(763) 485-6687
Minneapolis, MN
chand861@umn.edu

Nirshal Chandra Sekar

Robotics Engineer

Portfolio: nirshalchandrasekar
github.com/NirshalChandraSekar
linkedin.com/in/nirshal-chandra-sekar

Graduate student specializing in robotics perception and manipulation, with hands-on experience building deep learning-based perception systems for real-world robotic applications. Strong background in 3D vision, point cloud processing, RGB-D sensing, object detection, semantic segmentation, multi-view geometry, and sensor fusion, with practical experience integrating perception into robot manipulation, imitation learning and control pipelines. Proficient in PyTorch, OpenCV, ROS/ROS2, and simulation environments, with exposure to both research-driven development and industry workflows.

Skills

Programming Languages	Python, C/C++, JavaScript, MATLAB
Libraries / Frameworks	OpenCV, PyTorch, Open3D, PyBullet, MuJoCO, NumPy, scikit-learn
Tools / Platforms	Wandb, Git/GitHub, Linux, Docker, ROS/ROS2, Gazebo, NVIDIA Isaac Sim, Blender

Education

PhD in Computer Science, *University of Minnesota, Twin Cities* — Sep 2025 — Present
Master's in Robotics, *University of Minnesota, Twin Cities* — Sep 2023 — May 2025
B.Tech in Mechanical Engineering, *Vellore Institute of Technology, Vellore* — 2019 — 2023

Technical Experience

Graduate Research Assistant — **Jan 2024 — Present**
Robotics: Perception and Manipulation Lab, University of Minnesota — *Minneapolis, MN*

- Developing **voxel-based 3D scene understanding pipelines** for robotic perception, enabling structured spatial representations that support downstream **manipulation, grasp planning, and reasoning**.
- Built an end-to-end **Grasp Imitation** pipeline that extracts **object-centric visual features** and **3D human hand-pose trajectories** from monocular video, enabling robots to imitate grasp behaviors directly from **human demonstrations**.
- Engineered a **segmentation-guided grasp generation system** integrating **Segment Anything (SAM)**, **Contact GraspNet**, and **RealSense L515 LiDAR** to produce physically feasible grasp candidates from perception outputs.
- Designed a **human demonstration–guided object-part grasping network** that leverages visual cues from a **single demonstration video** to infer graspable object parts on **previously unseen objects**.
- Performed precise **multi-camera intrinsic and extrinsic calibration** to align multiple RGB-D sensors, ensuring accurate cross-view geometry for manipulation experiments.

Software Engineering Intern — **May 2024 — Dec 2024**
Nilfisk — *Brooklyn Park, MN*

- Utilized **NVIDIA Isaac Sim's Replicator** to generate synthetic datasets for detection of industrial scraps in factory environments.
- Validated sim-to-real transfer by evaluating trained detection models on real-world data, achieving a **92% mAP**.
- Developed **custom annotation tool** using the **Segment Anything Model** to support object detection and segmentation pipelines.
- Streamlined data annotation workflows to improve labeling efficiency and data quality for machine learning models.

Technical Projects

Bi-Manual Manipulation using Diffusion Policy Github
- Designed and executed a **vision-based (CNN) diffusion policy** for bottle uncorking using PyTorch and dual UR5e arms.
- Trained on 188 teleoperated demos via L515 LiDAR and D405 stereo cameras, using ROS for data collection and action execution.
- Achieved a **74.7% task completion rate** across 30 rollouts, showcasing effective deployment of learned policies.

Depth Video Diffusion for Robot Policy Learning Github
- Built a **conditional depth video diffusion model from scratch** using a lightweight **3D U-Net** with spatial and temporal attention.
- Designed a **multi-modal conditioning pipeline** combining **Sentence-BERT** text embeddings and a CNN-based **RGB-D** encoder.
- Integrated **FiLM-based conditioning** to modulate 3D U-Net feature maps with fused text and visual embeddings.
- Implemented a scalable **DDPM training and sampling pipeline** with **multi-GPU parallelization**.
- Evaluated generation quality using **Fréchet Video Distance (FVD)** and achieved 10-frame inference in **1.5 minutes per GPU**.

3D Semantic Reconstruction Paper
- Performed **3D semantic reconstruction** using **Structure from Motion (SfM)** and **Multi-View Stereo (MVS)** with **COLMAP**.
- Conducted **2D semantic segmentation** with **YOLOv8** and linked 2D points to 3D points via a voting process.
- Generated a fully labeled **3D triangle mesh model** with **86% semantic labeling accuracy** across the reconstructed surface.