Jay Darshan Vakil

+1 (206)294-9533 | jdvakil@gmail.com | jdvakil.github.io | github.com/jdvakil | linkedin.com/in/jdvakil | Jay Vakil

Robotics, Robot learning, Motion planning, Embodied AI

Education

Bachelor of Science in Electrical Engineering, *University of Washington* | Seattle, WA *Sept 2018 - Mar 2022*

- Minor in Mathematics
- **Courses:** Object-Oriented programming with Java, Software Engineering, Data structures and algorithm, Linear/discrete signal processing, Microprocessor system design, AC/DC circuit analysis, Digital circuit analysis, Electrodynamics, Transistors and amplifiers, Electrical testing, Biomedical instrumentation.

Research Experience

Facebook AI Research, *Robotic System Engineer* | Fremont, CA *July 2023 - Present*

- **Mentors:** Dr Christopher Paxton, Dr Franziska Meier
- **Research Focus:** Mobile manipulation, Visual language models, Instance exploration and navigation
- Conducted a comprehensive empirical investigation on leveraging pre-trained visual representations for training real-world task execution policies, uncovering insights about the transferability from simulation to real-world environments, achieving groundbreaking results in indoor navigation, and showcasing the advantages of data augmentation and fine-tuning for real-world performance improvement. Evaluated the performance of five different PVRs across two policy-learning paradigms (visual imitation and reinforcement learning) and on three distinct robotic platforms.
- Developed a mobile manipulation agent capable of autonomously planning and exploring environments, while simultaneously recording instances of objects of interest in a 3D map. Utilizing a visual-language model, the agent can accurately identify objects and their corresponding instances based on open-vocabulary requests. Subsequently, the robot can navigate to the specified object, pick it up, and then transport it to a target location for placement.

Facebook AI Research, *Robotic System Engineer* | Pittsburgh, PA *April 2022 - July 2023*

- **Mentors:** Dr Vikash Kumar, Dr Christopher Paxton
- **Research Focus:** Visual imitation learning for robot manipulation, Offline reinforcement learning, Mobile manipulation
- Designed and constructed a distributed robotic arm cluster specifically tailored for large-scale experimentation and data collection. This cluster served as the hardware backbone for numerous impactful research projects.
- Researched a generalized universal agent proficient in solving challenging manipulation tasks within a kitchen-top environment. Our approach efficiently addressed over 38 tasks exhibiting 12 skills. By incorporating world priors from existing foundational models, we augmented the dataset, enhancing the agent's ability to generalize across different scenes. Outperformed prior methods by over 40%.
- Curated and released RoboSet, a scalable multi-task real-world robotic dataset for generalized offline learning.
- Developed a two-stage policy involving an instruction-based, fine-tuned language model to describe actionable skills and a multi-modal policy for predicting subsequent actions. This work was accepted for presentation at a CVPR 2023 workshop and CoRL 2023.

Publications

RoboHive: A Unified Framework for Robot Learning

Vikash Kumar, Rutav Shah, Gaoyue Zhou, Vincent Moens, Vittorio Caggiano, **Jay Vakil**, Abhishek Gupta, and Aravind Rajeswaran

- A comprehensive platform for Robot Learning and Embodied AI research, offering unified task interfaces, diverse environments, and expert benchmarks for algorithmic research in various domains.
- Accepted at NeurIPS 2023
- Website Paper

RoboAgent: Towards Sample Efficient Robot Manipulation with Semantic Augmentations and Action Chunking

Homanga Bharadhwaj, **Jay Vakil**, Mohit Sharma, Abhinav Gupta, Shubham Tulsiani, and Vikash Kumar

- An efficient system for training universal agents with multi-task manipulation skills using semantic augmentations and compact action representations, achieving superior performance and adaptability with only 7500 demonstrations.
- Accepted at workshops in NeurIPS 2023 and CoRL 2023. In submission to IEEE International Conference on Robotics and Automation (ICRA), 2023
- Website Paper

What do we learn from a large-scale study of pre-trained visual representations in sim and real environments?

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, and Oleksandr Maksymets

- Extensive research on using pre-trained visual representations (PVRs) for real-world task execution, involving multiple PVRs, policy-learning methods, and robots. Identified key insights, including the correlation between simulation and real-world performance and the success of PVRs in achieving zero-shot indoor navigation.
- In submission to IEEE International Conference on Robotics and Automation (ICRA), 2024
- Website Paper

Where are we in the search for an Artificial Visual Cortex for Embodied Intelligence?

Arjun Majumdar, Karmesh Yadav, Sergio Arnaud, Yecheng Jason Ma, Claire Chen, Sneha Silwal, Aryan Jain, Vincent-Pierre Berges, Tingfan Wu, **Jay Vakil**, Pieter Abbeel, Jitendra Malik, Dhruv Batra, Yixin Lin, Oleksandr Maksymets, Aravind Rajeswaran, Franziska Meier

- Conducted the largest study on pre-trained visual representations (PVRs) for Embodied AI, creating CORTEXBENCH with 17 tasks. Demonstrated that data scaling doesn't universally improve performance, but task-specific adaptation of VC-1 models achieved competitive or superior results on all benchmarks.
- Accepted at NeurIPS 2023
- Website Paper

SLAP: Spatial-Language Attention Policies

Priyam Parashar, Vidhi Jain, Xiaohan Zhang, **Jay Vakil**, Sam Powers, and Chris Paxton

- Address challenges in language-guided mobile manipulation, achieving an 80% success rate in real-world tasks with a single model, and a 30% improvement over prior work in challenging conditions with limited data, demonstrating robust execution of open-vocabulary task plans.
- Accepted at CoRL 2023
- Website Paper

Master Controller For High Energy Electron Source Part II

Jay Vakil, Esayas Abera, Cyrus Safi, Wayne Kimura

- Designed a master controller system for a high-energy electron source for particle-accelerator based on diamond wafer technology. System consisted of complex remote controllers and monitoring systems.
- Undergraduate capstone research project
- Presentation

Select media coverage

- **Carnegie Mellon University** Parenting a 3-Year-Old Robot
- **TechCrunch** Human toddlers are inspiring new approaches to robot learning
- **IEEE Spectrum** Video Friday: A 3-Year-Old Robot Your weekly selection of awesome robot videos
- **Hackaday** ROBOAGENT GETS ITS MT-ACT TOGETHER
- **TechBrief** Meet RoboAgent: Enabling Robots to Acquire Manipulation Abilities
- **Communications of the ACM** Parenting a Three-Year-Old Robot

Academic service

Reviewer

- 2024 IEEE International Conference on Robotics and Automation (ICRA 2024)

Skills

Programming	Python, C/C++, C#, Java, Catkin, CUDA, CMake, Matlab, Scripting (Bash), LaTeX, HTML, Vim, Linux, Tensorflow, Pytorch, Docker, OpenCV, AWS S3, Git
Robotics/Hardware	ROS 1/2, MuJoCo, FPGA, Arduino, Raspberry PI, PCB design, AtMega 2560, Sensor Interfacing, Sensor Fusion, Signal processing, Digital/Analog circuit design and simulation, Microprocessors, Embedded systems, Circuit testing, Transistor-level design, CAD, Systems on Chip (SoC), Electrical circuit testing, PCB design and testing.