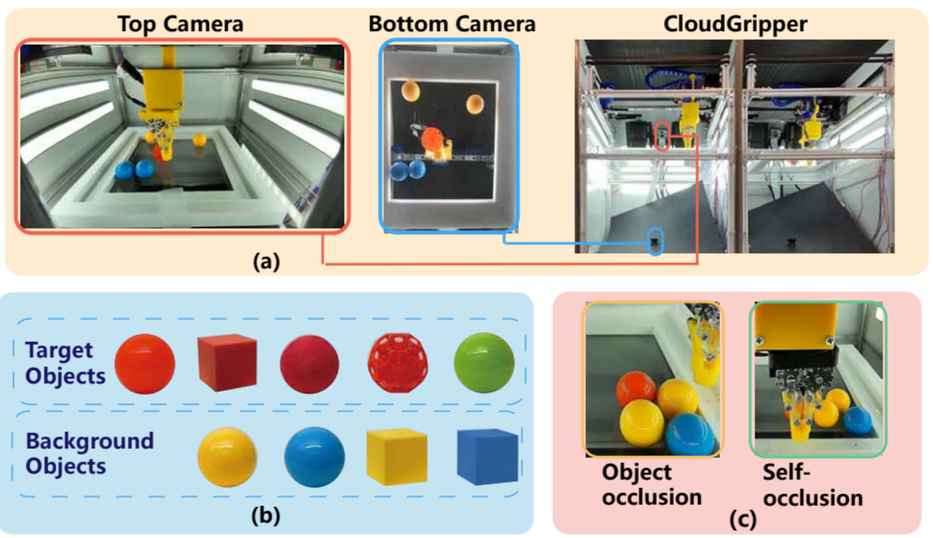
CloudGripper-Push-1K: Understanding the Generalization Gap of Physics and Background Attributes for Robotic Manipulation

Shutong Jin, Ruiyu Wang, Muhammad Zahid and Florian T. Pokorny
KTH Royal Institute of Technology

Motivation:

- As model and dataset sizes continue to scale in robot learning, the need to understand how the composition and properties of a dataset affect model performance becomes increasingly urgent to ensure cost-effective data collection and model performance.

- We introduce *CloudGripper-Push-1K*, a large real-world visual-based robot pushing dataset featuring various target objects (Ball, Cube, Foam, Icosahedron) with distinct physics attributes (color, friction coefficient, shape).



Object	Background	Training	Testing
Ball (Red)	Single	30030	2800
	Double	22860	1881
	Triple	20720	2570
	Quintuple	20000	1696
Cube	Single	21860	1889
	Double	32870	1500
	Triple	27800	2311
	Quintuple	21610	2180
Foam	Single	19450	2163
	Double	20440	2134
	Triple	20838	2100
	Quintuple	21430	2129
Icosahedron	Single	21670	2697
	Double	24980	2379
	Triple	21430	3150
	Quintuple	22070	1661
	Quintuple_Static	25360	3392
Ball (Green)	Single	21540	1660

[1] Zahid, Muhammad, and Florian T. Pokorny. "CloudGripper: An Open Source Cloud Robotics Testbed for Robotic Manipulation Research, Benchmarking and Data Collection at Scale." 2024 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2024.

***CloudGripper-Push-1K*:** visual-based real robot pushing dataset of **1.4TB** and in total **1278 hours** of videos collected with *CloudGripper* [1].

18 sub-datasets consisting various target objects (Ball, Cube, Foam, Icosahedron) with distinct physics attributes (color, friction coefficient, shape).

In each trajectory: 1. A high-resolution top-camera video; 2. An occlusion-free bottom-camera video; 3. trajectory of the target object in pixel space.