

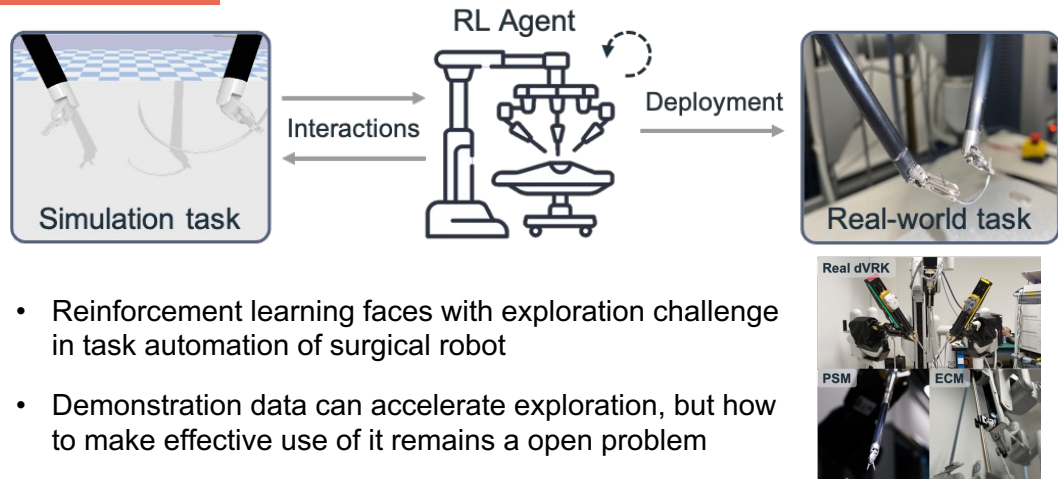


Demonstration-Guided Reinforcement Learning with Efficient Exploration for Task Automation of Surgical Robot



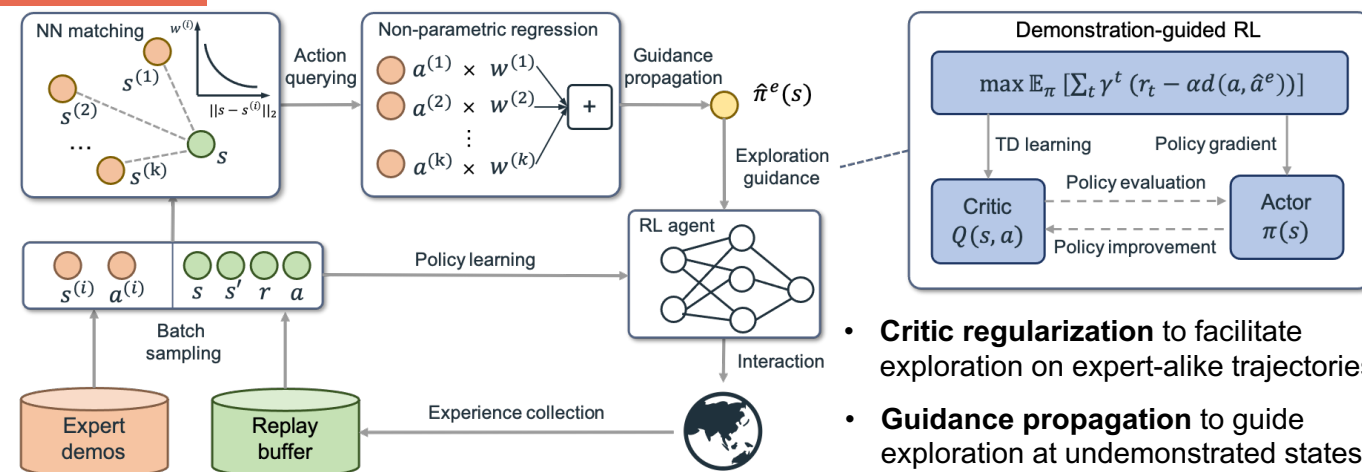
Tao Huang, Kai Chen, Bin Li, Yun-Hui Liu, Qi Dou
The Chinese University of Hong Kong

Background



- Reinforcement learning faces with exploration challenge in task automation of surgical robot
- Demonstration data can accelerate exploration, but how to make effective use of it remains an open problem

Method



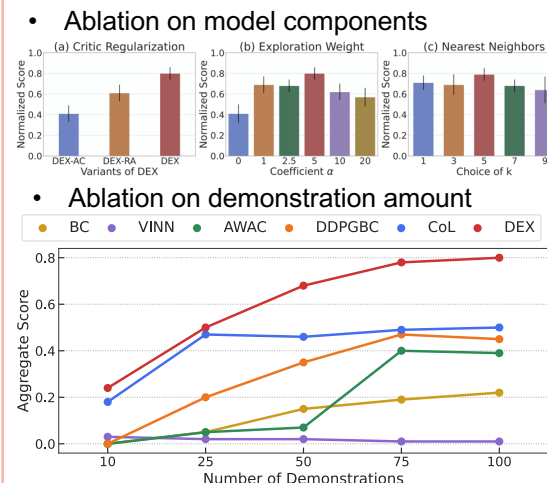
- Critic regularization** to facilitate exploration on expert-alike trajectories
- Guidance propagation** to guide exploration at undemonstrated states

Main Results

Task	Task Description	Reinforcement Learning		Imitation Learning			Demonstration-guided Reinforcement Learning				Ours
		SAC	DDPG	BC	SQIL	VINN	DDPGBC	AMP	CoL	AWAC	DEX
ECM	Aggregate	0.99 (±0.03)	0.99 (±0.02)	1.00 (±0.00)	0.24 (±0.06)	0.58 (±0.06)	1.00 (±0.00)	1.00 (±0.01)	1.00 (±0.00)	0.99 (±0.01)	1.00 (±0.00)
	ECMReach	1.00 (±0.06)	1.00 (±0.00)	1.00 (±0.00)	0.07 (±0.04)	0.49 (±0.10)	1.00 (±0.00)	0.99 (±0.02)	1.00 (±0.00)	1.00 (±0.00)	1.00 (±0.00)
	StaticTrack	0.92 (±0.14)	0.98 (±0.05)	1.00 (±0.00)	0.43 (±0.26)	0.56 (±0.10)	1.00 (±0.00)	0.97 (±0.03)	1.00 (±0.00)	1.00 (±0.00)	1.00 (±0.00)
	MisOrient	1.00 (±0.00)	1.00 (±0.00)	1.00 (±0.00)	0.56 (±0.10)	0.50 (±0.11)	0.99 (±0.02)	0.98 (±0.02)	0.99 (±0.02)	0.98 (±0.03)	0.99 (±0.02)
	ActiveTrack	0.79 (±0.08)	0.67 (±0.08)	0.95 (±0.01)	0.07 (±0.06)	0.92 (±0.09)	0.81 (±0.05)	0.94 (±0.01)	0.96 (±0.01)	0.51 (±0.12)	0.94 (±0.01)
PSM	Aggregate	0.00 (±0.00)	0.00 (±0.00)	0.40 (±0.05)	0.00 (±0.00)	0.02 (±0.02)	0.80 (±0.04)	0.00 (±0.00)	0.85 (±0.06)	0.46 (±0.19)	0.89 (±0.03)
	NeedleReach	1.00 (±0.00)	1.00 (±0.00)	1.00 (±0.00)	0.07 (±0.09)	0.89 (±0.06)	1.00 (±0.00)	0.99 (±0.02)	1.00 (±0.00)	0.94 (±0.20)	1.00 (±0.00)
	GauzeRetrieve	0.00 (±0.00)	0.00 (±0.00)	0.07 (±0.05)	0.00 (±0.00)	0.01 (±0.02)	0.63 (±0.11)	0.00 (±0.00)	0.71 (±0.16)	0.43 (±0.43)	0.73 (±0.12)
	NeedlePick	0.00 (±0.00)	0.00 (±0.00)	0.21 (±0.06)	0.00 (±0.00)	0.02 (±0.02)	0.91 (±0.05)	0.00 (±0.00)	0.96 (±0.05)	0.26 (±0.33)	0.94 (±0.05)
BI-PSM	PegTransfer	0.00 (±0.00)	0.00 (±0.00)	0.56 (±0.11)	0.02 (±0.05)	0.05 (±0.04)	0.48 (±0.22)	0.00 (±0.00)	0.23 (±0.23)	0.31 (±0.32)	0.73 (±0.20)
	Aggregate	0.00 (±0.00)	0.00 (±0.00)	0.08 (±0.04)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.39 (±0.11)
	NeedleRegrasp	0.00 (±0.00)	0.00 (±0.00)	0.09 (±0.03)	0.01 (±0.01)	0.01 (±0.02)	0.05 (±0.08)	0.00 (±0.00)	0.04 (±0.07)	0.00 (±0.00)	0.63 (±0.19)
BiPegTransfer	0.00 (±0.00)	0.00 (±0.00)	0.09 (±0.05)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.00 (±0.00)	0.01 (±0.02)	0.00 (±0.00)	0.18 (±0.14)	
Overall	-	0.46 (±0.03)	0.45 (±0.01)	0.68 (±0.02)	0.02 (±0.02)	0.24 (±0.03)	0.83 (±0.05)	0.48 (±0.01)	0.87 (±0.03)	0.58 (±0.08)	0.92 (±0.02)

- Our method significantly outperforms prior RL-based approaches on the surgical robot learning tasks from SurRoL, especially on complex bi-manual tasks

Ablation



Robot Evaluation

