(+1) 412-626-8797
vidhij@andrew.cmu.edu
vidhijain.github.io
linkedin.com/in/vidhijain96/

Vidhi Jain

Education

2022–Present **Doctorate in Robotics, Carnegie Mellon University**.
Robotics Institute, Advisor: Yonatan Bisk

2019–2021 **Masters of Science in Robotics, Carnegie Mellon University**.
Robotics Institute, *GPA 4.0* Advisor: Katia Sycara

2014–2018 **Bachelor of Engineering (Honors) in Computer Science, BITS Pilani**.
CGPA 9.06/10.00, Distinction, Advisor (off-campus thesis): Aaron Courville

Experience

Sep'22–Present **Connecting Language to Actions & the World (CLAW) Lab**, *CMU Pittsburgh, PA, USA*.
PhD student with Yonatan Bisk
- Trained motion policies in simulated tasks with rich physical context and language instruction diversity.
- Developing language grounding in VLMs through information-seeking actions and their outcomes.

Jun'23–Present **Google DeepMind**, *Mountain View, CA, USA*.
Student Researcher with Debidatta Dwibedi
- Training models for learning video-conditioned robot policies.
- Extending models for visual grounding with conditional caption generation to question answering.

Sep'21–Aug'22 **Meta (Facebook)**, *Menlo Park, CA, USA*.
AI resident with Akshara Rai and Yixin Lin
- Developed procedural generation of feasible trajectories for loading the simulated dishwasher in AI Habitat Replica Synthetic Apartment 0 Kitchen. Code
- Trained Encoder-decoder Transformer-based high-level policy for learning preferences in loading dishwasher from a single demonstration as prompt. Paper
- Demonstrated the transfer to real hardware with a Robot arm (Franka-Emika) to load dishes in drawers. Website

Aug'18–Aug'19 **Microsoft Research**, *Bengaluru, India*.
Research Fellow with Amit Deshpande and Navin Goyal
- **Divergence minimization in GANs** Investigated the properties of Jensen Shannon and Wasserstein divergences between the given and the generated distributions during training neural network generators, to discuss the learning dynamics like convergence and generalization.
- **Unsupervised learning from Information-theoretic perspective** Experimented with techniques like Contrastive Estimation, Predictive Coding, and Mutual Information to explain the empirical performance of unsupervised learning strategies and the existing gap compared to the supervised performance.

Jan'18–Jun'18 **Montreal Institute for Learning Algorithms (MILA)**, *Montreal, QC, Canada*.
Research Intern with Aaron Courville
- **Out-of-Distribution Detection in Generative Models** Experimented on autoregressive generative models and Variational Autoencoders, to detect any 'out-of-distribution (OOD)' sample at deployment. Thesis
- **Distillation in Generative Models** Investigated probabilistic knowledge distillation in autoregressive generative models and looked into the challenges in minimizing only Kullback Leibler (KL) divergence between the two distributions modeled by teacher and student network for a high-dimensional output in generative models.

May'17–Jul'17 **Simon Fraser University (SFU)**, *Burnaby, BC, Canada*.
MITACS Globalink Research Intern with Oliver Schulte
- **Learning Bayesian Networks for Relational Databases** Remodelled existing codebase for project *FactorBase* to integrate cross-table sufficient statistics or population variables in the contingency tables for learning ground-level sufficient statistics. Code | Invited Paper

May'16–Jul'16 **Forschungszentrum Informatik (FZI)**, *Karlsruhe, Germany*.
Research Intern with York Sure-Vetter and Matthias Frank
- **API annotation platform using Semantic Web** Designed and developed a customized Semantic MediaWiki for annotation of APIs for dynamic integration of provenance information in the decision support system of Project *BigGIS*; Presented Technical Poster at Grace Hopper Celebration India GHCI'16, Bengaluru, India; Extended Abstract | Poster

Preprint

Jan 2024 **Vid2Robot: End-to-end Video-conditioned Policy Learning with Cross-Attention Transformers**.
Vidhi Jain, Maria Attarian, Nikhil J Joshi, Ayzaan Wahid, Danny Driess, Quan Vuong, Pannag R Sanketi, Pierre Sermanet, Stefan Welker, Christine Chan, Igor Gilitschenski, Yonatan Bisk, Debidatta Dwibedi, February 2024. Paper | Website

Dec 2023 **Towards General-Purpose Robots via Foundation Models: A Survey and Meta-Analysis.**.
Yafei Hu*, Quanting Xie*, **Vidhi Jain***, Jonathan Francis, Jay Patrikar, Nikhil Keetha, Seungchan Kim, Yaqi Xie, Tianyi Zhang, Shibo Zhao, Yu Quan Chong, Chen Wang, Katia Sycara, Matthew Johnson-Roberson, Dhruv Batra, Xiaolong Wang, Sebastian Scherer, Zsolt Kira, Fei Xia, Yonatan Bisk, December 2023. Paper | Website

Sep 2023 **FlexCap: Generating Rich, Localized, and Flexible Captions in Images**.
Debidatta Dwibedi, **Vidhi Jain**, Jonathan Tompson, Andrew Zisserman, Yusuf Aytar. September 2023. Paper | Website

Publications

Nov 2023 **How to Prompt Your Robot: A PromptBook for Manipulation Skills with Code as Policies**.
Montserrat Gonzalez Arenas, Ted Xiao, Sumeet Singh, **Vidhi Jain**, Allen Z. Ren, Quan Vuong, Jake Varley, Alexander Herzog, Isabel Leal, Sean Kirmani, Dorsa Sadigh, Vikas Sindhwani, Kanishka Rao, Jacky Liang, Andy Zeng. *Accepted at* ICRA 2024 and *Presented at* CoRL 2023 Workshop TGR Poster, Atlanta, USA. Paper

Nov 2023 **Open X-Embodiment: Robotic Learning Datasets and RT-X Models**.
Open X-embodiment collaboration. *Accepted at* ICRA 2024 and *Presented at* LangRob @ CoRL 2023 Poster, Atlanta, USA. Paper | Website

Jun 2023 **SLAP: Spatial-Language Attention Policies**.
Priyam Parasher, **Vidhi Jain**, Xiaohan Zhang, Jay Vakil, Sam Powers, Yonatan Bisk, Chris Paxton. *In Proceedings of* Conference on Robot Learning (CoRL) 2023, Atlanta, USA., Paper | Website

Jun 2023 **HomeRobot: Open-Vocabulary Mobile Manipulation**.
Sriram Yenamandra, Arun Ramachandran, Karmesh Yadav, Austin S Wang, Mukul Khanna, Theophile Gervet, Tsung-Yen Yang, **Vidhi Jain**, Alexander Clegg, John M Turner, Zsolt Kira, Manolis Savva, Angel X Chang, Devendra Singh Chaplot, Dhruv Batra, Roozbeh Mottaghi, Yonatan Bisk, Chris Paxton. *In Proceedings of* Conference on Robot Learning (CoRL) 2023, Atlanta, USA. Paper | Website | Competition @ NeurIPS 2023

Dec 2022 **Transformers are Adaptable Task Planners**.
Vidhi Jain, Yixin Lin, Eric Undersander, Yonatan Bisk, Akshara Rai. *In Proceedings of* Conference on Robot Learning (CoRL) 2022, Auckland, New Zealand. Paper | Website | Video | Code

Workshop Presentations

Dec 2022 **MAEA: Multimodal Attribution for Embodied AI**.
Vidhi Jain, Jayant Sravan Tamarapalli, Sahiti Yerramilli, Yonatan Bisk. *Presented at* NeurIPS 2022, 5th Robot Learning Workshop: Trustworthy Robotics (RLW 2022), New Orleans, USA. Short Paper | Website

Oct 2020 **Predicting Strategies in Simulated Search and Rescue**.
Vidhi Jain, Rohit Jena, Huao Li, Tejus Gupta, Dana Hughes, Michael Lewis, Katia Sycara. *Presented at* NeurIPS AI+HADR 2020, Virtual. Video | Preprint

Oct 2020 **Learning Embeddings that Capture Spatial Semantics for Indoor Navigation**.
Vidhi Jain, Shishir Patil, Prakhar Agarwal, Katia Sycara. *Presented at* NeurIPS 2020 Object Representations for Learning and Reasoning (ORLR 2020), Virtual. Video | Preprint | Code

Sep 2020 **Learning to Navigate in Unseen Cluttered Structured Environments**.
Vidhi Jain, Ganesh Iyer, Katia Sycara. NeurIPS 2020 Women In Machine Learning (WiML 2020), Virtual. Poster

Jul 2020 **Coping with Sample Inefficiency in Deep Reinforcement Learning (DRL) in embodied AI**.
Vidhi Jain, Simin Liu, Ganesh Iyer. ICML'20 Women In Machine Learning (WiML) Un-Workshop, Virtual. Discussion | Presentation

Aug 2017 **Model Selection Scores for Multi-Relational Bayesian Networks**.
Sajjad Gholami, Oliver Schulte,**Vidhi Jain**, Qiang Zhao. IJCAI 2017 Declarative Learning Based Programming (DeLBP). Invited Paper

Academic Projects

Mar 2020 **Learning Diverse Goal-Conditioned Policies for Frontier Selection in Navigation**. *with Ganesh Iyer* Decomposed the task of map exploration into modular differentiable policies that can be combined hierarchically for navigation strategy: (1) implemented the 'global' policy for proposing frontier locations that lead to high coverage, and (2) created sub-maps from a planner for the 'local' policy trained for low-level control. Report

Nov 2019 **Towards Zero-Shot Alignment and Retrieval for Forensic Detection** *with Amrit Setlur and Aditya Pratapa* Designed contrastive learning-based noise removal architecture for matching probed shoe prints with the reference image. Report

Tech Reports

Aug 2021 **Towards Explainable Embodied AI.** Masters Thesis. Report

Jun 2018 **Investigating the viability of Generative Models for Novelty Detection.** Bachelors Thesis. Report

Selected Awards and Honors

Aug 2020 **CIFAR Deep Learning Reinforcement Learning Summer School**

Apr 2020 **Young Researcher at Heidelberg Laureate Forum (HLF '20)** among 224 researchers worldwide.

Jul 2019 **J N Tata Endowment Scholarship for Higher Studies, India**

Jun 2019 **K. C. Mahindra Scholarships for Post-Graduate Studies Abroad**

May 2017 **ACM-W Scholarship for Attendance at Research Conferences**

Mar 2017 **Citi Women Leader Award (CWLA), Mumbai, India**, Awarded one year of study scholarship covering upto INR 400,000 i.e. USD 6000, among the top 3 out of 1200 applicants

Jun 2016 **GE Foundation Scholar-Leaders Program (GEFSLP)**, Awarded USD 2000 for two years of study, among top 5 scholars selected in India

Mar 2016 **National Finalists at Microsoft's 'Build the Shield', Hyderabad, India**, Ranked 15^{th} nationwide

2014–15 **Merit Scholarship** by Dean, BITS Pilani, *Top 1% among the batch of about 900 students*

2013 **Awarded Kishore Vaigyanik Protsahan Yojna (KVPY) Fellowship** by Dept of Science and Tech., Govt.of India

2010 **Awarded National Talent Search Exam (NTSE) Scholarship** by NCERT, India

2010 Ranked Top 1% in Regional State in **National Standard Examination in Junior Science (NSEJS)**

Talks

Jun 2023 **Transformers are Adaptable Task Planners**.
CoRL Spotlight talk
Google RL Reading Group

Dec 2022 **MAEA: Multimodal Attribution Framework for Embodied AI**.
NeurIPS 2022, New Orleans. Conference Video

Oct 2021 ***3 brushes to paint your research canvas***.
SAiDL (Virtual) Video | Event Poster

Dec 2020 **Predicting Human Strategies in Simulated Search and Rescue** .
NeurIPS 2020 Virtual Video

Dec 2020 **Learning Embeddings that Capture Spatial Semantics for Indoor Navigation**.
Neurips 2020 Virtual Video

May 2019 ***Tutorial on Deep Learning with PyTorch***.
PyLadies Bangalore Presentation | Event Poster

Jan 2019 ***The One in Asankhya Project***.
by Sukriti Paul and Mansi Goyal Video | Blog

Aug 2018 ***Invited speaker for 'Research in undergraduate studies'***.
IIIT-Bangalore ACM Student Chapter Video

Jul 2017 **Invited speaker for discussion on 'Innovation by Young India'**.
telecasted on national news channel NDTV India Video

Professional Service and Leadership

Sep'23 **Reviewer** at IEEE International Conference on Robotics and Automation (ICRA).

Jan'20–Present **Reviewer** at International Conference on Machine Learning (ICML), Neural Information Processing and Signal (NeurIPS) and International Conference on Learning Representations (ICLR).

Jun'20–Jul'20 **Workshop Session Organizer** Coping with Sample Inefficiency in Deep Reinforcement Learning (DRL) in embodied AI, ICML'20 WiML Un-Workshop, Virtual. Report | Presentation

Sep'17–May'18 **Co-Founder and Vice Chair** | BITS Pilani Association for Computing Machinery-Women Chapter, *Started with 16 on-campus girls, organized project brainstorming and mentoring sessions*

May'15–Aug'15 **Technical Volunteer** | ARISE Impact, non-profit organization *Developed Android application for Audio-based learning for visually challenged*

Teaching Experience

Jan'24-May'24 **Teaching Assistant** | 11777 Multimodal Machine Learning by Yonatan Bisk. Mentoring 5 teams with 3-5 students each for semester-long research projects. Gave a guest lecture on *how to think of multimodality in Embodied AI and Robotic Manipulation*.

Jan'23–May'23 **Teaching Assistant** | 16720 Computer Vision by Deva Ramanan. Gave a hands-on lecture on *Training vision models in PyTorch* to 106 students enrolled in the course, prepared homework, and graded.

Jan'17–May'17 **Teaching Assistant** | CS F111 Computer Programming by Vishal Gupta, *Mentored 35 students for 4 hours per week for Code programming lab sessions.*

Aug'14–May'15 **Teaching Volunteer** | Computer Literacy Programme, National Service Scheme, *Tutored 12-15 rural natives for 3 hours per week on computer applications.*

References

Yonatan Bisk, Assistant Professor, Carnegie Mellon University, ybisk at andrew.cmu.edu

Debidatta Dwibedi, Research Scientist, Google DeepMind, Mountain View, debidatta at google.com

Akshara Rai, Research Scientist, Meta (previously Facebook), Menlo Park, akshararai at meta.com