DILRAJ DEVGUN

New York, NY · dilrajsinghdevgun@gmail.com · 425-279-3444 · www.thejarlid.com

EDUCATION

University of Washington
BS Computer Science *GPA: 3.75*

Seattle, WA
Aug 2015 - Jun 2019

EXPERIENCE

Meta
Senior Software Engineer - XR Eyes

New York, NY
Oct 2023 - Present

- Productionized advanced Eye Tracking Computer Vision/Machine Learning (CV/ML) solutions for next-generation AR/VR devices, enabling novel UX/input modalities through gaze-based interactions
- Led development and initial device bring-up for the Eye Tracking Engine to enable on-device real-time inference to provide gaze data for critical use cases such as user input and third-party integrations
- Implemented Executorch (PyTorch Edge) support in our neural network library to enable optimized on-device inference of ML models for use cases such as Eye/Hand Tracking and LLMs.
- Optimized memory allocations for low-level computer vision algorithms and developed performance profiling tools to identify optimizations, leveraging hardware acceleration to meet memory, latency, and power KPIs for real-time eye-tracking
- Debugged ML model generalization issues during production deployment by implementing debugging tools and performing analysis in notebooks to compare tensors during on/offline execution, identifying training data generalization problems and resolving them through image augmentation
- Implemented various AI agent workflows to automate unit test generation on our Eye Tracking Engine to increase coverage from 50% to >80%

Senior Software Engineer - Instagram Web Server
Oct 2022 - Oct 2023

- Enhanced the performance, scalability, and reliability of Instagram's infrastructure to support billions of global users as part of Web Server team
- Improved IG Django server fleet efficiency by >3% by implementing immortal instances into CPython runtime and optimizing refcount behavior; this is now upstream in official CPython source (PEP-683)
- Optimized threading and fiber memory on IG's server fleet by 66% (600MB/Host) by profiling system memory allocations and fine-tuning pool sizes

Senior Software Engineer - AR Glasses CV/ML Firmware Team
Apr 2021 - Oct 2022

- Developed custom silicon firmware to accelerate on-device Computer Vision and Machine Learning for Orion and future AR glasses
- Adapted Structured Light Depth algorithms from research implementations into high-performance production C/C++ to produce depth maps from raw sensor data in an embedded system

Microsoft
Software Engineer

Redmond, WA
Aug 2019 - Apr 2021

- Software development for custom ASIC on the HoloLens 2 and IVAS Project
- Tripled frame rate for IVAS cameras by implementing shared memory buffer libraries in an embedded memory constrained environment
- Maintained C++ tools for recording and replaying sensor data streams to debug runtime algorithms

PROJECTS

Vibe Dine

www.vibedine.ai

- Built a full-stack application enabling users to query their saved google maps places through natural language using a Next.js/React frontend and a FastAPI backend
- Architected an ingestion pipeline to scrape and index data into Vector PostgreSQL Databases and designed an agentic search/retrieval algorithm that aggregates results to power a conversational AI recommendation agent
- Implemented AI agent leveraging OpenAI LLMs, LangGraph orchestration, and custom tools for retrieval and web search

SKILLS

Proficient Programming Languages: C, C++, Java, Python, Swift, Objective-C, C#