SRECHARAN SELVAM

📞 +1 917-495-3410 ✉️ sselvam@andrew.cmu.edu 💼 linkedin ⭘ srecharan.com ⭘ GitHub

Education

Carnegie Mellon University **May 2025**
Master of Science, Research: Machine Learning *Pittsburgh, PA*
Relevant Coursework: Machine Learning, Deep Learning, Advanced Computer Vision, Visual Recognition GPA - 4/4

Experience

Kantor Lab, Carnegie Mellon University [GitHub ↗] **Aug 2023 – Present**
Graduate Research Assistant *Pittsburgh, PA*
- Engineered real-time vision system for autonomous leaf grasping integrating YOLOv8 segmentation, RAFT-Stereo depth estimation & custom CNN (GraspPointCNN), featuring self-supervised data generation and weighted decision framework.
- Built self-supervised data generation pipeline using geometric features (SDF, gradients), eliminating manual annotation.
- Finetuned YOLOv8 on custom dataset of **900+** plant images, achieving **68% mAP@[0.5:0.95]** for leaf mask generation.
- Optimized RAFT-Stereo pipeline for precise grasping, attaining sub-pixel accuracy (**<0.5px**) on 1080p stereo pairs.
- Trained GraspPointCNN with spatial attention for grasp point selection, reaching **97.09%** recall and **92.59%** precision.
- Developed a confidence-weighted decision system merging geometric scoring algorithms (70-90%) and CNN predictions (10-30%), improving leaf grasping success rate by **4.66%** over scoring-only baseline across **150** test cases.

Hanon Systems [GitHub ↗] **Jan 2023 – June 2023**
Machine Learning Engineer Intern *Chennai, India*
- Architected a real-time 3D hand tracking and gesture recognition system in Unity AR, enabling over **50** automotive technicians to practice precise component placement and assembly procedures for virtual HVAC systems.
- Implemented Extended Kalman Filter for robust 3D hand tracking, with < **7.5mm** ground truth tracking accuracy.
- Designed geometric analysis pipeline for static hand gesture recognition (grab, pinch, point, open palm) with **97% acc.**
- Trained custom GRU neural network to recognize dynamic hand motions (swipe, circle, wave) attaining < **30ms** latency.
- Optimized hand tracking using ONNX, delivering **33%** faster inference with **50%** smaller model size than MediaPipe.
- Integrated WebSockets between Python backend and Unity frontend, driving **30Hz** data streaming with **<1KB** packets.

Vee Ess Engineering [GitHub ↗] **Jul 2022 – Dec 2022**
Computer Vision Engineer Intern *Chennai, India*
- Developed a multi-camera vision system using YOLOv5 to enhance high-value material recovery amidst low-value debris through material sorting and conveyor speed modulation, demonstrating **96.8% mAP@[0.5:0.95]** detection accuracy.
- Accelerated dataset creation using fine-tuned Mask R-CNN for auto mask generation, creating **43k+** segmented images.
- Boosted detection robustness through custom data augmentation pipeline, incorporating synthetic object compositing.
- Minimized false positives in material counting through worker-interaction filtering, combining ROI-based MOG2 (Mixture of Gaussian) background subtraction with YOLOv5 person tracking and worker-induced occlusion handling.
- Achieved **<15ms** inference latency with **74.5% mAP@[0.5:0.95]** worker detection precision for safety monitoring.

Projects

Vision-Language Understanding via Retrieval & Feedback Alignment [GitHub ↗] **Jan 2025 – May 2025**
- Established a Vision-Language (V-L) Understanding architecture for industrial tool recognition with safety guidance generation, using Qwen-2.5-VL-7B & Llama-3.2-11B-Vision models fine-tuned on custom dataset with safety annotations.
- Implemented PEFT (LoRA) across V-only, L-only, V+L strategies, boosting tool recognition precision from **35%** to **72%**.
- Built RAG pipeline (FAISS & Sentence Transformers) to augment safety guidance, achieving ∼**90%** information accuracy.
- Incorporated RLHF using Group Relative Policy Optimization (GRPO) to align fine-tuned VLMs with structured safety information generation, resulting in **80–85%** of RAG's information accuracy with **30%** lower inference latency.
- Crafted LLM-based scoring system (Gemini API) to assess accuracy/completeness of domain-specific safety instructions.

Multi-Model Stock Prediction with NLP & Automated Trading [GitHub ↗] **Oct 2024 – Feb 2025**
- Built end-to-end automated trading system combining stock prediction (bidirectional LSTM + XGBoost) with real-time FinBERT market sentiment analysis (news/Reddit/SEC), yielding **55-65%** directional accuracy for multiple timeframes.
- Created a dynamically-weighted ensemble model using LSTM (attention-based time-series) and XGBoost (**35+** market features) with adaptive learning rate, attaining MAE of **0.27-0.44%** across intraday timeframes (5m, 15m, 30m, 1h).
- Leveraged FinBERT to process **360+** financial texts, boosting ∼**5%** prediction accuracy using 12 temporal features.
- Simulated trading via Tradier API with risk management, showing **58.5%** win rate and **+0.32%** net return.

Skills

Frameworks: PyTorch, TensorFlow, OpenCV, scikit-learn, pandas, ONNX, Transformers, FAISS, TRL, Unsloth, ROS2
ML Training: SFT, RLHF (GRPO, DPO), RAG, VLM/LLM Fine-tuning, PEFT, Distributed Training, LLM Evaluation
Languages: Python, C++, JavaScript, SQL **Deployment**: AWS (SageMaker, EC2), GCP(Vertex AI), Azure, Docker

Publications

Srecharan Selvam, Abhisesh Silwal, George Kantor **"Self-Supervised Learning for Robotic Leaf Manipulation: A Hybrid Geometric-Neural Approach"** [To be submitted to **CVPR 2026**] (http://arxiv.org/abs/2505.03702 ↗)