

Ashish Sukumar

Worcester, MA, USA

☎ +1 (508)-502-8692 • ✉ asukumar1@wpi.edu • 🌐 yamiportfolio.netlify.app • 🌐 Yami1106
in ashish-sukumar-086663272

Professional Summary

Computer Vision and Deep Learning engineer pursuing M.S. in Robotics Engineering at WPI (GPA 3.83, May 2027). Experienced building end-to-end systems spanning **object detection, monocular depth estimation, optical flow, visual-inertial odometry, 3D scene reconstruction (NeRF, SfM), homography estimation, and autonomous driving perception**. Strong foundations in **PyTorch, camera geometry, sensor fusion, and multi-modal learning**. Published researcher (ICIoT 2025) with industry experience at e-Yantra, IIT Bombay. Seeking Summer 2026 US Deep Learning / Computer Vision internship with full-time conversion potential.

Education

Worcester Polytechnic Institute

M.S. Robotics Engineering

GPA: 3.83/4.00 Expected Graduation: May 2027

Relevant Coursework: Computer Vision (RBE 549), Deep Learning for Perception, Motion Planning, Robot Dynamics, Foundations of Robotics

SRM Institute of Science and Technology

B.Tech Computer Science & Engineering

CGPA: 9.54/10 Excellence Award for Outstanding Student & Best Class Representative

Worcester, MA

2025–Present

Chennai, India

2021–2025

Technical Skills

Languages: Python (primary), C++, C, Bash

DL Frameworks: PyTorch (primary), TensorFlow, Keras, scikit-learn, NumPy; CUDA, TensorBoard

Computer Vision: OpenCV, YOLOv8, YOLO-World, ByteTrack, Depth Anything V2 (ViT-L), HybridNets, RAFT Optical Flow, EasyOCR, YOLOv8-Pose, Camera Calibration, Epipolar Geometry, Homography Estimation

3D Vision & SLAM: NeRF, SfM, COLMAP, Bundle Adjustment, Pose Estimation, Stereo Vision, S-MSCKF, EKF, Sensor Fusion, Point Clouds

Architectures: CNN, ResNet / DenseNet / ResNeXt, ViT / Transformer, Spatial Transformer Networks, Bi-LSTM with Attention, Multi-scale Encoders, Gated Fusion

3D Rendering: Blender (scripted scene generation, Python API), Differentiable Rendering

Robotics: ROS2, ROS, MoveIt, OMPL, Webots, Gazebo, RViz

Tools: Docker, Git, Linux, TensorBoard, Jupyter, Tableau, Fusion 360, AutoCAD

Experience

e-Yantra, IIT Bombay

Junior Project Technical Assistant

- Developed **computer vision and image processing** pipelines in **Python and OpenCV** for autonomous robot perception within competitive robotics environments (eYSRC national competition).
- Built simulation environments with autonomous task execution pipelines in **Webots**, integrating perception with decision-making for navigation and obstacle avoidance.
- Formally recognised by **Prof. Kavi Arya (PI, e-Yantra, IIT Bombay)**; conducted two technical workshops on Embedded Systems and Robotics; produced structured documentation deployed nationally.

Chennai, India

Jun 2024–Feb 2025

Robocare Lab, WPI

Voluntary Research Assistant

- Contributed to **multimodal perception and HRI pipelines** for the SoftBank **Pepper Robot**, integrating speech, gesture, and visual feedback via **ROS2**.
- Configured and tested onboard camera, microphone, and joint actuator systems for controlled lab experiments.

Worcester, MA

Oct–Dec 2025

WPI Small Business Digitization Initiative (SBDI)

Digital Consultant

- Independently scoped, designed, and delivered end-to-end digital solutions for small business clients, managing full project lifecycle with minimal supervision.

Worcester, MA

Oct–Dec 2025

Key Projects

Einstein Vision: Built a modular **autonomous driving perception pipeline** processing dashcam video through **YOLOv8 + ByteTrack** multi-object detection and tracking, **Depth Anything V2 (ViT-L)** monocular depth estimation, **HybridNets** lane detection, and **RAFT optical flow** for vehicle state estimation. Implemented **ego-motion compensation** using Sampson distance and homography-based verification to distinguish moving vs. parked vehicles; added brake light detection via HSV thresholding and **3-phase collision prediction**. Rendered full 3D scenes procedurally in **Blender** with color-coded vehicle states, traffic signals, and **YOLOv8-Pose** pedestrian visualization across 13 real-world driving sequences.

Deep VIO: Implemented full **S-MSCKF** classical VIO pipeline achieving **0.12 m RMSE ATE** on EuRoC MAV benchmark. Designed three deep learning models (vision-only, IMU-only, **visual-inertial gated fusion**) using **multi-scale CNN encoders** and **bi-directional LSTM with temporal attention**. Built a synthetic Blender VIO dataset (20 scenes, 10K poses each at 100 Hz). Applied **pose graph optimization** with loop-closure and smoothness constraints, reducing test ATE from 1.94 m to **1.18 m (39% improvement)**.

NeRF: Implemented complete **NeRF** from scratch in **PyTorch**: positional encoding (sinusoidal, 63-dim), hierarchical coarse-fine sampling (64+128 pts/ray), volume rendering, and white-background compositing. Achieved **27.42 dB PSNR / 0.9084 SSIM** (Lego) and 25.75 dB / 0.799 (Ship). Extended to custom real-world datasets using **COLMAP**-based SfM for camera pose extraction.

HomographyNet: Implemented supervised (**8.7 px corner MAE**, MSE loss, AdamW + OneCycleLR) and self-supervised (**17.4 px MAE**, Charbonnier photometric loss) variants. Unsupervised pipeline uses **TensorDLT** (differentiable SVD) + **Spatial Transformer Network** for end-to-end gradient flow. Built hybrid panorama stitching system combining classical Shi-Tomasi + SSD matching with DL homography.

SfM Pipeline: Built full incremental SfM: **SIFT** features, **RANSAC**-based Fundamental matrix, Essential matrix decomposition, cheirality triangulation, **PnP-RANSAC** registration, and **bundle adjustment**. Reconstructed **1,318 3D points** across 4 cameras from 5 real images.

AutoCalib: Implemented Zhang's camera calibration from scratch: normalized DLT homography estimation, intrinsic recovery via SVD ($\mathbf{B} = \mathbf{K}^{-T}\mathbf{K}^{-1}$), and two-parameter radial distortion. Demonstrated **Rodrigues axis-angle parameterization** outperforms direct rotation matrix: **RMS 0.509 px vs 0.634 px**, 85 vs 163 parameters.

Pb-Lite + CNNs: Implemented **Pb-Lite** boundary detector using DoG (32), Leung-Malik (48), and Gabor (32) filter banks; K-Means texton / brightness / color clustering; half-disk χ^2 gradient computation fused with Canny/Sobel. Trained and benchmarked **ResNet (84%)**, **DenseNet (82%)**, **ResNeXt (75%)**, **baseline CNN (70%)** on CIFAR-10 with full ablations.

Fire Detection: Designed a neural network for **fire detection via color-pattern analysis**, enabling earlier detection than traditional smoke/heat sensors. Integrated into autonomous smart-bot with real-time sensor fusion and navigation. **Published in IJERCSE and at ICIoT 2025.**

Publications

Ashish S., Jeya R., Rithish R., Donipart R., & Ganesh K. (2025). *Fire Detection and Risk Prediction for Smart Safety: A Neural Network-Driven IoT Approach*. IJERCSE.

Ashish S., Jeya R., Rithish R., Donipart R., & Ganesh K. (2025). *Fire Aware Smart-Bot with AI Responsive System*. ICIoT 2025.

Certifications & Achievements

- Red Hat Certified System Administrator (RHCSA)
- Oracle Cloud Infrastructure (OCI) Certified
- 2nd Place – Project Expo, SRMIST