Habitat-Web: Learning Embodied Object-Search Strategies from Human Demonstrations at Scale

Ram Ramrakhya[1] Eric Undersander[2] Dhruv Batra[1,2] Abhishek Das[2]
[1]Georgia Institute of Technology [2]Meta AI Research
[1]{ram.ramrakhya,dbatra}@gatech.edu [2]{eundersander,abhshkdz}@fb.com

Abstract

*We present a large-scale study of imitating human demonstrations on tasks that require a virtual robot to search for objects in new environments – (1) ObjectGoal Navigation (*e.g. 'find & go to a chair'*) and (2)* PICK&PLACE *(*e.g. 'find mug, pick mug, find counter, place mug on counter'*). First, we develop a virtual teleoperation data-collection infrastructure – connecting Habitat simulator running in a web browser to Amazon Mechanical Turk, allowing remote users to teleoperate virtual robots, safely and at scale. We collect* $80k$ *demonstrations for* OBJECTNAV *and* $12k$ *demonstrations for* PICK&PLACE*, which is an order of magnitude larger than existing human demonstration datasets in simulation or on real robots. Second, we use this data to answer the question – how does large-scale imitation learning (IL) (which has not been hitherto possible) compare to reinforcement learning (RL) (which is the status quo)? On* OBJECTNAV*, we find that IL (with no bells or whistles) using* $70k$ *human demonstrations outperforms RL using* $240k$ *agent-gathered trajectories. This effectively establishes an 'exchange rate' – a single human demonstration appears to be worth* ~ 4 *agent-gathered ones. Finally, accuracy* vs. *training data size plots show promising scaling behavior, suggesting that simply collecting more demonstrations is likely to advance the state of art further. On* PICK&PLACE*, the comparison is starker – IL agents achieve* $\sim 18\%$ *success on episodes with new object-receptacle locations when trained with* $9.5k$ *human demonstrations, while RL agents fail to get beyond* 0%*. Overall, our work provides compelling evidence for investing in large-scale imitation learning.*
Project page: ram81.github.io/projects/habitat-web.

1. Introduction

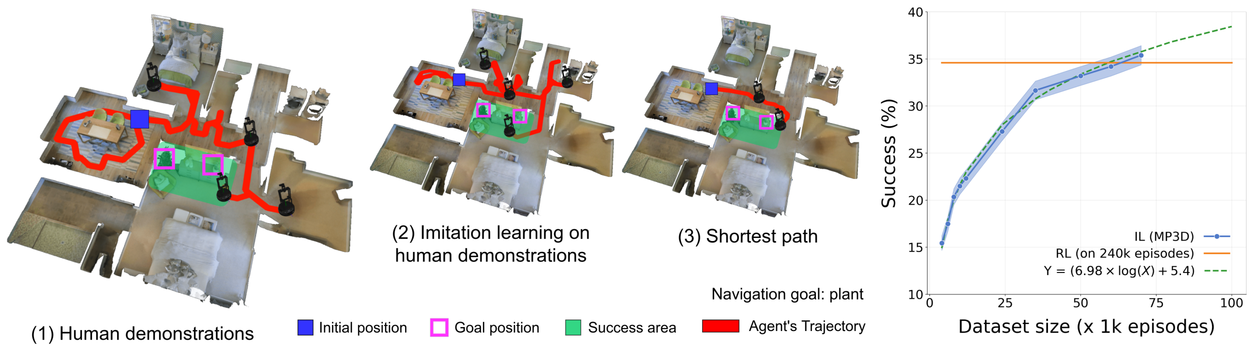
General-purpose robots that can perform a diverse set of embodied tasks in a diverse set of environments *have* to be good at visual exploration. Consider the canonical example of asking a household robot, *'Where are my keys?'*. To answer this (assuming the robot does not remember the answer from memory), the robot would have to search the house, often guided by intelligent priors – *e.g.* peeking into the washroom or kitchen might be sufficient to be reasonably sure the keys are not there, while exhaustively searching the living room might be much more important since keys are more likely to be there. While doing so, the robot has to internally keep track of where all it has been to avoid redundant search, and it might also have to interact with objects, *e.g.* check drawers and cabinets in the living room (but not those in the washroom or kitchen!).

This example illustrates fairly sophisticated exploration, involving a careful interplay of various implicit objectives (semantic priors, exhaustive search, efficient navigation, interaction, *etc.*). Many recent tasks of interest in the embodied AI community – *e.g.* ObjectGoal Navigation [1, 2], rearrangement [3, 4], language-guided navigation [5, 6] and interaction [7], question answering [8–12] – involve some flavor of this visual exploration. With careful reward engineering, reinforcement learning (RL) approaches to these tasks have achieved commendable success [13–17]. However, engineering the 'right' reward function so that the learned policy exhibits desired behavior is unintuitive and frustrating (even for domain experts), expensive (requiring multiple rounds of retraining under different rewards), and not scalable to new tasks or behaviors. For complex tasks (*e.g.* object rearrangement or tasks specified in open-ended natural language), RL from scratch may not even get off the ground.

In this work, we advance the alternative research agenda of imitation learning [18] – *i.e.* collecting a large dataset of human demonstrations (that implicitly capture intelligent behavior we wish to impart to our agents) and learning policies directly from these human demonstrations.

First, we develop a safe scalable virtual teleoperation data-collection infrastructure – connecting the Habitat simulator running in a browser to Amazon Mechanical Turk (AMT). We develop this in way that enables collecting human demonstrations for a variety of tasks being studied within the Habitat [19, 20] ecosystem (*e.g.* PointNav [2], OBJECTNAV [1, 2], ImageNav [21], VLN-CE [6], MultiON [22], *etc.*).

We use this infrastructure to collect human demonstration datasets for 2 tasks requiring visual search – 1) ObjectGoal Navigation (*e.g. 'find & go to a chair'*) and 2) PICK&PLACE (*e.g. 'find mug, pick mug, find counter, place*



(a) (b)

Figure 1. a) Example ObjectNav 1) human demonstration, 2) agent trained on human demonstrations, and 3) shortest path. Notice how humans demonstrate sophisticated exploration behavior to succeed at this task in unseen environments, which is hard to engineer into the right reward for an RL agent and is unlikely to be captured in shortest path demonstrations. An agent trained on human demonstrations learns this exploration and object-search behavior. b) Success on the ObjectNav MP3D-val split *vs*. no. of human demonstrations for training.

on counter'). In total we collect $92k$ human demonstrations, $80k$ demonstrations for ObjectNav and $12k$ demonstrations for Pick&Place. In contrast, the largest existing datasets have 3-$10k$ human demonstrations in simulation [23–25] or on real robots [26, 27], an order of magnitude smaller. The first thing this data provides is a 'human baseline' with sufficiently tight error-bars to be taken seriously. On the ObjectNav validation split, humans achieve $93.7 \pm 0.1\%$ success and $42.5 \pm 0.5\%$ Success Weighted by Path Length (SPL) [2] (*vs*. 34.6% success and 7.9% SPL for the 2021 Habitat ObjectNav Challenge winner [15]). On Pick&Place unseen environments test split, humans achieve 94.9% success and 20.5% SPL (*vs*. 8.3% success and 4.1% SPL for the IL w/ Human Demos). The success rate (93.7% and 94.9%) suggests the task is largely doable for humans (but not 100%). The SPL (42.5% and 20.5%) suggests that even humans need to explore significantly.

Beyond scale, the data is also rich and diverse in the strategies that humans use to solve the tasks. Fig. 1 shows an example trajectory of an AMT user controlling a LoCoBot looking for a 'plant' in a new house – notice the peeking into rooms, looping around the dining table – all of which is (understandably) absent from the shortest path to the goal.

We use this data to answer the question – how does large-scale imitation learning (IL) (which has not been hitherto possible) compare to large-scale reinforcement learning (RL) (which is the status quo)? On ObjectNav, we find that IL (with no bells or whistles) using only $70k$ human demonstrations outperforms RL using $240k$ agent-gathered trajectories (Table 1). This effectively establishes an 'exchange rate' – a single human demonstration appears to be worth $\sim$4 agent-gathered ones. Finally, the accuracy *vs*. training-data-size plot (Fig. 1b) shows promising scaling behavior, suggesting that simply collecting more demonstrations is likely to advance the state of art further. On Pick&Place, the comparison is even starker – IL-agents achieve $\sim$18% success on episodes with new object-receptacle locations, while RL agents fail to get beyond 0% (Table 2).

	Team / Method	Success (↑)	SPL (↑)
1)	DD-PPO baseline [13, 15]	6.2%	2.1%
2)	Active Exploration (Pre-explore)	8.9%	4.1%
3)	SRCB-robot-sudoer	14.4%	7.5%
4)	SemExp [28]	17.9%	7.1%
5)	Red Rabbit (6-Act Base) [15]	24.5%	6.4%
6)	Red Rabbit (6-Act Tether) [15]	21.1%	8.1%
7)	ExploreTillSeen + THDA [16]	21.1%	8.8%
8)	IL w/ $70k$ Human Demos	**27.8%**	**9.9%**
9)	Humans*	93.7%	20.5%

Table 1. Results on Habitat ObjectNav Challenge test-std [29]. * denotes results from val split.

			VAL		TEST	
		Method	Success % (↑)	SPL % (↑)	Success % (↑)	SPL % (↑)
New Inits.	1)	IL w/ Shortest Paths	1.9	1.8	1.7	1.6
	2)	IL w/ Human Demos	17.6 ±0.8	9.7 ±0.3	17.5	9.8
	3)	Humans	87.2	21.8	89.1	21.9
New Instr.	4)	IL w/ Shortest Paths	1.3	1.2	1.1	1.0
	5)	IL w/ Human Demos	15.9 ±0.2	8.4 ±0.4	15.1	8.3
	6)	Humans	85.0	21.0	86.1	20.5
New Envs.	7)	IL w/ Shortest Paths	–	–	0.2	0.3
	8)	IL w/ Human Demos	–	–	8.3	4.1
	9)	Humans	–	–	94.9	20.5

Table 2. Pick-and-place results on splits constructed with unseen initializations in seen environments (1-3), with unseen instructions (4-6), and with unseen environments (7-9).

On both tasks, we find that demonstrations from humans are essential; imitating shortest paths from an oracle produces neither accuracy nor the strategic search behavior. In hindsight, this is perfectly understandable – shortest paths (*e.g*. Fig. 1(a3)) do not contain any exploration but the task requires the agent to explore. Essentially, a shortest path is inimitable, but imitation learning is invaluable. Overall, our work provides compelling evidence for investing in large-scale imitation learning of human demonstrations.

References

[1] M. Savva, A. X. Chang, A. Dosovitskiy, T. Funkhouser, and V. Koltun, "MINOS: Multimodal indoor simulator for navigation in complex environments," *arXiv preprint arXiv:1712.03931*, 2017. 1

[2] P. Anderson, A. X. Chang, D. S. Chaplot, A. Dosovitskiy, S. Gupta, V. Koltun, J. Kosecka, J. Malik, R. Mottaghi, M. Savva, and A. R. Zamir, "On evaluation of embodied navigation agents," *arXiv preprint arXiv:1807.06757*, 2018. 1, 2

[3] D. Batra, A. X. Chang, S. Chernova, A. J. Davison, J. Deng, V. Koltun, S. Levine, J. Malik, I. Mordatch, R. Mottaghi, M. Savva, and H. Su, "Rearrangement: A Challenge for Embodied AI," *arXiv preprint arXiv:2011.01975*, 2020. 1

[4] L. Weihs, M. Deitke, A. Kembhavi, and R. Mottaghi, "Visual room rearrangement," in *CVPR*, 2021. 1

[5] P. Anderson, Q. Wu, D. Teney, J. Bruce, M. Johnson, N. Sünderhauf, I. Reid, S. Gould, and A. van den Hengel, "Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments," in *CVPR*, 2018. 1

[6] J. Krantz, E. Wijmans, A. Majumdar, D. Batra, and S. Lee, "Beyond the nav-graph: Vision-and-language navigation in continuous environments," in *ECCV*, 2020. 1

[7] M. Shridhar, J. Thomason, D. Gordon, Y. Bisk, W. Han, R. Mottaghi, L. Zettlemoyer, and D. Fox, "ALFRED: A Benchmark for Interpreting Grounded Instructions for Everyday Tasks," in *CVPR*, 2020. 1

[8] A. Das, S. Datta, G. Gkioxari, S. Lee, D. Parikh, and D. Batra, "Embodied Question Answering," in *CVPR*, 2018. 1

[9] E. Wijmans, S. Datta, O. Maksymets, A. Das, G. Gkioxari, S. Lee, I. Essa, D. Parikh, and D. Batra, "Embodied Question Answering in Photorealistic Environments with Point Cloud Perception," in *CVPR*, 2019. 1

[10] A. Das, F. Carnevale, H. Merzic, L. Rimell, R. Schneider, J. Abramson, A. Hung, A. Ahuja, S. Clark, G. Wayne, and F. Hill, "Probing emergent semantics in predictive agents via question answering," in *ICML*, 2020. 1

[11] A. Das, *Building agents that can see, talk, and act*. PhD thesis, Georgia Institute of Technology, 2020. 1

[12] L. Yu, X. Chen, G. Gkioxari, M. Bansal, T. L. Berg, and D. Batra, "Multi-target embodied question answering," in *CVPR*, 2019. 1

[13] E. Wijmans, A. Kadian, A. Morcos, S. Lee, I. Essa, D. Parikh, M. Savva, and D. Batra, "DD-PPO: Learning near-perfect pointgoal navigators from 2.5 billion frames," in *ICLR*, 2020. 1, 2

[14] J. Ye, D. Batra, E. Wijmans, and A. Das, "Auxiliary tasks speed up learning pointgoal navigation," in *CoRL*, 2020. 1

[15] J. Ye, D. Batra, A. Das, and E. Wijmans, "Auxiliary Tasks and Exploration Enable ObjectNav," in *ICCV*, 2021. 1, 2

[16] O. Maksymets, V. Cartillier, A. Gokaslan, E. Wijmans, W. Galuba, S. Lee, and D. Batra, "THDA: Treasure Hunt Data Augmentation for Semantic Navigation," in *ICCV*, 2021. 1, 2

[17] D. S. Chaplot, D. Gandhi, A. Gupta, and R. Salakhutdinov, "Object goal navigation using goal-oriented semantic exploration," in *NeurIPS*, 2020. 1

[18] S. Schaal, "Learning from demonstration," in *NIPS*, 1996. 1

[19] M. Savva, A. Kadian, O. Maksymets, Y. Zhao, E. Wijmans, B. Jain, J. Straub, J. Liu, V. Koltun, J. Malik, *et al.*, "Habitat: A platform for embodied AI research," in *ICCV*, 2019. 1

[20] A. Szot, A. Clegg, E. Undersander, E. Wijmans, Y. Zhao, J. Turner, N. Maestre, M. Mukadam, D. S. Chaplot, O. Maksymets, A. Gokaslan, V. Vondrus, S. Dharur, F. Meier, W. Galuba, A. Chang, Z. Kira, V. Koltun, J. Malik, M. Savva, and D. Batra, "Habitat 2.0: Training home assistants to rearrange their habitat," in *NeurIPS*, 2021. 1

[21] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *ICRA*, 2017. 1

[22] S. Wani, S. Patel, U. Jain, A. X. Chang, and M. Savva, "MultiON: Benchmarking Semantic Map Memory using Multi-Object Navigation," in *NeurIPS*, 2020. 1

[23] A. Padmakumar, J. Thomason, A. Shrivastava, P. Lange, A. Narayan-Chen, S. Gella, R. Piramuthu, G. Tur, and D. Hakkani-Tur, "TEACh: Task-driven Embodied Agents that Chat," *arXiv preprint arXiv:2110.00534*, 2021. 2

[24] J. Thomason, M. Murray, M. Cakmak, and L. Zettlemoyer, "Vision-and-dialog navigation," in *CoRL*, 2020. 2

[25] M. Hahn, J. Krantz, D. Batra, D. Parikh, J. M. Rehg, S. Lee, and P. Anderson, "Where are you? localization from embodied dialog," *arXiv preprint arXiv:2011.08277*, 2020. 2

[26] A. Mandlekar, D. Xu, J. Wong, S. Nasiriany, C. Wang, R. Kulkarni, L. Fei-Fei, S. Savarese, Y. Zhu, and R. Martín-Martín, "What matters in learning from offline human demonstrations for robot manipulation," in *CoRL*, 2021. 2

[27] F. Ebert, Y. Yang, K. Schmeckpeper, B. Bucher, G. Georgakis, K. Daniilidis, C. Finn, and S. Levine, "Bridge data: Boosting generalization of robotic skills with cross-domain datasets," *arXiv preprint arXiv:2109.13396*, 2021. 2

[28] D. S. Chaplot, D. Gandhi, A. Gupta, and R. Salakhutdinov, "Object goal navigation using goal-oriented semantic exploration," in *NeurIPS*, 2020. 2

[29] H. Team, "Habitat challenge, 2020." `https://aihabitat.org/challenge/2020`, 2020. 2