

# Kaivalya Shah

Computer Vision – Robotics – ROS2 – Edge AI

Satellite, Ahmedabad, Gujarat | +91 79840 81876 | kaivalyashah192@gmail.com | GitHub | LinkedIn

## EDUCATION

### Pandit Deendayal Energy University (PDEU)

B.Tech in Computer Science & Minor in Robotics; CGPA: 8.7

Gandhinagar, Gujarat, India

July 2022 – Present

### C.N. Vidhya Vihaar (GSEB) – Gujarati Medium

Class 12: 76%, Class 10: 81% – JEE Mains Percentile: 95.5

Ahmedabad, Gujarat, India

schooling 2010 – 2022

## TECHNICAL SKILLS

- **Programming Languages:** Python, C++, Java, R, MATLAB
- **Libraries & Frameworks:** Asyncio, ZeroMQ, PyTorch, TensorFlow, Open3D, Flask, NumPy
- **Computer Vision & AI:** CLIP, ViT, VLM/LLMs, NeRF, LoFTR
- **Theory:** Robotics and Control Systems
- **Edge Device Optimization:** Jetson Orin NX
- **Hardware:** Arduino, ESP32, Raspberry Pi
- **Technologies:** ROS, Gazebo, SLAM, Nav2, CUDA, WebSockets
- **Tools:** Docker, Conda, Git, Fusion 360 (URDF)

## EXPERIENCE

### Computer Vision Intern

Sastra Robotics (Startup) @ IITGN

Sep 2024 – Present

Gandhinagar, India

- Optimized multiple AI vision modules for edge device deployment on Jetson Orin NX
- Developed Modular Gripper System achieving real-time communication and node-like architecture with TCP and Web sockets
- Integrated low-level motor control between ESP32 and Ethernet Module via UART/Serial and socket communications

### Research Intern

IITGN Robotics Lab

May 2024 – Jul 2024

Gandhinagar, India

- Led development of a real-time 6D pose estimation pipeline with less than 5 cm error on live objects
- Created 3D reconstruction toolchain using RGB-D input, improving scene fidelity by 20% over baseline

### ROS Developer

iNav (Startup) @ IIC/PDEU

May 2023 – Jul 2023

Gandhinagar, India

- Developed ROS packages for autonomous car prototype

## PROJECTS

### Gripper System with (SDK + UI)

Modular Gripper Control and Vision System

ZeroMQ, Asyncio, Multi-Threading  
Well optimized and tested industrial coding

- Modular, node-based system for controlling and monitoring a robotic gripper
- Real-time visual feedback and perception for interactive robotic manipulation
- UI interfaces and SDK for users to control and system visualization

### Grasp Detection for Multiple Objects

6-DOF Grasp Generation on Edge Device

PyTorch/torchvision, Open3D(TSDF), Docker

- 6-DOF grasp detection in cluttered scenes using point clouds from a RealSense depth camera
- Learning-based grasp candidate generation and quality evaluation via dockerized API on an edge device
- Neural surface representations to enable robust any-view grasping in dynamic environments

### Live Pose – 6DOF Pose Estimation System + ROS package

Real-Time Pose Estimation using Depth Camera

Transformers, Pytorch, Pyrealsense2, Docker

30+ GitHub Stars ★

- Python package for live 6-DOF pose estimation using Intel RealSense depth cameras using CAD models
- Cross-platform support with Dockerized deployment on Jetson and native Windows compatibility

## Object Reconstruction

NeRF, LoFTR, XMem  
15+ GitHub Stars ★

### 3D Reconstruction from RGB-D Data

- Pipeline for object-level 3D reconstruction using RGB-D data from RealSense depth cameras
- Interactive mask-based object segmentation and boundary selection for accurate shape recovery

## Real-Time Voice Activity Detection (VAD)

Silero, PyTorch, SoundDevice

### Microphone-Based Speech Detection System

- Real-time voice activity detection tool using a lightweight deep learning model for microphone input
- Terminal-based live speech/silence feedback with support for customization and chunk-level analysis
- Modularity to allow future integration of noise suppression, speaker diarization, or audio logging

## Object Property Detection

VLMs, LLMs, CLIP, SAM2

### Physical Property Estimation from Visual Data

- System to identify objects and estimate material, dimensions, weight, and surface friction from images
- Compared two visual reasoning pipelines for balancing accuracy, runtime, and resource efficiency
- Fast, on-device property inference for robotic grasp planning and manipulation tasks

## Multi-Purpose Differential Drive Robot

ROS2, Gazebo, SLAM Toolbox, YOLO, MiDaS, RViz

### ROS2-Based Simulation and Perception System

- Versatile ROS2 package for simulating a differential drive robot in Gazebo with modular sensors
- Integrated SLAM, RGB-D vision, and point cloud generation with perception modules like YOLO and MiDaS
- Seamless teleoperation and real-time environment mapping for indoor navigation tasks

## ROS2 Perception Modules – YOLO and MiDaS Integration

ROS2, YOLOv8, MiDaS, RViz

### Real-Time Vision Pipelines for Robotics

- Modular ROS2 packages for real-time object detection and monocular depth estimation using YOLOv8 and MiDaS
- Integrated vision nodes with simulated robot camera streams from the bot\_spawn package in a Gazebo environment
- Published detection and depth topics for visualization in RViz, enabling plug-and-play support for robotic perception pipelines

## Hardware Projects

Embedded C, Arduino, ESP32/ESP8266, PID Control, MediaPipe

### Microcontroller-Based Robotics and Control Systems

- Gesture-controlled car using computer vision and ESP32–ESP8266 communication for wireless motor control [GitHub]
- Wall maze solver using a PID algorithm with obstacle avoidance [GitHub]
- Line Following Maze solver robot capable of solving printed mazes with sharp turns and junction decisions [GitHub]
- Simulator for line-following robot logic to test algorithms virtually before hardware deployment [GitHub]

## ACHIEVEMENTS

---

- Tech-Head @ CRETUS - Robotics and Automation Club, PDEU
- 2nd Prize winner - Maze Solving Bot @ MindBend, SVNIT, Surat
- Runner Up - Line Following Bot @ Techfest, IITB, Bombay
- Runner Up - Hexapod @ Robofest, Gujarat

## CERTIFICATIONS

---

- NPTEL – Understanding Incubation and Entrepreneurship
- NPTEL – Basic and Intermediate Level Spoken Sanskrit
- Industrial Drones: Theory and PID Controllers