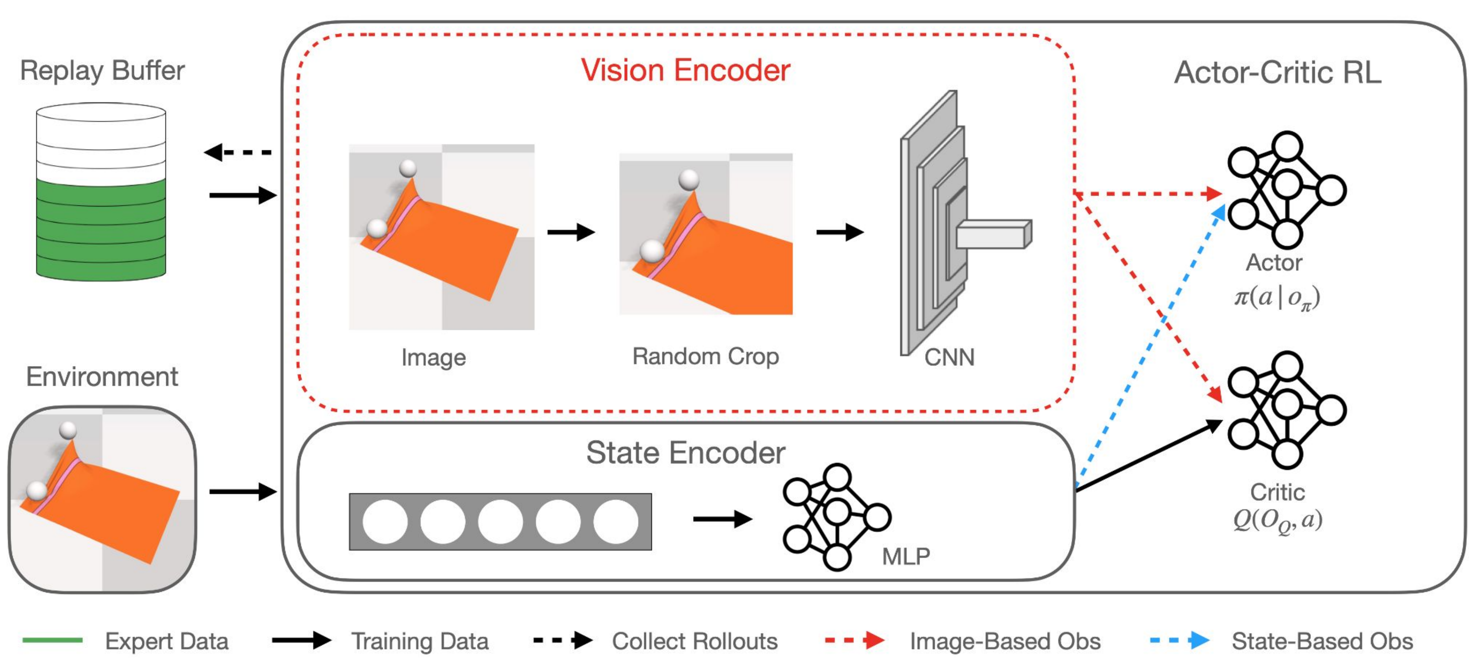
Learning Deformable Manipulation from Expert Demonstrations (DMfD)

Gautam Salhotra*, I-Chun Arthur Liu*, Marcus Dominguez-Kuhne, Gaurav S. Sukhatme

Why deformable manipulation?

- Autonomous deformable manipulation is important for versatile robots performing household activities
  - Cloth folding
  - Cooking
  - Bed covering
- However, it has many challenges
  - Reduced observability and controllability
    - High-dimensional configuration space
    - Complex object dynamics
  - Self-occlusion

Our Method (DMfD)



- Solve deformable manipulation tasks with expert demonstrations
  - Works with **state or image inputs**
  - Works for **1-D or 2-D deformables** (tested on Softgym[1])
- DMfD exceeds baseline performance by up to **12.9%** for state-based tasks and up to **33.44%** on image-based tasks,
  - Comparable or better robustness than SOTA
- Additionally, **two new challenging environments** for folding a 2D cloth using image-based observations

Advantage-weighted formulation with entropy regularization

- Advantage-weighted samples in replay buffer, to encourage policy to stay close to stored expert actions [2]

$$\mathcal{L}_A = \mathop{\mathbb{E}}_{\boldsymbol{s},\boldsymbol{a}\sim\mathcal{B}}\left[\log \pi_\theta(\boldsymbol{a}|\boldsymbol{s}) \exp\left(\frac{1}{\lambda}A^\pi(\boldsymbol{s},\boldsymbol{a})\right)\right]$$

- Entropy regularization to explore online

$$\mathcal{L}_E = \mathop{\mathbb{E}}_{\boldsymbol{s},\boldsymbol{a},\boldsymbol{o}\sim\mathcal{B}}[\alpha \log \pi_\theta(\boldsymbol{a}|\boldsymbol{o}) - Q(\boldsymbol{s},\boldsymbol{a})]$$

- Actor loss that balances the two

$$\mathcal{L}_\pi = (1-w_E)\mathcal{L}_A + w_E\mathcal{L}_E$$

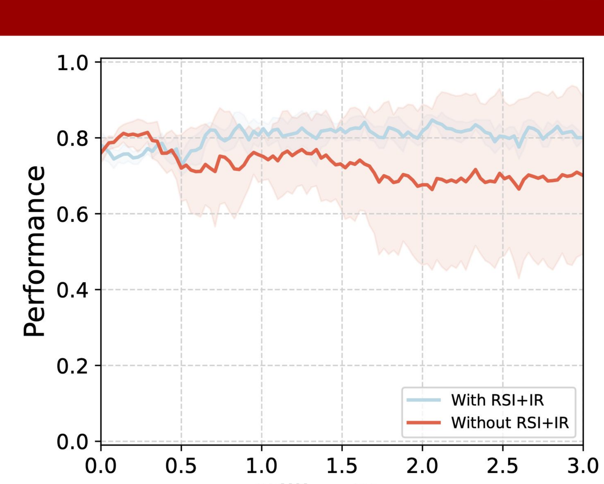
Reference state initialization

- Reset the agent to states seen by experts (hard to reach) [3]
- Compare state trajectories of expert and agent (imitation)

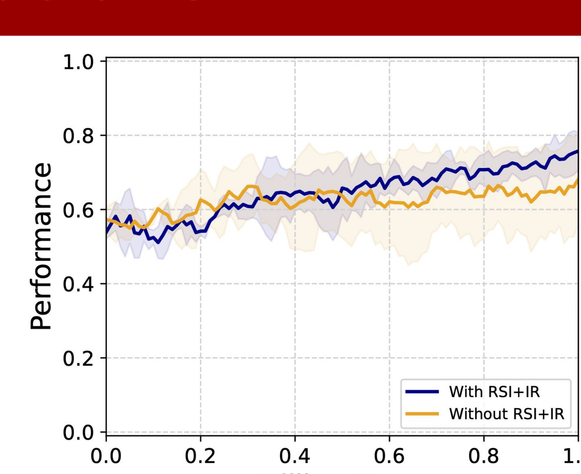
Privileged Critic

- Critic always has state information via the state encoder
- Stabilizes the critic, better value estimation

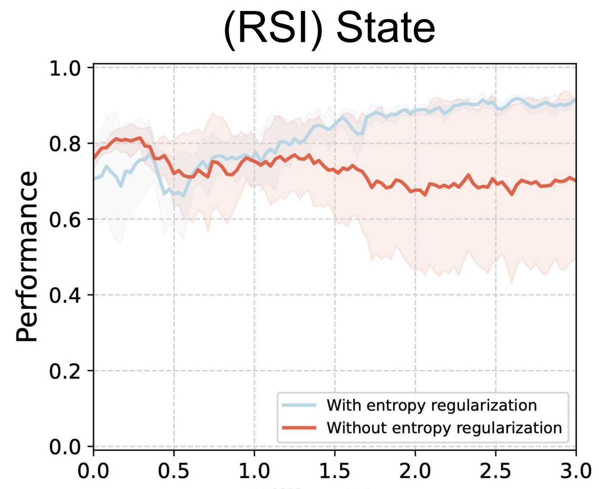
Ablations



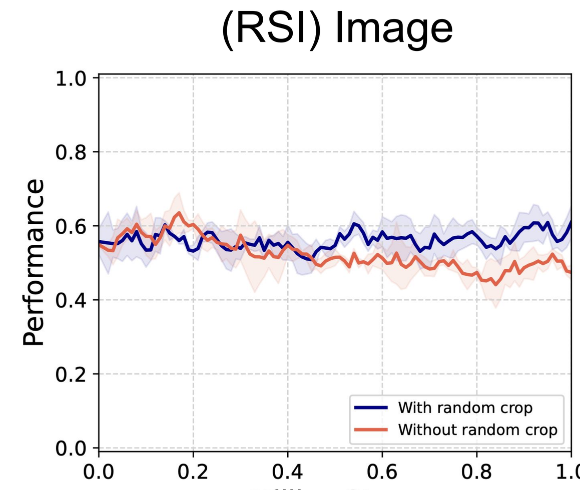
Reference State Initialization (RSI) State



Reference State Initialization (RSI) Image

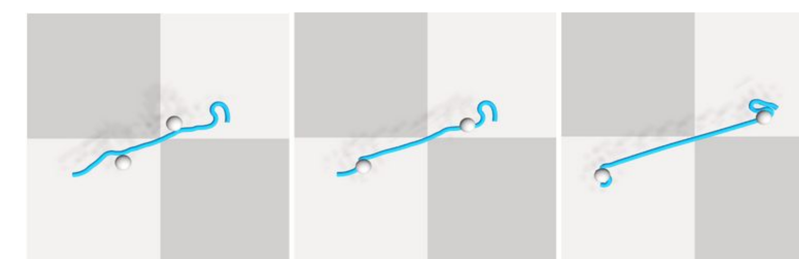


Entropy Regularization

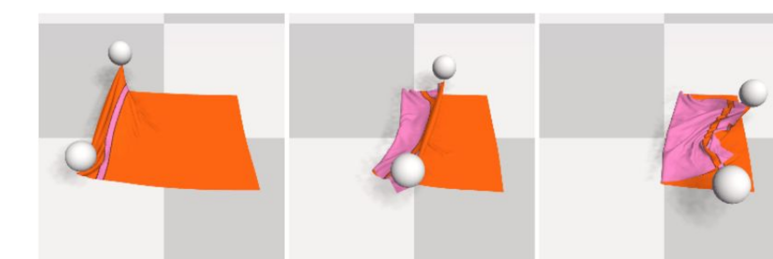


Random Crop of Images

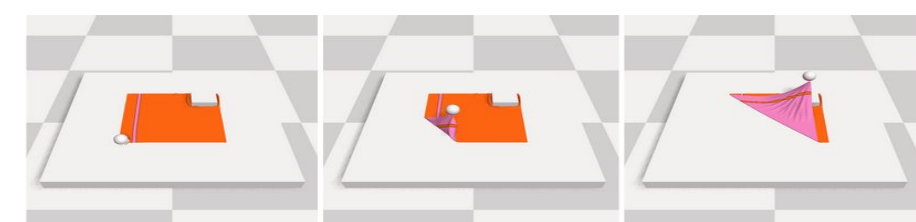
Environments



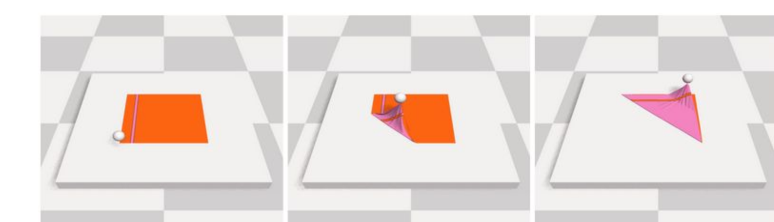
a) Straighten Rope



b) Fold Cloth along Edge



c) Fold Cloth Diagonal (pinned)

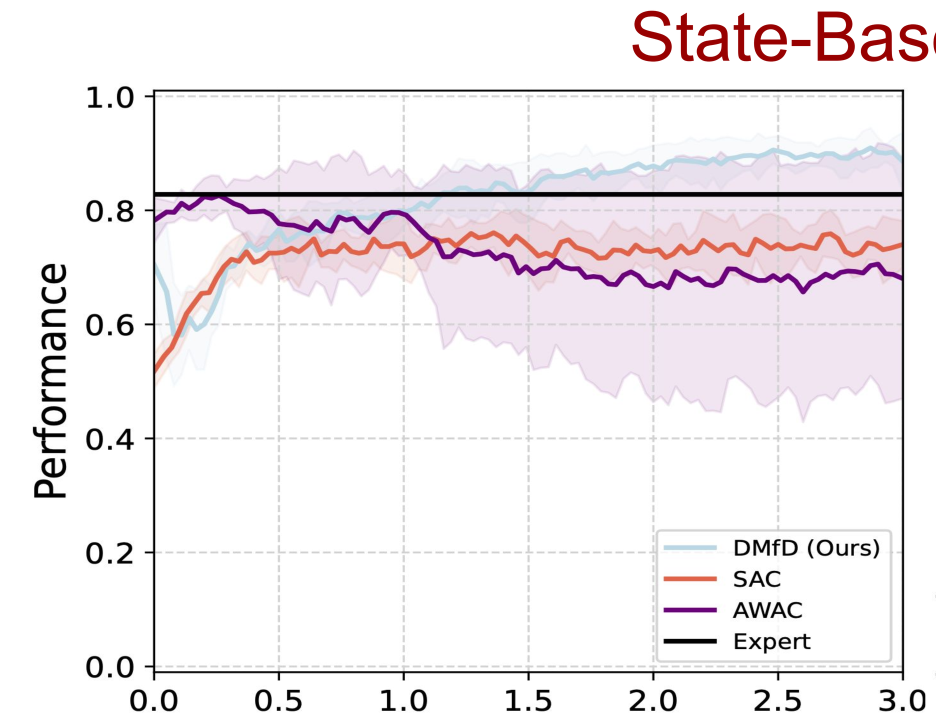


d) Fold Cloth Diagonal (unpinned)

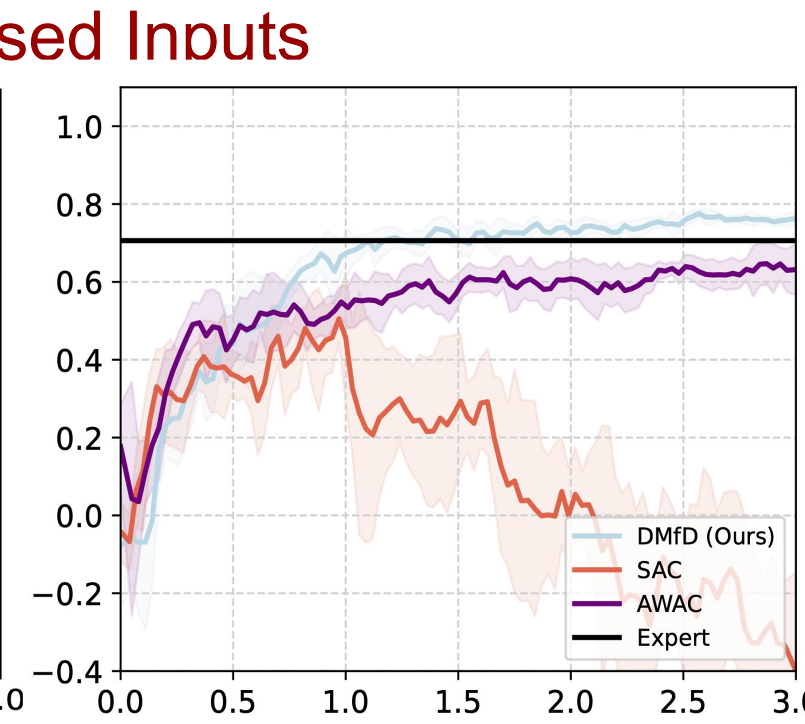
Rollouts of DMfD in solving manipulation tasks from SoftGym, such as straightening 1D ropes and folding 2D cloths, with image inputs. Additionally, we introduce the Cloth Fold Diagonal tasks.

Results

State-Based Inputs

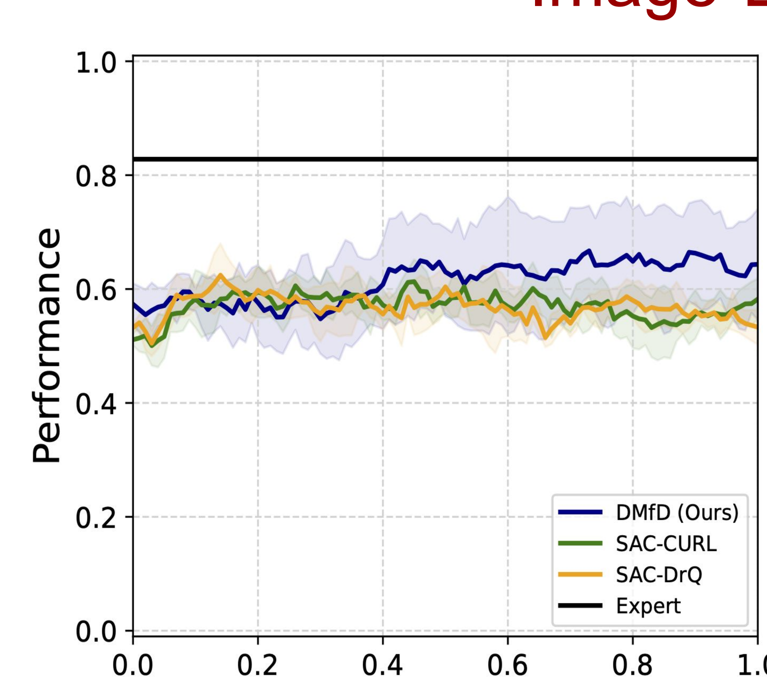


Straighten Rope State

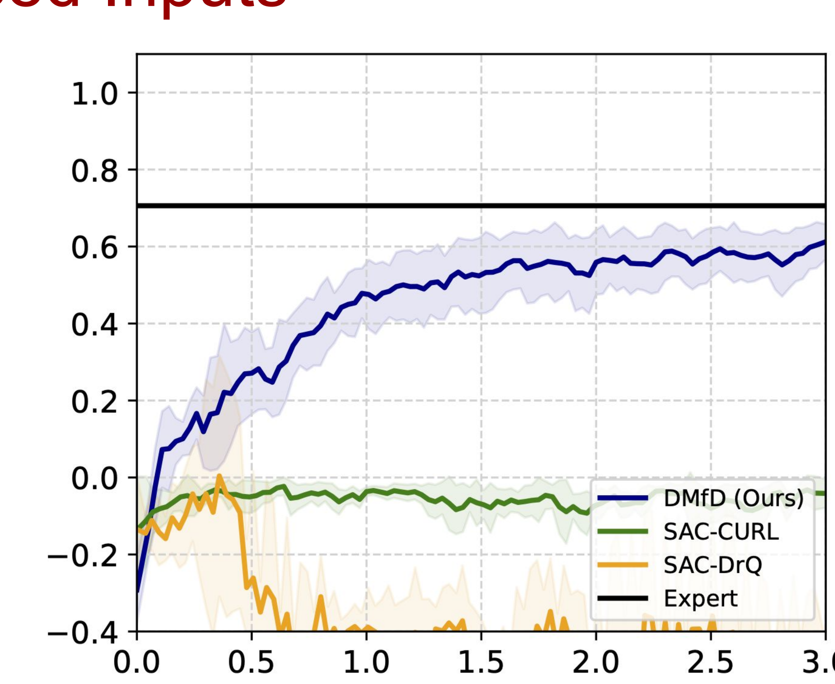


Fold Cloth State

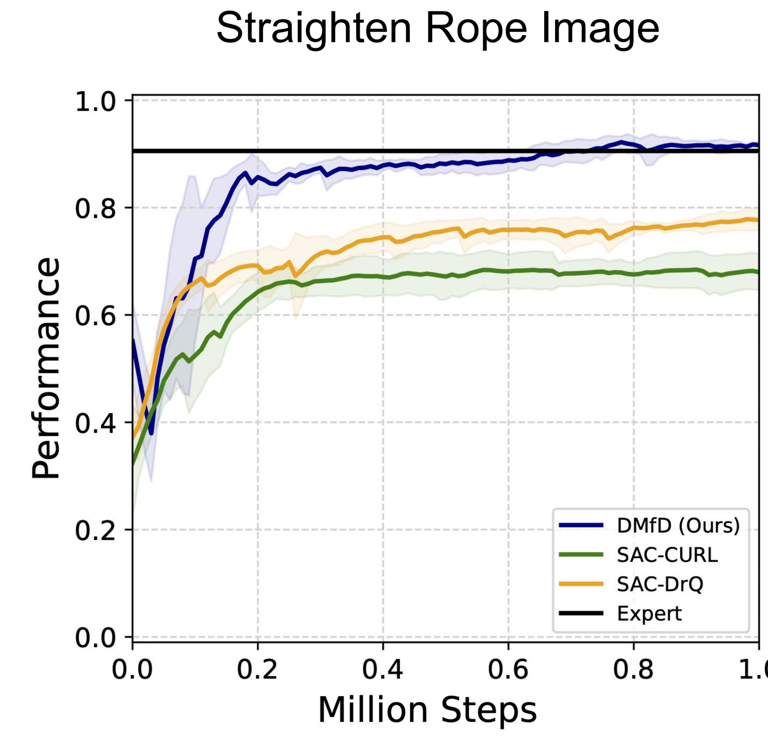
Image-Based Inputs



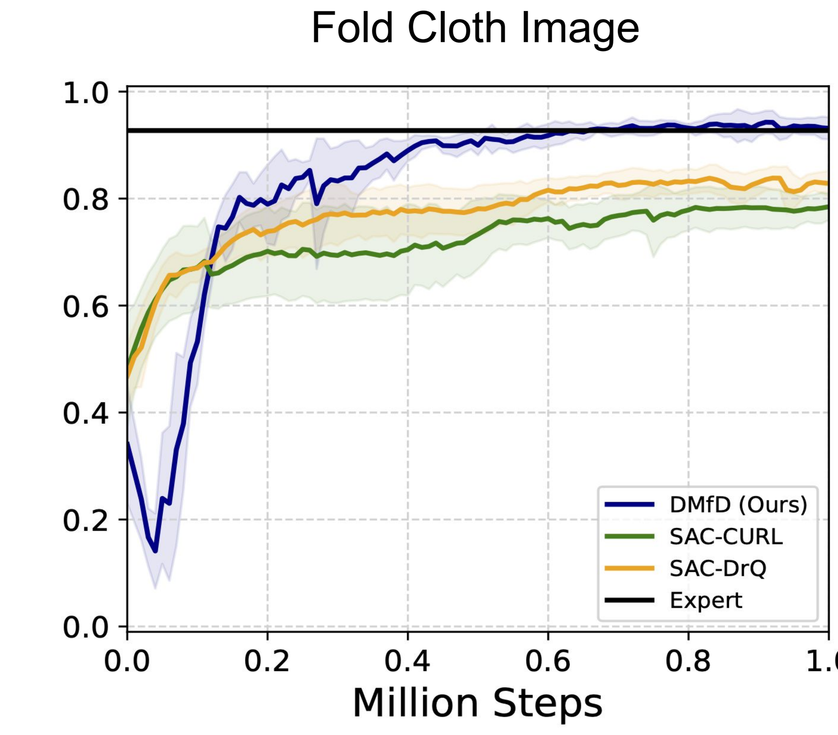
Straighten Rope Image



Fold Cloth Image

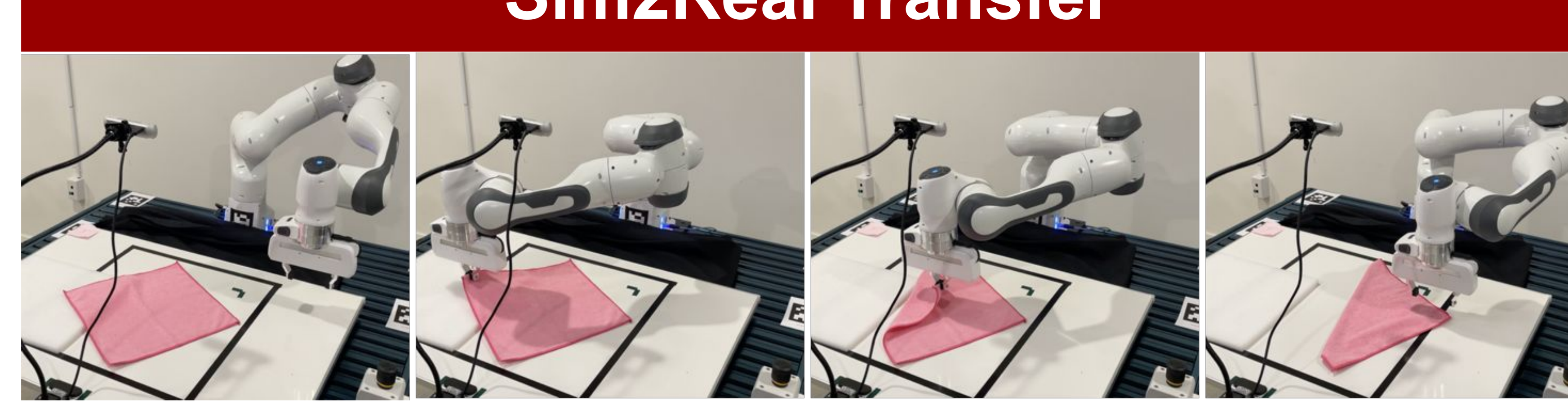


Fold Cloth Diagonal Pinned



Fold Cloth Diagonal Unpinned

Sim2Real Transfer



Sim2Real transfer with minimal (~6%) gap

References

1. Xingyu Lin, Yufei Wang and D. Held, "Softgym: Benchmarking deep reinforcement learning for deformable object manipulation," Conference on Robot Learning (CoRL), 2020.
2. X. B. Peng, A. Kumar, G. Zhang, and S. Levine, "Advantage-weighted regression: Simple and scalable off-policy reinforcement learning," arXiv preprint arXiv:1910.00177, 2019
3. X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills," ACM Transactions on Graphics, vol.37, no.4, pp.1–14, 2018.