




Thomas Hayama


Machine Learning Engineer

 thomashayama.com

 thomashayama

 thomas-hayama

 thomashayama@gmail.com

 +1(513)307-5640

ABOUT ME

Applied AI engineer with 3+ years of experience developing production and prototype systems. Experienced with integrating large language models into intelligent tools and deploying scalable services for LLM-based applications. Strong theoretical foundation backed by practical research experience.

LOCATED IN SUNNYVALE, CA

EDUCATION

CALIFORNIA INSTITUTE OF TECHNOLOGY

BS COMPUTER SCIENCE

MINOR INFORMATION &

DATA SCIENCE

2022 | Cum. GPA: 4.0 / 4.0

Focus in Machine Learning/Robotics

SKILLS

PROGRAMMING

Python (Expert)

SQL • LaTeX (Experienced)

C++ • MatLab • Java (Familiar)

LIBRARIES/Frameworks

PyTorch • Pandas • sklearn
SQLite • SQLAlchemy • MLflow
ROS[2] • Gazebo • OpenCV
FastAPI • Flask • Streamlit
Matplotlib • NumPy • Numba
Scalene

TOOLS/PLATFORMS

Docker [Compose]
Kubernetes • AWS
Harbor • Rancher • Airflow
Temporal • RagFlow • MCP
Jenkins • Sagemaker

MISC.

Portuguese (Native)
Autodesk Inventor • Fusion360
Solidworks • VHDL
V8 Max Climbing Grade

EXPERIENCE

THE AEROSPACE CORPORATION (DATA SCIENCE AND AI DEPT.)

Machine Learning Engineer

August 2022 – Current

- RAG-Enabled Chatbot with Table Understanding - Created a custom document chunking microservice API on kubernetes, used **ColBERT** for embedding/retrieval, and **Command-r/GPT-4o** as the frontend LLM.
- Arbitrary-length document generation - Used a hierarchy structure generation method to create long-form (larger than context length) documents with generated **LLM** prompts.
- High Frequency Anomaly Detection - Used a mix of classical and ML techniques to detect out of family behavior in live and historical high frequency rocket data. **PyTorch, Kolmogorov-Smirnov, FFT, CNN**
- Visual Odometry for non-cooperative RPOD - Created pipelines for simulated images of space objects in **Blender**. Implemented various deep learning and classical algorithms for 6/3/2DoF pose estimation for real-time and offline use. **ROS2, OpenCV, ICP, PnP, ORB, Superpoint, Superglue/Lightglue**

THE AEROSPACE CORPORATION (DATA SCIENCE AND AI DEPT.)

Technical Intern

June 2021 – September 2021

- Applied **Unsupervised** and **Semi-Supervised** learning methods for high frequency anomaly detection in rocket data. **PyTorch**
- Modified an automotive LiDAR SLAM library (**MULLS-SLAM C++**) for use in space. Removed assumptions to allow for 6DoF rather than ~3DoF estimation.

AUTONOMOUS ROBOTICS AND CONTROL LAB (CALTECH)

Research Fellow

June 2020 – October 2020

- Used **PyTorch** to train a model to classify thermal images of people and vehicles using only visible wavelength data (**Blind Domain Adaptation**) to 95.7%. Experimented with different preprocessing techniques to simulate thermal data with visible data. Used **variational autoencoders** to better transform from the visible to the thermal domain with satellite images.

PROJECTS

PUZZLE SOLVING ROBOT ARM (GROUP) | PYTHON, ROS

thomashayama.com/puzzle_bot

- Given the solved image, solves a shuffled puzzle with a custom 5 DoF arm, vacuum end effector, and **Realsense** depth camera.
- Used **OpenCV** for piece, corner, and edge detection for position and orientation.
- Millimeter precision accomplished with **PID** tuning and small relative motions such that puzzle pieces fit together. Cartesian motion control with inverse kinematics via Newton-Raphson

SLAM* ROBOT (INDIVIDUAL) | PYTHON, ROS, GAZEBO

thomashayama.com/slam_bot

- Probabilistic Roadmap (**PRM**) + **A*** navigation using a 2-wheeled robot.
- Localization via **ICP** with a 2d pointcloud from a **Realsense** depth camera.
- Voxel based mapping, with 'probability' buildup over time from pointcloud ray tracing.

FERMENTATION TRACKER AND OPTIMIZER (INDIVIDUAL) | PYTHON

thomashayama.com/alcopt

- Web app for tracking wine/mead fermentations, recording tasting reviews, and optimizing taste/quality. Recommender system using **collaborative filtering** and **Bayesian optimization**.
- Self hosted **streamlit** site with Google OAuth. Hosted in a **docker** container in a linux VM on my Proxmox machine.