

# Raman Kumar Jha

✉ [ramanjha@tamu.edu](mailto:ramanjha@tamu.edu) | [🌐 jhakrraman.github.io](https://github.com/jhakrraman) | [🌐 jhakrraman](https://www.linkedin.com/in/jhakrraman) | [in jhakrraman](https://www.in.jhakrraman) | [📧 Raman Jha](#)

## Education

---

### Texas A&M University

Doctor of Philosophy in Computer Engineering,

**Relevant Coursework:** Multi Robot Systems, Deep Learning, Multimodal Computer Vision

College Station, USA

Aug 2026 – May 2030

### New York University

Master of Science in Computer Engineering, GPA 3.7/4

**Relevant Coursework:** Foundation of Robotics, Robot Perception, Reinforcement Learning & Optimal Control for Robotics, Embodied Learning & Vision, Deep Learning; Plan, Learn & Control Autonomous Space Robots

New York, USA

Aug 2024 – May 2026

### Savitribai Phule Pune University

Bachelor of Engineering in Computer Engineering, GPA 8.8/10

**Relevant Coursework:** Principles of Programming Languages, Data Structures & Algorithms; Discrete Mathematics, Artificial Intelligence, Machine Learning, Data Science & Big Data Analytics, Deep Learning, Image Processing

Pune, India

Jul 2019 – Jun 2023

## Technical Skills

---

**Programming:** Python, C/C++, Matlab, Octave,  $\text{\LaTeX}$

**ML/AI Frameworks:** PyTorch, TensorFlow, OpenCV, Scikit-Learn, Keras, Stable Baselines3

**Simulation & Tools:** RViz, Gazebo, Isaac, Habitat, Linux, Git/GitHub, Docker, Anaconda, HuggingFace

**Robotics & Perception:** ROS, SLAM, PCL, Sensor Fusion, Camera Calibration, Motion Planning

**LLM & NLP:** LangChain, PEFT (LoRA/QLoRA), RAG, NLTK, Prompt Engineering, Multimodal Learning

**Data Analysis:** NumPy, Pandas, Matplotlib, SciPy, Seaborn

**Cloud, and MLOps:** AWS, GCP, Azure, Hugging Face, Wandb

## Publications

---

### [1] VCS-SLAM: Visibility-Consistent and Surface-Coupled Semantic SLAM

Raman Jha, Shuaihang Yuan, Yi Fang.

Robotics and Automation Letters(RAL) 2026 [Paper](#)

### [2] Adaptive Keyframe Selection for Scalable 3D Scene Reconstruction in Dynamic Environments.

Raman Jha, Yang Zhou, Giuseppe Loianno.

International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS) 2026, [Paper](#).

### [3] RT-X NEt: Cross Attention Network using RGB and Thermal Images for Low-Light Image Enhancement.

Raman Jha, Adithya Lenka, Mani Ramanagopal, Aswin C. Sankaranarayanan, Kaushik Mitra.

International Conference on Image Processing (ICIP) 2025, [Paper](#).

### [4] A Contrastive Meta-Learning Approach with Isotropic Sparse Decomposition for Scalable Audio-Visual Learning.

Dhruv Dixit, Paritosh Pandey, Raman Jha, Pranshu Bhatnagar, Shashank Mouli Satapathy.

International Conference on Computing, Communication and Learning (CoCoLe) 2024, [Paper](#).

### [5] Detection of Vehicle Lights and Classification of Its State.

Raman Kumar Jha

International Conference on Advanced Computing Technologies and Applications (ICACTA) 2023, [Paper](#).

### [6] Prediction of Bankruptcy of a Company Using Machine Learning Techniques.

V. Kothuru, Raman Kumar Jha, S. Ranjit, B. V. Kumar Reddy, S. Roy, S. Sudheer.

International Conference on Electronics and Sustainable Communication Systems (ICESC) 2022, [Paper](#).

## Experience

---

### Embodied AI and Robotics Lab, New York University

Research Assistant(MS Thesis)

New York, USA

Aug 2025 – May 2026

- Mentored by [Shuaihang Yuan](#), and advised by [Prof. Yi Fang](#) for MS thesis.

- Addressed robot object goal navigation and semantic mapping challenges in visually complex environments.
- Developed a physically consistent semantic mapping module to mitigate semantic ghosts and occlusion leakage.
- Implemented a visibility-aware model and surface-coupled regularization to align semantic boundaries.
- Enforced temporal consistency and visibility gating to reduce the Occlusion Leakage Rate (OLR) and enhance spatial reasoning for robust autonomous navigation.

#### Agile Robotics and Perception Lab, New York University

Research Assistant

New York, USA

Oct 2024 – July 2025

- Mentored by [Yang Zhou](#) and supervised by [Prof. Giuseppe Loianno](#).
- Worked on adaptive keyframe selection for scene reconstruction in dynamic environments.
- Developed a Transformer-based architecture to regress 3D point maps directly from 2D images.
- Optimized keyframe selection by implementing an error-based metric and momentum-based keyframe control.
- Reduced the required number of keyframes by 5% while improving the system's reconstruction on a drone.

#### Computational Imaging Lab, IIT Madras

Research Associate

Chennai, India

Aug 2023 – June 2024

- Advised by [Prof. Kaushik Mitra](#) and [Prof. Aswin C. Sankaranarayanan](#).
- Developed a cross-attention network for low-light image enhancement and restoration in autonomous systems.
- Used thermal images as a guide for RGB images to improve the performance of the enhancement algorithm.
- Curated and released the V-TIEE dataset, consisting of 50 paired real-world RGB-Thermal sequences.
- Achieved 2 dB performance increase over traditional RGB-only networks for low-light image enhancement.

#### SeiSei.ai

Founding Machine Learning Intern

Jaipur, India

April 2023 – July 2023

- Designed a Generative AI pipeline for automated video transformation, visual face dubbing, and voice conversion.
- Created a system to detect and extract the landmarks from the face of a particular individual in a video.
- Implemented a facial landmark detection system to drive a GAN-based FreeVC model for voice cloning and a DInet model for visual dubbing.
- Improved inference performance by 17% and successfully integrated the pipeline into the company's production.

#### Dalhousie University

Mitacs Globalink Research Intern / Research Report

Halifax, Canada

July 2022 – Oct 2022

- Supervised by [Prof. Andrew Rutenberg](#).
- Worked on classifying health variables into low and high dimensions for predicting aging and death.
- Used Variational Autoencoders and Recurrent Neural Networks for model development and testing.
- Trained and tested on the [English Longitudinal Study of Ageing \(ELSA\)](#) dataset.
- Developed latent space models improving performance by 14% and designed a modified separate filter to calculate the loss function for each health variable.

#### Centre of Visual Information Technology Lab (CVIT), IIIT Hyderabad

Applied Research Fellow

Hyderabad, India

Jan 2022 – June 2022

- Supervised by [Prof. Ravi Kiran Sarvadevabhatla](#) and [Prof. C V Jawahar](#).
- Worked on detecting vehicle lights and classifying their states by training multiple object detection models, including YOLOv7, YOLOv5, YOLOv4, Faster R-CNN, and MobileNet.
- Used [Indian Driving Dataset \(IDD\)](#) for training and testing of the detection and classification system.
- YOLO V7 achieved best results with 88.4% mAP and was fine-tuned to improve detection of small lights.
- Applied color thresholding principle for the red and white light recognition in diverse environments.

## Projects

### Lunar Rover Navigation & Control | [GitHub](#)

Oct 2025 – Dec 2025

- Simulated a six-wheeled lunar rover for autonomous navigation in a GPS-denied, low-gravity environment.
- Targeted robust waypoint traversal while maintaining chassis stability and mitigating risks on uneven terrain.
- Implemented an EKF-based sensor fusion for state estimation and designed a stability-aware proportional control.
- Maintained a 0.6–0.9 stability margin and quantified a 55m localization drift caused by wheel slip in lunar regolith.

### Self-Drive: Virtual Integrated Project | [GitHub](#)

Jan 2025 – May 2025

- Planned an autonomous driving system for real-world navigation in complex maze, and in real-world.
- Designed and integrated perception, planning, and control modules for robust decision-making.
- Utilized DINOv2-based lane detection and object recognition, and optimized A\* for efficient path planning.
- Achieved 95% navigation accuracy in real-world maze environments with safe and reliable performance.

### Diffusion-Based Model Predictive Control | [GitHub](#)

Feb 2025 – May 2025

- Designed a control strategy to improve robotic trajectory optimization in uncertain and dynamic environments.

- Enhanced MPC by integrating probabilistic modeling to handle uncertainty in unseen conditions.
- Combined diffusion models with MPC and implemented probabilistic trajectory generation using diffusion.
- Achieved a 20% improvement in trajectory optimization performance over traditional MPC approaches.

#### **CLIP Model Benchmarking on Flickr8K | [GitHub](#)**

Apr 2025

- Designed CLIP vision-language models for zero-shot image retrieval and classification on Flickr8K dataset.
- Objective was to assess how encoders impact retrieval quality and response to both concrete and abstract prompts.
- Constructed and pretrained three model pipelines, tested on custom prompts to evaluate performance.
- Larger models excelled at retrieval tasks, however smaller ones trained faster but ,performed worse on queries.

#### **LoRA Fine-Tuning for Text Classification | [GitHub](#)**

Mar 2025

- Adapted a pre-trained transformer (RoBERTa) for a text classification task in a resource-constrained environment.
- Aimed to maximize classification accuracy while minimizing additional parameter usage using PEFT techniques.
- Fine-tuned RoBERTa with LoRA adapters using low learning rate (1e-4), weight decay, and optimized batch size.
- Achieved 94.06% validation accuracy (final loss: 0.16) and 85.9% Kaggle test accuracy on final competition.

#### **2D Quadrotor Control Using Reinforcement Learning | [GitHub](#)**

Oct 2024 – Dec 2024

- Designed a control solution for a 2D quadrotor to reach target positions while avoiding obstacles.
- Implemented a reinforcement learning framework for stable navigation under dynamic constraints.
- Trained a PPO-based policy and engineered rewards balancing target accuracy, stability, and obstacle avoidance.
- Achieved robust flight control with smooth trajectories, minimal oscillations, and reliable obstacle avoidance.

#### **American Sign Language Recognizer | [GitHub](#)**

Jun 2022 – Jul 2022

- Built an American Sign Language recognizer using real-time hand gesture detection.
- Applied the YOLOv5 object detection framework to real-time camera images for accurate sign recognition.
- Recognized five distinct American Sign Language signs, including Hello, I Love You, Yes, No, and Thank You.
- Achieved 97.4% recognition accuracy across all American Sign Language sign classes during real-time camera-based testing and evaluation.

#### **Driver Drowsiness Detector | [GitHub](#)**

May 2021 – Jun 2021

- Developed a driver drowsiness detector based on real-time eye-state analysis.
- Trained a CNN-based model on real-time camera image sequences to detect signs of drowsiness.
- Classified the driver's state as awake or drowsy using extracted eye condition features and blink patterns.
- Achieved 85.2% detection accuracy for driver drowsiness classification across real-time camera-based testing and evaluation.

## **Teaching Experience**

---

### **Teaching Assistant**

Spring 2026

#### **Computer Vision for Sciences and Engineering**

New York University

- Assisted students with paper readings, research critiques, and developing effective research presentation skills.
- Mentored students through structured project guidance, including proposals and feasibility studies.
- Evaluated weighted presentations, peer-review assignments, and initial and final project reports.

### **Teaching Assistant**

Fall 2025

#### **Advanced Topics in Computer Vision**

New York University

- Assisted students with their presentations, paper reviews, project queries, and coursework.
- Conducted office hours for the students to help them with their projects in Computer Vision.
- Graded student presentations, paper reviews, and course projects submitted through Brightspace.

### **Teaching Assistant**

Spring 2025

#### **Physics Informed Machine Learning**

New York University

- Mentored students with their homework, quizzes, and projects, and addressed doubts on Piazza.
- Held office hours and tutorial sessions for the students in Physics-Informed Machine Learning.
- Evaluated student homework, quizzes, assignments, and midterms for the course on Brightspace.

### **Teaching Assistant**

Spring 2025

#### **Chemical Engineering Computation**

New York University

- Assisted students with their homework, assignments, and in-class doubts on Python programming.
- Held office hours and guided students through their mini projects and final course projects.
- Graded homework, quizzes, assignments, and midterms for Chemical Engineering Computation on Brightspace.

### **Teaching Assistant**

Spring 2024

#### **Modern Computer Vision**

IIT Madras

- Conducted tutorial classes on the basics of programming and deep learning for students.
- Evaluated student assignments and projects for the computer vision course through Moodle.
- Mentored students on their course projects, competitions, and hackathon participation.

## Teaching Assistant Machine Learning

Spring 2023  
Savitribai Phule Pune University

- Graded student assignments, projects, and exams for the machine learning course.
- Held office hours and assisted students with their homework, projects, and hackathons.
- Assisted students and answered their questions on Sowl about reports and final projects.

## Achievements and Extracurriculars

---

|                                                                                                              |         |
|--------------------------------------------------------------------------------------------------------------|---------|
| 1 <b>Judge</b> , <a href="#">Hackathon LLUS X PULSE</a> hosted by Korean International Students Organization | 04/2026 |
| 2 <b>Winner</b> , <a href="#">Spark Hackathon</a> hosted by NVIDIA, and sponsored by ACER                    | 04/2026 |
| 3 <b>Outstanding Reviewer</b> , <a href="#">British Machine Vision Conference (BMVC)</a>                     | 11/2024 |
| 4 <b>Recipient</b> , merit-based scholarship of \$7000 per year from New York University                     | 04/2024 |
| 5 <b>Finalist</b> , <a href="#">Smart India Hackathon</a>                                                    | 08/2022 |
| 6 <b>Recipient</b> , <a href="#">Mitacs Globalink Research Fellowship</a>                                    | 03/2022 |
| 7 <b>Winner</b> , <a href="#">Codeheat contest</a> , FOSSASIA                                                | 09/2021 |
| 8 <b>Recipient</b> , Script Fellowship, Script organization                                                  | 08/2021 |

## Reviewer Experience

---

**Reviewer:** BMVC 2026, NeurIPS 2025, BMVC 2025, ICLR 2024, BMVC 2024, ICVGIP 2023

## Leadership and Extra-curriculars

---

- 1 **Treasurer**, [Hindu Student Union NYU](#) (July 2025 – May 2026). Directed financial operations, including budget forecasting and allocation, while securing resources for diverse events to promote community engagement.
- 2 **Treasurer**, [IEEE NYU Student Branch](#) (Jun 2025 – May 2026). Managed financial activities, budget planning, and coordinated funding for technical events and workshops for the club.
- 2 **Steward**, [GSOC NYU](#) (May 2025 – May 2026). Represented as a steward for NYU Tandon School of Engineering, and helped students with queries, grievances, and issues related to their on-campus jobs.
- 3 **Mitacs Globalink Ambassador**, [Mitacs Globalink Research Internship](#) (Mar 2023 – Jun 2023). Assisted and guided upcoming research interns regarding the Mitacs Globalink Research Fellowship.
- 4 **Machine Learning Lead**, [Google Developer Student Club](#) (Aug 2021 – Jul 2022). Conducted events, seminars, tutorials, and guided juniors for research in Machine Learning.