

Environment

Intervention

RLPD Update

CTDE



Human

Policy



Weights

Feature Extraction

Proprio Encoder

Image Encoder

Critic

Multi-Head

$Q(s, a)$

$Q_{task}(s_c, a_c)$

$Q_{grasp}(s_d, a_d)$

Q_{sub}

Actor
Agent 0 Agent 1

Arm

Gripper

...

Trajectories

Batch

Replay Buffer

Transitions

Interventions

Policy Buffer

Demos

Interventions

Demo Buffer

- rollout
- intervention
- episode
- store transition
- training
- update weight

Asynchronous Training

