# SREEHARSHA PARUCHURI

(412) 273-5793 ◇ sparuchu@cs.cmu.edu ◇ LinkedIn ◇ Website

## 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 (11-785)
- 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: Methods in Artificial Intelligence, Applied Optimization, Mobile Robotics, Data Structures and Algorithms

## EXPERIENCE

**Mach9** | Computer Vision and Generative AI | Website — San Francisco, CA
Perception Software Engineering Intern — May 2025 – Aug 2025

- **Multimodal Painted Symbol Extraction:** Designed a **self-correcting vector-field** and DBSCAN-based clustering algorithm to preserve **instance consistency** from multi-view panoptic masks. Accelerated per-point transformations across Pointcloud, RGB, and rasterized BEV frames by **50**× via a custom **CUDA kernel**.
- **Vision–Language Model Inference**: Designed a orientation robust RAG pipeline for classification of rasterized symbols using Gemini text embeddings and GPT-o3, achieving an F1 score of **85%** on **20k** annotated samples.
- **Uncertainty Estimation:** Developed methods to quantify uncertainty in DETR-style vectorized 3D polyline predictions via self-calibration, variance heads, and Bayesian dropout, enabling smoother QA workflow for customers.

**Tata Consultancy Services Research** | Reinforcement Learning and Multimodal Learning — Kolkata, India
Pre-Doctoral Research Fellow — Jul 2022 - Jul 2024

- **Audio-Visual Navigation**: Led the development of an embodied AI agent with **multimodal sensing**, training an **online reinforcement learning policy** with a novel class-agnostic reward, reducing path length by 21%.
- **CLIP-Enhanced Scene Graphs**: Designed and trained a **contrastive-learning** framework to compute **visual–language embeddings**, leveraging **graph neural networks** to model object–region relationships in indoor environments.
- **Open Vocabulary Manipulation, NeurIPS 23**: Developed an **active SLAM** exploration algorithm conditioned on a probabilistic semantic map to maximize area-coverage and led to an improvement of 60% in task success.

**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, **Visual-SLAM** and navigation stack to simulate, build, test, and deploy (**Sim2Real**) an autonomous robot; **finished runners-up** out of 140 teams.
- **Real-Time 3D Scene Understanding**: Adapted stereo and monocular depth algorithms for an autonomous driving setup. Developed a ROS package for multi-view **bundle adjustment**.

## PROJECTS

- **Augmented-Reality and Robot Assisted Knee Surgery** | Website — CMU
  Gathered and analyzed requirements from user studies, market competition, and sponsors to inform system development. Processed 3D and RGB information from the **Apple Vision Pro** to detect bone models in the environment via ICP.
- **3D Foundation-Models for Monocular Video reconstruction** | Report — CMU
  Implemented **semantic-geometric feature fusion** using cross-attention between foundation model embeddings **(DINOv2, Depth Anything)** in a hierarchical state representation to recover camera extrinsics. Devised an adaptive keyframe selection strategy for confidence-aware pointmap refinement using a **DUST3R-style** architecture.
- **Neural-Assisted Depth Disparity Estimation** — Hackathon
  Developed a **coarse-to-fine network architecture** for the OAK-D Pro that improved real-world disparity estimation quality while respecting strict onboard compute and FPS constraints. Ranked in the **Top 25 teams** internationally.
- **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, Pytorch Lightning, Pytorch3D, TensorFlow, Scikit-learn, OpenCV, Unity 3D, LanceDB, XCode