

# Augmentation Enables One-Shot Generalization In Learning From Demonstration for Contact-Rich Manipulation

Xing Li<sup>1,2</sup> Manuel Baum<sup>1,2</sup> Oliver Brock<sup>1,2</sup>

**Abstract**—We introduce a Learning from Demonstration (LfD) approach for contact-rich manipulation tasks, i.e., tasks in which the manipulum’s motion is constrained by contact with the environment. Our approach is motivated by the insight that even a large number of demonstrations will often not contain sufficient information to obtain a general policy for the task. To obtain general policies, our approach *augments* the information contained in a single demonstration. This autonomous augmentation is based on the insight that environmental constraints play a central role in generalization. We validate our approach in real-world experiments with mechanisms with multiple, interdependent articulations, including latch locks, chain locks, and drawers with handles. The extracted policies, obtained from a single *augmented* human demonstration, generalize to different mechanisms of the same type and in varying environmental settings.

## I. INTRODUCTION

Learning from demonstration [1] (LfD) aims to extract general policies from human demonstrations. But do demonstrations necessarily contain the necessary information? We argue that for contact-rich manipulation tasks this is *not* the case. Manipulation tasks are contact-rich if the manipulum’s motion is constrained by features in the environment. Examples include the manipulation of articulated objects, such as drawers or scissors, but also sliding motions between objects, such as a box being pushed across a table.

Demonstrations contain instance-specific information. This information does not generalize and must be discarded or altered for producing good policies [2]. Demonstrations also contain task-general information; this is the information we want to extract and turn into a robot policy [2]. But this information is not always sufficient for several reasons.

First, robust and general policies for contact-rich manipulation must include information about the *environmental constraints* (EC) [3] present in the task [2]. Demonstrations can contain such information as forces acting in directions other than the direction of motion. However, humans are such expert manipulators that they often follow environmental constraints without generating significant force signals for a constraint [4]. This means that demonstrations often lack EC information important for generalization. For example, when a human opens a drawer, they rarely generate forces orthogonal to the direction of motion. This makes it chal-

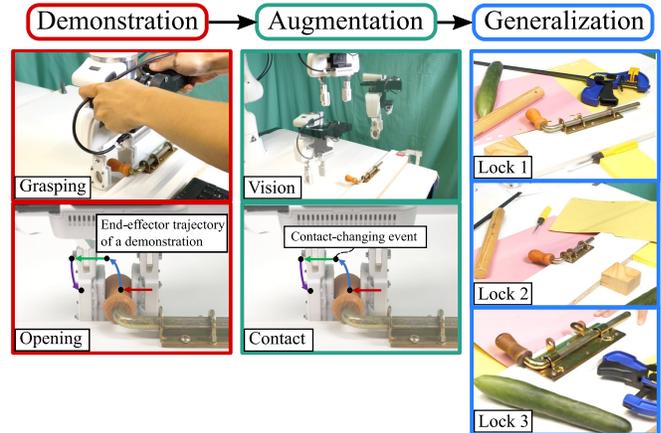


Fig. 1: Our approach to learning from demonstration takes a single human demonstration (left, red) and augments it autonomously with information required to produce a general policy (middle, green). The augmentation process works for complex contact-constrained motions (green, bottom) and also for visual servoing, for example, during grasping (green, top). The resulting policies generalize robustly, here shown for three different locks with changing backgrounds (blue, right).

lenging to distinguish a drawer (EC: prismatic joint) from a free-floating object moving on a linear trajectory.

Second, to perform a demonstration, humans might consider *perceptual features* that are not available to the robot, due to differing perception capabilities. It is not practical to consider all possible perceptual information in LfD. Whatever the selection is, it will likely exclude relevant information, for example about tight clearances and small environmental features not easily distinguishable in the presence of sensor noise. In addition, suitability of features might differ between humans and robots due to their different embodiment and different lower-level control abilities.

Third, humans draw on extensive prior manipulation knowledge. It is not possible to extract this from demonstrations on a single task. Yet, this knowledge might be required to generalize demonstrations into a good policy.

This shows that demonstrations cannot be guaranteed to contain sufficient information for obtaining general and robust policies. In our experience, this is very frequently the case. To address this problem, we propose a method for automatic *augmentation* of demonstrations to fill in missing details. *Augmentation*, a process illustrated in Fig. 1, relies on prior knowledge about environmental constraints [3] and

<sup>1</sup> Robotics and Biology Laboratory, Technische Universität Berlin

<sup>2</sup> Science of Intelligence, Research Cluster of Excellence, Berlin

We gratefully acknowledge funding provided by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC 2002/1 “Science of Intelligence” – project number 390523135.

interactive perception [5] to produce general policies from a single demonstration. Of course, our method is only able to augment certain types of information; further research will be needed to fully close the gap that results from the three problems mentioned above. Still, we show that our method is capable to extract generalizable policies from a single demonstration with a subsequent augmentation phase. It can reliably open different mechanical latches (Fig. 1) after one demonstration on just one latch. Similarly, it can generalize single demonstrations to open different variations of drawers and chain locks.

## II. RELATED WORK

Learning from demonstration aims to teach robots manipulation skills from human demonstrations [6] and is often cast as a supervised-learning problem [7]–[12]. The LfD paradigm commonly assumes that skills can be extracted solely from demonstrations. This assumption seems to hold for tasks that involve little contact, but human demonstrations often do not contain enough force information to extract contact-rich skills [4]. This finding is aligned with our previous work, where we found that humans need to purposely demonstrate high forces to reveal physical constraints [2]. Suomalainen et al. report a similar finding [13]. One approach to increase force information in demonstrations is to impair the teacher, e.g. by blindfolding [14]. We suggest to abstain from such impairment and to let the robot gather the required information itself by augmentation. Note that the idea of augmentation is similar to [15] which augments human demonstrations through interactive perception to disambiguate articulated motion models present in the task. However, the scenario considered in our work differs from [15] in that our approach disambiguates contact-changing events that are required for the successful operation of articulated mechanisms with multi-DoF.

Combining Reinforcement Learning (RL) with LfD can be seen as a form of augmentation. Exploration in these integrated approaches gathers additional data to refine policies learned by LfD with RL-based optimization [16]–[22]. However, RL is often data-hungry due to domain-agnostic exploration strategies. In contrast, our approach is tailored to contact-rich manipulation tasks. It augments demonstrations by exploring contact situations around demonstrated motions. Note that our domain-specific exploration approach, utilizing environmental constraints as priors, can also be applied in RL to achieve more efficient and safe explorations [23].

In contactless tasks, augmentation was applied to improve visual servoing policies [24], [25]. Augmentation allows better convergence, because the robot gathers additional training images from unseen camera perspectives. However, these visual servoing policies are challenged by distractors and changes in the environment, because they perform regression directly on input images. We employ and augment a similar visual servoing strategy, but our vision-based policy incorporates a grasping-specific attention mechanism that only focuses on task-relevant visual features, thereby significantly enhancing the policy’s generalization ability.

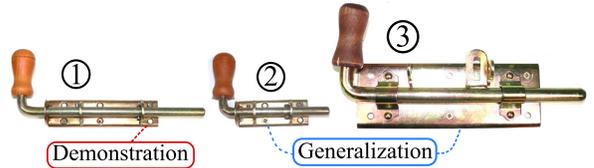


Fig. 2: Locks used in the experiment: The policy derived from a demonstration with lock 1 generalizes to locks 2 and 3. Generalization is possible because all locks contain the same vision- and contact-based environmental constraints, even though the geometric and visual parameters are different.

## III. AUGMENTING DEMONSTRATIONS FOR CONTACT-RICH MANIPULATION

When learning to manipulate complex mechanisms with multiple degrees of freedom (multi-DoF mechanisms), it is especially crucial to identify environmental constraints imposed by the kinematics of manipulated mechanisms. Different mechanisms with the same functionality often share the same environmental constraints. Thereby, policies based on these environmental constraints rather than absolute positions can be transferred to various instances of the same type [2]. This insight is exemplified by the locks shown in Fig. 2. This section describes how to extract general policies from a demonstration and which missing information should be augmented to ensure a successful extraction.

### A. Policy Representation

To manipulate multi-DoF mechanisms, robots must actuate a sequence of mechanical DoF, which can be seen as a hybrid control problem. Different stages of such tasks are governed by different environmental constraints. We mirror this hybrid structure in our control strategy and implement policies as hybrid automata [26] implementing sequences of compliant controllers that are guided by environmental constraints.

To reproduce compliant motions, we employ adaptive compliant controllers from our previous work [2]. These controllers maintain a belief about a currently feasible motion direction  $\vec{m}$  and follow that direction using velocity impedance control. They adaptively track  $\vec{m}$  to follow environmental constraints that change smoothly. If environmental constraints change abruptly, it will trigger a transition to the next controller in the hybrid automaton.

Abrupt changes in mechanical constraints can happen in two ways. Either  $\vec{m}$  is not feasible anymore (*making contact* event) or a spatial direction other than  $\vec{m}$  was formerly constrained and now becomes available for motion (*breaking contact* event). To detect this second kind of event, our controllers exert an additional force  $\vec{f}$  (rendered as a desired velocity) during motion. They can detect if a direction orthogonal to  $\vec{m}$  becomes unconstrained when this force leads to motion. This is an example of interactive perception [5], where forceful interaction induces additional sensory events that are otherwise unavailable.

To summarize, we construct hybrid automata as manipulation policies. To fully instantiate such a hybrid automaton, we need feasible motion directions  $\vec{m}$  and desired force direction  $\vec{f}$  to instantiate adaptive compliant controllers, as well as corresponding contact-changing events as transition functions. Fig. 3 illustrates that the policy extracted from one mechanism is able to generalize to new instances with the same type, as the compliant motions and contact-changing events are identical across various instances.

### B. Instantiation from Demonstration

To fully instantiate our hybrid automaton, we must identify the number of controllers  $K$ , as well as the feasible motion directions  $\vec{m}_{1:K}$  and the force directions  $\vec{f}_{1:K}$ . However, human demonstrations generally will not contain sufficient force information to reliably extract  $\vec{f}_{1:K}$ . We therefore use the initial demonstration’s end-effector positions to extract  $K$  and  $\vec{m}_{1:K}$ . To do so, we first segment the demonstrated position trajectory  $Tr = (p_1, p_2, \dots, p_N)$  into sub-trajectories  $D_1, D_2, \dots, D_K$  using a change point detection algorithm [27]. For each sub-trajectory  $D_k$ , we apply Singular Value Decomposition (SVD) to the recorded positions to get three eigenvectors  $V_k \in \mathbb{R}^{3 \times 3}$ . The first eigenvector  $\vec{v}_{k,1}$  is the estimated motion direction  $\vec{m}_k$ . This way, we obtain a sequence of motion directions that can be used to imitate the compliant motions with adaptive compliant controllers.

To achieve generalization, terminations of these compliant motions should not be dependent on absolute positions but rather on contact-changing events. However, as argued in the introduction, demonstrations are unlikely to always contain the required force directions  $\vec{f}$ , which are crucial to detect contact-changing events. The following subsection will explain how to complete this missing information via augmentation.

### C. Augmentation for Contact-rich Manipulation

The goal of augmentation is to reveal useful force directions in which the robot can exert a force to maintain contact, so that a salient contact event occurs appropriately to trigger a transition. To do so, we search for such a useful force direction  $\vec{f}_k$ . We iterate through a set of hypothesized force directions  $\vec{f}_{k,h}$  and test if they trigger a salient contact event at the desired switching position  $p_k^e$ . Each hypothesis is tested by first moving to the starting position of the segment  $p_k^s$  and then executing the adaptive compliant controller with the motion direction  $\vec{m}_j$  together with the hypothesized  $\vec{f}_{k,h}$ . We discard the hypothesized  $\vec{f}_{k,h}$  if the EE position deviates more than  $\delta = 10 \text{ mm}$  from the demonstrated trajectory ( $\vec{f}_{k,h}$  conflicts with motion direction) or a contact event is triggered before the end-point  $p_k^e$  of the trajectory is reached. Otherwise, the first hypothesis that is not omitted is saved as the desired  $\vec{f}_k$ .

We consider the following five vectors as sensible hypotheses for  $\vec{f}_k$ . Although this set is likely incomplete, it captures several common contact situations in mechanical manipulation tasks.

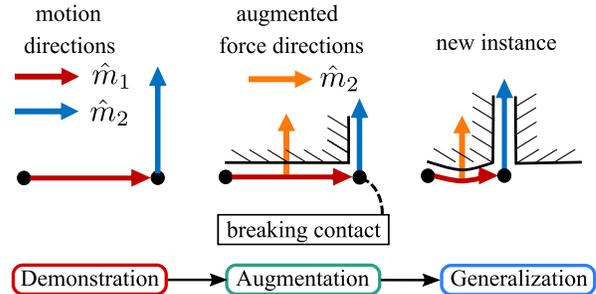


Fig. 3: This 2D scenario illustrates the idea of augmentation for contact-based ECs. We first extract two motion directions from a demonstration (left). It is challenging to extract the underlying contact situations solely from this demonstration. However, with augmentation, we discover that maintaining contact in the future motion direction (orange arrow) will reveal a *breaking contact* event that explains the transition between two demonstrated motion directions (middle). Our policy based on the augmentation results ensures generalization across various instances, as it exploits the contact-based ECs rather than focus on absolute positions (right).

- 1)  $\vec{m}_{k+1}$ : The motion direction of the next segment may become feasible, but may currently be constrained.
- 2)  $\vec{m}_{k-1}$ : The motion direction of segment  $k - 1$  may have been constrained in the current segment, but this constraint may disappear in a relevant configuration.
- 3)  $\vec{v}_{k,3}$ : The direction of the third eigenvector  $\vec{v}_{j,3}$  from SVD (smallest variance). The demonstration might have low variance in this direction due to a physical constraint.
- 4)  $-\vec{v}_{k,3}$ : This is the opposing direction of 3).
- 5)  $\vec{0}$ : Exert zero force. This is a fallback and the default in case no other  $\vec{f}_{k,h}$  is permissible.

Once we have identified a suitable force direction  $\vec{f}_k$ , the robot moves to the next segment’s start position  $p_{k+1}^s$  and searches suitable force directions  $\vec{f}_{k+1, \dots, K}$  in the same way.

When force directions  $\vec{f}_{1:K}$  have been gathered for all  $K$  segments, we construct the switching events for the hybrid automaton. Controllers switch either due to a *making contact* event where the motion direction  $\vec{m}_k$  is not feasible anymore, or when the robot could move into the force direction  $\vec{f}_k$  for at least  $\delta = 10 \text{ mm}$  (*breaking contact* event).

In summary, we extract the feasible motion direction  $\vec{m}$  of these controllers from a demonstration and then complete this with a desired force direction  $\vec{f}$  and contact-changing events through augmentation. We now have the necessary information to instantiate a hybrid automaton. The encoded sequence of compliant motions represents a policy that generalizes across object instances, as we will see in the experimental evaluation.

## IV. AUGMENTING DEMONSTRATION FOR VISION-BASED ENVIRONMENTAL CONSTRAINTS

The idea of augmentation has previously been applied for visual servoing [24], [25]. An additional contribution

of our work is to improve generalization of this approach by introducing a grasping-specific attention mechanism. Our experimental evaluation requires such visual servoing, as the robot must move from a pre-grasp position to a grasp position to enable manipulation. To achieve the pre-grasp poses, we move the end-effector such that a wrist-mounted camera sees a graspable handle in a desired position and orientation.

Visual servoing requires visual features to guide end-effector motions from a wide set of states to a goal state. To be robust, servoing should succeed also from states not visited during a demonstration. Limited generalization can be achieved by using invariant visual features [28], but large deviations from demonstrations still pose a problem.

We propose to use *augmentation* to directly gather camera input from additional camera perspectives to achieve generalization. This idea has been successfully applied in prior work [24], [25]. In these approaches, the robot autonomously gathers training data  $D = \{I_i, {}^G T_i\}_i$  of camera images  $I_i$  and transformations  ${}^G T_i$  from end-effector  $E$  to target grasp pose  $G$ . Then it trains a regression model to predict desired end-effector motion given input images. This approach generalizes well beyond demonstrated poses [25]. However, previous implementations of this idea [24], [25] used CNN-based models on the full input image, which we observed to be susceptible to clutter and changes in the background. To achieve robustness against these issues, we employ a Faster-RCNN [29] based approach that focuses computation on relevant parts in the image.

Our approach follows [25] in its general data collection and computation scheme, but we do not map full input images  $I_i$  to end-effector relative graph poses  ${}^G T_i$ . Instead, our approach (depicted in Fig. 4) divides grasp pose estimation into two subsequential components: 1) a detector that proposes Regions of Interest (ROI) as bounding boxes; 2) an orientation estimator that only focuses on visual features within ROI. For estimating a grasp pose, our network first detects the area of the grasping location as an ROI. The center point  $c = (c_x, c_y)$  of the bounding box is the grasp position in the camera frame. We can then calculate the 3d grasp position  $p = (p_x, p_y, p_z)$  using a depth camera. Then, the network predicts the grasp orientation using the image cropped to the ROI’s bounding box. We consider the tasks where a mechanism is mounted on a surface, and the end-effector is constrained to be perpendicular to the surface. Hence, our network only has to predict the yaw angle  $\Psi \in \mathbb{R}$  in the end-effector frame. We implement this network by extending the framework of Faster-RCNN [29]. In addition to a bounding box regressor and a classifier, we add a layer to estimate  $\Psi$  for each potential object in the image.

Training our network requires a bounding box and a rotation angle  $\Psi$  as labels. Given an input image and a ground-truth grasp pose in end-effector frame  ${}^G T$ , we project a rectangle defined by the volume in the robot’s open gripper into the camera image. The corners of this rectangle correspond to the blue circles in Fig. 4. Then we fit this rectangle which is not axis aligned with an axis-aligned rectangle and use it as the target for the ROI’s bounding box.

The rotation between both rectangles is used as the target for the orientation estimator.

The key advantage of our approach is that the grasp pose estimator concentrates on visual features within bounding boxes, breaking undesired correlations with visual features outside these boxes. This enables the policy to generalize to different objects sharing similar task-relevant visual features and ensures robustness against environmental changes e.g. visual distractors.

## V. EXPERIMENTS AND RESULTS

We test the capabilities of our approach in real-world experiments on contact-rich manipulation tasks with various articulated mechanisms. The experiments are conducted on a 7-DOF Panda arm with a 1-DOF gripper. We attached a RealSense D435i depth camera to the end-effector and an ATI-mini 40 Force/Torque sensor to the wrist. Videos for experiments are available at the: [project page](#).

### A. Opening Latch Locks

In this experiment, our goal is to test the generalization and robustness of the proposed approach for opening latch locks. Given a demonstration with lock 1 (see Fig. 2), our approach augments this demonstration to obtain an EC-based policy and test it on all three locks. As the lock opening task involves two sub-tasks: grasping and opening, we divided this experiment into three parts. We evaluate the grasping and opening sub-tasks separately in the first two parts. Lastly, we demonstrate the overall efficacy of our approach by showing that it allows a robot to grasp and open (full-task) a lock in a dynamic and cluttered environment.

1) *Grasping*: In this experiment, We provide a kinesthetic demonstration of grasping with lock 1. We record around 50 images during the demonstration. Note that there is no significant rotation in the demonstration. We then augment this demonstration by 2500 camera images collected in a cone-shaped volume, pointing towards the demonstrated grasp point. We use augmented data to train our network.

To evaluate the effect of augmentation, we compare our algorithms with the following baselines which do not use augmentation: **SIFT-based visual servoing**: matches SIFT [28] features from the input image to the target image captured in the demonstrated grasp pose, then uses an image-based visual servoing controller that moves the robot towards the grasp pose [30]; **w/o augmentation**: directly trains our network only on the demonstration data without augmentation.

During the test, we place the lock at a randomized location on the table. To test the controller’s reactivity, we change the pose of the lock while the robot approaches the lock. We run 10 grasping trails for each lock. A grasp trial is successful if the robot grasps the knob and lifts it.

The results of the various grasping approaches are summarized in Table I. Among the methods tested, **SIFT-based visual servoing** is able to grasp the knob 3/10 times, where most failures result from mismatched features between the observed and target image. These mismatches are caused by the small size (approximately 23cm) of the lock and the

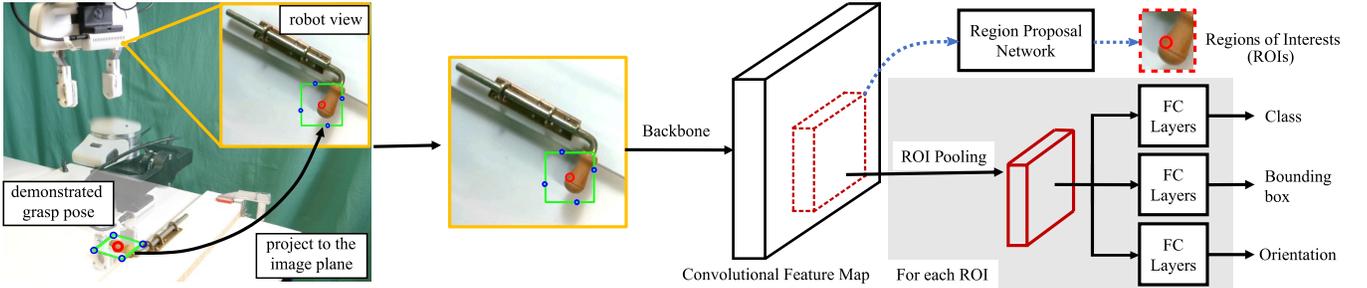


Fig. 4: This figure shows how to generate training data and the architecture of the neural network. We crop an image based on a demonstrated grasp position and back-project it to the image plane to create a bounding box. We then derive the orientation of the object based on the robot’s proprioception. The bounding box, orientation, and the known class are used to train a neural network. This network detects a task-relevant region of interest in the first stage while discarding distractions. It simplifies the training of the pose estimation and improves the robustness against environmental changes e.g. visual distractors.

Method	Lock 1	Lock 2	Lock 3
SIFT-based visual servoing	30%	-	-
W/o augmentation	0%	-	-
Our approachn	100%	100%	100%

TABLE I: Comparison of success rates for grasping locks. Augmentation significantly improves the success rate.

near-homogeneous color of the lock’s body. This approach cannot generalize to grasp lock 2 and lock 3, as these locks differ in size and appearance from lock 1. The other baseline **w/o augmentation** fails to grasp all locks in all trials (0/10). All failure cases happen when the lock is rotated, and the network cannot predict the orientation correctly, as the demonstrated data does not contain data for rotated images. This baseline is not able to apply to other locks due to the limited training data.

Our approach, by contrast, robustly grasps lock 1 in this dynamic environment. The grasping performance is significantly improved with the additional information provided by augmentation. It exemplifies that a demonstration alone is not informative enough to infer a reliable policy for grasping. Augmentation overcomes this problem by efficiently completing the missing information in a demonstration i.e. images from different perspectives. Furthermore, the policy extracted from lock 1 can reliably grasp lock 2 and lock 3 with a 100% success rate, even though the appearance of locks’ backbones differs significantly. We achieve the generalization by only focusing on task-relevant visual features.

2) *Opening*: In the second part of the experiment, we validate our approach to open locks after a grasp was established. We first provide a demonstration to open lock 1. The algorithm segments this demonstration into 4 segments and computes motion directions  $\vec{m}_{1:4}$  as described in Section III-B. For each segment, it also hypothesizes a set of force directions  $\vec{f}_{1:4,h}$  as described in Section III-C. We plot  $\vec{m}_{1:4}$  and  $\vec{f}_{1:4,h}$  in Fig 5. Our approach then validates these hypotheses and constructs a hybrid automaton as described in Section III-C. The resulting hybrid automaton is shown in

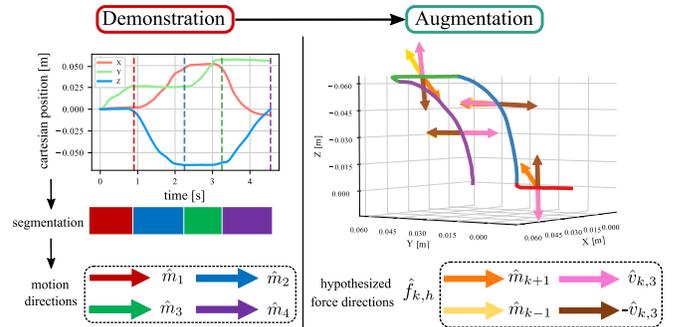


Fig. 5: Based on a EE trajectory of a demonstration with lock 1, we divide it into four distinct segments and compute a motion direction for each segment (left). We then derive a set of force directions based on the motion directions and positions for each segment (right).

Fig. 6.

To assess the impact of augmentation, we use the method developed in [2] as a baseline, which extracts the required motion direction  $\vec{m}$  and force direction  $\vec{f}$  directly from the demonstration. This approach is identical to the in-contact behavior of our approach, except it does not perform augmentation.

We run 10 trials per lock for each method. As clearly seen in Table II, the approach without augmentation performs poorly and only succeeds 5/30 times (17%) on all locks. Specifically, the baseline approach was never able to open lock 1. All failed trials for lock 1 are because the robot fails to pass the pin through the narrow slot, as shown in Fig. 7.

In contrast, our approach reliably passed through the narrow slot as it exerted a force  $\vec{f}$  directed towards that narrow passage. The algorithm could only select this force direction due to augmentation. As a result of augmentation, the learned policy succeeds 30/30 times (100%).

These results show that naive demonstrations cannot guarantee sufficient force information to reason environmental constraints present in the task. Nevertheless, our augmentation phase completes the missing information by actively

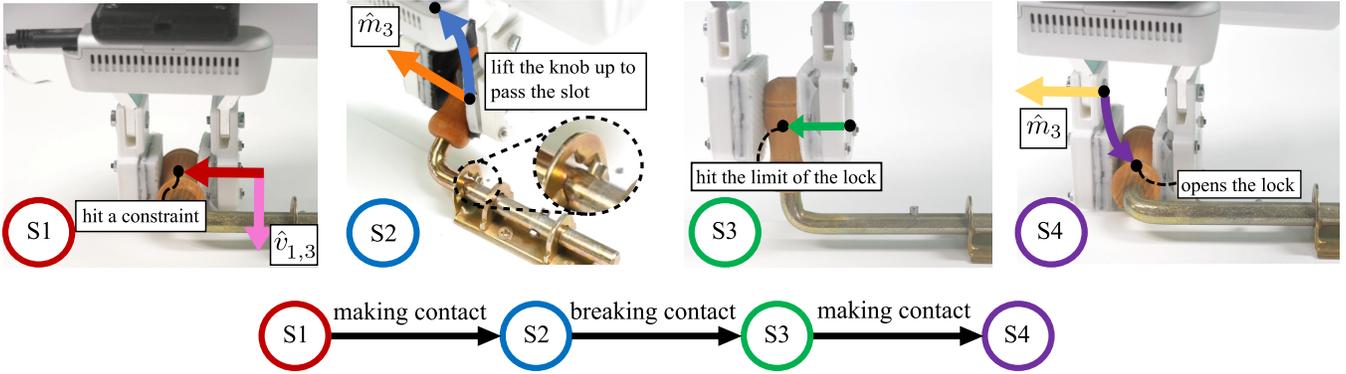


Fig. 6: This figure shows the augmentation results of latch lock 1. The real images correspond to four segmented compliant motions with valid augmented force directions (see Fig. 5). The first compliant motion (S1) involves sliding the knob to the left while keeping contact in the  $\hat{v}_{1,3}$  direction. Then the robot encounters a contact in the motion direction and switches to S2, in which the robot lifts the knob and maintains contact in the future motion direction  $\hat{m}_3$ . This leads the robot to enter the narrow slot and triggers a *breaking contact* event. Next, the robot transitions to the sliding motion (S3) and detects a *making contact* event by reaching the limit of the lock. Note that in S3, none of the hypothesized force directions is valid. Therefore, the augmented force direction for S3 is a zero-vector. Finally, the robot simply lowers the knob and opens the lock.

Method	Lock 1	Lock 2	Lock 3	Avg.
W/o augmentation [2]	0%	30%	20%	17%
Our approach	100%	100%	100%	100%

TABLE II: Comparison of success rates for opening three locks. The performance gap between the baseline approach and our approach shows that augmentation is key to successfully extracting a reliable manipulation policy for physically operating mechanisms.

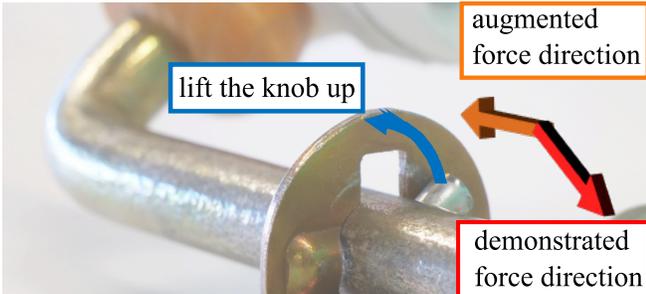


Fig. 7: The robot can successfully traverse the slot by lifting the knob upwards (blue) and applying force along the augmented force direction (orange), which is orthogonal to the slot direction. By contrast, the robot fails to detect the narrow slot and gets stuck when it maintains contact in the demonstrated force direction (red).

uncovering these environmental constraints, leading to the successful extraction of a generalizable policy.

3) *Full-Task Execution*: After individually examining grasping and opening, we test the method’s performance in grasping and opening locks as a whole task. We challenge the robustness of our approach in a dynamic and cluttered environment (see supplementary video). Concretely, we place

Method	Lock 1	Lock 2	Lock 3	Avg.
Our Approach	100%	100%	70%	90%

TABLE III: Success rates of full-task executions including grasping and opening locks. All three failures in lock 3 were due to the robot failing to grasp the knob.

distractors on the table and change the position of the lock when the robot is approaching. We run 10 trials per lock and calculate the success rate. The result is depicted in Table III. Our approach achieves 100% for locks 1 and 2, even in this highly uncertain environment. The success rate of lock 3 decreases to 70% due to the false-negative detection for the knob, which exposes the limitation of the generalization ability of our approach for grasping. We will discuss this limitation in Section VI. These experimental results prove that the augmentation paradigm can yield generalizable and robust policies, even using only a single demonstration.

### B. Opening Drawers

The previous experiments established that augmentation can improve the generalization and robustness of contact-rich manipulation behavior. We will now perform additional experiments to show that augmentation also enables to solve other problems than latch locks. We will begin with opening the two drawers in Fig. 8.

These drawers can only be opened after a revolte handle is rotated, and these drawers differ in the required opening angle for those handles. We record a demonstration on the drawer with a smaller opening angle and train our method as described in Section III. The visual augmentation phase gathers additional images for 3 minutes. The contact augmentation phase discovers two segments. The first segment is to rotate the handle while the robot exerts a contact force

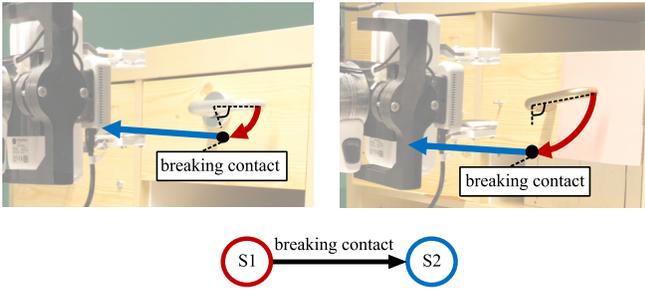


Fig. 8: The hybrid automaton extracted from the left drawer can generalize to the right drawer even if they have different rotation angles. When the robot opens drawers, it first pushes down on the handle and complies with the contact-based EC provided by the revolute joint of the handle (S1). Simultaneously, the robot maintains contact in the pulling direction by employing an augmented force direction. This contact generates a *breaking contact* event that enables the robot to transition to the subsequent complaint pulling action (S2). Overall, our approach reliably accomplishes the task (grasping and opening) with a 100% success rate.

to pull on the drawer, and the second phase is to pull open the drawer. A *breaking contact* event necessarily occurs at the transition between these phases.

To compute the success rate, we run 10 trials to open this drawer in a dynamic environment (i.e. human disturbances) with changes in the background. We achieved a 100% success rate using our approach on both the original as well as the transferred drawer that requires further rotation of the handle. These results demonstrate our method’s robustness in operating various mechanisms from a single augmented demonstration. They also show the ability to yield behavior generalizing to unseen but similar object instances by using environmental constraints as the representation of the manipulation policy.

### C. Opening Chain Locks

We performed additional experiments with chain locks (see Fig. 9) to exemplify that our approach can apply to various mechanisms. For the purpose of simplicity, we 3-D print a larger-sized knob and use it for all chain locks. Again, we obtained a grasping policy similar to the latch lock experiment. We explain here the contact-rich policy in more detail.

The demonstration of the opening chain lock 1 consists of two compliant motions: sliding the knob to the side and pulling it outwards. By augmentation, the robot finds a force direction, which makes it pull on the knob while it slides it along the slit in the lock. This facilitates the detection of a breaking contact event that triggers the transition from sliding to pulling.

Overall, our approach succeeds 10/10 times for chain lock 1 and 9/10 times for chain lock 2 and 3, respectively. We achieve this generalization by augmenting contact-changing events for the contact-rich policy, which exploits the contact-based ECs. The failure in chain lock 2 is due to the

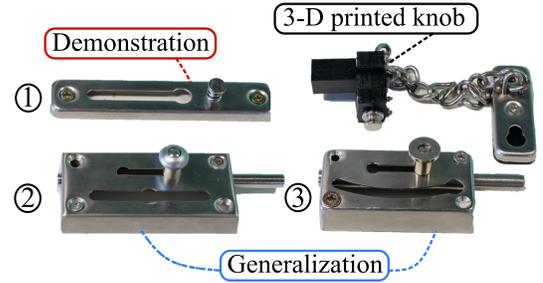


Fig. 9: We conduct experiments on chain locks with a 3D-printed knob. A demonstration of chain lock 1 generalizes to chain lock 2 and 3, achieving a 93% success rate. The transferability of the policy is achieved by exploiting the contact-based ECs.

robot past the slot while sliding the knob, and our control policy cannot recover from such failures. This motivates us to investigate corrective demonstrations to extract failure recovery behaviors to enhance the robustness of the contact-rich control policy [32]. The failure in chain lock 3 is caused that the 3D-printed knob is not fixed with the chain after several trials. Nonetheless, this additional experiment provides further evidence that our environmental-constraints-based policy generalizes well to unseen yet similar instances.

## VI. LIMITATIONS

The proposed approach has two major limitations. First, it is restricted to *prehensile* manipulation with articulated mechanisms, which are *constrained* by the environment. In this context, our approach can augment vision- and contact-based ECs that physically exist in the environment but not task constraints e.g. reference frames [33]. However, we believe that the ideas presented here can form the basis for an overarching LfD approach. Towards this end, we plan to extend our approach by incorporating human corrections into the demonstration process to reveal information that is not augmentable by the robot alone, such as additional types of task constraints and failure recovery strategies [34].

As a second limitation, only relevant for augmentation in the context of visual servoing, manifests itself in the performance decrease for lock 3. The extracted policy for grasping cannot reliably generalize to objects with substantially different visual appearances. This is not surprising for an approach based on visual appearance. However, we believe this limitation can be overcome by incorporating 3D information into the servoing process [35], [36].

## VII. CONCLUSION

We present a novel LfD approach for contact-rich manipulation tasks. Our approach takes as input a single human demonstration, by itself insufficient to produce a general policy. The method then autonomously gathers additional task-relevant information, following the insight that generalization for contact-rich tasks depends on structuring policies based on environmental constraints. Our extensive real-world experiments show the transfer of learned policies

to previously unseen but similar mechanisms that vary in size and appearance. The high success rate across all of our experiments supports this paper’s main insights: Environmental constraints are the appropriate representation for contact-rich manipulation policies. Demonstrations do not contain sufficient information to fully instantiate these policies. Augmentation identifies the missing information and is able to extract general policies from only a single human demonstration.

## REFERENCES

- [1] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, “Recent advances in robot learning from demonstration,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 297–330, 2020.
- [2] X. Li and O. Brock, “Learning from demonstration based on environmental constraints,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10938–10945, 2022.
- [3] C. Eppner, R. Deimel, J. Álvarez-Ruiz, M. Maertens, and O. Brock, “Exploitation of environmental constraints in human and robotic grasping,” *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 1021–1038, 2015.
- [4] M. Suomalainen, Y. Karayiannidis, and V. Kyrki, “A survey of robot manipulation in contact,” *Robotics and Autonomous Systems*, vol. 156, p. 104224, 2022.
- [5] J. Bohg, K. Hausman, B. Sankaran, O. Brock, D. Kragic, S. Schaal, and G. S. Sukhatme, “Interactive perception: Leveraging action in perception and perception in action,” *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1273–1291, 2017.
- [6] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, “Robot programming by demonstration,” in *Springer handbook of robotics*. Springer, 2008, pp. 1371–1394.
- [7] D. A. Pomerleau, “Alvinn: An autonomous land vehicle in a neural network,” *Advances in neural information processing systems*, vol. 1, pp. 305–313, 1988.
- [8] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, “Deep imitation learning for complex manipulation tasks from virtual reality teleoperation,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 5628–5635.
- [9] R. Rahmatizadeh, P. Abolghasemi, L. Bölöni, and S. Levine, “Vision-based multi-task manipulation for inexpensive robots using end-to-end learning from demonstration,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 3758–3765.
- [10] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, “Learning and generalization of motor skills by learning from demonstration,” in *IEEE International Conference on Robotics and Automation*, 2009, pp. 763–768.
- [11] H. Ito, K. Yamamoto, H. Mori, and T. Ogata, “Efficient multitask learning with an embodied predictive model for door opening and entry with whole-body control,” *Science Robotics*, vol. 7, no. 65, p. eaax8177, 2022. [Online]. Available: <https://www.science.org/doi/abs/10.1126/scirobotics.aax8177>
- [12] S. Scherzinger, A. Roennau, and R. Dillmann, “Learning human-inspired force strategies for robotic assembly,” *arXiv preprint arXiv:2303.12440*, 2023.
- [13] M. Suomalainen, F. J. Abu-Dakka, and V. Kyrki, “Imitation learning-based framework for learning 6-d linear compliant motions,” *Autonomous Robots*, vol. 45, no. 3, pp. 389–405, 2021.
- [14] D. Ehlers, M. Suomalainen, J. Lundell, and V. Kyrki, “Imitating human search strategies for assembly,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 7821–7827.
- [15] K. Hausman, S. Niekum, S. Osentoski, and G. S. Sukhatme, “Active articulation model estimation through interactive perception,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 3305–3312.
- [16] A. Nair, B. McGrew, M. Andrychowicz, W. Zaremba, and P. Abbeel, “Overcoming exploration in reinforcement learning with demonstrations,” in *IEEE international conference on robotics and automation (ICRA)*, 2018, pp. 6292–6299.
- [17] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, “Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations,” in *Proceedings of Robotics: Science and Systems (RSS)*, 2018.
- [18] Y. Zhu, Z. Wang, J. Merel, A. Rusu, T. Erez, S. Cabi, S. Tunyasuvunakool, J. Kramár, R. Hadsell, N. de Freitas, and N. Heess, “Reinforcement and imitation learning for diverse visuomotor skills,” in *Proceedings of Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018.
- [19] M. Akbulut, E. Oztop, M. Y. Seker, X. Hh, A. Tekden, and E. Ugur, “Acnmp: Skill transfer and task extrapolation through learning from demonstration and reinforcement learning via representation sharing,” in *2021 PMLR Conference on Robot Learning*, pp. 1896–1907.
- [20] P. Englert and M. Toussaint, “Learning manipulation skills from a single demonstration,” *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 137–154, 2018.
- [21] S. Hoppe, Z. Lou, D. Hennes, and M. Toussaint, “Planning approximate exploration trajectories for model-free reinforcement learning in contact-rich manipulation,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4042–4047, 2019.
- [22] J. Yamada, J. Collins, and I. Posner, “Efficient skill acquisition for complex manipulation tasks in obstructed environments,” *arXiv preprint arXiv:2303.03365*, 2023.
- [23] J. Lundell, R. Krug, E. Schaffernicht, T. Stoyanov, and V. Kyrki, “Safe-to-explore state spaces: Ensuring safe exploration in policy search with hierarchical task optimization,” in *International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 132–138.
- [24] X. Fu, Y. Liu, and Z. Wang, “Active learning-based grasp for accurate industrial manipulation,” *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 4, pp. 1610–1618, 2019.
- [25] E. Johns, “Coarse-to-fine imitation learning: Robot manipulation from a single demonstration,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 4613–4619.
- [26] M. Egerstedt, “Behavior based robotics using hybrid automata,” in *Hybrid Systems: Computation and Control (HSCC): Third International Workshop*. Springer, 2000, pp. 103–116.
- [27] A. Fod, M. J. Mataríć, and O. C. Jenkins, “Automated derivation of primitives for movement classification,” *Autonomous robots*, vol. 12, pp. 39–54, 2002.
- [28] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, pp. 91–110, 2004.
- [29] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *Advances in neural information processing systems*, vol. 28, 2015.
- [30] F. Chaumette and S. Hutchinson, “Visual servo control part I: Basic approaches,” *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.
- [31] S. Ross, G. Gordon, and D. Bagnell, “A reduction of imitation learning and structured prediction to no-regret online learning,” in *International conference on artificial intelligence and statistics (AISTATS)*, 2011, pp. 627–635.
- [32] S. Niekum, S. Osentoski, G. Konidaris, S. Chitta, B. Marthi, and A. G. Barto, “Learning grounded finite-state representations from unstructured demonstrations,” *The International Journal of Robotics Research*, vol. 34, no. 2, pp. 131–157, 2015.
- [33] A. L. P. Ureche, K. Umezawa, Y. Nakamura, and A. Billard, “Task parameterization using continuous constraints extracted from human demonstrations,” *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1458–1471, 2015.
- [34] C. Celemin, R. Pérez-Dattari, E. Chisari, G. Franzese, L. de Souza Rosa, R. Prakash, Z. Ajanović, M. Ferraz, A. Valada, J. Kober, et al., “Interactive imitation learning in robotics: A survey,” *Foundations and Trends in Robotics*, vol. 10, no. 1-2, pp. 1–197, 2022.
- [35] J. Gao, Z. Tao, N. Jaquier, and T. Asfour, “K-VIL: Keypoints-based visual imitation learning,” *IEEE Transactions on Robotics*, pp. 1–21, 2023.
- [36] B. Eisner, H. Zhang, and D. Held, “FlowBot3D: Learning 3D articulation flow to manipulate articulated objects,” in *Robotics: Science and Systems (RSS)*, 2022.