

Avnish Patel

+1-(857)693-9988 | patel.avni@northeastern.edu | Boston, MA

[Portfolio](#) | [Github](#) | [LinkedIn](#)

Education

Northeastern University, Boston, MA

Sept 2022-Dec 2024

Master Of Science in Robotics | Electrical & Computer Engineering Concentration

Coursework: Robot Sensing, Autonomous Field Robotics, Advanced Computer Vision

Publication ([Link](#))

- B. Maxwell, A. Patel, "Logarithmic Lenses: Exploring Log RGB Data for Image Classification", CVPR 2024

Experience

Research Student – Computer Vision Lab | Northeastern University, Boston

May-Dec 2023

- Researched raw log RGB data's impact on deep networks like ResNet-18, improving classification performance and robustness to intensity and color variations, on novel RAW10 dataset (10k DNG & JPG images each, 10 categories)

Research Assistant - AirLab | Carnegie Mellon University, PA ([GitHub](#))

July 2025-Present

- Integrated a Relative Pose Graph pipeline in ROS2 in C++ using GTSAM Fixed-Lag Smoother in the IMU Preintegration module of the open-source SuperOdometry codebase, to improve ATE/RPE and reduce long term drift in SLAM

Medtronic

Jan 2024-Apr 2025

Surgical R&T Machine Learning Engineer, Boston

- Ground Truth (G.T.) Generation (Python) with Camera Calibration

- Built an end-to-end SLAM pipeline using stereo-rectified, semantically segmented images in DROID-SLAM for dense depth estimation in surgical videos and scene understanding of anatomy and surgical tools
- Optimized camera trajectory using GTSAM non-linear optimization and refining 3D reconstruction with Bundle Adjustment and LightGlue feature matching and developed a Open3D based offline rendering pipeline
- Developed a real-time Ground Truth pose estimation pipeline using OptiTrack camera capture and robot kinematics with PnP and ROMA feature detection for training deep learning models on instrument articulation
- Automated Endoscope Stereo Camera Calibration using Zhang's method on a Charuco board with a UR3 Robot Arm
- Implemented a custom pipeline to generate synthetic images with varying focal lengths and distortions, enabling regression of camera intrinsic parameters using a pretrained Transformer model
- Deep Learning (PyTorch)
- Segmented hernias in 10,000 medical images using Swin Base Transformer with PyTorch DDP for multi-GPU training, tracked experiments via MLflow, and deployed the optimized ONNX/TensorRT model on NVIDIA Holoscan
- Developed a YOLOv8-based pipeline for precise detection of surgical instrument tips from medical images in real-time
- Applied Monocular Depth Estimation to get metric distance between two instruments from an image by Depth Anything
- Implemented a PyTorch wrapper with Optical Flow on FAST API using Unimatch, deploying models to ONNX and TensorRT for 10x reduction in real time annotation of medical image frames with 1-second latency

Kisan Drip Irrigation Pvt Ltd

Aug 2020-Aug 2022

Artificial Intelligence Engineer, India

- Integrated ElasticFusion: RGB-D SLAM with C++ to align multi-view point clouds from Intel RealSense D455 cameras, enabling accurate 3D Reconstruction for pipe inspection and defect analysis
- Experimented with PointNet-based deep learning models in Python for point cloud classification to enhance complex defect identification, achieving a 30% improvement over traditional 2D vision methods
- Implemented YOLO-based object detection to localize drippers in pipe assemblies, enabling precise hole punching
- Deployed the 3D vision pipeline as a containerized FastAPI service integrated into on-premises manufacturing workflows

CUDA: Flash Attention in C++ (currently implementing other algorithms in CUDA) ([GitHub](#))

Aug-Sept 2024

- Optimized forward pass of Flash Attention with CUDA reducing memory usage and increasing speed significantly

3D Gaussian Splatting for Scene Representation and Rendering (PyTorch)

Jan-Feb 2025

- Implemented a 3D Gaussian Splatting pipeline with differentiable rendering to train 3D scene representations from multi-view posed images and achieve novel view synthesis on the Cow dataset

3D Object Detection & Tracking with Late Sensor Fusion on Waymo Dataset (PyTorch) ([GitHub](#))

Feb-Mar 2025

- Built a 3D object detection and multi-object tracking LiDAR-Camera Late Fusion pipeline using FPN-ResNet with Bird's Eye View (BeV) from LiDAR, along with Extended Kalman Filter (EKF) based filtering and nearest-neighbor data association
- Generated a globally consistent semantic 3D map by fusing LIDAR point cloud data with vehicle pose transformations
- Augmented Llama 3 (via Ollama) to generate real-time spoken natural language descriptions of the 3D environment

Real-Time SLAM Mapping with ROS2, RPLIDAR & IMU Integration on Jetson Nano (C++)

May-June 2025

- Deployed 2D SLAM using ROS2 Humble Cartographer on NVIDIA Jetson Nano with RPLIDAR and IMU integration with Allan variance to correct drift, enabling real-time environment mapping and localization