Richard Wang

408-431-1437 richardwang007@gmail.com

Summary

Extensive experience with state-of-the-art deep learning research and full-stack ML development. Expert in image-to-image, object detection, and time-series machine learning theory and development. Self-motivated, always curious, hardworking, fast learner.

Education

California Institute of Technology (Caltech)

B.S. Computer Science; Class of 2023; Core GPA: 3.9

COURSEWORK:

Computer Science: Computing Systems, ML and Data Mining, Programming Methods, Decidability and Tractability, Data Structures, Compilers and Interpreters, Computer Architecture, Expert Systems, Learning Systems, Vision Neuromechanics Theory and Computation, GPU Programming, Algorithms, Information Theory

Math: Discrete Math, Abstract Algebra, Calculus on Manifolds, Differential Equations, Probability and Stats, Number Theory, Applied Linear Algebra, Probability Models, Mathematical Options Pricing

Experience

MACHINE LEARNING ENGINEER, PERFORMANCE STAR LLC – 2023 - PRESENT

Skills: Generative Models, Time Series, LLM, Fine-tuning, Retrieval, Transformers, ONNX, CUDA, Docker, Kubernetes, MLFlow, Tensorflow, Pytorch, E2E, Distributed Computing, ML Theory

- Part of small-team startup; contributed to SoTA technology that led to acquisition by TEL.
- Developed full ML data pipelines, training, and inference pipelines, servers, and API from scratch using Docker, Kubernetes, MLflow, Tensorflow, Tensorflow Serving, and Flask.
- Led research in generative models for high-dimensional, sparse time-series data from semiconductor fabrication; predictive maintenance and anomaly detection for on-device real time inference.
- Implemented novel physics embedding models and used in LLM fine-tuning with text-prompting for natural language description of complex physical processes.
- Developed novel mathematical learning for modeling chaotic dynamical systems with contrastive learning and spectral analysis
- Research and development into Dynamic Mode Decomposition (DMD), Complex Valued Neural Networks (CVNN), Wavelet analysis, VAEs, transformers, and other SoA deep learning.
- Deployed SoTA models for large-scale user base. Production software for internal and consumer use.

DATABASE ANALYTICS INTERN, BLUESKY – 2023

Skills: SQL, NoSQL, Time Series, Looker, Hex, Cloud Platforms (GCloud, AWS), RESTful

- Contributor at seed-stage startup backed by Greylock, Foundation Capital, and etc.
- Automated query optimization on Snowflake, ML for time series modeling on query metadata.
- Analyzed business metrics with Looker; query-cost analysis with Hex.

MACHINE LEARNING RESEARCH INTERN, MIT LINCOLN LABS – 2022

Skills: Pytorch, Hydra, DVC, Image Segmentation, Video ML, Image-to-image, Security Clearance

- Research and analysis on various SoA semantic segmentation ML models for government projects.
- Designed and implemented image-to-image models using Pytorch.
- Implemented ML pipeline with Hydra and DVC.

MACHINE LEARNING RESEARCH INTERN, THERMO FISHER SCIENTIFIC – 2021

Skills: 3D and 2D vision, Visual Odometry, Visual Reconstruction, Image Techniques, VMWare

- Developed machine learning solutions for non-identical image-mask homography estimation and image manipulation AI.
- Implemented virtual machine automation for ML ops deployment.

DATA ANALYSIS INTERNSHIP, UNIVERSITY OF MICHIGAN ANN ARBOR – 2018-19

Skills: High-throughput screening, PCR, Unsupervised Learning, Wet/Dry Lab

- Analyzed high-throughput compound screening data using statistical methods.
- Discovered a novel compound for cancer therapy and won National Regeneron Scholar Award.

Awards and Publications

- 2024 Publication and patent in preparation: *Siamese Learning and Dynamic Mode Decomposition in Semi-supervised Predictive Maintenance for Recipe-Agnostic Semiconductor Fabrication*
- 2022 National Association of Basketball Coaches (NABC) Academic Excellence Award
- 2021 Presentation at Thermo Fisher Scientific Innovation Day (Poster: *Homography Estimation for Non-Identical Image Mask Pairs*)
- 2019 NuevaHacks NEAR Protocol Best Util (NuevaHacks Prize Winner)
- 2019 Regeneron Scholar (Paper: *Identify Novel Natural Products with Inhibitory Effects on HOXA-9-induced Acute Leukemia through in vitro functional Assays*)
- 2019 Presentation at Pathobiology for Investigators, Students and Academicians Conference
- 2018, 2019 President's Award for Volunteer Work
- 2017 NorCal Sports TV Asian-American All Star (www.youtube.com/watch?v=E67m1DTT-Gk)
- 2016 US Chess Federation High School National Chess Team Champion (https://news.harker.org/harker-students-win-individual-and-team-national-chess-championships/)
- 2015 NorthSouth NorCal Top10 Basketball (https://www.northsouthbasketball.com/college-players)

Volunteer Work and Non-Profit

YOUTH BASKETBALL COACH, HOOPRIGHT – 2024-

- Coaching 6th grade boys basketball for AAU in the Bay Area.

GUITAR INSTRUCTOR, CANON MUSIC – 2024-

- Teaching youth guitar lessons in various genres in the Bay Area.

BASKETBALL COACH, CUPERTINO HOOPS – 2016-19; 2023-24

- Coaching 3rd-6th grade with a local basketball youth league.

Key Skills

Programming: Python, C/C++, PyTorch, Tensorflow, x86-64 assembly, Mathematica, JS, Java, SQL

Math: Functional Analysis, Advanced Algorithms and Discrete Maths, Complexity Theory, Probability and Statistics

Apps/Frameworks: IDA Pro, Wireshark, git, aircrack, MySQL, nodeJS, sqlite, LAMP, MLflow, Kubernetes, Docker, Jupyter

Projects

SELF DRIVING MODEL CAR

Engineered model car with self-driving capabilities using Intel Real-Sense depth camera and ROS. Autonomous navigation, on-device ML. In Python.

COMPILER + INTERPRETER

Compiler for Java as well as an interpreter for Java bytecode (basically a mini JVM). In C.

WEB PROXY

Multi-threaded web proxy in C++. Using a concurrent hash table to implement a LRU cache system.

ML SONAR FISH TRACKING

An Alaskan government project to detect, track, identify, and count fish via machine learning and motion tracking techniques. Deployed real-time ML solutions for in-water sonar data.

MINI OS

Simple monolithic operating system from scratch. File system, kernel threads, different programs. In C.

Personal

Bilingual: English (Native), Mandarin (Native)

Misc:

- Former Caltech Men's NCAA Basketball Team
- Former President of Cooking Club
- CTF and Hackathon Enthusiast
- Former Caltech Quant Trading Club
- Former Caltech Chess Club and National Champion with Harker High School Chess Club
- Guitar and Piano