SREEHARSHA PARUCHURI

(412) 273-5793 ◇ sparuchu@cs.cmu.edu ◇ LinkedIn ◇ sreeharshaparuchur1.github.io

EDUCATION

Carnegie Mellon University, School of Computer Science — Pittsburgh, PA
Master of Science in Robotic Systems Development (MRSD) CGPA: 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 GCPA: 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 Generative AI — San Francisco, CA
Perception Software Engineering Intern — May 2025 - Aug 2025

- **Multimodal Painted Symbol Extraction:** Accelerated per-point transformations across Pointcloud, RGB, and rasterized BEV frames by 50× via a custom **CUDA kernel**. Designed a **self-correcting vector-field** and DBSCAN-based clustering algorithm to preserve panoptic instancing from multi-view images across **2D-3D unprojections**.
- **Vision Language Model Inference**: Built an orientation robust **RAG pipeline** for classification of 2D painted symbols using **Gemini text embeddings** and **GPT-o3**, achieving an F1 score of 91% on 10k annotated samples.
- **Uncertainty Estimation:** Created a tool to quantify uncertainty in a **DETR**-style outdoor polyline feature detector via self-calibration, and Bayesian dropout, enabling smoother QA workflow for customers.

TCS Research - Visual Computing and Embodied AI Group | Reinforcement Learning — Kolkata, India
Pre-Doctoral Research Fellow — Jul 2022 - Jul 2024

- **Audio-Visual Navigation**: Developed a **novel on-policy Actor–Critic** model trained with PPO that fused semantic cues from temporal binaural audio with images to improve mapless indoor navigation for embodied AI.
- **CLIP-Enhanced Scene Graphs**: Trained a **contrastive-learning** pipeline to compute **visual–language embeddings**, leveraging graph neural networks to model object–region relationships in cluttered indoor environments.
- **NeurIPS 23 Open Vocabulary Manipulation**: Engineered a **semantic-aware active SLAM** strategy using probabilistic occupancy grids to guide exploration to maximize area-coverage that increased task success by 60%.

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

- **Autonomous Sanitization Robot**: Spearheaded development of the computer vision and **Visual-SLAM** stack to simulate, test, and deploy (**Sim2Real**) algorithms for marker detection in HSV and localization via loop-closures.
- **Real-Time 3D Scene Understanding**: Implemented classical stereo and learning-based monocular depth estimation algorithms for autonomous driving on Indian roads. Developed a ROS package for multi-view **bundle adjustment**.

PROJECTS

- **Augmented-Reality and Robot Assisted Knee Surgery** | Link — CMU
  Achieved a sub-4 mm drilling accuracy with a KUKA MED7 arm through multi-stage pointcloud registration (SAM2 + ICP). Integrated **Apple Vision Pro** for dynamic bone tracking and real-time surgeon-in-the-loop long-horizon planning.
- **3D Foundation-Models for reconstruction from monocular video** | Link — CMU
  Combined DINOv2 semantics and ViT depth encodings to create a rich local-global state representation with adaptive **keyframe selection** to recover camera extrinsics through a view-consistency loss and confidence-aware pointmap refinement through spatio-temporal losses using a **DUST3R-style** architecture.
- **CMU Vision-Language-Autonomy Challenge** — CMU
  Built a ROS state machine for a **vision-language-navigation** model to parse queries using a modified A* planner with **scene-graph** augmented Gemini 2.5 Pro prompts to reason and plan in an unseen environment under a time budget.
- **Music, Mental Health, and Representation Learning** | 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 mental health related subreddits during COVID-19.

SKILLS

- **Programming**: Python, C++, MATLAB, CUDA Numba, ROS2, Java, Go, Protobuf, Django, Swift
- **Frameworks**: Pytorch (Lightning, 3D), Jax, Scikit-learn, OpenCV, OpenAI, Google Genai, Unity 3D, XCode