Survey of learning-based approaches for robotic in-hand manipulation

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

[1]AI-WEINBERG AI Experts, Tel-Aviv, Israel, [2]Faculty of Mechanical Engineering, Technion - Israel Institute of Technology, Haifa, Israel, [3]School of Mechanical Engineering, Tel-Aviv University, Tel-Aviv, Israel
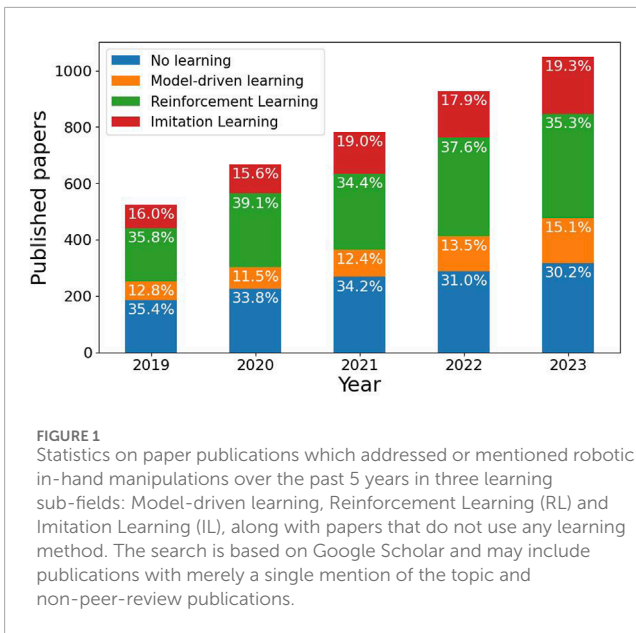
Human dexterity is an invaluable capability for precise manipulation of objects in complex tasks. The capability of robots to similarly grasp and perform in-hand manipulation of objects is critical for their use in the ever changing human environment, and for their ability to replace manpower. In recent 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.

KEYWORDS

in-hand manipulation, dexterous manipulation, model learning, reinforcement learning, imitation learning

1 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 (Lehman et al., 2010), assembly in production lines (Kang et al., 2021), and robotic assistance for sick and disabled users (Petrich et al., 2021). During the COVID-19 pandemic, there has been a significant need for autonomous and complex

FIGURE 1
Statistics on paper publications which addressed or mentioned robotic in-hand manipulations over the past 5 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 publications with merely a single mention of the topic and non-peer-review publications.

robot manipulators (Kroemer et al., 2021). 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 (Sun et al., 2021).

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. These approaches have significant implications for in-hand manipulation, as they each offer unique advantages in improving robotic capabilities. They offer complementary contributions to improving robotic in-hand manipulation. Model-based methods provide a foundation for understanding system dynamics, RL enables self-learning of optimal policies, and IL allows for learning from human expertise. Figure 1 shows the increase in paper publications over the past 5 years with regard to these three subfields. We note that the search is based on Google Scholar and aims only to show a trend. Results may include publications with merely a single mention of the topic without actual scientific contribution and non-peer-review publications. In this study, we survey in-hand manipulation approaches, tasks and applications that use either of these subfields with substantial contribution to the topic.

Previous surveys focused on specific aspects of robotic manipulation such as the use of contact (Suomalainen et al., 2022), space applications (Papadopoulos et al., 2021), handling of deformable objects (Herguedas et al., 2019), multi-robot systems (Feng et al., 2020) and manipulation in cluttered environments (Mohammed et al., 2022). Some other surveys discussed learning approaches for general manipulation such as imitation learning (Fang et al., 2019), deep-learning (Han et al., 2023) and general trends (Billard and Kragic, 2019; Kroemer et al., 2021; Cui and Trinkle, 2021). However, 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 1 provides a summary of prominent state-of-the-art work including key properties. These properties will be defined, introduced and discussed in the next section which provides an overview of in-hand manipulation. 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 2). 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 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.

2 Overview on in-hand manipulations

Robotic in-hand manipulation involves physical interaction between a robotic end-effector, an object and often with the environment (Cruciani et al., 2020). 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.

2.1 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 (Okamura et al., 2000). Dexterous in-hand manipulation is, therefore, the manipulation of an

FIGURE 2
Taxonomy of robotic in-hand manipulation.
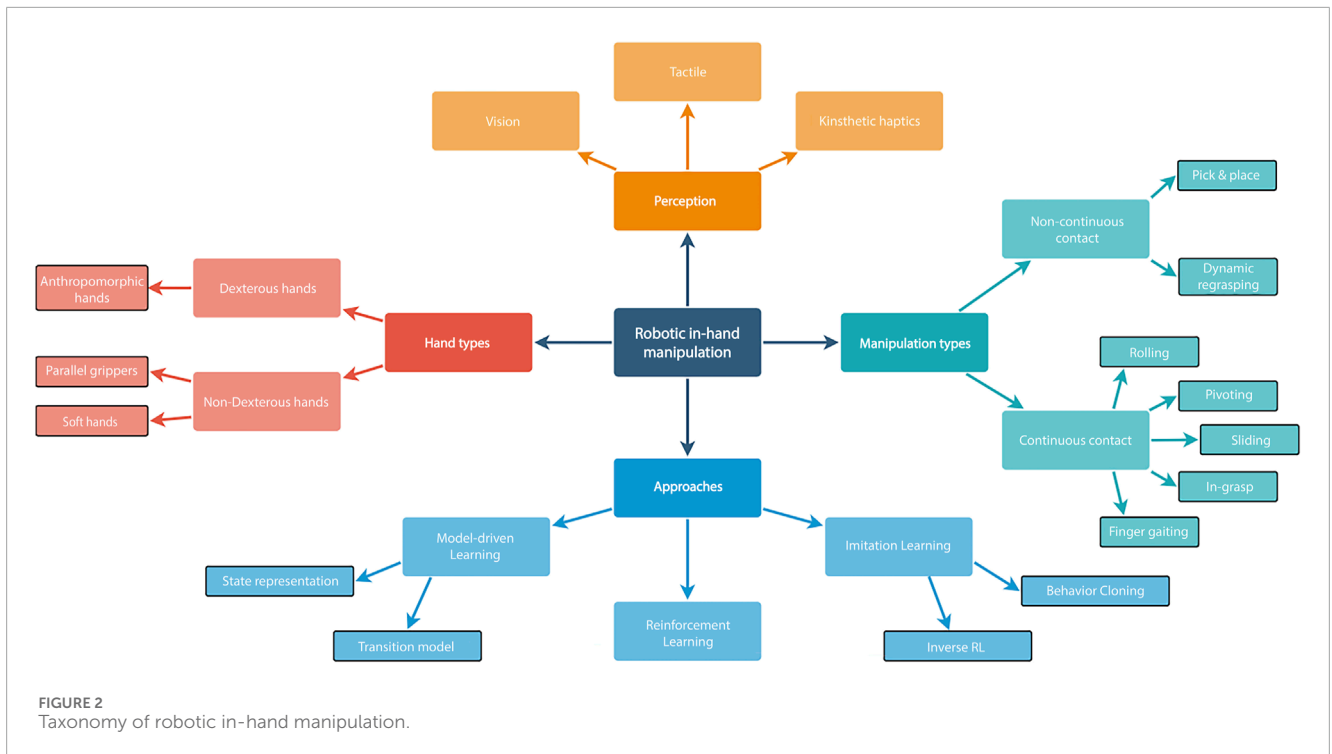
object in the hand by using its own mechanics (Mason and Salisbury, 1985). Naturally, dexterous in-hand manipulation requires a high number of DOF and includes, in most cases, anthropomorphic hands. Contrary to dexterous hands, non-dexterous ones have a low number of DOF and, thus, do not have an intrinsic capability to manipulate objects by themselves and require some extrinsic involvement.

2.2 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.

2.2.1 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.

- *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 (Han et al., 1997). 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 (Sintov and Shapiro, 2016; Bhatt et al., 2021). 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 (Viña B. et al., 2015), initiating external contact (Dafle et al., 2014) or generating dynamic motions of the robotic arm (Sintov et al., 2016; Cruciani and Smith, 2017).

- *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 (Shi et al., 2017a). The object slides along the links of the hand to a desired position due to forces from the fingers or external forces (Shi et al., 2020).

- *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 (Ma and Dollar, 2011). During manipulation, finger contacts and a stable grasp are maintained. However, sliding or contact rolling may occur (Calli et al., 2018b).

- *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 (Rus, 1999; Xu et al., 2007; Fan et al., 2017; Bhatt et al., 2021). 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

TABLE 1 Summary of state-of-the-art work on learning approaches for in-hand manipulation (in alphabetical order).

Paper	Learning type			Manipulation	Dex./Non-Dex	Rigid/Soft	Perception		
	Model	RL	IL	type	hand	hand	Vision	Tactile	Kines
Allshire et al. (2022)		✓		In-Grasp	Dex	Rigid	✓		✓
Antonova et al. (2017)		✓		Pivoting	Non-Dex	Rigid	✓		
OpenAI et al. (2019); Andrychowicz et al. (2020)		✓		In-Grasp	Dex	Rigid	✓		✓
Arunachalam et al. (2023b)			✓	Mix	Dex	Rigid	✓		
Azulay et al. (2022a)	✓			In-grasp	Non-Dex	Soft		✓	✓
Azulay et al. (2022b)		✓		In-Grasp	Non-Dex	Soft			✓
Calli et al. (2018b)	✓			In-Grasp	Non-Dex	Soft	✓		✓
Chen et al. (2021a, 2022)		✓		Mix	Dex	Rigid	✓		
Chen et al. (2023a)		✓		F. Gait	Dex	Rigid	✓		✓
Cruciani and Smith (2017, 2018)		✓		Pivoting	Non-Dex	Rigid	✓		
Deng and Zhang (2020)			✓	Mix	Dex	Rigid	✓		✓
Dimou et al. (2023)	✓			In-Grasp	Dex	Rigid	✓		
Falco et al. (2018)		✓		In-Grasp	Dex	Soft	✓	✓	
Funabashi et al. (2019a, 2020b, 2019b)		✓		Rolling	Dex	Rigid		✓	✓
Garcia-Hernando et al. (2020)			✓	Mix	Dex	Rigid	✓		
Gupta et al. (2021)		✓		Mix	Dex	Rigid	✓		
Handa et al. (2023)		✓		Mix	Dex	Rigid	✓		
Huang et al. (2021)		✓		F. Gait	Dex	Rigid	✓		
Jain et al. (2019)		✓	✓	In-Grasp	Dex	Rigid	✓	✓	✓
Khandate et al. (2022, 2023)		✓		F. Gait	Dex	Rigid		✓	✓

TABLE 1 (*Continued*) Summary of state-of-the-art work on learning approaches for in-hand manipulation (in alphabetical order).

Paper	Learning type			Manipulation	Dex./Non-Dex	Rigid/Soft	Perception		
	Model	RL	IL	type	hand	hand	Vision	Tactile	Kines
Kimmel et al. (2019)	✓			In-Grasp	Non-Dex	Soft	✓		✓
Korthals et al. (2019)		✓		Mix	Dex	Rigid	✓	✓	✓
Kumar et al. (2016)			✓	Mix	Dex	Rigid		✓	✓
Li et al. (2014)			✓	In-grasp	Dex	Rigid	✓	✓	✓
Li et al. (2020)	✓	✓		In-Grasp	Dex	Rigid	✓		✓
Luo et al. (2023)	✓			Mix	Dex	Rigid	✓		✓
Melnik et al. (2021)		✓		Mix	Dex	Rigid	✓	✓	✓
Morgan et al. (2020)	✓			In-Grasp	Non-Dex	Soft	✓		
Morgan et al. (2021b)	✓	✓		Mix	Non-Dex	Soft	✓		
Nagabandi et al. (2020)	✓	✓		Mix	Dex	Rigid	✓		
Orbik et al. (2021)			✓	Mix	Dex	Rigid	✓		✓
Park et al. (2024)	✓			In-Grasp	Non-Dex	Soft		✓	
Pitz et al. (2023)		✓		In-Grasp	Dex	Rigid			✓
Qi et al. (2022)		✓		F. Gait	Dex	Rigid			✓
Qi et al. (2023)		✓		F. Gait	Dex	Rigid	✓	✓	✓
Radosavovic et al. (2021)			✓	Mix	Dex	Rigid	✓		
Rajeswaran et al. (2017)			✓	Mix	Dex	Rigid	✓		✓
Shin and Jeon (2024)			✓	Rolling	Dex	Rigid	✓	✓	✓
Sievers et al. (2022)		✓		In-Grasp	Dex	Rigid			✓
Sintov et al. (2019, 2020a)	✓			In-Grasp	Non-Dex	Soft	✓		✓
Solak and Jamone (2019)			✓	In-Grasp	Dex	Rigid			✓
Solak and Jamone (2023)			✓	In-Grasp	Dex	Rigid		✓	✓
Srinivasan et al. (2020)		✓		In-Grasp	Dex	Rigid	✓		✓

(Continued on the following page)

TABLE 1 (*Continued*) Summary of state-of-the-art work on learning approaches for in-hand manipulation (in alphabetical order).

Paper	Learning type			Manipulation	Dex./Non-Dex	Rigid/Soft	Perception		
	Model	RL	IL	type	hand	hand	Vision	Tactile	Kines
Tao et al. (2023)		✓		Mix	Dex	Rigid	✓		
Toskov et al. (2023)	✓			Pivoting	Non-Dex	Rigid		✓	
Toledo et al. (2021)		✓		Pivoting	Non-Dex	Rigid	✓		
Van Hoof et al. (2015)		✓		In-Grasp	Non-Dex	Soft		✓	
Veiga et al. (2020)		✓		In-Grasp	Dex	Rigid		✓	
Wang et al. (2020)	✓			Pivoting	Non-Dex	Rigid			✓
Wei et al. (2023)			✓	Sliding	Dex	Rigid	✓		
Yang et al. (2023)		✓		Mix	Dex	Rigid		✓	
Yin et al. (2023)		✓		Mix	Dex	Rigid		✓	✓
Yuan et al. (2020)			✓	Rolling	Non-Dex	Rigid	✓		✓
Yuan et al. (2023)		✓		Mix	Dex	Rigid	✓	✓	✓

requires sufficiently many DOFs to manipulate the grasped object between two grasp configurations while maintaining stable grasps.

2.2.2 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 and then picked up again at a new grasp configuration (Lozano-Perez et al., 1987; Tournassoud et al., 1987). 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 (Sintov and Shapiro, 2017). 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.

2.3 In-hand manipulation with non-dexterous hands

2.3.1 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 (Guo et al., 2017). 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 (Tournassoud et al., 1987). In pick-and-place, the object is placed on a surface and picked up again in a different grasp configuration (Zeng et al., 2018). 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 (Ma and Dollar, 2011; Billard and Kragic, 2019). 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 (Dafle et al., 2014; Chavan-Dafle et al., 2020) or performing dynamic manipulation. For instance, pivoting can be done by intrinsic slippage control or extrinsic dynamic manipulation of the arm (Viña B. et al., 2015; Sintov et al., 2016; Cruciani and Smith, 2018). Slippage control leverages gravity and tunes finger contact force of the parallel gripper (Costanzo et al., 2021). Costanzo
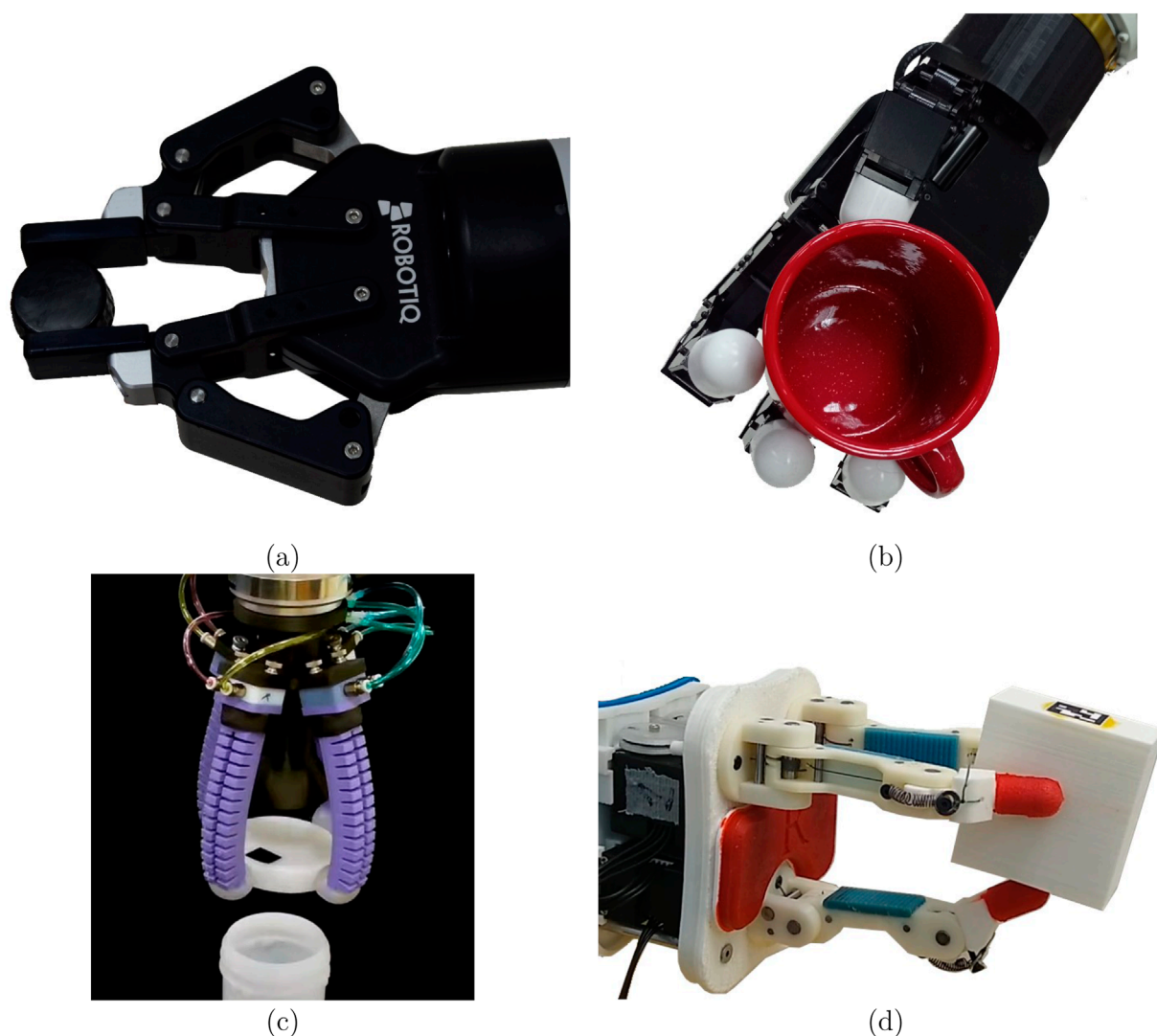
FIGURE 3
Various dexterous and non-dexterous hands. **(A)** Non-dexterous parallel jaw gripper model 2F-85 by Robotiq. The gripper has only a single DOF for opening and closing on an object. **(B)** The four-finger dexterous and anthropomorphic Allegro hand with 16 DOF. **(C)** A four finger non-dexterous soft hand operated by pneumatic bending actuators (Abondance et al., 2020). **(D)** Underactuated compliant hand model-O from the Yale OpenHand project (Ma and Dollar, 2017). Images **(A, B, D)** were taken by the authors.

(2021) exploited a dual-arm system and tactile feedback to allow controlled slippage between the object and parallel grippers. The work of Shi et al. (2017b) controlled the force distribution of a pinch grasp to predict sliding directions. Similarly, Chen Y. et al. (2021) 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 (Cruciani et al., 2018). While jaw grippers have only one DOF, some work has been done to augment their intrinsic manipulation capabilities. These robotic hands, equipped with additional functionalities beyond the traditional single DOF parallel gripper, can no longer be categorized as simple grippers. Seminal work by Nagata (1994) 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 (Terasaki and Hasegawa, 1998). Zhao J. et al. (2020) augmented a jaw gripper tip with a two DOF transmission mechanism to re-orient and translate randomly placed screws. Zuo et al. (2021) added a linear actuation along each of the two fingers to enable translation and twist of a grasped object. Similarly, a rolling mechanism was added to the gripper by Chapman et al. (2021) 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 (Ma and Dollar, 2016). Taylor et al. (2020) included a pneumatic braking mechanism 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 (Nahum and Sintov, 2022). These innovative designs offer enhanced in-hand manipulation capabilities without requiring complex additional controls or software, opening up new avenues for research and development within the field of in-hand manipulation. Traditional control and motion planning methods for these systems often lack the flexibility to generalize to diverse tasks and objects. Learning-based approaches offer a promising solution for enhancing the capabilities of these mechanisms.

2.3.2 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 (Zhou et al., 2018). 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, under-actuated anthropomorphic robotic hand (Deimel, 2014). 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) (Abondance et al., 2020). Another pneumatic hand with reconfigurable fingers and an active palm was designed to enable in-hand dexterity while maintaining low mechanical complexity (Pagoli et al., 2021). Batsuren and Yun (2019) presented a soft robotic gripper for grasping various objects by mimicking in-hand manipulation. 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 (Dollar and Howe, 2010; Liu et al., 2020). 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 (Odhner and Dollar, 2011). 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 (Ma and Dollar, 2017). Along with good grasping capabilities, precise in-grasp manipulation was shown possible (Odhner and Dollar, 2015). Using visual servoing along with a linear approximation of the hand kinematics, closed-loop control of a two-finger hand was demonstrated (Calli and Dollar, 2016) and later used to track paths planned with an optimization-based model-free planner (Calli et al., 2018a). However, a precise analytical model for soft hands is not easy to acquire due to the compliance and inherent fabrication uncertainties. Therefore, data-based models were proposed and will be discussed later.

2.4 In-hand manipulation with dexterous hands

Mason and Salisbury (1985) 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 (Llop-Harillo et al., 2019) (Figure 3B). Early work on dexterous anthropomorphic hands includes a three-finger and 11-DOF hand (Okada, 1979), the four-finger Utah/MIT hand (Jacobsen et al., 1986), and later the Barrett and DLR hands (Townsend, 2000; Butterfass et al., 2001). 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 (Mouri, 2002). On the other hand, the Robotnaut hand was designed for space usage and included flex shafts for bending the fingers (Lovchik and Diftler, 1999). A hand from Karlsruhe used 13 flexible fluidic actuators for a lightweight design (Schulz et al., 2001). The UB hand is a five-finger anthropomorphic hand that used elastic hinges to mimic human motion (Lotti et al., 2004). Beyond anthropomorphic designs, few non-anthropomorphic dexterous hands have been proposed, incorporating multiple fingers in various designs (Hammond et al., 2012). However, most attempts to design a non-anthropomorphic multi-finger hand adhere to under-actuation, limiting their dexterity (Molnar and Menguc, 2022).

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 by Furukawa et al. (2006) 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 (Sundaralingam and Hermans, 2018). Recent work by Pfanne et al. (2020) used impedance control for stable finger gaiting over various objects with a dexterous multi-finger hand.

Multi-finger anthropomorphic hands are commonly employed in the development of bionic prostheses as they resemble the human hand (Cordella et al., 2016). They are usually operated by Electromyographic (EMG) signals to reduce the cognitive burden on the user (Starke et al., 2022). While these hands are often highly dexterous and have multimodal information from various sensors (Stefanelli et al., 2023), their use is commonly limited to pick and place tasks (Marinelli et al., 2022). Hence, the learning methods explored in this paper offer potential avenues for advancing the capabilities of various hands including prosthetic ones with in-hand manipulation tasks.

2.5 Perception

Humans use both visual feedback and touch perception for interacting with the environment and, in particular, manipulate objects within their hand (Robles-De-La-Torre and Hayward, 2001). Such sensory modules have been widely explored in robotics, both individually and combined.

2.5.1 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 (Garrido-Jurado et al., 2014), AprilTags (Olson, 2011) or reflective markers for a Motion Capture system (MoCap) Azulay et al. (2022a). These markers provide instant pose recognition of a rigid object without the need for its geometry recognition (Kalaitzakis et al., 2021). However, the requirement to apply them on an object prevents spontaneous unplanned interaction with an object. Specifically with reflective markers, the work is limited to a room or lab where the MoCap system is located. In general, vision-based 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 (Calli and Dollar, 2017) and hand state representation (Sintov et al., 2019) during in-grasp manipulation of an object with an underactuated hand.

While fiducial markers offer immediate pose estimation, their reliance on predefined visual patterns limits their applicability in real-world environments. To address this, visual perception, combined with learning-based methods, is often employed for robust object recognition and pose estimation. 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 (Rad and Lepetit, 2017; Billings and Johnson-Roberson, 2019; Kokic et al., 2019). 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 by Furukawa et al. (2006) 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 (Higo et al., 2018). A work by OpenAI used three RGB cameras to train a model for pose estimation of a cube manipulated by the Shadow hand (Andrychowicz et al., 2020). Ichnowski et al. (2021) presented Dex-NeRF, a novel approach that enables grasping using Neural Radiance Field (NeRF) technique. NeRF receives five-dimensional vectors as 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. Commonly used depth cameras include Intel's RealSense and StereoLab's ZED, where the latter leverages GPU capabilities for advanced spatial perception. For instance, an RGB-D camera was used to estimate the pose of objects before and during grasp by a soft hand (Choi et al., 2017). Similar work involved a depth camera to demonstrate robust pose estimation of objects grasped and partly occluded by a two-finger underactuated hand (Wen et al., 2020). Although visual perception can provide an 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.

2.5.2 Haptics

Information from haptic sensors is acquired through direct contact with objects by either tactile sensing (Yousef et al., 2011) or internal sensing of joint actuators known as Kinesthetic (or Proprioception) haptics (Carter and Fourney, 2005). 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 (Bimbo et al., 2016) or high-resolution optical sensors (Yuan et al., 2017; Sun et al., 2022). 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 (Delgado et al., 2017). Optical tactile sensors work by projecting a pattern of light onto a surface and observing the distortion of that pattern caused by contact using an internal camera. Different sensors utilize different cameras with sensing resolution of up to 2592×1944. This distortion provides information about the shape, texture, and pressure applied to the surface (Taylor et al., 2022; Azulay et al., 2024). Lambeta et al. (2020)used these sensors to learn in-hand manipulation models. 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 (He et al., 2015). While haptics provides valuable information regarding the state of contact with the environment, traditional analytical methods are often insufficient for processing this data. Hence, learning-based approaches have emerged as a promising solution for extracting meaningful information from haptic sensor data.

2.6 Simulation of in-hand manipulation

Simulating in-hand and dexterous manipulation is a critical aspect of robotic research, offering a controlled environment for developing and testing advanced control algorithms. High-fidelity simulators like MuJoCo (Todorov et al., 2012) and Isaac Gym (Makoviychuk et al., 2021) allow researchers to model complex interactions between robotic hands and objects, enabling the study of tasks such as reorienting a cube (Andrychowicz et al., 2017), opening doors (Rajeswaran et al., 2017) or dynamically adjusting grasps on irregular objects (Agarwal et al., 2023). For instance, MuJoCo's ability to model soft contacts and Isaac Gym's high-speed parallel simulations make them valuable tools for training and evaluating robotic manipulation strategies. The use of simulations in dexterous manipulation research is invaluable. It enables large-scale experimentation and rapid iteration, eliminating the risks or costs associated with physical testing. Researchers can explore complex manipulation tasks with multimodal sensing, including tactile and visual inputs (Yuan et al., 2023), in a controlled and scalable setting. Ultimately, these simulations drive the development of more adaptive robotic systems capable of human-like dexterity in unstructured environments.

Despite their advantages, these simulators face significant challenges in accurately replicating real-world physics, particularly in modeling friction, soft deformation and contact forces (Haldar et al., 2023). This sim-to-real gap can result in behaviors learned in simulation failing to transfer seamlessly to physical robots due to unmodeled dynamics and sensor noise. Furthermore, while rigid body dynamics are often well-represented, simulators struggle with soft materials and deformable objects, which is crucial for tasks like manipulating cloth or delicate items. More specifically, the simulation of underactuated hands is still a challenge.

2.7 Datasets of in-hand manipulation motions

Learning models require a significant amount of data to achieve sufficient accuracy. Data in many applications is inherently high-dimensional, often consisting of multimodal signals like visual and haptic data. Simulators, such as mentioned above, provide an environment to collect such data. However, the reality gap is often too large and the acquisition of real-world data is necessary. However, acquiring the data may be exhausting, expensive and even dangerous. Hence, practitioners often disseminate their collected data for the benefit of the community and for potential benchmarking (Khazatsky et al., 2024). For example, RealDex is a dataset focused on capturing authentic dexterous hand motions with human behavioral patterns based on tele-operation (Liu et al., 2024). The RUM dataset includes data of real in-hand manipulation of various objects with adaptive hands (Sintov et al., 2020b). A prominent dataset is the YCB object and model set (Calli et al., 2015a), aimed to provide a standard set of object for benchmarking general manipulation tasks including in-hand ones (Cruciani et al., 2020). Some datasets are simulation based such as the DexHand (Nematollahi et al., 2022) where the data is comprised of RGB-D images of a Shadow Hand robot manipulating a cube. Overall, publicly available datasets are an important tool to promote standardized objects, tasks and evaluation metrics to benchmark and compare different approaches to robotic in-hand manipulation.

3 Model-driven learning for in-hand manipulation

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.

3.1 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 (Azulay et al., 2022a; Andrychowicz et al., 2020; Funabashi et al., 2019b; Sodhi et al., 2021; Dimou et al., 2023). It is also effective in combining data from multiple sensors or information sources (Qi et al., 2022), and is often used by robots to merge information from different modalities, such as vision and haptic feedback (Andrychowicz et al., 2020). Without a compact and meaningful representation of the object-hand state, the robot may struggle to perform successful and efficient manipulations (Azulay et al., 2022a).

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 (Su et al., 2012). Often, haptic perception is used alongside vision to refine initial pose estimation (Bimbo et al., 2013). Contact sensing is the common approach for pose estimation during manipulation (Azulay et al., 2022a). Such sensing has been generally achieved using simple force or pressure sensors (Tegin and Wikander, 2005; Cheng et al., 2009; Wettels et al., 2009). As such, Koval et al. (2013) used contact sensors and particle filtering to estimate the pose of an object during contact manipulation. Park et al. (2024) used soft sensors in a pneumatic finger and a neural-network to estimate the angles of the finger. 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 (Bimbo et al., 2016). 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. (2021) used data from an optical-based tactile sensor to estimate the pose of an object being pushed, while others (Lepora and Lloyd, 2020; Psomopoulou et al., 2021) explored the use of these sensors for estimating the relative pose of an object during a grasp. In a work by Wang et al. (2020), 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. (2020) 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. (2023) 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. (2019b) 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.

3.2 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) (Bellman, 1957). 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 (Nagabandi et al., 2020). While a hand transition model is generally available analytically in rigid hands where the kinematics are known (Ozawa et al., 2005), 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. (2019) 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 (Kimmel et al., 2019; Sintov et al., 2020a). Recently, Morgan et al. (2020) 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 (Morgan et al., 2019). The work by Wen et al. (2020) 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 (Morgan et al., 2021b). 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. (2018b) 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 integrated allocentric visual perception along with four tactile modules, that combine pressure, magnetic, angular velocity and gravity sensors, on two underactuated fingers (Fonseca et al., 2019). These sensors were used to train a pose estimation model. Lambeta et al. (2020) explored a tactile-based transition model for marble manipulation using a self-supervised detector with auto-encoder architectures. Azulay et al. (2022a) 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. (2023) 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.

3.3 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 from unlabeled data, where the learning algorithm is able to infer the desired behavior from the structure of the data itself. This can be particularly useful for in-hand manipulation, as it allows the robot to learn about the various states and transitions of an object without the need for explicit human supervision. Exploration, on the other hand, refers to the process of actively seeking out and interacting with the environment in order to collect useful data. In the context of in-hand manipulation, exploration can involve the robot trying out different grasping and manipulation strategies in order to learn what works best for a given object and task. By actively seeking out and interacting with the environment in this way, the robot can learn about the various states and transitions of the object through trial and error, and utilize this knowledge to improve its manipulation performance. Together, self-supervision and exploration can be powerful tools for learning transitions in in-hand manipulation, as they allow the robot to learn from its own experiences and actively gather information about the object being manipulated and its surroundings.

The collection process to generate a state transition model for a robotic system requires active exploration of the high-dimensional state space (Kroemer et al., 2021). The common strategy is to exert random actions (Calli et al., 2018a) in the hope of achieving sufficient and uniform coverage of the robot's state space. In practice, some regions are not frequently visited and consequentially sparse. In systems such as compliant hands (Sintov et al., 2020c) or

object throwing (Zeng et al., 2020), each collection episode starts approximately from the same state and, thus, data is dense around the start state while sparser farther away. Therefore, acquiring state transition models for robotic systems requires exhausting and tedious data collection along with system wear, i.e., the transition function $f(\mathbf{x}_t, \mathbf{a}_t)$ is difficult to evaluate.

Active sampling is an alternative strategy where actions that are more informative for a specific task are taken (Wang et al., 2018). However, acquiring a general model of the robot requires the exploration of the entire feasible state space. Bayesian Optimization is the appropriate tool to identify key locations for sampling that would provide increased model accuracy. However, having knowledge of sampling locations does not guarantee the ability to easily reach them. Reaching some state-space regions may require exerting complicated maneuvers. The right actions that will drive the system to these regions for further exploration are usually unknown, particularly in preliminary stages with insufficient data. That is, we require a good model in order to learn a good model.

4 In-hand manipulation with 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 (Singh et al., 2022). 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.

4.1 Brief RL overview

Assuming that a given system is a Markov Decision Process (MDP), the next state depends solely on the current state and desired action according to a forward transition dynamics while receiving a reward. In model-based RL, a transition model can be learned independently of the policy learning as described in Section 3.2. When the system is stochastic due to uncertainties and limited observability, a Partial Observed Markov Decision Process (POMDP) is considered. As the true state cannot be fully observed in such a case, an observation space is introduced. Therefore, the agent receives an observation when reaching the next state with some probability. In both MDP and POMDP, the goal is, therefore, to learn a policy which maximizes the expected reward. A wider review

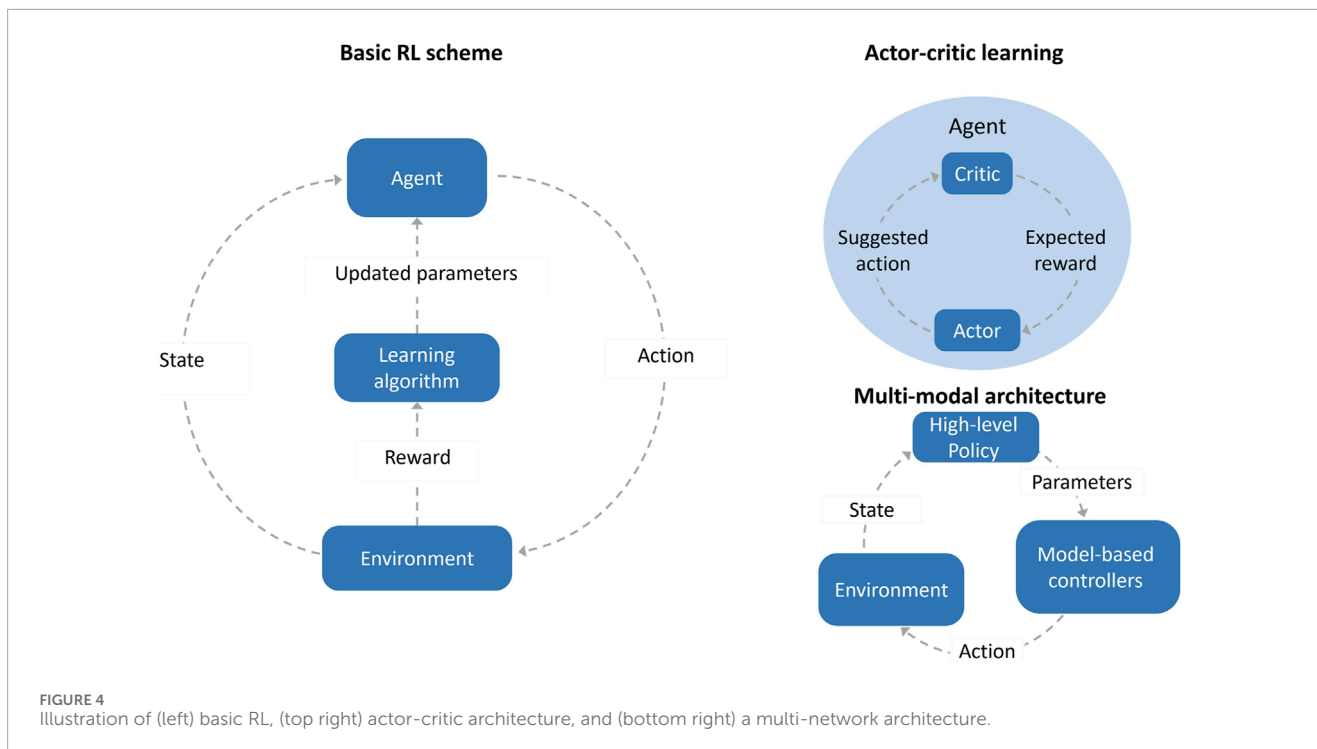of key concepts and methods can be found in the work of Nguyen and La (2019).

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 (Mnih et al., 2013; 2015) including robot manipulation (Levine et al., 2016; Nguyen and La, 2019) and specifically in-hand dexterous manipulation (Jain et al., 2019; Andrychowicz et al., 2020). Examples of straight-forward implementation of an RL algorithm include the work of Antonova et al. (2017) which addressed the pivoting problem with a parallel gripper. The RL policy was trained while relying on fast tracking with a camera. However, the trained policy yielded excessive back and forth motions. Cruciani and Smith (2017) 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 (Watkins and Dayan, 1992). In an extension work, Cruciani and Smith (2018) integrated path planning with the RL policy for the robotic arm to perform more complex tasks. In another example, Van Hoof et al. (2015) 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.

4.2 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 (Pourpanah et al., 2023). 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 (Huang et al., 2021).

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. (2019a) 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

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

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 (Kalashnikov et al., 2018; Zhu et al., 2019), 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 (Zhao W. et al., 2020). 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 (Höfer et al., 2021). This is especially relevant in the case of in-hand manipulation tasks which tend to heavily involve hard-to-model contact dynamics (Funabashi et al., 2020b; Liarokapis and Dollar, 2016). 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 (Van Baar et al., 2019). In this approach, various system parameters in the simulation are constantly varied in order to improve robustness to modeling errors. OpenAI et al. (2019); Andrychowicz et al. (2020) 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. (2022), 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 (Pitz et al., 2023). Beyond directly modifying the dynamics in domain randomization, applying small random forces to the grasped object was shown by Allshire et al. (2022) to improve the robustness of the resulting policy in in-grasp manipulation of the TriFinger hand. Recently, Handa et al. (2023) 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 (Makoviychuk et al., 2021). Using a GPU-based simulator reduces the amount of computational resources and costs.

As opposed to domain randomization, Qi et al. (2022) 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 (Qi et al., 2023). 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. (2023) 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.

4.3 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. (2018) 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 by Srinivasan et al. (2020), 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 (Falco et al., 2018). 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 (Gupta et al., 2021). 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.

4.4 Multi-network architecture

Multi-network architectures, such as actor-critic (Lillicrap et al., 2016) or teacher-student (Zimmer et al., 2014), 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. (2021a), Chen et al. (2022), Chen et al. (2023a) 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 (Morrison et al., 2020) or the YCB (Calli et al., 2015b) 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 (Toledo et al., 2021). 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 (Morgan et al., 2021a) and object insertion (Azulay et al., 2022b). Recently, Tao et al. (2023) 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. (2020), 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. (2020) 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.

4.5 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 (Wang et al., 2021). 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 (Chen et al., 2022). 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. (2022b) 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 (Allshire et al., 2022). 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.

4.6 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. (2019), 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. (2019) 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. (2019) 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 (Melnik et al., 2021).

While Melnik et al. used tactile information directly as part of the state vector, Funabashi et al. (2020b,a) 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,
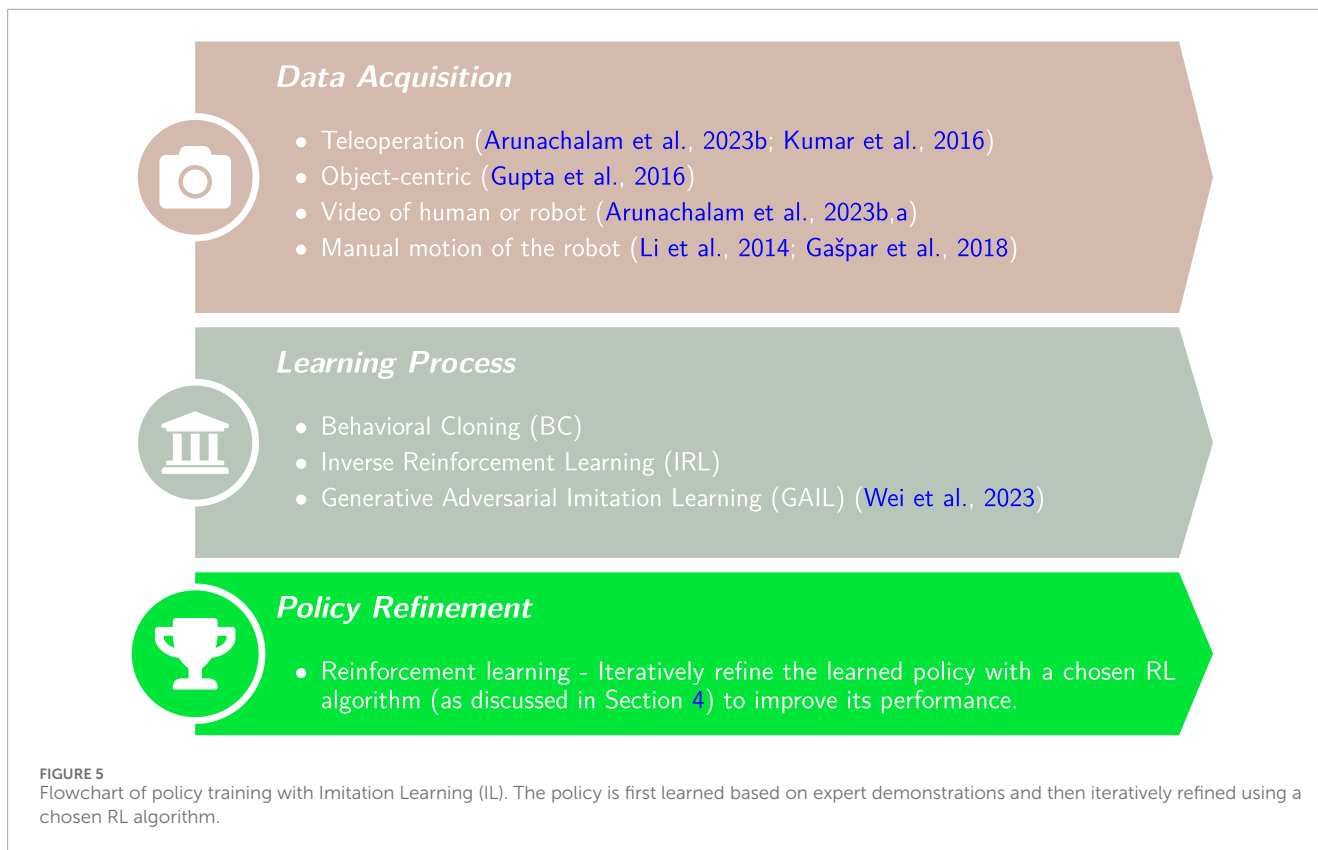
Yang et al. (2023) used a tactile array across a multi-finger hand. The array was embedded using Graph Neural Network which provides an object state during the manipulation and used for model-free RL. Recently, Khandate et al. (2022) implemented model-free RL to reorient an object through finger gaiting 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 (Khandate et al., 2023). 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 (Chen Y. et al., 2023). 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.

5 Imitation learning for in-hand manipulation

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 (Rapisarda, 2019). 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 (Duan et al., 2017; Fang et al., 2019). 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) (Zheng et al., 2022). 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 (Arora and Doshi, 2021).

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

FIGURE 5
Flowchart of policy training with Imitation Learning (IL). The policy is first learned based on expert demonstrations and then iteratively refined using a chosen RL algorithm.

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.

5.1 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 (Zhang et al., 2018). However, remote controls are unnatural and quite infeasible in teleoperation of dexterous robotic hands. In a more natural approach, Arunachalam et al. (2023b) 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 (Arunachalam et al., 2023a). 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. (2016) 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 (2023) designed a wearable robotic hand for IL teleoperation such that the expert has tactile feedback during demonstrations.

In a different approach by Gupta et al. (2016), 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 (Li et al., 2014; Shin and Jeon, 2024; Gašpar et al., 2018). 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.

5.2 Learning process

The process for learning from the demonstrations is commonly conducted by either BC or IRL (Hussein et al., 2017). In BC, the agent is required to directly take the strategy of the expert observed in the demonstrations (Arunachalam et al., 2023b). 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. (2021) 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. (2023) 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 (Yuan et al., 2020). The demonstration data was extracted from a handcrafted controller and shown to improve performance.

Rajeswaran et al. (2017) 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. (2019) 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. (2019) 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 h to 2–3 h 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 (Ross et al., 2011). IRL, on the other hand, directly learns the reward function of the demonstrated expert policy which prioritizes some actions over others (Arora and Doshi, 2021). 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 (Orbik et al., 2021). 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 (2020) 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 (Cai et al., 2019).

IL was also proposed for in-hand manipulation without the use of RL. Solak and Jamone (2019) proposed the use of Dynamical Movement Primitives (DMP) (Ijspeert et al., 2013). 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 (Solak and Jamone, 2023).

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) (Ho and Ermon, 2016) was proposed and combines IL with Generative Adversarial Networks (GAN) (Goodfellow et al., 2020). 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 (Wei et al., 2023). 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.

6 Discussion

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. A summary comparison of the three learning approaches discussed in this paper is given in Table 2. These tools encapsulate the ability to model complex and noisy systems such as a dexterous robotic hand

TABLE 2 Comparison of key components in learning methods for robotic in-hand manipulation.

Feature	Model-based learning	Reinforcement learning	Imitation learning
Learning Approach	Learns a model of the environment and uses it to plan actions	Learns a policy directly through trial and error, interacting with the environment	Learns a policy by imitating demonstrated behaviors
Data Requirements	Large amounts of data to train the model accurately	Large amounts of data for exploration and learning	Demonstrations of desired behaviors
Sample Efficiency	Generally, more sample efficient than reinforcement learning due to model-based planning	Can be sample inefficient, especially in complex and unstructured environments	Relatively sample efficient, as it directly learns from a set of demonstrations
Exploration	Exploration is often done through simulation or carefully designed exploration strategies	Exploration is a key component, often achieved through techniques like ε-greedy	Limited exploration, as it learns from fixed demonstrations
Generalizability	Can generalize to unseen environments if the model is accurate	Generalizability depends on the diversity of training data and the learning algorithm	Generalizability is limited by the diversity of demonstrations
Suitability for In-Hand Manipulation	Well-suited for tasks with complex dynamics and uncertainties, as the model can provide predictions and plan accordingly	Can be challenging due to the complexity of in-hand manipulation and the need for precise control demonstrated behaviors	Can be effective for tasks with clear demonstrations and limited variability

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 (Monastirsky et al., 2023) and diffusion policies (Chi et al., 2023). These advanced methods are versatile and can be applied to either of the three learning paradigms: model-based learning, RL and IL.

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 in terms of data efficiency, computationally light-weight and generalizable to different hardware, tasks and environments. Specifically, future research should prioritize the development of more realistic simulation environments tailored for soft and adaptive robotic hands.

4. Tactile sensing. While visual perception technology is quite mature, the use of high-resolution tactile sensing is relatively new. In general, Table 1 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 (Wang et al., 2022) 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. Collected datasets typically consist of several tens of thousands of samples tailored to the specific task at hand. The field therefore necessitates a paradigm shift toward solutions capable of seamless adaptation or generalization to novel tasks or objects. A large, standard and unified dataset of in-hand manipulation in-the-wild assembled by many researchers would be invaluable for advancing generalization.

7 Conclusion

This paper provides a comprehensive survey of various learning-based approaches for robotic in-hand manipulation, focusing on model-based methods, reinforcement learning (RL), and imitation learning (IL). Each of these methodologies has demonstrated significant progress in enabling robotic systems to perform dexterous in-hand manipulation tasks, which are essential for robots to operate effectively in complex human environments. Despite these advancements, several challenges remain, such as the need for higher data efficiency, improved sim-to-real transfer and better generalization across different objects and tasks.

While RL has revealed success due to its ability to generate solutions with minimal human intervention. Key findings indicate that RL policies often struggle with generalization and multi-task scenarios. Similarly, model-based approaches offer precision but can be limited by the complexity of dynamic environments. Imitation learning provides a promising avenue by leveraging expert demonstrations, but it requires extensive data collection, and its performance is highly dependent on the quality of the demonstrations. In addition to the challenges and future research suggestions discussed in Section 6, advancements should also be made in more applicative directions such as: enhance the generalization of models to be agnostic to the robotic hand with versatility to various tasks, through few-shot or zero-shot learning; augment the capabilities of prosthetic hands to perform

more complex tasks that usually involve in-hand manipulation; explore simplistic multimodal sensing while efficiently integrating these modalities; and, utilize the significant potential in human demonstration and continuous learning during human-robot collaboration, where robots can learn from human demonstrations and adapt to human preferences. By addressing these challenges, future research can push the boundaries of robotic dexterity, enabling robots to perform more sophisticated tasks autonomously.

Author contributions

AW: Conceptualization, Formal Analysis, Investigation, Supervision, Writing–original draft. AS: Conceptualization, Formal Analysis, Investigation, Software, Visualization, Writing–original draft, Writing–review and editing. OA: Conceptualization, Formal Analysis, Investigation, Validation, Visualization, Writing–original draft. AS: Conceptualization, Formal Analysis, Investigation, Supervision, Writing–original draft, Writing–review and editing.

Funding

Conflict of interest

Author AW was employed by AI-WEINBERG AI Experts.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abondance, S., Teeple, C. B., and Wood, R. J. (2020). A dexterous soft robotic hand for delicate in-hand manipulation. *IEEE Robotics Automation Lett.* 5, 5502–5509. doi:10.1109/LRA.2020.3007411

Agarwal, A., Uppal, S., Shaw, K., and Pathak, D. (2023). "Dexterous functional grasping," in *7th annual conference on robot learning.*

Allshire, A., Mittal, M., Lodaya, V., Makoviychuk, V., Makoviichuk, D., Widmaier, F., et al. (2022). "Transferring dexterous manipulation from GPU simulation to a remote real-world trifinger," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 11802–11809.

Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., et al. (2017). Hindsight experience replay (HER).

Andrychowicz, O. M., Baker, B., Chociej, M., Jozefowicz, R., McGrew, B., Pachocki, J., et al. (2020). Learning dexterous in-hand manipulation. *Int. J. Robotics Res.* 39, 3–20. doi:10.1177/0278364919887447

Antonova, R., Cruciani, S., Smith, C., and Kragic, D. (2017). Reinforcement learning for pivoting task. *arXiv preprint arXiv:1703.00472*

Arora, S., and Doshi, P. (2021). A survey of inverse reinforcement learning: challenges, methods and progress. *Artif. Intell.* 297, 103500. doi:10.1016/j.artint.2021.103500

Arunachalam, S. P., Guzey, I., Chintala, S., and Pinto, L. (2023a). "Holo-dex: teaching dexterity with immersive mixed reality," in *IEEE international conference on robotics and automation (ICRA)*, 5962–5969.

Arunachalam, S. P., Silwal, S., Evans, B., and Pinto, L. (2023b). "Dexterous imitation made easy: a learning-based framework for efficient dexterous manipulation," in *IEEE international conference on robotics and automation (ICRA)*, 5954–5961. doi:10.1109/ICRA48891.2023.10160275

Azulay, O., Ben-David, I., and Sintov, A. (2022a). Learning haptic-based object pose estimation for in-hand manipulation control with underactuated robotic hands. *IEEE Trans. Haptics* 16, 73–85. doi:10.1109/TOH.2022.3232713

Azulay, O., Curtis, N., Sokolovsky, R., Levitski, G., Slomovik, D., Lilling, G., et al. (2024). Allsight: a low-cost and high-resolution round tactile sensor with zero-shot learning capability. *IEEE Robotics Automation Lett.* 9, 483–490. doi:10.1109/lra.2023.3333701

Azulay, O., Monastirsky, M., and Sintov, A. (2022b). Haptic-based and $SE(3)$-aware object insertion using compliant hands. *IEEE Robotics Automation Lett.* 8, 208–215. doi:10.1109/lra.2022.3224670

Batsuren, K., and Yun, D. (2019). Soft robotic gripper with chambered fingers for performing in-hand manipulation. *Appl. Sci.* 9, 2967. doi:10.3390/app9152967

Bellman, R. (1957). A Markovian decision process. *Ind. Uni. Math. J.* 6, 679–684. doi:10.1512/iumj.1957.6.56038

Bhatt, A., Sieler, A., Puhlmann, S., and Brock, O. (2021). "Surprisingly robust in-hand manipulation: an empirical study," in *Robotics: science and systems*.

Billard, A., and Kragic, D. (2019). Trends and challenges in robot manipulation. *Science* 364, eaat8414. doi:10.1126/science.aat8414

Billings, G., and Johnson-Roberson, M. (2019). Silhonet: an rgb method for 6d object pose estimation. *IEEE Robotics Automation Lett.* 4, 3727–3734. doi:10.1109/LRA.2019.2928776

Bimbo, J., Luo, S., Althoefer, K., and Liu, H. (2016). In-hand object pose estimation using covariance-based tactile to geometry matching. *IEEE Robotics Automation Lett.* 1, 570–577. doi:10.1109/LRA.2016.2517244

Bimbo, J., Seneviratne, L. D., Althoefer, K., and Liu, H. (2013). "Combining touch and vision for the estimation of an object's pose during manipulation," in *IEEE/RSJ int* (NY, USA: Conf. on Intel. Rob. & Sys).

Butterfass, J., Grebenstein, M., Liu, H., and Hirzinger, G. (2001). DLR-Hand II: next generation of a dextrous robot hand. *IEEE Int. Conf. Robotics Automation* 1, 109–114. doi:10.1109/ROBOT.2001.932538

Cai, Q., Hong, M., Chen, Y., and Wang, Z. (2019). On the global convergence of imitation learning: a case for linear quadratic regulator

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

Calli, B., and Dollar, A. M. (2017). "Vision-based model predictive control for within-hand precision manipulation with underactuated grippers," in *IEEE international conference on robotics and automation*, 2839–2845.

Calli, B., Kimmel, A., Hang, K., Bekris, K., and Dollar, A. (2018a). "Path planning for within-hand manipulation over learned representations of safe states," in *International symposium on experimental robotics* (Springer), 437–447.

Calli, B., Singh, A., Walsman, A., Srinivasa, S., Abbeel, P., and Dollar, A. M. (2015a). "The ycb object and model set: towards common benchmarks for manipulation research," in *International conference on advanced robotics (ICAR)* (IEEE), 510–517.

Calli, B., Srinivasan, K., Morgan, A., and Dollar, A. M. (2018b). "Learning modes of within-hand manipulation," in *IEEE international conference on robotics and automation (ICRA)*, 3145–3151. doi:10.1109/ICRA.2018.8461187

Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., and Dollar, A. M. (2015b). Benchmarking in manipulation research: using the yale-cmu-berkeley object and model set. *IEEE Robot. Autom. Mag.* 22, 36–52. doi:10.1109/mra.2015.2448951

Carter, J., and Fourney, D. (2005). "Research based tactile and haptic interaction guidelines," in *Proceedings of guidelines on tactile and haptic interactions*, 84–92.

Chapman, J., Gorjup, G., Dwivedi, A., Matsunaga, S., Mariyama, T., MacDonald, B., et al. (2021). "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)*, 6941–6947.

Chavan-Dafle, N., Holladay, R., and Rodriguez, A. (2020). Planar in-hand manipulation via motion cones. *Int. J. Robotics Res.* 39, 163–182. doi:10.1177/0278364919880257

Chen, T., Tippur, M., Wu, S., Kumar, V., Adelson, E., and Agrawal, P. (2023a). Visual dexterity: in-hand reorientation of novel and complex object shapes. *Sci. Robotics* 8, eadc9244. doi:10.1126/scirobotics.adc9244

Chen, T., Xu, J., and Agrawal, P. (2021a). "A simple method for complex in-hand manipulation," in *Annual conference on robot learning*, 3.

Chen, T., Xu, J., and Agrawal, P. (2022). "A system for general in-hand object re-orientation," in *Conference on robot learning*.

Chen, Y., Geng, Y., Zhong, F., Ji, J., Jiang, J., Lu, Z., et al. (2023b). Bi-dexhands: towards human-level bimanual dexterous manipulation. *IEEE Trans. Pattern Analysis Mach. Intell.* 46, 2804–2818. doi:10.1109/TPAMI.2023.3339515

Chen, Y., Prepscius, C., Lee, D., and Lee, D. D. (2021b). Tactile velocity estimation for controlled in-grasp sliding. *IEEE Robotics Automation Lett.* 6, 1614–1621. doi:10.1109/lra.2021.3058931

Cheng, M.-Y., Tsao, C.-M., Lai, Y.-T., and Yang, Y.-J. (2009). "A novel highly-twistable tactile sensing array using extendable spiral electrodes," in *IEEE international conference on micro electro mechanical systems*, 92–95.

Chi, C., Feng, S., Du, Y., Xu, Z., Cousineau, E., Burchfiel, B., et al. (2023). "Diffusion policy: visuomotor policy learning via action diffusion," in *Robotics: science and systems (RSS)*.

Choi, C., Del Preto, J., and Rus, D. (2017). "Using vision for pre- and post-grasping object localization for soft hands," in *International symposium on experimental robotics*. Editors D. Kulić, Y. Nakamura, O. Khatib, and G. Venture (Cham), 601–612.

Cordella, F., Ciancio, A. L., Sacchetti, R., Davalli, A., Cutti, A. G., Guglielmelli, E., et al. (2016). Literature review on needs of upper limb prosthesis users. *Front. Neurosci.* 10, 209. doi:10.3389/fnins.2016.00209

Costanzo, M. (2021). Control of robotic object pivoting based on tactile sensing. *Mechatronics* 76, 102545. doi:10.1016/j.mechatronics.2021.102545

Costanzo, M., De Maria, G., and Natale, C. (2021). "Dual-arm in-hand manipulation with parallel grippers using tactile feedback," in *International conference on advanced robotics (ICAR)*, 942–947.

Cruciani, S., and Smith, C. (2017). "In-hand manipulation using three-stages open loop pivoting," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 1244–1251. doi:10.1109/IROS.2017.8202299

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

Cruciani, S., Smith, C., Kragic, D., and Hang, K. (2018). "Dexterous manipulation graphs," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 2040–2047.

Cruciani, S., Sundaralingam, B., Hang, K., Kumar, V., Hermans, T., and Kragic, D. (2020). Benchmarking in-hand manipulation. *IEEE Robotics Automation Lett.* 5, 588–595. doi:10.1109/LRA.2020.2964160

Cui, J., and Trinkle, J. (2021). Toward next-generation learned robot manipulation. *Sci. Robotics* 6, eabd9461. doi:10.1126/scirobotics.abd9461

Dafle, N. C., Rodriguez, A., Paolini, R., Tang, B., Srinivasa, S. S., Erdmann, M., et al. (2014). "Extrinsic dexterity: in-hand manipulation with external forces," in *IEEE international conference on robotics and automation (ICRA)*, 1578–1585.

Deimel, R. (2014). "A novel type of compliant, underactuated robotic hand for dexterous grasping," in *Robotics: science and systems*, 1–9.

Delgado, A., Jara, C., and Torres, F. (2017). In-hand recognition and manipulation of elastic objects using a servo-tactile control strategy. *Robotics Computer-Integrated Manuf.* 48, 102–112. doi:10.1016/j.rcim.2017.03.002

Deng, Z., and Zhang, J. W. (2020). Learning synergies based in-hand manipulation with reward shaping. *CAAI Trans. Intell. Technol.* 5, 141–149. doi:10.1049/trit.2019.0094

Dimou, D., Santos-Victor, J., and Moreno, P. (2023). Robotic hand synergies for in-hand regrasping driven by object information. *Aut. Robots* 47, 453–464. doi:10.1007/s10514-023-10101-z

Dollar, A. M., and Howe, R. D. (2010). The highly adaptive sdm hand: design and performance evaluation. *Int. J. Robotics Res.* 29, 585–597. doi:10.1177/0278364909360852

Duan, Y., Andrychowicz, M., Stadie, B., Jonathan Ho, O., Schneider, J., Sutskever, I., et al. (2017). "One-shot imitation learning,". *Advances in neural information processing systems*. Editors I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. (NY, USA: Curran Associates, Inc.), 30.

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

Falco, P., Attawia, A., Saveriano, M., and Lee, D. (2018). On policy learning robust to irreversible events: an application to robotic in-hand manipulation. *IEEE Robotics Automation Lett.* 3, 1482–1489. doi:10.1109/lra.2018.2800110

Fan, Y., Gao, W., Chen, W., and Tomizuka, M. (2017). Real-time finger gaits planning for dexterous manipulation. *IFAC-PapersOnLine* 50, 12765–12772. doi:10.1016/j.ifacol.2017.08.1831

Fang, B., Jia, S., Guo, D., Xu, M., Wen, S., and Sun, F. (2019). Survey of imitation learning for robotic manipulation. *Int. J. Intelligent Robotics Appl.* 3, 362–369. doi:10.1007/s41315-019-00103-5

Feng, Z., Hu, G., Sun, Y., and Soon, J. (2020). An overview of collaborative robotic manipulation in multi-robot systems. *Annu. Rev. Control* 49, 113–127. doi:10.1016/j.arcontrol.2020.02.002

Fonseca, V., Alves de Oliveira, T. E., and Petriu, E. (2019). Estimating the orientation of objects from tactile sensing data using machine learning methods and visual frames of reference. *Sensors* 2285. doi:10.3390/s19102285

Funabashi, S., Isobe, T., Ogasa, S., Ogata, T., Schmitz, A., Tomo, T. P., et al. (2020a). "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)*, 9166–9173.

Funabashi, S., Ogasa, S., Isobe, T., Ogata, T., Schmitz, A., Tomo, T. P., et al. (2020b). "Variable in-hand manipulations for tactile-driven robot hand via cnn-lstm," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 9472–9479.

Funabashi, S., Schmitz, A., Ogasa, S., and Sugano, S. (2019a). Morphology specific stepwise learning of in-hand manipulation with a four-fingered hand. *IEEE Trans. Industrial Inf.* 16, 433–441. doi:10.1109/tii.2019.2893713

Funabashi, S., Schmitz, A., Sato, T., Somlor, S., and Sugano, S. (2019b). "Versatile in-hand manipulation of objects with different sizes and shapes using neural networks," in *IEEE-RAS international conference on humanoid robots 2018-novem*, 768–775. doi:10.1109/HUMANOIDS.2018.8624961

Furukawa, N., Namiki, A., Taku, S., and Ishikawa, M. (2006). "Dynamic regrasping using a high-speed multifingered hand and a high-speed vision system," in *IEEE international conference on robotics and automation*, 181–187. doi:10.1109/ROBOT.2006.1641181

Garcia-Hernando, G., Johns, E., and Kim, T.-K. (2020). "Physics-based dexterous manipulations with estimated hand poses and residual reinforcement learning," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 9561–9568. doi:10.1109/IROS45743.2020.9340947

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., and Marín-Jiménez, M. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit.* 47, 2280–2292. doi:10.1016/j.patcog.2014.01.005

Gašpar, T., Nemec, B., Morimoto, J., and Ude, A. (2018). Skill learning and action recognition by arc-length dynamic movement primitives. *Robotics Aut. Syst.* 100, 225–235. doi:10.1016/j.robot.2017.11.012

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2020). Generative adversarial networks. *Commun. ACM* 63, 139–144. doi:10.1145/3422622

Guo, M., Gealy, D. V., Liang, J., Mahler, J., Goncalves, A., McKinley, S., et al. (2017). "Design of parallel-jaw gripper tip surfaces for robust grasping," in *IEEE international conference on robotics and automation (ICRA)*, 2831–2838.

Gupta, A., Eppner, C., Levine, S., and Abbeel, P. (2016). "Learning dexterous manipulation for a soft robotic hand from human demonstrations," in *IEEEqRSJ international conference on intelligent robots and systems*, 3786–3793. doi:10.1109/IROS.2016.7759557

Gupta, A., Yu, J., Zhao, T. Z., Kumar, V., Rovinsky, A., Xu, K., et al. (2021). "Reset-free reinforcement learning via multi-task learning: learning dexterous manipulation behaviors without human intervention," in *IEEE international conference on robotics and automation (ICRA)*, 6664–6671. doi:10.1109/ICRA48506.2021.9561384

Haldar, S., Pari, J., Rai, A., and Pinto, L. (2023). Teach a robot to fish: versatile imitation from one minute of demonstrations. *arXiv Prepr. arXiv:2303.01497*.

Hammond, F. L., Weisz, J., de la Llera Kurth, A. A., Allen, P. K., and Howe, R. D. (2012). "Towards a design optimization method for reducing the mechanical complexity of underactuated robotic hands," in *2012 IEEE international conference on robotics and automation*, 2843–2850. doi:10.1109/ICRA.2012.6225010

Han, D., Mulyana, B., Stankovic, V., and Cheng, S. (2023). A survey on deep reinforcement learning algorithms for robotic manipulation. *Sensors* 23, 3762. doi:10.3390/s23073762

Han, L., Guan, Y., Li, Z., Shi, Q., and Trinkle, J. (1997). Dextrous manipulation with rolling contacts. *IEEE Int. Conf. Robotics Automation* 2, 992–997. doi:10.1109/robot.1997.614264

Handa, A., Allshire, A., Makoviychuk, V., Petrenko, A., Singh, R., Liu, J., et al. (2023). "Dextreme: transfer of agile in-hand manipulation from simulation to reality," in *IEEE international conference on robotics and automation (ICRA)*, 5977–5984. doi:10.1109/ICRA48891.2023.10160216

He, J., Pu, S., and Zhang, J. (2015). Haptic and visual perception in in-hand manipulation system. *IEEE Int. Conf. Robotics Biomimetics (ROBIO)*, 303–308. doi:10.1109/ROBIO.2015.7418784

Herguedas, R., López-Nicolás, G., Aragüés, R., and Sagüés, C. (2019). "Survey on multi-robot manipulation of deformable objects," in *IEEE international conference on emerging technologies and factory automation (ETFA)*, 977–984. doi:10.1109/ETFA.2019.8868987

Higo, R., Yamakawa, Y., Senoo, T., and Ishikawa, M. (2018). "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)*, 6609–6614. doi:10.1109/IROS.2018.8593538

Ho, J., and Ermon, S. (2016). Generative adversarial imitation learning. *Corr. abs/1606.03476*. doi:10.5555/3157382.3157608

Höfer, S., Bekris, K., Handa, A., Gamboa, J. C., Mozifian, M., Golemo, F., et al. (2021). Sim2real in robotics and automation: applications and challenges. *IEEE Trans. automation, Sci. Eng.* 18, 398–400. doi:10.1109/tase.2021.3064065

Huang, W., Mordatch, I., Abbeel, P., and Pathak, D. (2021). Generalization in dexterous manipulation via geometry-aware multi-task learning. *arXiv*.

Hussein, A., Gaber, M. M., Elyan, E., and Jayne, C. (2017). Imitation learning: a survey of learning methods. *ACM Comput. Surv. (CSUR)* 50, 1–35. doi:10.1145/3054912

Ichnowski, J., Avigal, Y., Kerr, J., and Goldberg, K. (2021). Dex-nerf: using a neural radiance field to grasp transparent objects. *arXiv Prepr. arXiv:2110.14217*.

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput.* 25, 328–373. doi:10.1162/NECO_a_00393

Jacobsen, S., Iversen, E., Knutti, D., Johnson, R., and Biggers, K. (1986). Design of the Utah/m.i.t. dextrous hand. *IEEE Int. Conf. Robotics Automation* 3, 1520–1532. doi:10.1109/ROBOT.1986.1087395

Jain, D., Li, A., Singhal, S., Rajeswaran, A., Kumar, V., and Todorov, E. (2019). *Learning deep visuomotor policies for dexterous hand manipulation*. IEEE International Conference on Robotics and Automation ICRA, 3636–3643doi:10.1109/icra.2019.8794033

Kalaitzakis, M., Cain, B., Carroll, S., Ambrosi, A., Whitehead, C., and Vitzilaios, N. (2021). Fiducial markers for pose estimation: overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers. *J. Intelligent & Robotic Syst.* 101, 71. doi:10.1007/s10846-020-01307-9

Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., et al. (2018). "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Conference on robot learning (PMLR)*, 651–673.

Kang, T., Yi, J.-B., Song, D., and Yi, S.-J. (2021). High-speed autonomous robotic assembly using in-hand manipulation and re-grasping. *Appl. Sci.* 11, 37. doi:10.3390/app11010037

Khandate, G., Haas-Heger, M., and Ciocarlie, M. (2022). "On the feasibility of learning finger-gaiting in-hand manipulation with intrinsic sensing," in *IEEE international conference on robotics and automation (ICRA)*, 2752–2758. doi:10.1109/ICRA46639.2022.9812212

Khandate, G., Shang, S., Chang, E. T., Saidi, T. L., Adams, J., and Ciocarlie, M. (2023). "Sampling-based exploration for reinforcement learning of dexterous manipulation," in *Robotics: Science and systems (daegu, Republic of Korea)*.

Khazatsky, A., Pertsch, K., Nair, S., Balakrishna, A., Dasari, S., Karamcheti, S., et al. (2024). DROID: a large-scale in-the-wild robot manipulation dataset. *arXiv*.

Kimmel, A., Sintov, A., Tan, J., Wen, B., Boularias, A., and Bekris, K. E. (2019). "Belief-space planning using learned models with application to underactuated hands," in *International symposium on robotics research (ISRR)*.

Kokic, M., Kragic, D., and Bohg, J. (2019). "Learning to estimate pose and shape of hand-held objects from rgb images," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 3980–3987. doi:10.1109/IROS40897.2019.8967961

Korthals, T., Melnik, A., Leitner, J., and Hesse, M. (2019). "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*, 1–6.

Koval, M. C., Dogar, M. R., Pollard, N. S., and Srinivasa, S. S. (2013). Pose estimation for contact manipulation with manifold particle filters. *IEEE/RSJ Int. Conf. Intel. Rob. Sys*, 4541–4548. doi:10.1109/IROS.2013.6697009

Kroemer, O., Niekum, S., and Konidaris, G. (2021). A review of robot learning for manipulation: challenges, representations, and algorithms. *J. Mach. Learn. Res.* 22, 1395–1476. doi:10.5555/3546258.3546288

Kumar, V., Gupta, A., Todorov, E., and Levine, S. (2016). Learning dexterous manipulation policies from experience and imitation. doi:10.1177/ToBeAssigned

Lambeta, M., Chou, P.-W., Tian, S., Yang, B., Maloon, B., Most, V. R., et al. (2020). Digit: a novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation. *IEEE Robotics Automation Lett.* 5, 3838–3845. doi:10.1109/lra.2020.2977257

Lehman, A. C., Wood, N. A., Farritor, S., Goede, M. R., and Oleynikov, D. (2010). Dexterous miniature robot for advanced minimally invasive surgery. *Surg. Endosc.* 25, 119–123. doi:10.1007/s00464-010-1143-6

Lepora, N., and Lloyd, J. (2020). Optimal deep learning for robot touch: training accurate pose models of 3d surfaces and edges. *IEEE Robotics Automation Mag.* 27, 66–77. doi:10.1109/mra.2020.2979658

Levine, S., Finn, C., Darrell, T., and Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *J. Mach. Learn. Res.* 17, 1334–1373. doi:10.5555/2946645.2946684

Li, M., Yin, H., Tahara, K., and Billard, A. (2014). "Learning object-level impedance control for robust grasping and dexterous manipulation," in *IEEE international conference on robotics and automation*, 6784–6791. doi:10.1109/ICRA.2014.6907861

Li, T., Srinivasan, K., Meng, M. Q.-H., Yuan, W., and Bohg, J. (2020). "Learning hierarchical control for robust in-hand manipulation," in *IEEE international conference on robotics and automation (ICRA)*, 8855–8862. doi:10.1109/ICRA40945.2020.9197343

Liarokapis, M., and Dollar, A. M. (2016). "Post-contact, in-hand object motion compensation for compliant and underactuated hands," in *IEEE international symposium on robot and human interactive communication* (NY, USA: RO-MAN), 986–993.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., et al. (2016). "Continuous control with deep reinforcement learning," in *International conference on learning representations (ICLR)*.

Liu, H., Zhao, L., Siciliano, B., and Ficuciello, F. (2020). Modeling, optimization, and experimentation of the paragripper for in-hand manipulation without parasitic rotation. *IEEE Robotics Automation Lett.* 5, 3011–3018. doi:10.1109/LRA.2020.2974419

Liu, Y., Yang, Y., Wang, Y., Wu, X., Wang, J., Yao, Y., et al. (2024). Realdex: towards human-like grasping for robotic dexterous hand. 6859, 6867. doi:10.24963/ijcai.2024/758

Llop-Harillo, I., Pérez-González, A., Starke, J., and Asfour, T. (2019). The anthropomorphic hand assessment protocol (ahap). *Robotics Aut. Syst.* 121, 103259. doi:10.1016/j.robot.2019.103259

Lotti, F., Tiezzi, P., Vassura, G., Biagiotti, L., and Melchiorri, C. (2004). Ubh 3: an anthropomorphic hand with simplified endo-skeletal structure and soft continuous fingerpads. *IEEE Int. Conf. Robotics Automation* 5, 4736–4741. doi:10.1109/ROBOT.2004.1302466

Lovchik, C., and Diftler, M. (1999). The robonaut hand: a dexterous robot hand for space. *IEEE Int. Conf. Robotics Automation* 2, 907–912. doi:10.1109/ROBOT.1999.772420

Lozano-Perez, T., Jones, J., Mazer, E., O'Donnell, P., Grimson, W., Tournassoud, P., et al. (1987). Handey: a robot system that recognizes, plans, and manipulates. *IEEE Int. Conf. Robotics Automation* 4, 843–849. doi:10.1109/ROBOT.1987.1087847

Luo, Y., Li, W., Wang, P., Duan, H., Wei, W., and Sun, J. (2023). Progressive transfer learning for dexterous in-hand manipulation with multi-fingered anthropomorphic hand

Ma, R. R., and Dollar, A. M. (2011). "On dexterity and dexterous manipulation," in *International conference on advanced robotics (ICAR)*, 1–7.

Ma, R. R., and Dollar, A. M. (2016). "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.*

Ma, R. R., and Dollar, A. M. (2017). Yale openhand project: optimizing open-source hand designs for ease of fabrication and adoption. *IEEE Robotics Autumation Magezine* 24, 32–40. doi:10.1109/mra.2016.2639034

Makoviychuk, V., Wawrzyniak, L., Guo, Y., Lu, M., Storey, K., Macklin, M., et al. (2021). Isaac gym: high performance gpu-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470*

Marinelli, A., Boccardo, N., Tessari, F., Di Domenico, D., Caserta, G., Canepa, M., et al. (2022). Active upper limb prostheses: a review on current state and upcoming breakthroughs. *Prog. Biomed. Eng.* 5, 012001. doi:10.1088/2516-1091/acac57

Mason, M. T., and Salisbury, J. K. (1985). *Robot hands and the mechanics of manipulation*. Cambridge, MA, USA: MIT Press.

Melnik, A., Lach, L., Plappert, M., Korthals, T., Haschke, R., and Ritter, H. (2019). "Tactile sensing and deep reinforcement learning for in-hand manipulation tasks," in *IROS workshop on autonomous object manipulation*.

Melnik, A., Lach, L., Plappert, M., Korthals, T., Haschke, R., and Ritter, H. (2021). Using tactile sensing to improve the sample efficiency and performance of deep deterministic policy gradients for simulated in-hand manipulation tasks. *Front. Robotics AI* 8, 538773. doi:10.3389/frobt.2021.538773

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., et al. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi:10.1038/nature14236

Mohammed, M. Q., Kwek, L. C., Chua, S. C., Al-Dhaqm, A., Nahavandi, S., Eisa, T. A. E., et al. (2022). Review of learning-based robotic manipulation in cluttered environments. *Sensors* 22, 7938. doi:10.3390/s22207938

Molnar, J., and Menguc, Y. (2022). "Toward handling the complexities of non-anthropomorphic hands," in *CHI conference on human factors in computing systems* (LA, USA: Association for Computing Machinery).

Monastirsky, M., Azulay, O., and Sintov, A. (2023). Learning to throw with a handful of samples using decision transformers. *IEEE Robotics Automation Lett.* 8, 576–583. doi:10.1109/LRA.2022.3229266

Morgan, A. S., Hang, K., Bircher, W. G., and Dollar, A. M. (2019). "A data-driven framework for learning dexterous manipulation of unknown objects," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 8273–8280.

Morgan, A. S., Hang, K., and Dollar, A. M. (2020). Object-agnostic dexterous manipulation of partially constrained trajectories. *IEEE Robotics Automation Lett.* 5, 5494–5501. doi:10.1109/lra.2020.3007467

Morgan, A. S., Nandha, D., Chalvatzaki, G., D'Eramo, C., Dollar, A. M., and Peters, J. (2021a). "Model predictive actor-critic: accelerating robot skill acquisition with deep reinforcement learning," in *IEEE international conference on robotics and automation (ICRA)*, 6672–6678.

Morgan, A. S., Wen, B., Liang, J., Boularias, A., Dollar, A. M., and Bekris, K. (2021b). "Vision-driven compliant manipulation for reliable, high-precision assembly tasks," in *Robotics: science and systems*.

Morrison, D., Corke, P., and Leitner, J. (2020). Egad! an evolved grasping analysis dataset for diversity and reproducibility in robotic manipulation. *IEEE Robotics Automation Lett.* 5, 4368–4375. doi:10.1109/lra.2020.2992195

Mouri, T. (2002). "Anthropomorphic robot hand: Gifu hand iii," in *International conf. Control* (Automation and Systems), 1288–1293.

Nagabandi, A., Konolige, K., Levine, S., and Kumar, V. (2020). "Deep dynamics models for learning dexterous manipulation," in *Conference on robot learning (PMLR)*, 1101–1112.

Nagata, K. (1994). Manipulation by a parallel-jaw gripper having a turntable at each fingertip. *IEEE Int. Conf. Robotics Automation* 2, 1663–1670. doi:10.1109/ROBOT.1994.351352

Nahum, N., and Sintov, A. (2022). Robotic manipulation of thin objects within off-the-shelf parallel grippers with a vibration finger. *Mech. Mach. Theory* 177, 105032. doi:10.1016/j.mechmachtheory.2022.105032

Nematollahi, I., Rosete-Beas, E., Azad, S. M. B., Rajan, R., Hutter, F., and Burgard, W. (2022). *T3vip: transformation-based 3d video prediction*. IEEE/RSJ International Conference on Intelligent Robots and Systems IROS.

Nguyen, H., and La, H. M. (2019). "Review of deep reinforcement learning for robot manipulation," in *IEEE international conference on robotic computing (IRC)*, 590–595. doi:10.1109/IRC.2019.00120

Odhner, L. U., and Dollar, A. M. (2011). "Dexterous manipulation with underactuated elastic hands," in *IEEE int. Conf. On rob. And aut*, 5254–5260.

Odhner, L. U., and Dollar, A. M. (2015). Stable, open-loop precision manipulation with underactuated hands. *Int. J. Robotics Res.* 34, 1347–1360. doi:10.1177/0278364914558494

Okada, T. (1979). Object-handling system for manual industry. *IEEE Trans. Syst. Man, Cybern.* 9, 79–89. doi:10.1109/tsmc.1979.4310152

Okamura, A., Smaby, N., and Cutkosky, M. (2000). An overview of dexterous manipulation. *IEEE Int. Conf. Robotics Automation* 1, 255–262. doi:10.1109/robot.2000.844067

Olson, E. (2011). "AprilTag: a robust and flexible visual fiducial system," in *IEEE international conference on robotics and automation (ICRA)*, 3400–3407.

OpenAI, Akkaya, I., Andrychowicz, M., Chociej, M., Litwin, M., McGrew, B., et al. (2019), *Solving Rubik's cube with a robot hand*, 1–51.

Orbik, J., Agostini, A., and Lee, D. (2021). "Inverse reinforcement learning for dexterous hand manipulation," in *IEEE international conference on development and learning (ICDL)*, 1–7. doi:10.1109/ICDL49984.2021.9515637

Ozawa, R., Arimoto, S., Nakamura, S., and Bae, J.-H. (2005). Control of an object with parallel surfaces by a pair of finger robots without object sensing. *IEEE Trans. Robotics* 21, 965–976. doi:10.1109/tro.2005.852263

Pagoli, A., Chapelle, F., Corrales, J. A., Mezouar, Y., and Lapusta, Y. (2021). A soft robotic gripper with an active palm and reconfigurable fingers for fully dexterous in-hand manipulation. *IEEE Robotics Automation Lett.* 6, 7706–7713. doi:10.1109/LRA.2021.3098803

Papadopoulos, E., Aghili, F., Ma, O., and Lampariello, R. (2021). Robotic manipulation and capture in space: a survey. *Front. Robotics AI* 8, 686723. doi:10.3389/frobt.2021.686723

Park, W., Park, S., An, H., Seong, M., Bae, J., and Jeong, H. E. (2024). A sensorized soft robotic hand with adhesive fingertips for multimode grasping and manipulation. *Soft Robot.* 11, 698–708. doi:10.1089/soro.2023.0099

Petrich, L., Jin, J., Dehghan, M., and Jagersand, M. (2021). Assistive arm and hand manipulation: how does current research intersect with actual healthcare needs?

Pfanne, M., Chalon, M., Stulp, F., Ritter, H., and Albu-Schäffer, A. (2020). Object-level impedance control for dexterous in-hand manipulation. *IEEE Robotics Automation Lett.* 5, 2987–2994. doi:10.1109/LRA.2020.2974702

Pitz, J., Röstel, L., Sievers, L., and Bäuml, B. (2023). "Dextrous tactile in-hand manipulation using a modular reinforcement learning architecture," in *IEEE international conference on robotics and automation (ICRA)*, 1852–1858. doi:10.1109/ICRA48891.2023.10160756

Pourpanah, F., Abdar, M., Luo, Y., Zhou, X., Wang, R., Lim, C. P., et al. (2023). A review of generalized zero-shot learning methods. *IEEE Trans. Pattern Analysis Mach. Intell.* 45, 4051–4070. doi:10.1109/TPAMI.2022.3191696

Psomopoulou, E., Pestell, N., Papadopoulos, F., Lloyd, J., Doulgeri, Z., and Lepora, N. F. (2021). A robust controller for stable 3d pinching using tactile sensing. *IEEE Robotics Automation Lett.* 6, 8150–8157. doi:10.1109/lra.2021.3104057

Qi, H., Kumar, A., Calandra, R., Ma, Y., and Malik, J. (2022). "In-hand object rotation via rapid motor adaptation," in *Conference on robot learning (CoRL)*.

Qi, H., Yi, B., Suresh, S., Lambeta, M., Ma, Y., Calandra, R., et al. (2023). "General in-hand object rotation with vision and touch,". *Conference on robot learning*. Editors J. Tan, M. Toussaint, and K. Darvish, 229, 2549–2564.

Rad, M., and Lepetit, V. (2017). "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)*, 3848–3856. doi:10.1109/ICCV.2017.413

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

Rajeswaran, A., Kumar, V., Gupta, A., Vezzani, G., Schulman, J., Todorov, E., et al. (2017). Learning complex dexterous manipulation with deep reinforcement learning and demonstrations

Rapisarda, C. (2019). Reinforcement learning for dexterity transfer between manipulators

Robles-De-La-Torre, G., and Hayward, V. (2001). Force can overcome object geometry in the perception of shape through active touch. *Nature* 412, 445–448. doi:10.1038/35086588

Ross, S., Gordon, G., and Bagnell, D. (2011). "A reduction of imitation learning and structured prediction to no-regret online learning," in *International conference on artificial intelligence and Statistics. 15 of Proceedings of machine learning research.* Editors G. Gordon, D. Dunson, and M. Dudík, 627–635.

Rus, D. (1999). In-hand dexterous manipulation of piecewise-smooth 3d objects. *Int. J. Robotics Res.* 18, 355–381. doi:10.1177/02783649922066268

Schulz, S., Pylatiuk, C., and Bretthauer, G. (2001). A new ultralight anthropomorphic hand. In *IEEE Int. Conf. Robotics Automation.* vol. 3, 2437–2441. doi:10.1109/ROBOT.2001.932988

Shi, J., Weng, H., and Lynch, K. M. (2020). In-hand sliding regrasp with spring-sliding compliance and external constraints. *IEEE Access* 8, 88729–88744. doi:10.1109/ACCESS.2020.2991382

Shi, J., Woodruff, J. Z., Umbanhowar, P. B., and Lynch, K. M. (2017a). Dynamic in-hand sliding manipulation. *IEEE Trans. Robotics* 33, 778–795. doi:10.1109/tro.2017.2693391

Shi, J., Woodruff, J. Z., Umbanhowar, P. B., and Lynch, K. M. (2017b). Dynamic in-hand sliding manipulation. *IEEE Trans. Robotics* 33, 778–795. doi:10.1109/TRO.2017.2693391

Shin, K. J., and Jeon, S. (2024). Nonprehensile manipulation for rapid object spinning via multisensory learning from demonstration. *Sensors* 24, 380. doi:10.3390/s24020380

Sievers, L., Pitz, J., and Bäuml, B. (2022). "Learning purely tactile in-hand manipulation with a torque-controlled hand," in *IEEE international conference on robotics and automation (ICRA)*, 2745–2751. doi:10.1109/ICRA46639.2022.9812093

Singh, B., Kumar, R., and Singh, V. P. (2022). Reinforcement learning in robotic applications: a comprehensive survey. *Artif. Intell. Rev.* 55, 945–990. doi:10.1007/s10462-021-09997-9

Sintov, A., Kimmel, A., Bekris, K. E., and Boularias, A. (2020a). "Motion planning with competency-aware transition models for underactuated adaptive hands," in *2020 IEEE international conference on robotics and automation (ICRA)* (IEEE), 7761–7767. doi:10.1109/ROBOT.1994.351352

Sintov, A., Kimmel, A., Wen, B., Boularias, A., and Bekris, K. (2020b). "Tools for data-driven modeling of within-hand manipulation with underactuated adaptive hands," in *Conference on learning for dynamics and control. PMLR), vol. 120 of Proceedings of machine learning research.* Editors A. M. Bayen, A. Jadbabaie, G. Pappas, P. A. Parrilo, B. Recht, C. Tomlin, et al. 771–780.

Sintov, A., Kimmel, A., Wen, B., Boularias, A., and Bekris, K. (2020c). Tools for data-driven modeling of within-hand manipulation with underactuated adaptive hands. *Proc. Conf. Learn. Dyn. & Cont* 120, 771–780.

Sintov, A., Morgan, A. S., Kimmel, A., Dollar, A. M., Bekris, K. E., and Boularias, A. (2019). Learning a state transition model of an underactuated adaptive hand. *IEEE Robotics Automation Lett.* 4, 1287–1294. doi:10.1109/LRA.2019.2894875

Sintov, A., and Shapiro, A. (2016). "Swing-up regrasping algorithm using energy control," in *IEEE international conference on robotics and automation (ICRA)*, 4888–4893.

Sintov, A., and Shapiro, A. (2017). Dynamic regrasping by in-hand orienting of grasped objects using non-dexterous robotic grippers. *Robotics Computer-Integrated Manuf.* 50, 114–131. doi:10.1016/j.rcim.2017.09.009

Sintov, A., Tslil, O., and Shapiro, A. (2016). Robotic swing-up regrasping manipulation based on the impulse–momentum approach and cLQR control. *IEEE Trans. Robotics* 32, 1079–1090. doi:10.1109/tro.2016.2593053

Sodhi, P., Kaess, M., Mukadam, M., and Anderson, S. (2021). "Learning tactile models for factor graph-based estimation," in *IEEE international conference on robotics and automation.*

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

Solak, G., and Jamone, L. (2023). Haptic exploration of unknown objects for robust in-hand manipulation. *IEEE Trans. Haptics* 16, 400–411. doi:10.1109/TOH.2023.3300439

Srinivasan, K., Eysenbach, B., Ha, S., Tan, J., and Finn, C. (2020). Learning to be safe: deep RL with a safety critic. *arXiv preprint arXiv:2010.14603*

Starke, J., Weiner, P., Crell, M., and Asfour, T. (2022). Semi-autonomous control of prosthetic hands based on multimodal sensing, human grasp demonstration and user intention. *Robotics Aut. Syst.* 154, 104123. doi:10.1016/j.robot.2022.104123

Stefanelli, E., Cordella, F., Gentile, C., and Zollo, L. (2023). Hand prosthesis sensorimotor control inspired by the human somatosensory system. *Robotics* 12, 136. doi:10.3390/robotics12050136

Su, Z., Fishel, J., Yamamoto, T., and Loeb, G. (2012). Use of tactile feedback to control exploratory movements to characterize object compliance. *Front. neurorobotics* 6, 7. doi:10.3389/fnbot.2012.00007

Sun, H., Kuchenbecker, K., and Martius, G. (2022). A soft thumb-sized vision-based sensor with accurate all-round force perception. *Nat. Mach. Intell.* 4, 135–145. doi:10.1038/s42256-021-00439-3

Sun, J., King, J. P., and Pollard, N. S. (2021). Characterizing continuous manipulation families for dexterous soft robot hands. *Front. Robotics AI* 8, 645290. doi:10.3389/frobt.2021.645290

Sundaralingam, B., and Hermans, T. (2018). "Geometric in-hand regrasp planning: alternating optimization of finger gaits and in-grasp manipulation," in *IEEE international conference on robotics and automation (ICRA)*, 231–238.

Suomalainen, M., Karayiannidis, Y., and Kyrki, V. (2022). A survey of robot manipulation in contact. *Robotics Aut. Syst.* 156, 104224. doi:10.1016/j.robot.2022.104224

Tao, L., Zhang, J., Bowman, M., and Zhang, X. (2023). "A multi-agent approach for adaptive finger cooperation in learning-based in-hand manipulation," in *IEEE international conference on robotics and automation (ICRA)*, 3897–3903. doi:10.1109/ICRA48891.2023.10160909

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

Taylor, I. H., Dong, S., and Rodriguez, A. (2022). "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)*, 10781–10787.

Tegin, J., and Wikander, J. (2005). Tactile sensing in intelligent robotic manipulation—a review. *Indus. Robot An Int. Jou.* 32, 64–70. doi:10.1108/01439910510573318

Terasaki, H., and Hasegawa, T. (1998). Motion planning of intelligent manipulation by a parallel two-fingered gripper equipped with a simple rotating mechanism. *IEEE Trans. Robotics Automation* 14, 207–219. doi:10.1109/70.681241

Todorov, E., Erez, T., and Tassa, Y. (2012). "Mujoco: a physics engine for model-based control," in *2012 IEEE/RSJ international conference on intelligent robots and systems* (IEEE), 5026–5033.

Toledo, L. V. O., Giardini Lahr, G. J., and Caurin, G. A. P. (2021). "In-hand manipulation via deep reinforcement learning for industrial robots," in *Multibody mechatronic systems.* Editors M. Pucheta, A. Cardona, S. Preidikman, and R. Hecker (Cham: Springer International Publishing), 222–228.

Toskov, J., Newbury, R., Mukadam, M., Kulic, D., and Cosgun, A. (2023). "In-hand gravitational pivoting using tactile sensing," in *Proceedings of the 6th conference on robot learning. (PMLR), vol. 205 of Proceedings of machine learning research.* Editors K. Liu, D. Kulic, and J. Ichnowski, 2284–2293.

Tournassoud, P., Lozano-Perez, T., and Mazer, E. (1987). Regrasping. *IEEE Int. Conf. Robotics Automation* 4, 1924–1928. doi:10.1109/ROBOT.1987.1087910

Townsend, W. (2000). The barretthand grasper – programmably flexible part handling and assembly. *Industrial Robot An Int. J.* 27, 181–188. doi:10.1108/01439910010371597

Van Baar, J., Sullivan, A., Cordorel, R., Jha, D., Romeres, D., and Nikovski, D. (2019). "Sim-to-real transfer learning using robustified controllers in robotic tasks involving complex dynamics," in *IEEE international conference on robotics and automation (ICRA)*, 6001–6007.

Van Hoof, H., Hermans, T., Neumann, G., and Peters, J. (2015). "Learning robot in-hand manipulation with tactile features," in *IEEE-RAS international conference on humanoid robots 2015-decem*, 121–127. doi:10.1109/HUMANOIDS.2015.7363524

Veiga, F., Akrour, R., and Peters, J. (2020). Hierarchical tactile-based control decomposition of dexterous in-hand manipulation tasks. *Front. Robotics AI* 7, 521448. doi:10.3389/frobt.2020.521448

Viña, B. F. E., Karayiannidis, Y., Pauwels, K., Smith, C., and Kragic, D. (2015). "In-hand manipulation using gravity and controlled slip," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 5636–5641.

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

Wang, S., Lambeta, M., Chou, P.-W., and Calandra, R. (2022). Tacto: a fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors. *IEEE Rob. & Aut. Let.* 7, 3930–3937. doi:10.1109/lra.2022.3146945

Wang, X., Chen, Y., and Zhu, W. (2021). A survey on curriculum learning. *IEEE Trans. Pattern Analysis Mach. Intell.* 44, 4555–4576. doi:10.1109/tpami.2021.3069908

Wang, Z., Garrett, C. R., Kaelbling, L. P., and Lozano-Pérez, T. (2018). Active model learning and diverse action sampling for task and motion planning. *IEEE/RSJ Int. Conf. Int. Rob. & Sys*, 4107–4114. doi:10.1109/iros.2018.8594027

Watkins, C. J., and Dayan, P. (1992). Q-learning. *Mach. Learn.* 8, 279–292. doi:10.1023/a:1022676722315

Wei, D., Sun, G., Ren, Z., Li, S., Shao, Z., Li, X., et al. (2023). In-hand re-grasp manipulation with passive dynamic actions via imitation learning. arXiv preprint arXiv:2309.15455

Wei, D., and Xu, H. (2023). A wearable robotic hand for hand-over-hand imitation learning. arXiv preprint arXiv:2309.14860

Wen, B., Mitash, C., Soorian, S., Kimmel, A., Sintov, A., and Bekris, K. E. (2020). "Robust, occlusion-aware pose estimation for objects grasped by adaptive hands," in *IEEE international conference on robotics and automation (ICRA)*, 6210–6217. doi:10.1109/ICRA40945.2020.9197350

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

Xu, J., Koo, T. J., and Li, Z. (2007). "Finger gaits planning for multifingered manipulation," in *IEEE/RSJ international conference on intelligent robots and systems*, 2932–2937. doi:10.1109/IROS.2007.4399189

Yang, L., Huang, B., Li, Q., Tsai, Y.-Y., Lee, W. W., Song, C., et al. (2023). Tacgnn: learning tactile-based in-hand manipulation with a blind robot using hierarchical graph neural network. *IEEE Robotics Automation Lett.* 8, 3605–3612. doi:10.1109/LRA.2023.3264759

Yin, Z.-H., Huang, B., Qin, Y., Chen, Q., and Wang, X. (2023). "Rotating without seeing: towards in-hand dexterity through touch," in *Robotics: science and systems*.

Yousef, H., Boukallel, M., and Althoefer, K. (2011). Tactile sensing for dexterous in-hand manipulation in robotics—a review. *Sensors Actuators A Phys.* 167, 171–187. doi:10.1016/j.sna.2011.02.038

Yuan, S., Shao, L., Yako, C. L., Gruebele, A., and Salisbury, J. K. (2020). "Design and control of roller grasper v2 for in-hand manipulation," in *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 9151–9158. doi:10.1109/IROS45743.2020.9340953

Yuan, W., Dong, S., and Adelson, E. H. (2017). Gelsight: high-resolution robot tactile sensors for estimating geometry and force. *Sensors* 17, 2762. doi:10.3390/s17122762

Yuan, Y., Che, H., Qin, Y., Huang, B., Yin, Z.-H., Lee, K.-W., et al. (2023). Robot synesthesia: in-hand manipulation with visuotactile sensing. *arXiv Prepr. arXiv:2312.01853*.

Zeng, A., Song, S., Lee, J., Rodriguez, A., and Funkhouser, T. (2020). Tossingbot: learning to throw arbitrary objects with residual physics. *IEEE Trans. Robotics* 36, 1307–1319. doi:10.1109/TRO.2020.2988642

Zeng, A., Song, S., Yu, K.-T., Donlon, E., Hogan, F. R., Bauza, M., et al. (2018). "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)*, 3750–3757.

Zhang, T., McCarthy, Z., Jow, O., Lee, D., Chen, X., Goldberg, K., et al. (2018). "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *IEEE international conference on robotics and automation (ICRA)*, 5628–5635.

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

Zhao, W., Queralta, J. P., and Westerlund, T. (2020b). "Sim-to-real transfer in deep reinforcement learning for robotics: a survey," in *IEEE symposium series on computational intelligence (SSCI)*, 737–744.

Zheng, B., Verma, S., Zhou, J., Tsang, I. W., and Chen, F. (2022). Imitation learning: progress, taxonomies and challenges. *IEEE Trans. Neural Netw. Learn. Syst.*, 1–16. doi:10.1109/LRA.2018.2869644

Zhou, J., Yi, J., Chen, X., Liu, Z., and Wang, Z. (2018). Bcl-13: a 13-dof soft robotic hand for dexterous grasping and in-hand manipulation. *IEEE Robotics Automation Lett.* 3, 3379–3386. doi:10.1109/LRA.2018.2851360

Zhu, H., Gupta, A., Rajeswaran, A., Levine, S., and Kumar, V. (2019). "Dexterous manipulation with deep reinforcement learning: efficient, general, and low-cost," in *IEEE international conference on robotics and automation (ICRA)*, 3651–3657.

Zimmer, M., Viappiani, P., and Weng, P. (2014). "Teacher-student framework: a reinforcement learning approach," in *AAMAS workshop autonomous robots and multirobot systems*.

Zuo, S., Li, J., and Dong, M. (2021). Design, modeling, and manipulability evaluation of a novel four-dof parallel gripper for dexterous in-hand manipulation. *J. Mech. Sci. Technol.* 35, 3145–3160. doi:10.1007/s12206-021-0636-7