SREEHARSHA PARUCHURI

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

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Robotic Systems Development (MRSD) GPA: 4.11/4.0 — May 2026

- Teaching: Introduction to Deep Learning.
- Coursework: Learning for 3D Vision, Generative Artificial Intelligence, Deep Reinforcement Learning and Control.

International Institute of Information Technology (IIIT-H) — Hyderabad, India
Bachelor of Technology in Electronics and Communication Engineering (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.

EXPERIENCE

Mach9 | Computer Vision and Machine Learning Systems — San Francisco, CA
Perception Software Engineering Intern — May 2025 - Aug 2025

- **Multimodal Painted Symbol Extraction:** Engineered 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 & LLMs (**Gemini**, **GPT-o3**). Designed this software as a **high-throughput asyncio system** deployed on AWS, employing **semaphores** to stably manage 200+ concurrent API calls and prevent rate-limiting.
- **Uncertainty Estimation:** Created a tool to quantify uncertainty in a **DETR**-style model via self-calibration, and Bayesian dropout, enabling semi-automation of QA workflow. Developed with full unit testing with PyTest for quick **CI/CD** integration.

TCS Research - Visual Computing and Embodied AI Group | Reinforcement Learning — Kolkata, India
Pre-Doctoral Research Fellow — 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 human demonstrations to bootstrap the RL policy, **reducing online training time by 30%**.

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

- **Embedded Robotics & 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**.
- **Depth Estimation for Autonomous Driving:** Adapted and evaluated both **classical vision** (**Structure-from-Motion**) and **neural network-based** disparity-estimation and refinement techniques on data from autonomous vehicles on Indian roads.

PROJECTS

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.

IROS '25 Vision-Language-Autonomy Challenge — CMU

- Secured **3rd Place** by building a **modular state machine** for a vision-language navigation system with **active perception** to decompose a **natural-language query** into subtasks to reason about semantic and spatial relationships and act in an unseen environment.
- Guided **Gemini 2.5 Pro** planner reasoning through Chain-of-Thought prompting, explicit mathematical conditioning and scene graphs.

3D Foundation Models for Neural Reconstruction from Monocular Video | 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** with fused DINOv2 and ViT embeddings to recover per-frame camera pose.

Multimodal Representation Learning for Language and Audio | Publication — IIIT-H

- Applied **BERT-based sentiment analysis** and **k-means clustering** to uncover nuanced links between language and acoustic music features in data scraped from user discourse on mental health related subreddits during COVID-19.

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.