Karmesh Yadav

ykarmesh@gmail.com | www.karmeshyadav.com | ykarmesh | KarmeshYadav

Education

Georgia Institute of Technology — *Atlanta*
PhD in Computer Science — *Aug 2023 - Jul 2026 (Expected)*
- Advisors: Dr. Zsolt Kira and Dr. Dhruv Batra

Carnegie Mellon University — *Pittsburgh*
Masters in Robotic Systems Development — *Aug 2018 - May 2020*
- Capstone Project Advisor: Dr. John Dolan

Indian Institute of Technology, Guwahati — *Guwahati*
B.Tech. in Mechanical Engineering — *July 2013 - May 2017*
- Undergrad research advisor: Dr. S.K. Dwivedi

Professional Experience

Yutori — *San Francisco*
Intern, Technical Staff, AI — *Aug 2024 - Nov 2024*
- Developed deployment pipelines for web agents to operate on real websites while effectively avoiding bot detection.
- Created data filtering pipeline and trained VLM-based agents for web navigation tasks.

Fundamental AI Research (FAIR), Meta AI — *Menlo Park*
AI Resident — *Aug 2021 - Jun 2023*
- Researched self-supervised pretraining techniques for learning useful visual representations for embodied agents.
- Released the HM3D-Semantics (HM3DSem) dataset and the Open-Vocabulary Mobile Manipulation (OVMM) benchmark based on the Habitat Simulator.

ISEE Inc. — *Boston*
Senior Robotics Engineer — *Jul 2020 - Aug 2021*
- Explored deep uncertainty estimation techniques for predicting the closed loop tracking performance of an AV controller. Estimated the collision prob. of the AV w.r.t. obstacles in an occupancy grid.
- Improved the trajectory optimization planner and robustified its collision checking. This led to an increased confidence in its performance and resulted in its deployment on the AV.

ISEE Inc. — *Boston*
Software Development Intern — *May 2019 - Aug 2019*
- Built toolboxes to automate the system identification and calibration procedure of Isee's vehicles. Also researched and implemented various vehicle and tire models for control application in AVs.

Autonomous Driving Team, Mathworks — *Hyderabad*
Intern — *Aug 2017 - Nov 2017*
- Worked on improving the localization module of the autonomous vehicle by fusing ORB-SLAM's output with other sensors like GPS, IMU and Wheel Odometer

Publications

Conference Publication

Gunshi Gupta*, **Karmesh Yadav***, Yarin Gal, Dhruv Batra, Zsolt Kira, Cong Lu, Tim G. J. Rudner. Pre-trained Text-to-Image Diffusion Models Are Versatile Representation Learners for Control. **Spotlight (Selection Rate: 326/15671 = 2.1%)** at Neural Information Processing Systems (NeurIPS), 2024 and **Oral** at Workshop on Generative Models for Decision Making (GenAI4DM) at ICLR 2024.

Arjun Majumdar*, Anurag Ajay*, Xiaohan Zhang*, Sriram Yenamandra, Mikael Henaff, Alexander Sax, Sneha Silwal, Paul Mcvay, Oleksandr Maksymets, Sergio Arnaud, Pranav Putta, **Karmesh Yadav**, Qiyang Li, Ben Newman, Mohit Sharma,

Vincent-Pierre Berges, Shiqi Zhang, Pulkit Agrawal, Dhruv Batra, Yonatan Bisk, Mrinal Kalakrishnan, Franziska Meier, Chris Paxton, Aravind Rajeswaran. OpenEQA: Embodied Question Answering in the Era of Foundation Models. Conference on Computer Vision and Pattern Recognition (CVPR), 2024 and **Spotlight** at ICRA Workshop on Mobile Manipulation and Embodied Intelligence, 2024

Sneha Silwal*, **Karmesh Yadav***, Tingfan Wu*, Jay Vakil*, Arjun Majumdar*, Sergio Arnaud*, Claire Chen, Vincent-Pierre Berges, Dhruv Batra, Aravind Rajeswaran, Mrinal Kalakrishnan, Franziska Meier, Oleksandr Maksymets. What do we learn from a large-scale study of pre-trained visual representations in sim and real environments? International Conference on Robotics and Automation (ICRA), 2024 and **Spotlight** at ICRA Workshop on Mobile Manipulation and Embodied Intelligence, 2024

Sriram Yenamandra*, Arun Ramachandran*, **Karmesh Yadav***, Austin Wang, Mukul Khanna, Theophile Gervet, Tsung-Yen Yang, Vidhi Jain, Alexander William Clegg, John Turner, Zsolt Kira, Manolis Savva, Angel Chang, Devendra Singh Chaplot, Dhruv Batra, Roozbeh Mottaghi, Yonatan Bisk, Chris Paxton. HomeRobot: Open-Vocabulary Mobile Manipulation. Conference on Robot Learning (CoRL), 2023.

Arjun Majumdar*, **Karmesh Yadav***, Sergio Arnaud*, Yecheng Jason Ma, Claire Chen, Sneha Silwal, Aryan Jain, Vincent-Pierre Berges, Pieter Abbeel, Jitendra Malik, Dhruv Batra, Yixin Lin, Oleksandr Maksymets, Aravind Rajeswaran, Franziska Meier. Where are we in the search for an Artificial Visual Cortex for Embodied Intelligence? Neural Information Processing Systems (NeurIPS), 2023.

Jacob Krantz, Theophile Gervet, **Karmesh Yadav**, Austin Wang, Chris Paxton, Roozbeh Mottaghi, Dhruv Batra, Jitendra Malik, Stefan Lee, Devendra Singh Chaplot. Navigating to Objects Specified by Images.
International Conference on Computer Vision (ICCV), 2023.

Karmesh Yadav*, Ram Ramrakhya*, Santhosh Kumar Ramakrishnan*, Theo Gervet, John Turner, Aaron Gokaslan, Noah Maestre, Angel Xuan Chang, Dhruv Batra, Manolis Savva, Alexander William Clegg, Devendra Singh Chaplot. Habitat-Matterport 3D Semantics Dataset. Conference on Computer Vision and Pattern Recognition (CVPR), 2023. **Highlight Paper (Selection Rate: 235/9155 = 2.5%)**

Justin Wasserman, **Karmesh Yadav**, Girish Chowdhary, Abhinav Gupta, Unnat Jain. Last-Mile Embodied Visual Navigation. Conference on Robot Learning (CoRL), 2022.

Gunshi Gupta*, **Karmesh Yadav***, Liam Paull. Look-ahead meta learning for continual learning. Neural Information Processing Systems (NeurIPS), 2020. **Oral (Selection Rate: 105/9454 = 1.1%)**

Vignesh Prasad*, **Karmesh Yadav***, Rohitashva Singh Saurabh, Swapnil Daga, Nahas Pareekutty, K Madhava Krishna, Balaraman Ravindran, Brojeshwar Bhowmick. Learning to Prevent Monocular SLAM Failure using Reinforcement Learning. Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), 2018.

Workshop Publication

Karmesh Yadav, Ram Ramrakhya, Arjun Majumdar, Vincent-Pierre Berges, Sachit Kuhar, Dhruv Batra, Alexei Baevski, and Oleksandr Maksymets. Offline visual representation learning for embodied navigation. Reincarnating Reinforcement Learning (RRL) at ICLR, 2023.

Preprint

Gunshi Gupta, **Karmesh Yadav**, Zsolt Kira, Yarin Gal, Rahaf Aljundi. Memo: Training Memory-Efficient Embodied Agents with Reinforcement Learning. (In submission) 2025

Ahmad Elawady, Gunjan Chhablani, Ram Ramrakhya, **Karmesh Yadav**, Dhruv Batra, Zsolt Kira, Andrew Zsolt. ReLIC: A recipe for 64k steps In-Context Reinforcement Learning for Embodied AI. arXiv preprint arXiv:2410.02751 (2024).

Karmesh Yadav*, Arjun Majumdar*, Ram Ramrakhya, Naoki Yokoyama, Alexei Baevski, Zsolt Kira, Oleksandr Maksymets, Dhruv Batra. OVRL-V2: A simple state-of-art baseline for ImageNav and ObjectNav. arXiv preprint arXiv:2303.07798 (2023).

Service & Teaching

Workshops and Competitions

May 2024	**Workshop on Vision-Language Models for Mobile Manipulation**, Organizer	*Yokohama*
Dec 2023	**Open Vocabulary Mobile Manipulation Challenge, NeurIPS**, Organizer	*New Orleans*
Jun 2023	**Habitat Navigation Challenge, CVPR**, Lead Organizer	*Vancouver*
Dec 2022	**Habitat Rearrangement Challenge, NeurIPS**, Organizer	*New Orleans*
Jun 2022	**Habitat ObjectNav Challenge, NeurIPS**, Lead Organizer	*New Orleans*

Peer Review

NeurIPS'23, ICLR'24, ICLR'25, CVPR'25

Mentoring

2022-2023	**Sriram Yenamandra**, Masters Student	*Georgia Tech*
2022-2023	**Arun Ramachandran**, Masters Student	*Georgia Tech*

Teaching

Spring'24	**Deep Learning**, Teaching Assistant	*Georgia Tech*