Survey of Learning Approaches for Robotic In-Hand Manipulation

Abraham Itzhak Weinberg[1], Alon Shirizly[2], Osher Azulay[3] and Avishai Sintov[3]

***Abstract*—Human dexterity is an invaluable capability for precise manipulation of objects in complex tasks. The capability of robots to similarly grasp and perform in-hand manipulation of objects is critical for their use in the ever changing human environment, and for their ability to replace manpower. In recent decades, significant effort has been put in order to enable in-hand manipulation capabilities to robotic systems. Initial robotic manipulators followed carefully programmed paths, while later attempts provided a solution based on analytical modeling of motion and contact. However, these have failed to provide practical solutions due to inability to cope with complex environments and uncertainties. Therefore, the effort has shifted to learning-based approaches where data is collected from the real world or through a simulation, during repeated attempts to complete various tasks. The vast majority of learning approaches focused on learning data-based models that describe the system to some extent or Reinforcement Learning (RL). RL, in particular, has seen growing interest due to the remarkable ability to generate solutions to problems with minimal human guidance. In this survey paper, we track the developments of learning approaches for in-hand manipulations and, explore the challenges and opportunities. This survey is designed both as an introduction for novices in the field with a glossary of terms as well as a guide of novel advances for advanced practitioners.**

***Index Terms*—in-hand manipulation, dexterous manipulation, reinforcement learning, imitation learning**

I. Introduction

Robot in-hand manipulation has long been considered challenging. However, it has undergone rapid development in recent years. With the vast industrial development and increase in demand for domestic usage, significant growth in interest in this field can be predicted. Evidently, we witness an increase in research papers, as shown in Figure 1, along with algorithms for solving versatile tasks. Just to mention a few, research has sought solutions for real-world tasks such as medical procedures [1], assembly in production lines [2], and robotic assistance for sick and disabled users [3]. During the COVID-19 pandemic, there has been a significant need for autonomous and complex robot manipulators [4]. In this study, we survey various in-hand manipulation tasks of robotic hands and advanced learning approaches for achieving them. As commonly done by the robotics community and as shown in Figure 2, we distinguish between two main categories of robot in-hand manipulation: *dexterous* and *non-dexterous* in-hand manipulations. To the best of our knowledge, we can conclude that the former approach is more prolific in algorithms and the number of published papers. In addition, we divide the types of manipulations into ones that have continuous and non-continuous contacts during execution. The continuous approach has more techniques than the other as it normally uses dexterous robotic hands with higher Degrees-Of-Freedom (DOF) in comparison to non-continuous approaches [5].

Efforts for learning in-hand manipulation can be classified into three subfields: model-based methods, Reinforcement Learning (RL) and Imitation Learning (IL). Model-based methods focus on the supervised learning of the dynamics of a system or state representation. On the other hand, RL provides a reward function that embeds an implicit directive for the system to self-learn an optimal policy for completing a task. Similarly, IL requires a policy to imitate human expert demonstrations. Figure 1 shows the increase in paper publications over the past five years with regard to these three subfields. In this study, we survey in-hand manipulation approaches, tasks and applications that use either of these subfields.

To the best of the author's knowledge, this is the first survey of learning approaches for robotic in-hand manipulation. Hence, this study offers multiple contributions. Papers were classified and grouped into meaningful clusters. The survey can help researchers to efficiently locate relevant research in a desired class and perceive what has already been achieved. Table I provides a summary of prominent state-of-the-art work including key properties. Practitioners can use our survey to estimate the added value of their research and compare it with previous studies. We also explain the relationships between the different subfields to showcase a wide perspective of the topic. In addition, our work can be seen as a survey of survey papers, similar to the references of some other survey papers.

This study adopts a top-down approach. First, we provide a technical overview of in-hand manipulation, including the types of manipulation, hands and sensing modalities (Section II). This overview provides an understanding of the relevant hardware, manipulations and common terms to be used later. Next, we survey the subject from a high-level perspective and later zoom into

[1] AI-WEINBERG AI Experts, Israel.
[2] Department of Mechanical Engineering, Technion - Israel Institute of Technology, Israel.
[3] School of Mechanical Engineering, Tel-Aviv University, Israel.

The imitation learning is what we are looking for, as we have motions

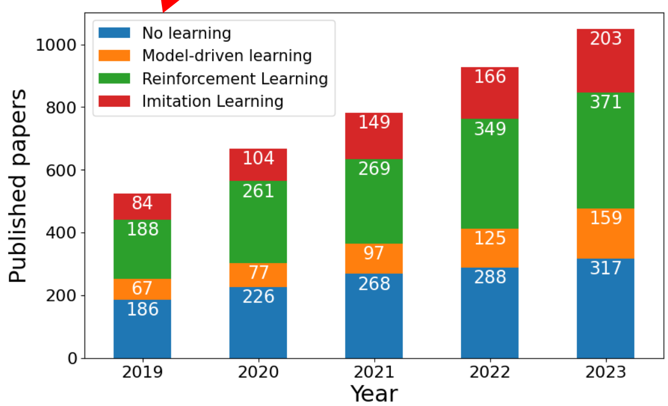

Fig. 1. Statistics on paper publications which addressed or mentioned robotic in-hand manipulations over the past five years in three learning sub-fields: Model-driven learning, Reinforcement Learning (RL) and Imitation Learning (IL), along with papers that do not use any learning method. The search is based on Google Scholar and may include non peer-review publications.

more detailed sub-topics. In addition, we discuss popular tasks in each field. For each task, we often found several approaches while comparing the benefits of one over the other. Finally, we provide insights into future challenges and open problems that should be addressed by the robotics community.

II. Overview on in-hand manipulations

Robotic in-hand manipulation involves physical interaction between a robotic end-effector, an object and often with the environment. The properties of an end-effector define its ability to manipulate the object including: sensory perception, number of DOF, kinematics and friction. In this section, we provide an overview of various types of in-hand manipulations and the robotic hand types that are capable to exert them. In addition, we discuss the common perception and control methods used in these manipulations.

A. Dexterous and non-dexterous manipulation

The conventional paradigm is to distinguish between dexterous and non-dexterous hands. Generally, dexterous manipulation is the cooperation of multiple robot arms or fingers to manipulate an object [66]. Dexterous in-hand manipulation is, therefore, the manipulation of an object in the hand by using its own mechanics [67]. Naturally, dexterous in-hand manipulation requires a high number of DOF and includes, in most cases, anthropomorphic hands. Contrary to dexterous hands, non-dexterous ones have a low number of DOF and, thus, do not have an intrinsic capability to manipulate objects by themselves and require some extrinsic involvement.

B. Types of in-hand manipulations

We now introduce the types of in-hand manipulations commonly addressed in the literature and distinguish between those that maintain and do not maintain continuous contact with the object. These types can be referred to as both dexterous and non-dexterous manipulations as will be discussed later.

1) In-hand manipulations that maintain continuous contact: Initiating object motion within a robotic hand poses some risk of losing control and potentially dropping it. Hence, the majority of in-hand manipulations perform the motion while maintaining sufficient contact with the object, as it is the safest approach. Nevertheless, a prominent condition for a successful manipulation is that grasp stability is guaranteed throughout the motion. Following are the key in-hand manipulation types that maintain contact.

The followings are for cont. grasping

- *Rolling.* Rolling manipulation is the ability to rotate an object within the hand through rolling contact. This is caused by finger motion during contact while the object is held against another part of the hand, usually static, such as another finger or palm [70]. Rolling manipulation is limited to objects of certain geometries and is most feasible with round ones.
- *Pivoting.* Pivoting is the reorientation of an object between two fingers with respect to the hand [71]. The rotation point is commonly the pinching point of the fingers and the reorientation is conducted to some desired angle. The pivoting operation can be done by utilizing gravity [72], initiating external contact [73] or generating dynamic motions of the robotic arm [18, 74].
- *Sliding.* In in-hand sliding manipulation, controlled slip is initiated in order to vary the relative position of the object with regards to the hand [75]. The object slides along the links of the hand to a desired position due to forces from the fingers or external forces [76].
- *In-Grasp.* While pivoting, sliding and rolling aim to change either position or orientation of the object, in-grasp exploits kinematic redundancy of the hand to vary both the position and orientation of the object [77]. During manipulation, finger contacts and a stable grasp are maintained. However, sliding or contact rolling may occur [14].
- *Finger Gaiting.* An approach analogous to gait where the hand's DOF are exploited to switch between contact locations while maintaining a force-closure grasp [78–80]. Hence, the approach is commonly performed quasi-statically where the motion is performed relatively slow to reduce dynamic effects. Finger gaiting may be considered quite wasteful, as it requires sufficiently many DOFs to manipulate the grasped object between two grasp configurations while maintaining stable grasps.

followings are for non-cont. grasping

2) In-hand manipulations that do not maintain continuous contact:

- *Pick and place.* While not usually considered as in-hand manipulation, pick-and-place is worth mentioning since it is the most common. The method designates a work area near the robotic arm, where the grasped object can be placed in a controlled manner

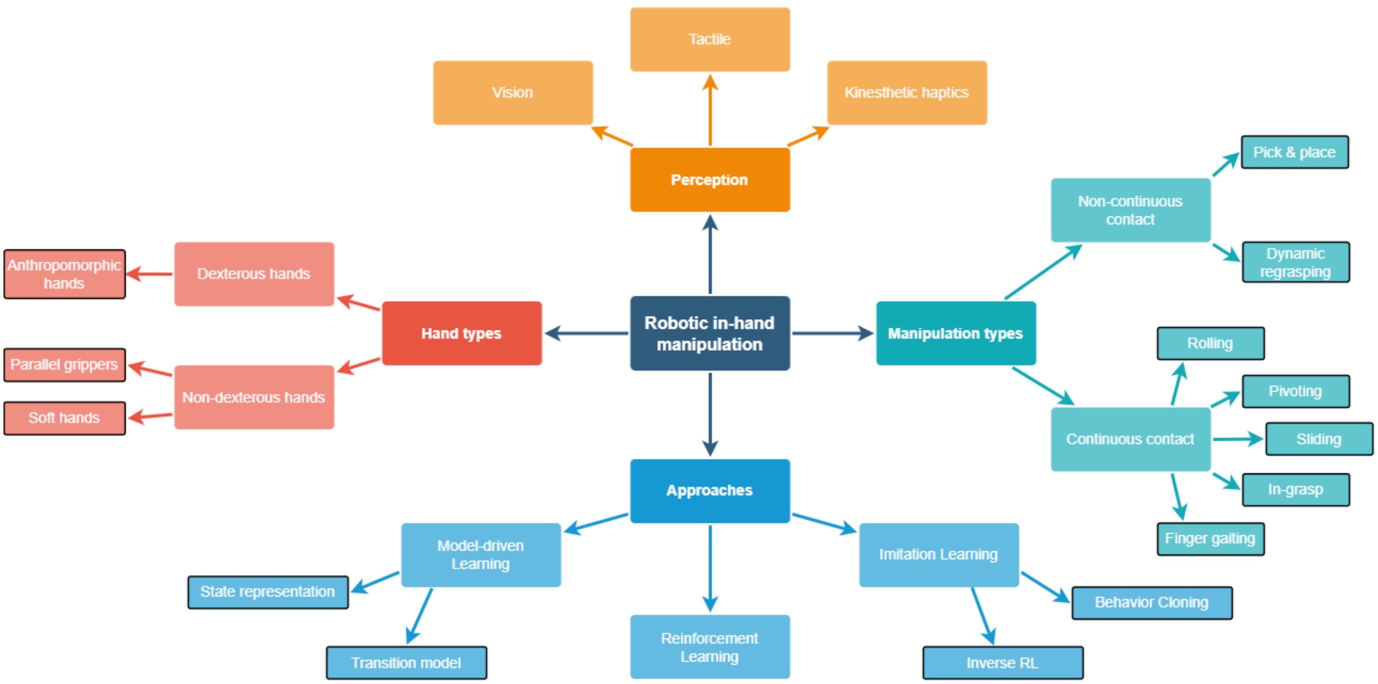

Fig. 2. Taxonomy of robotic in-hand manipulation.

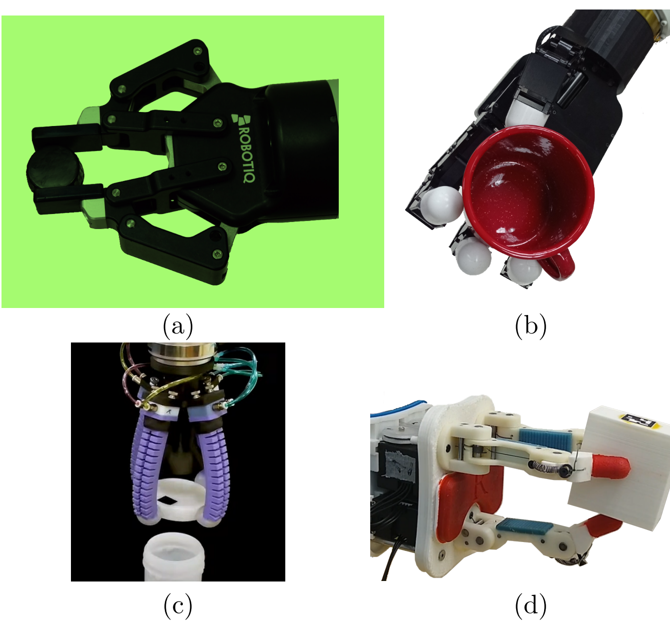

Fig. 3. Various dexterous and non-dexterous hands. (a) Parallel jaw gripper model 2F-85 by Robotiq. (b) The four-finger Allegro hand with 16 DOF. (c) A four finger soft hand operated by pneumatic bending actuators [68]. (d) Underactuated hand model-O from the Yale OpenHand project [69].

and then picked up again at a new grasp configuration [81, 82]. This approach consumes valuable production time and occupies a substantial work area.

- *Dynamic Regrasping.* In this approach, the robot initiates an intended loss of grasp stability through the set of dynamical motions. In most cases, the object is thrown or released into midair and later caught in different grasp configurations. Hence, the hand loses contact (fully or partly) with the object and regains contact by catching it at the final contact points [83]. Such a method has the advantage of fast manipulation and may require a low number of DOF. However, in contrast to manipulations that maintain contact, the success rate for dynamic regrasping may be lower as object stability is not maintained throughout the motion.

C. *In-hand manipulation with Non-dexterous hands*

1) *Parallel grippers:* The most common and ubiquitous non-dexterous robotic hand is the *parallel or jaw gripper* seen in Figure 3a. Parallel grippers are widely used due to their simplicity, durability and low cost. They can precisely grasp almost any object of the same scale and, therefore, are ubiquitous in industrial applications of material handling [84]. Parallel grippers normally have only one DOF for opening and closing the jaws. Hence, they do not have independent in-hand manipulation capabilities. Consequently, solutions for in-hand manipulation with parallel grippers often involve the discrete manipulation approach of *pick-and-place* [82]. In pick-and-place, the object is placed on a surface and picked up again in a different grasp configuration [85]. However, the picking and placing can be slow and demands a large surface area around the robot. Hence, approaches for in-hand manipulation with parallel grippers, that do not involve picking and placing, are also divided to *extrinsic and intrinsic dexterity* [77, 86]. The former compensates for the lack of gripper DOF and involves actions of the entire robotic arm for either pushing the object against an obstacle [73, 87] or performing dynamic manipulation. For instance, pivoting can be done by intrinsic slippage control

our base story

TABLE I
SUMMARY OF STATE-OF-THE-ART WORK ON LEARNING APPROACHES FOR IN-HAND MANIPULATION (IN ALPHABETICAL ORDER)

Paper		Learning type			Manipulation type	Dex./Non-Dex. hand	Rigid/Soft hand	Perception		
		Model	RL	IL				Vision	Tactile	Kines.
Allshire et al.	[6]		✓		In-Grasp	Dex.	Rigid	✓		✓
Anatova et al.	[7]		✓		Pivoting	Non-Dex.	Rigid	✓		
Andrychowicz et al.	[8, 9]		✓		In-Grasp	Dex.	Rigid	✓		✓
Arunachalam et al.	[10, 11]			✓	Mix	Dex.	Rigid	✓		
Azulay et al.	[12]	✓			In-grasp	Non-Dex.	Soft		✓	✓
Azulay et al.	[13]		✓		In-Grasp	Non-Dex.	Soft			✓
Calli et al.	[14]	✓			In-Grasp	Non-Dex	Soft	✓		✓
Chen et al.	[15, 16]		✓		Mix	Dex.	Rigid	✓		
Chen et al.	[17]		✓		F. Gait	Dex.	Rigid	✓		✓
Cruciani and Smith	[18, 19]		✓		Pivoting	Non-Dex.	Rigid	✓		
Deng and Zhang	[20]			✓	Mix	Dex.	Rigid	✓		✓
Dimou et al.	[21]	✓			In-Grasp	Dex.	Rigid	✓		
Falco et al.	[22]		✓		In-Grasp	Dex.	Soft	✓	✓	
Funabashi et al.	[23–25]		✓		Rolling	Dex.	Rigid		✓	✓
Garcia-Hernando et al.	[26]			✓	Mix	Dex	Rigid	✓		
Gupta et al.	[27]		✓		Mix	Dex.	Rigid	✓		
Handa et al.	[28]		✓		Mix	Dex.	Rigid	✓		
Huang et al.	[29]		✓		F. Gait	Dex.	Rigid	✓		
Jain et al.	[30]		✓	✓	In-Grasp	Dex.	Rigid	✓	✓	✓
Khandate et al.	[31, 32]		✓		F. Gait	Dex.	Rigid		✓	✓
Kimmel et al.	[33]	✓			In-Grasp	Non-Dex.	Soft	✓		✓
Korthals et al.	[34]		✓		Mix	Dex.	Rigid	✓	✓	✓
Kumar et al.	[35]			✓	Mix	Dex.	Rigid		✓	✓
Li et al.	[36]			✓	In-grasp	Dex.	Rigid	✓	✓	✓
Li et al.	[37]	✓	✓		In-Grasp	Dex.	Rigid	✓		✓
Luo et al.	[38]	✓			Mix	Dex.	Rigid	✓		✓
Melnik et al.	[39]		✓		Mix	Dex.	Rigid	✓	✓	✓
Morgan et al.	[40]	✓			In-Grasp	Non-Dex.	Soft	✓		
Morgan et al.	[41]	✓	✓		Mix	Non-Dex.	Soft	✓		
Nagabandi et al.	[42]	✓	✓		Mix	Dex.	Rigid	✓		
Orbik et al.	[43]			✓	Mix	Dex.	Rigid	✓		✓
Pitz et al.	[44]		✓		In-Grasp	Dex.	Rigid			✓
Qi et al.	[45]		✓		F. Gait	Dex.	Rigid			✓
Qi et al.	[46]		✓		F. Gait	Dex.	Rigid	✓	✓	✓
Radosavovic et al.	[47]			✓	Mix	Dex.	Rigid	✓		
Rajeswaran et al.	[48]			✓	Mix	Dex.	Rigid	✓		✓
Sievers et al.	[49]		✓		In-Grasp	Dex.	Rigid			✓
Sintov et al.	[50, 51]	✓			In-Grasp	Non-Dex.	Soft	✓		✓
Solak et al.	[52]			✓	In-Grasp	Dex.	Rigid			✓
Solak et al.	[53]			✓	In-Grasp	Dex.	Rigid		✓	✓
Srinivasan et al.	[54]		✓		In-Grasp	Dex.	Rigid	✓		✓
Tao et al.	[55]		✓		Mix	Dex.	Rigid	✓		
Toskov et al.	[56]	✓			Pivoting	Non-Dex.	Rigid		✓	
Toledo et al.	[57]		✓		Pivoting	Non-Dex.	Rigid	✓		
Van Hoof et al.	[58]		✓		In-Grasp	Non-Dex.	Soft		✓	
Veiga et al.	[59]		✓		In-Grasp	Dex.	Rigid		✓	
Wang et al.	[60]	✓			Pivoting	Non-Dex.	Rigid			✓
Wei et al.	[61]			✓	Sliding	Dex.	Rigid	✓		
Yang et al.	[62]		✓		Mix	Dex.	Rigid		✓	
Yin et al.	[63]		✓		Mix	Dex.	Rigid		✓	✓
Yuan et al.	[64]			✓	Rolling	Non-Dex.	Rigid	✓		✓
Yuan et al.	[65]		✓		Mix	Dex	Rigid	✓	✓	✓

or extrinsic dynamic manipulation of the arm [19, 72, 74]. Slippage control leverages gravity and tunes finger contact force of the parallel gripper [88]. Costanzo [89] exploited a dual-arm system and tactile feedback to allow controlled slippage between the object and parallel grippers. The work of Shi et al. [90] controlled the force distribution of a pinch grasp to predict sliding directions. Similarly, Chen et al. [91] controlled the sliding velocity of an object grasped by a parallel gripper.

In intrinsic manipulation, the available DOF of the gripper are exploited for manipulating the grasped object [92]. While jaw grippers have only one DOF, some work has been done to augment their intrinsic manipulation capabilities. Seminal work by Nagata [93] proposed six gripper mechanisms with an additional one DOF at the tip, each having the ability to either rotate or slide an object in some direction. Similarly, a passively rotating mechanism was integrated into the fingers of the gripper allowing the object to rotate between the fingers by gravity [94]. In [95], a jaw gripper tip was augmented with a two DOF transmission mechanism to re-orient and translate randomly placed screws. In [96], linear actuation was added along each of the two fingers to enable translation and twist of a grasped object. Similarly, a rolling mechanism was added to the gripper in [97] in order to manipulate a flat cable. In-hand manipulation was also enabled for a minimal underactuated gripper by employing an active conveyor surface on one finger [98]. In [99], a pneumatic braking mechanism was included in a parallel gripper in order to transition between object free-rotation and fixed phases. The above augmentation methods for parallel grippers are limited to one manipulation direction and yield bulky mechanisms that complicate the hardware. However, a simple vibration mechanism was recently proposed to enable $SE(2)$ sliding motion of a thin object between the jaws of a parallel gripper [100].

2) Soft hands: Soft hands are robotic grippers that are comprised of soft or elastic materials. Due to their soft structure, they usually provide passive compliance upon interaction with the environment [101]. Hence, they can grasp objects of varying sizes and shapes without prior knowledge. A class of soft hands is the *pneumatic-based hands* where stretchable fingers can be inflated to generate a grasp. For instance, RBO Hand 2 is a compliant, underactuated anthropomorphic robotic hand [102]. Each finger of the hand is made of cast silicon wrapped with inelastic fabric. When inflated, the fabric directs the stretch of the fingers toward a compliant grasp. In-hand manipulation with pneumatic-based hands was demonstrated for which heuristic finger gait enabled continuous object rotation (Figure 3c) [68]. Another pneumatic hand with reconfigurable fingers and an active palm was designed to enable in-hand dexterity while maintaining low mechanical complexity [103]. In [104], a soft robotic gripper for grasping various objects by mimicking in-hand manipulation was presented. It consists of three fingers, where each of them contains three air chambers: two side chambers for twisting in two different directions and one middle chamber for grasping. The combination of these air chambers makes it possible to grasp an object and rotate it.

An important class commonly referred as a soft hand is the *Underactuated* or *Compliant* hand [105, 106]. While the links of such a hand are generally rigid, each finger has compliant joints with springs where a tendon wire runs along its length and is connected to an actuator (Figure 3d). Such a structure enables a two or more finger hand to passively adapt to objects of uncertain size and shape through the use of compliance [107]. They can, therefore, provide a stable and robust grasp without tactile sensing or prior planning, and with open-loop control. In addition, due to the low number of actuators, they enable a low-cost and compact design. Recently, open-source hardware was distributed for scientific contributions and can be easily modified and fabricated by 3D printing [69]. Along with good grasping capabilities, precise in-grasp manipulation was shown possible [108]. Using visual servoing along with a linear approximation of the hand kinematics, closed-loop control of a two-finger hand was demonstrated [109] and later used to track paths planned with an optimization-based model-free planner [110]. However, a precise analytical model for such hands is not easy to acquire due to the compliance and inherent fabrication uncertainties. Therefore, data-based models were proposed and to be discussed later on.

D. In-hand manipulation with Dexterous hands

Mason and Salisbury [67] claimed that rigid hands can acquire controllability of an object with at least three fingers of three joints each. Such a hand control is termed dexterous manipulation and the hand is a *dexterous hand*. Naturally, grippers that satisfy this dexterity condition are bio-inspired or anthropomorphic [111] (Figure 3b). Early work on dexterous anthropomorphic hands includes a three-finger and 11-DOF hand [112], the four-finger Utah/MIT hand [113], and later the Barrett and DLR hands [114, 115]. Furthermore, extensive work was done on five-finger anthropomorphic hands. Similar to the DLR hand, The Gifu Hand used 16 built-in servo-motors in the joints [116]. On the other hand, the Robotnaut hand was designed for space usage and included flex shafts for bending the fingers [117]. A hand from Karlsruhe used 13 flexible fluidic actuators for a lightweight design [118]. The UB hand is a five-finger anthropomorphic hand that used elastic hinges to mimic human motion [119].

While recent work on in-hand manipulation with dexterous hands is based on learning approaches, earlier and few recent ones have proposed non-data-driven methods. For instance, the work in [120] proposed a high-speed dynamic regrasping strategy with a multi-fingered hand based on visual feedback of the manipulated object. A different work introduced a planning framework for an anthropomorphic hand to alternate between finger gait and in-grasp manipulations [121]. Recent work used impedance control for stable finger gaiting over various objects with a dexterous multi-finger hand [122].

E. Perception

Humans use both visual feedback and touch perception for interacting with the environment and, in particular, manipulate objects within their hand [123]. Such sensory modules have been widely explored in robotics, both individually or combined.

1) Vision: Different variations of visual perception are used to observe a manipulated object and estimate its pose in real time. The easiest application is the positioning of *fiducial markers* such as ArUcO [124] or AprilTags [125]. These markers provide instant pose recognition of a rigid object without the need for geometry recognition [126]. However, the requirement to apply them on an object prevents spontaneous unplanned interaction with an object. In addition, the markers are required to be continuously visible to the camera. Hence, they are commonly for manipulation of specific known objects, or for prototyping. For instance, fiducial markers were used in visual servoing [127] and hand state representation [50] during in-grasp manipulation of an object with an underactuated hand.

Visual pose estimation, which is based on geometry recognition of the object, is usually based on an RGB (monocular) camera, depth camera or both (RGB-D). With RGB data, much work has been done to regress 2D images to spatial pose of objects [128–130]. Nevertheless, in simpler applications where the object is known, it can be segmented using image processing tools. For instance, a high-speed vision system was used in [120] to track a cylinder thrown and caught by a multi-finger hand. Similarly, a high-speed camera was used to solve a Rubik's cube with a fast multi-finger hand [131]. A work by OpenAI used three RGB cameras to train a model for pose estimation of a cube manipulated by the Shadow hand [9]. Ichnowski et al. [132] presented Dex-NeRF, a novel approach that enables grasping using Neural Radiance Field (NeRF) technique. NeRF receives five-dimensional vectors as an input and can be used for grasping transparent objects. The RGB values are calculated using an Artificial Neural Network (ANN) only after the initial stages.

Contrary to RGB cameras, 3D sensing such as stereo cameras, laser scanners and depth cameras enable direct access to the distance of objects in the environment. RGB-D sensing, in particular, provides an additional point cloud corresponding to the spatial position of objects in view. For instance, an RGB-D camera was used to estimate the pose of objects before and during grasp by a soft hand [133]. Similar work involved a depth camera to demonstrate robust pose estimation of objects grasped and partly occluded by a two-finger underactuated hand [134]. While visual perception can provide accurate pose estimation of a manipulated object, it requires a line of sight. Hence, it cannot function in fully occluded scenes and may be sensitive to partial occlusions. Haptic-based approaches can, therefore, provide an alternative or complementary solution.

2) Haptics: Information from haptic sensors is acquired through direct contact with objects by either tactile sensing or internal sensing of joint actuators known as Kinesthetic (or Proprioception) haptics [135]. Traditionally, tactile refers to information received from touch sensing, while kinesthetic refers to internal information of the hand sensed through movement, force or position of joints and actuators. While kinesthetic haptics can be easier to measure, tactile sensing is the leading haptic-based sensing tool for object recognition and in-hand manipulation. State-of-the-art tactile sensors include force sensors on fingertips, arrays of pressure sensors [136] or high-resolution optical sensors [137–140]. With these sensors, robotic hands can continuously acquire information about the magnitude and direction of contact forces between them and the manipulated object during interaction. An array of pressure sensors was used for servo control of the Shadow hand in in-hand manipulation tasks of deformable objects [141]. Another work used optical sensors to learn in-hand manipulation models [142]. To combine the advantages of haptics and visual perception, some work has been done with both to explore the hand-object interaction during in-grasp manipulation [143].

III. Model-Driven Learning

The establishment of control policies for in-hand manipulation remains challenging regardless of gripper, object or task properties. Various contact models and hand configurations have been used in the literature to develop kinematic and dynamic models for in-hand manipulation as described in previous sections. In order to execute in-hand manipulation tasks with these models, detailed knowledge of the object-hand interaction is required. For most robotics scenarios, however, such information cannot be reasonably estimated using conventional analytical methods since precise object properties are often not a priori known. Model learning offers an alternative to careful analytical modeling and accurate measurements for this type of system, either through robot interaction with the environment or human demonstrations. Learning a model can be done explicitly using various Supervised Learning (SL) techniques, or implicitly by maximizing an objective function. In this section, we will focus on the former technique of supervised learning while the latter will be covered in the following section.

A. *Learning state representation*

Learning a state representation for in-hand robotic manipulations refers to the process of developing a mathematical model that describes the various states of the hand-object system during manipulation. This model can be used to represent the position, orientation, velocity, and other physical properties of an object. Furthermore, the model can be used to predict object response to certain actions.

State representation is an important building block where the object-hand configuration is sufficiently described at any given time. For example, if a robot is trying to roll an object within the hand, it may use some state representation to measure and track the object's pose, and use this information to exert informed actions. In SL, an ANN is commonly used to extract relevant features from data and learn useful features from high-dimensional observation spaces [9, 12, 21, 25, 144]. It is also effective in combining data from multiple sensors or information sources [45], and is often used by robots to merge information from different modalities, such as vision and haptic feedback [9]. Without a compact and meaningful representation of the object-hand state, the robot may struggle to perform successful and efficient manipulations [12].

Haptic perception is commonly used to learn various features of an object in uncertain environments so as to grasp and manipulate it. Such information may include stiffness, texture, temperature variations and surface modeling [145]. Often, haptic perception is used alongside vision to refine initial pose estimation [146]. Contact sensing is the common approach for pose estimation during manipulation [12]. Such sensing has been generally achieved using simple force or pressure sensors [147–149]. As such, Koval et al. [150] used contact sensors and particle filtering to estimate the pose of an object during contact manipulation. In recent years, optical sensor arrays have become more common due to advancements in fabrication abilities and due to their effectiveness in covering large contact areas [136]. The softness of the sensing surface allows the detection of contact regions as it deforms while complying with the surface of the object. Changes in images captured by an internal camera during contact are analyzed.

Several works have used data from optical-based tactile sensors and advanced deep-learning networks to estimate the relative pose of an object during contact manipulation. For example, Sodhi et al. [144] used data from an optical-based tactile sensor to estimate the pose of an object being pushed, while others [151, 152] explored the use of these sensors for estimating the relative pose of an object during a grasp. In [60], the use of tactile sensing for in-hand physical feature exploration was also explored in order to achieve accurate dynamic pivoting manipulations. Wang et al. [60] also used optical tactile sensors on a parallel gripper in order to train a model to predict future pivoting angles given some control parameters. Toskov et al. [56] addressed the pivoting with tactile sensing and trained a recursive ANN for estimating the state of the swinging object. The model was then integrated with the gripper controller in order to regulate the gripper-object angle. Funabashi et al. [25] learned robot hand-grasping postures for objects with tactile feedback enabling manipulation of objects of various sizes and shapes. These works demonstrate the potential of haptic perception and learning techniques for improving the accuracy and efficiency of in-hand manipulation tasks. In practice, tactile sensing provides valuable state information which is hard to extract with alternative methods.

B. Learning hand transition models

A common solution for coping with the unavailability of a feasible model is to learn a *transition model* from data. Robot learning problems can typically be formulated as a Markov Decision Process (MDP) [153]. Hence, a *transition model*, or *forward model*, is a mapping from a given state $\mathbf{x}_t \in \mathcal{X}$ and action $\mathbf{a}_t \in \mathcal{A}$ to the next state $\mathbf{x}_{t+1}$, such that $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{a}_t)$. Subsets $\mathcal{X}$ and $\mathcal{A}$ are the state and action spaces of the system, respectively. Such models are commonly obtained through non-linear regression in a high-dimensional space. Often, the forward model is described as a probability distribution function, i.e., $P(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{a}_t)$, in order to represent uncertainties in the transition.

Learning transition models for in-hand manipulation tasks typically involves understanding how changes in the robot's state are caused by its actions [42]. While a hand transition model is generally available analytically in rigid hands where the kinematics are known [154], analytical solutions are rarely available for compliant or soft hands. As far as the authors' knowledge, the major work on learning transition models involves such hands. Attempts to model compliant hands usually rely on external visual feedback. For example, Sintov et al. [50] proposed a data-based transition model for in-grasp manipulation with a compliant hand where the state of the hand involves kinesthetic features such as actuator torques and angles along with the position of the manipulated object acquired with visual feedback. The extension of this work used a data-based transition model in an asymptotically optimal motion planning framework in the space of state distributions, i.e., in the belief space [33, 51]. Recently, Morgan et al. [40] proposed an object-agnostic manipulation using a vision-based Model Predictive Control (MPC) by learning the manipulation model of a compliant hand through an energy-based perspective [155]. The work in [134] used a depth camera to estimate the pose of an object grasped and partly occluded by the two fingers of an underactuated hand. While the work did not consider manipulation, an extension proposed the use of the depth-based 6D pose estimation to control precise manipulation of a grasped object [41]. The authors leverage the mechanical compliance of a 3-fingered underactuated hand and utilize an object-agnostic offline model of the hand and the 6D pose tracker using synthetic data. While not strictly a transition model, Calli et al. [14] trained a model to classify transitions and identify specific modes during in-grasp manipulations of an underactuated hand. By using visual and kinesthetic perception, state and future actions are classified to possible modes such as object sliding and potential drop.

While the above methods focus on pure visual perception for object pose estimation, tactile sensors were used in recent work independently or combined with vision. Recent work [156] integrated allocentric visual perception along with four tactile modules, that combine pressure, magnetic, angular velocity and gravity sensors, on two underactuated fingers. These sensors were used to train a pose estimation model. Lambeta et al. [142] explored a tactile-based transition model for marble manipulation using a self-supervised detector with auto-encoder architectures. Azulay et al. [12] tackled the problem of partial or fully occluded in-hand object pose estimation by using an observation model that maps haptic sensing on an underactuated hand to the pose of the grasped object. Moreover, an MPC approach was proposed to manipulate a grasped object to desired goal positions solely based on the predictions of the observation model. A similar forward model with MPC was proposed by Luo et al. [38] for the multi-finger dexterous Allegro hand. Overall, these approaches demonstrate the potential of using external visual feedback fused with tactile sensing to learn transition models for in-hand manipulation tasks of various hands with uncertainty.

C. Self-supervision and Exploration for Learning Transitions

Self-supervision and exploration are important techniques for learning transitions of in-hand robotic manipulation. Self-supervision refers to the process of learning

[redacted]

IV. Reinforcement Learning

Reinforcement learning (RL) is one of the main paradigms of machine learning, akin to supervised and unsupervised learning. RL models learn to take optimal actions within some environments by maximizing a given reward (Figure 4). In contrast to model-driven learning, most RL algorithms collect data during the learning process. Often, the learning is done in simulated environments in order to avoid tedious work and wear of the real robot. RL policies are functions mapping current states to optimal actions and a distinction is commonly made between on- and off-policy learning [160]. Both approaches commonly approximate the value function which is an expected cumulative reward defined for states and actions. In on-policy learning, data collection is guided by the intermediate policy learned by the agent. The value function is learned directly by the policy. Therefore, a balance must be kept between exploration of unvisited action-state regions, and exploitation of known regions in order to maximize reward. In off-policy methods, on the other hand, the value function of the optimal policy is learned independent of the actions conducted by the agent during training.

A. Formal RL Definition

We briefly present the formal definition of RL. A wider review of key concepts and methods can be found in the work of Nguyen et al. [161]. Assuming that a given system is a Markov Decision Process (MDP), the system can be defined by the tuple $\{\mathcal{S}, \mathcal{A}, P, \mathcal{R}, \delta\}$ where $\mathcal{S}$ and $\mathcal{A}$ are the state and action spaces, respectively. In MDP, the next state depends solely on the current state $\mathbf{s}_t \in \mathcal{S}$ and desired action $\mathbf{a}_t \in \mathcal{A}$ according to a forward transition dynamics $P(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t)$ while receiving a reward $r_t \in \mathcal{R}$. Scalar $\delta > 0$ is the discount factor which reflects the significance of the reward over time. In model-based RL, transition model P can be learned independently of the policy learning as described in Section III-B.

When the system is stochastic due to uncertainties and limited observability, a Partial Observed Markov Decision Process (POMDP) is considered. The POMDP can be defined by the tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{O}, P, \mathcal{Z}, \mathcal{R}, \delta\}$, where $\mathcal{S}$ is now the unobserved state space and $\mathcal{O}$ is the observation space. As the true state cannot be fully observed, an observation space $\mathcal{O}$ is introduced to describe the set of all possible observations. Therefore, the agent receives an observation $\mathbf{o}_{t+1} \in \mathcal{O}$ when reaching the next state $\mathbf{s}_{t+1}$ with probability $\mathcal{Z}(\mathbf{o}_{t+1}|\mathbf{s}_{t+1})$. In both MDP and POMDP, the goal is, therefore, to learn a policy $\pi : \mathcal{S} \to \mathcal{A}$ or $\pi : \mathcal{O} \to \mathcal{A}$, respectively, which maximizes the expected reward $\mathbb{E}\left\{\sum_{t=0}^{H} \delta^t r_t\right\}$, where $H \in \mathbb{R}^+$ is the horizon. By modeling the system as a POMDP and using RL, the agent has the potential to minimize the true error within a sequence of actions rather than applying greedy ones, such as in supervised learning policies.

Deep reinforcement learning integrates the RL learning paradigm with deep ANNs serving as the policy or value function approximation. Such integration has revealed significant capabilities and led to the success of many reinforcement learning domains [162, 163] including robot manipulation [161, 164] and specifically in-hand dexterous manipulation [9, 30]. Examples of straight-forward implementation of an RL algorithm include the work of Anatova et al. [7] which addressed the pivoting problem with a parallel gripper. The RL policy was trained while relying

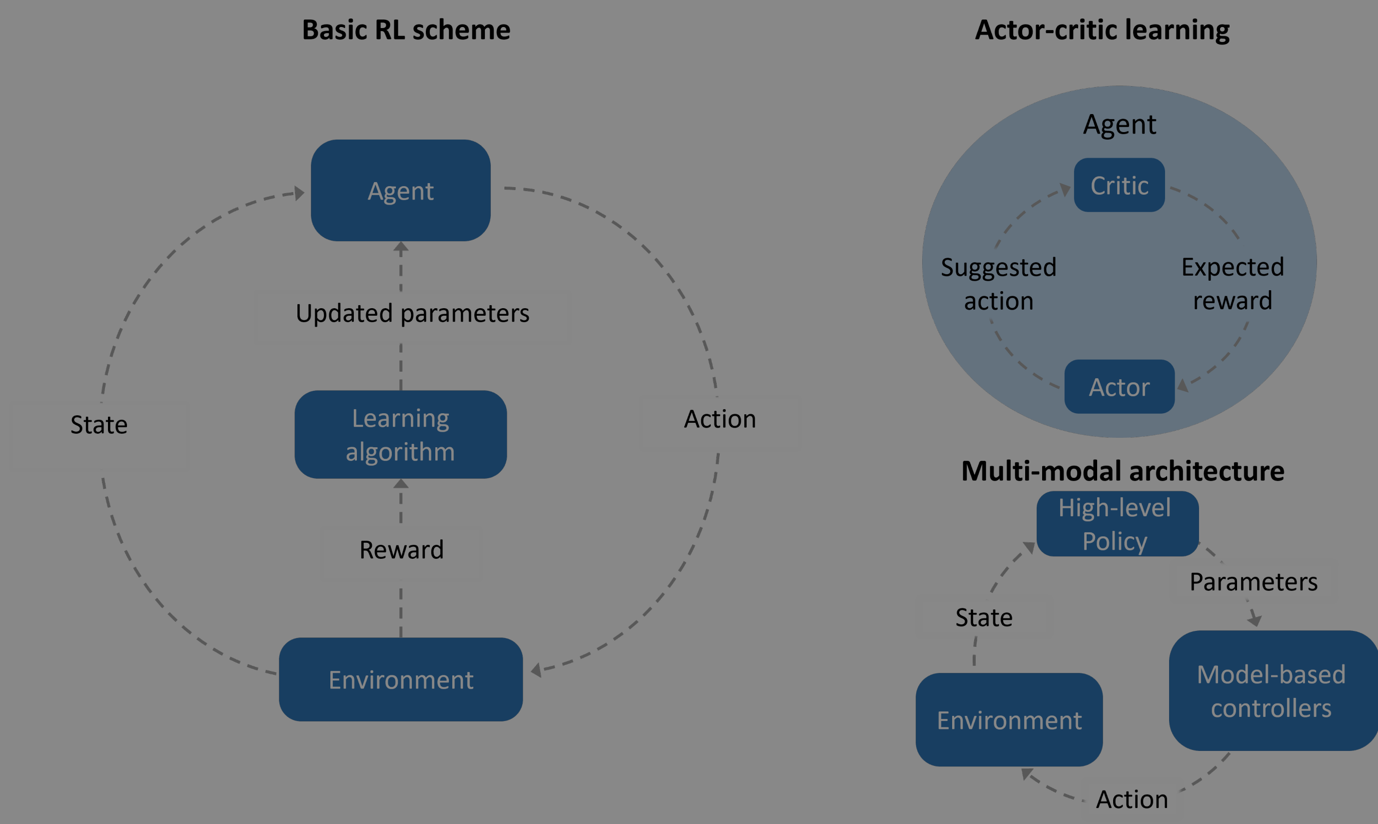

Fig. 4. Illustration of (left) basic RL, (top right) actor-critic architecture, and (bottom right) a multi-network architecture.

on fast tracking with a camera. However, the trained policy yielded excessive back and forth motions. Cruciani and Smith [18] coped with this limitation by employing a three-stage manipulation in which the robot learns to control the velocity and opening of the gripper. An RL policy was acquired through Q-Learning [165]. In an extension work, Cruciani and Smith [19] integrated path planning with the RL policy for the robotic arm to perform more complex tasks. In another example, Van Hoof et al. [58] was the first to employ RL on a two-finger underactuated hand by utilizing its compliance and tactile sensing.

Direct implementation of RL algorithms usually only works in specific and limited applications. Hence, this section presents a comprehensive survey of advanced approaches and current research in the field, with a focus on the unique challenges of in-hand manipulation. First, we describe transfer learning challenges and approaches focusing on the sim-to-real problem. Next, the problem of episodic resetting in real-life experiments is discussed. Finally, we explore the topic of multi-level control systems and actor-critic learning schemes.

B. Transfer learning and sim-to-real problems

Often, ANN based controllers must train extensively for each new task before being able to perform successfully, requiring long training periods and extensive computation resources. Specifically for in-hand manipulation, performing tasks with trained policies on new objects may be challenging. Transfer learning can be divided into few-shot, one-shot and zero-shot learning [166]. Few-shot and one-shot transfer learning requires few instances or a single instance, respectively, of a new task in order to tune the previously trained model. On the other hand, a trained model in zero-shot transfer can instantly perform tasks not included in the training stage. Thus, the ability to learn and generalize to new tasks in few-shots or less is highly beneficial. In in-hand manipulation, the transfer of a model often refers to the generalization to new objects not included in the training [29].

The common approach in few- and one-shot transfer learning is to share weights and data between different tasks, objects and hands. Funabashi et al. [23] demonstrated the ability to pre-train a policy to perform a stable rolling motion with only three fingers of the Allegro hand and then transfer to utilize all four fingers. This was shown to be possible given identical finger morphologies. It was shown that pre-training can be done with data gathered even from random motions such that, afterward, training for specific tasks can be done in one-shot transfer.

While training RL models on real robots yields highly successful controllers [167, 168], it is also expensive in time and resources, or can pose danger with some robots. Furthermore, it tends to require extensive human involvement as discussed in the next section. Consequently, simulations are an emerging approach for policy training as they enable rapid and efficient collection of a massive amount of data. While training on simulations can be beneficial, transferring a robot policy trained in simulation to a real robot remains a challenge [169]. Compared to real-world systems that are usually uncertain and noisy, simulations are naturally more certain and simplified. This gap is commonly known as the sim-to-real problem and can significantly reduce the performance of policies trained in the simulation domain and transferred to the application domain in the real world [170]. This is especially relevant in the case of in-hand manipulation tasks which tend to heavily involve hard-to-model contact dynamics [24, 171]. Hence, the resulting controllers are often sensitive to small errors and external disturbances.

The most common approach for bridging the reality gap in a sim-to-real problem is *domain randomization* [172]. In this approach, various system parameters in the simulation are constantly varied in order to improve robustness to modeling errors. Andrychowicz et al. [8, 9] proposed the Automatic Domain Randomization (ADR) approach where models are trained only in simulation and can be used to solve real-world robot manipulation problems. Specifically, a Rubik's cube was solved by performing finger gaiting and rolling manipulations with the anthropomorphic Shadow hand. ADR automatically generates a distribution over randomized environments. Control policies and vision state estimators trained with ADR exhibit vastly improved sim-to-real transfer.

In the work of Sievers et al. [49], a PyBullet simulation of the DLR hand was used to train an off-policy for in-grasp manipulation solely using tactile and kinesthetic sensing. Domain randomization was used for sim-to-real transfer to the real hand. An extension of the work has demonstrated zero-shot sim-to-real transfer while focusing on 24 goal orientations [44]. Beyond directly modifying the dynamics in domain randomization, applying small random forces to the grasped object was shown by Allshire et al. [6] to improve the robustness of the resulting policy in in-grasp manipulation of the TriFinger hand. Recently, Handa et al. [28] have taken the domain randomization approach to reorient a cube within the four-finger Allegro hand using RGB-D perception. As opposed to Pybullet and similar simulators which are based on CPU computations, Allshire et al. and Handa et al. used the GPU-based Nvidia Isaac Gym simulator [173]. Using a GPU-based simulator reduces the amount of computational resources and costs.

As opposed to domain randomization, Qi et al. [45] used an adaptation module learning to cope with the sim-to-real problem. The module trained through supervised learning to approximate the important properties of the system based solely on kinesthetic sensing. An RL policy is trained to take actions for finger gaiting with a multi-finger hand based on the approximations and on real-time state observations. An extension of the work added visual and tactile perception while also including a Transformer model for embedding past signals [46]. Qi et al. have also used the Isaac Gym simulator since it excels in contact modeling. However, Isaac Gym and most other simulators tend to provide unreliable contact force values. To cope with this limitation, Yin et al. [63] simulated 16 tactile sensors across a four-finger Allegro (i.e., fingertips, fingers and palm) while considering only binary signals of contact or no contact. Due to this configuration, a trained policy is shown to successfully ease the sim-to-real transfer.

C. *Episodic resetting*

Learning robot tasks in the real world often requires sufficient experience. In many systems, this is commonly achieved with frequent human intervention for resetting the environment in between repeating episodes, for example when the manipulated object is accidentally dropped. It is particularly relevant in the case of in-hand manipulation where resetting may be more complex due to large uncertainties in failure outcomes. Removing the costly human intervention will improve sample collection and, thus, decrease learning time. Eysenbach et al. [174] proposed a general approach for training a reset policy simultaneously with the task policy. For instance, a robot manipulator can be trained to reset the environment within the policy training allowing a more autonomous and continuous learning process. As shown in [54], the resulting reset policy can be used as a critic for the task controller in order to discern unsafe task actions that will lead to irreversible states, where reset is inevitable. Specifically in this work, a model learns to identify actions that a Shadow hand may exert while attempting to reorient a cube through rolling and finger gaiting, without the risk of dropping the cube entirely. Preventing the reach of these irreversible states increases the safety of the controller, and can also be used to induce a curriculum for the forward controller.

Another approach to avoid irreversible states is by the addition of a reactive controller designed specifically for intervening only when the robot state is in the close neighborhood of such irreversible states [22]. In this work, Falco et al. used a compliant prosthetic hand in the in-grasp manipulation of objects based on visual perception with an added reactive controller connected to tactile sensors. The goal of the reactive controller is to avoid object slipping. The nominal control method can, thus, be trained with the goal of not only succeeding in the given task but also minimizing the intervention of the reactive controller.

While episodic resetting is often considered a burden, it can instead be considered an opportunity. When training multi-task capabilities for in-hand manipulation, failure in one task may cause a need for a resetting of the grasp. Rather than using human intervention or an additional control system, the reset can instead be viewed as another manipulation task [27]. For example, an unsuccessful attempt at a rolling motion which leads to a wrong object pose, may require the learning of a sliding task to fix the pose. Thus, task training ending in success or failure can both be chained to further learning of other tasks. This results in a reset-free learning scheme.

D. Multi-network architecture

Multi-network architectures, such as actor-critic [175] or teacher-student [176], are often beneficial in improving the learning process. In the more common actor-critic structure, an actor network is trained as the policy while the critic network is trained to estimate the value function. Such structures try to cope with the inherent weaknesses of single-network structures. That is, actor-only models tend to yield high variance and convergence issues while critic-only models have a discretized action space and, therefore, cannot converge to the true optimal policy. In teacher-student architectures, on the other hand, knowledge distillation enables the transfer of knowledge from an unwieldy and complex model to a smaller one. As such, a teacher model is an expert agent that has already learned to take optimal actions whereas the student model is a novice agent learning to make optimal decisions with the guidance of the teacher.

Chen et al. [15–17] used an asymmetric teacher-student training scheme with a teacher trained on full and privileged state information. Then, the teacher policy is distilled into a student policy which acts only based on limited and realistic available information. Policies for object reorientation tasks were trained with a simulated Shadow hand on either the EGAD [177] or the YCB [178] benchmark object sets and tested on the other. Results have exhibited zero-shot transfer to new objects. While the teacher-student approach utilizes privileged information during training, the actor-critic approach manages the learning by continuous interaction between the two models. This was demonstrated with the Proximal Policy Optimization (PPO) algorithm for in-hand pivoting of a rigid body held in a parallel gripper, using inertial forces to facilitate the relative motion [57]. Moreover, combining actor-critic methods with model-based methods can result in improved learning. The learned model can be used within a model predictive controller to reduce model bias induced by the collected data. This was demonstrated by an underactuated hand to perform finger gaiting [179] and object insertion [13]. Recently, Tao et al. [55] proposed to consider the multi-finger hand during a reorientation task as a multi-agent system where each finger or palm is an agent. Each agent has an actor-critic architecture while only the critic has a global observation of all agents. The actor, on the other hand, has only local observability of neighboring agents. In such a way, the hand does not have centralized control and can adapt to changes or malfunctions.

In a different multi-modal architecture approach proposed by Li et al. [37], various control tasks required for in-hand manipulation are divided into multiple hierarchical control levels (Figure 4). This allows the use of more specialized tools for each task. In the lower level, traditional model-based controllers robustly execute different manipulation primitives. On the higher level, a learned policy orchestrates between these primitives for a three-finger hand to robustly reorient grasped objects in a planar environment with gravity. A similar approach was taken by Veiga et al. [59] where low-level controllers maintain a stable grasp using tactile feedback. At the higher level, an RL is trained to perform in-grasp manipulation with a multi-finger dexterous hand.

E. Curriculum Learning

Often, directly training models with data from the entire distribution may yield insufficient performance. Hence, Curriculum Learning (CL) is a training strategy where the model is gradually exposed to increasing task difficulty for enhanced learning efficacy [180]. Such a process imitates the meaningful learning order in human curricula. Adding CL to guide the development of necessary skills can aid policies to learn difficult tasks that tend to have high rates of failure [16]. In this example, researchers modify the behavior of gravity in simulations according to the success rate to aid the learning of gravity-dependent manipulation tasks. This method allows the robot to first successfully learn a skill and then move on to increasingly harder and more accurate problems, slowly reaching the actual desired skill. Azulay et al. [13] trained an actor-critic model to insert objects into shaped holes while performing in-grasp manipulations with a compliant hand. The work has exhibited object-based CL where simple objects were first introduced to the robot followed by more complex ones. Other uses of CL can involve guiding the exploration stages of solution search using reward shaping [6]. The work shows that it is possible to improve early exploration by guiding the model directly to specific regions using specific reward functions as priors. Those regions however may not hold actual feasible solutions, and it may be necessary to reduce the effects of the reward functions in later stages of the learning process.

F. Tactile information

While visual perception is a prominent approach for feedback in RL, it may be quite limited in various environments and the object is often occluded by the hand. On the other hand, tactile sensing provides direct access to information regarding the state of the object. Nevertheless, data from tactile sensing is often ambiguous, and information regarding the object is implicit. Yet, the addition of tactile sensory is widely addressed as it can improve the learning rate. For instance, in the work of Korthals et al. [34], tactile sensory information for the Shadow hand increased the sampling efficiency and accelerated the learning process such that the number of epochs for similar performance was significantly decreased. Jain et al. [30] have shown that the integration of tactile sensors increases the learning rate when the object is highly occluded. This was demonstrated in various manipulation tasks including in-hand manipulation of a pen by a simulated anthropomorphic hand. Melnik et al. [181] compared multiple sensory methods, including continuous versus binary (i.e., touch or no touch) tactile signals and, higher versus lower sensory resolutions. The results from this comparative study have shown that using tactile information is beneficial to the learning process, compared to not having such information. However, the specific method that gave the best result was dependent on the learned manipulation task [39].

While Melnik et al. used tactile information directly as part of the state vector, Funabashi et al. [24, 182] used a higher-resolution sensory array without visual perception. The model coped with the increased dimension of the output tactile information by considering the relative spatial positioning of the sensors. Similarly, Yang et al. [62] used a tactile array across a multi-finger hand. The array was embedded using Graph Neural Network which provides an object state during the manipulation and used for model-free RL. Recently, Khandate et al. [31] implemented model-free RL to reorient an object through finger gaiting with a multi-finger dexterous hand while only using kinesthetic and tactile sensing. In an extension work, Khandate et al. offered the use of sampling-based motion planning in order to sample useful parts of the manipulation space and improve the exploration [32]. Tactile sensing provides valuable information and often can fully replace visual perception. However, policies based on tactile perception usually require an excessive amount of real-world experience in order to reach sufficient and generalized performance.

In summary, research in the field of RL for robotic in-hand manipulation is growing and achieving increasing success in recent years. While showcasing promising performance in specific tasks, RL policies still perform poorly in multi-task scenarios and struggle to generalize with zero- or few-shot learning [183]. Major challenges currently being faced include the control of highly dexterous hands with high amounts of sensory information, the transfer of a model learned in simulation to a real robotic system, and the transfer of learning specific tasks on well-known objects to other tasks and unknown objects. Major advances are being achieved using multi-level control structures, domain and dynamic adaptation, and the combination of model-based and model-free methods to gain the benefits of both. While exciting advances have been made in RL, the field continues to explore the challenges of data efficiency and adaptability to new domains.

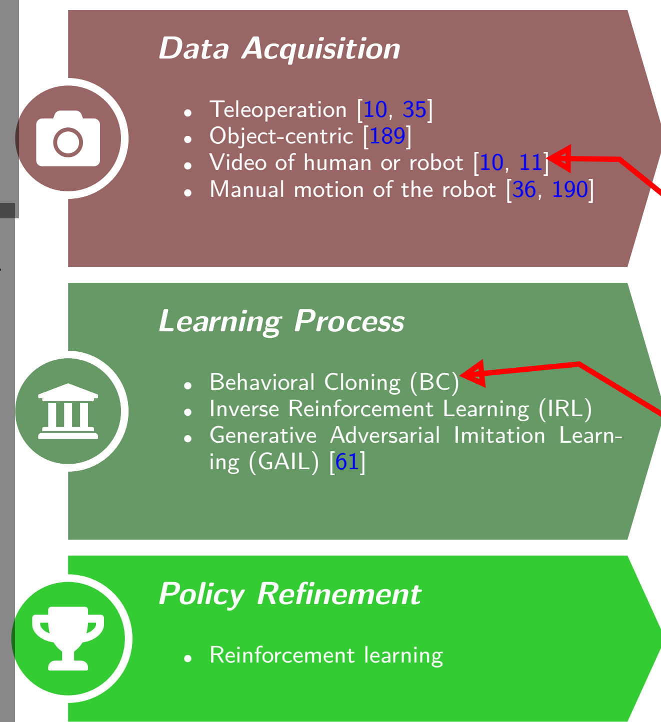

Fig. 5. Flowchart of policy training with Imitation Learning (IL).

This is what we are looking for[1]

V. Imitation Learning

As discussed in the previous section, training RL policies for real robots from scratch is usually time-consuming and often infeasible due to the lack of sufficient data [184]. A prominent approach for coping with these challenges is Imitation Learning (IL). Instead of learning a skill without prior knowledge, IL aims to learn from expert demonstrations [185, 186]. Prior knowledge from the expert can then be optimized for the agent through some learning framework such as RL. While IL is often considered a sub-field of RL, we provide a distinct focus due to its importance and wide work. IL can be categorized into two main approaches: Behavioral Cloning (BC) and Inverse Reinforcement Learning (IRL) [187]. In BC, a policy is trained in a supervised learning fashion with expert data to map states to actions. IRL, on the other hand, extracts the reward function from the expert data in order to train an agent with the same preferences [188].

the procedure[2]

In both BC and IRL, a policy is learned with some prior. This is in contrast to RL where the policy is learned from scratch based on the agent's own experience. Hence, IL requires an initial process of data acquisition as illustrated in Figure 5. First, data is collected from demonstrations of an expert. Demonstrations can be acquired in various mediums such as recording human motion or recording proprioceptive sensing of the robot during manual jogging. In the next step, IL usually involves either learning a policy to directly imitate the demonstration (BC) or feature extraction from the data (IRL). The last step is further policy refinement through conventional RL. From an RL perspective, IL usually reduces the learning time by bootstrapping the learning process using an approximation of the expert's policy.

Now the details [2-1]

A. *Date Acquisition*

Data is collected from an expert demonstrator while conducting the desired task. The motion of the expert is recorded through some set of sensors such that the learning agent can later observe and learn to imitate. There are various approaches to demonstrate and record the motions and their choice may affect the learning process.

One data acquisition approach is to teleoperate the robot throughout the task using designated tools such as a remote control [191]. However, remote controls are unnatural and quite infeasible in teleoperation of dexterous robotic hands. In a more natural approach, Arunachalam et al. [10] used a visual hand pose estimation model (i.e., skeleton) to approximate keypoints on the human hand during reorientation of an object. The user can also use a VR set in order to have the point-of-view of the robot [11]. In these examples, a policy is learned for an anthropomorphic robot hand by using simpler nearest-neighbors search in the data. The action in the demonstration data which has a state closest to the current state is exerted. Similarly, Kumar et al. [35] recorded the proprioceptive state of a virtual anthropomorphic robotic hand during teleoperation with the CyberGlove worn by an expert user. With the glove, the joint angles along with tactile information are recorded. The recorded tasks are then used to train and evaluate in-hand manipulation with a five-finger dexterous hand for reorientation tasks. In a similar approach, Wei and Xu [192] designed a wearable robotic hand for IL teleoperation such that the expert has tactile feedback during the demonstrations.

In a different approach by Gupta et al. [189], only information regarding the motion of a manipulated object is collected while ignoring the motions of the human expert. Hence, an object-centric policy is learned while selecting the most relevant demonstration for each initial state in the training. In a different approach, the demonstrator manually moves the robot by contacting and pushing it to perform the task [36, 190]. During the demonstration, the robot collects kinesthetic data from the joints. While the approach is simple, it is usually applied to robotic arms with a single serial kinematic chain. It is quite infeasible for a human to synchronously move a dexterous and multi-contact robotic hand to perform a complex in-hand manipulation task. Nevertheless, simpler tasks with non-dexterous hands may by possible while the authors have not found prior work.

B. Learning Process

The process for learning from the demonstrations is commonly conducted by either BC or IRL [193]. In BC, the agent is required to directly take the strategy of the expert observed in the demonstrations [10]. The agent will exert an action taken by the expert when in a similar state. Hence, demonstration data is usually recorded in the form of state-action pairs which is easy to learn. Then, a policy is learned in a supervised learning manner. However, state-action pairs can be difficult to obtain from, for instance, video data. To cope with this problem, Radosavovic et al. [47] proposed the state-only imitation learning (SOIL) approach where an inverse dynamics model can be trained to extract the actions chosen based on the change in the state perceived from videos. The inverse dynamics model and the policy are trained jointly. SOIL enables learning from demonstrations originating from different but related settings. While not an IL approach, Yuan et al. [65] considered a trained teacher policy as an expert and used BC to distill it to a student in the training of in-hand manipulation with vision and tactile sensing. In a different work, BC was used to control a unique design of a gripper having actuated rollers on its fingertips [64]. The demonstration data was extracted from a handcrafted controller and shown to improve performance.

Rajeswaran et al. [48] compared methods of RL to solve complex manipulation tasks, with and without incorporating human demonstrations. The authors suggested a method of incorporating demonstrations into policy gradient methods for manipulation tasks. The proposed Demonstration Augmented Policy Gradient (DAPG) method uses pre-training with BC to initialize the policy and an augmented loss function to reduce ongoing bias toward the demonstration. The results in the paper showcase that DAPG policies can acquire more human-like motion compared to RL from scratch and are substantially more robust. In addition, the learning process is considerably more sample-efficient. Jain et al. [30] extended the work by exploring the contribution of demonstration data to visuomotor policies while being agnostic about the data's origin. Demonstrations were shown to improve the learning rate of these policies in which they can be trained efficiently with a few hundred expert demonstration trajectories. In addition, tactile sensing was found to enable faster convergence and better asymptotic performance for tasks with a high degree of occlusions. While Rajeswaran et al. and Jain et al. demonstrated the approach only in simulations, Zhu et al. [168] demonstrated the use of DAPG on a real-robot in complex dexterous manipulation tasks. The results have shown a decrease of training time from 4-7 hours to 2-3 hours by incorporating human demonstrations.

Few studies on in-hand manipulation have used BC due to the significant effort required to collect sufficient demonstration data. While simple to implement, BC usually requires large amounts of data for sufficient performance [194]. IRL, on the other hand, directly learns the reward function of the demonstrated expert policy which prioritizes some actions over others [188]. IRL learns the underlying reward function of the expert which is the best definition of a task. Once acquired the reward function, an optimal policy can be trained to maximize such a reward using a standard RL algorithm. While general work on IRL is wide for various robotic applications, not much work has been done that combines IRL with in-hand manipulation. A single work demonstrated the IRL approximation of the reward function using expert samples of desired behaviours [43]. However, the authors have argued that the learned reward functions are biased towards the demonstrated actions and fail to generalize. Randomization and normalization were used to minimize the bias and enable generalization between different tasks.

While not directly IRL, Deng and Zhang [20] utilized reward-shaping to improve the RL training of in-hand manipulation with a dexterous hand. By observing hand synergies of a human demonstrator, a limited and low-dimensional state space was constructed. Using reward-shaping allows the inclusion of multiple levels of knowledge, from the standard extrinsic reward to hand synergies-based reward and an uncertainty-based reward function that is aimed at directing efficient exploration of the state space. Learning using all three reward functions is shown through simulations to improve learning. The minor use of IRL to address in-hand manipulation problems may be explained by its tendency to provide ill-behaved reward functions and unstable policies [195].

IL was also proposed for in-hand manipulation without the use of RL. Solak et al. [52] proposed the use of Dynamical Movement Primitives (DMP) [196]. The approach shows that a multi-finger dexterous hand can perform a task based on a single human demonstration while being robust to changes in the initial or final state, and is object-agnostic. However, the former property may yield object slip and compromise grasp stability. Hence, an extended work proposed haptic exploration of the object such that the manipulation is informed by surface normals and friction at the contacts [53].

While traditional IRL has shown high performance in a wide range of tasks, it only provides a reward function that implicitly explains the experts' behaviour but does not provide the policy dictating what actions to take. Hence, the agent will still have to learn a policy through RL training in a rather expensive process. To address this problem, the Generative Adversarial Imitation Learning (GAIL) [197] was proposed and combines IL with Generative Adversarial Networks (GAN) [198]. Similar to GAN, GAIL incorporates a generator and a discriminator. While the generator attempts to generate a policy that matches the demonstrations, the discriminator attempts to distinguish between data from the generator and the original demonstration data. Training of GAIL is, therefore, the minimization of the difference between the two. Consequently, GAIL is able to extract a policy from the demonstration data. Recently, the use of GAIL was proposed for in-hand manipulation by a dexterous hand [61]. The approach was shown to perform significantly better than BC or direct RL training. GAIL has the potential to improve and expedite policy learning of more complex in-hand manipulation tasks, and should be further explored.

VI. Discussion & Conclusions

In-hand manipulation is one of the most challenging topics in robotics and an important aspect for feasible robotic applications. Traditional analytical methods struggle to estimate object properties and noisy sensory information. With in-hand manipulation reaching a bottleneck using these traditional methods, researchers are leveraging advancements in deep learning and reinforcement learning to unlock new levels of dexterity. These tools encapsulate the ability to model complex and noisy systems such as a dexterous robotic hand equipped with various sensors. Nevertheless, current research still faces significant challenges:

1) *Data efficiency.* Learning models is essential for understanding changes in the robot's state caused by its actions during in-hand manipulation. While analytical solutions are available for rigid hands, compliant or soft hands rely on external visual feedback. However, collecting data can be challenging due to the high-dimensional state space and the need to explore the entire feasible space. Future work by researchers should address methods to reduce the required size of training data by making models more general to various applications. For instance, Bayesian optimization can assist in identifying key sampling locations, but reaching some regions may require complex maneuvers, making it necessary to have a good prior model to learn a better one.
2) *Sim-to-real transfer.* Learning policies in simulation is a prominent approach to improve data efficiency in robot training. While significant progress has been made to address the sim-to-real problem, simulations hardly represent the real world and trained policies work poorly on the real system. Hence, large efforts should be put into closing the reality gap by generating better simulations and, incorporating advanced data-based models that can generalize better. Examples of the latter include decision transformers [199] and diffusion policies [200].
3) *Soft robotic hands.* High-dexterous hands such as anthropomorphic ones have been demonstrated in multiple complex in-hand manipulation tasks. However, they are highly expensive making their adaptation to real-world tasks not possible. Consequently, an abundance of research and development has been put in recent years on soft robotic hands that are typically low-cost to manufacture. However, these hands cannot be modeled or controlled analytically and learning approaches are the common paradigm. As discussed previously, common solutions require a significant amount of data and are usually specific to a single hand and task. Therefore, the robotics community should promote efficient learning approaches that can be generalized to different hardware, tasks and environments.
4) *Tactile sensing.* While visual perception technology is quite mature, the use of high-resolution tactile sensing is relatively new. In general, Table I clearly shows the dominance of visual perception over tactile sensing in research. Highly capable tactile sensors can provide vital information regarding the contact state including position, forces, torsion, shape and texture. Nevertheless, they often require a large amount of real-world data in order to perform well. Simulations such as TACTO [201] address this problem by simulating tactile interactions. However, these remain quite far from reality and cannot provide reliable load sensing. Practitioners should work toward better tactile simulators along with distillation approaches for efficient sim-to-real transfer.
5) *Learning from Demonstrations.* IL with expert demonstrations has proved to be efficient for shortening the data-hungry training phase of RL. However, hardware and methods for collecting demonstration

data generally lack the ability to capture the entire state space of the hand-object system. For instance, visual perception is incapable of observing the intrinsic and contact state of the system. Furthermore, IL models focus on task completion and fail to address strategy learning with efficient data utilization. Future work should facilitate efficient platforms for collecting high-dimensional data in the real world. In addition, learning methods should require a small amount of data from the expert user in order to generalize well to various scenarios of the tasks.

6) *Task generalization.* The prevailing paradigm in in-hand manipulation focuses on crafting task-specific or narrowly applicable policies, which hinders broader applicability. The field therefore necessitates a paradigm shift toward solutions capable of seamless adaptation or generalization to novel tasks or objects.

Overall, learning approaches such as supervised learning, haptic perception and self-supervision, along with the combination of visual and tactile sensing, show promise in enhancing the accuracy and efficiency of in-hand robotic manipulation tasks. These techniques enable robots to acquire a comprehensive understanding of the object-hand system and improve their manipulation capabilities. Many challenges remain to be met by the academic community in order for high-ability robots to be used in real-world applications.

References

[1] A. C. Lehman, N. A. Wood, S. Farritor, M. R. Goede, and D. Oleynikov, "Dexterous miniature robot for advanced minimally invasive surgery," *Surgical Endoscopy*, vol. 25, pp. 119–123, 2010.

[2] T. Kang, J.-B. Yi, D. Song, and S.-J. Yi, "High-speed autonomous robotic assembly using in-hand manipulation and re-grasping," *Applied Sciences*, vol. 11, no. 1, p. 37, 2021.

[3] L. Petrich, J. Jin, M. Dehghan, and M. Jagersand, *Assistive arm and hand manipulation: How does current research intersect with actual healthcare needs?* 2021. arXiv: 2101.02750 [cs.RO].

[4] O. Kroemer, S. Niekum, and G. Konidaris, "A review of robot learning for manipulation: Challenges, representations, and algorithms," *The Journal of Machine Learning Research*, vol. 22, no. 1, pp. 1395–1476, 2021.

[5] J. Sun, J. P. King, and N. S. Pollard, "Characterizing continuous manipulation families for dexterous soft robot hands," *Frontiers in Robotics and AI*, vol. 8, p. 645 290, 2021.

[6] A. Allshire, M. Mittal, V. Lodaya, V. Makoviychuk, D. Makoviichuk, F. Widmaier, M. Wüthrich, S. Bauer, A. Handa, and A. Garg, "Transferring dexterous manipulation from GPU simulation to a remote real-world trifinger," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 2022, pp. 11 802–11 809.

[7] R. Antonova, S. Cruciani, C. Smith, and D. Kragic, "Reinforcement learning for pivoting task," *arXiv preprint arXiv:1703.00472*, 2017.

[8] OpenAI *et al.*, "Solving Rubik's Cube with a Robot Hand," pp. 1–51, 2019. arXiv: 1910.07113.

[9] O. M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, *et al.*, "Learning dexterous in-hand manipulation," *The International Journal of Robotics Research*, vol. 39, no. 1, pp. 3–20, 2020.

[10] S. P. Arunachalam, S. Silwal, B. Evans, and L. Pinto, "Dexterous imitation made easy: A learning-based framework for efficient dexterous manipulation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5954–5961. DOI: 10.1109/ICRA48891.2023.10160275.

[11] S. P. Arunachalam, I. Guzey, S. Chintala, and L. Pinto, "Holo-dex: Teaching dexterity with immersive mixed reality," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5962–5969.

[12] O. Azulay, I. Ben-David, and A. Sintov, "Learning haptic-based object pose estimation for in-hand manipulation control with underactuated robotic hands," *IEEE Transactions on Haptics*, pp. 1–12, 2022. DOI: 10.1109/TOH.2022.3232713.

[13] O. Azulay, M. Monastirsky, and A. Sintov, "Haptic-based and $SE(3)$-aware object insertion using compliant hands," *IEEE Robotics and Automation Letters*, vol. 8, no. 1, pp. 208–215, 2022.

[14] B. Calli, K. Srinivasan, A. Morgan, and A. M. Dollar, "Learning modes of within-hand manipulation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 3145–3151. DOI: 10.1109/ICRA.2018.8461187.

[15] T. Chen, J. Xu, and P. Agrawal, "A simple method for complex in-hand manipulation," in *Annual Conference on Robot Learning*, vol. 3, 2021.

[16] T. Chen, J. Xu, and P. Agrawal, "A system for general in-hand object re-orientation," in *Conference on Robot Learning*, 2022.

[17] T. Chen, M. Tippur, S. Wu, V. Kumar, E. Adelson, and P. Agrawal, "Visual dexterity: In-hand reorientation of novel and complex object shapes," *Science Robotics*, vol. 8, no. 84, 2023. DOI: 10.1126/scirobotics.adc9244.

[18] S. Cruciani and C. Smith, "In-hand manipulation using three-stages open loop pivoting," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1244–1251. DOI: 10.1109/IROS.2017.8202299.

[19] S. Cruciani and C. Smith, "Integrating path planning and pivoting," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 6601–6608.

[20] Z. Deng and J. W. Zhang, "Learning synergies based in-hand manipulation with reward shap-

ing," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 3, pp. 141–149, 2020.

[21] D. Dimou, J. Santos-Victor, and P. Moreno, "Robotic hand synergies for in-hand regrasping driven by object information," *Autonomous Robots*, vol. 47, pp. 453–464, 2023.

[22] P. Falco, A. Attawia, M. Saveriano, and D. Lee, "On Policy Learning Robust to Irreversible Events: An Application to Robotic In-Hand Manipulation," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1482–1489, 2018. DOI: 10.1109/lra.2018.2800110.

[23] S. Funabashi, A. Schmitz, S. Ogasa, and S. Sugano, "Morphology specific stepwise learning of in-hand manipulation with a four-fingered hand," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 433–441, 2019.

[24] S. Funabashi, S. Ogasa, T. Isobe, T. Ogata, A. Schmitz, T. P. Tomo, and S. Sugano, "Variable in-hand manipulations for tactile-driven robot hand via cnn-lstm," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9472–9479.

[25] S. Funabashi, A. Schmitz, T. Sato, S. Somlor, and S. Sugano, "Versatile In-Hand Manipulation of Objects with Different Sizes and Shapes Using Neural Networks," *IEEE-RAS International Conference on Humanoid Robots*, vol. 2018-Novem, pp. 768–775, 2019. DOI: 10.1109/HUMANOIDS.2018.8624961.

[26] G. Garcia-Hernando, E. Johns, and T.-K. Kim, "Physics-based dexterous manipulations with estimated hand poses and residual reinforcement learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9561–9568. DOI: 10.1109/IROS45743.2020.9340947.

[27] A. Gupta, J. Yu, T. Z. Zhao, V. Kumar, A. Rovinsky, K. Xu, T. Devlin, and S. Levine, "Reset-free reinforcement learning via multi-task learning: Learning dexterous manipulation behaviors without human intervention," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 6664–6671. DOI: 10.1109/ICRA48506.2021.9561384.

[28] A. Handa *et al.*, "Dextreme: Transfer of agile in-hand manipulation from simulation to reality," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5977–5984. DOI: 10.1109/ICRA48891.2023.10160216.

[29] W. Huang, I. Mordatch, P. Abbeel, and D. Pathak, *Generalization in dexterous manipulation via geometry-aware multi-task learning*, 2021. arXiv: 2111.03062.

[30] D. Jain, A. Li, S. Singhal, A. Rajeswaran, V. Kumar, and E. Todorov, "Learning Deep Visuomotor Policies for Dexterous Hand Manipulation," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3636–3643, 2019. DOI: 10.1109/icra.2019.8794033.

[31] G. Khandate, M. Haas-Heger, and M. Ciocarlie, "On the feasibility of learning finger-gaiting in-hand manipulation with intrinsic sensing," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2022, pp. 2752–2758. DOI: 10.1109/ICRA46639.2022.9812212.

[32] G. Khandate, S. Shang, E. T. Chang, T. L. Saidi, J. Adams, and M. Ciocarlie, "Sampling-based Exploration for Reinforcement Learning of Dexterous Manipulation," in *Robotics: Science and Systems*, Daegu, Republic of Korea, 2023.

[33] A. Kimmel, A. Sintov, J. Tan, B. Wen, A. Boularias, and K. E. Bekris, "Belief-space planning using learned models with application to underactuated hands," in *International Symposium on Robotics Research (ISRR)*, 2019.

[34] T. Korthals, A. Melnik, J. Leitner, and M. Hesse, "Multisensory assisted in-hand manipulation of objects with a dexterous hand," in *Proceedings of 2019 ICRA Workshop on ViTac: Integrating Vision and Touch for Multimodal and Cross-modal Perception*, 2019, pp. 1–2.

[35] V. Kumar, A. Gupta, E. Todorov, and S. Levine, "Learning Dexterous Manipulation Policies from Experience and Imitation," 2016. DOI: 10.1177/ToBeAssigned. arXiv: 1611.05095v1.

[36] M. Li, H. Yin, K. Tahara, and A. Billard, "Learning object-level impedance control for robust grasping and dexterous manipulation," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 6784–6791. DOI: 10.1109/ICRA.2014.6907861.

[37] T. Li, K. Srinivasan, M. Q.-H. Meng, W. Yuan, and J. Bohg, "Learning hierarchical control for robust in-hand manipulation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 8855–8862. DOI: 10.1109/ICRA40945.2020.9197343.

[38] Y. Luo, W. Li, P. Wang, H. Duan, W. Wei, and J. Sun, *Progressive transfer learning for dexterous in-hand manipulation with multi-fingered anthropomorphic hand*, 2023. arXiv: 2304.09526 `[cs.RO]`.

[39] A. Melnik, L. Lach, M. Plappert, T. Korthals, R. Haschke, and H. Ritter, "Using tactile sensing to improve the sample efficiency and performance of deep deterministic policy gradients for simulated in-hand manipulation tasks," *Frontiers in Robotics and AI*, p. 57, 2021.

[40] A. S. Morgan, K. Hang, and A. M. Dollar, "Object-agnostic dexterous manipulation of partially constrained trajectories," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5494–5501, 2020.

[41] A. S. Morgan, B. Wen, J. Liang, A. Boularias, A. M. Dollar, and K. Bekris, "Vision-driven compliant manipulation for reliable, high-precision as-

sembly tasks," in *Robotics: Science and Systems*, 2021.

[42] A. Nagabandi, K. Konolige, S. Levine, and V. Kumar, "Deep dynamics models for learning dexterous manipulation," in *Conference on Robot Learning*, PMLR, 2020, pp. 1101–1112.

[43] J. Orbik, A. Agostini, and D. Lee, "Inverse reinforcement learning for dexterous hand manipulation," in *IEEE International Conference on Development and Learning (ICDL)*, 2021, pp. 1–7. DOI: 10.1109/ICDL49984.2021.9515637.

[44] J. Pitz, L. Röstel, L. Sievers, and B. Bäuml, "Dextrous tactile in-hand manipulation using a modular reinforcement learning architecture," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 1852–1858. DOI: 10.1109/ICRA48891.2023.10160756.

[45] H. Qi, A. Kumar, R. Calandra, Y. Ma, and J. Malik, "In-Hand Object Rotation via Rapid Motor Adaptation," in *Conference on Robot Learning (CoRL)*, 2022.

[46] H. Qi, B. Yi, S. Suresh, M. Lambeta, Y. Ma, R. Calandra, and J. Malik, "General in-hand object rotation with vision and touch," in *Conference on Robot Learning*, J. Tan, M. Toussaint, and K. Darvish, Eds., vol. 229, 2023, pp. 2549–2564.

[47] I. Radosavovic, X. Wang, L. Pinto, and J. Malik, "State-only imitation learning for dexterous manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 7865–7871.

[48] 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," 2017. arXiv: 1709.10087.

[49] L. Sievers, J. Pitz, and B. Bäuml, "Learning purely tactile in-hand manipulation with a torque-controlled hand," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2022, pp. 2745–2751. DOI: 10.1109/ICRA46639.2022.9812093.

[50] A. Sintov, A. S. Morgan, A. Kimmel, A. M. Dollar, K. E. Bekris, and A. Boularias, "Learning a State Transition Model of an Underactuated Adaptive Hand," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1287–1294, 2019. DOI: 10.1109/LRA.2019.2894875.

[51] A. Sintov, A. Kimmel, K. E. Bekris, and A. Boularias, "Motion planning with competency-aware transition models for underactuated adaptive hands," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 7761–7767.

[52] G. Solak and L. Jamone, "Learning by demonstration and robust control of dexterous in-hand robotic manipulation skills," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 8246–8251. DOI: 10.1109/IROS40897.2019.8967567.

[53] G. Solak and L. Jamone, "Haptic exploration of unknown objects for robust in-hand manipulation," *IEEE Transactions on Haptics*, vol. 16, no. 3, pp. 400–411, 2023. DOI: 10.1109/TOH.2023.3300439.

[54] K. Srinivasan, B. Eysenbach, S. Ha, J. Tan, and C. Finn, "Learning to be safe: Deep RL with a safety critic," *arXiv preprint arXiv:2010.14603*, 2020.

[55] L. Tao, J. Zhang, M. Bowman, and X. Zhang, "A multi-agent approach for adaptive finger cooperation in learning-based in-hand manipulation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 3897–3903. DOI: 10.1109/ICRA48891.2023.10160909.

[56] J. Toskov, R. Newbury, M. Mukadam, D. Kulic, and A. Cosgun, "In-hand gravitational pivoting using tactile sensing," in *Proceedings of The 6th Conference on Robot Learning*, K. Liu, D. Kulic, and J. Ichnowski, Eds., ser. Proceedings of Machine Learning Research, vol. 205, PMLR, 2023, pp. 2284–2293.

[57] L. V. O. Toledo, G. J. Giardini Lahr, and G. A. P. Caurin, "In-hand manipulation via deep reinforcement learning for industrial robots," in *Multibody Mechatronic Systems*, M. Pucheta, A. Cardona, S. Preidikman, and R. Hecker, Eds., Cham: Springer International Publishing, 2021, pp. 222–228.

[58] H. Van Hoof, T. Hermans, G. Neumann, and J. Peters, "Learning robot in-hand manipulation with tactile features," *IEEE-RAS International Conference on Humanoid Robots*, vol. 2015-Decem, pp. 121–127, 2015, ISSN: 21640580. DOI: 10.1109/HUMANOIDS.2015.7363524.

[59] F. Veiga, R. Akrour, and J. Peters, "Hierarchical tactile-based control decomposition of dexterous in-hand manipulation tasks," *Frontiers in Robotics and AI*, vol. 7, 2020, ISSN: 2296-9144. DOI: 10.3389/frobt.2020.521448.

[60] C. Wang, S. Wang, B. Romero, F. Veiga, and E. Adelson, "Swingbot: Learning physical features from in-hand tactile exploration for dynamic swing-up manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5633–5640.

[61] D. Wei, G. Sun, Z. Ren, S. Li, Z. Shao, X. Li, N. Tsagarakis, and S. Ma, *In-hand re-grasp manipulation with passive dynamic actions via imitation learning*, 2023. arXiv: 2309.15455 [cs.RO].

[62] L. Yang, B. Huang, Q. Li, Y.-Y. Tsai, W. W. Lee, C. Song, and J. Pan, "Tacgnn: Learning tactile-based in-hand manipulation with a blind robot using hierarchical graph neural network," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3605–3612, 2023. DOI: 10.1109/LRA.2023.3264759.

[63] Z.-H. Yin, B. Huang, Y. Qin, Q. Chen, and X. Wang, "Rotating without seeing: Towards in-hand dexterity through touch," in *Robotics: Science and Systems*, 2023.

[64] S. Yuan, L. Shao, C. L. Yako, A. Gruebele, and J. K. Salisbury, "Design and control of roller grasper v2 for in-hand manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9151–9158. DOI: 10.1109/IROS45743.2020.9340953.

[65] Y. Yuan, H. Che, Y. Qin, B. Huang, Z.-H. Yin, K.-W. Lee, Y. Wu, S.-C. Lim, and X. Wang, *Robot synesthesia: In-hand manipulation with visuotactile sensing*, 2023. arXiv: 2312.01853 [cs.RO].

[66] A. Okamura, N. Smaby, and M. Cutkosky, "An overview of dexterous manipulation," in *IEEE International Conference on Robotics and Automation.*, vol. 1, 2000, pp. 255–262.

[67] M. T. Mason and J. K. Salisbury, *Robot Hands and the Mechanics of Manipulation*. Cambridge, MA, USA: MIT Press, 1985.

[68] S. Abondance, C. B. Teeple, and R. J. Wood, "A dexterous soft robotic hand for delicate in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5502–5509, 2020. DOI: 10.1109/LRA.2020.3007411.

[69] R. R. Ma and A. M. Dollar, "Yale openhand project: Optimizing open-source hand designs for ease of fabrication and adoption," *IEEE Robotics and Autumation Magezine*, vol. 24, pp. 32–40, 2017.

[70] L. Han, Y. Guan, Z. Li, Q. Shi, and J. Trinkle, "Dextrous manipulation with rolling contacts," in *IEEE International Conference on Robotics and Automation*, vol. 2, 1997, 992–997 vol.2.

[71] A. Sintov and A. Shapiro, "Swing-up regrasping algorithm using energy control," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4888–4893.

[72] F. E. Viña B., Y. Karayiannidis, K. Pauwels, C. Smith, and D. Kragic, "In-hand manipulation using gravity and controlled slip," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 5636–5641.

[73] N. C. Dafle, A. Rodriguez, R. Paolini, B. Tang, S. S. Srinivasa, M. Erdmann, M. T. Mason, I. Lundberg, H. Staab, and T. Fuhlbrigge, "Extrinsic dexterity: In-hand manipulation with external forces," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 1578–1585.

[74] A. Sintov, O. Tslil, and A. Shapiro, "Robotic swing-up regrasping manipulation based on impulse-momentum approach and cLQR control," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1079–1090, 2016.

[75] J. Shi, J. Z. Woodruff, P. B. Umbanhowar, and K. M. Lynch, "Dynamic in-hand sliding manipulation," *IEEE Transactions on Robotics*, vol. 33, no. 4, pp. 778–795, 2017.

[76] J. Shi, H. Weng, and K. M. Lynch, "In-hand sliding regrasp with spring-sliding compliance and external constraints," *IEEE Access*, vol. 8, pp. 88 729–88 744, 2020. DOI: 10.1109/ACCESS.2020.2991382.

[77] R. R. Ma and A. M. Dollar, "On dexterity and dexterous manipulation," in *International Conference on Advanced Robotics (ICAR)*, 2011, pp. 1–7.

[78] D. Rus, "In-hand dexterous manipulation of piecewise-smooth 3d objects," *The International Journal of Robotics Research*, vol. 18, no. 4, pp. 355–381, 1999.

[79] J. Xu, T. J. Koo, and Z. Li, "Finger gaits planning for multifingered manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 2932–2937. DOI: 10.1109/IROS.2007.4399189.

[80] Y. Fan, W. Gao, W. Chen, and M. Tomizuka, "Real-Time Finger Gaits Planning for Dexterous Manipulation," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 12 765–12 772, 2017, ISSN: 24058963. DOI: 10.1016/j.ifacol.2017.08.1831.

[81] T. Lozano-Perez, J. Jones, E. Mazer, P. O'Donnell, W. Grimson, P. Tournassoud, and A. Lanusse, "Handey: A robot system that recognizes, plans, and manipulates," in *IEEE International Conference on Robotics and Automation*, vol. 4, 1987, pp. 843–849. DOI: 10.1109/ROBOT.1987.1087847.

[82] P. Tournassoud, T. Lozano-Perez, and E. Mazer, "Regrasping," in *IEEE International Conference on Robotics and Automation*, vol. 4, 1987, pp. 1924–1928. DOI: 10.1109/ROBOT.1987.1087910.

[83] A. Sintov and A. Shapiro, "Dynamic regrasping by in-hand orienting of grasped objects using non-dexterous robotic grippers," *Robotics and Computer-Integrated Manufacturing*, vol. 50, pp. 114–131, 2017, ISSN: 0736-5845.

[84] M. Guo, D. V. Gealy, J. Liang, J. Mahler, A. Goncalves, S. McKinley, J. A. Ojea, and K. Goldberg, "Design of parallel-jaw gripper tip surfaces for robust grasping," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 2831–2838.

[85] A. Zeng *et al.*, "Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 3750–3757.

[86] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, 2019.

[87] N. Chavan-Dafle, R. Holladay, and A. Rodriguez, "Planar in-hand manipulation via motion cones," *International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 163–182, 2020. DOI: 10.1177/0278364919880257. arXiv: arXiv:1810.00219v1.

[88] M. Costanzo, G. De Maria, and C. Natale, "Dual-arm in-hand manipulation with parallel grippers

using tactile feedback," in *International Conference on Advanced Robotics (ICAR)*, 2021, pp. 942–947.

[89] M. Costanzo, "Control of robotic object pivoting based on tactile sensing," *Mechatronics*, vol. 76, p. 102 545, 2021.

[90] J. Shi, J. Z. Woodruff, P. B. Umbanhowar, and K. M. Lynch, "Dynamic In-Hand Sliding Manipulation," *IEEE Transactions on Robotics*, vol. 33, no. 4, pp. 778–795, 2017. DOI: 10.1109/TRO.2017.2693391.

[91] Y. Chen, C. Prepscius, D. Lee, and D. D. Lee, "Tactile velocity estimation for controlled in-grasp sliding," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1614–1621, 2021.

[92] S. Cruciani, C. Smith, D. Kragic, and K. Hang, "Dexterous manipulation graphs," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 2040–2047.

[93] K. Nagata, "Manipulation by a parallel-jaw gripper having a turntable at each fingertip," in *IEEE International Conference on Robotics and Automation*, 1994, 1663–1670 vol.2.

[94] H. Terasaki and T. Hasegawa, "Motion planning of intelligent manipulation by a parallel two-fingered gripper equipped with a simple rotating mechanism," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 2, pp. 207–219, 1998.

[95] J. Zhao, X. Jiang, X. Wang, S. Wang, and Y. Liu, "Assembly of randomly placed parts realized by using only one robot arm with a general parallel-jaw gripper," 2020, pp. 5024–5030.

[96] S. Zuo, J. Li, and M. Dong, "Design, modeling, and manipulability evaluation of a novel four-dof parallel gripper for dexterous in-hand manipulation," *Journal of Mechanical Science and Technology*, vol. 35, pp. 1–16, 2021.

[97] J. Chapman, G. Gorjup, A. Dwivedi, S. Matsunaga, T. Mariyama, B. MacDonald, and M. Liarokapis, "A locally-adaptive, parallel-jaw gripper with clamping and rolling capable, soft fingertips for fine manipulation of flexible flat cables," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 6941–6947.

[98] R. R. Ma and A. M. Dollar, "In-Hand Manipulation Primitives for a Minimal, Underactuated Gripper With Active Surfaces," in *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. Volume 5A: 40th Mechanisms and Robotics Conference, 2016.

[99] I. H. Taylor, N. Chavan-Dafle, G. Li, N. Doshi, and A. Rodriguez, "Pnugrip: An active two-phase gripper for dexterous manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9144–9150.

[100] N. Nahum and A. Sintov, "Robotic manipulation of thin objects within off-the-shelf parallel grippers with a vibration finger," *Mechanism and Machine Theory*, vol. 177, p. 105 032, 2022.

[101] J. Zhou, J. Yi, X. Chen, Z. Liu, and Z. Wang, "Bcl-13: A 13-dof soft robotic hand for dexterous grasping and in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3379–3386, 2018. DOI: 10.1109/LRA.2018.2851360.

[102] R. Deimel, "A Novel Type of Compliant , Underactuated Robotic Hand for Dexterous Grasping," in *Robotics: Science and Systems*, 2014, pp. 1–9.

[103] 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. DOI: 10.1109/LRA.2021.3098803.

[104] K. Batsuren and D. Yun, "Soft robotic gripper with chambered fingers for performing in-hand manipulation," *Applied Sciences*, vol. 9, no. 15, 2019. DOI: 10.3390/app9152967.

[105] A. M. Dollar and R. D. Howe, "The highly adaptive sdm hand: Design and performance evaluation," *The International Journal of Robotics Research*, vol. 29, no. 5, pp. 585–597, 2010.

[106] H. Liu, L. Zhao, B. Siciliano, and F. Ficuciello, "Modeling, optimization, and experimentation of the paragripper for in-hand manipulation without parasitic rotation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3011–3018, 2020. DOI: 10.1109/LRA.2020.2974419.

[107] L. U. Odhner and A. M. Dollar, "Dexterous manipulation with underactuated elastic hands," in *IEEE Int. Conf. on Rob. and Aut.*, 2011, pp. 5254–5260.

[108] L. U. Odhner and A. M. Dollar, "Stable, open-loop precision manipulation with underactuated hands," *The International Journal of Robotics Research*, vol. 34, no. 11, pp. 1347–1360, 2015.

[109] B. Calli and A. M. Dollar, "Vision-based precision manipulation with underactuated hands: Simple and effective solutions for dexterity," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 1012–1018.

[110] B. Calli, A. Kimmel, K. Hang, K. Bekris, and A. Dollar, "Path planning for within-hand manipulation over learned representations of safe states," in *International Symposium on Experimental Robotics*, Springer, 2018, pp. 437–447.

[111] I. Llop-Harillo, A. Pérez-González, J. Starke, and T. Asfour, "The anthropomorphic hand assessment protocol (ahap)," *Robotics and Autonomous Systems*, vol. 121, p. 103 259, 2019. DOI: https://doi.org/10.1016/j.robot.2019.103259.

[112] T. Okada, "Object-handling system for manual industry," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 2, pp. 79–89, 1979.

[113] S. Jacobsen, E. Iversen, D. Knutti, R. Johnson, and K. Biggers, "Design of the utah/m.i.t. dextrous hand," in *IEEE International Conference on*

Robotics and Automation, vol. 3, 1986, pp. 1520–1532. DOI: 10.1109/ROBOT.1986.1087395.

[114] W. Townsend, "The barretthand grasper – programmably flexible part handling and assembly," *Industrial Robot: An International Journal*, vol. 27, pp. 181–188, 2000. DOI: 10.1108/01439910010371597.

[115] J. Butterfass, M. Grebenstein, H. Liu, and G. Hirzinger, "Dlr-hand ii: Next generation of a dexterous robot hand," in *IEEE International Conference on Robotics and Automation*, vol. 1, 2001, 109–114 vol.1. DOI: 10.1109/ROBOT.2001.932538.

[116] T. Mouri, "Anthropomorphic robot hand : Gifu hand iii," in *International Conf. Control, Automation and Systems*, 2002, pp. 1288–1293.

[117] C. Lovchik and M. Diftler, "The robonaut hand: A dexterous robot hand for space," in *IEEE International Conference on Robotics and Automation*, vol. 2, 1999, 907–912 vol.2. DOI: 10.1109/ROBOT.1999.772420.

[118] S. Schulz, C. Pylatiuk, and G. Bretthauer, "A new ultralight anthropomorphic hand," in *IEEE International Conference on Robotics and Automation*, vol. 3, 2001, 2437–2441 vol.3. DOI: 10.1109/ROBOT.2001.932988.

[119] F. Lotti, P. Tiezzi, G. Vassura, L. Biagiotti, and C. Melchiorri, "Ubh 3: An anthropomorphic hand with simplified endo-skeletal structure and soft continuous fingerpads," in *IEEE International Conference on Robotics and Automation*, vol. 5, 2004, pp. 4736–4741. DOI: 10.1109/ROBOT.2004.1302466.

[120] N. Furukawa, A. Namiki, S. Taku, and M. Ishikawa, "Dynamic regrasping using a high-speed multifingered hand and a high-speed vision system," in *IEEE International Conference on Robotics and Automation*, 2006, pp. 181–187. DOI: 10.1109/ROBOT.2006.1641181.

[121] B. Sundaralingam and T. Hermans, "Geometric in-hand regrasp planning: Alternating optimization of finger gaits and in-grasp manipulation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 231–238.

[122] M. Pfanne, M. Chalon, F. Stulp, H. Ritter, and A. Albu-Schäffer, "Object-level impedance control for dexterous in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2987–2994, 2020. DOI: 10.1109/LRA.2020.2974702.

[123] G. Robles-De-La-Torre and V. Hayward, "Force can overcome object geometry in the perception of shape through active touch," *Nature*, vol. 412, pp. 445–448, 2001.

[124] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[125] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 3400–3407.

[126] M. Kalaitzakis, B. Cain, S. Carroll, A. Ambrosi, C. Whitehead, and N. Vitzilaios, "Fiducial markers for pose estimation: Overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers," *Journal of Intelligent & Robotic Systems*, vol. 101, 2021. DOI: 10.1007/s10846-020-01307-9.

[127] B. Calli and A. M. Dollar, "Vision-based model predictive control for within-hand precision manipulation with underactuated grippers," in *IEEE International Conference on Robotics and Automation*, 2017, pp. 2839–2845.

[128] M. Rad and V. Lepetit, "Bb8: A scalable, accurate, robust to partial occlusion method for predicting the 3d poses of challenging objects without using depth," in *IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 3848–3856. DOI: 10.1109/ICCV.2017.413.

[129] G. Billings and M. Johnson-Roberson, "Silhonet: An rgb method for 6d object pose estimation," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3727–3734, 2019. DOI: 10.1109/LRA.2019.2928776.

[130] M. Kokic, D. Kragic, and J. Bohg, "Learning to estimate pose and shape of hand-held objects from rgb images," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 3980–3987. DOI: 10.1109/IROS40897.2019.8967961.

[131] R. Higo, Y. Yamakawa, T. Senoo, and M. Ishikawa, "Rubik's cube handling using a high-speed multi-fingered hand and a high-speed vision system," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 6609–6614. DOI: 10.1109/IROS.2018.8593538.

[132] J. Ichnowski, Y. Avigal, J. Kerr, and K. Goldberg, "Dex-nerf: Using a neural radiance field to grasp transparent objects," *arXiv preprint arXiv:2110.14217*, 2021.

[133] C. Choi, J. Del Preto, and D. Rus, "Using vision for pre- and post-grasping object localization for soft hands," in *International Symposium on Experimental Robotics*, D. Kulić, Y. Nakamura, O. Khatib, and G. Venture, Eds., Cham, 2017, pp. 601–612.

[134] B. Wen, C. Mitash, S. Soorian, A. Kimmel, A. Sintov, and K. E. Bekris, "Robust, occlusion-aware pose estimation for objects grasped by adaptive hands," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 6210–6217. DOI: 10.1109/ICRA40945.2020.9197350.

[135] J. Carter and D. Fourney, "Research based tactile and haptic interaction guidelines," in *Proceedings of Guidelines On Tactile and Haptic Interactions*, 2005, pp. 84–92.

[136] J. Bimbo, S. Luo, K. Althoefer, and H. Liu, "In-hand object pose estimation using covariance-based

tactile to geometry matching," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 570–577, 2016. DOI: 10.1109/LRA.2016.2517244.

[137] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[138] H. Sun, K. Kuchenbecker, and G. Martius, "A soft thumb-sized vision-based sensor with accurate all-round force perception," *Nature Machine Intelligence*, vol. 4, pp. 135–145, 2022.

[139] I. H. Taylor, S. Dong, and A. Rodriguez, "Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2022, pp. 10 781–10 787.

[140] O. Azulay, N. Curtis, R. Sokolovsky, G. Levitski, D. Slomovik, G. Lilling, and A. Sintov, "Allsight: A low-cost and high-resolution round tactile sensor with zero-shot learning capability," *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 483–490, 2024.

[141] A. Delgado, C. Jara, and F. Torres, "In-hand recognition and manipulation of elastic objects using a servo-tactile control strategy," *Robotics and Computer-Integrated Manufacturing*, vol. 48, pp. 102–112, 2017. DOI: https://doi.org/10.1016/j.rcim.2017.03.002.

[142] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, *et al.*, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.

[143] J. He, S. Pu, and J. Zhang, "Haptic and visual perception in in-hand manipulation system," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 303–308. DOI: 10.1109/ROBIO.2015.7418784.

[144] P. Sodhi, M. Kaess, M. Mukadam, and S. Anderson, "Learning tactile models for factor graph-based estimation," in *IEEE International Conference on Robotics and Automation*, 2021.

[145] Z. Su, J. Fishel, T. Yamamoto, and G. Loeb, "Use of tactile feedback to control exploratory movements to characterize object compliance," *Frontiers in neurorobotics*, vol. 6, p. 7, 2012.

[146] J. Bimbo, L. D. Seneviratne, K. Althoefer, and H. Liu, "Combining touch and vision for the estimation of an object's pose during manipulation," in *IEEE/RSJ Int. Conf. on Intel. Rob. & Sys*, 2013.

[147] J. Tegin and J. Wikander, "Tactile sensing in intelligent robotic manipulation—a review," *Indus. Robot: An Int. Jou.*, vol. 32, 2005. DOI: 10.1108/01439910510573318.

[148] M.-Y. Cheng, C.-M. Tsao, Y.-T. Lai, and Y.-J. Yang, "A novel highly-twistable tactile sensing array using extendable spiral electrodes," in *IEEE International Conference on Micro Electro Mechanical Systems*, 2009, pp. 92–95.

[149] N. Wettels, J. A. Fishel, Z. Su, C. H. Lin, G. E. Loeb, and L. SynTouch, "Multi-modal synergistic tactile sensing," in *Tactile sensing in humanoids—Tactile sensors and beyond workshop, 9th IEEE-RAS international conference on humanoid robots*, 2009.

[150] M. C. Koval, M. R. Dogar, N. S. Pollard, and S. S. Srinivasa, "Pose estimation for contact manipulation with manifold particle filters," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys.*, 2013, pp. 4541–4548. DOI: 10.1109/IROS.2013.6697009.

[151] N. Lepora and J. Lloyd, "Optimal deep learning for robot touch: Training accurate pose models of 3d surfaces and edges," *IEEE Robotics and Automation Magazine*, vol. 27, no. 2, 2020.

[152] E. Psomopoulou, N. Pestell, F. Papadopoulos, J. Lloyd, Z. Doulgeri, and N. F. Lepora, "A robust controller for stable 3d pinching using tactile sensing," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8150–8157, 2021.

[153] R. Bellman, "A markovian decision process," *Ind. Uni. Math. J.*, vol. 6, pp. 679–684, 4 1957, ISSN: 0022-2518.

[154] R. Ozawa, S. Arimoto, S. Nakamura, and J.-H. Bae, "Control of an object with parallel surfaces by a pair of finger robots without object sensing," *IEEE Transactions on Robotics*, vol. 21, no. 5, pp. 965–976, 2005.

[155] A. S. Morgan, K. Hang, W. G. Bircher, and A. M. Dollar, "A data-driven framework for learning dexterous manipulation of unknown objects," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 8273–8280.

[156] V. Fonseca, T. E. Alves de Oliveira, and E. Petriu, "Estimating the orientation of objects from tactile sensing data using machine learning methods and visual frames of reference," *Sensors*, p. 2285, 2019.

[157] A. Sintov, A. Kimmel, B. Wen, A. Boularias, and K. Bekris, "Tools for data-driven modeling of within-hand manipulation with underactuated adaptive hands," in *Proceedings of the Conf. on Learning for Dyn. & Cont.*, vol. 120, Berkeley, CA, USA, 2020, pp. 771–780.

[158] A. Zeng, S. Song, J. Lee, A. Rodriguez, and T. Funkhouser, "Tossingbot: Learning to throw arbitrary objects with residual physics," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1307–1319, 2020. DOI: 10.1109/TRO.2020.2988642.

[159] Z. Wang, C. R. Garrett, L. P. Kaelbling, and T. Lozano-Pérez, "Active model learning and diverse action sampling for task and motion planning," in *IEEE/RSJ Int. Conf. on Int. Rob. & Sys.*, 2018, pp. 4107–4114.

[160] B. Singh, R. Kumar, and V. P. Singh, "Reinforcement learning in robotic applications: A compre-

hensive survey," *Artif. Intell. Rev.*, vol. 55, no. 2, pp. 945–990, 2022. DOI: 10.1007/s10462-021-09997-9.

[161] H. Nguyen and H. M. La, "Review of Deep Reinforcement Learning for Robot Manipulation," in *IEEE International Conference on Robotic Computing (IRC)*, 2019, pp. 590–595. DOI: 10.1109/IRC.2019.00120.

[162] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[163] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, p. 529, 2015.

[164] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, 2016.

[165] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[166] F. Pourpanah, M. Abdar, Y. Luo, X. Zhou, R. Wang, C. P. Lim, X.-Z. Wang, and Q. M. J. Wu, "A review of generalized zero-shot learning methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4051–4070, 2023. DOI: 10.1109/TPAMI.2022.3191696.

[167] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, *et al.*, "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Conference on Robot Learning*, PMLR, 2018, pp. 651–673.

[168] H. Zhu, A. Gupta, A. Rajeswaran, S. Levine, and V. Kumar, "Dexterous manipulation with deep reinforcement learning: Efficient, general, and low-cost," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 3651–3657.

[169] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: A survey," in *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 737–744.

[170] S. Höfer, K. Bekris, A. Handa, J. C. Gamboa, M. Mozifian, F. Golemo, C. Atkeson, D. Fox, K. Goldberg, J. Leonard, *et al.*, "Sim2real in robotics and automation: Applications and challenges," *IEEE Transactions on automation, Science and engineering*, vol. 18, no. 2, pp. 398–400, 2021.

[171] M. Liarokapis and A. M. Dollar, "Post-contact, in-hand object motion compensation for compliant and underactuated hands," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 986–993.

[172] J. Van Baar, A. Sullivan, R. Cordorel, D. Jha, D. Romeres, and D. Nikovski, "Sim-to-real transfer learning using robustified controllers in robotic tasks involving complex dynamics," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 6001–6007.

[173] V. Makoviychuk *et al.*, *Isaac gym: High performance gpu-based physics simulation for robot learning*, 2021.

[174] B. Eysenbach, S. Gu, J. Ibarz, and S. Levine, "Leave no trace: Learning to reset for safe and autonomous reinforcement learning," in *International Conference on Learning Representations (ICLR)*, OpenReview. net, 2018.

[175] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *International Conference on Learning Representations (ICLR)*, 2016.

[176] M. Zimmer, P. Viappiani, and P. Weng, "Teacher-student framework: A reinforcement learning approach," in *AAMAS Workshop Autonomous Robots and Multirobot Systems*, 2014.

[177] D. Morrison, P. Corke, and J. Leitner, "Egad! an evolved grasping analysis dataset for diversity and reproducibility in robotic manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4368–4375, 2020.

[178] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The ycb object and model set: Towards common benchmarks for manipulation research," in *International Conference on Advanced Robotics (ICAR)*, IEEE, 2015, pp. 510–517.

[179] A. S. Morgan, D. Nandha, G. Chalvatzaki, C. D'Eramo, A. M. Dollar, and J. Peters, "Model predictive actor-critic: Accelerating robot skill acquisition with deep reinforcement learning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 6672–6678.

[180] X. Wang, Y. Chen, and W. Zhu, "A survey on curriculum learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[181] A. Melnik, L. Lach, M. Plappert, T. Korthals, R. Haschke, and H. Ritter, "Tactile sensing and deep reinforcement learning for in-hand manipulation tasks," in *IROS Workshop on Autonomous Object Manipulation*, 2019.

[182] S. Funabashi, T. Isobe, S. Ogasa, T. Ogata, A. Schmitz, T. P. Tomo, and S. Sugano, "Stable in-grasp manipulation with a low-cost robot hand by using 3-axis tactile sensors with a cnn," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9166–9173.

[183] Y. Chen, Y. Geng, F. Zhong, J. Ji, J. Jiang, Z. Lu, H. Dong, and Y. Yang, "Bi-dexhands: Towards human-level bimanual dexterous manipulation," *IEEE Transactions on Pattern Analysis and*

Machine Intelligence, pp. 1–15, 2023. DOI: 10.1109/TPAMI.2023.3339515.

[184] C. Rapisarda, *Reinforcement learning for dexterity transfer between manipulators*, 2019.

[185] Y. Duan, M. Andrychowicz, B. Stadie, O. Jonathan Ho, J. Schneider, I. Sutskever, P. Abbeel, and W. Zaremba, "One-shot imitation learning," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30, Curran Associates, Inc., 2017.

[186] B. Fang, S. Jia, D. Guo, M. Xu, S. Wen, and F. Sun, "Survey of imitation learning for robotic manipulation," *International Journal of Intelligent Robotics and Applications*, vol. 3, no. 4, pp. 362–369, 2019.

[187] B. Zheng, S. Verma, J. Zhou, I. W. Tsang, and F. Chen, "Imitation learning: Progress, taxonomies and challenges," *IEEE Transactions on Neural Networks and Learning Systems*, no. 99, pp. 1–16, 2022.

[188] S. Arora and P. Doshi, "A survey of inverse reinforcement learning: Challenges, methods and progress," *Artificial Intelligence*, vol. 297, p. 103 500, 2021. DOI: https://doi.org/10.1016/j.artint.2021.103500.

[189] A. Gupta, C. Eppner, S. Levine, and P. Abbeel, "Learning dexterous manipulation for a soft robotic hand from human demonstrations," in *IEEEqRSJ International Conference on Intelligent Robots and Systems*, 2016, pp. 3786–3793. DOI: 10.1109/IROS.2016.7759557.

[190] T. Gašpar, B. Nemec, J. Morimoto, and A. Ude, "Skill learning and action recognition by arc-length dynamic movement primitives," *Robotics and Autonomous Systems*, vol. 100, pp. 225–235, 2018, ISSN: 0921-8890. DOI: https://doi.org/10.1016/j.robot.2017.11.012.

[191] 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.

[192] D. Wei and H. Xu, *A wearable robotic hand for hand-over-hand imitation learning*, 2023. arXiv: 2309.14860 [cs.RO].

[193] A. Hussein, M. M. Gaber, E. Elyan, and C. Jayne, "Imitation learning: A survey of learning methods," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.

[194] 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*, G. Gordon, D. Dunson, and M. Dudík, Eds., ser. Proceedings of Machine Learning Research, vol. 15, 2011, pp. 627–635.

[195] Q. Cai, M. Hong, Y. Chen, and Z. Wang, *On the global convergence of imitation learning: A case for linear quadratic regulator*, 2019. arXiv: 1901.03674 [cs.LG].

[196] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2013. DOI: 10.1162/NECO_a_00393.

[197] J. Ho and S. Ermon, "Generative adversarial imitation learning," *CoRR*, vol. abs/1606.03476, 2016. arXiv: 1606.03476.

[198] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020. DOI: 10.1145/3422622.

[199] M. Monastirsky, O. Azulay, and A. Sintov, "Learning to throw with a handful of samples using decision transformers," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 576–583, 2023. DOI: 10.1109/LRA.2022.3229266.

[200] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," in *Robotics: Science and Systems (RSS)*, 2023.

[201] S. Wang, M. Lambeta, P.-W. Chou, and R. Calandra, "Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors," *IEEE Rob. & Aut. Let.*, vol. 7, pp. 3930–3937, 2022.