Example-Driven Model-Based Reinforcement Learning for Solving Long-Horizon Visuomotor Tasks Bohan Wu, Suraj Nair, Li Fei-Fei[†], Chelsea Finn[†] († Equal Advising) **Stanford University**

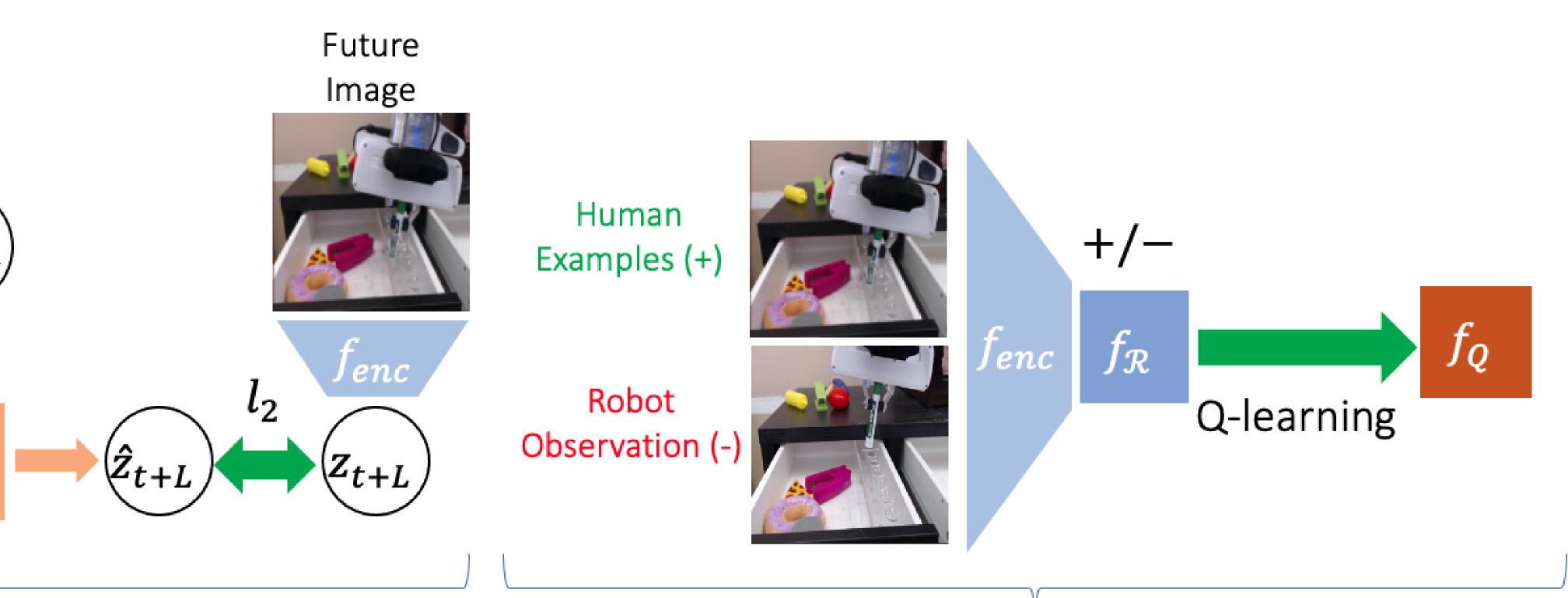
Long-Horizon Vision-Based Manipulation		Example-E
Demands learning a repertoire	e of visuomotor skills that are:	A framework for learning grounde
Robust High Success Rates	Persistent Closed-loop and Reactive	Step 1: Learn a repertoir
		Step 2: Specify model of
Key Insight		Step 3: Run symb
Use human-provided example images as supervision to learn a repertoire of skills, groundings, and success detectors		Learni
 Provides supervision signal for Enables skill grounding in lon Encourages closed-loop visuo Performance 	g-horizon task planning	Image Reconstruction $\widehat{a_{t:t+L-1}}$
<image/>	<image/>	<section-header></section-header>
(a) Robot Environment (b) Trai Object		۲ Trained on data acros
Successful trials (out of 20) and success rates (%)		Learns a low-o
EMB	R EMBR EMBR w/o f_{vae} ,	
(Our	s) w/o $f_{\mathcal{T}}$ $f_{\mathcal{T}}$ (Qt-Opt)	Real-Robot Experiments: I
Avg. Skill Success 96.59	% 89.1% 85.9%	
Avg. Long-Horizon Task Success	% 70.0% 58.3%	inize Des I Cabinet
	Scalable Deep Reinforcement Sobotic Manipulation, CoRL'18	
	ikeaways	P S S
Q-function is important for low-level skill performance		arke arke
Robustness of replanning w/ model is critical for		5 ž
long-horizon performance Future Work		
Future	S VVUIK	e si si si se
Reduce the amount of human supervision		bjec
• Expand task scope and handle partial observation		Be De la

Long-Horizon Vision-Based Manipulation	n Example-E
Demands learning a repertoire of visuomotor skills t	hat are: A framework for learning ground
RobustPersistentHigh Success RatesClosed-loop and Read	active Step 1: Learn a repertoir
	Step 2: Specify model
Key Insight	Step 3: Run symb
Use human-provided example images as supervision to repertoire of skills, groundings, and success detectors	b learn a Learni
 Provides supervision signal for reward learning in RL Enables skill grounding in long-horizon task planning Encourages closed-loop visuomotor control Performance Evaluation 	Deconstruction
<image/>	<image/>
	Test jects
Successful trials (out of 20) and success rates (%	
EMBR EMBR EMBR w	$/o f_{vae},$
$(Ours)$ w/o $f_{\mathcal{T}}$ $f_{\mathcal{T}}$ (Qt-	
Avg. Skill Success 96.5% 89.1% 85.9	
Avg. Long-Horizon Task Success 85.0% 70.0% 58.39	
* Kalashnikov et al. QT-Opt: Scalable Deep Reinfo Learning for Vision-Based Robotic Manipulation, (
Key Takeaways	P S S
 Q-function is important for low-level skill performance Robustness of replanning w/ model is critical for long-horizon performance 	(- 0) = (- 0)
Future Work	
 Reduce the amount of human supervision Expand task scope and handle partial observation 	Rearrang Objects

³Generalization to new environment setup

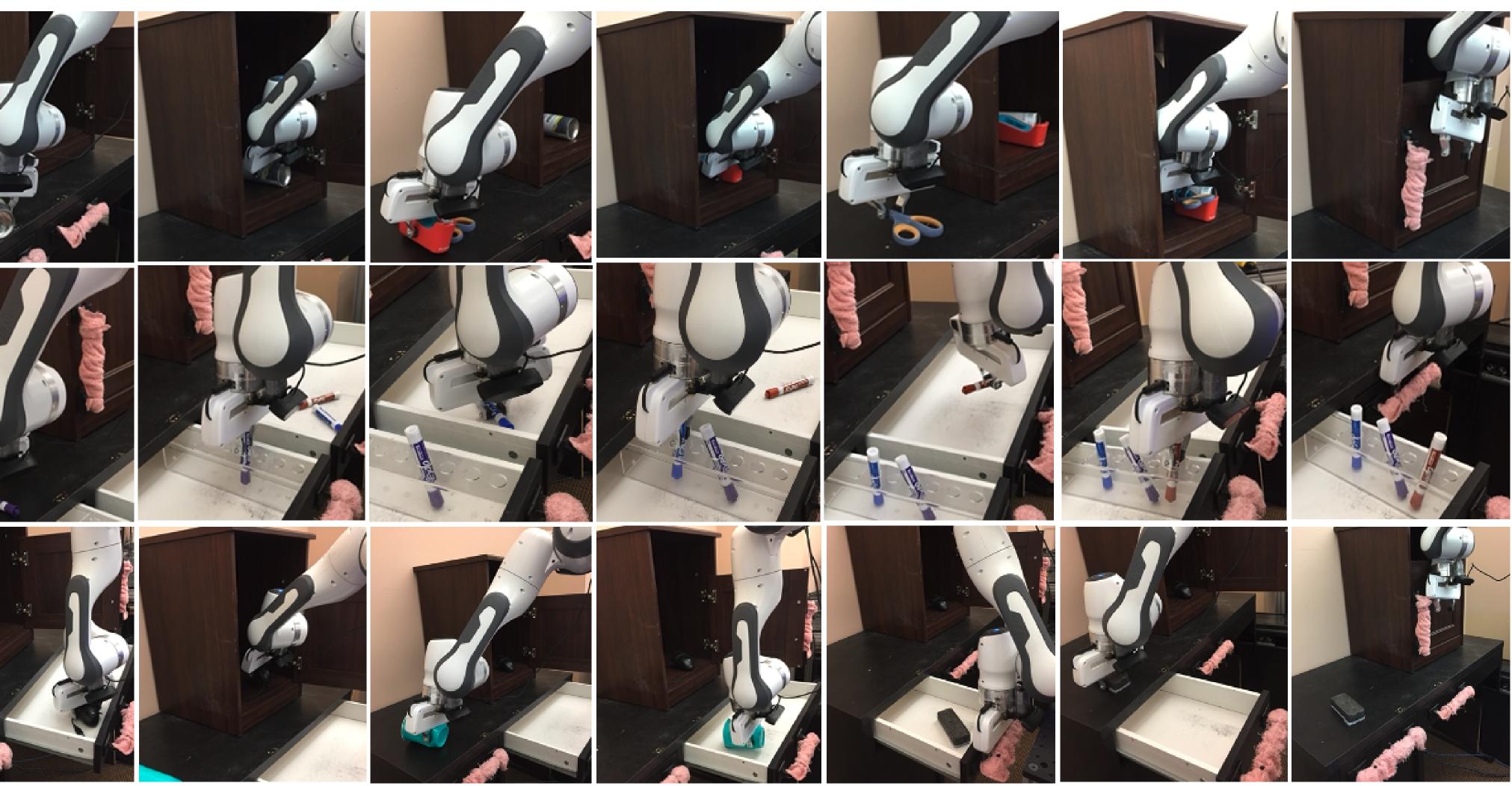
-Driven Model-Based Reinforcement Learning (EMBR)

ded visuomotor skills sequenced by symbolic planners to complete long-horizon tasks. ore of skills with example-driven model-based reinforcement learning (EMBR) over skills, using image classifiers for representing pre- and post-conditions bolic planner with the repertoire of visuomotor skills learned in Step 1 ning a Repertoire of Visuomotor Skills with EMBR



oss all skills

-dimensional latent space with a VAE and a latent dynamics model Obtains rewards by learning image classifiers Learns Q-functions for model-based control Long-Horizon Tasks with Novel Objects from Raw Image Observations





Trained on skill-specific data