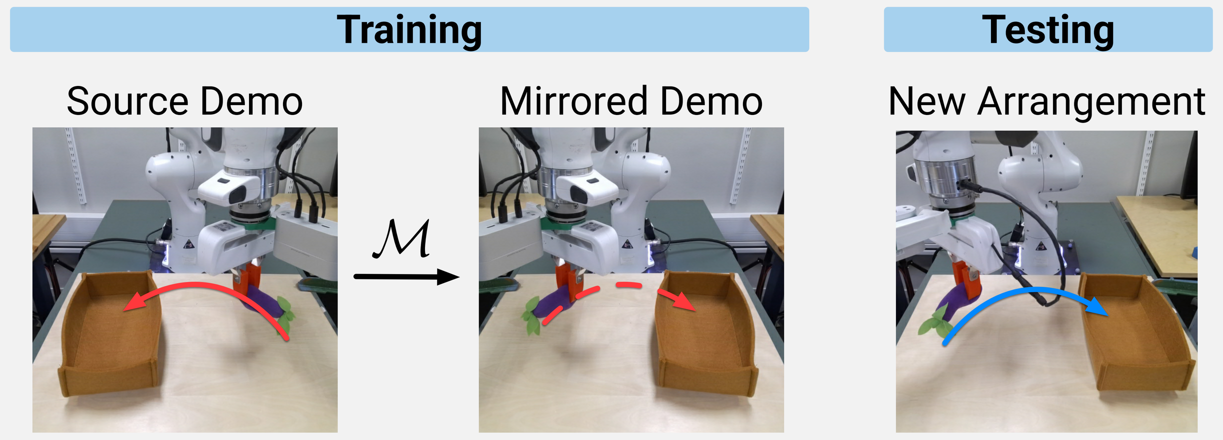
MirrorDuo

Reflection-Consistent Visuomotor Learning from Mirrored Demonstration Pairs

Zheyu Zhuang*, Ruiyu Wang*, Giovanni Luca Marchetti, Florian T. Pokorny, Danica Kragic

*KTH Royal Institute of Technology, Stockholm, Sweden. * Equal Contribution*

Get One Demo, Get One for Free



MirrorDuo captures a simple idea: moving left implies knowing how to move right. We turn this intuition into a framework for image-based imitation learning in the $\mathrm{SE}(3)$ state-action space.

- Workspace Generalization

One-sided demos can be mirrored to cover the opposite side, effectively expanding the workspace.

- Data Efficiency

For **both-sides** demonstrations, MirrorDuo boosts data efficiency, fewer demos are needed to span the same workspace.

How Does MirrorDuo Work?

Base formulation mirrors end-effector poses (and actions) $\mathbf{X}_\mathrm{H}$ in the camera frame $\mathbf{X}_\mathrm{C}$ using a reflection operator $\mathbf{E} = \mathrm{diag}([-1,1,1,1])$. Mapping back to the world frame gives $\mathbf{X}_\mathrm{H}^* = \mathbf{X}_\mathrm{C}\,\mathbf{E}\,\mathbf{X}_\mathrm{C}^{-1}\,\mathbf{X}_\mathrm{H}\,\mathbf{E}$.

Remove Camera Extrinsics

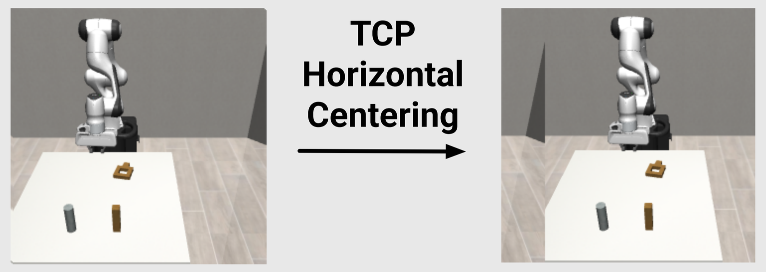
Absolute Pose → Delta Poses

$$\Delta\mathbf{X}_{\mathrm{H}_t} = \mathbf{X}_{\mathrm{H}_0}^{-1}\mathbf{X}_{\mathrm{H}_t}$$

Relative Poses unchanged

$$\delta\mathbf{X}_{\mathrm{H}_t} = \mathbf{X}_{\mathrm{H}_0}^{-1}\mathbf{X}_{\mathrm{H}_t}$$

Center TCP to the Flipping Axis



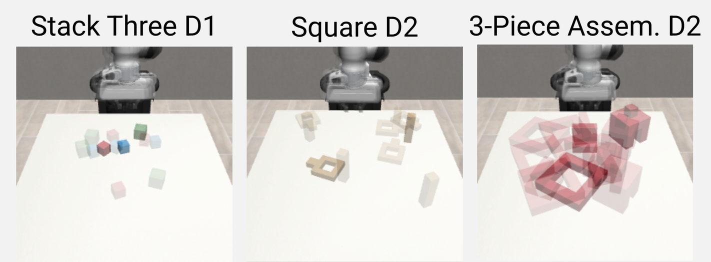
Align Initial Rotation to the World Frame

Ensure the initial rotation remains unchanged under mirroring:

$$\mathbf{R}^* = \mathbf{ERE}.$$

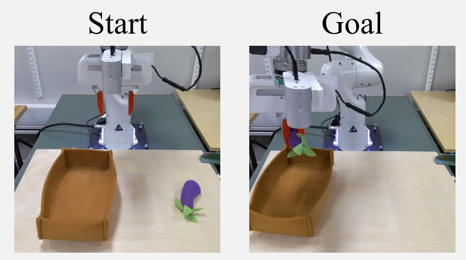
These steps enable the policy to learn trajectories that are immediately deployable in the mirrored setup, starting from the same initial pose.

How Effective Is MirrorDuo?



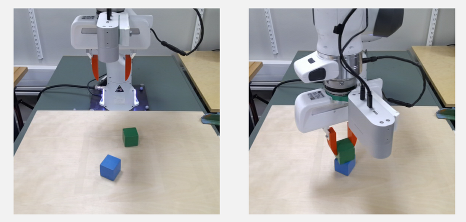
	Method	Low Demos	Medium Demos	High Demos
Delta	MirrorDiff. $+\mathcal{O}$	13.1	13.3	3.2
	DiffPo $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	17.8	15.3	2.2
	BC-RNN $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	0.3	2.7	15.6
Relative	MirrorDiff. $+\mathcal{O}$	13.6	20.7	17.5
	DiffPo. $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	10.6	14.4	9.1
	BC-RNN $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	6.1	9.3	12.0

Table.1 Average **absolute success rate improvement** over three tasks, relative to each baseline, for both absolute and relative actions. $(\mathcal{M},\mathcal{O},\mathcal{P})$ denote MirrorAug, Random Overlay, and Pretrained Encoder.



	In-domain	# M-Demos	
		0	5
MirrorDiff. $+\mathcal{O}$	76.7	0.0	73.3
DiffPo. $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	**86.7**	**20.0**	**83.3**
DiffPo. $+\mathcal{O},\ \mathcal{P}$	83.3	0.0	3.3

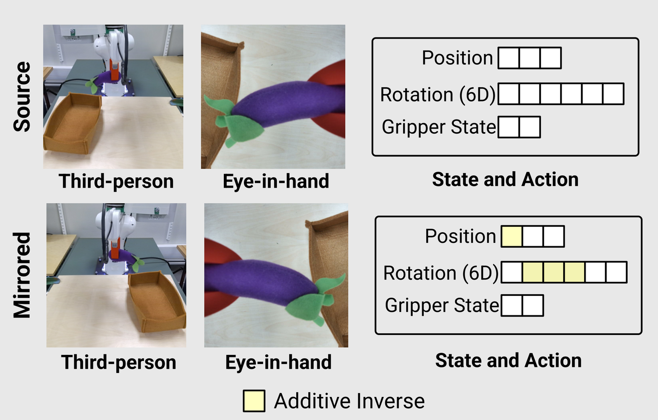
Table.2 Generalization to Mirrored Setup. Success rates (out of 30 trials) in original and mirrored setups. M-demos denote demonstrations in the mirrored setup.



	# Demos	
	200	300
MirrorDiff. $+\mathcal{O}$	43.0	60.0
DiffPo. $+\mathcal{M},\ \mathcal{O},\ \mathcal{P}$	**66.7**	**73.3**
DiffPo. $+\mathcal{O},\ \mathcal{P}$	40.0	53.3

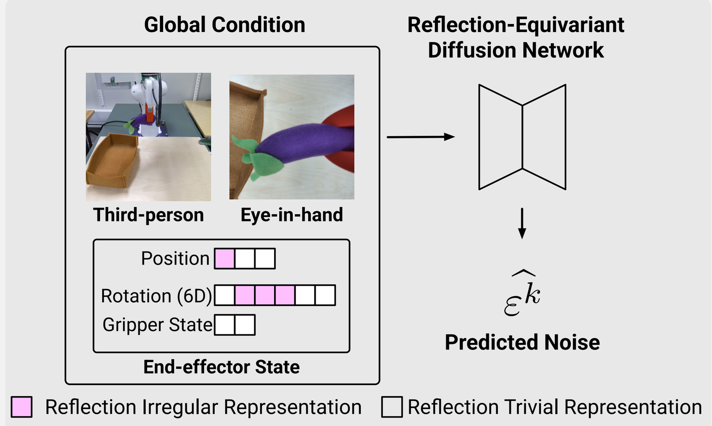
Table.3 In-domain Data Efficiency. Success rates (30-trial average) on block stacking

Dual Realization



Data Augmentation (*MirrorAug*)

Synthesize mirrored samples (Image, Proprioception, Action)



Reflection-Equivariant Diffusion Policy (*MirrorDiffusion*)

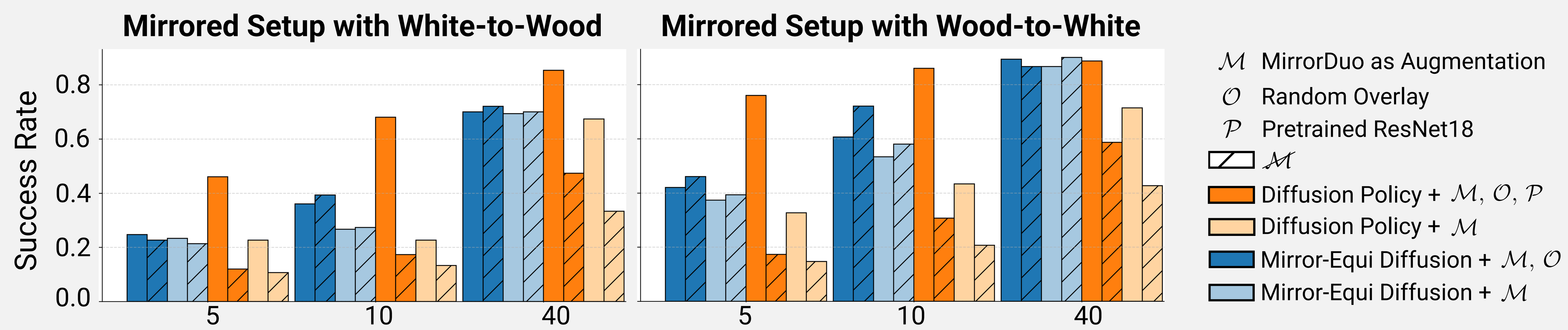
Use reflection-symmetry as inductive bias for Policy

Visual Asymmetry as Distractors

Visual asymmetry handling is closely tied to robustness against backgrounds, distractors, and lighting variations.

Even simple robustness techniques, such as finetuning ImageNet-pretrained encoders or applying Random Overlay, yield significant gains. With enhanced visual generalization, MirrorDuo greatly reduces the real demos needed for transfer to the opposite side.

Mirrored Setup for Evaluation



Number of Demos from Mirrored Setup Added to 200 Demos

Project Page

Personal page