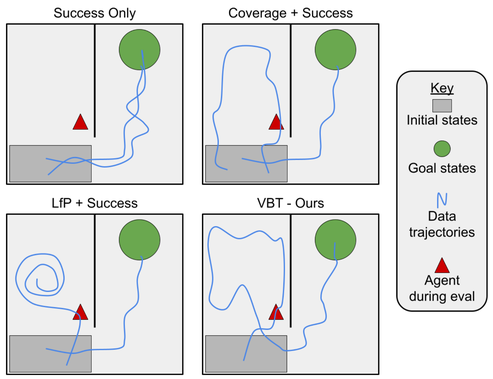
Visual Backtracking Teleoperation: A Data Collection Protocol for Offline Image-Based RL

David Brandfonbrener[1], Stephen Tu[2], Avi Singh[2], Stefan Welker[2], Chad Boodoo[2], Nikolai Matni[2,3], Jake Varley[2]

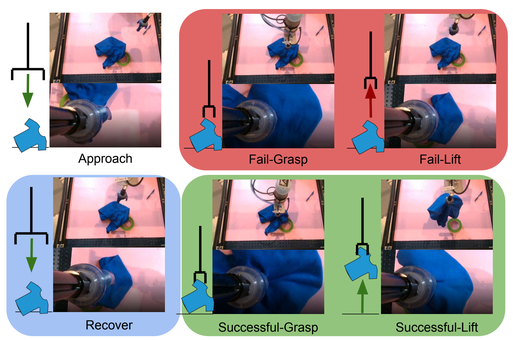
1 NYU 2 Robotics at Google 3 Penn

Motivation

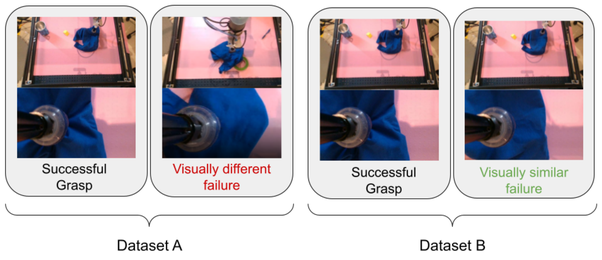
- Behavior cloning (BC) from successful demos can be **brittle**.
- Including **failure and recovery** in the dataset and running offline RL can reduce brittleness by learning to recover.
- But, naively adding failures can lead to **overfitting**.



VBT method



Intution about why VBT reduces overfitting:



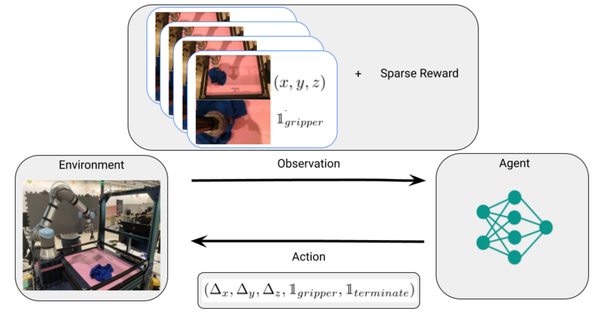
Visual Backtracking Teleoperation (VBT) is a protocol to **collect data** specifically for image-based offline RL.

VBT improves **sample efficiency** relative to standard datasets by improving **coverage** and preventing visual **overfitting.**

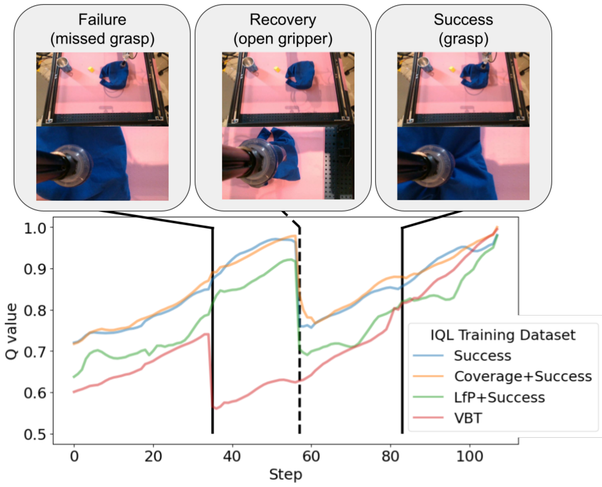
Take a picture to see the project website.'

Experimental results

Real robot setup for visual grasping:



VBT data leads to better visual grasping Q functions:



VBT leads to better visual grasping policies:

Dataset	Policy	Task Success
Success	BC	$66 \pm 4\%$
Success	AWAC	$67 \pm 3\%$
Success	IQL	$69 \pm 3\%$
Coverage+Success	AWAC	$52 \pm 4\%$
Coverage+Success	IQL	$64 \pm 3\%$
LfP+Success	AWAC	$62 \pm 4\%$
LfP+Success	IQL	$58 \pm 4\%$
VBT-Ours	BC	$73 \pm 3\%$
VBT-Ours	AWAC	$73 \pm 3\%$
VBT-Ours	**IQL**	$\mathbf{79 \pm 3\%}$

TABLE I

EVALUATION OF GRASPING ON AN AB TEST OF 1437 TOTAL EPISODES. ERROR BARS REPORT STANDARD ERROR.

References

I. Kostrikov, A. Nair, and S. Levine, "Offline reinforcement learning with implicit q-learning," arXiv preprint arXiv:2110.06169, 2021.
A. Nair, A. Gupta, M. Dalal, and S. Levine, "Awac: Accelerating online reinforcement learning with offline datasets," arXiv preprint arXiv:2006.09359, 2020.
S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp. 627–635, 2011.