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, Visual Question Answering, Assistive Technology

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". under review.
- 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, University of Colorado Boulder — Aug 2024 - Present
Graduate Research Assistant; Advised by Prof. Danna Gurari — *Boulder, CO*

- Developing models to recognize ambiguity in visual questions, focusing on data from people who are blind.

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 parameter-efficient fine-tuning for CLIP, focusing on the impact of the model's backbone and modalities on its adaptability to downstream tasks across varied domain shifts.
- Worked on model development for visual question answering, with a focus on complex scenes and integrating LiDAR as a new modality alongside RGB and text.
- Configured a Kubernetes infrastructure with CUDA-enabled pods to efficiently run and monitor experiments in a controlled environment.

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 - Present
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 a CI/CD pipeline 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.

Research Projects

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.

DeepCoffee | *Python, TypeScript, OpenAI, React.js, Material UI* April 2022 - June 2022

- Built a DeepCoffee Model with an accuracy of **83.2%** in coffee flavor prediction task using GPT-3 text-davinci-002 Engine fine-tuned with limited coffee flavor data of only **1300** entries.

Class Projects

Minimal Photoshop | *Python, NumPy, OpenCV, Scikit-Image, Scikit-Learn* Apr - June 2023

- Implemented a sharpening feature using NumPy and OpenCV, applying Gaussian blur for initial smoothing and subsequent difference calculation to achieve image sharpening.
- Conducted Experiments on comparing effectiveness of Sobel Operators(Magnitude, Orientation) and Canny Edge Detection on detecting edges.
- Developed a panorama function by employing a pipeline that included Harris corner detection for feature point identification, SIFT descriptor formation, RANSAC image alignment, and subsequent image transformations.

ECS-150-FS | *C/C++, GNU Make, Linux* May 2023

- Developed kernel-level components to support file system(FAT-Like) operations, including file creation, deletion, listing, and data retrieval, using C/C++ and Linux.
- Enhanced Makefile rules to maintain generality while implementing precise dependency tracking and automated dependency file generation, ensuring efficient, adaptable, and accurate builds.

AI Tonight | *Python, PyTorch, Hugging Face* Feb - Mar 2022

- Pre-Processed Short Jokes Dataset consisting 0.2M lines of jokes to improve model performance.
- Built a novel NLP model that generates jokes with pre-trained GPT-2 Model fine-tuned with Short Jokes Dataset, using Python, PyTorch and Hugging Face.

Technical Skills

Languages: Python, TypeScript/JavaScript, 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
Database: SQLite, PostgreSQL, MongoDB, Firebase, Realm, Redis

Service

Organizer VizWiz Grand Challenge Workshop at CVPR 2025

Reviewer ICLR 2024

Mentor CU Prospect Match program 2025