

Gaurav Raut

240-413-4120 | gauraut14@gmail.com | [linkedin.com/in/gauraut](https://www.linkedin.com/in/gauraut) | github.com/gauraut | gauravrout.com

Domain Skills: Robot Perception, Computer Vision, Sensor Fusion, Calibration, Deep Learning, Artificial Intelligence
EDUCATION

University of Maryland

Master of Engineering, Robotics

College Park, MD

Aug. 2021 – May 2023

University of Pune

Bachelor of Engineering, Mechanical

Pune, India

Aug. 2016 – May 2020

PROFESSIONAL EXPERIENCE

ArcBest

Robotics Software Engineer, Perception II

Fort Smith, AR

June 2024 – Present

- Engineered a scalable, cloud-based VLM inference pipeline (FastAPI, Docker, Kubernetes) that leverages GPT-4V prompts to classify pallet condition and stackability with 85% accuracy and real-time performance.
- Integrated RT-DETR for detection and ByteTrack for data association to track pedestrians, forklifts, and other dynamic objects on a custom warehouse dataset, achieving a MOTA of 85.2%, IDF1 of 88.4%, and real-time performance at 22 FPS on NVIDIA Jetson Orin NX.
- Established a state machine-based perception pipeline for dimensioning and barcode scanning on edge devices
- Built an autonomous barcode detection pipeline using a PTZ camera for 3-DOF scanning
- Worked on autonomous stacking of automotive freights using LiDAR-Camera fusion
- Led the summer interns as a mentor for the project of large-scale 3D reconstruction

Robotics Software Engineer, Perception I

May 2023 – June 2024

- Implemented a pallet detection network to detect and localize pallets and their pockets
- Established Azure-based MLDC pipeline to store datasets, train models, and deploy
- Worked on detecting automotive parts freight with abnormal sizes using 3D point cloud data.
- Made CUDA accelerated point cloud filtering pipeline using cuPCL and custom-made CUDA compatible filters.
- Built algorithms for dimensioning of standard and non-standard freights.

Robotics Software Engineer, Perception Intern

May 2022 – May 2023

- Edge device deployment of deep learning algorithms using ONNX.
- Implemented a complete ROS-compatible stack for real-time ground points segmentation.
- Built algorithms for safety alert in the robot's 3D environment.
- Developed algorithms using computer vision and deep learning to extract 3D information of objects.
- Created a custom dataset to train deep learning model using transfer learning.
- Founded and implemented methods for 2D-3D extrinsic calibration for ToF cameras.

University of Maryland

Independent Researcher with Dr. Pratap Tokekar

College Park, MD

Aug. 2022 – Dec. 2022

- Researched various multi-object tracking paradigms.
- Developed a Kalman filtering-based tracking-by-detection algorithm using YOLOv5 and SORT techniques.
- Tested and evaluated the performance of novel Transformer-based tracking network on MOTs metrics.

TECHNICAL SKILLS

Languages: Python, C, C++, MATLAB, Bash

Libraries and Tools: PyTorch, TensorFlow, OpenCV, PCL, Open3D, ROS/ROS2, git, gtest, Docker, CUDA, ONNX, Kubernetes

Hardware: 3D/2D LiDAR, Stereo and ToF depth cameras, RGB cameras

PUBLICATIONS

- Raut, Gaurav, and Apoorv Singh. "Generative AI in Vision: A Survey on Models, Metrics and Applications." arXiv preprint arXiv:2402.16369 (2024).
- [*ECCV'24*] Singh, Apoorv, Gaurav Raut, and Alka Choudhary. "Multi-agent Collaborative Perception for Robotic Fleet: A Systematic Review." arXiv preprint arXiv:2405.15777 (2024).
- [*ICMI'24*] Raut, Gaurav, and Advait Patole. "End-to-End 3D Object Detection using LiDAR Point Cloud." 2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI). IEEE, 2024.
- Sunil Ranmale, Gaurav Raut, Chaitali Sawant, Jyotiraman De, Viraj Kalyani, Akash Patil. *Apparatus And A Method For Measuring Slackness*. Patent, IN 201921045820, May 14, 2021.

RELEVANT PROJECTS

VLM-Enabled Multi-View AV Scene Understanding Pipeline | *Docker, AWS, GPT-4V, nuScenes*

- Built a serverless Docker/AWS pipeline for synchronized left, front, and right camera feeds, using GPT-4V and nuScenes to classify weather, time-of-day, and detect traffic anomalies in real time ($\geq 85\%$ accuracy).

Aerial Multi-Object Tracking with RT-DETR & ByteTrack | *Python, PyTorch, CUDA, ROS2*

- Developed drone-based tracking pipeline using RT-DETR and ByteTrack on the VisDrone aerial dataset, achieving MOTA of 72.5%, IDF1 of 76.3%, and 18 FPS real-time inference on NVIDIA Jetson Orin NX.
- Tracked multiple classes (pedestrians, vehicles) at varying altitudes, boosting overall tracking accuracy by 12% over baseline.

PFE3DNet: End-to-End 3D Object Detection using LiDAR point cloud | *Python, PyTorch* [Paper link](#)

- Created a 3D Object detection network using Pillar Feature Encoder, a custom ResNet backbone, and 3D SSD detector to output 3D bounding box and classes.
- Network outperformed the original PointPillars implementation by 5.61 mAP points, achieving SOTA performance on the KITTI 3D Object detection dataset.

Structure from Motion | *Python, PyTorch*

- Developed an SfM module that reconstructs the 3D scene and estimates the ego-motion using image sequences.
- Modified and improved the unsupervised SfMLearner network by implementing epipolar photometric loss, batch normalization, and data augmentation.

Automatic Manga Colorization | Trained a U-Net/GAN-based image colorization network on a custom dataset.

Stereo Vision | *Python* Built a module to construct depth and disparity maps from two vantage points.

Face Swap | *Python, TensorFlow* Implemented a face swapping module using classical (Delaunay Triangulation and Thin plate spline) and Deep Learning (Position Map Regression Network).

AutoCalib | *Python* Constructed a camera calibration module using Zhang's technique and non-linear optimization.

AR Tag Detection and Re-projection for Augmented Reality Applications | *Python*

MyAutoPano | *Python* | Fabricated an automatic panorama generation tool using classical (Feature matching and Homography estimation) and deep learning (HomographyNet) techniques.