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, Assitive Technology

Education

University of Colorado Boulder — Boulder, CO
Ph.D. in Computer Science — *Aug 2024 - May 2029*

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

Publications and Technical Report

- **Z. Li**, Z. Liu, J. Zhang, "Parameter-Efficient Fine-Tuning for Vision-Language Models", Under Review
- Z. Lai, **Z. Li**, L. Cerny Oliveira, J. Chauhan, B. N. Dugger, C-N. 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
- M. Suresh, **Z. Li**, "DeepCoffee : Coffee Flavors Prediction Using Deep Learning", Technical Report for ECS 289G Advanced Deep Learning

Research Experience

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

- Improving algorithms and datasets from existing general pre-trained vision and vision-language models to fairly and accurately assist visually-impaired people to understand their surroundings.

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

- Launched RoastPic, a coffee analytic app that uses computer vision algorithms and ML models to get size, color, and defects information from coffee beans image; secured **$250K** in seed funding.
- Presented poster and demoed at Specialty Coffee Expo 2022, 2023, and 2024; and delivered 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 in CLIP, focusing on the impact of model backbone, methods, and modalities on adaptability to downstream tasks across varied domain shifts.
- Worked on research in Visual Question Answering with a focus on complex scenes, integrating LiDAR as a new modality alongside RGB and text for improved analysis and understanding.
- 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 an **26.26%** accuracy improvement over Zero-Shot CLIP when fine-tuned with only **10%** of labeled data.

Volunteering

ICLR 2024 - Tiny Paper Track — Reviewer

Teaching Interest

CSCI 3308: Software Development Methods and Tools — Aug 2024 - Dec 2024

ECS 289G: Advanced Deep Learning — June 2022 - June 2022

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.
- Customized Material UI's autocomplete component to enable enhanced auto complete function via OpenAI API and DeepCoffee model, using user input as prompt and model-predicted information as autocomplete results.
- Summarized the project into technical report, taking primary responsibility for the method and result sections.

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.

Fitness Tracker | *JavaScript, HTML/CSS, Express, REST API, Passport.js, GCP, SQLite3* Mar 2021

- Implemented RESTful API to perform CRUD operations on fitness-related activity records, took advantage of Express.js to perform SQL query in response of request.

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