

Huy Hoang

352-235-4466 | huyhoang.developer@gmail.com | [linkedin.com/in/huyhoanggia](https://www.linkedin.com/in/huyhoanggia) | github.com/hwii-florescent | huyhg.dev

EDUCATION

University of Florida

Bachelor of Science in Computer Engineering | GPA: 3.9/4.0

Expected Graduation: May 2027

Gainesville, FL

TECHNICAL SKILLS AND CERTIFICATIONS

Languages: Python, Go, C++, Java, Shell, SQL, Protobuf, TypeScript, JavaScript

Frameworks: FastAPI, Flask, Spring, Node.js, React Native, React

Developer Tools: AWS, Terraform, Kubernetes, Kafka, MySQL, Docker, Jenkins, OpenSearch, Cadence

AI/ML: Model Context Protocol (MCP), LangChain, PyTorch, TensorFlow

Libraries: OpenCV, NumPy, pandas, COLMAP

Certifications: NVIDIA Fundamentals of Deep Learning

EXPERIENCE

Uber Technologies Inc.

May 2026 – July 2026

Software Engineering Intern

- Led search improvements project for an A/B experimentation platform end-to-end, implemented OpenSearch-native sort and filter on Go gRPC endpoints for 120k+ experiments with 27k+ searches/day
- Ran a zero-downtime schema migration in Go across protobuf, MySQL, Kafka, and OpenSearch indices, dual-writing on the live path and backfilling 120k+ production experiments through idempotent Cadence workflows at 99.8% completion
- Built a Go service exposing 7 experiment-management tools over Model Context Protocol (MCP), integrating it into an agentic debugging pipeline that root-causes misconfigured experiments for 1k+ services running experiments
- Architected a Python LLM reviewer for A/B experiment configs across 500 diffs/day, encoding past-incident patterns as lazy-loaded rule packs that cut agent context 50%, and built the eval harness measuring 94% precision and 86% recall

University of Florida – Agriculture Precision Lab

Sep. 2024 – Sep. 2025

Machine Learning Research Assistant

- Developed GPU-accelerated PyTorch pipelines for dense feature matching, increasing match reliability by 24% in dense foliage and improving multi-view tracking stability
- Architected a pipeline to consolidate 1K+ per-frame multi-view COLMAP reconstructions into a unified 3D field-scale model
- Applied monocular depth estimation to densify point clouds, achieving 90%+ data completeness for volumetric analysis

FPT Corporation

May 2025 – Aug. 2025

Software Engineer Intern

- Implemented RESTful APIs in Java Spring Boot with WSO2 mediation policies, deployed with AWS EC2, Docker, and monitored via Grafana and Prometheus, achieving 98% request accuracy and improving response time by 20ms
- Configured and integrated 12+ enterprise microservices using WSO2 API Gateway deployed on Kubernetes, reducing cross-service failures by 20% and increasing system resilience in production clusters
- Developed custom Java WSO2 handlers for payload encryption/decryption and token validation using JWT and OAuth2, improving gateway throughput by 25% and enforcing secure, authenticated API communication across teams

FPT Information System

May 2024 – Aug. 2024

Site Reliability Engineer Intern

- Automated AWS EKS infrastructure provisioning with Terraform and GitOps, improving deployment efficiency by 30%
- Built CI/CD pipelines with Jenkins and Docker, enabling instant rollbacks and cutting deployment time by 90%
- Deployed Prometheus monitoring across 20+ EC2 instances, cutting network setup from 2h to 15min and crash rates by 30%

University of Florida (Programming, DSA, Algorithm Abstract, Dlog, Databases)

Aug. 2024 – Present

Teaching Assistant

- Led mini-lectures and structured office hours for 50+ students; section grades averaged 15% higher than other section

PROJECTS

Horus' Eyes AI Platform (1st Place, SwampHacks) | Terraform, AWS, YOLOv8, Flask, Docker

- Awarded 1st Place for Best Use of Terraform by automating cloud infrastructure for high-throughput ML model deployment
- Engineered an inference pipeline leveraging YOLOv8 and DPT for object and text detection to assist visually impaired users
- Deployed a low-latency Flask + AWS Bedrock pipeline on EC2, executing real-time Text-to-Speech generation

Perch – Agent Development Environment | Rust, WebSockets, React, TypeScript, SQLite, AI harness

- Built a Rust agent runtime integrating 10+ AI harnesses with WebSocket streaming, SQLite history, and PTY terminals
- Engineered SSH/tmux federation with Tailscale for durable remote agents with automatic reconnection and restart recovery
- Built a React/TypeScript interface over a typed WebSocket protocol, packaged with Tauri and enabling mobile agent control

LEADERSHIPS

Software and Electrical Team Lead | Society of Asian Scientists and Engineers (SASE) Engineering Team