robot learning

universität freiburg

Dream to Remember: Continual Steering of Generalist Robot Policies with World Models

Bahey Tharwat[*1], Iman Nematollahi[*1], Marvin Alles[2], Xingyuan Zhang[2], Akshay L Chandra[1], Abhinav Valada[1]

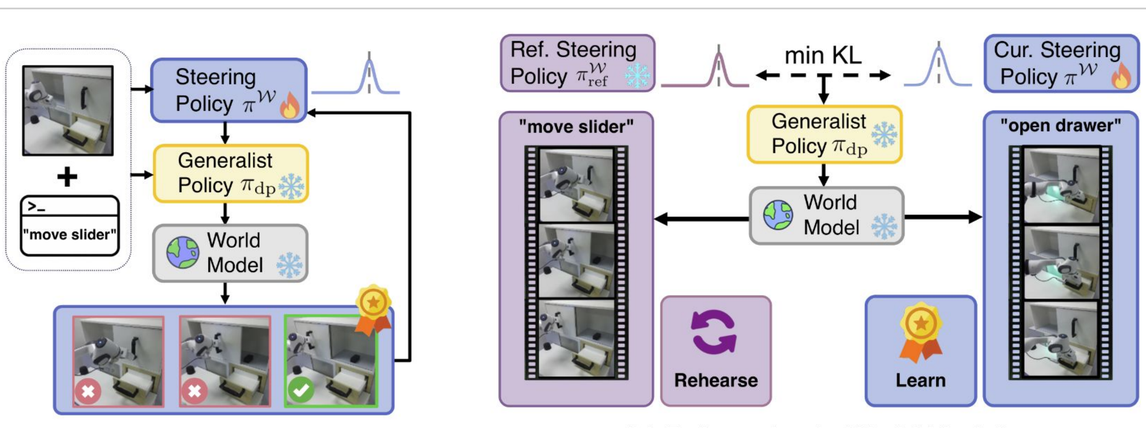
[1] University of Freiburg [2] Technical University of Munich [*] Equal contribution

https://d2r.cs.uni-freiburg.de

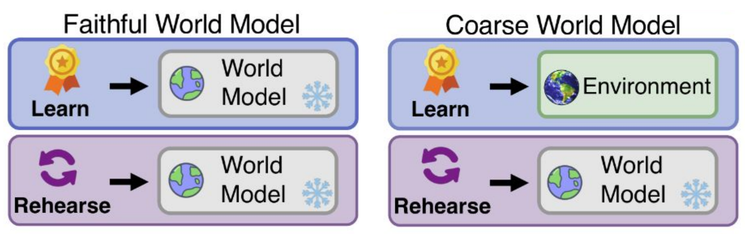
Motivation

- **Catastrophic Forgetting Bottleneck**: Sequential fine-tuning of robot policies degrades the performance of prior skills.
- **Dream to Remember (D2R)**: A framework to continually steer frozen generalist robot policies to learn new tasks while preserving earlier skills through imagined replay.

D2R Framework



(a) Steering in World Model (b) Rehearsing in World Model



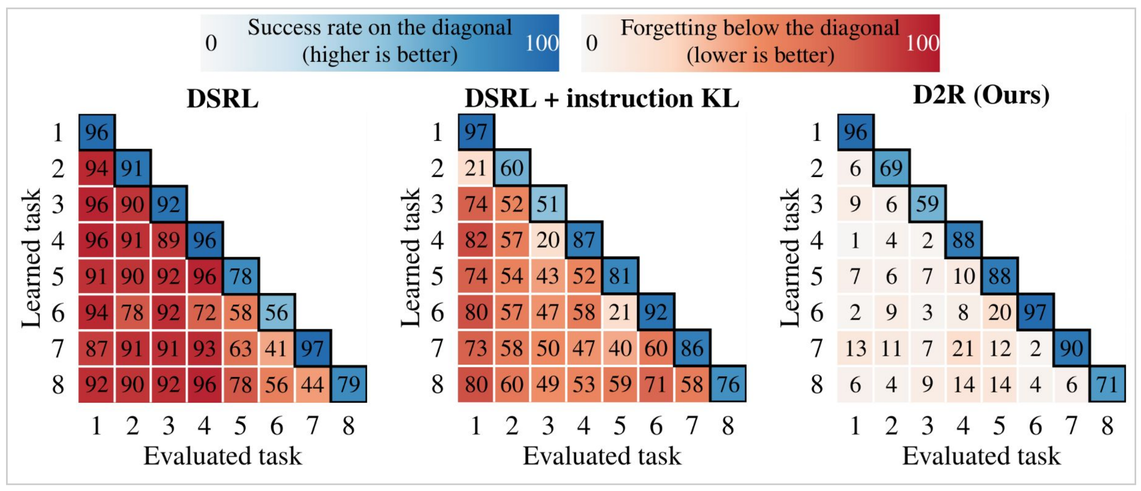
(c) Simulator or Memory

- **Steering in Imagination:** A lightweight steering module learns initial noise vectors fed to the frozen policy via training in imagination.
- **World Model Rehearsal:** Generates imagined rollouts from initial frames, anchoring the policy to previous behavior with a KL regularizer.
- **Simulator vs. Memory:** High-fidelity world models optimize the policy directly, but coarse models act purely as generative memory.

Continual Steering

Continual learning of eight unseen tasks in the LIBERO benchmark.

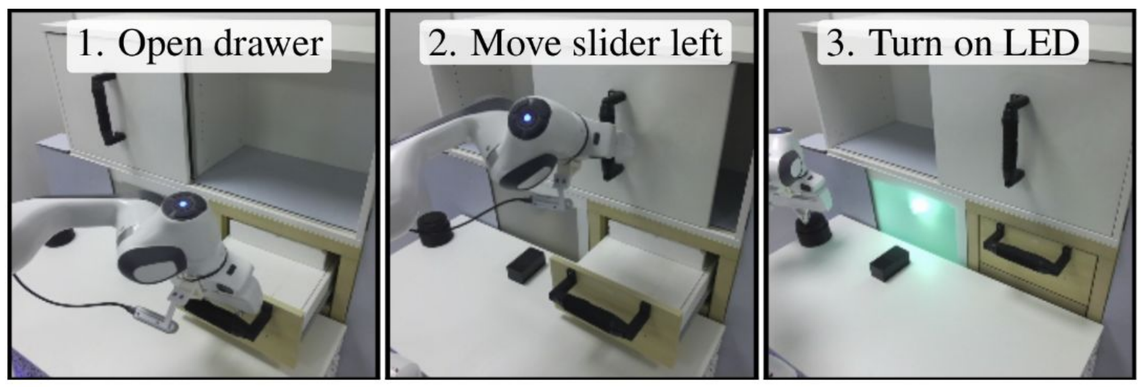
Model	Success↑	Retention↑	Forgetting↓	Storage↓
Base Policy (π_0)	10.4	0.0	0.0	-
LoRA	1.0	-41.6	61.4	-
DSRL	16.9	-30.8	78.3	-
DSRL+Instruct KL	24.9	-23.3	61.4	-
DSRL+Replay KL	73.9	**+2.0**	11.6	60k frames
D2R ★	**74.9**	+1.6	**8.3**	**400 frames**



- **D2R** matches replay buffer continual learning performance, while cutting storage 150 times by generating imagined rollouts.

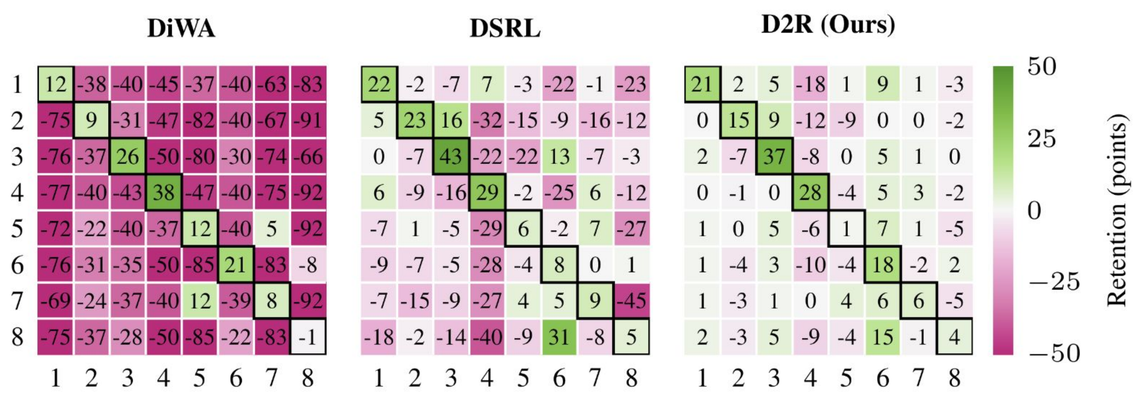
Real-world Experiments

- **Continual Steering in Imagination**: Steering the policy within a learned world model without collecting a single robot expert demonstration.
- **Prior Skills Retention**: Improves +30-point performance over the base policy across all three tasks, while forgetting occurs in DSRL.



Steering in Imagination

- D2R matches weight fine-tuning on a task, while preserving performance across other tasks.



Key Takeaways

- **Continual Steering** of frozen policies reduce forgetting while match fine-tuning performance.
- **World Models as a Memory** anchor prior skills through on-demand imagined rollouts.
- **Dual Roles.** A faithful world model supports policy learning in imagination, whereas a coarse model remains effective for rehearsal memory.