Domain Randomization for sim-to-real Readiness in Imitation Learning

WAT.ai Onboarding Task – Challenge #4

Authors: Mohammad Hasani, Sayon Ramanan

Date: February 4, 2026

Overview

We collected 30 successful demonstrations without domain randomization and 20 with domain randomization (dynamics, lighting, camera variations). Both policies were trained using behavioral cloning for 100 epochs.

Results

Both policies achieved 100% object reach rate and 100% grasp attempt rate on training seeds, indicating successful learning of the approach and grasp motions. However, neither achieved task completion (0% success rate), with all failures classified as "dropped_or_missed."

Metric	No DR (Train)	No DR (Test)	DR (Train)	DR (Test)
Success Rate	0%	0%	0%	0%
Reached Object	100%	100%	100%	95%
Attempted Grasp	100%	100%	100%	95%
Avg Gripper Closed	61.7%	31.2%	61.7%	48.5%

Metric	No DR (Train)	No DR (Test)	DR (Train)	DR (Test)
Avg Total Movement	151.8	172.7	151.8	399.2
Failure Mode	dropped_or_missed	dropped_or_missed	dropped_or_missed	dropped_or_missed

Key Findings

Gripper consistency improved with DR:

- No DR: 61.7% → 31.2% on test seeds (30.5% drop)
- DR: 61.7% → 48.5% on test seeds (13.2% drop)
- **DR reduced generalization gap by 17.3%**

Generalization Analysis

The key difference appeared in gripper consistency on unseen test seeds. The no-DR policy's gripper activity dropped from 61.7% to 31.2% (a 30.5% decrease), while the DR policy only dropped from 61.7% to 48.5% (a 13.2% decrease). This suggests domain randomization helped maintain more consistent gripper behavior when encountering new scenarios.

Failure Analysis

Both policies failed at the placement stage rather than reaching or grasping. This is likely due to observation format inconsistencies. the object name varies between seeds (Cereal, Milk, Can, etc.), causing the policy to confuse object position with bin position during the placement phase.

Conclusion

While full task success was not achieved, domain randomization reduced the generalization gap in gripper behavior by 17.3%. With a corrected observation format that handles varying object names consistently, we expect these policies would show improved task completion rates.

Demo:
https://drive.google.com/file/d/1DTx7WhceKGI64HQ2TuuA03IgqEjvph7v/view?usp=sharing