

# AVNISH PATEL

+1 8576939988 | [Email](#) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#) | Boston, MA

## PUBLICATION [\[LINK\]](#)

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

## EXPERIENCE

### Research Assistant - AirLab | [Report](#) | Carnegie Mellon University, Remote July 2025 – Present

- Integrated *Relative Pose Graph Optimization* in ROS2 in C++ using *GTSAM Fixed-Lag Smoother* in the *IMU Preintegration* module of the *multi-modal IMU-LiDAR sensor fusion* to reduce long term drift in SLAM
- Achieved 35.8% lower ATE and 52.5% RPE on the *SubT-MRS Laurel Cavern dataset* with *Velodyne LiDAR*
- Executed trajectory mapping using *Livox LiDAR* and *IMU* sensors on the *Unitree G1* robot, applying a *low-pass filter* to mitigate *IMU bias* and enhance mapping accuracy
- Computed *Allan deviation* from *rosbag* data to estimate accelerometer and gyroscope noise parameters (*random walk*, *white noise*) for integration into the *Unitree G1* robot mapping pipeline

### Medtronic | Surgical R&T Machine Learning Engineer, Boston Jan. 2024 – Apr. 2025

Ground Truth Generation (Python) with Camera Calibration & Deep Learning (PyTorch)

- 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*
- 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*
- Implemented *YOLO-based object detection* to localize drippers in pipe assemblies, enabling precise hole punching
- Applied *Monocular Depth Estimation* using *Depth Anything* model to get metric distance between two instruments
- Implemented a *PyTorch* wrapper with *Optical Flow* on *FAST API* using *Unimatch*, deploying models to *ONNX/TensorRT* for 10x reduction in real time annotation of medical image frames with 1-second latency

### Research Student – Computer Vision Lab | Northeastern University, Boston May. 2023 – 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)

### Kisan Drip Irrigation Ltd. | Artificial Intelligence Engineer, India Aug. 2020 – Aug. 2022

- 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 the *YOLO-based object detection* to localize drippers in pipe assemblies, enabling precise hole punching
- Deployed the *YOLO pipeline* as a *Docker containerized FastAPI* service integrated into manufacturing workflows

## PROJECTS

### 3D Detection & Tracking with Late Fusion | PyTorch | [GitHub](#) June 2025 – July 2025

- Built a *3D object detection* and *multi-object tracking LiDAR-Camera Late Fusion pipeline* on the *Waymo Dataset* using *FPN-ResNet* along with *Extended Kalman Filter (EKF)* based filtering
- Generated a *globally consistent semantic 3D map* by fusing *LIDAR point cloud data* with vehicle pose transforms
- Augmented *Llama 3 (via Ollama)* to generate real-time spoken natural language descriptions of the 3D environment

### 3D Gaussian Splatting for Scene Representation and Rendering | PyTorch May 2025 – June 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*

## EDUCATION

### Northeastern University

Masters of Science in Robotics, ECE Concentration

Boston, MA

Sep. 2022 – Dec. 2024

- Relevant Coursework:** Autonomous Field Robotics, Advanced Computer Vision