

Hanson Li

971-275-7450 | hanson.li@wustl.edu | linkedin.com/in/hanson-li1 | github.com/SmallChungus1 | U.S. Citizen

EDUCATION

Washington University in St. Louis

St. Louis, MO

M.S. in Computer Science – GPA: 3.9/4.0

May 2026

B.S. in Computer Science and Economics – GPA: 3.6/4.0

Dec 2024

Relevant Coursework: Data Structures and Algorithms, Machine Learning, Database Management, Computer Vision

EXPERIENCE

Software Engineer Intern

June 2025 – Aug 2025

CarMax

Dallas, TX

- Developed a Python-based web application enabling data scientists to configure A/B tests of vehicle recommendation models; utilized Azure App Service, CosmosDB, and integrated SSO authentication with Azure Entra ID
- Built embedding-based vehicle recommendation and semantic search proof of concepts leveraging a vector database paired with a vision-language model finetuned on 5k+ vehicle image-text pairs for vehicle embedding generation
- Wrote extensive unit and integration tests using pytest and mock frameworks, increasing project code coverage to 80%+

Machine Learning Intern

Jan 2025 – Apr 2025

Norfolk Southern Railway

Atlanta, GA

- Fine-tuned an object detection model using real and synthetic data for rail-car wheel defects, achieving 0.95+ mAP; detected two previously unknown defect classes via parallelized offline inference over 18M+ images
- Prototyped synthetic data generation methods with Diffusion Models and GANs; successfully trained a FastGAN model with 30 images to generate new classes of rail-car wheel defects
- Built an active-learning data pipeline that identifies and uploads low confidence inference images to CVAT for labeling with SQLite, CVAT API, and Cron orchestration; achieved automation of 300+ image uploads to CVAT daily

Machine Learning Developer

Sept 2024 – Apr 2025

Engineer Test Kitchen (Student Organization)

St. Louis, MO

- Trained YOLO-based dense staves counting model under low-data constraints (40 training images), achieving 0.8+ mAP with domain-relevant pretraining
- Developed a containerized model inference web app. Accelerated inference speed by 19× on CPU using ONNX Runtime

Machine Learning Research Assistant

May 2024 – Present

Washington University School of Medicine

St. Louis, MO

- Trained 3D-ResNets and shallow ML models for radiology brain scan quality classification, achieving AUC of 0.92 and recall of 0.94. Incorporated radiomics and motion-artifact related features
- Developed a 100% accurate containerized OCR application to extract perfusion biomarkers from 800+ CT sessions

Software Engineer Intern

May 2024 – July 2024

Alpha and Omega Semiconductor Limited

San Jose, CA

- Developed and shipped 3 Java-based logistical dashboard web apps supporting analytics of 100k+ rows OracleDB data
- Built a reusable fuzzy string-matching utility for vendor name matching, enhancing downstream analytics accuracy and addressed lack of text-processing libraries in company's legacy Java environment

PROJECTS

LoRA-tuned LLM with RAG for Rust Q&A | [GitHub](#) | Python, LangChain, Hugging Face, ChromaDB, Llama.cpp

- Fine-tuned Qwen3 0.6B with LoRA on 5k synthetically generated Rust question-answer (Q&A) pairs; quantized the LLM and integrated with RAG to perform grounded Q&A using Rust textbooks through a Streamlit front-end

iOS Motion Analytics App | [GitHub](#) | Swift, Python, PyTorch, PySpark, Airflow, Kafka, Azure EventHub, Snowflake

- Developed an iOS App with CoreMotion and CoreLocation libraries to stream device motion and location data to Snowflake data warehouse via an Azure EventHub data ingestion layer and a Spark-based data processing pipeline
- Trained a transformer encoder model from user collected data to detect risky driving behavior, achieving 97%+ validation accuracy and integrated ML model into an Airflow-orchestrated email analytics pipeline

TECHNICAL SKILLS

Languages: Python, Java, SQL, JavaScript, Bash

ML Tools: PyTorch, TensorFlow, LangChain, Hugging Face, Scikit-Learn, OpenCV, ChromaDB, ONNX, Docker

Full Stack/Data Tools: Node.js, Flask, FastAPI, Django, MongoDB, MySQL, NumPy, Pandas, Spark, Airflow, Kafka, AWS