https://robotics-transformer-x.github.io

Open X-Embodiment: Robotic Learning Datasets and RT-X Models

Open X-Embodiment Collaboration

(hover to display full author list)

For technical questions, please file a bug at the github repo. For any other inquiries, please email open-x-embodiment@googlegroups.com.

Contributing datasets: if you are interested in contributing datasets to the Open X-Embodiment dataset, please fill out the Dataset Enrollment Form.


Paper


Blogpost


Code


Data

[Submitted on 13 Oct 2023 (v1), last revised 14 May 2025 (this version, v9)]

Open X-Embodiment: Robotic Learning Datasets and RT-X Models

Open X-Embodiment Collaboration, Abby O'Neill, Abdul Rehman, Abhinav Gupta, Abhiram Maddukuri, Abhishek Gupta, Abhishek Padalkar, Abraham Lee, Acorn Pooley, Agrim Gupta, Ajay Mandlekar, Ajinkya Jain, Albert Tung, Alex Bewley, Alex Herzog, Alex Irpan, Alexander Khazatsky, Anant Rai, Anchit Gupta, Andrew Wang, Andrey Kolobov, Anikait Singh, Animesh Garg, Aniruddha Kembhavi, Annie Xie, Anthony Brohan, Antonin Raffin, Archit Sharma, Arefeh Yavary, Arhan Jain, Ashwin Balakrishna, Ayzaan Wahid, Ben Burgess-Limerick, Beomjoon Kim, Bernhard Schölkopf, Blake Wulfe, Brian Ichter, Cewu Lu, Charles Xu, Charlotte Le, Chelsea Finn, Chen Wang, Chenfeng Xu, Cheng Chi, Chenguang Huang, Christine Chan, Christopher Agia, Chuer Pan, Chuyuan Fu, Coline Devin, Danfei Xu, Daniel Morton, Danny Driess, Daphne Chen, Deepak Pathak, Dhruv Shah, Dieter Büchler, Dinesh Jayaraman, Dmitry Kalashnikov, Dorsa Sadigh, Edward Johns, Ethan Foster, Fangchen Liu, Federico Ceola, Fei Xia, Feiyu Zhao, Felipe Vieira Frujeri, Freek Stulp, Gaoyue Zhou, Gaurav S. Sukhatme, Gautam Salhotra, Ge Yan, Gilbert Feng, Giulio Schiavi, Glen Berseth, Gregory Kahn, Guangwen Yang, Guanzhi Wang, Hao Su, Hao-Shu Fang, Haochen Shi, Henghui Bao, Heni Ben Amor, Henrik I Christensen, Hiroki Furuta, Homanga Bharadhwaj, Homer Walke, Hongjie Fang, Huy Ha, Igor Mordatch, Ilija Radosavovic, Isabel Leal, Jacky Liang, Jad Abou-Chakra, Jaehyung Kim, Jaimyn Drake, Jan Peters, Jan Schneider, Jasmine Hsu, Jay Vakil et al. (194 additional authors not shown)

Large, high-capacity models trained on diverse datasets have shown remarkable successes on efficiently tackling downstream applications. In domains from NLP to Computer Vision, this has led to a consolidation of pretrained models, with general pretrained backbones serving as a starting point for many applications. Can such a consolidation happen in robotics? Conventionally, robotic learning methods train a separate model for every application, every robot, and even every environment. Can we instead train generalist X-robot policy that can be adapted efficiently to new robots, tasks, and environments? In this paper, we provide datasets in standardized data formats and models to make it possible to explore this possibility in the context of robotic manipulation, alongside experimental results that provide an example of effective X-robot policies. We assemble a dataset from 22 different robots collected through a collaboration between 21 institutions, demonstrating 527 skills (160266 tasks). We show that a high-capacity model trained on this data, which we call RT-X, exhibits positive transfer and improves the capabilities of multiple robots by leveraging experience from other platforms. More details can be found on the project website this https URL.

https://github.com/google-deepmind/open_x_embodiment

google-deepmind / open_x_embodiment

Type / to search

Code | Issues 54 | Pull requests 1 | Actions | Projects | Security | Insights

open_x_embodiment Public

Watch 26 | Fork 100 | Star 1.6k

main | 1 Branch | 0 Tags | Go to file t | Add file | Code

sicxu-google Merge pull request #105 from annie-zhang12/main 9eeb68b · last month 27 Commits

colabs	delete delete the installation of numpy 1.25.2 and add a war...	5 months ago
imgs	Add colabs.	2 years ago
models	No public description	last year
CONTRIBUTING.md	Initial commit.	2 years ago

About

No description, website, or topics provided.

Readme
Apache-2.0 license
Contributing
Activity
Custom properties
1.6k stars
26 watching

README Contributing Apache-2.0 license

Open X-Embodiment

Open X-Embodiment aims to provide all open-sourced robotic data in the same unified format, for easy downstream consumption.

The first publication using the Open X-Embodiment dataset is `Open X-Embodiment: Robotic Learning Datasets and RT-X Models`

Dataset Access

Releases

No releases published

Packages

No packages published

Contributors 4

- ayzaan Ayzaan Wahid
- annie-zhang12 Annie Zhang
- quanvuong Quan Vuong
- sicxu-google

Languages

Jupyter Notebook 99.8% Python 0.2%

https://docs.google.com/spreadsheets/d/1rPBD77tk60AEIGZrGSODwyyzs5FgCU9Uz3h-3_t2A9g/edit?gid=0#gid=0

This browser version is no longer supported. Please upgrade to a supported browser. Learn More

Open X-Embodiment Dataset Overview

File Edit View Insert Format Data Tools Extensions Help

Share

100% View only

A1

For all data release details, models, Colabs etc, please check out our Github Repo

To filter the datasets based on attributes of your choice:

(1) select the columns you want to filter by (select row 15 and below)

(2) select "Data" --> "Filter views" --> "Create New Filter View" (you can also use one of the views we prepared)

(3) click the little filter symbol next to the column you want to filter by and choose your filtering condition (you can check the example filter views if you're unsure)

(4) once you selected your filters and are happy with the remaining datasets, copy the list of dataset names below (it auto-updates to reflect the remaining datasets) and pas

(5) for convenience, we also provide a correponsing list of citations that you can directly copy into your bib file and latex document to appropriately credit the used datasets

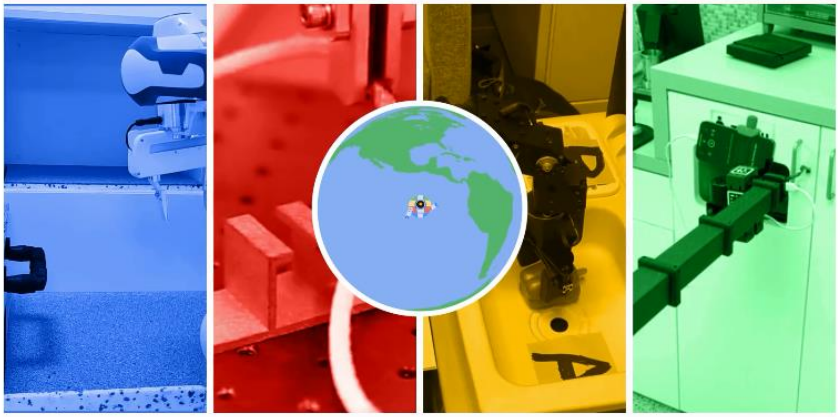
# Total Episodes:	2,419,193
Current Download Size (GB):	8964.94
Dataset Download List:	['fractal20220817_data', 'kuka', 'bridge', 'taco_play', 'jaco_play', 'berkeley_cable_routing', 'roboturk', 'nyu_door_opening_surprising_effectiveness', 'viola', '
Citation List (copy into bib file):	@article{brohan2022rt,
Cite cmd (copy into Latex file):	\citep{brohan2022rt, kalashnikov2018qt, walke2023bridgedata, rosete2022tacorl, mees23hulc2, dass2023jacoplay, luo2023multistage, mandlekar2019sc

Dataset	Robot	# Episodes	File Size (GB)	Robot Morpholo	Gripper	Action Space	# RGB Cams	# Depth Cams	# Wrist Cams	Language Annotatio
RT-1 Robot Action	Google Robot	73,499	111.06	Mobile Manipula	Default	EEF Position	1	1	0	Templated
QT-Opt	Kuka iiwa	580,392	778.02	Single Arm	Default	EEF Position	1	0	0	None
Berkeley Bridge	WidowX	25,460	387.49	Single Arm	Default	EEF Position	4	1	1	Natural
Freiburg Franka Play	Franka	3,242	47.77	Single Arm	Custom 3D printed	EEF Position	2	2	2	Templated
USC Jaco Play	Jaco 2	976	9.24	Single Arm	Default	EEF Position	2	0	1	Templated
Berkeley Cable Routing	Franka	1,482	4.67	Single Arm	Default	EEF velocity	3	0	2	None
Roboturk	Sawyer	2,144	45.39	Single Arm	Default	EEF Position	2	1	0	Templated
NYU VINN	Hello Stretch	435	7.12	Mobile Manipula	Default	EEF Position	1	0	1	None
Austin VIOLA	Franka	135	10.4	Single Arm	Default	EEF Position	2	0	1	Templated
Berkeley Autolab UR5	UR5	896	76.39	Single Arm	Robotiq 2F-85	EEF Position	2	1	1	Templated
TOTO Benchmark	Franka	901	127.66	Single Arm	Default	Joint position	1	0	0	None

Dataset Overview Dataset Metadata Copy of Form Responses 1

Contributing datasets: if you are interested in contributing datasets to the Open X-Embodiment dataset, please fill out the Dataset Enrollment Form.

Paper Blogpost Code Data



Abstract

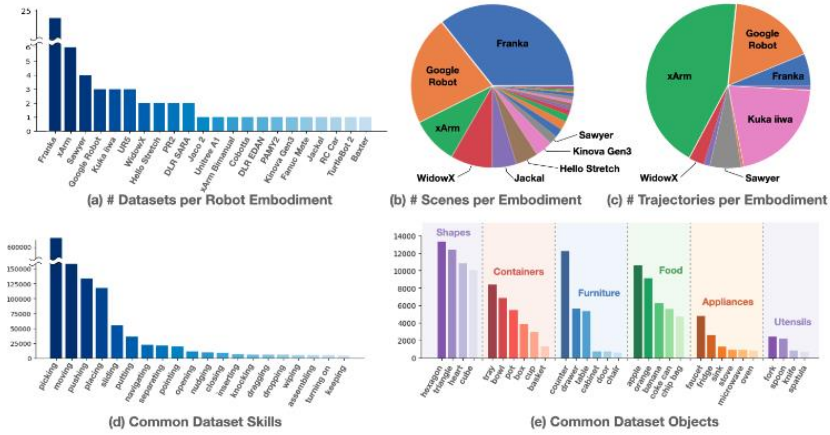
Large, high-capacity models trained on diverse datasets have shown remarkable successes on efficiently tackling downstream applications. In domains from NLP to Computer Vision, this has led to a consolidation of pretrained models, with general pretrained backbones serving as a starting point for many applications. Can such a consolidation happen in robotics? Conventionally, robotic learning methods train a separate model for every application, every robot, and even every environment. Can we instead train "generalist" X-robot policy that can be adapted efficiently to new robots, tasks, and environments? In this paper, we provide datasets in standardized data formats and models to make it possible to explore this possibility in the context of robotic manipulation, alongside experimental results that provide an example of effective X-robot policies. We assemble a dataset from 22 different robots collected through a collaboration between 21 institutions, demonstrating 527 skills (160266 tasks). We show that a high-capacity model trained on this data, which we call RT-X, exhibits positive transfer and improves the capabilities of multiple robots by leveraging experience from other platforms.

move red pepper to tray pick ice cream move red pepper to A

RT-2-X (55B): one of the biggest models to date performing unseen tasks in academic labs

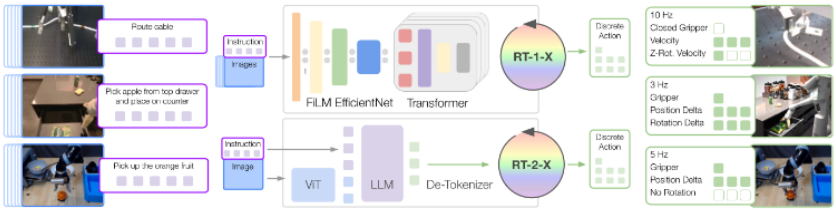
Dataset Overview

We introduce the Open X-Embodiment Dataset, the largest open-source real robot dataset to date. It contains 1M+ real robot trajectories spanning 22 robot embodiments, from single robot arms to bi-manual robots and quadrupeds.



The dataset was constructed by pooling 60 existing robot datasets from 34 robotic research labs around the world. Our analysis shows that the number of visually distinct scenes is well-distributed across different robot embodiments and that the dataset includes a wide range of common behaviors and household objects. For a detailed listing of all included datasets, see this Google Sheet.

Model Overview



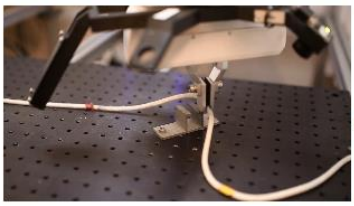
We train two models on the robotics data mixture: (1) RT-1, an efficient Transformer-based architecture designed for robotic control, and (2) RT-2, a large vision-language model co-fine-tuned to output robot actions as natural language tokens.

Both models output robot actions represented with respect to the robot gripper frame. The robot action is a 7-dimensional vector consisting of x, y, z, roll, pitch, yaw, and gripper opening or the rates of these quantities. For data sets where some of these dimensions are not exercised by the robot, during training, we set the value of the corresponding dimensions to zero.

We refer to the RT-1 model trained using the robotic data mixture as RT-1-X, and the RT-2 model trained using the robotic data mixture as RT-2-X.

Results

RT-1-X evaluation on in-distribution skills

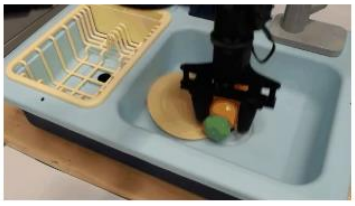


At UC Berkeley (RAIL)

At University of Freiburg (AiS)

At NYU (CILVR)

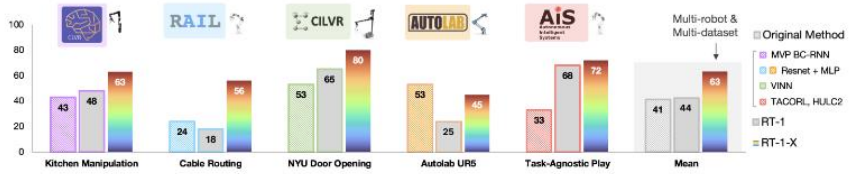
At UC Berkeley (AUTOLab)



At Stanford (IRIS)

At USC (CLVR)

RT-1-X performing diverse tasks in 6 academic labs



RT-1-X models outperform RT-1 or Original Methods trained on individual datasets by 50% in the small-data domain

Original Method refers to the model developed by the creators of the dataset trained only on that respective dataset. The Original Method constitutes a reasonable baseline insofar as it can be expected that the model has been optimized to work well with the associated data. The lab logos indicate the physical location of real robot evaluation, and the robot pictures indicate the embodiment used for the evaluation.

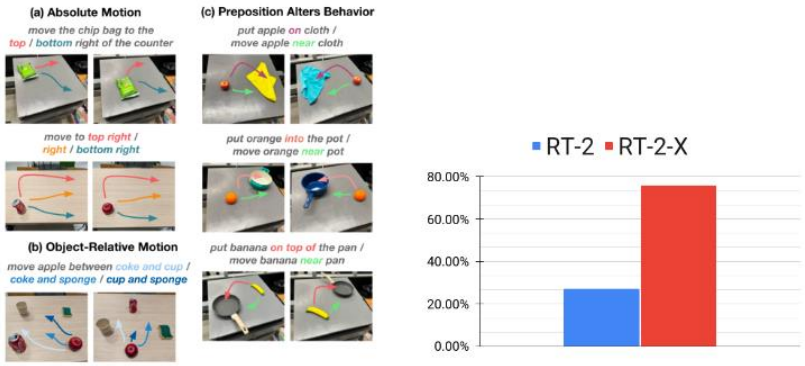
RT-2-X evaluation on emergent skills

move apple near cloth

move apple on cloth

move apple between can & orange

RT-2-X modulates low-level behaviors based on small changes in prepositions (see “on” vs “near” above) and demonstrates understanding of spatial relationships between objects



RT-2-X outperforms RT-2 by 3x in emergent skill evaluations

RT-2-X demonstrates skills that the RT-2 model was not capable of previously, including better spatial understanding in both the absolute and relative sense. Small changes in preposition in the task string can also modulate low-level robot behavior. The skills used for evaluation are illustrated in the figure above.

Citation

If you're using the Open X-Embodiment dataset and RT-X in your research, please cite. If you're specifically using datasets that have been contributed to the joint effort, please cite those as well. We provide a dataset spreadsheet with citation for each dataset for your convenience.

Acknowledgements

We would like to thank John Guilyard for the amazing animations used for this website. The authors would like to acknowledge Yuheng Kuang, Ning Hou, Utsav Malla, Sarah Nguyen, Rochelle Dela Cruz, Justice Carbajal, Brianna Zitkovich, Emily Perez, Elio Prado, Jodilyn Peralta, Tran Pham, Deeksha Manjunath, Samuel Wan, Jaspiar singh and the greater Google DeepMind team for their feedback and contributions. The authors would like to thank Sanah Choudhry, Michael Griessel and Jon Small for their legal advice.

The website template was borrowed from Jon Barron.