interface with ML services, and employed Docker for environment containerization.
- Built CI/CD pipelines with GitHub Actions for automated testing and deployment of APIs as containers on AWS ECS, resulting in a **90%** efficiency improvement over manual processes.
- Oversaw GraphQL schema, defining fields, queries, and mutation types for user, analysis history, and calibration sheet collections, optimizing CRUD operations for internal tools.

Collimate Mar 2021 - Jan 2022
Software Engineer *Davis, CA*

- Collaborated with an engineering team of 11 to implement a real-time chat application aimed at helping students find classmates during Covid-19; helped **300+** students find classmates under remote instruction.
- Developed a Homepage UI for both Android and iOS, encompassing side navigation menus and Class chat interfaces, utilizing React Native as framework; TypeScript, Kotlin, and Swift as programming languages.

Open-Source Projects

Molmo-0.5B | *Python, PyTorch* Mar 2025 - May 2025

- Integrated Qwen-0.5B as the language encoder in a Molmo-style VLM architecture and trained from scratch on the PixMo dataset, achieving comparable visual grounding performance to 7B models on pointing tasks.

SAM-Quest | *Unity, C-sharp, ONNX* Mar 2025 - May 2025

- Deployed an ONNX-formatted Segment Anything Model (SAM) on Meta Quest 3. Built an interactive segmentation app enabling users to segment objects via ray-based point selection in AR.

MetaCLIP | *Python, PyTorch* Oct 2023 - Nov 2023

- Contributed to MetaCLIP github repository by submitting a pull request that added an automated feature for downloading and loading trained checkpoints including ViT-B32/B16/L14-400M and 2.5B.

Technical Skills

Languages: Python, C/C++, Java, HTML/CSS, SQL, Bash, regex
Technology: PyTorch, scikit-learn, sciki-image, OpenCV, Hugging Faces, Weights & Biases, Pandas, Docker, Kubernetes, AWS, React, D3.js, GraphQL, Django, FastAPI
Database: SQLite, PostgreSQL, MongoDB, Firebase, Realm, Redis

Service

Organizer VizWiz Grand Challenge Workshop at CVPR 2025

Reviewer ICLR '24, ICLR '25, ACM TCH

Mentor CU Prospect Match program 2025