Jonathan Muhire

(+1) 572-205-5073 — muhirejonathan123@gmail.com — jonathanmuhire.com — github.com/Jonathan-321

Education

Oklahoma Christian University Expected Apr. 2026
B.S. in Computer Science & M.S. in Artificial Intelligence (Concurrent Degrees) Edmond, OK
GPA: 3.89/4.0 — Completing both degrees in 4 years through advanced standing
Relevant Coursework: Deep Learning, Computer Vision, Robotics, NLP, Distributed Systems, Linear Algebra, Statistical Methods

Experience

Neotix Robotics May 2025 – Present
Founder & Lead Engineer New Haven, CT

- Founded robotics data infrastructure startup; secured $20,000 Yale Tsai City Ventures grant to build scalable data collection and training systems for embodied AI.
- Architected distributed data pipeline processing large-scale multimodal robotics data (RGB-D, proprioception, tactile) using MinIO/LakeFS for Git-like version control and reproducibility.
- Built ROS2-based teleoperation framework for multiple MyCobot manipulators; integrated SAM2 and ORB-SLAM3 into real-time computer vision pipeline.
- *Stack:* Python, C++, ROS2, PyTorch, Docker, MinIO/S3, LakeFS, OpenCV, SAM2, ORB-SLAM3

Scale AI May 2024 – Aug. 2024
AI Safety & Code Evaluation Specialist Remote

- Evaluated frontier language models (GPT-4, Claude) for code generation at $13.8B AI data platform serving OpenAI, Meta, and Microsoft.
- Conducted adversarial testing and red-teaming of model outputs; identified security vulnerabilities including SQL injection, memory leaks, and race conditions in generated code.
- Designed evaluation frameworks for multi-turn coding conversations, specializing in complex algorithmic and async tasks; developed rubrics for improved model steering accuracy.

Projects

smolVLA: Efficient Vision-Language-Action Models June 2025 – Aug. 2025

- Implemented knowledge distillation from OpenVLA (7B parameters) to create efficient VLA for manipulation using structured pruning and quantization techniques.
- Built training infrastructure integrating distributed training (PyTorch DDP) and sim-to-real evaluation; focused on optimizing inference for real-time robotic deployment.

Bimanual Teleoperation & SLAM Pipeline June 2025 – Aug. 2025

- Engineered teleoperation system for collecting dual-arm manipulation demonstrations; implemented ORB-SLAM3 for 7-DOF trajectory extraction from video streams.
- Built annotation tooling for labeling task keyframes and manipulation primitives; created a foundational dataset for training diffusion policies.

RenAIssance: Document Understanding Pipeline May 2025 – Aug. 2025
Google Summer of Code 2025

- Selected for GSoC 2025; built deep learning pipeline for manuscript digitization; fine-tuned LayoutLMv3 Vision Transformers for historical document layout analysis.
- Optimized the inference pipeline for production use, achieving significant reduction in processing time per page using batching and model quantization.

ArtExtract: Large-Scale Image Classification Sept. 2024 – Dec. 2024

- Built CNN-LSTM architecture for artwork classification and style detection across multiple artistic periods; trained on a large-scale art dataset with ResNet50 backbone.
- Implemented neural style transfer and similarity search using learned image embeddings.

Technical Skills

Languages: Python, C++, JavaScript/TypeScript, SQL, MATLAB
ML/AI: PyTorch, TensorFlow, Hugging Face, JAX, Model Compression, Vision Transformers, LLMs, Distributed Training
Robotics: ROS/ROS2, MuJoCo, PyBullet, NVIDIA Isaac Gym, Real-Time Control, SLAM
Computer Vision: OpenCV, SAM2, ORB-SLAM3, Object Detection, 3D Reconstruction
Infrastructure: Docker, AWS S3, MinIO, LakeFS, Git, PostgreSQL, CI/CD