

# THOMAS YEOH

+1 (669)-799-7334 | tyeh@berkeley.edu | <https://tyeh9.github.io/tyeh/>

## Education

**University of California, Irvine**  
*Master of Computer Science*

Expected: Dec 2026

**University of California, Berkeley**  
*B.S. Electrical Engineering and Computer Science*

Graduated: May 2025

- Coursework: Data Structures, Algorithms, Computer Architecture, Optimization models, Discrete Mathematics, Probability Theory, Computer Vision, Computer Graphics, Machine Learning, Databases
- Activities & Awards: Cal Leadership Scholarship, Google Developers Club (VP of Strategy)

## Skills

**Languages**: Python, Java, C++, HTML, CSS, JavaScript, SQL, R, Bash, Assembly, Scheme

**Technologies**: Linux, GCP, PyTorch, Pandas, NumPy, Git, Flask, React, Node.js, Streamlit, Docker, Jupyter

## Experiences

**Berkeley Operations & Behavioral Analytics Lab**  
*Research Assistant*

Berkeley, CA

Jan 2024 – May 2025

- Performed data analysis and manipulation using Pandas and NumPy on datasets to support empirical findings.
- Translated mathematical models into code, optimized the data pipeline, and sped up pre-computation by 44%.
- Actively involved in using HPC cluster computers, creating tutorials, and adding documentation for lab members.

**Keysight Technologies**  
*R&D Software Engineer Intern*

Penang, Malaysia

May 2024 – August 2024

- Deployed and developed a text classification app for service orders with React, reducing analysis time by 25%.
- Built an end-to-end data pipeline with REST API and Flask to efficiently process over 120,000 data points daily.
- Utilized various libraries (Keras, Sklearn, Seaborn) for advanced data modeling and analysis on large datasets.
- Experimented with ML methods: XGBoost, SVM, K-means, logistic regression, and Facebook fastText.

## Projects

**RookieDB** (Java, SQL)

- Implemented a SQLite-like (RDBMS) database using policies such as multi-granularity locking and ARIES recovery.
- Utilized data structures like B+ tree indices and join algorithms for efficient query execution.

**Live Traffic Monitor** (GCP, Docker, Streamlit)

- Deployed a web app via GCP Run and Docker to fetch snapshots of live traffic and store them in Cloud Storage.
- Analyzed and displayed congestion levels on a Streamlit app via OpenAI's vision API.

**Image Denoising with U-Net and Diffusion Models** (Python, PyTorch)

- Developed and trained a U-Net for single-step image denoising on noisy datasets.
- Implemented diffusion-based iterative denoising for enhanced noise removal in high-noise images.
- Visualized model performance across training epochs, including varying noise levels for quality analysis.

**Automatic Image Stitching** (Python)

- Developed an image stitching system using Harris corner detection and feature matching.
- Implemented 4-point RANSAC to compute homographies, producing panoramas from overlapping images.
- Automated mosaic generation based on Brown et al.'s research paper for accurate image alignment.