Alok Raj

+91-8709414158
Github Website
Alok Raj
alokraj.xml@gmail.com

Education

- **Indian Institute of Technology (ISM) Dhanbad** — Dhanbad, India
  Bachelors of Technology in Computer Science and Engineering; CGPA: 8.59 / 10.00 or 3.44 / 4.00 — *Class of 2026*
- **BR DAV Public School** — Begusarai, India
  High School(12th); Percentage: 97.2% — *Class of 2022*

Publications

GRIM: Task-Oriented Grasping with Conditioning on Generative Examples Paper

*Shailesh, **Alok Raj**, Nayan Kumar, Priya Shukla, Andrew Melnik, Michael Beetz, Gora Chand Nandi*
*Accepted at the **Association for the Advancement of Artificial Intelligence (AAAI)*** *2026*

Search-TTA: A Multi-Modal Test-Time Adaptation Framework for Visual Search in the Wild Paper

*Derek Ming Siang Tan, Shailesh, Boyang Liu, **Alok Raj**, Qi Xuan Ang, Weiheng Dai, Tanishq Duhan, Jimmy Chiun, Yuhong Cao, Florian Shkurti, Guillaume Adrien Sartoretti*
*Accepted at the **Conference on Robot Learning (CoRL)*** *2025*

Experience

- **MARMoT Lab, NUS** — Remote
  Research Internship Under Prof. Guillaume A Sartoretti — *Feb 2025 – Oct 2025*
  - **Project**: *Long Horizon Task and Motion Planning*.
  - **Description**:
    - Working on long-horizon **loco-manipulation** and policy mobilization frameworks in **RoboCasa**.
    - Fine-tuning and benchmarking SOTA VLMs like Qwen for **Visual-Question-Answering** tasks in kitchen environments.
  - **Project (CoRL 2025)**: *Search-TTA*: A Multi-Modal Test-Time Adaptation Framework for Visual Search in the Wild
  - **Contributions**:
    - Adapted VAS/PSVAS RL frameworks & developed a **Dijkstra-based evaluation method** using model predictions & exploration penalties.
    - Implemented & evaluated **meta-learning Test-Time Adaptation (TTA)**, improving Out-of-Distribution performance on iNaturalist dataset.
- **Samsung R&D Institute India-Bangalore** — Bangalore, India
  Research and Development Intern — *May 2025 - July 2025*
  - **Project**: Voice Biometrics for low-compute devices such as smart watches.
  - **Description**:
    - Developed an on-device speaker verification system for low-compute devices, utilizing modern attention-based architectures.
    - Implemented model **quantization** and optimized the system for on-device authentication.
    - Hybrid inference architecture to infer on multiples devices based on device constraints.
- **Center of Intelligent Robotics, IIIT Allahabad** — Remote
  Research Internship Under Prof. G.C. Nandi & Andrew Melnik — *Dec 2024 - May 2025*
  - **Project**: Task-Oriented Grasping using Generative conditioning (*GRIM* Framework)
  - **Description**:
    - Developed the **GRIM memory creation pipeline**, involving single-view 3D hand-object reconstruction using foundation models (**VLMs**, **Genie (Text-to-3D)**).
    - Created the **hybrid alignment strategy** for matching retrieved memory instances to scene objects, utilizing **DINOv2 PCA features** for coarse alignment and **ICP** with Chamfer distance for refinement.
    - Training-free task-oriented grasping by transfer of grasp poses from aligned, generatively-created 3D examples to novel objects.
- **Clutterbot Technologies** — Bangalore, India
  Machine Learning Intern — *May 2024 - July 2024*
  - **Project**: Addressed challenge of limited labeled data via Self-Training with Distillation and Curriculum Learning.
  - **Description**:
    - Self-Training, using Co-DETR, to expand the dataset with unannotated images.
    - Curriculum learning trained DAMO-YOLO-M, distilled to DAMO-YOLO-Tiny for robot deployment.
    - Improved mAP50 from 34% to 42% and evaluated performance with TIDE.
    - Deployed on robot using Nvidia DeepStream and integrated with ROS2.

- **Robotics and Automation Lab, IIT (ISM)** [video] Dhanbad, India
  Research Intern: Under Prof. Arun Dayal Udai *December 2023*
  - **Project**: Development of in-house **Quadrupedal Robot** for Mining Application.
    - Developed ROS based framework for a in-house developed Quadrupedal Robot.
    - **Reinforcement Learning** based control policy.

Selected Projects

- **Multimodal Price Regressor** [ppt] [code]
  Amazon ML Challenge 2025 *Oct 2025*
  · Achieved **3rd** place on the public leaderboard and **5th** on the private leaderboard in the national Amazon ML Challenge 2025.
  · Developed a multimodal solution for smart product pricing, leveraging both image and text data to predict optimal price points.
  · Feature extraction pipeline using pretrained embeddings from **Qwen-3**, **Siglip2**, and **DinoV3**.
  · Designed an **neural network with modality-specific tower MLPs and a final regressor**, using Log-based MSE loss function for right-skewed price data.

- **Mobile-Swarm-Navigation** [video folder] [code]
  Inter-IIT Tech Meet 13.0 - BharatForge *Nov 2024 - Dec 2024*
  · **Project**: Create a Centralised Intelligence for Dynamic Swarm Navigation.
  · Scalable ROS2 based robot swarm for autonomous exploration and navigation in a dynamic environment.
  · Database management system for task allocation for the swarm with **Agentic LLM based Tool-Calling**.
  · Dynamic environmental mapping with Instance Segmentation and Stereo Depth.

- **Panoramic Dental X-ray Anomaly Detection** [code]
  Active Growth Partners - ML Intern Project *Aug 2024 - Nov 2024*
  · Built and deployed a dental disease detection system with disease segmentation, achieving mAP@50 of 31%.
  · Developed a Flask API for real-time processing, integrated with **AWS EC2** and **S3** for scalability.

- **Autonomous Driving NXP-B3RB Buggy** [link]
  NXP-AIM Self Driving Car Design Challenge: Under Prof. Subhrangsu Mandal *Aug 2024 - Oct 2024*
  · Developed an autonomous driving system, for a B3RB-buggy, achieving a 1:42 (min:sec) track time.
  · Integrated LiDAR and camera for lane detection, obstacle avoidance, and traffic sign recognition.
  · Trained YOLOv5s, optimized with **INT8 quantization** for NPU, achieving real-time 7 Hz inference.

- **Hologlyph Bots** [video] [code]
  E-Yantra 2023 *Aug 2023 - Jan 2024*
  · Designed holonomic drawing robots, developing PID control with inverse kinematics on an ESP-32 (Micro-ROS).
  · Simulated and deployed the 3-bot swarm, using an overhead camera with Aruco detection for pose tracking.

Skills

- **Programming**: C++, Python, Linux, Git, SSH
- **Simulation/Visualization**: Isaac Gym, Gazebo, Mujoco, Sapien, Open3D, RoboCasa
- **Frameworks/Libraries**: ROS/ROS2, PyTorch, AWS

Relevant Coursework

- **Computer Science**: Data Structure & Algorithms, Database Management System, Optimization Techniques
- **Machine Learning**: Reinforcement Learning, Self-Supervised Learning, Convolutional Neural Networks, Transformers, SSMs, VLMs

Honors and Awards

- 5th Position: Amazon ML Challenge 2025
- 6th Position: InterIIT Tech Meet 13.0 for Rigbettlelabs
- Winner: NXP-AIM Regional Finale and Finalist: Grand Finale
- 3rd Position: Robowars(BattleBots) at TechKriti 2024 (Annual Tech Fest of IIT Kanpur)
- 3rd Position: Robowars(BattleBots) at Concetto 2024 (Annual Tech Fest of IIT Dhanbad)
- Received the Excellent Academic Performance Award (AISSCE 2022).

Extra-Curricular Activities

- **Club Coordinator**: RoboISM - The official Robotics and AI club of IIT ISM Dhanbad.
- **Joint Event Coordinator**: NVCTI - The innovation cell of IIT ISM Dhanbad.