SREEHARSHA PARUCHURI

(412) 273-5793 | sparuchu@cs.cmu.edu | linkedin.com/in/sreeharshaparuchur1 | sreeharshaparuchur1.github.io | Google Scholar

EXPERIENCE

Mach9 (YC S21) | Perception Software Engineering Intern — San Francisco, CA
Focus: Computer Vision and Machine Learning Systems — May 2025 - Aug 2025

- **Multimodal Painted Symbol Extraction:** Designed and integrated a multi-stage feature extractor in the **HD Map generation** pipeline, delivering a **50x acceleration** on geometric transformations across RGB-Pointcloud-BEV frames via a custom **CUDA kernel**, and preserving panoptic instancing across multi-view 2D-3D correspondences with a **self-correcting vector-field** and DBSCAN.
- **VLM Inference & Scalable Deployment:** Achieved **91% F1 score** on MUTCD classification of painted symbols in BEV by developing an orientation-robust **RAG pipeline** using VLM embeddings & LLM reasoning (Gemini, GPT-o3). Deployed this software as a high-throughput asyncio system on AWS, sustaining **200+ concurrent requests** via semaphore-based rate control.
- **Transformer Uncertainty Estimation:** Created a tool to quantify **epistemic uncertainty** in polyline prediction through self-calibration on a DETR-style model, reducing customer QA time. Delivered with full **PyTest coverage** and CI/CD-ready integration.

TCS Research - Visual Computing and Embodied AI Group | Applied Scientist — Kolkata, India
Focus: Reinforcement Learning and Audio Perception — Jul 2022 - Jul 2024

- **Audio-Visual RL for Embodied AI:** Led research on developing a novel Actor-Critic approach for **mapless navigation** to an audiogoal in unseen environments trained using PPO and an **Entropy Loss** to encourage exploration in a dense-reward setting. Constructed a state-space encoding a **scene graph** created from depth and binaural **audio spectrogram features** fused using cross-attention.
- **Imitation Learning & Data-Collection:** Setup a **Behaviour Cloning** pipeline within the audio-enabled HabitatSim simulator. This system collected and trained on human demonstrations to bootstrap the RL policy, **reducing online training time by 30%**.

Robotics Research Center (RRC, IIIT-H) | Research Assistant — Hyderabad, India
Focus: Computer Vision and Robotics — Jan 2020 - Jun 2022

- **Geometric Vision & Classical SLAM:** Spearheaded the **sim-to-real** development of a Visual SLAM stack for an autonomous sanitization robot. Processed real-time data from RGBD cameras for marker detection and loop-closures **onboard an Intel NUC**.

PROJECTS

Online Reinforcement Learning for Robotic Foundation Models for Manipulation | Link — CMU

- **Fine-tuned OpenVLA-OFT with GRPO + LoRA**, enabling task adaptation beyond SFT on the sparse-reward LIBERO benchmark.
- Boosted task success from **80% to 98%**, preserving a **100Hz control frequency** by training a decoupled stochastic policy head.

Apple Vision Pro Guided Precision Robotics for Knee Replacement Surgery | Link — CMU

- Achieved **sub-4mm precision** with a KUKA MED7 arm using **SAM2** and ICP to register DICOM imagery with RealSense data.
- Architected a thread-safe **error-recovery system** using **DINOv3** features for visual feedback to counter bone movement in operation.

Physical Stability-Aligned Autoregressive 3D Generation with Direct Preference Optimization | Link — CMU

- Curated a sequence-length-balanced Text2Brick preference dataset enabling **Rejection Fine-Tuning** without human supervision.
- Post-trained LLaMA-1B-Instruct with DPO + Quantized LoRA, reducing **regenerations by 57%** while preserving prompt alignment.

Keyframe selection for reconstruction from video using 3D Foundation Models | Link — CMU

- Engineered a DUST3R-style pipeline for neural pointmap refinement from video, using a **hierarchical keyframe selection** strategy to populate a **Transformer-based memory module** with fused DINOv2 and ViT embeddings to recover per-frame camera pose.

IROS '25 Workshop: Vision-Language-Navigation Challenge | Link — CMU

- Secured **3rd Place** by building a **VLM-based planner** wrapped in a **modular state machine** to interpret natural-language queries in unseen environments, reasoning over semantic and spatial relationships via scene-graph conditioning under the competition time budget.

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Robotics, GPA: 4.10/4.0 — May 2026

- Teaching: Introduction to Deep Learning, Robot Autonomy and Manipulation
- Coursework: Learning for 3D Vision, Generative Artificial Intelligence, Deep Reinforcement Learning, Diffusion and Flow Matching.

International Institute of Information Technology (IIIT-H) — Hyderabad, India
Bachelor of Technology with Honours, Major CGPA: 9.02/10 — Jul 2022

- Awards: Deans Merit List, Undergraduate Research Award.
- Coursework: Statistics in Artificial Intelligence, Applied Optimization, Mobile Robotics, Data Structures and Algorithms.

SKILLS

Languages & Tools: Python, C++, CUDA-Numba, ROS2, Protobuf, Docker, CI/CD PyTest, AWS, GCP, ONNX, Swift, Cursor.
Libraries & Simulators: PyTorch (Lightning, 3D), JAX, OpenCV, Scikit-learn, OpenAI, HabitatSim, Unity 3D, MuJoCo, gRPC.