

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: 8.76

Gandhinagar, Gujarat, India

July 2022 – Present

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

JEE Mains Percentile: 95.5

Ahmedabad, Gujarat, India

schooling 2010 – 2022

TECHNICAL SKILLS

- **Programming Languages:** Python, C++, Java, R, MATLAB
- **Frameworks:** PyTorch, TensorFlow, LangGraph, Open3D
- **AI & CV:** CLIP, ViT, VLM/LLMs, SAM, GroundingDINO
- **Theory:** Machine Learning, Deep Learning, Controls
- **Edge Device Optimization:** Jetson Orin NX
- **Hardware:** Dynamixel Motors, ESP32, Raspberry Pi
- **Technologies:** ROS 2, CUDA, SLAM, WebSockets
- **Tools:** Docker, Conda, Git, Fusion 360 (URDF)

EXPERIENCE

Lead Engineer

DexSent Robotics - Startup @ IITGN Robotics Lab

August 2024 – Present

Gandhinagar, India

- Designed agentic AI system for vision-grounded humanoid task planning.
- 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

SRIP Research Intern – Agentic AI for Robotics

IITGN Robotics Lab

May 2025 – Jul 2025

Gandhinagar, India

- Built agentic AI pipeline integrating GPT-4, VLMs, and LangGraph for natural language robot control
- Enabled bi-manual execution and fallback re-planning using multi-turn reasoning
- Developed modular ROS 2 nodes and Python SDK for educational robot platform

PROJECTS

Vision-Guided Robotic Manipulation (VGR Module)

End-to-End Vision-to-Action Pipeline for Industrial Tasks

ROS, OpenCV, PyKDL

- ROS-based module that converts camera detections into calibrated robot poses using pixel-to-mm mapping and pose transforms
- UDP-based trigger system and adaptive planning node for sequential multi-object pick-and-place, CNC tending, and kitting
- Real-time vision bridge, trajectory planners, and homing logic to enable autonomous operation with DexSent's robotic arm

Shoonya Recycling – Battery Classification Vision System

Multi-Level AI Classification for Mobile Li-ion Batteries

PyTorch, PaddleOCR, EfficientNet

- **94.4% accuracy** in Level-1 classification (Polymer, Non-Polymer, Waste) using dual-camera vision inspection.
- **Multi-branch pipeline** logo detection (YOLOv8), OCR with brand lexicon (PaddleOCR), and ROI-CNN (EfficientNet-B0).
- Fusion logic for **Level-2 brand classification**, targeting **85%+ accuracy** across 10+ mobile brands.
- Optimized runtime with ONNX/TensorRT for conveyor deployment, ensuring **200 ms latency** at industrial throughput.

Agentic AI for Half-Humanoid Robot

Reasoning-Driven Bimanual Robot Intelligence

LangGraph, GPT-4, VLMs

Agent-based multi-turn execution pipeline

- Designed dynamic reasoning loop using LangGraph linking GPT, VLM, and control APIs
- Enabled natural language control with fallback logic, memory, and task decomposition
- Vision-grounded manipulation using VLMs (SAM + GroundingDINO) with bimanual execution