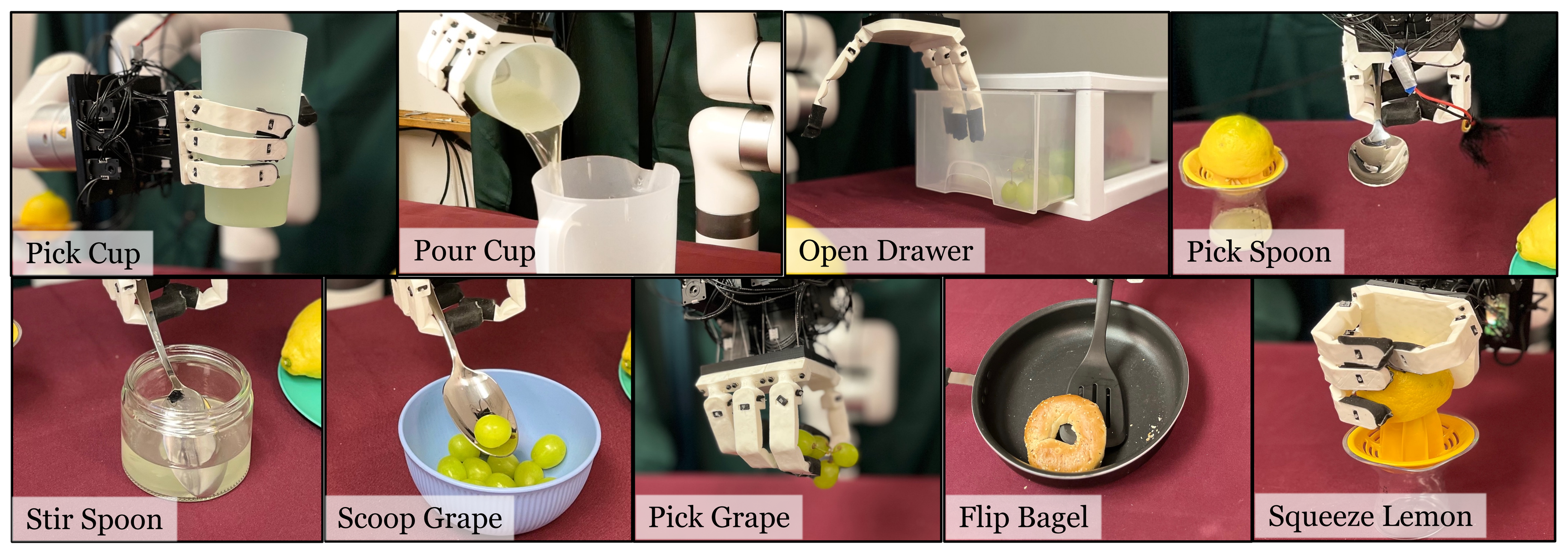
DEFT: Dexterous Fine-Tuning for Hand Policies

Carnegie Mellon University

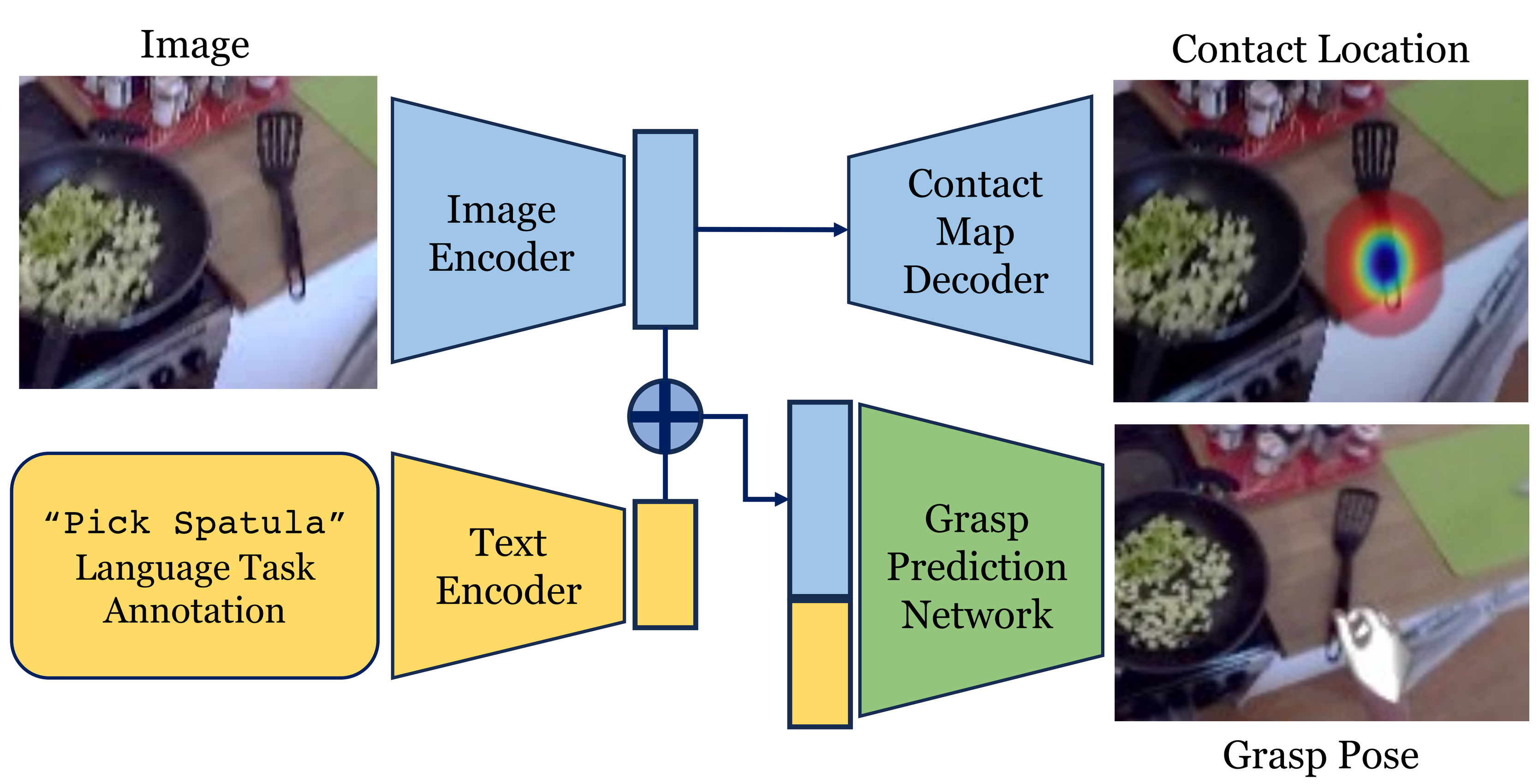
Aditya Kannan*, Kenneth Shaw*, Shikhar Bahl, Pragna Mannam, Deepak Pathak

dexterous-finetuning.github.io

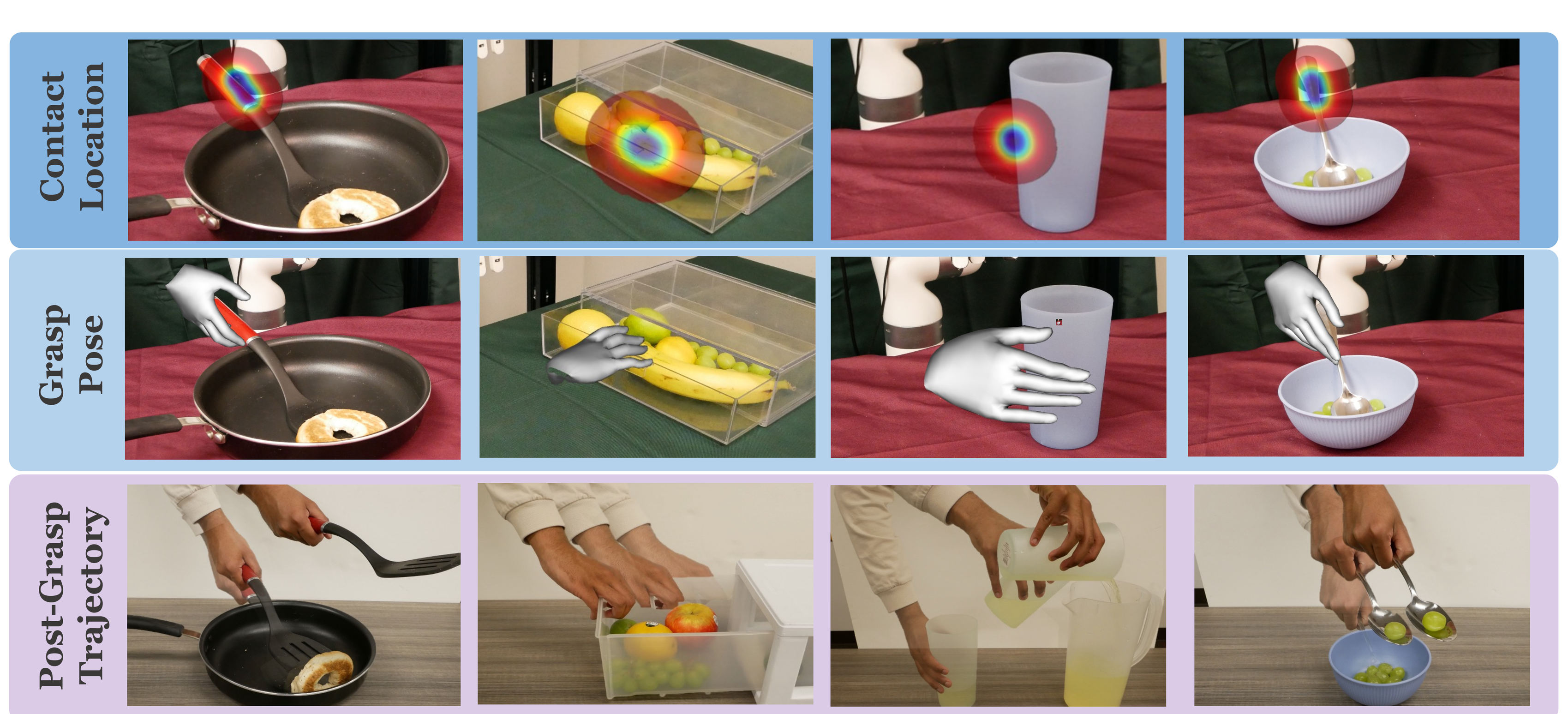
How can we *learn* and *improve* dexterous policies in the real world directly?



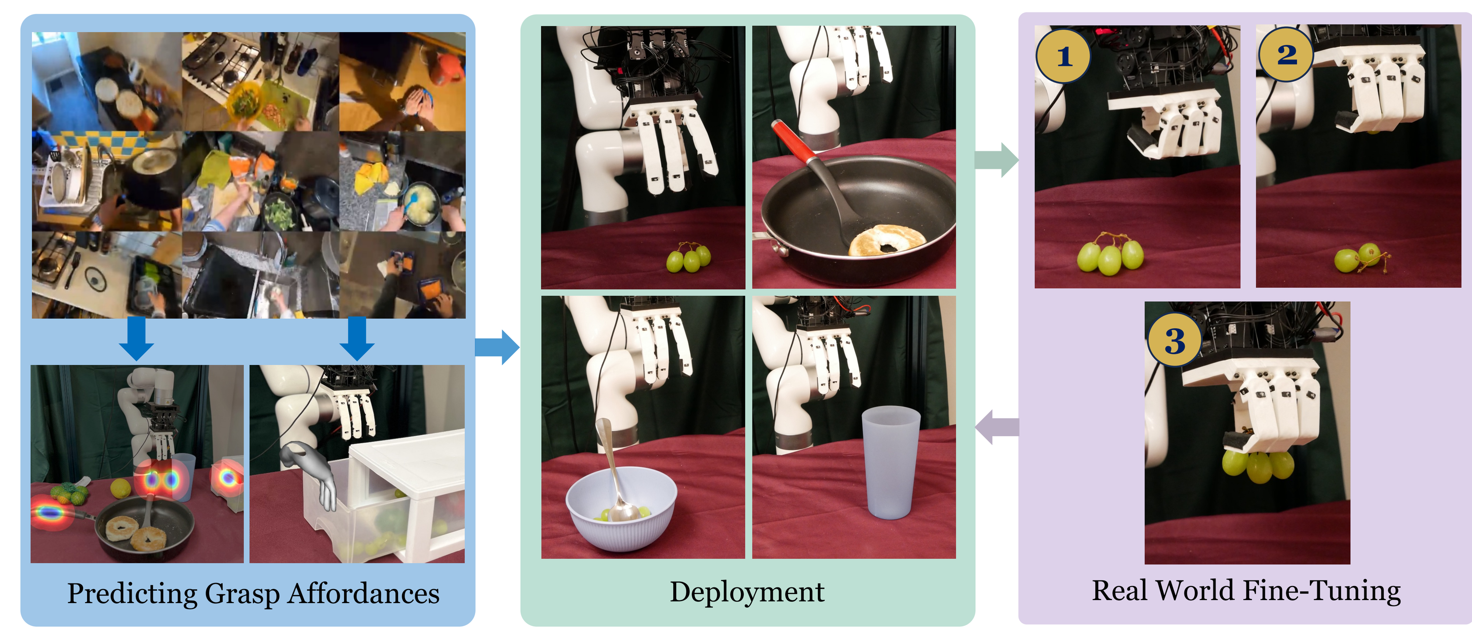
Nine tasks, including manipulating **tools** and soft objects.



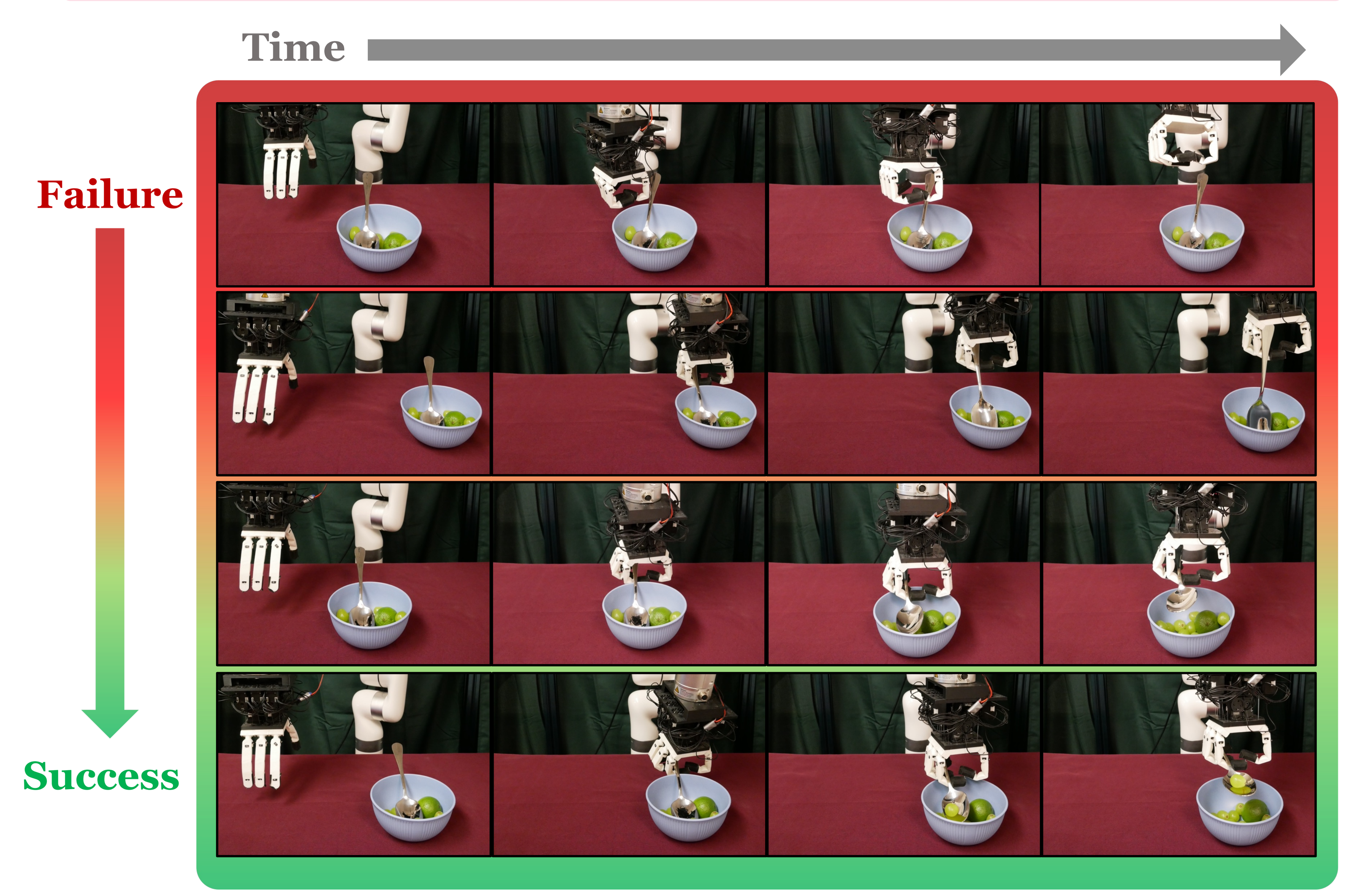
Our affordance model predicts wrist location and hand pose.



We predict *where* (**top**) and *how* (**middle**) to grasp.
We use human video for post-grasp trajectory (**bottom**).



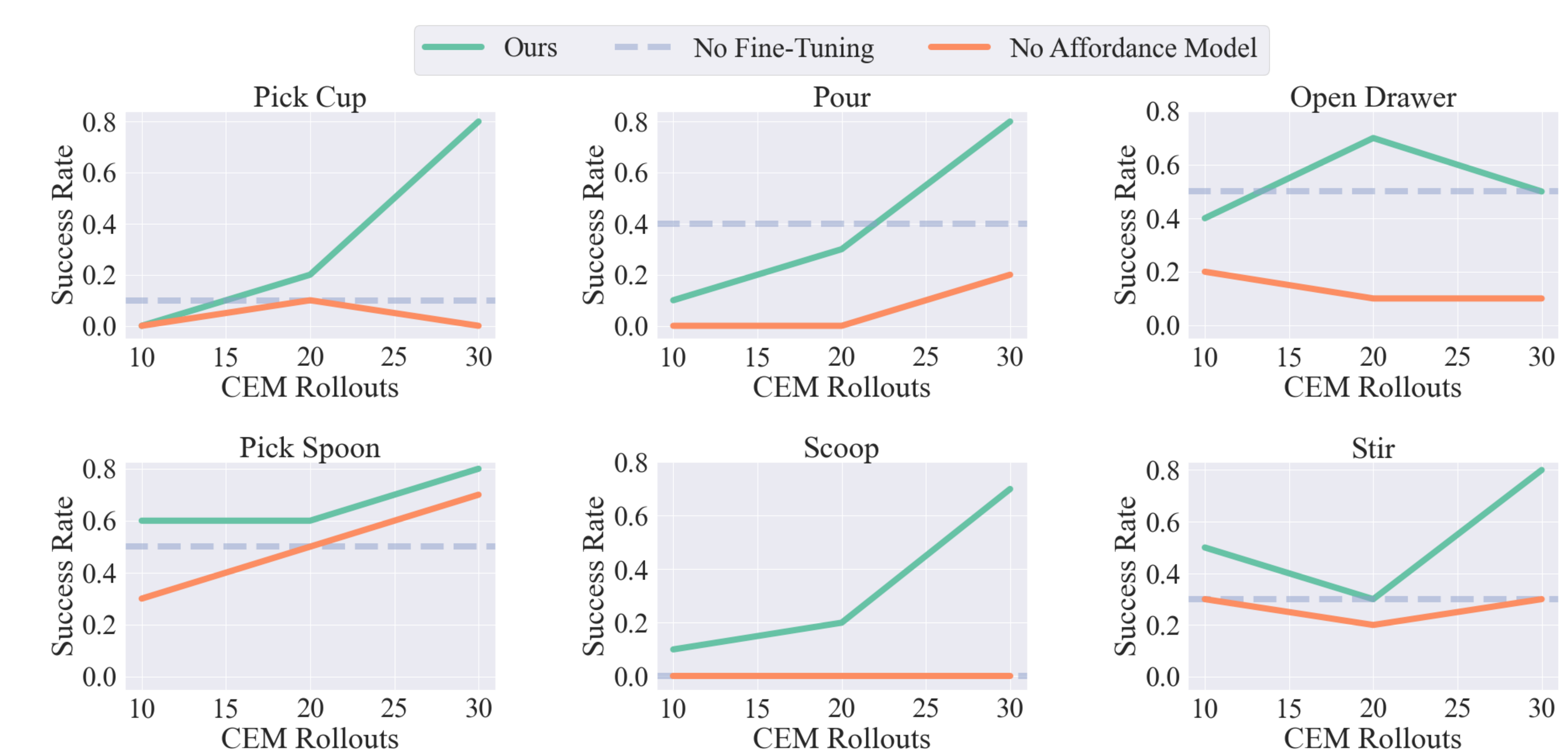
We train a grasp affordance model from human videos and retarget it to a Soft Hand.



We fine-tune the grasp affordance model in the real world.

Left: Workspace Setup. **Right**: Objects used in experiments.

DEFT improves beyond the grasp affordance prior over 30 iterations of CEM.



Method	Pick cup train	Pick cup test	Pour cup train	Pour cup test	Open drawer train	Open drawer test	Pick spoon train	Pick spoon test	Scoop Grape train	Scoop Grape test	Stir Spoon train	Stir Spoon test
Real-World Only	0.0	0.1	0.2	0.1	0.1	0.0	0.7	0.3	0.0	0.0	0.3	0.0
Affordance Model Only	0.1		0.4		**0.5**		0.5		0.0		0.3	
DEFT	**0.8**	**0.8**	**0.8**	**0.9**	**0.5**	**0.4**	**0.8**	**0.6**	**0.7**	**0.3**	**0.8**	**0.5**

No Fine-Tuning is a zero-shot application of our affordance model; No Affordance Model uses a heuristic as the prior.

Method	Pour Cup train	Pour Cup test	Open Drawer train	Open Drawer test	Pick Spoon train	Pick Spoon test
Reward Function:						
R3M Reward	0.0	0.0	0.4	**0.5**	0.5	0.4
Resnet18 Imagenet Reward	0.1	0.2	0.3	0.1	0.4	0.2
Policy Ablation:						
DEFT w/ MLP	0.0	0.0	0.5	0.0	0.6	0.5
DEFT w/ Transformer	0.4	0.5	**0.6**	0.1	0.4	0.5
DEFT w/ Direct Parameter est.	0.1	0.1	0.1	0.0	0.3	0.0
DEFT	**0.8**	**0.9**	0.5	0.4	**0.8**	**0.6**

Ablations for (1) reward function type, (2) model architecture, and (3) parameter estimation approach.