



Robust Visual Sim-to-Real Transfer for Robotic Manipulation



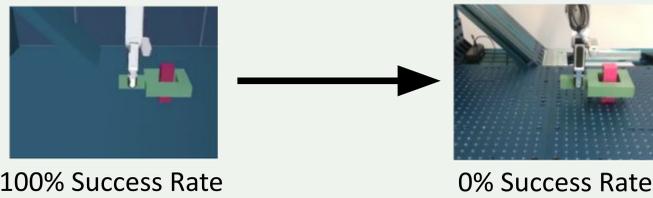
Ricardo Garcia, Robin Strudel, Shizhe Chen, Etienne Arlaud, Ivan Laptev, Cordelia Schmid

Inria Paris, École normale supérieure, PSL



Motivation

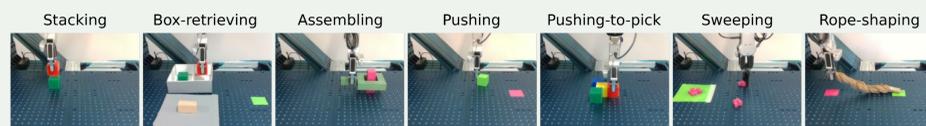
- Large visual sim-to-real gap:



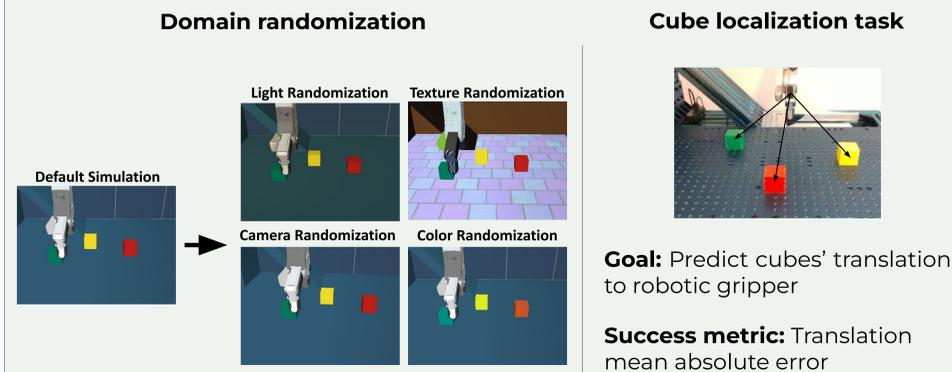
- Domain randomization is a method to close the visual gap, but previous works:

- Focus on disembodied tasks such as pose estimation and object detection.
- What is the correct range of randomization?
- How well the chosen randomization works across a varied set of tasks and scene visual variations?

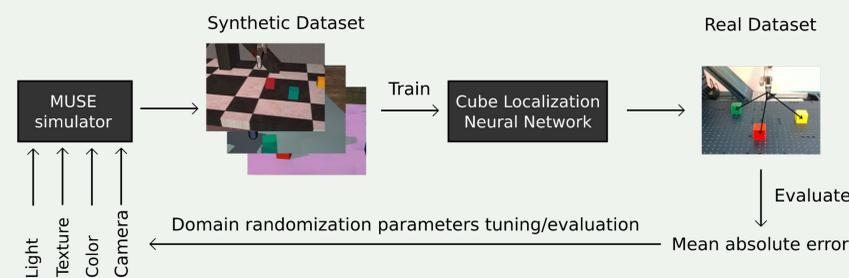
7 challenging robotic manipulation tasks



Method



Systematic analysis of domain randomization



Sim-to-Real Policy Transfer

Policy transferring results correlate with the proxy task results:

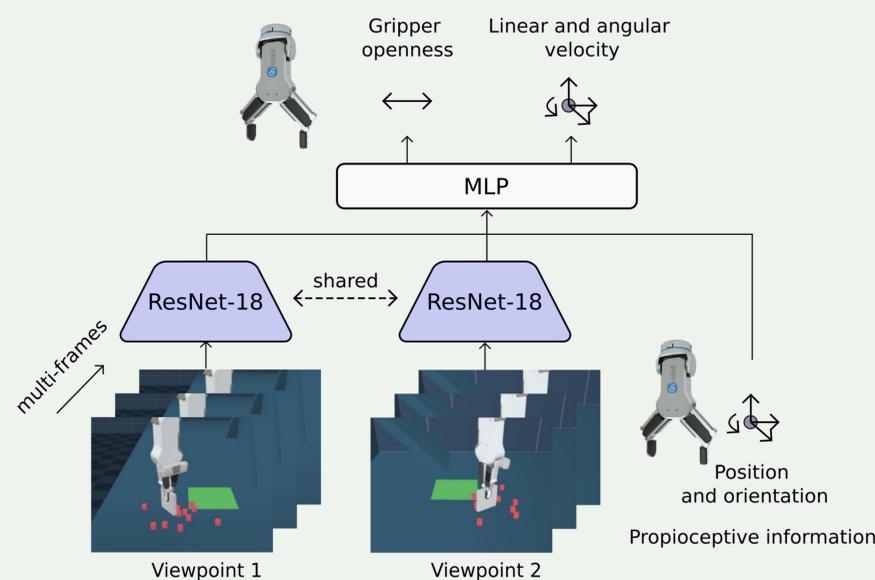
img aug	Texture ACG	Light	Obj color	Camera	Success rate							Average
					Stacking	Box-retrieving	Assembling	Pushing	Pushing-to-pick	Sweeping	Rope-shaping	
✓	×	×	×	×	0/20	0/20	0/20	0/20	0/20	0/20	0/20	0/20
✓	✓	×	×	×	15/20	16/20	7/20	13/20	11/20	9/20	8/20	11.3/20
✓	✓	×	✓	×	17/20	15/20	13/20	18/20	13/20	11/20	11/20	14.0/20
✓	✓	✓	✓	✓	20/20	17/20	16/20	18/20	17/20	19/20	18/20	18.6/20

Robustness to Visual Variations

Transferred policy is robust under default and five variations of the real scene appearance:

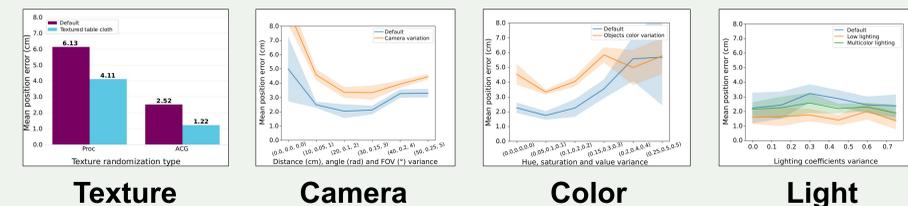
	DR	Real	DR+Real
(a) Default	20/20	20/20	20/20
(b) Textured table cloth	20/20	1/20	20/20
(c) Low lighting	19/20	17/20	20/20
(d) Multicolor lighting	16/20	14/20	17/20
(e) Object colors variation	18/20	19/20	20/20
(f) Camera variation	11/20	8/20	12/20
Average	17.3/20	13.2/20	18.2/20

Visuomotor Policy



Proxy Task Results

Study of each randomization component range separately:



Study of the combination of the chosen randomization ranges:

Training data	Img aug	Texture ACG	Light	Obj color	Camera	Error (cm)
Synt	Real	×	×	×	×	7.55
-	×	✓	×	×	×	6.92
-	×	×	×	×	×	2.52
20k	-	×	✓	✓	×	2.66
-	×	×	✓	✓	×	1.62
-	×	×	✓	✓	✓	1.33
100k	-	✓	✓	✓	✓	0.95
-	✓	✓	✓	✓	✓	0.48
-	750	×	×	×	×	0.72
100k	750	✓	✓	✓	✓	0.14

- From 7.55cm to a final error of **0.95cm** (20k synthetic images) and **0.48cm** (100k synthetic images).

- Texture and object color randomization has the biggest impact.

Conclusions

Systematically explore domain randomization methods to reduce visual sim-to-real gap for robotic manipulation:

- Cube localization is a good proxy task to study and choose domain randomization parameters.
- Chosen parameters show good performance over seven diverse manipulation tasks.
- Transferred policies demonstrate robustness to visual scene variations.

https://di.ens.fr/willow/research/robust_s2r/

