Qasim Ali

muhammad.qasim.ali06@gmail.com +1-647-807-7243 Toronto, Canada Scholar LinkedIn

Education

Master of Applied Science (MASc), Systems Design Engineering

University of Waterloo

09/2023 – 08/2025

Supervisors: Alex Wong , Javad Shafiee

Awards: Engineering Excellence MASc Fellowship ($50k)

Research Focus: Enhancing embodied reasoning with scene representations.

- **Generative AI & Video Prediction**: Trained a small generative video model for Humanoid Robots.
- **External Memory for VLMs**: Created an agentic system for Vision-Language Models to execute spatial reasoning: process scenes into JSON representations and interact with them using function-calling APIs.
- **Representation Learning for RL**: Applied representation learning strategies for goal-conditioned reinforcement learning (maze and manipulation tasks).

Mechatronics Engineering

McMaster University - B.Eng

2018 – 2023

GPA: 10.9/12, 3.9/4

Courses: Databases, Software Development, Operating Systems, Embedded Systems I & II, Predictive & Intelligent Control, Robotics

Skills

Programming Languages: *Python, C, C++*

Frameworks & Tools Used: *Pytorch, Jax, Matlab, Simulink, Git, AWS, Unreal Engine, ROS, SQL, Canoe, Trace32, ARM-Cortex Microcontroller, Verilog, Tensorflow*

Extracurricular Activities

Parking Lot Drone

Undergrad Capstone Project

09/2022 – 04/2023

Built a drone that autonomously surveys parking lots for occupancy levels. Sample flight videos .

Autonomous Vehicle Team

McMaster EcoCar

10/2021 – 02/2023

Wrote an autonomous control algorithm for vehicles at V2I-connected intersections. Presented the algorithm at contest.

Work Experience

Software Engineering Intern

Magna Electronics

05/2021 – 08/2022

Worked with the **embedded software** team to develop advanced self-driving features on automotive platforms.

Languages/Tools: C/C++, Vehicle ECUs, Visual Studio, Canoe, AWS, Unreal Engine, Trace32

- Debugged and tested complex detection algorithms and state machines across both simulated (offline) and real-time (online) scenarios.
- Wrote communication interfaces to connect live vehicle signals with driving algorithms.
- Designed and implemented the team's first end-to-end AWS pipeline to run object detection KPIs in the cloud; presented the solution to management.
- Developed a visualization tool in Unreal Engine to display detected objects and lane boundaries.

Publications

Humanoid World Models: Open World Foundation Models for Humanoid Robotics

ICML 2025 Workshop: Building Physically Plausible World Models

- Trained video-based world models using a 100-hour dataset of humanoid video and multiple GPUs.
- Experimented with variety of generative models: diffusion, flow matching, and masked transformer.
- Reduced transformer size by 50% using parameter reduction strategies.

GraphPad: Inference-Time 3D Scene Graph Updates for Embodied Question Answering

CVPR 2025 Workshop: Bridging Language, Vision and Action in 3D Environments

- Implemented an RGBD-to-scene-graph pipeline.
- Designed an agentic reasoning pipeline and function-calling APIs for step-by-step spatial reasoning.
- Improved upon baseline by **+3.0%** and **5x** fewer images on embodied question answering benchmark.

MR-CRL: Leveraging Predictive Representations for Contrastive Goal-Conditioned Reinforcement

RLC 2025 Workshop: Reinforcement Learning Beyond Rewards

- Developed an algorithm for learning stable, predictive representations that forecast the results of actions.
- Improved performance on some goal-conditioned RL benchmarks (maze solving and puzzle manipulation).