SHIHUA YU

youseihuayu@gmail.com | (206) 454-0097 | linkedin.com/in/shihuy

PROFESSIONAL SUMMARY

AI/ML Engineer with hands-on experience in computer vision, natural language processing, and multi-agent systems. Skilled in developing and evaluating deep learning models for multilingual AI applications, 3D robotic simulation, and large-scale data analysis. Proficient in Python, PyTorch, and cloud-based ML pipelines. Proven ability to bridge research and production, with experience in on-device AI model optimization, dataset curation, and cross-functional collaboration with R&D teams. Currently pursuing an MS in AI with a focus on Computer Vision at Northeastern University.

EDUCATION

Master of Science in Artificial Intelligence – Computer Vision *Boston, MA*
Northeastern University Jan. 2026 – Present

- Coursework: Deep Learning, Computer Vision, Machine Learning, Probabilistic Graphical Models

Bachelor of Science in Computer Science Engineering *Seattle, WA*
University of Washington Sep. 2021 – Aug. 2023

- Relevant Coursework: Data Structures & Algorithms, Linear Algebra, Probability & Statistics, Software Engineering

WORK EXPERIENCE

AI/ML Test & Evaluation Engineer *Bellevue, WA*
OnePlus – Global Consumer Electronics Oct. 2023 – Present

- Led on-device AI model evaluation for multilingual voice recognition and image translation across 5+ languages, improving inference accuracy and edge-case coverage.
- Curated and labeled multimodal datasets (text, speech, image) for model fine-tuning; designed systematic evaluation benchmarks for AI feature releases (OnePlus 13/13R).
- Performed root-cause analysis on AI/ML inference failures using log analysis tools (QXDM, QCAT), collaborating with R&D to optimize model-hardware integration on Snapdragon chipsets.
- Built automated testing pipelines for AI feature validation across 5G environments, reducing regression cycle time by 40%.
- Delivered executive-level reports with data-driven KPIs on AI model performance, risk assessment, and release readiness.

Machine Learning Engineer Intern *Bellevue, WA*
WinGsrobotics – Robotics Operating System Jun. 2023 – Aug. 2023

- Developed sensor-driven motion prediction algorithms for robotic arm control using Go and Python, integrating real-time sensor fusion for path planning.
- Built a React-based 3D simulation platform for robotic task visualization, enabling low-code model testing and improving simulation accuracy for ML-driven control models.
- Supported end-to-end integration of AI perception modules (object detection, pose estimation) with hardware control systems.
- Coordinated agile sprints across software, hardware, and AI teams to meet delivery milestones.

Automation & ML Testing Engineer Intern *Bellevue, WA*
T-Mobile – 5G Innovation Lab Mar. 2022 – Jun. 2022

- Programmed robotic arms for automated UI interaction testing on Android devices, applying computer vision techniques for screen element detection and verification.
- Integrated automation frameworks with lab instruments (R&S CMX500, Keysight UXM) for real-time data collection and anomaly detection in 5G network testing.
- Developed data analysis scripts in Python for performance metrics extraction, statistical analysis, and defect pattern recognition.

PROJECTS

AI Language Model Evaluation Platform

OnePlus R&D Mar. 2024 – Oct. 2024

- Designed and executed evaluation pipelines for on-device large language models, testing multilingual NLP capabilities including voice-to-text, text-to-image, and cross-lingual translation.
- Analyzed model edge cases and failure modes using statistical methods; provided actionable insights that improved model robustness by 25%.

Wi-Fi Device ML-Based Testing & Optimization

Personal Project Nov. 2024 – Present

- Developed ML-assisted test case generation for Wi-Fi products (802.11ac/ax), leveraging pattern recognition to identify firmware anomalies and optimize connectivity.
- Automated data-driven triage workflows for defect classification and root-cause analysis with third-party vendors.

TECHNICAL SKILLS

Languages: Python, C++, SQL, Golang, Kotlin, Swift

ML/AI Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, spaCy, OpenCV, scikit-learn

Tools & Platforms: AWS (SageMaker, EC2, S3), Docker, Git/GitHub, Tableau, Azure DevOps, CI/CD Pipelines

Core Competencies: Computer Vision, NLP, Deep Learning, Multi-Agent Systems, A/B Testing, Data Visualization, Statistical Analysis, Model Evaluation & Optimization

Languages Spoken: English, Mandarin Chinese, Korean