Evaluating a Movable Palm in Caging Inspired Grasping using a Reinforcement Learning-based Approach

Luke Beddow, Helge Wurdemann, and Dimitrios Kanoulas

Abstract— **In this paper, we study the effectiveness of using a rigid movable palm for grasping varied objects, on a caging inspired gripper with three flexible fingers. This rigid palm extends to actively exert downwards force on objects, in contrast with existing methods, which combine movable palms with negative pressure to exert lifting forces on objects. We compare grasping with and without the palm, whilst also changing finger stiffness and fingertip angle, to analyse the effect on grasp success rate and stability over 24 design permutations. Reinforcement learning was used to train a unique grasping controller in every design case, aiming to achieve optimal grasping as the basis for comparison. Validation in both simulation and the real world was completed for every permutation. We demonstrated that the using palm improved success rates on average by 11% in simulation, 13% in the real world, and achieved a best real world success rate of 96% on 18 YCB benchmark food objects. Grasp stability against disturbances in three axes improved by 15% on average when using the palm. Our investigation determined fingertip angle had a large effect, whereas finger stiffness was less important.**

I. Introduction

Grasping varied items with a single gripper is a challenging design problem, with applications such as logistics and grocery packing. RightHand Robotics commercial logistics gripper, RightPick [1], uses a movable suction cup palm, actuating three fingers to secure objects following establishment of suction grasps. Suction is commonly used because it offers a simple and versatile way to grasp a large variety of objects. However, suction is not always applicable, such as in the case of grocery item grasping. Fruits and vegetables rarely have flat surfaces and exhibit varied textures, making them challenging for suction grippers [2], [3]. In fact, grocery items are challenging for many grippers due to their highly diverse weights and sizes as well as their propensity to bruising, which requires force limiting. Suction is also less effective against horizontal disturbances, because it applies vertical force, hence its combination with stabilising fingers.

The RightPick, as well as many researchers proposing suction grippers with added fingers, make the suction cup palm movable [4]–[7], to ensure the ideal relative position of the palm and the fingers for objects of different sizes. Therefore, in addition to exerting a holding force, the fingers and palm

The authors are with the Departments of Computer Science and Mechanical Engineering, University College London, Gower Street, WC1E 6BT, London, UK. {luke.beddow, h.wurdemann, d.kanoulas}@ucl.ac.uk. Corresponding author: *Luke Beddow*

This work is supported by the UCL EPSRC DTP in Fundamental Engineering [2425110] and the the UKRI Future Leaders Fellowship [MR/V025333/1] (RoboHike).

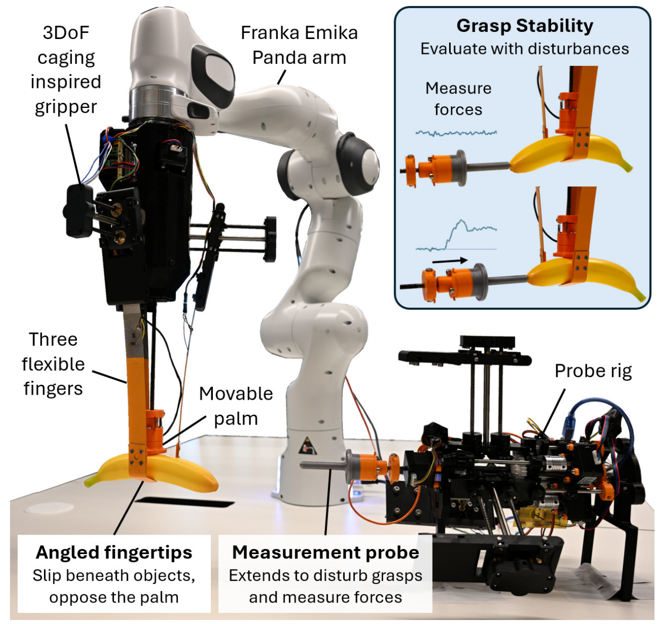


Fig. 1. Experimental setup, with gripper mounted on the robot arm to autonomously perform grasps. Objects were grasped by surrounding them with three flexible fingers and a movable palm, all equipped with force sensing. A fixed probe rig was then used to disturb and push objects out of grasp whilst measuring the force, in order to evaluate grasp stability.

can surround and stabilise the object. This can potentially cage the object, which refers to trapping an object within the gripper so that it cannot escape. Rather rely on caging or frictional forces, existing research combines movable palms with negative pressure, either the aforementioned suction or granular jamming [7], [8], to exert an upwards holding force. Intuitively, this addresses an issue of movable palms exerting downwards forces on objects, and pushing them away from the gripper and out of grasp. However, when combined with caging, we propose that a rigid palm exerting downwards force in this manner can push objects into the fingers in order to actually improve grasp stability.

Our previous work [9], presented a three degrees of freedom (DoF) caging inspired gripper with three flexible fingers and a rigid movable palm, which actively pushed on objects. Force sensing on the fingers and palm was used alongside a reinforcement learning grasping controller to achieve grasps on grocery objects whilst limiting surface forces. This design could cage many objects, but relied on friction to secure objects that could not be caged. This friction came from the fingers bending and squeezing objects, and also the palm pressing onto objects. However, it was not clear whether the palm improved grasp success rate or stability. Without the palm, objects could still be subject to an energy-bound cage, given that gravity prevents escape upwards. In addition, failure cases were observed where the palm pushed objects out of grasp. No investigation by the present authors or other researchers has determined the efficacy of a rigid movable palm, actively exerting downwards force in this manner.

In the current work, we compared grasping success rate and stability with and without using the movable palm. We also varied finger stiffness and fingertip angle, the two key factors influencing how the fingers provide an opposition force to the palm during grasping, resulting in 24 design permutations. A reinforcement learning-based approach was used, training a suite of controllers, one for each design permutation, to optimise grasping in every case to ensure a fair comparison. We firstly validate in simulation, as in similar reinforcement learning design comparison approaches [10]–[12], but go further to provide real world validation as well. We compared grasping success rate, and evaluated grasp stability along three different disturbance axes, using the 18 YCB benchmark food items [13]. We ran real world experiments on all 24 permutations, totalling 1752 grasps, with the gripper and experimental setup shown in Figure 1.

The contribution of this work is an exploration of the efficacy of using a movable palm for caging inspired grasping, using a reinforcement learning-based approach. We present an extensive real world validation, measuring grasp success rate and grasp stability with and without the palm, including the effect of finger stiffness and fingertip angle. The effectiveness of training grasping controllers with reinforcement learning as a means to compare designs is demonstrated.

This paper is organised as follows. Related work is highlighted in Section II. The gripper and grasp design features are given in Section III. The reinforcement learning exploration approach is explained in Section IV. The experimental procedure for validation is detailed in Section V, before results, discussion, and conclusions in Sections VI–VIII.

II. Related Work

Caging in 3D is an unsolved problem [14], hence caging grasping often considers top down 2D cages, most frequently using rods or pins to squeeze objects and secure them with friction [15], [16]. In these cases, a movable palm would inhibit grasping by pushing objects out of grasp. Backus et al. [17] demonstrated how this 2D caging approach was improved with curling fingertips to get underneath objects, similar to the angled fingertips used in the present work.

Movable palms for fingered grippers have been combined with negative pressure, either suction cups [1], [4], [5] or granular jamming [8], where vacuum is applied to a bag of grains to make it rigid. In both of these cases, the objective is to exert an upwards, holding force on objects. Deng et al. [6] applied reinforcement learning to aid grasping with their two-finger gripper with movable suction cup palm. Pagoli et al. [7] tested both suction and jamming for their soft gripper and used the lifting force to help with in-hand manipulation. Teeple et al. [18] presented a rigid movable palm, again for in-hand manipulation, but avoided exerting downwards forces by orientating their gripper upwards or sideways. In contrast, the rigid movable palm investigated in our work exerts a downwards force on the object, to complete a cage, and we evaluate the effect on grasp success and stability.

Reinforcement learning has been used for design optimisation of parts subject to loads [19], [20]. It has also been used to effectively search design spaces [21]–[23] to discover optimal designs. However, these approaches required an expert evaluator, finite element analysis, to provide ground truth in regards to optimality. Other approaches rely not on an expert, but evaluation of performance in simulation. Schaff et al. [10] and Ha [11] demonstrated learning of design parameters jointly with a control policy. They evaluated using simulated performance of the resultant policy, as we do in the present work. Belmonte-Baeza et al. [12] applied similar approaches to quadrupedal robots, combining a learned locomotion policy with an evolutionary algorithm for iterating the design parameters. Rather than iterate design parameters, the present work fixed a set of design permutations beforehand and then evaluated those by learning controllers for each in simulation. In contrast to previous works, we validated performance in the real world in addition to simulation.

Training grasping controllers with reinforcement learning is a popular approach [24], and has often been done in simulation [25]. We utilised a simulated training method from our previous work [9] which learns grasping based only on force sensing. Similar approaches have used rigid commercial grippers such as the Barrett hand [26], [27], Allegro hand [28], and Shadow hand [29], with a focus on enveloping grasps, similar to the caging grasping in this work, but not involving a movable palm.

III. Design Variations and Disturbances

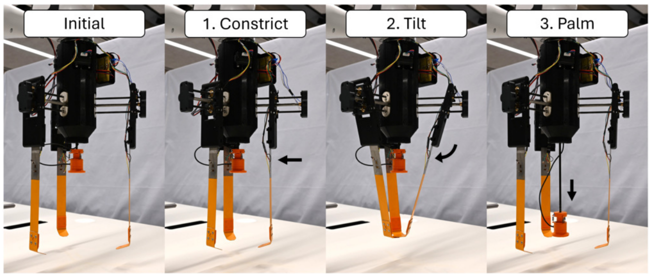
The gripper design is introduced, alongside variations to fingertip angle and finger stiffness. We detail our method of measuring grasp stability with disturbances along three axes.

A. Grasping Principle and Gripper Design

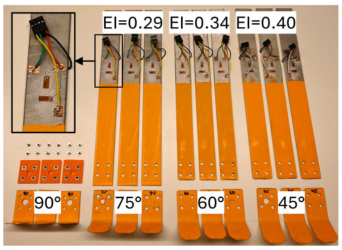
This subsection summarises the gripper design and grasping approach from previous work [9]. The gripper has three flexible fingers and a movable palm, and is actuated in 3DoF. The gripper and actuations are shown in Figure 2-a, with the mechanically coupled fingers being able to tilt and constrict to surround and slip underneath objects. The palm can extend to push down on objects, and contains a load cell to measure contact force. The gripper fingers are made of stainless steel and intended to bend, being instrumented with strain gauges for force sensing. The fingers have an angled bend at the tip, which can slip underneath objects and prevent them falling out of grasp. Figure 3 shows a caged object, the apple, which is surrounded on all sides by the fingers and palm; as well as an object that cannot be caged, the pringles can. This object requires friction to prevent it sliding out along its long axis.

B. Varying Fingertip Angle and Finger Stiffness

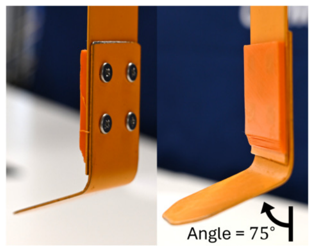
During grasping, the palm actively applies a downwards force on objects, which must be opposed by an upwards force



(a) Gripper motions for each of the three degrees of freedom



(b) All fingers and fingertips



(c) Assembled 75° fingertip

Fig. 2. Overview of: a) Gripper DoF; b) Fingers and fingertips used for experiments, with stiffnesses (EI) and fingertip angles labelled, as well as zoom in on finger strain gauges (same configuration underneath, not shown); c) An assembled 75° fingertip, with the angular dimension indicated.

from the gripper fingers to keep objects from being pushed directly out of grasp. Two key design parameters influence this upwards force in how they change grasp geometry and frictional forces: the fingertip angle and the finger stiffness.

When fingertips get beneath objects, they provide reaction force to act in opposition to the palm. The fingertip angle affects the proportion of vertical and horizontal reaction force. A larger angle, up to $90°$, provides more direct vertical force in opposition to the palm. A shallower angle, like $45°$, results in greater horizontal force, liable to push the fingers to bend and open up such that the object may fall out of grasp. However, a shallower angle may improve the fingertips ability to slip underneath objects in the first place.

Finger stiffness also influences opposition to the palm, by adjusting the amount of force required to bend the fingers and change the geometry of the grasp, i.e., break the cage. When the fingers tilt, their bending directly opposes the palm, and as discussed, horizontal components of fingertip reaction forces also lead to bending. Stiffer fingers deflect less to better maintain a cage around objects; however, they allow less adaptation to objects whilst remaining within force limits, which could reduce grasp quality.

We vary both the fingertip angle and finger stiffness in our grasping experiments (Section V), in order to evaluate how these two variables affect grasping with and without the palm. Figure 2-b shows the three sets of fingers and four sets of fingertips used for comparison, labelling the different angles and stiffness values. All were manufactured from 304 stainless steel, which had appropriate elastic properties for bending, and along with the even spray painted coat, aimed to provide a consistent friction coefficient.

The fingers are cantilever beams, with beam stiffness against an applied load, EI. This is the product of the Youngs Modulus for 304 stainless steel, $E = 193$GPa [30], and cross-sectional 2^{nd} moment of area, I. We varied I to achieve three different finger stiffnesses. For our rectangular cross-section fingers, $I = t^3w/12$ where t is thickness (in the bending direction) and w is width. Hence, we used three fingers with $t \times w$ (all mm): 0.86×28.0; 0.96×24.0; 0.96×28.0; resulting in $EI = [0.29, 0.34, 0.40]\text{Nm}^2$. These have corresponding yield point end loads of $[3.2, 3.4, 3.9]$N. We were interested in grasping with contact forces from $1-$3N, therefore these stiffnesses effectively span between our minimum (yield just above 3N) and a reasonable maximum, 40% stiffer, beyond which deflection at 3N reduces below 30mm and may hinder adaptation to objects.

Fingertips with varied angles were designed to screw onto the end of any of the fingers. As shown in Figure 2-c, the fingertips used a plastic front plate and four screws to fasten them to the end of a finger. This connection was convenient, low profile, and very stable. Fingertips were all 28mm wide and, following the bend, 35mm long. Their front edges, which contacted objects, were smoothly filleted to a 7.5mm radius. Bend angle tolerance was $\pm 2°$, and did not change following the experiments. We produced four sets of fingertips, at angles: $[45, 60, 75, 90]°$. This range was chosen based on pilot testing in simulation (Section IV), where grasping success reduced noticeably outside this range.

C. Measuring Grasp Stability with Disturbances

Evaluation of grasp stability was done using disturbances applied along three axes. We chose two axes in the horizontal plane, labelled X and Y, and also the vertical axis Z (coincident with gravity), all of which are shown in Figure 3.

The two horizontal axes were chosen to characterise the range of disturbances that could be tolerated. The X axis was chosen as the most stable axis of grasp, disturbing the object in a direction which could be directly opposed by a gripper finger. The Y axis was chosen as the least stable axis of grasp, disturbing in a direction to push the object between two gripper fingers. For objects that are not caged, this means the disturbance is opposed by no gripper fingers, instead requiring friction to resist, as shown in Figure 3.

The Z axis corresponded to disturbing objects vertically, directly out of grasp. We used the palm itself to push objects out of grasp, using the embedded load cell to measure the force required. Measurement uncertainty was ± 39mN, at the 95% confidence interval (CI) [9].

The X and Y disturbances were applied to grasps by fixed "probe rig", shown in Figure 1. It featured a stepper motor driven measurement probe which could extend linearly up to 165mm via a non-backdrivable leadscrew. The probe contained a load cell to measure force, and used a 30mm diameter 3D printed finger to push on objects. In fact, this probe rig was a repurposed previous prototype of the gripper itself, and the probe a repurposed movable palm, with the same design as the gripper palm. Force measurements had uncertainty ± 31mN (95% CI). The measurement procedure during testing is given in the upcoming Section V-B.

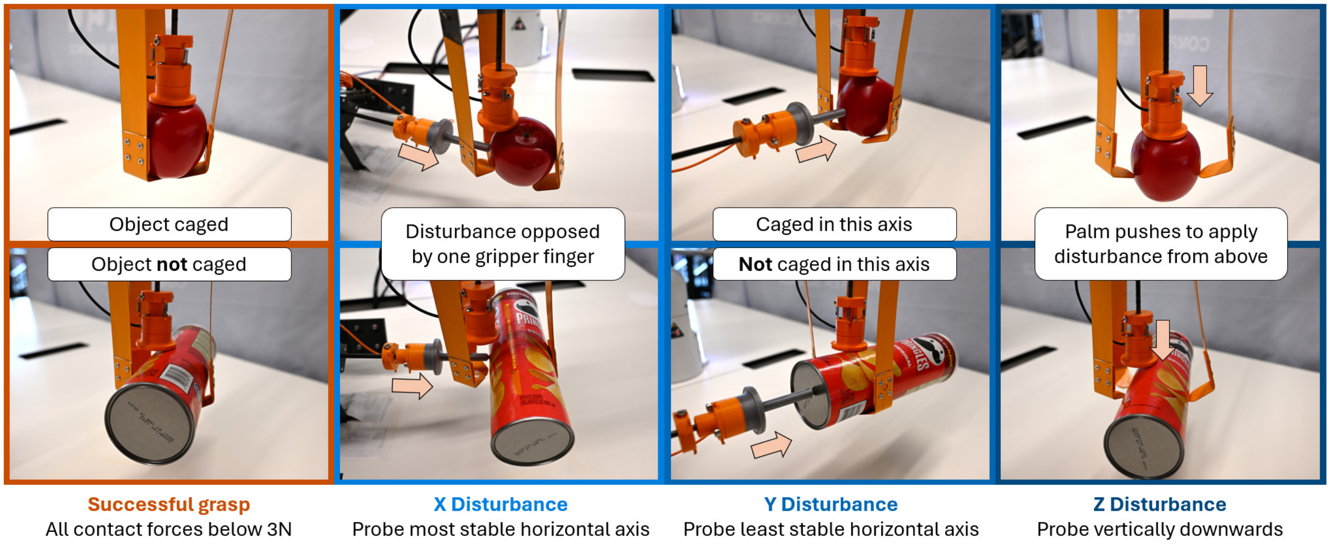


Fig. 3. Demonstration of objects which can and cannot be completely caged, as well as the corresponding disturbance directions used during experimental evaluation of grasp stability. The X direction was chosen as the most stable horizontal direction, Y as the least stable, and Z to act vertically.

IV. Reinforcement Learning-based Exploration

We evaluated how the movable palm affects grasping using a reinforcement learning-based approach. The central idea was to compare multiple design permutations to determine which resulted in the most effective grasping. Therefore, in order to make it a fair test, grasping should be optimised for every permutation, and so a controller was needed that could maximise the chance of success in each and every case. We consider that reinforcement learning can be used to train high quality controllers for grasping [31], [32]. Then, these trained controllers can be evaluated by testing grasping in simulation and the real world, in order to compare the efficacy of different designs. Using a data-driven approach such as reinforcement learning removes any constraint that these controllers should grasp in the same way, or human bias that may expect that certain strategies would be the best. Learning can exploit the unique features of each permutation, and ideally present the optimal strategy. In practice, optimality is not guaranteed, and training procedures can introduce bias.

A. Reinforcement Learning for Grasping Controllers

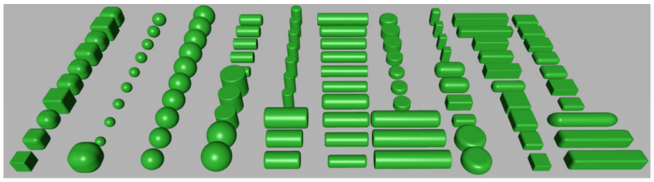
Training grasping controllers for each design permutation of the caging inspired gripper was done using a reinforcement learning grasping approach outlined in our previous work [9], which is briefly summarised in this subsection. Simulation of the gripper in MuJoCo was used to train with proximal policy optimisation [33] a neural network to output the next grasping action, based on an observation of the current state. There were four continuous actions, the three gripper DoF (as shown in Figure 2-a), and an action to change the gripper height, actuated by the Franka Emika Panda robot arm, upon which the gripper was mounted. The arm also provided the vertical force at the wrist, which in addition to the force sensor data from the fingers and palm, as well as the positions of the 4DoF, made up the state observations for force feedback grasping. The simulated training set was composed of 1500 elementary objects, such as spheres, cuboids, and cylinders, with randomised densities and friction coefficients. Noise was added to every aspect of the state observation to enable generalisation from elementary objects to real, irregular objects. The flexible fingers of the gripper were accurately simulated depending on their stiffness, and the controller must respect force limits, so episodes terminated if any contact forces exceeded 5N.

B. Simulated Training for Design Permutations

We applied this reinforcement learning-based approach to a fixed design space, rather than actively exploring the design space with reinforcement learning [10], [11] or evolutionary algorithms [12]. We discretised our design space by varying designs along three parameters:

- 2 palm conditions: with the palm, without the palm.
- 3 finger stiffnesses: $EI = [0.29, 0.34, 0.40]\text{Nm}^2$.
- 4 fingertip angles: $[45, 60, 75, 90]°$.

This gives a total of $2 \times 3 \times 4 = 24$ design permutations, for which each would have grasping controllers trained in simulation. We added to the MuJoCo environment the capability to adjust the gripper designs to meet our requirements for all 24 design permutations. For example, when training a controller for the design permutation that uses the palm, has a fingertip angle of $60°$, and a finger stiffness of $EI = 0.29\text{Nm}^2$, all of these features would be accurately recreated in the physics simulator. Adding the capability for grasping without using the palm introduced two changes to the existing training method (Section IV-A). When grasping with the palm, there were four output actions, whereas without the palm this reduced to three. The state observation also changed, the palm position was no longer included and neither was the force data from the load cell in the palm. Trainings took $20-30$ hours as CPU processes.



(a) The 100 elementary objects in the simulation test set.

(b) The 18 YCB benchmark [13] food objects.

Fig. 4. Objects used for testing a) in simulation and b) in the real world.

V. Experimental Method

An experiment was performed to evaluate grasping with and without the palm, measuring grasp success rate in simulation and real life, and measuring grasp stability in real life. This experiment was composed of 24 tests, covering all the design permutations given in Section IV-B, which varied palm usage, finger stiffness, and fingertip angle.

A. Test Procedure

Each of the 24 tests was conducted in the same manner. The design permutation was set in simulation for 10 training runs. The grasping controller with the highest simulated success rate was then chosen for real robot evaluation. The object sets differed between simulated and real testing, but the definition of a trial and a successful grasp was the same.

1) Object Sets: For simulated testing, the object set was a fixed 100 unseen objects from the training set, shown in Figure 4-a. For real testing we used the 18 food objects from the YCB benchmark set [13], shown in Figure 4-b. Grocery items can be approximated as elementary objects, and the YCB set covers a variety of shapes and sizes, but the irregularities would challenge the controllers to generalise.

2) Grasping Trial: Objects were placed directly beneath the gripper, with $\pm$15mm of uniformly random position noise and in any random rotation about the vertical which avoided collision with the gripper fingers. The gripper fingertips began 10mm above the table surface, with an additional $\pm$5mm of noise. Trials ended upon completion of a successful grasp, if the 250 action limit was exceeded, or if the object became ungraspable, for example rolling out of reach.

3) Successful Grasp: The gripper must support the entire weight of the object, whilst the average contact force from all three fingers, and the force from the palm (if in use), both exceeded 1N. No individual finger or palm force could exceed 4N. The gripper must have raised higher than 20mm above the table surface. Successful grasps were automatically detected using sensor data. In real life, successful grasps were then human validated and included the robot arm lifting an additional 30mm to confirm the object remained in grasp.

In simulation, 20 trials were performed on each of the 100 objects, then the success rate percentage (SR%) was reported. This was the number of successful grasps as a percentage of the total number of trials.

In real life, the number of trials was not fixed for each of the 18 YCB objects. Grasp stability was assessed using disturbances applied along the three axes (shown in Figure 3 and procedure explained in upcoming Section V-B). Therefore three separate successful grasps were required to test all three disturbances. Hence, trials continued for each object until three successful grasps had been achieved, or there had been five failed grasps. This meant each object had $3-7$ trials, depending on performance, and the minimum number of trials to cover all 18 objects was 54. The actual 24 tests conducted averaged 73 trials each, with a total of 1752 trials. The SR% was calculated by taking the average of the success rates per object, to ensure that every object was weighted equally with respect to the final SR%.

B. Disturbance Measurement Approach

Grasp stability was measured by applying disturbances in three axes (X, Y, Z), as introduced in Section III-C and shown in Figure 3. The X and Y disturbances were applied using the probe rig, and the Z disturbance using the palm.

For each disturbance, the probe, either from the rig or the palm itself, would extend until the point of first contact with the object. Then, it extended in 2mm increments, measuring the force required to continue pushing forwards, in order to get the maximum force that the grasp was able to tolerate. The probing would finish if: 1) the maximum force required exceeded 10N; 2) the object fell out of grasp; 3) the probe reached 50mm from the point of first contact; 4) finger bending beyond the yield load was detected; or 5) to prevent an unfair test in cases where the probe would contact the gripper fingers or the object would touch the table.

Disturbance application was partially automated, with the human operator fine-tuning the position to apply the probe. For some objects, such as the yellow mustard bottle, the initial object orientation affected the resulting grasp and probe axes for X and Y. Hence, to improve grasp consistency and disturbance comparability for these objects between different tests, they were constrained to an initial random orientation within only a $60°$ range before trials where X and Y disturbances were applied, rather than the full $360°$.

The data for grasp stability were averaged over all of the objects probed in each test. Readings were capped at 10N for the purpose of taking averages. Since three successful grasps were required to probe all three disturbances, some tests did not have certain probes on certain objects. These cases were excluded, with the average only being done across the taken measurements. Following the experiments, each test averaged probes on $16.2/18$ objects for each of X, Y, and Z.

TABLE I

AVERAGED RESULTS OVER ALL 24 TESTS.

Use palm	Simulation success rate %	Real success rate %	Tolerated disturbances (N) X	Y	Z
Yes	**89**	**82**	**3.56**	**2.56**	**6.50**
No	78	69	3.32	2.00	5.87

TABLE II

MAXIMUM RESULTS FROM ALL 24 TESTS.

Use palm	Simulation success rate %	Real success rate %	Tolerated disturbances (N) X	Y	Z
Yes	**94**	**96**	**4.32**	**3.26**	**8.90**
No	87	91	4.00	2.88	8.34

VI. RESULTS

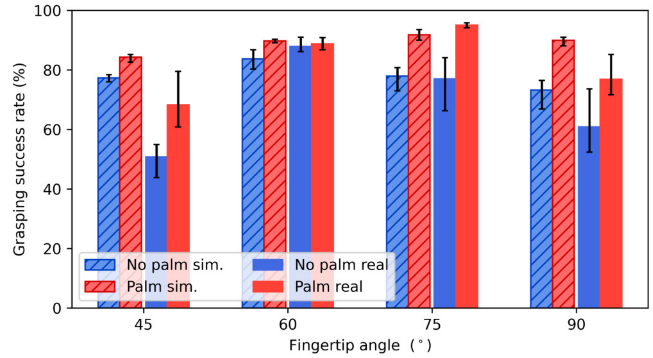
The overall averages from all 24 tests are shown in Table I. Using the palm resulted in an average improvement of: 11% for simulation success rate, 13% for real success rate, 7% for X disturbance tolerance, 28% for Y tolerance, and 11% for Z tolerance. Therefore, the palm improved grasping across all metrics, on average by 16%. Table II shows the maximum results from the 24 tests, in which the palm again improved all metrics, by 8% on average. The highest real world success rate of 96% was achieved using the palm.

The success rate results from all 24 tests are shown in Figure 5-a. Using the palm resulted in an improvement for all fingertip angles, both in simulation and in real life. Real grasp success rates were best when using the palm at 75°, achieving 94%, 95%, and 96% over the three finger stiffnesses. Real grasp success rates without the palm were best at 60°, averaging 88% and peaking at 91%, almost equal to with the palm averaging 89%, and peaking also at 91%.

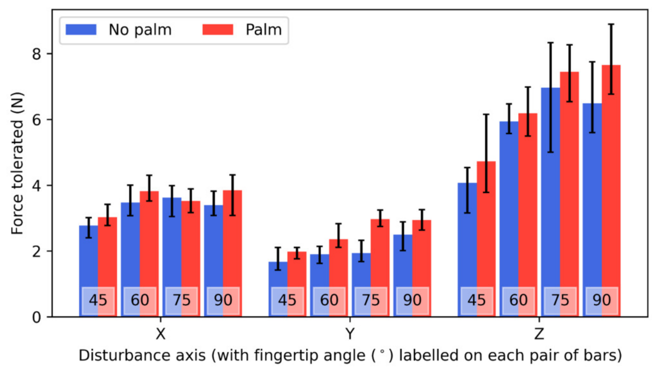
The simulation and real world results were in agreement, especially for the 60° and 75° fingertip angles, which had an average simulation to real (sim2real) gap of only 3%. Across all the fingertip angles, the sim2real gap was 8% with the palm and 13% without the palm. 45° fingertips without the palm had the largest gap and lowest real success rates.

Comparing success rate between the three finger stiffnesses, 0.34Nm^2 averaged the highest real grasp success rate, 85% with the palm and 72% without the palm. Lowest was 0.29Nm^2, 79% with the palm and 68% without the palm.

Figure 5-b shows that disturbance tolerance in the three axes, X, Y, and Z (defined Section III-C), improved on average by 15% when using the palm. The sum of disturbance tolerance for each controller, $S = X + Y + Z$, averaged as S = 12.6N with the palm and S = 11.2N without the palm. The highest real success rate palm controller (96%, 75° fingertips) tolerated $S = 15.4$N; whereas, the best controller not using the palm (91%, 60° fingertips) only tolerated $S = 11.1$N. Stability increased with fingertip angle, 45° tolerated $S = 9.1$N on average, up to 90° tolerating $S = 13.4$N. Stability was similar across finger stiffnesses, with $EI = [0.29, 0.34, 0.40]$Nm^2 tolerating $S = [11.6, 11.5, 12.7]$N.



(a) Grasping success rate without the palm (blue) and with the palm (red)



(b) Disturbance tolerance without the palm (blue) and with the palm (red)

Fig. 5. Comparing grasp success and grasp stability with and without the palm over the 24 tests. Each bar presents the average over the three finger rigidities, with the range given by the black markers atop each bar. Simulation (sim.) results use cross-hatched bars, real results use solid bars.

VII. DISCUSSION

The best real world grasping success rate of 96% was achieved using the palm, compared to the best of 91% without the palm. The results clearly show the palm improved grasping: simulation success rates were higher, real world success rates were higher, disturbance tolerance improved along all three axes, and sim2real transfer had a smaller gap.

Policy behaviour was similar with and without the palm. In both cases, fingertips aimed to slip underneath objects using a combination of tilt and constriction. Once objects were in grasp force feedback was used to maintain stability whilst objects were lifted, as well as detect slips, which were often recovered. Figure 6 demonstrates grasps with and without the palm, where similarities can be seen in finger position. However, objects grasped without the palm were visibly less constrained and moved more in grasp. Policies that used the palm avoided pushing objects into the table by initially holding back the palm until finger contact. The palm proceeded cautiously and often retreated immediately after first contact, then tailoring the grasp to the object size with reduced motions that would not push it out of grasp.

Figure 7 illustrates the most common failure modes. Loop failure occurred when policies continuously output ineffective actions, such as closing the fingers even though they were already at their limit. This arose from insufficient

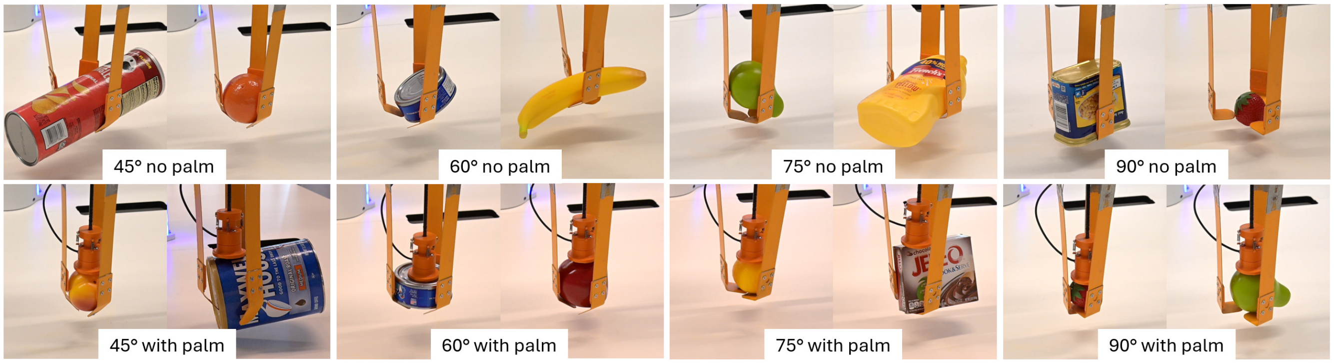


Fig. 6. Successful grasps achieved with and without the palm, across all four fingertip angles, on the YCB benchmark food [13] objects used for testing.

Fig. 7. The four most common failure modes observed during testing.

sensor data or states outside the training distribution, such as when grasping objects like the strawberry which were smaller than any training objects, which may fail to cause enough finger bending to be detected. This failure mode did not occur when using the palm, which provided an alternate means to detect objects and diversified the state space with an additional sensor. Tilt failure, where objects overbalanced when lifted up, was also rare when using the palm, as objects were better caged and less able to escape. Drop failure, where objects slipped through fingers and were not picked back up, was more common using the palm, which did sometimes push objects out of grasp. Roll failure was equally common with and without the palm. This almost exclusively occurred with 90° fingertips, which pointed upwards when the fingers tilted, and frequently rolled spherical objects out of grasp.

The palm improved grasp reliability and stability for two reasons. Firstly, the use of palm improved grasp quality. By pressing the object into the fingers, it completed the cage to prevent tilt failures, and also repositioned objects to ensure consistent contact with all the fingers, to increase frictional forces. Average finger forces were increased by using the palm, 1.43N compared to 1.28N, in addition to the average palm force of 1.79N. The 28% improvement in Y disturbance stability shows this extra friction was important in stabilising grasps on objects that could not be caged. Secondly, the palm improved grasp strategy by contributing an additional force sensor. Force interactions between the fingers and palm provided richer grasp information and more diverse states. This improved learning and hence simulation success rates. It also improved generalisation and sim2real transfer, preventing loop failures and being observed to provide greater capacity for error detection and recovery.

The palm did have limitations. Whilst the palm increased the maximum force tolerated against the Z disturbance, it also exerted an initial force. When comparing the change between the maximum and initial force, grasps using the palm tolerated 16% less on average. Hence, palm grasps may have less potential to grasp the heaviest objects. The palm requires a sufficient holding force from the fingers to become effective, as shown by the increased cases of drop failure. A limitation of using the palm to measure Z stability was that grasp geometry changed during the measurement and grasps without the palm then may have benefited from the presence and caging effect of the palm during the measurement.

Fingertip angle was shown to play an important role. The simulation results correctly indicated best performance with the palm at 75° and without the palm at 60°. In fact, learning was so effective that these were the only two cases of improved average performance in real life. This was likely also due to the non-identical simulation and real life test sets having unequal difficulties for different policies. Sim2real transfer was excellent for 60° and 75°, only 3% difference on average. However, the gap widened to 18% for 45° and 90° fingertips. The observed reason was higher incidence of loop failures and drop failures for 45°, and large number of roll failures for 90°. Loop failures indicate poor generalisation to objects smaller than seen at training time, and many drop failures were related to this too, or to insufficient vertical holding force from the lower angle. Roll failures showed lack of generalisation to irregular objects, as the perfectly regular training objects would not unbalance when lifted up by 90° fingertips, whereas irregular real objects did.

The main limitation of the reinforcement learning-based approach was variability in sim2real transfer, as illustrated by the generalisation issues with 45° and 90° fingertips, demonstrating why our real world validation was essential. However, in every fingertip angle case except 45° without the palm, one of the three tested policies did achieve good transfer no more than 5% below the simulation results (see the range atop each bar in Figure 5-a), showing that good learning and transfer was reasonably reliable.

Comparing between the three finger stiffnesses, average success rate varied by 7% and stability by 11%. 0.34Nm^2

had the highest success rates in simulation and the real world, which may indicate a good balance between adapting to objects whilst maintaining a firm grasp. 0.40Nm^2 was the most stable, as expected since it presented the strongest cage. Controllers learned to grasp with average finger forces which varied only by 7%, despite stiffness varying by 40%, suggesting that frictional forces were more influential for stability than stiffness.

VIII. Conclusion

We investigated the efficacy of using a movable palm, which actively exerts downwards force on objects, for caging inspired grasping. We considered 24 design permutations, varying usage of the palm, fingertip angle, and finger stiffness. A reinforcement learning approach was applied to train controllers for every permutation, before comparison of simulated and real world grasping success rate. The best real world success rate of 96% on 18 YCB food objects was achieved with the palm, compared to 91% without the palm. The palm improved simulated grasp success rates by an average of 11%, real grasp success rate by 13%, and reduced the sim2real gap from 13% to 8%. We measured grasp stability by applying disturbances along three axes, demonstrating a 15% average increase in maximum force tolerated when using the palm. Grasping with the palm had highest success rate with a 75° fingertip angle, whilst without the palm was best using 60° fingertips, and a 45° angle was worst in both cases. Finger stiffness had less effect on success rate and stability than fingertip angle, with the most rigid fingers ($EI = 0.40\text{Nm}^2$) the most stable, and higher grasping success on average being achieved with $EI = 0.34\text{Nm}^2$.

Future work will consider use of our reinforcement learning-based approach for exploring design features such as number of fingers, fingertip actuation, or palm design.

References

[1] RightHand Robotics, "RightPick," www.righthandrobotics.com, accessed Feb. 29, 2024.

[2] E. Ardissone, S. Ulrich, and A. Kirchheim, "Design and evaluation of an automatic decision system for gripper selection in order picking," *Logistics Journal: Proceedings*, vol. 2023, no. 1, 2023.

[3] F. Angelini *et al.*, "SoftHandler: An integrated soft robotic system for handling heterogeneous objects," *IEEE Robotics and Automation Magazine*, vol. 27, no. 3, pp. 55–72, 2020.

[4] L. Kang, J.-T. Seo, S.-H. Kim, W.-J. Kim, and B.-J. Yi, "Design and implementation of a multi-function gripper for grasping general objects," *Applied Sciences*, vol. 9, no. 24, 2019.

[5] G. Zhong, Y. Hou, and W. Dou, "A soft pneumatic dexterous gripper with convertible grasping modes," *International Journal of Mechanical Sciences*, vol. 153-154, pp. 445–456, 2019.

[6] Y. Deng *et al.*, "Deep reinforcement learning for robotic pushing and picking in cluttered environment," in *2019 IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 619–626.

[7] A. Pagoli, F. Chapelle, J. A. Corrales, Y. Mezouar, and Y. Lapusta, "A soft robotic gripper with an active palm and reconfigurable fingers for fully dexterous in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7706–7713, 2021.

[8] Y. Li *et al.*, "A highly adaptive robotic gripper palm with tactile sensing," in *2022 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2022, pp. 130–135.

[9] L. Beddow, H. Wurdemann, and D. Kanoulas, "Reinforcement learning grasping with force feedback from modeling of compliant fingers," *IEEE/ASME Transactions on Mechatronics*, 2024, in press.

[10] C. Schaff, D. Yunis, A. Chakrabarti, and M. R. Walter, "Jointly learning to construct and control agents using deep reinforcement learning," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 9798–9805.

[11] D. Ha, "Reinforcement learning for improving agent design," *Artificial Life*, vol. 25, no. 4, pp. 352–365, 11 2019.

[12] A. Belmonte-Baeza, J. Lee, G. Valsecchi, and M. Hutter, "Meta reinforcement learning for optimal design of legged robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 12 134–12 141, 2022.

[13] B. Calli *et al.*, "The YCB object and model set: Towards common benchmarks for manipulation research," in *2015 International Conference on Advanced Robotics (ICAR)*, 2015, pp. 510–517.

[14] S. Makita and W. Wan, "A survey of robotic caging and its applications," *Advanced Robotics*, vol. 31, no. 19-20, pp. 1071–1085, 2017.

[15] X. Li, Y. Qian, R. Li, X. Niu, and H. Qiao, "Robust form-closure grasp planning for 4-pin gripper using learning-based attractive region in environment," *Neurocomputing*, vol. 384, pp. 268–281, 2020.

[16] J. Su, C. Liu, and Y. Meng, "Immobilizing caging grasps of convex polyhedrons with a four-pin gripper," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7683–7690, 2021.

[17] S. B. Backus and A. M. Dollar, "An adaptive three-fingered prismatic gripper with passive rotational joints," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 668–675, 2016.

[18] C. B. Teeple, G. R. Kim, M. A. Graule, and R. J. Wood, "An active palm enhances dexterity of soft robotic in-hand manipulation," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 11 790–11 796.

[19] N. K. Brown, A. P. Garland, G. M. Fadel, and G. Li, "Deep reinforcement learning for engineering design through topology optimization of elementally discretized design domains," *Materials & Design*, vol. 218, p. 110672, 2022.

[20] F. Dworschak, S. Dietze, M. Wittmann, B. Schleich, and S. Wartzack, "Reinforcement learning for engineering design automation," *Advanced Engineering Informatics*, vol. 52, p. 101612, 2022.

[21] J. P. Chhabra and G. P. Warn, "A method for model selection using reinforcement learning when viewing design as a sequential decision process," *Structural and Multidisciplinary Optimization*, vol. 59, pp. 1521–1542, 2019.

[22] M. E. Ororbia and G. P. Warn, "Design synthesis through a markov decision process and reinforcement learning framework," *Journal of Computing and Information Science in Engineering*, vol. 22, no. 2, p. 021002, 07 2021.

[23] K. Yonekura and H. Hattori, "Framework for design optimization using deep reinforcement learning," *Structural and Multidisciplinary Optimization*, vol. 60, pp. 1709–1713, 2019.

[24] R. Newbury *et al.*, "Deep learning approaches to grasp synthesis: A review," *IEEE Transactions on Robotics*, pp. 1–22, 2023.

[25] A. Lobbezoo, Y. Qian, and H.-J. Kwon, "Reinforcement learning for pick and place operations in robotics: A survey," *Robotics*, vol. 10, no. 3, 2021.

[26] H. Merzić, M. Bogdanović, D. Kappler, L. Righetti, and J. Bohg, "Leveraging contact forces for learning to grasp," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 3615–3621.

[27] Y. Chebotar, K. Hausman, O. Kroemer, G. S. Sukhatme, and S. Schaal, "Generalizing regrasping with supervised policy learning," in *2016 International Symposium on Experimental Robotics*. Springer, 2017, pp. 622–632.

[28] V. Kumar, T. Hermans, D. Fox, S. Birchfield, and J. Tremblay, "Contextual reinforcement learning of visuo-tactile multi-fingered grasping policies," in *NeurIPS: Robot learning workshop*, 2019.

[29] H. Liang *et al.*, "Multifingered grasping based on multimodal reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1174–1181, 2022.

[30] D. Peckner and I. Bernstein, *Handbook of Stainless Steels*. New York: McGraw-Hill Book Company, 1977.

[31] J. Mahler *et al.*, "Learning ambidextrous robot grasping policies," *Science Robotics*, vol. 4, no. 26, 2019.

[32] D. Kalashnikov *et al.*, "QT-Opt: Scalable deep reinforcement learning for vision-based robotic manipulation," *arXiv:1806.10293*, 2018.

[33] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv:1707.06347*, 2017.