Shayan Gerami

+1 (872) 314-4947 | sgera@uic.edu | linkedin.com/in/shane-gerami | Portfolio

Education

University of Illinois Chicago (UIC) – Chicago, IL **Aug 2023 – Expected Dec 2026**
B.S. in Information and Decision Sciences, Minor: Computer Science
GPA: 3.67/4.00 | Dean's List | Honors College

Relevant Coursework: Data Structures & Algorithms, Machine Learning and Data Mining, Linear Algebra, Discrete Mathematics, Multivariable Calculus, Advanced Statistics & Probability, and Linear Optimization.

Technical Skills

Programming: Python, C++, C, SQL, R
ML/CV Libraries: PyTorch, TensorFlow, scikit-learn, OpenCV, Stone Soup, NumPy, SciPy, pandas
Edge AI & Systems: TensorFlow Lite, ONNX, TensorRT, CUDA, CuPy, Jetson Orin Nano, Linux, Docker, Git, Bash, pytest

Work Experience

Software Engineer Intern, Epsilon Systems Solutions, Inc. (C5I), Dayton, OH **May 2026– Aug 2026**

- **Computer vision and GPU deployment:** Built and fine-tuned a PyTorch U-Net with an ImageNet-pretrained MobileNetV2 encoder for semantic water segmentation; exported models through ONNX and deployed TensorRT-optimized on-device inference on NVIDIA Jetson Orin Nano.
- **CUDA kernel development and acceleration:** Developed custom CUDA kernels and CuPy implementations for per-pixel, camera-pose-aware polarization correction, accelerating image-processing throughput 7× from approximately 1 to 7 FPS on Jetson Orin Nano.
- **Multimodal sensor fusion:** Engineered a real-time, multithreaded Python acquisition pipeline for a drone imaging system, integrating a polarimetric camera, high-rate IMU, and GPS through hardware-triggered synchronization and timestamp alignment.
- **Detection and tracking pipelines:** Developed a target-detection and multi-object tracking pipeline using CA-CFAR, OpenCV image processing, connected components, watershed segmentation, and Stone Soup Kalman filtering and data association to maintain persistent tracks across frames.

Research Experience

Research Assistant, Machine Learning and Embedded AI, RUIC Lab, UIC **Apr 2025– Present**

- Design and optimize computer vision models for embedded platforms constrained by compute, memory, and energy.
- Build training and benchmarking pipelines to evaluate accuracy, operation count, and model size.
- Optimize on-device inference through quantization, sparsity, low-MAC architectures, and memory-aware design.
- Investigate robust ML inference under adversarial attacks, hardware bit flips, and intermittent power.

Publications

First Author — ACM GLSVLSI 2026, *"SENTRY: Spiking Event Reasoning for Selective Deep Inference in Event-Driven Edge Vision"*
First Author — IEEE/ACM DAC 2026, *"Late Breaking Results: ROAST: Reverse-training Offset Attack on Spatial Sampling Topologies"*
First Author — IEEE ACSOS 2025, *"SunSift: Solar-Powered Intelligent Sensing Through Informative Sample Selection"*
First Author — IEEE JETCAS 2025, *"Pixel-Wise Intelligence: Heatmap-Guided Adaptive LBP for In-Sensor Vision Processing"*
Co-Author — IEEE TCASAI 2025, *"ARISE: Adaptive Resource-aware Intelligent Sensing Enabling Intermittent Learning"*
Co-Author — IEEE TC 2025, *"BISen: A Robust Framework for Efficient CNN Inference on Battery-Free Intelligent Sensory Nodes"*

Projects

SunSift: Battery-Free Computer Vision for Edge Devices | Edge AI **IEEE ACSOS 2025**

- Built an end-to-end model-conversion and on-device inference pipeline for battery-free computer vision using C++, exporting PyTorch models through ONNX to TensorFlow Lite Micro on energy-harvesting microcontrollers.
- Implemented fault-tolerant intermittent inference using power-aware scheduling, state recovery, selective MRAM checkpointing, and a six-state voltage-aware controller.

SCENE: Sparse Video Reconstruction and Object Detection | Edge Vision **Manuscript in preparation**

- Implemented system-level evaluation and YOLOV++ object detection on reconstructed ImageNet-VidVRD videos, benchmarking accuracy and resource tradeoffs against full-frame, running-average, and ViBe baselines.
- Evaluated sparse-update configurations across five video resolutions, reducing modeled data transfer from 2.76 to 0.80 MB/frame while matching baseline detection accuracy at 82% mAP@0.5 after fine-tuning.