

# **Discovering and Achieving Goals via World Models**



## 4 Diverse Exploration



Images discovered by the explorer policy on the kitchen environment



describe world model, explorer, achiever



Explorer is trained on imagined latent state rollouts of the world model to maximize the disagreement objective

Achiever policy is conditioned on goals and trained on imagined (2) rollouts to minimize a distance function. Goals are sampled randomly from replay buffer images.

Trains a temporal distance function, using the imagined rollouts of the achiever and predicst the number of time steps between states.

# 5 Evaluation on Manipulation Envs

### Kitchen evaluation



### RoboBin evaluation



Russell Mendonca<sup>1</sup>\*, Oleh Rybkin<sup>2</sup>\*, Kostas Daniilidis<sup>2</sup>, Danijar Hafner<sup>3</sup>, Deepak Pathak<sup>1</sup> \* equal contribution <sup>1</sup>Carnegie Mellon University <sup>2</sup>University of Pennsylvania <sup>3</sup> University of Toronto

	0.94	1	0.89	0.56	0.94	0.44	0.33
	1	1	1	0.5	0.5	0	0
	1	0.83	0	0	0.17	0	0
	0	0.67	0	0	0	0	0
	0.23	0.3	0	0	0	0	0
	0.71	0	0	0.1	0	0	0
Ņ	right push	front place	front push	back both	push n front for both	n front bot	push n back

### 3 Goal Image Benchmark



**RoboYoga** : Goals correspond to different body poses for Walker and Quadruped, such as lying down, standing up, and balancing.

**RoboBins :** Sawyer robotic arm with two bins, and two differently colored blocks. Tasks include reaching, manipulating one block, manipulating both blocks.

**RoboKitchen:** Goals specify single or multiple tasks that the franka can do in the kitchen.







# 6 Evaluation on Pose Matching Envs

Ours+ mporal	0.97	0.8	1	0.17	0.3	0.37	0.2	0.2	0.07	0.17	0	0.1
Ours+ Cosine	1	0.9	0.93	0.97	0.87	0.57	0.77	0.93	0.9	0.63	0.57	0.1
DDL	1	0.97	0.5	0.7	0.4	0.53	0.2	0.03	0.13	0.17	0	0.1
DIAYN	0	0	0	0	0	0.03	0	0	0.03	0	0	0
GCSL	0.1	0.2	0	0	0	0	0	0	0	0	0	0
SkewFit	0.09	0.24	0	0	0	0	0	0	0	0	0	0
ijef	ront	Dack leo	5 UP Y	neels	land	feet	Dack a	nglebr	idge st	cand ,	poat	que
1.				cta	nd o.	100	SIDE		near		arar	

~~	Ours+ Temporal	0.63	0.4	0.33	0.23	0.4	0.33	0.33	0.4	0.27	0.27	0.07	0.03
	Ours+ Cosine	0.9	0.7	0.9	0.8	0.73	0.77	0.67	0.4	0.33	0.27	0.37	0.07
	DDL	0.57	0.2	0.63	0.57	0.37	0	0	0.03	0.13	0.1	0.03	0
	DIAYN	0.83	0.87	0.17	0	0	0	0	0	0	0	0.03	0
3 4 moles <sup>1e6</sup>	GCSL	0.63	0.73	0.13	0	0	0.17	0.27	0.27	0	0	0.03	0
— DIAYN	Skewfit	0.17	0.12	0.05	0	0.03	0.11	0.09	0.12	0.01	0	0	0
GCSL Skewfit	•	jie <sup>2</sup>	ile str	etch lie	side balan	ce2 sta	nd <sup>2</sup> s	cand at	cack bala	ile si	de 2 leo	5 UP P	oint

### Website with videos: https://orybkin.github.io/lexa/