IEEE Robotics & Automation Magazine (RAM) paper, presented at ICRA 2026, Vienna, Austria. Cite as RAM paper.

Leveraging Embodied Mechanical Intelligence for Learning Decluttering Tasks

Enrico Turco[1,*], Valerio Bo[2,*], Chiara Castellani[1], Gionata Salvietti[1,3], Monica Malvezzi[3], Domenico Prattichizzo[1,3] and Maria Pozzi[1,3]

***Abstract*—In this work, we investigate how a state-of-the-art grasp planner based on deep reinforcement learning performs when applied to a soft-rigid gripper in a decluttering task. The gripper, called Soft ScoopGripper, is endowed with a rigid scoop-shaped part that facilitates the interaction with the environment and with objects. We hypothesize that the clever design of such a gripper can *facilitate* the learning process, reducing the number of required training steps and eliminating the need for learning non-prehensile actions, such as pushing. To validate our hypothesis, we conducted experiments in both simulated and real-world environments, comparing the selected gripper with a rigid parallel-jaw gripper and a four-fingered soft gripper. Results show that the Soft ScoopGripper learns to effectively declutter scenes using a single action (grasping) instead of two (pushing and grasping). This is due to the fact that the scoop-shaped add-on allows to perform non-prehensile motions *during* the grasp action.**

I. INTRODUCTION

Learning-based approaches have led to important results in robotic grasping and manipulation, above all when it comes to the proficient use of rigid parallel-jaw grippers and suction cups in unstructured environments [1], [2]. Meanwhile, the development of a growing number of grippers with different structures, materials, and capabilities has shown that robot design features can fundamentally help the grasping process [3]. Compliant structures, for example, show superior capabilities in adapting to objects with different shapes, and can safely interact with their surroundings [4]. However, the interplay between learning algorithms and gripper design has rarely been studied since most of the works presenting learning-based approaches adopt parallel-jaw grippers.

This work falls within a wider research line aiming at answering the following question: Can we leverage the *mechanical intelligence*, or - according to [5] - the *mechanical computation*, of a robotic gripper to simplify the learning process in grasping scenarios?

To start investigating this problem, we decided to focus on a specific task, i.e., removing all objects from a cluttered scene (decluttering). In this context, we study the synergy between a state-of-the-art learning-based algorithm which outputs decluttering plans [6], and a robotic gripper that is particularly suitable for the chosen task [7]. Specifically, we used the Soft ScoopGripper (SSG) which differs from a standard parallel gripper as it features passively compliant fingers and a rigid embedded scoop-shaped structure enabling environmental constraints exploitation [8].

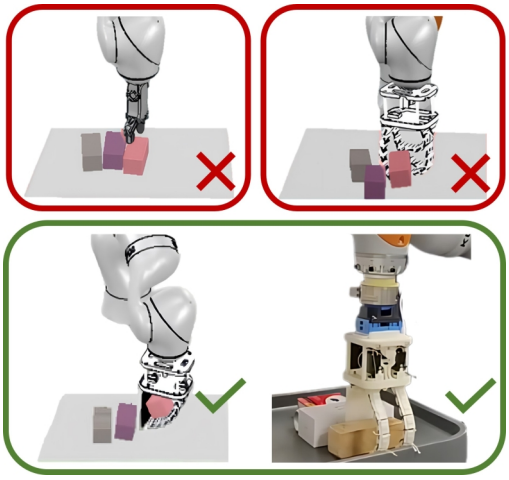


Fig. 1: Grasp attempts in a cluttered environment using three different grippers: rigid parallel gripper, soft gripper, and soft-rigid gripper with an embedded constraint (Soft ScoopGripper). The first two fail to grasp the object, while the third successfully completes the grasp. The Soft ScoopGripper is illustrated in both simulated and real-world scenarios.

This paper demonstrates that the SSG learns decluttering in a more successful and efficient way when adopting a learning policy based on a single action ("grasp-only"), with respect to a policy with two actions ("push-grasp"). This result is obtained both in simulation and with real-world experiments, and constitutes the main contribution of this paper. The simplification of the learning policy is made possible by the SSG *clever* design which allows non-prehensile push actions to be performed *while* grasping, and not as pre-grasp preparatory adjustments as for simpler grippers [6], [9].

A detailed analysis in simulation complements and confirms these results showing how different design parameters influence learning, as well as how other grippers (a rigid parallel-jaw gripper and a four-fingered soft gripper) perform with respect to the SSG (Fig. 1).

The rest of the paper is organized as follows. In Sec. II, we frame our work in the related literature, whereas the employed grippers and the RL framework are detailed in Sec. III. Sec. IV and Sec. V report the results of the training and testing phases conducted in simulated and real scenarios, respectively, along

[1]Istituto Italiano di Tecnologia, Center for Robotics and Intelligent Systems, Humanoids and Human Centered Mechatronics, Genoa, Italy. {name.surname}@iit.it

[2]Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain. {name.surname}@iri.upc.edu

[3]Università degli Studi di Siena, Dip. di Ingegneria dell'Informazione e Scienze Matematiche, Siena, Italy. {name.surname}@unisi.it

* The authors equally contributed to the paper.

This work was supported by the European Union by the Next Generation EU project ECS17 "THE - Tuscany Health Ecosystem" (PNRR MUR M4 C2 Inv. 1.5, CUP B63C22000680007, Spoke 9: Robotics and Automation for Health).

with their corresponding discussions. Lastly, in Sec. VI, we outline conclusions and future developments of the paper.

II. RELATED WORKS

Current research on robotic grasping frequently faces a mismatch between the availability of highly capable prototypes of robotic hands [4], and their actual use in state-of-the-art learning-based methods for perception and planning. Above all when a large amount of data for training is required, simple and commercially available grippers (parallel-jaw, suction) are usually preferred [2].

Focusing on the specific case of grasping in cluttered spaces, while targeted gripper designs have been proposed [10], [11], still the problem has been predominantly addressed from a software perspective, using vision-based planning and learning approaches [6], [9], [12]–[14]. Besides methods that directly generate 6-DoF pick and place poses for objects in clutter [13], other works leverage pushing actions to "funnel large amounts of uncertainty into smaller amounts" before a grasp [12]. This idea of planning sequences of non-prehensile and prehensile actions to move obstacles out of the way, or funnel uncertainty before grasping is at the basis of several deep reinforcement learning approaches which tackle grasping in clutter [6], [9], [14]. In [6], a synergy between pushing motions and grasping actions (referred to as Visual Pushing-Grasping, VPG) is presented, whereas in [9] grasping in clutter is achieved with a combination of manipulation primitives like clamping and shifting, and in [14] bimanual grasping is preceded by non-prehensile decluttering actions to pick objects in densely stacked shelves.

Exploiting non-prehensile actions is indeed a first step towards explicitly using the object-environment interaction to reduce the uncertainty in the manipulation task. However, since the grippers adopted in the above mentioned works are rigid, the physical interaction between the gripper and external objects is anyway minimized and the environment is primarily seen as a source of uncertainty or occlusion to overcome. This paper, instead, is based on the idea that "robust grasping performance is enabled by exploiting constraints present in the environment", as claimed in [15]. This principle is at the basis of soft robotic manipulation [3], and has its roots in the wider research line of embodied intelligence [5], [16].

Thanks to their passive adaptability and robustness to uncertainties, soft grippers promise to reduce the need for high-precision motion and accurate contact modeling and planning. However, most learning-based robotic grasping and manipulation methods have largely ignored the diverse capabilities that a purposefully designed and fabricated gripper morphology can offer [17], assuming the use of parallel-jaw or suction grippers [2], or relying on the hand anthropomorphism [18]. Only a few studies, in particular, consider learning with soft or underactuated hands, but are limited in scope, as they typically tackle top-grasp in simplified uncluttered scenarios [19], [20].

Framing our work according to the perspective outlined in [5], where intelligence is seen as a "computational phenomenon", here we study how mechanical computation coming from the robotic gripper design influences the way the robot learns complex grasping tasks. While seminal work [17] has already shown the prominent role of mechanical computation in in-hand manipulation tasks performed with continuous soft hands, here we narrow the problem down to the specific task of decluttering with a soft-rigid gripper.

III. MATERIALS AND METHODS

A. Grippers

In the following, we describe the three types of employed grippers: rigid, soft, and soft-rigid.

1) Rigid: The rigid gripper we simulated is the RG2 model (Onrobot, DK), a commercial parallel-jaw gripper with rigid fingers, typically employed for standard grasping tasks. This type of gripper is effective in structured environments but less adaptable to objects with different shapes.

2) Soft: The soft gripper we simulated is the Co-Gripper [21] which is specifically designed for safe human-robot interaction. It features an underactuated structure with four fingers, each composed of five phalanges connected by flexible joints. This design allows for compliance and adaptability, making it suitable for handling various objects.

3) Soft-rigid: The soft-rigid gripper that we simulated and used in real experiments is the Soft ScoopGripper (SSG), a non-anthropomorphic underactuated robotic gripper composed of two soft fingers and a flat surface (the "scoop") attached to the palm through a flexible hinge. The SSG is designed to exploit environmental constraints thanks to the interplay of its soft and rigid parts. In particular, the scoop can slide over surfaces (e.g., walls, tables) allowing the fingers to grasp objects that are in contact with them [8], [22]. The tendon-driven fingers have compliant joints and rigid phalanges. A differential system actuated by one motor permits their simultaneous flexion.

B. Problem Formulation and State Representation

As in [6], the problem of decluttering a scene is formulated as a Markov decision process (MDP). At time t, at any state s_t, the agent chooses and executes an action a_t according to a policy $\pi(s_t)$. Once reaching a new state s_{t+1}, the agent receives a corresponding reward $R_{a_t}(s_t, s_{t+1})$. The aim is to find an optimal policy π^* able to maximize the sum of the future rewards over an infinite-horizon, given by $R_t = \sum_{i=t}^{\infty} \gamma R_{a_i}(s_i, s_{i+1})$, where γ is the future discount.

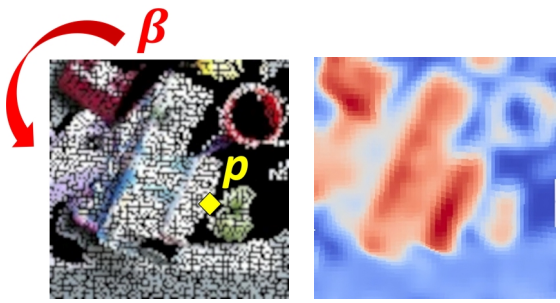
An off-policy Q-learning was adopted to train a greedy policy $\pi(s_t)$. This policy chooses the actions by maximizing the Q-function $Q_\pi(s_t, a_t)$, i.e., the measurement of expected reward for selecting an action a_t in a certain state s_t at time t. Let us define the temporal difference error δ_t between $Q_\pi(s_t, a_t)$ and a fixed target value y_t:

$$\delta_t = |Q(s_t, a_t) - y_t| \tag{1}$$

$$y_t = R_{a_t}(s_t, s_{t+1}) + \gamma \max_{a'}(Q(s_{t+1}, a')) \tag{2}$$

where a' is the set of all available actions. The training phase is based on iteratively minimizing δ_t.

We modeled each state s_t as a heightmap of the cluttered scene at time t. RGB-D images which come from a fixed-mount camera were firstly projected onto a 3D point cloud,



(a) (b)

Fig. 2: Examples of registered heightmaps. (a) Color heightmap that contains the RGB channels and is fed to the network. The best pixel p is in yellow and the angle to rotate the image β is in red. (b) Depth heightmap.

and then re-projected to give rise to two heightmap images represented in both channels (RGB and depth). We determined the size of the heightmap taking into account the robot reachability to execute the actions. We constrained the workspace to an area of $0.3 \times 0.3\ m^2$, leading to heightmaps of 100×100 pixels. Hence, each pixel spatially represents a 3D space of $3 \times 3\ mm^2$ in the agent's workspace.

C. Action Definition

Starting from the state representation s_t (i.e., the heightmap of the scene), we defined the actions as motion primitives μ. We back-projected a given pixel p selected by the network from the heightmap into a 3D location q, where the primitives are then executed. An example of a registered heightmap is shown in Fig. 2.

We defined two actions, i.e., pushing and grasping. In both actions the robot positions the gripper at the 3D location q with an orientation angle β determined by the network (Fig. 2a). The angle β expresses a rotation of the heightmap and is discretized in intervals of $22.5°$, resulting in 16 different orientations. For every gripper, the pushing action is performed by firstly closing the fingers to avoid hindrance during the motion. Then, the gripper moves above the target, and it pushes following a straight trajectory. To perform a grasp, the gripper is first positioned above the grasping location. Hence, it starts descending following a straight trajectory toward the object. Once the target is reached, the fingers close around the object. This grasping procedure is the same for all three grippers.

D. Learning Algorithm

To train the policy $\pi(s_t)$, we defined a fully convolutional network (FCN) for each action (ϕ_g for grasping and ϕ_p for pushing). Then, we modeled the Q-function using the two networks. Both FCNs take as input the heightmap images representation of the state s_t and generate a pixel-dense map of Q values with the same size and image resolution of the input image. Each Q value prediction at a pixel p represents the expected future reward of executing primitive μ at the 3D position q, where $p \in s_t$.

Each FCN consists of two 121-layer DenseNet [23] pre-trained on ImageNet [24], followed by channel-wise concatenation and 2 additional 1×1 convolutional layers interleaved with nonlinear activation functions (ReLU) and spatial batch normalization. The first DenseNet takes as input the RGB heightmap, whereas the other one takes the normalized depth heightmap. For each FCN, the input heightmaps are rotated in 16 different orientations, resulting in pixel-wise maps Q values as output (16 for pushing and 16 for grasping). The executed action a'_t is the primitive μ which maximizes the Q-function at a pixel p corresponding to the highest Q value among all pixel-wise maps. For what concerns rewards, a reward $R_g(s_t, s_{t+1}) = 1$ is assigned if a grasp is successful. To evaluate the success or failure of a grasp, we kept track of the torque value at the fingers' motor. If a certain threshold was overcome, the grasp was deemed successful, as it meant that the object was firmly held until the end of the grasping action. In pushing, we assigned a reward $R_p(s_t, s_{t+1}) = 0.5$ if the algorithm detected significant differences between the states at time instants t and $t+1$, making a comparison between two consecutive heightmaps.

The FCNs were trained using the Huber loss function

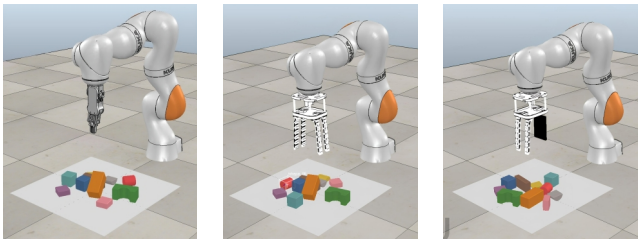
$$\mathcal{L}_i = \begin{cases} \frac{1}{2}(Q^{\theta_i}(s_i, a_i) - y_i^{\theta_i^-})^2, \text{ for } |Q^{\theta_i}(s_i, a_i) - y_i^{\theta_i^-}| < 1 \\ |Q^{\theta_i}(s_i, a_i) - y_i^{\theta_i^-}| - \frac{1}{2}, \text{ otherwise} \end{cases}, \tag{3}$$

where θ_i are the parameters of the neural network at iteration i, and the target network parameters θ_i^- are held fixed between individual updates. Only if the pixel p and the network ϕ are responsible for the executed action a_i the gradients pass through them. Otherwise, the pixels backpropagate with 0 loss. We adopted stochastic gradient descent with momentum to train the FCNs, with fixed learning rates of 10^{-4}, momentum of 0.9, and weight decay 2^{-5}. We used stochastic rank-based prioritization, approximated with a power-law distribution, to exploit prioritized experience replay. We adopted an ϵ-greedy exploration strategy, where ϵ is initialized at 0.5 and annealed to 0.1, while the future discount γ is constant at 0.5.

The models are trained in PyTorch with an NVIDIA GeForce RTX 3060 on an Intel CORE i7-11800H clocked at 2.30 GHz. The FCNs were trained with data gathered with simulated and real experimental setups. For both, we executed several experimental trials to first train and then test two policies, one using only the grasping action (grasp-only) and one with both grasping and pushing actions (push-grasp).

E. Evaluation Metrics

The performance of the policies in the testing phase both in simulation and in the real-world setup was evaluated with 3 metrics: completion rate (C.R.), action efficiency (A.E.), and grasp success rate (G.S.). The first one is computed as the ratio between the number of objects grasped before the end of the trial and the total number of objects. The second one is the number of grasped test objects over the number of actions performed until the end of the trial. The third one consists of the number of successful grasps over the number of performed grasp actions. Thus, in grasp-only policies, action efficiency and grasp success rate are equivalent.



(a) Rigid gripper (b) Soft gripper (c) Soft-rigid gripper

Fig. 3: Simulated training scenes with the three grippers. During scene initialization, the objects are placed randomly.

IV. SIMULATIONS

In simulation, we compared the results obtained using the rigid, the soft, and the soft-rigid gripper, for both training and testing phases. We also performed additional simulation experiments changing different design parameters of the SSG to explore the impact of each parameter on the algorithm's overall effectiveness. To collect training data and then to test the trained policies, we used a camera resolution of 960×540, we employed an LBR iiwa 7-DOF robot arm (KUKA AG), as that is the arm used in the real experiments, and we adopted the robot simulation platform CoppeliaSim [25] with MuJoCo as physical engine, to simulate soft-body point contacts more accurately. The adopted objects are the same used for training and testing in simulation in [6].

A. Results

1) Training: The training phase for the simulation consisted of decluttering scenes with 10 objects, whose shapes and colors were randomly chosen during experiments. Examples of simulated training scenes are shown in Fig. 3. For each gripper, we trained the two policies (grasp-only and push-grasp) five times. The training curves for the two adopted strategies are presented with different colors in Fig. 4, where the solid lines represent the average success rates across the five runs, and the shaded areas indicate the standard deviations. The curves represent the grasp success rate across 2500 training steps. To quantitatively assess learning performance, we computed two widely used metrics in reinforcement learning [26]: (*i*) the area under the curve (AUC), defined as the integral of the learning curve over time, and (*ii*) sample efficiency, quantified using the time-to-threshold metric, i.e., the number of training steps required to reach predefined success rate thresholds (50%, 60%, 70%, 80%). Results are summarized in Table I.

2) Testing: To test the trained networks, we adopted the test scenes shown in Fig. 5, where there are objects of different shapes placed in challenging arrangements. For the three grippers, we evaluated all the 4 scenes 10 times with the two networks, reaching a total of $3 \times 4 \times 10 \times 2 = 240$ simulated test trials. The obtained results with respect to the metrics introduced in Sec. III-E are shown in Fig. 6. In addition, the cumulative number of objects removed as a function of the number of executed actions is reported in Fig. 7 for each gripper-policy pair, across all test scenes.

B. Discussion of the training results

Analyzing the learning curves from the simulation with the three grippers (Fig. 4), it is evident that introducing soft and soft-rigid grippers in cluttered scenarios significantly influences the learning process compared to the rigid gripper.

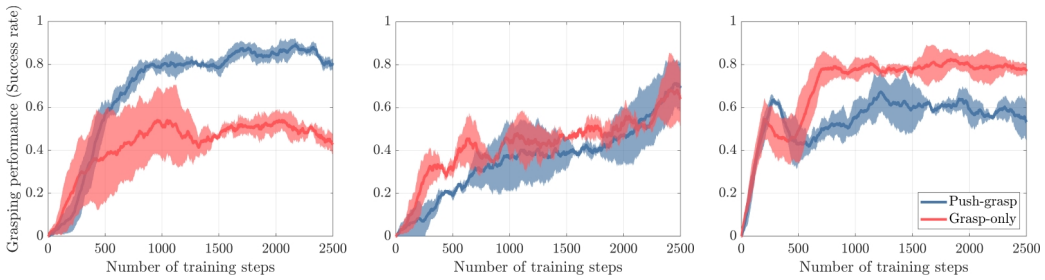
As already demonstrated in [6], when employing the rigid gripper (Fig. 4a), the push-grasp strategy shows a steady performance improvement, with the success rate increasing rapidly during the initial training phase. This suggests that the rigid gripper benefits significantly from the possibility to push objects, allowing it to reposition items into more favorable orientations for grasping. The success rate eventually stabilizes around 80%. This result is confirmed quantitatively by the AUC score of 1720.6 ± 106.5, and by the sample efficiency, reaching 50% grasp success in 442 ± 58 training steps on average and steadily progressing to 80% from step 996 ± 283. Moreover, the learning curves exhibit low variance across training runs, suggesting stable convergence. The high success rate can be attributed to the fact that pushing helps resolve more complex scenarios where objects are tightly packed or partially obstructed, making them easier to grasp. Conversely, the grasp-only strategy for RG2 is significantly less effective. Without the ability to reposition objects, the gripper struggles in cluttered scenarios, particularly when objects are not optimally oriented. This is also reflected in the substantially lower AUC (1070.3 ± 244.3) and a higher variance in early training (before step 1500), highlighting instability and sensitivity to initial conditions when push actions are unavailable. After a late 60% threshold crossing at 1599 ± 1067 steps and no further meaningful improvements, the final success rate plateaus at a much lower level (around 43%).

Concerning the soft gripper (Fig. 4b), learning is consistently slower than with the other grippers, and performance is notably more variable. Quantitatively, this is also reflected in the low AUC values: 894.4 ± 219.4 for push-grasp and 1060.1 ± 151.2 for grasp-only. The push-grasp strategy exhibits a slower start in terms of success rate improvement than the grasp-only strategy. However, after the initial learning phase, the performance steadily improves and eventually surpasses the grasp-only strategy. On average, it requires 1411 ± 663 steps to reach 50% success and 2319 ± 165 steps to reach 70%. Notably, only one out of five training runs managed to achieve 80% success in 2363 steps. This indicates that the push-grasp strategy benefits the soft gripper by allowing it to manipulate and reposition objects before attempting to grasp them, leading to more effective handling of complex and cluttered environments. The grasp-only strategy, instead, demonstrates a faster initial improvement in performance compared to the push-grasp strategy. The success rate climbs rapidly during the early phases of training, but then it slowly achieves 60% after an average of 2151 ± 206 steps. As training progresses, the performance fluctuates and does not show a clear trend of consistent improvement. Eventually, the success rate stabilizes around 65%, indicating that the soft gripper is somewhat effective at grasping directly but struggles with consistent success, particularly in more complex scenarios. The compliance of the soft gripper may allow it to initially

TABLE I: Area under the curve (AUC) and sample efficiency, computed as time-to-threshold at different thresholds for all grippers and strategies. **Bold** values highlight the best result per row.

Metric		RG2		Co-Gripper		SSG	
		Push-grasp	**Grasp-only**	**Push-grasp**	**Grasp-only**	**Push-grasp**	**Grasp-only**
AUC		1720.6 ± 106.5	1070.3 ± 244.3	894.4 ± 219.4	1060.1 ± 151.2	1354.9 ± 151.8	$\mathbf{1741.3 \pm 149.9}$
Time-to-threshold	50%	442 ± 58	1202 ± 766	1411 ± 663	1318 ± 751	201 ± 40	$\mathbf{168 \pm 27}$
	60%	527 ± 85	1599 ± 1067	2210 ± 334	2151 ± 206	$\mathbf{250 \pm 34}$	289 ± 126
	70%	702 ± 102	1073^*	$2319 \pm 165^*$	2301^*	1173 ± 48	$\mathbf{631 \pm 78}$
	80%	996 ± 283	-	2363^*	2368^*	–	$\mathbf{885 \pm 346}$

* Threshold not reached in all runs.



(a) Rigid gripper (b) Soft gripper (c) Soft-rigid gripper

Fig. 4: Learning curves of the simulated training phase considering only the grasping action (red) or both pushing and grasping actions (blue). The solid lines indicate average results across 5 trials and the shaded areas are the standard deviations.

(a) Test #1 (b) Test #2 (c) Test #3 (d) Test #4

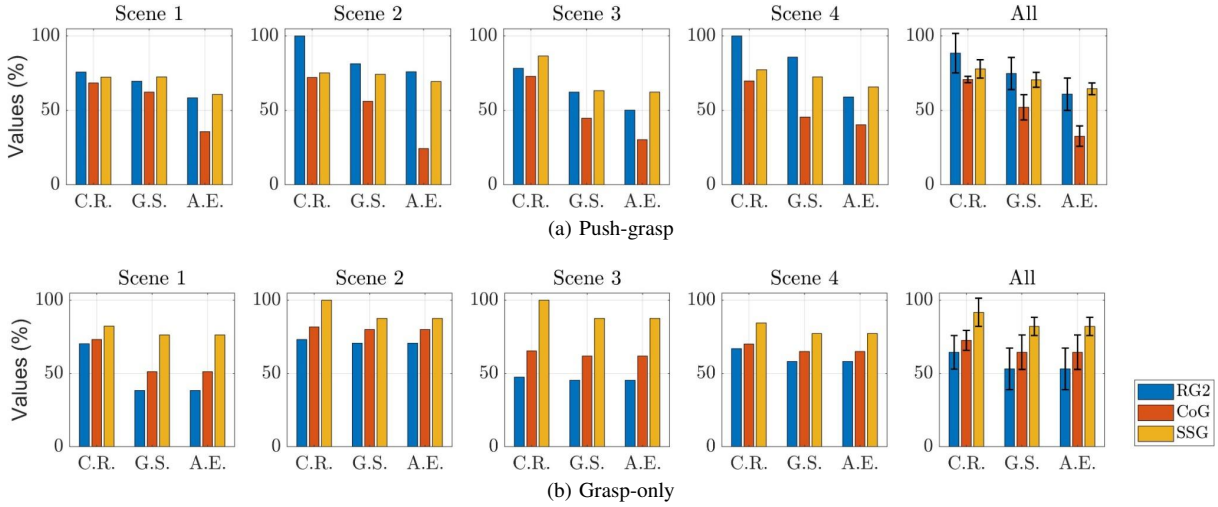
Fig. 5: Cluttered scenes adopted during the simulated test.

adapt well to various object shapes, but without the ability to reposition objects, it may encounter difficulties when objects are not optimally positioned for grasping.

When using the soft-rigid gripper (Fig. 4c), the push-grasp strategy exhibits a limited improvement in success rate compared to the grasp-only strategy. Although a 60% success rate is achieved in the first 250 ± 34 training steps, the performance plateaus soon after with minimal further gains. The final success rate stabilizes at a lower level, around 55%, resulting in a lower overall area under the curve of 1354.9 ± 151.8. Unlike the rigid and soft grippers, the soft-rigid gripper does not seem to benefit as much from the inclusion of pushing actions. The grasp-only strategy shows a strong initial performance, with the success rate rapidly increasing to 50% after only 168 ± 27 training steps. The success rate continues to improve steadily, eventually reaching and stabilizing around 80% at step 885 ± 346. The corresponding AUC (1741.3 ± 149.9) is the highest among all gripper–policy combinations. Furthermore, the learning curve exhibits low variance across training seeds, indicating stable and reproducible convergence. This suggests that the soft-rigid gripper is highly effective at directly grasping objects in cluttered environments without the need for additional manipulation actions such as pushing. Thanks to the scoop-shaped part, the SSG can separate objects without the need of a preliminary pushing action (see Fig. 1), leading to a high success rate within a small number of training steps. Unlike the other two grippers, for the SSG, a pushing effect is inherently embedded in the grasp execution as the scoop separates the target object from the neighbouring ones. This phase is not treated as a separate push primitive in our framework, as it is neither independently triggered by the policy nor separated in the control pipeline. Instead, it arises as a product of the SSG design, which allows the gripper to exploit environmental constraints during grasping.

C. *Discussion of the testing results*

The results from the simulated tests, shown in Fig. 6, provide a comprehensive comparison of the performance of three grippers across the two distinct policies. The findings highlight the strengths and weaknesses of each gripper in handling objects in cluttered environments and offer insights into the effectiveness of different manipulation strategies. The RG2 demonstrated a significant advantage when using the push-grasp policy (Fig. 6a), particularly due to its ability to reposition objects before attempting to grasp them. This capability resulted in a high overall completion rate of $88.5 \pm 13.3\%$ and a grasp success rate of $74.8 \pm 10.8\%$. However, the action efficiency of the RG2, although generally high, varied across different scenes ($60.9 \pm 10.9\%$), indicating that the gripper may require more corrective actions or additional attempts in



(a) Push-grasp

(b) Grasp-only

Fig. 6: Results of the simulated testing phase for: a) Push-grasp and b) Grasp-only, across the 4 different scenes. Each bar represents the performance in terms of Completion Rate (C.R.), Grasp Success (G.S.), and Action Efficiency (A.E.), expressed as percentages. The final plot in each figure presents the overall average performance across all scenes, including standard deviation bars, providing a summary of each gripper's effectiveness.

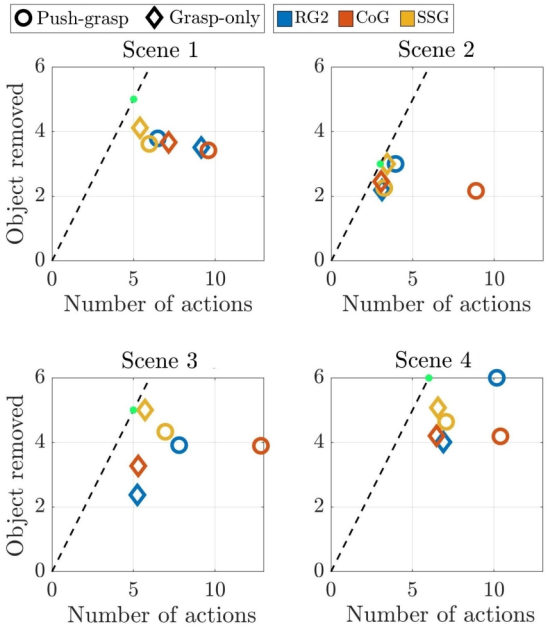


Fig. 7: Comparison between the number of actions performed and the number of objects removed across four test scenes for each gripper and policy. Each marker represents the outcome of a trained model: circles denote push-grasp policies, diamonds denote grasp-only policies, while gripper types are color-coded. The dashed diagonal line represents ideal efficiency, whereas the green dot denotes the ideal case. Points closer to the diagonal indicate higher action efficiency.

scenarios where the initial push does not optimally reposition the objects.

The Co-Gripper consistently underperformed compared to the other grippers, displaying the lowest average metrics. Its completion rate was at $70.8 \pm 2.1\%$, while its grasp success rate was $52.1 \pm 8.5\%$, and action efficiency was notably low at $32.7 \pm 6.8\%$. The compliant nature of the soft gripper may have contributed to difficulties in executing precise pushes, which are critical for repositioning objects effectively in cluttered scenarios. This led to its overall lower performance in both the completion rate and grasp success.

The SSG exhibited a more balanced performance across all three metrics. Notably, its action efficiency often matched or surpassed that of RG2 ($64.5 \pm 3.9\%$), indicating that the SSG could maintain high operational effectiveness without requiring excessive adjustments. Its completion rate ($77.8 \pm 6.1\%$) was on average slightly lower than that of the RG2, but with a smaller standard deviation, suggesting more reliable performance across different environments. The SSG grasp success rate ($70.6 \pm 4.9\%$) was also consistently high, reflecting its adaptability and precision in handling various objects. However, the SSG faced challenges in executing effective push actions, likely due to uneven contact or misalignment during the push, leading to fluctuations in performance. In certain cases, the scoop's design may limit the gripper's ability to exert optimal pressure along the object's surface during non-prehensile actions like pushing.

With the grasp-only policy (Fig. 6b), the performance among the three grippers shifted significantly. The RG2, which excelled with the push-grasp strategy, exhibited the weakest performance compared to the other grippers. Without the ability to reposition objects, the RG2 struggled to maintain high success rates, particularly in scenes featuring more complex object arrangements. Its overall completion rate dropped to $64.5 \pm 11.5\%$, and its grasp success rate fell to $53.2 \pm 14.2\%$, highlighting the strong dependence on push actions to effectively manage cluttered environments.

Conversely, the Co-Gripper showed an improvement under the grasp-only policy compared to its results with the push-grasp strategy. The soft gripper achieved a completion rate of $72.6 \pm 6.8\%$ and a grasp success rate of $64.5 \pm 11.8\%$. This improvement suggests that the gripper compliance is better

leveraged when it is not required to execute push actions. However, the Co-Gripper still faced challenges with more complex object arrangements (e.g., Test Scene #1).

The SSG obtained the best performance under the grasp-only policy, demonstrating versatility and effectiveness. A completion rate of $91.7 \pm 9.6\%$ and a grasp success rate of $82.2 \pm 6.1\%$ were achieved, and were the highest among the three grippers. The combination of soft fingers and rigid scoop allowed the SSG to have high success rates across all scenes. As observed in the training phase, thanks to its scoop, the SSG can separate objects while grasping them, and thus there is no more the need of performing the pushing action.

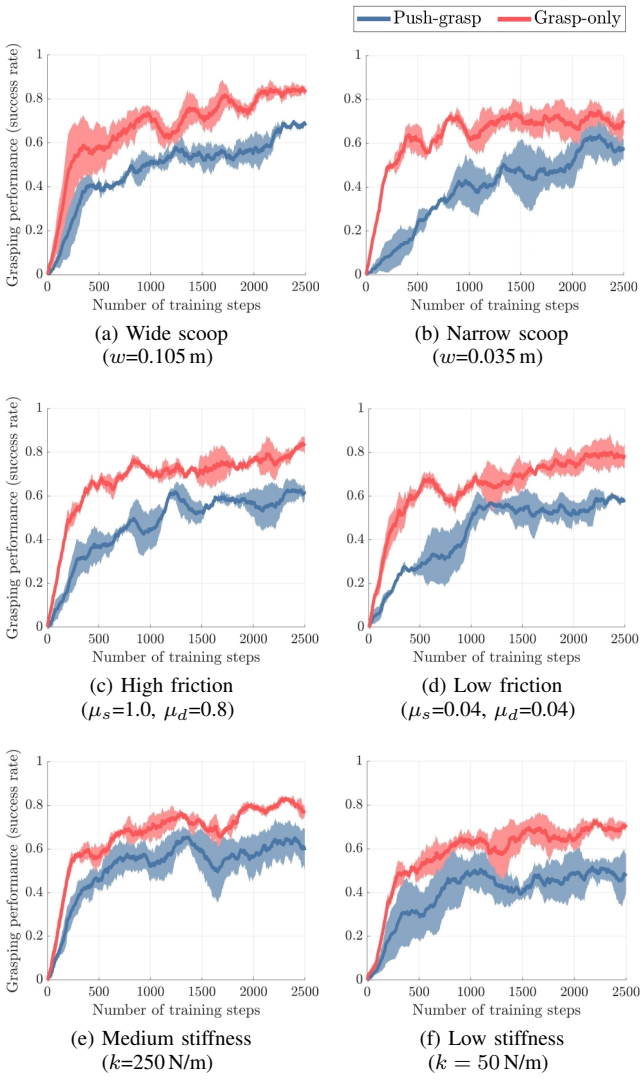
In summary, if we look at the rightmost column in Fig. 6, the results of the simulations reveal that the soft-rigid gripper outperforms the other two grippers in the grasp-only strategy, whereas the rigid gripper is the best performing in the push-grasp. The soft gripper achieves better results in the grasp-only strategy with respect to the push-grasp, but does not excel in either of the two strategies. These observations are further supported by Fig. 7, which compares the number of actions performed versus the number of objects removed for each gripper-policy combination across the four scenes. The dashed diagonal line represents the ideal scenario in which each action results in the successful removal of an object. The SSG mark, particularly with the grasp-only policy, consistently appears closest to the diagonal across scenes, indicating high action efficiency. In contrast, RG2 and especially Co-Gripper often fall farther from the diagonal, showing that more actions were required to remove fewer objects (e.g., the Co-Gripper under push-grasp policy in Scene 2).

D. Analysis of the SSG design parameters

To explore the impact of design parameters on learning and task performance of the SSG, we performed additional simulation analyses. Specifically, we considered three features of the SSG scoop: width w, surface friction coefficients μ_s and μ_d, and hinge joint stiffness k. Each parameter was varied independently, with the others kept constant at their default values that are: $w = 0.07$ m, $\mu_s = 0.5$, $\mu_d = 0.2$, $k = 1000$ N/m.

The conducted study considers six conditions: wide ($w = 0.105$ m) vs. narrow ($w = 0.035$ m) scoop, high ($\mu_s = 1$, $\mu_d = 0.8$) vs. low ($\mu_s = 0.04$, $\mu_d = 0.04$) friction, medium ($k = 250$ N/m) vs. low ($k = 50$ N/m) stiffness. Each configuration was trained under both grasp-only and push-grasp policies, and was repeated three times with different seeds to account for stochasticity in the learning process, collecting a total of $6 \times 2 \times 3 = 36$ simulation runs. Learning performance was quantitatively evaluated using the area under the curve and time-to-threshold metrics defined in Sec. IV-A.

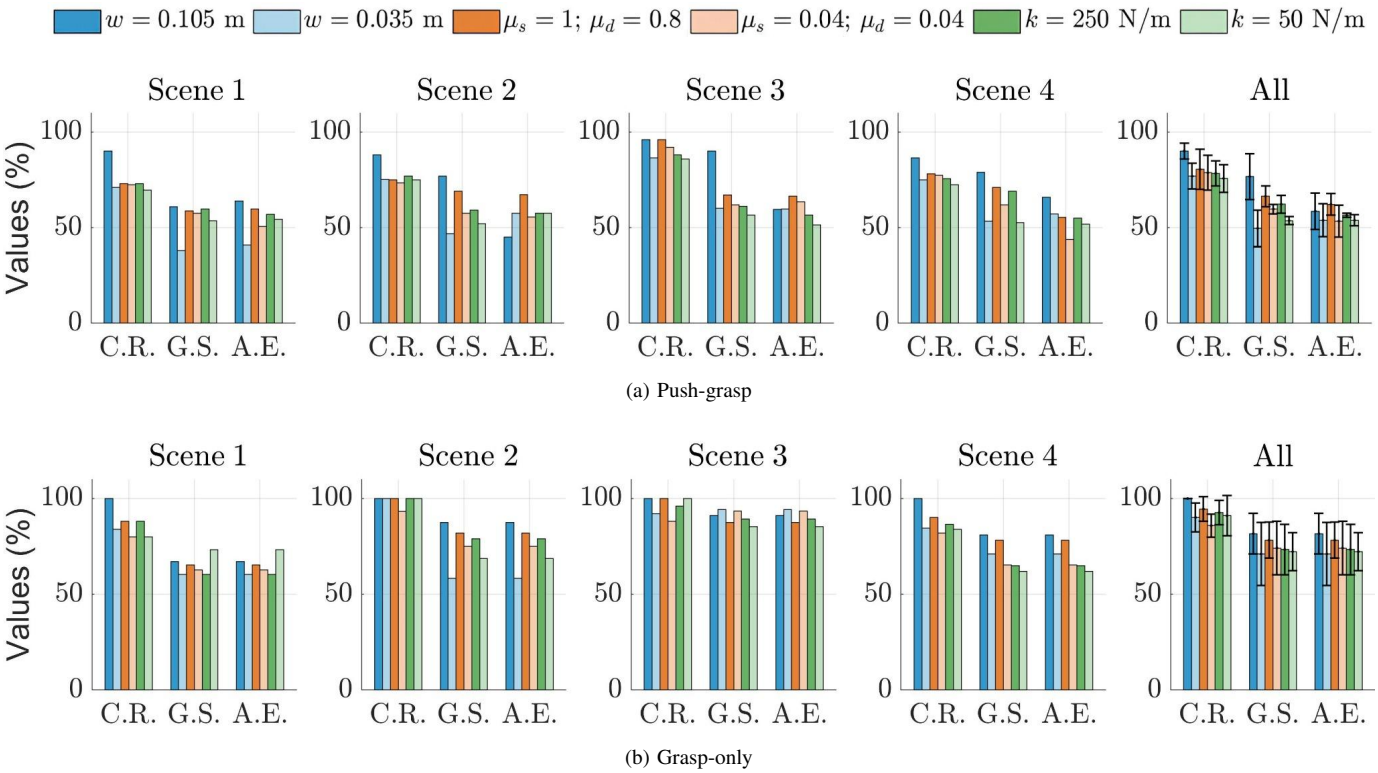
Each trained model was then tested on the four test scenes in Fig. 5. For each of the six conditions and each policy, we conducted 10 trials per scene, reaching a total of $6 \times 10 \times 4 \times 2 = 480$ simulated test episodes. Task performance was evaluated using the completion rate, grasp success, and action efficiency metrics defined in Sec. III-E.



(a) Wide scoop (w=0.105 m)
(b) Narrow scoop (w=0.035 m)
(c) High friction (μ_s=1.0, μ_d=0.8)
(d) Low friction (μ_s=0.04, μ_d=0.04)
(e) Medium stiffness (k=250 N/m)
(f) Low stiffness (k = 50 N/m)

Fig. 8: Learning curves for the push-grasp (blue) and the grasp-only (red) policies in the six considered conditions varying the design parameters of the SSG. Each plot shows the evolution of the grasp success rate over 2500 training steps.

1) Training - results and discussion: The evolution of the grasp success rate during training for all six conditions is shown in Fig. 8. More in detail, in the wide scoop condition (Fig. 8a), the grasp-only policy achieved 50% grasp success in 271 ± 136 steps and reached 80% at step 1798 ± 421. The AUC reached 1673.1 ± 155.5, significantly outperforming the push-grasp policy (AUC: 1200.7 ± 107.8). This suggests that widening the scoop enhances the gripper ability to clear clutter through grasping alone, thanks to a wider contact area. The narrow scoop configuration Fig. 8b still allowed grasp-only policies to achieve 50% success in 240 ± 31 and 80% at step 1783 ± 374. However, differently from the wide scoop case, the AUC is lower (1584.2 ± 120.2), and the average success rate throughout training remains below the 80% threshold. The push-grasp policy performed less effectively (AUC: 966.6 ± 179.7) and required 1083 ± 308 steps just to reach 50%, suggesting that narrower geometries limit the gripper's ability to exploit environmental constraints,



(a) Push-grasp

(b) Grasp-only

Fig. 9: Results of the simulated testing phase for: a) Push-grasp and b) Grasp-only, in the six conditions varying the SSG design parameters, across the 4 test scenes. Each bar reports performance in terms of Completion Rate (C.R.), Grasp Success (G.S.), and Action Efficiency (A.E.), expressed as percentages. The final plot in each figure presents the overall average performance across all scenes, including standard deviation bars.

resulting in slower learning.

In the high-friction condition (Fig. 8c), the grasp-only policy reached 50% success after 242 ± 51 steps and achieved 80% at step 1754 ± 306, achieving an AUC of 1671.0 ± 83.8. The push-grasp policy also benefited from high friction (AUC: 1380.8 ± 156.0), though with slower convergence and final lower success rate. The low-friction condition (Fig. 8d) caused small reductions in performance, with an AUC of 1602.7 ± 118.5 for the grasp-only policy which reached 50% after 276 ± 89 steps and 80% at 1671 ± 721 steps. The push-grasp success rates plateaued earlier, with an AUC of 1351.8 ± 171.3. These results show that while friction enhances contact stability, the difference among the tested conditions is not as evident as for other design parameters, thus suggesting that its overall effect on learning speed and final grasping proficiency is secondary.

Stiffness, instead, had a noticeable effect on learning dynamics. Reducing the hinge stiffness from k=1000 N/m to k=250 N/m (Fig. 8e) still allowed efficient learning under the grasp-only policy which reached 50% at 209 ± 25 steps and 80% by 1692 ± 514 steps, with AUC: 1665.1 ± 78.6. Conversely, push-grasp performance was lower (AUC: 1294.3 ± 184.1). When stiffness is further reduced at k=50 N/m, learning deteriorated (Fig. 8f). Grasp-only AUC dropped to 1447.2 ± 119.9, and the policy failed to reach 80% success. Push-grasp performance degraded even further (AUC: 993.1 ± 218.2), indicating poor sample efficiency and lower stability. These findings show that excessive compliance at the scoop hinge hinders the gripper ability to use the scoop to execute reliable grasps.

2) Testing - results and discussion: Results obtained in the conducted tests are reported in Fig. 9. With the push-grasp policy (Fig. 9a), the wide scoop yielded the best performance across all metrics (C.R.: $90.2\% \pm 4.1$, G.S.: $76.1\% \pm 11.9$, A.E.: $58.6\% \pm 9.4$). The wide geometry likely enhances the SSG ability during both pushing and grasping motions. The narrow scoop showed significant performance drops (G.S.: $49.6\% \pm 9.5$), likely due to reduced contact area and limited pushing efficacy. On the other hand, high friction improved grasp success ($66.5\% \pm 9.2$), while low friction reduced both G.S. and A.E. ($53.4\% \pm 8.2$), presumably due to slippage. Similarly, stiffness reductions affected performance. While $k = 250$ N/m maintained acceptable values (G.S.: $62.3\% \pm 4.6$), further lowering stiffness to 50 N/m led to very poor performance (G.S.: $53.2\% \pm 6.1$).

Under the grasp-only policy (Fig. 9b), the system was more robust to design changes. The wide scoop still performed best (G.S.: $81.6\% \pm 10.6$), followed by high friction (G.S.: $78.3\% \pm 13.4$). Interestingly, performance degradation is less critical when stiffness is reduced. While the $k = 50$ N/m stiffness condition showed decreased action efficiency ($72.3\% \pm 9.9$), it maintained high completion rates ($91.1\% \pm 0.5$) and grasp success ($72.3\% \pm 10.2$), suggesting that compliant scoops can still succeed when guided by a well-trained grasp policy.

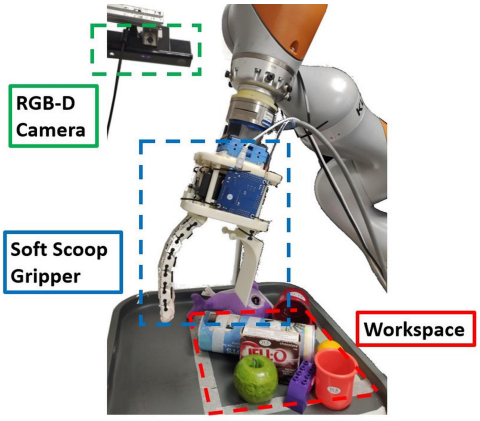


Fig. 10: Experimental setup. Red: workspace, green: RGB-D camera for object detection, blue: adopted soft-rigid gripper.

The conducted analyses demonstrate that the SSG learning dynamics and test performance are sensitive to design parameters, particularly in the push-grasp strategy. However, despite variations in the single conditions, the overall trends observed in Fig. 4 and in Fig. 6 are confirmed in Fig. 8 and Fig. 9, respectively. The behaviour of the SSG, which is more successful with the grasp-only policy rather than with the push-grasp one, remained consistent across different design conditions, suggesting that this behaviour relies more on the presence of the embedded scoop than on precise tuning of individual design parameters.

V. EXPERIMENTS

The results of the simulations showed that the SSG has a behaviour that is the opposite with respect to that of the rigid gripper and to evaluate whether this behaviour persists also in a real setup, we conducted real-world training and testing trials using the setup depicted in Fig. 10.

The experimental setup included a KUKA LBR iiwa robotic arm equipped with an ATI Gamma 6-axis force/torque sensor (ATI Industrial Automation, Inc.) at the wrist and the Soft ScoopGripper. A Kinect One RGB-D camera (Microsoft), statically mounted on a fixed shaft, was used to overlook the tabletop scenario and capture images of resolution 960×540. In both strategies, when the SSG had to grasp an isolated object, the gripper was tilted by a predefined angle so that the scoop could slide more easily underneath the object, as in [22]. Furthermore, a hybrid force-velocity controller was integrated into both the push and grasp actions to improve the safety and reliability of the manipulation process. In the push action, the controller maintained a constant force on the object surface while the scoop followed a straight trajectory along the table plane. During the grasp, the gripper started closing the fingers once it reached the table surface and the force sensed at the wrist of the robot exceeded a predefined threshold.

A total of 28 objects were used for training and testing in the real setup, most of them coming from the YCB dataset [27]. Their properties are listed in Table II. They were arranged in different ways to compose training and test scenes. An example of training scene is shown in Fig. 10, whereas the test scenes are shown in Fig. 11.

TABLE II: Dataset of objects.

	Object	Weight (g)	Size (mm)
Training objects	apple (YCB)	68	⌀75
	peach (YCB)	33	⌀64
	orange (YCB)	47	⌀63
	strawberry (YCB)	18	⌀44 × 55
	tennis ball (YCB)	58	⌀65
	baseball ball (YCB)	191	⌀96
	dolphin plush	84	80 × 200 × 90
	batman plush	102	140 × 70 × 140
	small cylinder	35	⌀40 × 130
	candy tube	12	⌀36 × 225
	scotch tape	75	⌀100 × 40
	spray can	118	⌀50 × 170
	chips can (YCB)	205	⌀75 × 250
	blue cup (YCB)	31	⌀90 × 74
	red cup (YCB)	21	⌀75 × 68
	metal mug (YCB)	90	⌀90 × 80
	pudding box (YCB)	187	35 × 110 × 89
	gelatin box (YCB)	97	28 × 85 × 73
	spring clamp (YCB)	59	90 × 115 × 27
	lego brick (YCB)	2.5	95 × 35 × 40
	plastic funnel	120	⌀120 × 130
Test objects	banana (YCB)	66	⌀35 × 175
	tuna fish can (YCB)	171	⌀85 × 33
	body spray	34	⌀45 × 175
	green cup (YCB)	17	⌀65 × 64
	mouse box	27	40 × 70 × 105
	wooden block	374	140 × 63 × 43
	mouse box	146	90 × 175 × 60

(a) Test #1 (b) Test #2 (c) Test #3 (d) Test #4

Fig. 11: Real-world test scenes: (a), (b), are composed of new objects, whereas (c), (d), are composed of objects used also in training scenes, but disposed in different ways.

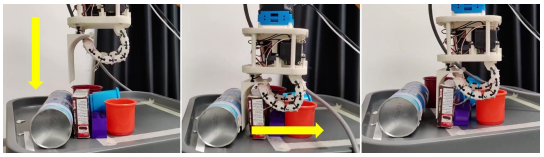
A. Results from Training

In the training phase with the real system, we let the robot declutter several different scenes, each including 12 out of the 21 training objects listed in Table II. The objects in the scene were randomly chosen and randomly placed in the robot workspace. Examples of successful pushing and grasping actions during training are shown in Fig. 12. The scene was changed either when all objects were successfully removed, or when the robot failed 10 consecutive actions.

The training curves of the two adopted strategies for the real setup are shown in Fig. 13. Each of the conducted training phases consisted in 2500 steps.

B. Results from Testing

To test the trained networks, the 4 scenes reported in Fig. 11 were decluttered with the SSG. Test scenes were composed either of completely new objects (Fig. 11a, Fig. 11b) or of objects used also in the training phase (Fig. 11c, Fig. 11d). It should be noted that since the number and the arrangement of the objects are completely different between training and



(a) Pushing action

(b) Grasping action

Fig. 12: Actions performed by the SSG. The arrows represent the motions of the gripper. (a) Push action: the gripper inserts the scoop between two objects and then performs the pushing. (b) Grasp action: the gripper moves towards the table after having reached the desired position above the object and then closes the fingers over it.

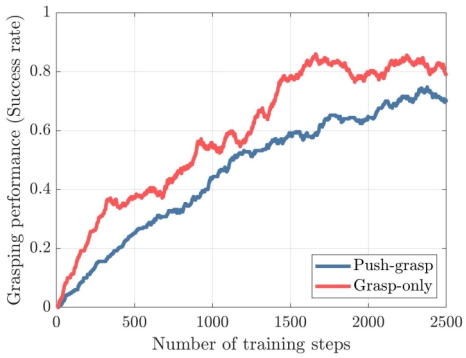


Fig. 13: Learning curves obtained in a real setup with the SSG in the training phase considering only the grasping action (red) or both pushing and grasping actions (blue).

testing, and since the whole learning algorithm is not based on the recognition of the single object but on the overall heightmap, this choice did not bias the results (see also Sec. V-C). Each test scene was tackled 10 times with the grasp-only network and 10 times with the push-grasp network, reaching a total of 80 test trials. The values obtained for the metrics introduced in Sec. III-E across different trials for the same scene are reported in rows 2-5 of Table III, whereas the mean and standard deviation among all trials for all scenes are in the last row.

C. Discussion

The trends observed in the real-world experiments align closely with those seen in the simulations.

In the training phase, the grasp-only strategy achieved higher success rates, consistently surpassing the performance of the push-grasp strategy. By the end of the training, the grasp-only strategy stabilized at a success rate of around 80%, while the push-grasp strategy reached a lower final success rate of approximately 68% (see Fig. 13). Thus, we confirmed in a real environment that the additional pushing action can be avoided when the gripper embeds "intelligent" components in its design. Specifically, in this work, a rigid, passive element, the scoop, attached to the hand via a flexible joint allows the robot to easily move the objects and/or adapt to them directly during the grasping action.

TABLE III: Results of the real testing phase performed with the SSG. All results are presented in percentage. Rows 2-5 represent the results obtained for each scene of Fig. 11, whereas the last row reports the average and standard deviation values.

Test Scene	Strategy	Completion rate	Action efficiency	Grasp success
1	Push-grasp	75.3	69.1	82.8
	Grasp-only	**82.0**	**85.1**	**85.1**
2	Push-grasp	68.9	65.4	78.5
	Grasp-only	**79.0**	**84.6**	**84.6**
3	Push-grasp	74.5	68.3	82.4
	Grasp-only	**81.4**	**85.4**	**85.4**
4	Push-grasp	74.9	66.0	81.5
	Grasp-only	**80.4**	**87.3**	**87.3**
All	Push-grasp	73.4 ± 2.9	67.2 ± 1.8	81.3 ± 1.9
	Grasp-only	$\mathbf{80.7 \pm 1.3}$	$\mathbf{85.6 \pm 1.2}$	$\mathbf{85.6 \pm 1.2}$

Concerning the results of the test trials, as it can be seen from Table III, all scenes present similar results. However, test scene #2 records the lowest values of the three metrics. We hypothesize that in this particular scene, it was harder for the network to discriminate the edges of the objects as they have similar heights. Also, the wooden block is rather heavy: so it was very difficult to grasp.

Overall, the grasp-only strategy outperforms the push-grasp network in all the adopted metrics. Especially the action efficiency disparity between the two strategies is remarkable. The number of necessary actions in the grasp-only trials is lower with respect to the push-grasp trials. This happens because the hand can directly grasp the objects, even in highly cluttered scenarios, without using non-prehensile actions. Conversely, the push-grasp strategy requires more actions to declutter scenes completely. A similar reasoning holds when comparing the completion rates, where it can be seen that the grasp-only policy leads to the complete clearing of the scene more frequently with respect to the push-grasp strategy, which fails more often because of the inefficiency of the push action. We also noticed that the grasp success rate is comparable for both strategies. This is due to the embodied mechanical intelligence of the hand, which can grasp objects even without a preparatory push action.

VI. Conclusion and Future Work

In this paper, we integrated a soft-rigid gripper able to safely and effectively interact with the environment, the Soft ScoopGripper, with an active exploration algorithm to empty a cluttered scene. We trained and tested two reinforcement learning policies both in simulation and in a real setup. The first policy was characterized by pushing and grasping actions, while the second relied only on the grasping action.

In simulation, we conducted a comparative analysis between a rigid, soft, and soft-rigid gripper, which revealed significant insights. As expected from previous work [6], the rigid gripper performed best with the push-grasp strategy, as it could move objects before grasping, but struggled with the grasp-only strategy due to its limited adaptability. The SSG proved to be the most effective with the grasp-only strategy, as its combination of soft and rigid elements allowed for direct grasping without the need for additional pushing actions. The soft gripper showed similar results with both strategies, but generally performed worse than the other two grippers in the push-grasp strategy, and worse than the soft-rigid gripper in the grasp-only. An additional simulation analysis was conducted to better understand the role of design parameters in the SSG training and testing performance. Results showed that scoop width and hinge stiffness affect learning speed and task success, particularly for push-grasp strategies, while grasp-only policies remained more robust to such variations. Despite these differences, the overall performance of the SSG remained consistent across design changes, suggesting that its effectiveness is driven more by the presence of the embedded scoop-shaped component than by the fine-tuning of specific design parameters.

The results obtained for the SSG in simulation were confirmed by the experimental results which showed that a simpler policy (i.e., grasp-only) learns and performs better in real test scenes than the two-actions policy (i.e., push-grasp). The objective of the conducted experiments was to confirm that the simplification of the learning process observed for the SSG in simulation translated to real environments. Thus, a comparison with the other two grippers or the evaluation of more challenging scenarios was beyond the scope of this paper. Indeed, the considered test scenes were chosen to ensure consistency with previous work, and complex situations with occlusion, packing or stacking of objects were not taken into account and are left for future investigation.

Building upon a previous study where scoop-like structures were added to commercial soft grippers [28], future work will study a RL framework that can be applied to any gripper that embeds a scoop, investigating whether the scoop-like element can foster transfer learning.

REFERENCES

[1] A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt, "Grasp pose detection in point clouds," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1455–1473, 2017.

[2] R. Newbury, M. Gu, L. Chumbley, A. Mousavian, C. Eppner, J. Leitner, J. Bohg, A. Morales, T. Asfour, D. Kragic, *et al.*, "Deep learning approaches to grasp synthesis: A review," *IEEE Transactions on Robotics*, 2023.

[3] A. Bicchi and O. Brock, "Editorial," *The International Journal of Robotics Research*, vol. 39, no. 14, pp. 1601–1603, 2020.

[4] C. Piazza, G. Grioli, M. Catalano, and A. Bicchi, "A century of robotic hands," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 1–32, 2019.

[5] O. Brock, "Intelligence as computation," in *IOP Conference Series: Materials Science and Engineering*, vol. 1321, no. 1. IOP Publishing, 2024, p. 012001.

[6] A. Zeng, S. Song, S. Welker, J. Lee, A. Rodriguez, and T. Funkhouser, "Learning synergies between pushing and grasping with self-supervised deep reinforcement learning," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4238–4245.

[7] G. Salvietti, Z. Iqbal, M. Malvezzi, T. Eslami, and D. Prattichizzo, "Soft hands with embodied constraints: The soft scoopgripper," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Montreal, Canada, May 2019, pp. 2758–2764.

[8] E. Turco, V. Bo, M. Pozzi, A. Rizzo, and D. Prattichizzo, "Grasp planning with a soft reconfigurable gripper exploiting embedded and environmental constraints," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, p. 5215–5222, 2021.

[9] L. Berscheid, P. Meißner, and T. Kröger, "Robot learning of shifting objects for grasping in cluttered environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 612–618.

[10] A. M. Sundaram, W. Friedl, and M. A. Roa, "Environment-aware grasp strategy planning in clutter for a variable stiffness hand," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 9377–9384.

[11] T. Nishimura, K. Akasaka, S. Ishikawa, and T. Watanabe, "Single-motor-driven (4+ 2)-fingered robotic gripper capable of expanding the workable space in the extremely confined environment," *IEEE Robotics and Automation Letters*, 2024.

[12] M. R. Dogar and S. S. Srinivasa, "A planning framework for non-prehensile manipulation under clutter and uncertainty," *Autonomous Robots*, vol. 33, no. 3, pp. 217–236, 2012.

[13] M. Gualtieri and R. Platt, "Learning 6-dof grasping and pick-place using attention focus," in *Conference on Robot Learning*. PMLR, 2018, pp. 477–486.

[14] A. Wu and D. Kruse, "In the wild ungraspable object picking with bimanual nonprehensile manipulation," *arXiv preprint arXiv:2409.15465*, 2024.

[15] R. Deimel, C. Eppner, J. Álvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," in *Robotics Research*. Springer, 2016, pp. 393–409.

[16] F. Iida and F. Giardina, "On the timescales of embodied intelligence for autonomous adaptive systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 6, no. 1, pp. 95–122, 2023.

[17] A. Bhatt, A. Sieler, S. Puhlmann, and O. Brock, "Surprisingly robust in-hand manipulation: An empirical study," in *Robotics Science and Systems (RSS)*, 2021.

[18] J. Ye, J. Wang, B. Huang, Y. Qin, and X. Wang, "Learning continuous grasping function with a dexterous hand from human demonstrations," *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 2882–2889, 2023.

[19] C. Choi, W. Schwarting, J. DelPreto, and D. Rus, "Learning object grasping for soft robot hands," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2370–2377, 2018.

[20] C. Della Santina, V. Arapi, G. Averta, F. Damiani, G. Fiore, A. Settimi, *et al.*, "Learning from humans how to grasp: a data-driven architecture for autonomous grasping with anthropomorphic soft hands," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1533–1540, 2019.

[21] G. Salvietti, Z. Iqbal, I. Hussain, D. Prattichizzo, and M. Malvezzi, "The co-gripper: a wireless cooperative gripper for safe human robot interaction," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4576–4581.

[22] E. Turco, V. Bo, M. Tavassoli, M. Pozzi, and D. Prattichizzo, "Learning grasping strategies for a soft non-anthropomorphic hand from human demonstrations," in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2022, pp. 934–941.

[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[24] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[25] E. Rohmer, S. P. N. Singh, and M. Freese, "V-rep: A versatile and scalable robot simulation framework," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1321–1326.

[26] M. E. Taylor and P. Stone, "Transfer learning for reinforcement learning domains: A survey." *Journal of Machine Learning Research*, vol. 10, no. 7, 2009.

[27] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set," *Robot. Automat. Mag.*, vol. 22, 2015.

[28] V. Bo, E. Turco, M. Pozzi, M. Malvezzi, and D. Prattichizzo, "Automated design of embedded constraints for soft hands enabling new grasp strategies," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11 346–11 353, 2022.