NRI: FND: Better robotic manipulation using state and action abstraction

PI: Robert Platt
Northeastern University

Objective

This work proposes equivariant methods for making reinforcement learning in SE(2) space more sample efficient.

Group-Invariant MDP

We focus on manipulation problems with embedded in the space. We define the Group-Invariant MDP as an MDP associated with a group : that satisfies the following assumptions:

Reward Invariance: The reward function is group invariant, i.e., where .

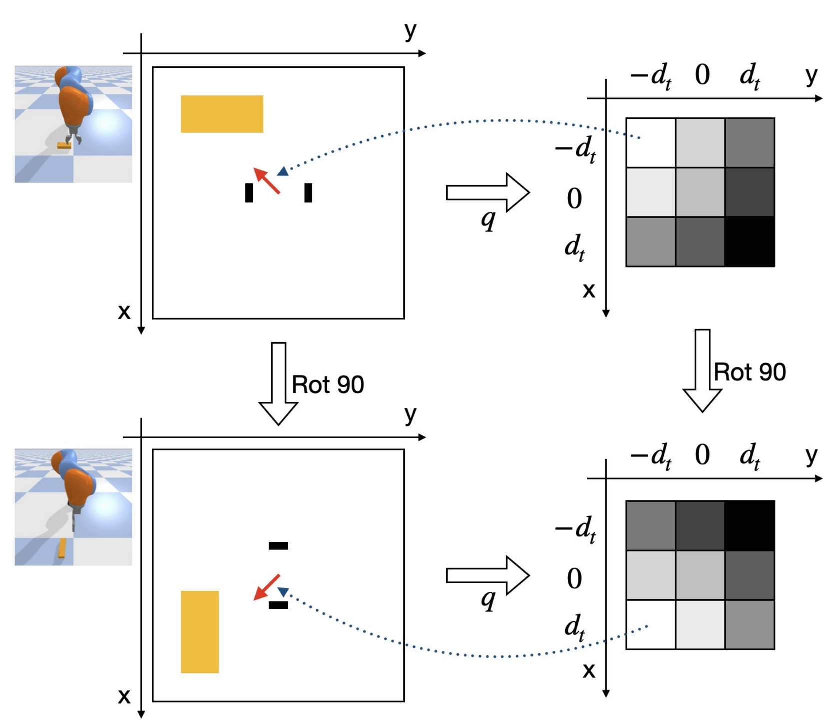
Transition Invariance: The transition function is group invariant, i.e., where .

Approach

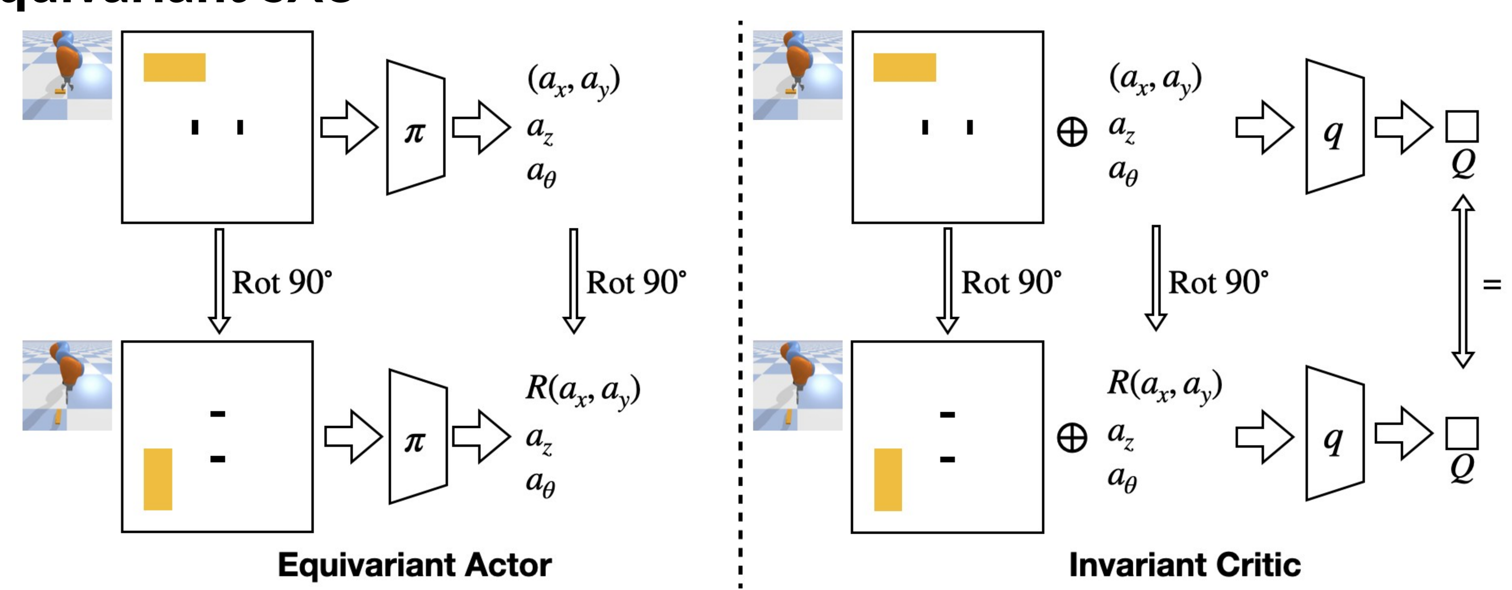
We propose two approaches, **Equivariant DQN** and **Equivariant SAC**.

Equivariant DQN

- Each pixel in the output represents a moving direction of the gripper in the xy plane.
- When the input image rotates (i.e., the state rotates), the output of Equivariant DQN rotates accordingly.
- Such equivariant properties are implemented using Steerable CNNs



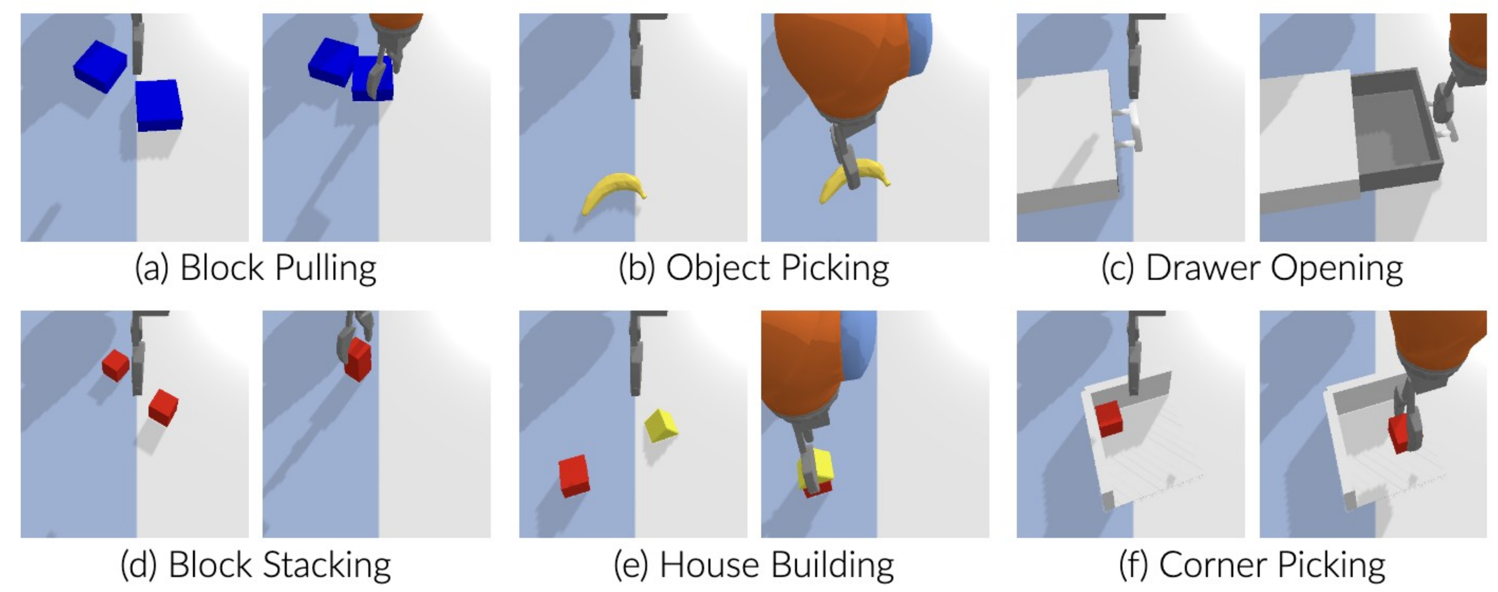
Equivariant SAC



Equivariant Actor: When the input state image rotates, the vector in the output rotates by the rotation matrix .

Invariant Critic: When the input state image and action rotate by the same amount, the output doesn't change.

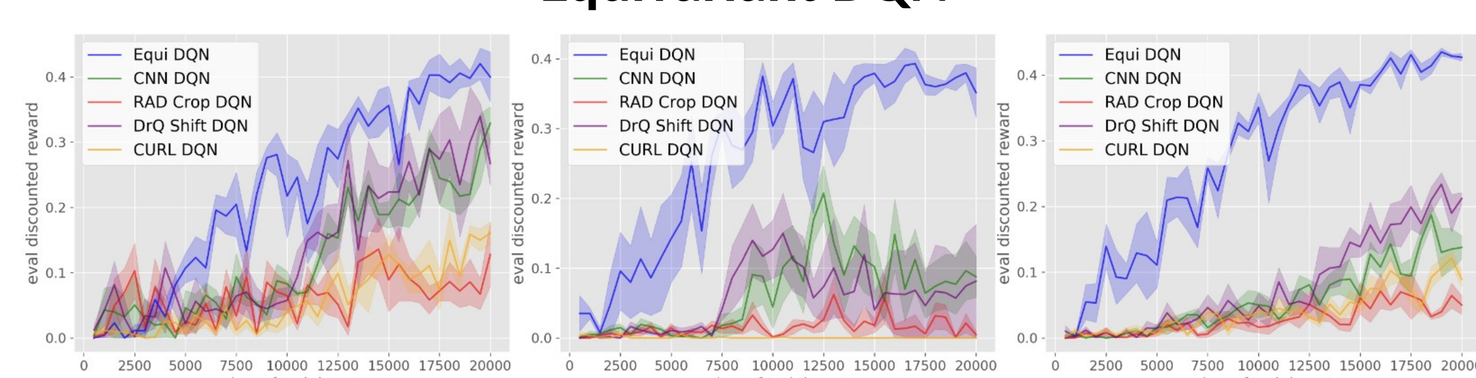
Experiments



(a) Block Pulling (b) Object Picking (c) Drawer Opening (d) Block Stacking (e) House Building (f) Corner Picking

The experimental environments. The left images show the initial state of each environment; the right images show the goal state.

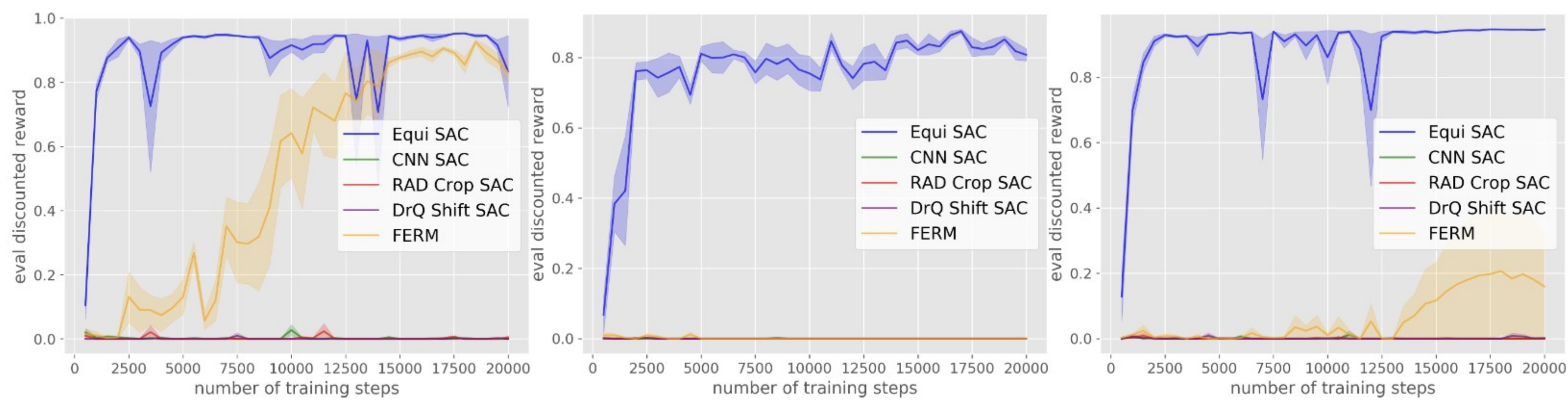
Equivariant DQN



(a) Block Pulling (b) Object Picking (c) Drawer Opening

Comparison of Equivariant DQN (blue) with baselines. Results are averaged over four runs. Shading denotes standard error.

Equivariant SAC

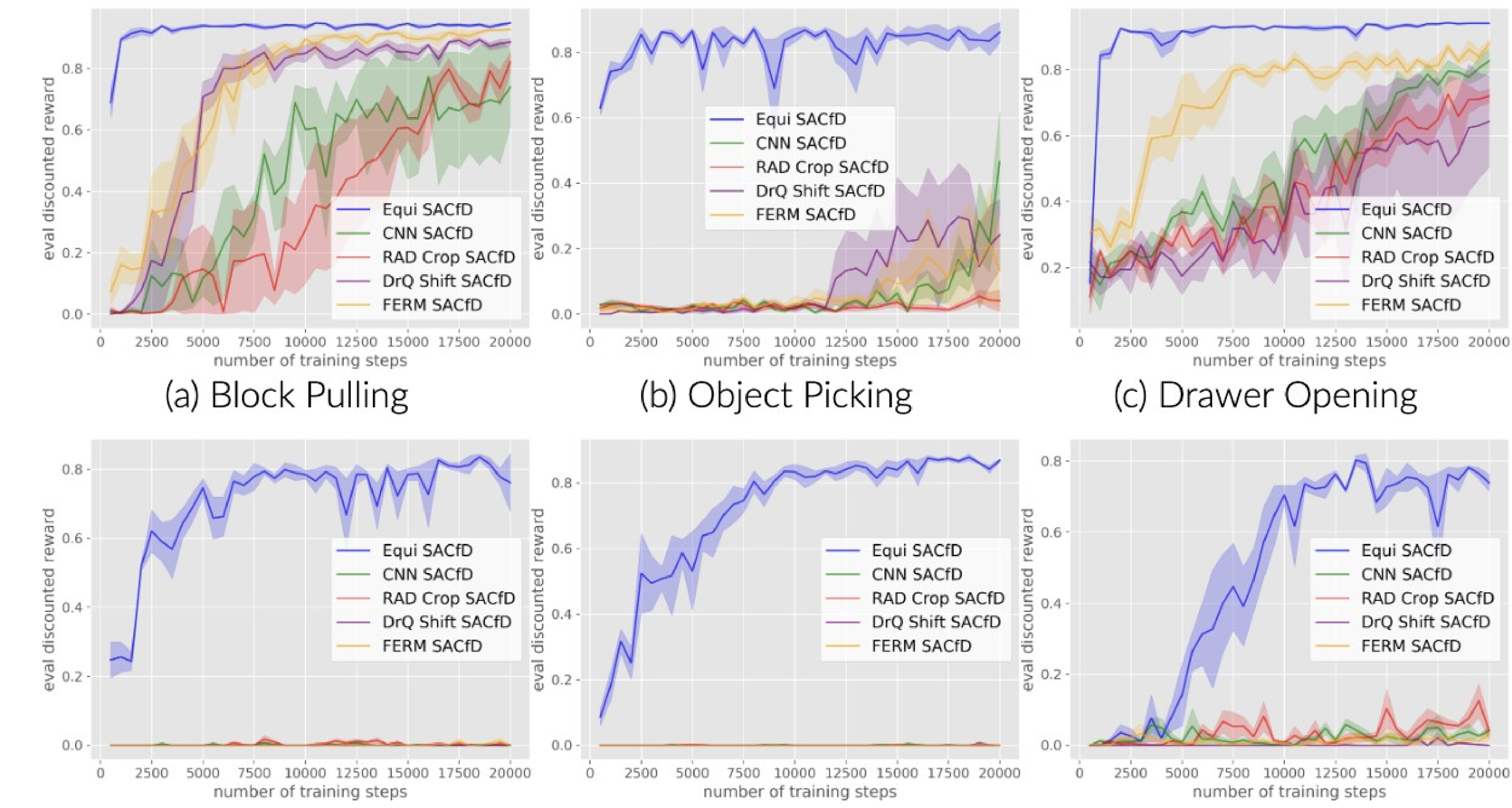


(a) Block Pulling (b) Object Picking (c) Drawer Opening

Comparison of Equivariant SAC (blue) with baselines. Results are averaged over four runs. Shading denotes standard error.

- In both DQN and SAC, our equivariant approach (blue) significantly outperforms all baselines.

Equivariant SACfD



(a) Block Pulling (b) Object Picking (c) Drawer Opening
(d) Block Stacking (e) House Building (f) Corner Picking

Comparison of Equivariant SACfD (blue) with baselines. Results are averaged over four runs. Shading denotes standard error.

- SACfD: An l2 loss is added to the actor to imitate the expert.
- In all environments, our equivariant approach (blue) significantly outperforms all baselines.

2022 NRI & FRR Principal Investigators' Meeting
April 19-21, 2022

Award ID#: 80NSSC19K1474