



Review

Human-like dexterous manipulation for anthropomorphic five-fingered hands: A review

Yayu Huang^{a,b}, Dongxuan Fan^{a,b}, Haonan Duan^{a,b}, Dashun Yan^{a,b}, Wen Qi^{a,b}, Jia Sun^a, Qian Liu^a, Peng Wang^{a,b,c,*}

^a Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

^b School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

^c Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong 999077, China

ARTICLE INFO

Article history:

Received 7 September 2024

Revised 8 December 2024

Accepted 7 January 2025

Available online 14 January 2025

Keywords:

Anthropomorphic five-fingered hands

Human-like dexterous manipulation

Robot learning

Review

ABSTRACT

Humans excel at dexterous manipulation; however, achieving human-level dexterity remains a significant challenge for robots. Technological breakthroughs in the design of anthropomorphic robotic hands, as well as advancements in visual and tactile perception, have demonstrated significant advantages in addressing this issue. However, coping with the inevitable uncertainty caused by unstructured and dynamic environments in human-like dexterous manipulation tasks, especially for anthropomorphic five-fingered hands, remains an open problem. In this paper, we present a focused review of human-like dexterous manipulation for anthropomorphic five-fingered hands. We begin by defining human-like dexterity and outlining the tasks associated with human-like robot dexterous manipulation. Subsequently, we delve into anthropomorphism and anthropomorphic five-fingered hands, covering definitions, robotic design, and evaluation criteria. Furthermore, we review the learning methods for achieving human-like dexterity in anthropomorphic five-fingered hands, including imitation learning, reinforcement learning and their integration. Finally, we discuss the existing challenges and propose future research directions. This review aims to stimulate interest in scientific research and future applications.

© 2025 The Author(s). Published by Elsevier B.V. on behalf of Shandong University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

In the past few decades, the development of biological sciences and the advancement of artificial intelligence technology have greatly fostered the growth of the robotics community. Robots are playing an increasingly crucial role in society and are widely used in industrial production [1–3], household services [4–6], engineering applications [7,8], and medical assistance [9–11]. While proficient in repetitive tasks such as assembly and object handling in structured settings, their capabilities in performing dexterous, human-like manipulations remain limited. Addressing this gap is crucial for elevating robots to a new level of functionality and intelligence (see Fig. 1).

Dexterous manipulation is crucial for advancing robotics by bridging the gap between mechanical functionality and adaptive, intelligent behavior [12]. This skill enables robots to operate effectively in complex, unstructured, and dynamic environments, enhancing their ability to perceive, learn, and adapt. **Human-like dexterous manipulation** refers to robots performing tasks

that demand high precision, fine motor control, and adaptability, closely mirroring human capabilities. It involves interacting with dynamic environments, executing tasks with dexterity, and adjusting to new challenges. By emulating human-like dexterity, robots can not only improve efficiency and precision but also integrate seamlessly into human-centric applications, fostering a more harmonious interaction between technology and humanity.

Central to achieving human-like dexterity is the development of **anthropomorphic five-fingered robotic hands**, designed to replicate the structure and functionality of human hands. Featuring five independent fingers with multiple DoFs, this design enables precise and flexible movement. As a result, robots can perform intricate tasks, such as delicate object manipulation, tool use, and collaborative interactions with humans. The anthropomorphic structure plays a crucial role in enhancing robot's interaction capabilities and provides the physical foundation for advanced control systems that support nuanced, dexterous handling across diverse scenarios. Such hands are indispensable for realizing human-like manipulation in robotics, serving as a cornerstone for high-precision, adaptable task execution.

Human-like dexterous manipulation and anthropomorphic five-fingered hands are inherently interdependent. Dexterous

* Corresponding author.

E-mail address: peng_wang@ia.ac.cn (P. Wang).



Fig. 1. Demonstration of dexterous manipulation. Human-level dexterous manipulation (top). Human-like dexterous manipulation of anthropomorphic five-fingered hands (bottom).

manipulation demands the precision and adaptability that anthropomorphic hands are uniquely equipped to provide, while the design of such hands is driven by the functional requirements of dexterous tasks. Together, they form a synergistic foundation, where anthropomorphic hands enable robots to perform complex manipulation tasks, and dexterous manipulation tasks, in turn, validate and guide the design evolution of these hands.

Achieving human-like dexterous manipulation with anthropomorphic five-fingered hands involves significant challenges in perception, learning, and structural design. Human-like dexterous manipulation requires precise sensory feedback and real-time perceptual integration to interact with complex, dynamic environments. While human dexterity leverages vision, proprioception, and touch to handle objects delicately and adaptively, robots face critical limitations. Current sensors, despite their ability to detect contact and force, struggle with seamless information integration and real-time interpretation of dynamic changes. These limitations impede robots' capacity to perform dexterous tasks with human-like precision. Bridging this perceptual gap is essential for enabling robots to adapt and operate with the required subtlety in real-world scenarios.

In addition to perception, another fundamental obstacle lies in learning. Effective human-like dexterous manipulation demands rapid learning and adaptability across diverse tasks, from basic grasping to complex tool use. Humans excel at this through observation, trial-and-error, and contextual adaptation, but robots lag significantly. Machine learning advancements have improved task-specific performance, yet generalizing across the variability of real-world contexts remains a formidable challenge. Robots also struggle to autonomously adapt to new situations or refine strategies based on feedback, limiting their performance in unstructured environments. Addressing these barriers is essential for achieving the required versatility and flexibility for

human-like dexterous manipulation.

Furthermore, structural design represents another critical challenge. The human hand's complex anatomy, featuring 27 bones, numerous joints, and highly adaptable muscles, enables fine motor control and diverse grasping techniques, essential for manipulating fragile or irregular objects. Replicating such dexterity in robotic hands requires advanced mechanical design, including high DoFs in the joints, force feedback integration, and compliant control strategies. Current robotic hands struggle with limitations in range of motion, precision in force application, and the ability to adapt to complex tasks. Additionally, designing hands that balance flexibility, durability, and cost remains a key challenge, as current solutions often involve trade-offs that hinder performance in real-world scenarios. Overcoming these limitations is essential for realizing robots capable of fine, adaptive manipulation in dynamic environments.

Table 1 presents an overview of previous reviews on the topics of "dexterous manipulation" and "anthropomorphic hands". These reviews focus on two main perspectives: one emphasizing how robots learn to manipulate their environment, with an emphasis on perception and learning, and the other exploring various aspects of anthropomorphic robot design, including mechanisms, actuation, sensing, and control. The literature includes several specific reviews that contribute to these perspectives. Xia et al. [23] conducted a review on perception methods for dexterous manipulation, while Kroemer et al. [18] delved into machine learning based manipulation techniques. Hua et al. [19] focused on robot learning with deep learning algorithms. Furthermore, targeted reviews were conducted on specific aspects of dexterous manipulation. For instance, Duan et al. [20] provided a comprehensive review on dexterous grasping methods, and Yu et al. [22] explored the use of deep reinforcement learning for multi-fingered robots. Taking a human perspective, Sobinov [21] presented a review of human hand dexterity and its

Table 1
The related reviews and corresponding topics.

References	Topic keywords	Journals	Year
Mattar [13]	Biomimetics, robotics hands, dexterous manipulation	Robotics and Autonomous Systems	2013
Melo et al. [14]	Mechanism, actuation, sensing, control	Ingeniería y Desarrollo	2014
Melchiorri et al. [15]	Mechanism, actuation, sensing, control	Springer Handbook of Robotics	2016
Piazza [16]	Development history, new trends, major applications	Annual Review of Control, Robotics, and Autonomous Systems	2019
Liu et al. [17]	Bioinspired embodiment, dexterous manipulation, intelligent sensing	IEEE Transactions on Industrial Informatics	2020
Kroemer et al. [18]	Robot learning, robot manipulation	Journal of Machine Learning Research	2021
Hua et al. [19]	Dexterous manipulation, robot learning	Sensors	2021
Duan et al. [20]	Deep learning, point cloud, robot grasping	Frontiers in Neurobotics	2021
Sobinov et al. [21]	Manual dexterity, neural mechanisms	Nature Reviews Neuroscience	2021
Yu et al. [22]	Dexterous manipulation, multi-fingered hand, reinforcement learning	Frontiers in Neurobotics	2022
Xia et al. [23]	Dexterous robotic manipulation, robotic hand, sensory perception	International Journal of Advanced Robotic Systems	2022
SamPATH et al. [24]	Dexterous hand, Learning-based manipulation, Robot manipulation	Cognitive Computation and Systems	2023

implications for inspiring anthropomorphic hands in robots. In a similar vein, Mattar [13] conducted a bio-inspired survey of robotic hands, and Liu et al. [17] employed a bio-inspired approach to enhance robotic dexterous manipulation. Recently, [24] reviewed the history of dexterous hands, provided an overview of existing multi-fingered hands, and discussed traditional and data-driven learning methods, as well as the challenges in manipulating these hands. These works collectively contribute to the ongoing progress and development of dexterous manipulation in robotics, and pave the way for the future development of more capable and versatile anthropomorphic hands.

Unlike the aforementioned articles that primarily discuss dexterous manipulation and anthropomorphic hands as separate topics, this review bridges the gap by emphasizing their intrinsic interdependence. By integrating the principles of human-like dexterous manipulation with the structural and functional design of anthropomorphic five-fingered hands, this paper provides a unified framework that not only deepens the understanding of their synergy but also offers a foundation for developing advanced methodologies that align with both functional and structural requirements.

This paper presents the first focused review on the promising topic of human-like dexterous manipulation for anthropomorphic five-fingered hands. It builds on prior reviews by addressing previously overlooked aspects, particularly in the domain of robot learning, and consolidates a substantial body of work to complement existing literature. Human-like dexterous manipulation has immense value across diverse real-world applications, from healthcare—where robotic hands perform delicate surgical tasks—to industrial assembly, enabling precise handling of intricate components. Its potential further extends to assistive robotics for individuals with disabilities and space exploration, where adaptability to dynamic environments is critical. Recognizing these applications, this review systematically examines the interplay between human-like dexterous manipulation and anthropomorphic five-fingered hands, offering insights for future advancements.

The remainder of the paper is structured as follows. Section 2 explores human-like dexterous manipulation, including its definition and typical tasks. Section 3 discusses anthropomorphism and anthropomorphic five-fingered hands, covering their definitions, topology, mechanical design, and evaluation metrics. Section 4 reviews learning methods for enabling dexterous manipulation in anthropomorphic hands. Challenges and future research directions are highlighted in Section 5, particularly from a robot learning perspective. Finally, Section 6 concludes the paper.

2. Human-like dexterity and dexterous manipulations

In the following subsections, we summarize several important aspects that define human-like dexterity and introduce the classification of dexterous manipulations at the human-level, as well as the human-like dexterous manipulation tasks at the robot-level.

2.1. Definitions of human-like dexterity

While there are various definitions for dexterous manipulation, a widely accepted consensus-based definition for human-like dexterous manipulation is currently unavailable. The challenge in defining human-like dexterous manipulation lies in distinguishing it from general dexterous manipulation. We define human-like dexterity as follows: **“Human-like dexterity refers to the capacity to approach or surpass human dexterity in manipulators, the ability to effectively share and utilize human infrastructure during task execution, and the skill to engage in natural interactions, integration, and collaboration with humans through learning.”** In summary, the key distinctions from general dexterity become evident when considering *tasks*, *manipulators*, and *learning*, as illustrated in Fig. 2.

Tasks, particularly those involving grasping, inherently exhibit complexity, requiring functional manipulation of tools or objects within confined spaces to ensure long-term accuracy and efficiency. These tasks predominantly involve manipulations at the fingertip level, imposing significant demands on the manipulator’s capabilities to effectively share and utilize human infrastructure in task execution.

Manipulators, much like our human hands, require rich perception, excellent kinematics, and coordination between fingers. Simple end-effectors are insufficient; dexterous manipulation necessitates complex, multi-fingered hands and advanced sensing. To master such intricacy, robots must embrace a remarkable level of learning capability.

Learning, in the context of dexterity, requires adapting to changes in the position and orientation of objects across different hand configurations. Unlike task-oriented programming with known object parameters in a single scene, dexterity is object-oriented, with different objects corresponding to distinct hand configurations. This necessitates learning through demonstration, experience, and exploration. Recent advances in learning capabilities empower robots to tackle even more demanding tasks effectively.

2.2. Human-like robot dexterous manipulation tasks

Human hand manipulation can be classified into two main categories: prehensile manipulations and non-prehensile manipulations [25]. Prehensile manipulations involve grasping and seizing objects to apply functional forces, typically stabilizing an object using the hand’s contact force without relying on external forces. Non-prehensile manipulations, on the other hand, involve using the whole hand or fingers to push or lift objects, encompassing actions like pushing, pulling, toppling, and rolling. Human-like robot dexterous manipulation tasks aim to replicate the abilities of human hand movements, including both prehensile and non-prehensile manipulations, as well as their hybrid combinations. Below, we describe several typical dexterous manipulation tasks.

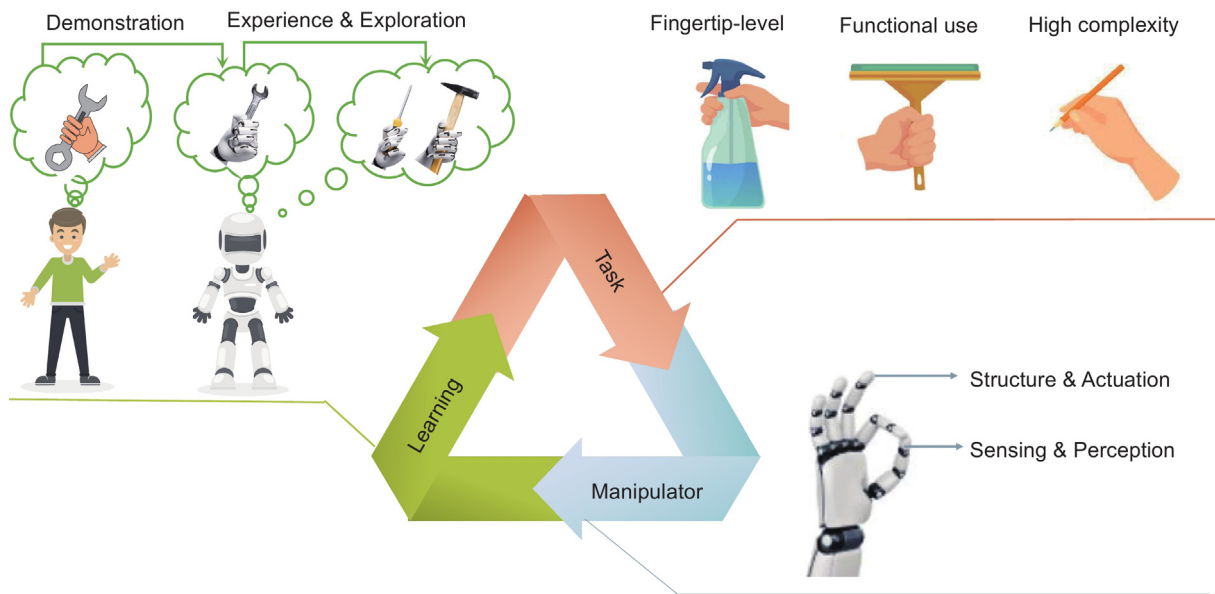


Fig. 2. Key elements in defining human-like hand dexterity: task, manipulator, and learning. The task is characterized by fine fingertip-level interactions, functional tool use, and high complexity. To execute these tasks effectively, manipulators with intricate structures and advanced actuation mechanisms, enhanced sensing, and perception capabilities are essential. Such manipulators facilitate advanced learning capabilities in robots, which often come from demonstration, experience, and exploration. These acquired skills empower robots to handle progressively complex tasks.

2.2.1. Grasping

Grasping, the act of securely holding an object with a hand, is a fundamental and widely used form of human-like dexterous manipulation that often serves as the initial step in complex tasks. This process presents several significant challenges that researchers are actively addressing. The first challenge is determining where to grasp, which involves detecting functional regions on the object's surface that are most suitable for grasping, a task that requires efficient algorithms to identify these regions [26–33]. The second challenge is how to grasp, which is about understanding and mimicking human manipulation classifications, such as power grasps and precision grasps, to perform tasks requiring adaptability and finesse [34–40]. The third challenge is keeping the grasp stable, which involves adjusting the force applied to the object to maintain its stability, a complex task that requires real-time feedback and control [41,42].

2.2.2. In-hand manipulation

To achieve increasingly human-like functions, focus has shifted to fingertip-level manipulations such as rotation, translation, regrasping, and so on, which are called in-hand manipulations. These tasks include object reorientation, where the object must be aligned with a target direction within specified tolerances, a common challenge that requires careful manipulation to ensure proper positioning [43–48]. Finger gaiting refers to the action of moving fingers in a coordinated manner to control the object's movement along the hand, similar to walking fingers [49–51]. Finger pivoting is the rotation of the object around a finger or a group of fingers, which is crucial for tasks that require the object to be turned in place [46]. Rolling involves the object rolling along the surface of the hand, which is essential for tasks requiring the object to be oriented in a specific way [52]. Sliding is the lateral movement of the object across the hand, which is important for tasks that require the object to be moved along a plane without rotation [53]. Regrasping involves changing the grip on the object, which is necessary when the initial grip is no longer suitable for the next phase of the task [54–56]. The challenges inherent in in-hand manipulation are threefold: balancing the object, which involves maintaining its stability to prevent it from

falling or slipping; making dynamic adjustments, which requires the system to adapt to the object's movement and the hand's position in real-time; and coordinating finger movements, which is essential for performing complex manipulations with precision.

2.2.3. Tool use

Tool use in robotics refers to a robot's ability to manipulate and utilize external objects or devices to perform tasks, which is a critical aspect of human-like dexterous manipulation. This capability is essential for robots to interact effectively with their surroundings and carry out complex tasks that require both dexterity and intelligence. For example, knocking involves applying controlled force to strike a surface, demonstrating precise force modulation [43–45]. Stirring with tools like spoons demands understanding the tool's purpose and the dynamics of containers and substances [57]. Scooping requires tools such as shovels to transfer materials, necessitating grip stability and adjustment to material properties [58]. Cutting tasks with scissors or knives require precise force to slice materials while maintaining alignment [37]. Pouring involves controlling tool angles and stability to avoid spillage [59,60]. Writing with pens or markers requires fine motor skills and environmental awareness [61], while screw-driving tasks focus on controlled torque and alignment [62]. Tool use involves more than mechanical action: precision and control are critical for tasks like knocking and cutting, where accurate force application prevents damage. Understanding the functionality of tools is also essential, as robots must grasp the specific purpose of each tool, such as recognizing that a pen creates marks on paper or a spoon is used for stirring liquids. Adaptability allows robots to handle various tools, each with distinct requirements for force, grip, and dexterity. Finally, environmental interaction challenges robots to interpret the properties of materials they manipulate, such as the viscosity of liquids during stirring or the resistance of materials when cutting.

2.2.4. Bi-manual manipulation

Bi-manual manipulation refers to the coordinated use of both hands to interact with and manipulate objects in the environment, enabling robots to perform more complex and nuanced

tasks. These actions are typically categorized into two types based on how the hands coordinate: loose coupling and tight coupling [63]. In loose coupling, the hands perform related tasks independently, such as holding a flashlight while using the other hand to examine an object, or picking up LEGO pieces with each hand and assembling them together [64]. In tight coupling, the movement of one hand directly influences the other, as when one hand holds an object steady while the other hand assembles parts, like screwing a bolt into place. Within tight coupling, actions can be either symmetric, such as lifting a box with both hands in the same direction [65–67], or asymmetric, like using one hand to pull a zipper while the other hand stabilizes the cloth, stabilizing a bottle with one hand while using the other to open it, or holding a fruit steady with one hand while using the other to slice it into segments [66,68–70]. Bi-manual manipulation involves several intricate aspects. One key consideration is the physical interaction between the hands and objects, which requires precise control over the forces and contacts to ensure successful task execution, particularly in dexterous tasks. Effective coordination and cooperation also depend on the precise timing and spatial arrangement of hand movements to avoid collisions and achieve a unified goal. Additionally, the wide diversity of tasks introduces further complexity, as robots must adapt to varying objects and environments, each requiring different manipulation strategies and levels of dexterity.

2.2.5. Teleoperation

Teleoperation is a key aspect of human–robot interaction, especially in scenarios requiring precise manipulation of objects in complex or hazardous environments. This approach is commonly used in tasks such as remote surgery, space exploration, and disaster recovery, where the operator is physically separated from the robot but still needs to perform dexterous actions. However, teleoperation presents several challenges that are currently the focus of active research. One of the most critical aspects of teleoperation is the human–machine interface, which captures human actions, such as hand movements or gestures, using either wearable devices like motion capture gloves or exoskeletons [71], or vision-based systems that track gestures via cameras [72,73]. Wearable devices provide high precision, while vision-based systems offer greater flexibility and eliminate the need for extra hardware, though they may struggle with accuracy in dynamic or cluttered environments. Another challenge is retargeting, where the operator’s movements must be mapped onto the robot’s hand configuration. This involves balancing between shaping the robot’s hand to resemble a human hand for intuitive control or prioritizing functional replication for greater precision. While mimicking the human shape may improve control, focusing on functionality allows for more adaptable and efficient manipulation, despite structural differences. Sophisticated algorithms are needed to ensure accurate movement replication while addressing these design trade-offs. Finally, effective feedback is crucial for precise control and immersion. Visual feedback, typically through cameras or VR systems [74,75], allows operators to monitor the environment, while haptic feedback simulates tactile sensations, helping operators feel forces or textures and improving their control over tasks.

3. Anthropomorphism and anthropomorphic five-fingered hands

Given the exceptional capabilities of the human hand, designers of robotic hands commonly pursue “anthropomorphism” to replicate human-like characteristics. The following subsections define anthropomorphism and distinguish it from dexterity. Additionally, we will explore the crucial factors involved in current robot topology and mechanical design processes, summarizing the criteria used to evaluate the anthropomorphic attributes of robots.

3.1. Definitions of anthropomorphism

“Anthropomorphism” derives from the Greek words “anthropos” (human) and “morphe” (shape or form). It refers to attributing uniquely human characteristics or behaviors to nonhuman entities. In the context of robotics, “anthropomorphism” is defined as the ability of a robot end-effector to mimic human hands, encompassing aspects such as shape, size, consistency, and general attributes (including color and temperature) [76].

Anthropomorphism and dexterity are two orthogonal concepts. Anthropomorphism pertains to the external properties of a robot, such as its physical appearance, without directly implying its functional capabilities. Dexterity, on the other hand, refers to the functional properties of a robot, specifically its ability to perform tasks and manipulations. While some robotic hands claim anthropomorphic features by resembling human hands in structure and DoFs, they may lack the necessary dexterity for performing complex grasping tasks efficiently. Conversely, robots lacking anthropomorphic features can demonstrate impressive dexterity, surpassing human-like capabilities in complex manipulations.

Despite anthropomorphism does not guaranteeing a high level of dexterity, it remains an ideal goal for designing robot end-effectors in the following scenarios:

- In human-oriented manipulation scenarios where tasks can be performed by either humans or robots, anthropomorphism aligns the size and shape of manipulated objects or tools with human manipulation requirements;
- In scenarios that require teleoperation, anthropomorphism enables the robot to directly reproduce human behavior through dedicated interface devices like data gloves;
- In scenarios for entertainment or assistance purposes, anthropomorphic robots, such as household robots, are specifically designed to emulate human-like behavior;
- In scenarios involving prosthetic devices or rehabilitation tasks, anthropomorphic robots play a crucial role in restoring the quality of human life.

3.2. Anthropomorphic five-fingered robot topology and mechanical design

The human hand is a marvel of intricate engineering, with bones, joints, muscles, tendons, and their connections working together seamlessly. This has driven advances in creating anthropomorphic five-fingered robots with similar hand functionality. The design of these robots encompasses numerous possibilities and concepts, influenced by key factors such as:

- **Mechanism:** Refers to the mechanical structure of the robotic hand, particularly the articulation of fingers and joints. Design considerations include determining the number of axes of rotation (DoF) for each joint and their parameters, including position and orientation.
- **Actuation:** Involves the actuators that drive the movements of the robotic hand. While electrical actuators are common, options like ultrasonic motors, pneumatic actuators, shape-memory alloy actuators, and twisted string actuators are also available. Actuation modes include in-site, where actuators are placed inside joints or connected links for streamlined mechanics but require size and mass management, and remote, where actuators drive joints externally, akin to biological structures, requiring attention to joint coupling.
- **Transmission:** The transmission system efficiently transfers motion from actuators to the fingers. Direct drive transmission is straightforward, while other options include chain transmission, gear transmission, tendon-based transmission, and belt and pulley transmission.

- **Sensing:** Sensors provide crucial real-time feedback to the robot's control system. Proprioceptive sensors monitor the robot's physical state (position, velocity, angle, etc.), while exteroceptive sensors gather environmental data (temperature, force, pressure, friction, etc.)
- **Control:** Dictates how the robot's movements and manipulations are orchestrated. Low-level control involves precise management of individual joints and actuators for specific finger movements and configurations, while high-level control encompasses strategic task planning, including sequence determination, object identification, localization, and optimal manipulation strategies.

Taking all the aforementioned factors into account, researchers have devoted significant efforts to developing experimental five-fingered robotic hands. In the quest to replicate the versatility and finesse of the human hand, the field of robotics has encountered a confluence of design philosophies. This exploration has led to the emergence of two primary dichotomies in robotic hand design: one distinguishing between rigid and soft hands based on their physical composition, and the other distinguishing among independent hands, coupled hands and underactuated hands based on their actuation strategy [16]. These classifications are not mutually exclusive but rather provide a framework for understanding the diverse approaches to achieving dexterity in robotic manipulation.

3.2.1. Rigid hands vs. soft hands

Rigid hands, characterized by fixed, non-flexible structures, are primarily designed to achieve high precision and exert significant force. These hands often adopt anthropomorphic designs to replicate the appearance and functionality of human hands, making them suitable for tasks that require fine-grained dexterity and repeatability. Early designs, based on rigid body dynamics and inverse kinematics, feature rigid body links, fixed DoFs joints, and minimal compliance. Such hands excel in applications that demand strong, repeatable, and precise grasps, such as assembly tasks, industrial manipulation, or any scenario where accuracy and force are paramount. However, the inherent rigidity of these designs limits their ability to interact with delicate or irregularly shaped objects, increasing the risk of damaging or inadequately grasping fragile items. The lack of mechanical flexibility also leads to higher control complexity, as adaptive strategies are required to compensate for the limited compliance, thereby increasing system complexity and the cognitive load on users. Additionally, in dynamic or uncontrolled environments, the rigid nature of these hands can pose safety risks during unexpected interactions, underscoring the need for greater compliance in robotic hand designs to ensure safe and secure operation.

Soft hands have emerged as a promising solution to address the limitations of rigid hands. Inspired by the human musculoskeletal system and invertebrate creatures, soft hands integrate flexible materials into joints, actuation systems, and contact surfaces, enhancing the hand's ability to adapt to diverse objects and environments. The primary advantage of soft hands lies in their high compliance and adaptability, which makes them ideal for handling irregular or fragile objects. Elastomers like silicone rubber, which can sustain large elastic deformations, offers advantages in ease of manufacturing, durability, and resistance to wear and tear. Furthermore, materials like Shape Memory Alloys (SMAs) and Shape Memory Polymers (SMPs) can change their properties in response to external stimuli, significantly expanding their functional capabilities. These advancements in soft robotics have led to the development of various soft robotic hands, including the RBO Hand 3 [77], Pisa/IIT SoftHand [78], Yale Multigrasp Hand [79], Bionic Hand [80], and SDM hand [81]. These hands are particularly suitable for unstructured environments where

high adaptability is needed, such as grasping deformable objects, interacting with the human body in rehabilitation tasks, or operating in hazardous environments where compliance is essential for safety. However, while soft hands excel in these applications, they generally sacrifice precision and force generation compared to rigid hands. As a result, they are less suitable for tasks that demand high levels of precision or strength, such as industrial manipulation or heavy-duty tasks.

3.2.2. Independent hands vs. coupled hands vs. underactuated hands

Independent hands represent one of the most intuitive strategies in robotic hand design, where the number of joints is equal to the number of DoFs. This design allows for precise control of each joint, providing high dexterity and accuracy in manipulation. For example, the UB Hand IV [82] employs a tendon network within each finger, driven by motors placed on the forearm, offering four DoFs per finger. A soft hand with 26 independent DoFs, proposed in [83], features M-shaped chambers as finger joints with vacuum actuation, capable of performing 32 of the 33 Feix taxonomy grasps [84]. While independent hands provide maximum precision and fine-grained control, they come with several challenges. These include increased complexity and cost due to the need for multiple actuators and motors, higher maintenance requirements, and the need for sophisticated control strategies to coordinate the actuators. Additionally, fitting multiple actuators into a compact hand design can be difficult, limiting the practical size of these hands.

Coupled hands offer a more compact and simplified actuation system compared to independent hands. These systems reduce the number of actuators by incorporating mechanical linkages that rigidly connect the movements of the finger joints. In these designs, the motion of one joint is directly influenced by the movement of others, simplifying the actuation mechanism and making the overall system more compact. A prominent example is the DLR/HIT Hand II [85], which uses super-flat BLDC motors and features three DoFs and four joints per finger, with the last two joints mechanically coupled. Similarly, Robonaut2's index and middle fingers feature a coupled design where the distal interphalangeal joint (DIP) maintains a fixed relationship with the proximal interphalangeal joint (PIP) [86]. The Schunk Five Finger Hand [87] adopts a coupling mechanism that goes beyond bending angles to include spreading angles. Additionally, the Integrated Linkage-Driven Dexterous Anthropomorphic (ILDA) hand [88] utilizes a four-bar linkage to couple the DIP joint's motion with the PIP joint's movement, enabling a more natural and human-like finger range while maintaining actuation efficiency. Although coupled hands limit individual finger control, they still offer commendable dexterity and are often used in applications where precision, reliability, and compact design are paramount. However, the trade-off is reduced flexibility in certain tasks that require independent finger motion.

Underactuated hands are designed to permit passive movements between DoFs, offering self-adaptive features that allow the joints to untangle dependencies under certain circumstances. This design allows independent motion in parts of the hand even when other parts encounter resistance. For instance, the Shadow Dexterous Hand [89] uses a single tendon to control both the PIP and DIP joints of the fingers. This allows the DIP joint to continue moving even if the PIP joint encounters obstruction, ensuring a high degree of adaptability. Similarly, the MCR-Hand III [90] features motors that actuate the PIP and MCP joints via spring linkages, while the DIP and PIP joints are coupled. Ryu et al. [91] employs four-bar linkages and gears to couple the DIP, PIP, and MCP joints, incorporating torsional springs to realize passive movements. Underactuated hands are particularly well-suited for tasks requiring compliance, as their ability to adapt

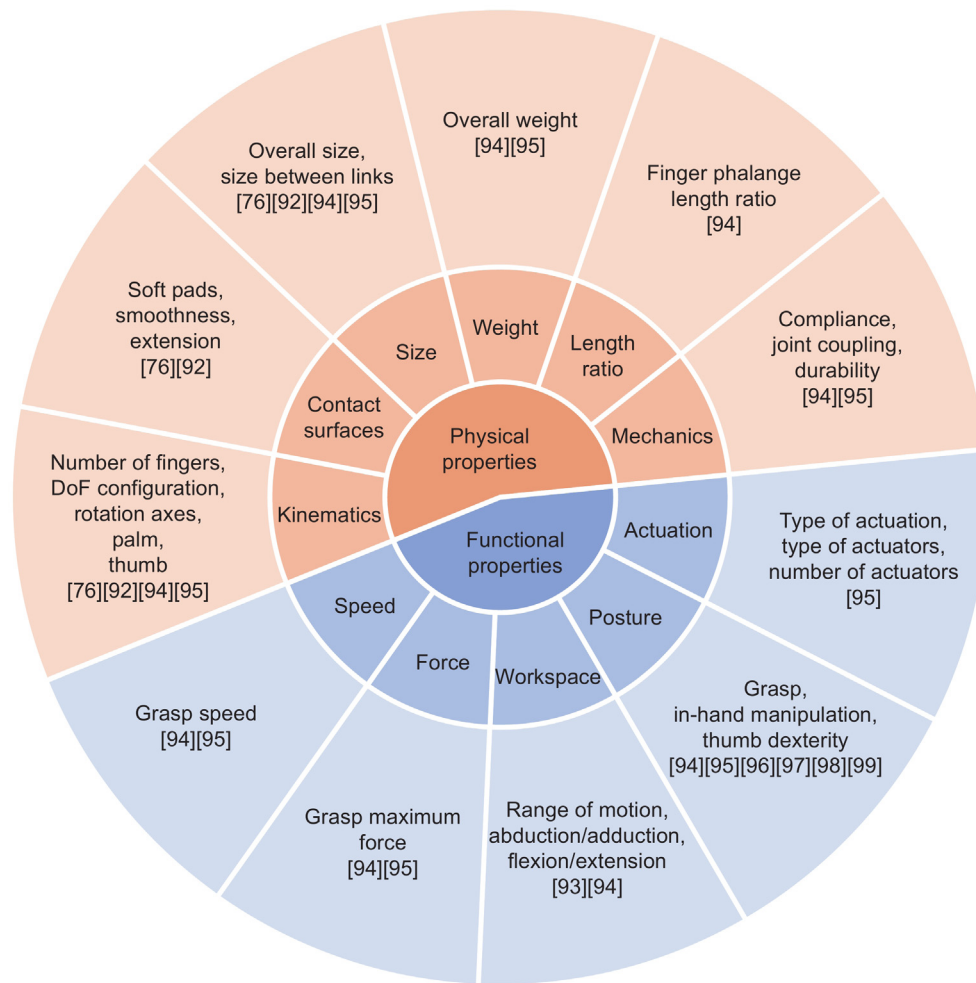


Fig. 3. Indicators to evaluate the level of anthropomorphism of the robotic hand. Physical properties, relating to the morphological characteristics of the robotic hand (orange). Functional properties, related to the functional performance of the robotic hand (blue).

to dynamic or uncertain environments makes them ideal for handling delicate, deformable, or irregular objects. However, this adaptability can come at the cost of precision and individual control over joint movement, which may limit their application in tasks requiring fine-grained dexterity and precision.

3.3. Evaluation of the anthropomorphism of robotic hands

Effective evaluation of anthropomorphic hands accomplishes two goals: assessing human-likeness in existing robotic hands and guiding the design of next-generation anthropomorphic hands. The comparison of hand features and capabilities has been carried out, considering their *physical properties* (such as size, weight, kinematics, etc.) and *functional properties* (including range of motion, speed, force, etc.), as illustrated in Fig. 3.

Earlier studies, such as [76,92], established anthropomorphism levels using weighted sums of kinematics, contact surfaces, and size. However, these criteria predominantly focus on basic morphological characteristics like the number of fingers and phalanges, lacking deeper discussion on critical similarities between robot and human hands such as workspace, grasp ability, speed, and strength. The study by Liarokapis et al. [93] proposes a method for quantifying anthropomorphism by comparing human and robot hands at two levels: finger phalanges workspaces and fingers base frames workspaces. Numerous comprehensive evaluations of robotic hand anthropomorphism have been undertaken. Liu et al. [94] introduced twelve indexes to quantify

prosthetic hand anthropomorphism through comparisons with human hands, covering factors like size, weight, range of motion, grasping speed, force, and posture. Belter et al. [95] conducted a qualitative assessment focusing on physical, actuation, and kinematic aspects, with special emphasis on thumb design due to its pivotal role in achieving human-like hand functionality.

Given the crucial role of manipulation in human hands, evaluating robotic hands' ability to perform manipulation tasks becomes imperative in assessing anthropomorphism. Zhou et al. [96] employed qualitative methods to evaluate robotic hand anthropomorphism through tasks such as grasping, in-hand manipulation, and thumb dexterity tests, focusing on the hand's success in executing these tasks. However, qualitative methods can be subjective, prompting researchers to strive for more objective quantitative assessments. Llop et al. [97] introduced a comprehensive protocol for systematically assessing robotic hand anthropomorphism, incorporating task-driven evaluations across twenty-six daily-life activity tasks, including eight grasp types, to quantitatively gauge the extent to which a robotic hand embodies anthropomorphism. Yet, experimental methods can be costly and time-consuming, making simulation methods appear as a reliable alternative. In [98], a simulation-based benchmarking approach was developed to assess anthropomorphic hands, leveraging human knowledge to generate effective grasping configurations and introducing new metrics for assessing similarity to human hand grasps. In a notable quantitative simulation study, Feix et al. [99] used the Gaussian process latent variable model

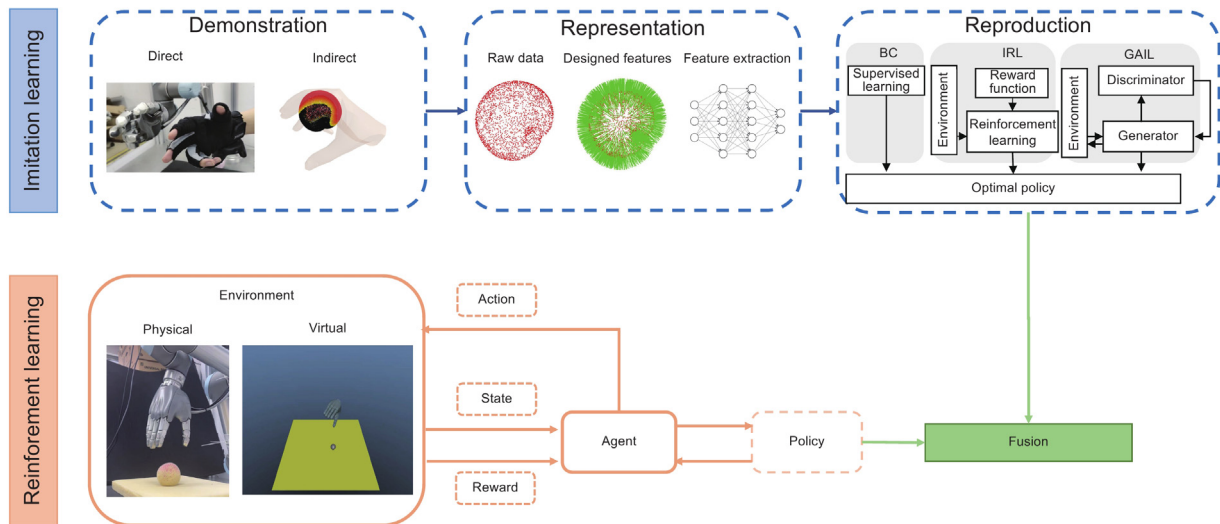


Fig. 4. Learning methods for anthropomorphic five-fingered hands fall into three categories: imitation learning (blue), reinforcement learning (orange), and a fusion of both (green).

(GP-LVM) to represent human and robotic fingertip workspaces in a low-dimensional manifold, quantifying anthropomorphic indices through calculations of similarity and overlap across twelve grasp types.

4. Learning human-like dexterous manipulation for anthropomorphic five-fingered hands

Learning stands undeniably as a fundamental component of dexterous manipulation, extending beyond a mere tool for specific robotic skills. Learning plays a pivotal role in facilitating dynamic adaptation to diverse, unstructured, and unpredictable scenarios [100]. Its significance extends beyond isolated skills, enabling skill generalization akin to the human capability to transition between tasks [101]. Moreover, learning empowers iterative refinement of intricate skills that progressively enhance dexterity, as opposed to conventional pre-programming [102]. As dexterity continues to improve, the intrinsic role of learning in manipulation has become indisputably fundamental in shaping adaptive, versatile, and human-like manipulative behavior.

Existing methods for human-like dexterous manipulation with anthropomorphic five-fingered hands are categorized as analytical or empirical based on success criteria [103]. Analytical methods involve solving constrained optimization problems by making assumptions about physical and geometric object properties, such as simplified contact models, Coulomb friction, and rigid-body modeling. Based on the assumptions, toolkits like GraspIt! [104] and SynGrasp [105] synthesize manipulation configurations that optimize objectives like grasp stability. However, analytical methods face limitations when dealing with unknown objects and real-time requirements. Empirical methods, also known as learning-based methods, emphasize object representation and perceptual processing, and are currently the mainstream approach for learning human-like dexterous manipulations with anthropomorphic five-fingered hands. They combine sensory input with prior knowledge, derived from human demonstrations or past experience, to generate and optimize candidate actions, yielding more general and robust performance than analytical methods. These empirical methods encompass imitation learning, reinforcement learning, and their integration (see Fig. 4).

4.1. Imitation learning for anthropomorphic five-fingered hands

Imitation learning (IL) involves replicating expert behavior to achieve optimal task performance. Through human demonstrations, the agent implicitly acquires prior knowledge about the world, which it uses to develop effective learning strategies. This approach enables robots to quickly acquire new skills and adapt to diverse environments with minimal demonstrations, enhancing their adaptability. Notably, IL seamlessly integrates with other algorithms like reinforcement learning, leveraging expert guidance to achieve excellent initial performance and increased learning efficiency.

Demonstration, representation, and reproduction constitute the entire process of IL. IL begins with capturing demonstrations through direct or indirect methods. The captured demonstrations are then processed to characterize the current state and environment. The represented features are used in algorithms to learn policies that mimic human behavior. Finally, the robot executes the learned policy and optimizes it based on performance.

4.1.1. Demonstration

Demonstration methods in IL are broadly categorized into direct and indirect approaches [106]. Direct demonstration involves obtaining samples directly from the robot primarily through two techniques: kinesthetic teaching and teleoperation teaching. In kinesthetic teaching, a human physically guides the robot to perform the desired actions, as shown in [107,108]. This method is intuitive and simple, making it suitable for straightforward tasks. However, it faces significant limitations when applied to robots with high DoFs, such as anthropomorphic robotic hands, where physically manipulating complex joints can become cumbersome and error-prone. Safety is another concern, particularly when teaching powerful or delicate robotic systems. Teleoperation teaching, on the other hand, leverages wearable devices, joysticks, or control panels to remotely guide robot actions [37, 60,109–111]. While this approach enhances safety by distancing the operator from the robot, it introduces challenges such as mechanical wear on the teleoperation devices and the potential for operator fatigue during extended sessions. Moreover, both direct methods often require task-specific setups and lack the inherent flexibility to adapt to diverse robotic configurations, thereby limiting their utility in generalized learning scenarios.

Indirect demonstration methods bypass the need for direct robot interaction and are broadly categorized into real-world

data collection and simulation-based data synthesis. The datasets generated from these methods exhibit wide applicability, supporting various research efforts in human-like dexterous manipulation, as shown in Table 2. In real-world methods, data is collected by observing human actions. Mocap systems track human hand motions with high accuracy using markers, as demonstrated in [112–116]. However, these systems require specialized equipment and calibration, limiting scalability. Markerless approaches address these challenges by estimating hand–object interactions from visual data. These approaches may use human annotation for higher accuracy, as seen in [117,118], or rely on automated vision-based estimation, as demonstrated in [119,120]. However, markerless methods can suffer from occlusions and reduced precision in complex scenarios, potentially compromising data quality. Simulation-based synthesis provides a scalable alternative for creating large datasets. Studies such as [121,122] use simple but effective strategies, generating random grasp poses in simulators and logging successful attempts into datasets. GraspIt!-based methods leverage optimization techniques like simulated annealing to generate stable grasps based on predefined metrics [118, 123–125], though their reliance on heuristics limits grasp diversity. Optimization methods for grasp generation are crucial for achieving stable and efficient robotic grasping. These methods can be classified into gradient-free and gradient-based optimization approaches. Gradient-free methods are particularly effective for non-convex landscapes, avoiding explicit gradient computation, which is useful when gradient information is noisy or unavailable. For example, [126,127] jointly optimize the weighted sum of multiple energy terms, such as contact stability, collision avoidance, and penetration minimization. These methods use the Metropolis-adjusted Langevin algorithm, which introduces stochasticity to explore the solution space more effectively and avoid being trapped in local optima. Similarly, [128] employs the Interior Point OPTimizer (IPOPT), a deterministic method well-suited for solving constrained optimization problems. IPOPT uses second-order information to navigate the optimization space efficiently, leading to robust solutions even for complex energy formulations. Gradient-based methods utilize grasp quality metric gradients to guide the optimization process, improving efficiency and precision. For example, [129] introduces a differentiable force closure estimator to optimize grasp stability, enabling the generation of diverse and stable grasps across different hand structures without pre-existing training data. By formulating the grasp synthesis process as a differentiable objective, this method efficiently computes gradients to refine grasp configurations. Despite their advantages, gradient-based methods face challenges, including low success rates, slow convergence, and difficulties with thin or low-quality object meshes. To address these issues, recent works have introduced advanced strategies. For instance, [130] improves optimization with robust penetration energy computation and better initialization techniques, leading to more reliable grasp outcomes. Additionally, [131] innovates in simulation frameworks by incorporating position-based dynamics integrators, which reduce computational complexity and accelerate the generation of realistic grasps.

4.1.2. Representation

Feature representation transforms observed demonstrations into an agent's internal format. Since real-world environments are intricate, gathered demonstrations often contain superfluous details, impeding subsequent IL. Hence, demonstrations need a concise, recognizable form applicable to the robot, termed a feature vector. This vector encompasses environment, object, and manipulator information, balancing training demands with computational efficiency. Raw data, rich in information and dimensions, can be directly used if adequate. However, it is often processed into useful information through crafted formulas, termed

designed features. Expert insights guide this task, adapted per the requirement. For instance, [137] retargets the obtained 3D positions of 21 skeletal human hand points to robot hand for dexterous manipulation. In [138], the depth camera recovers the point cloud, followed by segmentation, reduction, and calculation of normals. These surface normals, coupled with contact positions, characterize the visual information, forming a foundation for training in grasping. High-dimensional data often undergoes automatic mapping to low-dimensional space by feature extraction technology. For complex robots like anthropomorphic hands, original joint angles might be unwieldy, advocating mapping to synergy spaces, e.g., [139] maps hand-arm space to postural synergies subspace using Multiple Neural Networks. The similar idea can also be seen in [35,37,47,60,111,140].

4.1.3. Reproduction

The prime objective of IL is enabling robots to replicate and generate human behavior in novel, unseen scenarios, facilitated by IL algorithms. These algorithms, pivotal in achieving this goal, can be categorized into three main types: behavior cloning, inverse reinforcement learning, and adversarial imitation learning. We detail these algorithms below.

Behavior cloning. Behavior cloning (BC), a prominent algorithm within IL, embodies the principle of direct policy learning. The core idea is to map the teaching of the demonstrator to the actions of the learner. Since the demonstration data is provided with state–action labels, behavior cloning can be regarded as a subset of supervised learning. Essentially, the agent employs training data to construct a human behavior model that predicts actions based on states, using classification or regression methods to complete the task.

For cases where the learner's actions can be divided into discrete categories, classification methods are commonly employed [141]. Deng et al. [34,40] formulated grasping as a classification problem, where objects are assigned to one of six human-inspired grasp types: large wrap, small wrap, power, pinch, precision, and tripod. This classification directly guides the robot's grasp execution, providing a structured approach to mimic human manipulation strategies. Similarly, Li et al. [142] proposed a classification-based approach for grasp type selection, utilizing a neural network to determine grasp types such as power circular grasp, power heavy wrap grasp, power prismatic grasp, power thin grasp, precision circular grasp, and precision prismatic grasp based on perceived object features. This method enables the robot to achieve secure and stable grasps by aligning its actions with the physical characteristics of the object. Building on these principles, Santina et al. [143] used classification to sequence human-inspired action primitives, enabling robots to adaptively perform approach and grasp phases through the selection of one of ten predefined primitives. Differentiating from previous approaches, Hidalgo et al. [144] introduced a grasp posture predictor that operates at the point level, classifying each object surface point into medium wrap, tripod, or non-graspable areas. This granularity facilitates precise hand placement for effective grasp execution. Further enhancing adaptability, Rao et al. [39] introduced probabilistic classification, which predicts the likelihood of multiple grasp types based on detailed object attributes such as texture, stiffness, and fragility. By outputting a probability distribution, this approach emulates human decision-making under uncertainty, allowing the robot to flexibly adapt its grasping strategy to varying conditions. Classification methods, by mapping inputs to discrete action labels, excel in providing high-level action representations suited for tasks involving broad categorization or predefined action sets. These methods are computationally efficient and easier to train, making them appealing for straightforward robotic tasks. However, they lack the granularity

Table 2
Generalized human-like dexterous manipulation dataset.

Year	Dataset	Hand representation	Task	Method
2019	ObMan [123]	MANO	Grasp	Graspt!
2020	Contactpose [112]	MANO	Grasp	Mocap
2020	GRAB [113]	SMPL-X	Binamual manipulation	Mocap
2020	HO3D [119]	MANO	Grasp	Markerless (Estimation)
2020	DDGdata [124]	Shadow hand	Grasp	Graspt!
2021	KIT [114]	Master motor map (MMM)	Binamual manipulation	Mocap
2021	DexYCB [117]	MANO	Grasp	Markerless (human annotation)
2021	H2O [120]	MANO	Binamual manipulation	Markerless (estimation)
2022	OakInk [132]	MANO	Functional manipulation	Markerless
2022	HOI4D [133]	MANO	Functional manipulation	Markerless
2023	Arctic [115]	MANO	Binamual manipulation	Mocap
2023	HMDO [134]	Reconstructed hand model	Binamual manipulation	Markerless
2023	AffordPose [118]	MANO	Functional grasp	Markerless (human annotation) & Graspt!
2023	DexGraspNet [130]	Shadow hand	Grasp	Gradient-based optimization
2023	Fast-Grasp'D [131]	Shadow hand	Grasp	Gradient-based optimization
2023	GenDexGrasp [135]	Shadow Hand	Grasp	Gradient-based optimization
2024	TACO [116]	MANO	Binamual manipulation	Mocap & Markerless
2024	MultiGripperGrasp [125]	Shadow hand, Human hand	Grasp	Graspt! + Simulator
2024	Oakink2 [136]	MANO	Binamual manipulation	Mocap

necessary for low-level control, where precise, continuous action representations are critical.

Regression techniques enable precise control in robotic tasks by predicting continuous values for various parameters. These methods are categorized into probabilistic modeling, traditional machine learning, deep learning-based regression, and generative approaches, each with unique features and applications. Probabilistic methods, such as Gaussian Mixture Models (GMMs) and Gaussian Mixture Regression (GMR), have proven effective for modeling grasp trajectories. Huang et al. [128] used GMM-GMR for real-time grasp planning in dynamic environments. Katyara et al. [60] introduced “kernelized synergies”, where GMM captures synergistic trajectories, and GMR predicts reference trajectories for dexterous manipulation, adapting to new objects and conditions. These methods model human demonstrations probabilistically, ensuring smooth grasp predictions. Traditional machine learning methods focus on simpler regressors to predict grasp configurations from object features. Yuan et al. [26] used SVM-rank to predict configurations for a five-finger hand in a three-stage process, though it has limited scalability for high-dimensional data. Similarly, Gabellieri et al. [107] used a Decision Tree Regressor based on object point cloud dimensions, effectively transferring human expertise but limited by reliance on handcrafted features. Deep learning has enhanced regression models by learning from high-dimensional data. Ficuciello et al. [139] employed Multiple Neural Networks (MNN) to predict motor coefficients for anthropomorphic hands, enabling generalization across various objects. Song et al. [108] used multi-level CNNs to predict optimal grasp postures for dexterous hands. Duan et al. [145] introduced AHG-Net, an end-to-end framework that predicts hand configurations from single-view point clouds, enabling human-to-robot handovers. Ruppel and Zhang [146] integrated CNNs with trajectory optimization to teach dexterous manipulation from human demos, while Wu et al. [147] extended CNNs for grasp transfer across object categories with minimal labeled data. Emerging architectures, like Transformers, are advancing grasp planning by capturing global relationships in high-dimensional data. Xu et al. [148] introduced the Dexterous Grasp Transformer, which predicts diverse grasp poses from point clouds, modeling spatial and contextual dependencies more efficiently. Generative models, especially Conditional Variational Autoencoders (CVAEs), have transformed regression tasks by predicting diverse grasp configurations. Wei et al. [122] used a CVAE-based network to predict high-quality grasp configurations from single-view observations. Zhang et al. [126] enhanced multi-fingered grasping in clutter using CVAEs to produce semantic-rich contact maps. Shao et al. [127] applied Denoising Diffusion

Probabilistic Models (DDPMs) to generate bimanual grasp configurations, addressing the complexities of dual-hand manipulation. Regression methods address the limitations of classification by predicting continuous values, enabling precise and adaptable robotic manipulations essential for complex tasks. However, they often require significantly larger datasets to achieve robustness and are more sensitive to noise and variations in training data, which can affect their real-world performance.

BC provides a straightforward and efficient framework for IL by directly replicating expert demonstrations. Its simplicity enables rapid deployment in robotic systems, and its reliance on supervised learning makes it intuitive to implement. However, BC fundamentally lacks the ability to infer the intent behind demonstrated actions, focusing solely on reproducing observed behaviors. This limitation restricts its generalization to unseen scenarios or its adaptability to variations in task requirements. Moreover, its heavy dependence on the quantity and quality of expert data renders it vulnerable to biased or incomplete demonstrations. To overcome these challenges, approaches like inverse reinforcement learning aim to uncover the underlying objectives driving expert behavior, enabling more robust and flexible solutions.

Inverse reinforcement learning. Inverse reinforcement learning (IRL) is a typical branch of the IL method, which aims to evaluate actions by extracting a reward function that represent the intentions of the expert [149]. While both IRL and BC assume that state-action pairs are available, IRL learns a reward function that abstractly describes actions rather than directly mapping states to actions. This obtained reward function subsequently guides reinforcement learning to derive the ultimate control strategy. This approach is advantageous as reward functions often prove more concise than policies, particularly for plan-oriented tasks like dexterous manipulation in high dimensional space.

To address the complexities of designing task-specific reward functions for dexterous robotic hands, Orbik et al. [150] introduced an IRL algorithm that autonomously derives reward functions from human demonstrations. This approach eliminates the need for manual reward engineering, making it particularly valuable in tasks with high dimensionality and limited demonstration samples. The learned reward function serves as a proxy for the demonstrator’s expertise, enabling the robot to generalize behavior to unseen scenarios by inferring task-relevant priorities. However, due to the sparse distribution of samples in the expansive state-action space of dexterous tasks, traditional IRL methods often develop reward functions biased toward demonstrated actions, reducing generalization to unexplored areas. To

overcome this limitation, the authors enhanced the learning process by generating random samples to explore unvisited regions, normalizing reward values to stabilize training, and reducing the reward function's dimensionality to focus on task-relevant features. These strategies collectively improve the robustness, efficiency, and generalization of the reward function, allowing the robot to execute complex manipulation tasks such as grasping, in-hand manipulation, and tool use with greater autonomy and adaptability.

IRL, while providing a principled framework for uncovering reward functions, often relies on computationally intensive iterative processes that optimize policies based on inferred rewards. These processes can struggle to produce optimal results in high-dimensional spaces, where the complexity of the reward landscape amplifies computational demands. In contrast, generative adversarial imitation learning simplifies policy improvement by leveraging adversarial training to directly compare the behavior of experts and learners. This approach bypasses explicit reward inference, enabling a more direct and efficient alignment of policies with expert demonstrations.

Generative adversarial imitation learning. Generative adversarial imitation learning (GAIL) is a method that integrates Generative Adversarial Networks (GANs) with IL, aiming to directly learn policies from demonstrations without explicitly modeling a reward function [151]. The core idea of GAIL is framed as a competitive game between two components: the generator and the discriminator. The generator is trained to produce behaviors that mimic expert demonstrations, while the discriminator acts as a classifier to differentiate between the behaviors of the generator and the expert. Through iterative adversarial training, the generator's outputs are gradually optimized to become indistinguishable from expert behaviors, effectively aligning the learned policies with the demonstrated ones.

A practical example of GAIL's application is presented in [137]. This study focuses on imitating human hand movements, particularly in object manipulation, and transferring these to a 29-degree-of-freedom anthropomorphic robotic hand model. The process involves capturing human hand motion using an advanced hand pose estimator and retargeting the data through inverse kinematics and Particle Swarm Optimization (PSO), optimizing for task-specific objectives to bridge the gap between human and robotic hands. In this framework, GAIL is employed to learn a policy network from the retargeted demonstrations executed in a virtual environment. The generator network is trained to mimic the expert demonstrations, while the discriminator learns to differentiate between the expert and generated behaviors. Through adversarial optimization, the generator progressively aligns its policies with those of the expert. Key features considered during policy learning include the relative distance and velocity between the hand and the object, the local rotations and velocities of the hand's joints, and the minimum contact distance between the end-effectors and the object. The effectiveness of this GAIL-based framework is demonstrated by its ability to enable the robotic hand to autonomously grasp objects in unseen initial conditions within the virtual environment. This success underscores the potential of GAIL in learning human-like grasping and manipulation tasks, offering a promising approach for extending IL techniques to real-world robotic applications.

While GAIL offers advantages such as lower computational costs and enhanced generalization, dealing with high-dimensional models can pose challenges to the learning process. When attempting to generalize to similar conditions with added initial noises, the model's success rate can significantly decrease.

In summary, IL has become a key approach for enabling human-like dexterous manipulation with anthropomorphic five-fingered hands. IL methods such as BC, IRL, and GAIL each offer

distinct advantages but also come with limitations. BC is simple and intuitive but struggles with generalization and long-horizon tasks. IRL can infer reward functions and generalize better, yet it faces computational challenges and is sensitive to noisy demonstrations. GAIL improves upon this by using adversarial training to learn policies more efficiently, but still requires significant training data and computational resources. Table 3 summarizes various IL methods for anthropomorphic five-fingered hands. Despite their effectiveness, these methods still face challenges in scaling to complex, unstructured environments. As a result, reinforcement learning is gaining traction as it allows robots to refine their manipulation skills through direct interaction with the environment, offering a more flexible and adaptive approach to handling a wider range of tasks.

4.2. Reinforcement learning for anthropomorphic five-fingered hands

Humans inherently learn through interaction with their environment, a natural process that is observed even in infants who, without explicit instruction, develop perceptual and motor skills through trial and error. Reinforcement learning (RL) draws inspiration from this psychological principle, adapting it into the domain of artificial intelligence. In RL, an agent learns by interacting with a dynamic environment, receiving feedback in the form of a "reward", which guides its decision-making. This approach partitions the real world into two core components: the agent and the environment. The agent performs actions that influence the environment and receives rewards as feedback, driving its learning process.

Beyond the agent–environment interaction, RL frameworks incorporate four fundamental components: policy, reward function, value function, and an optional model of the environment. These elements collectively enable RL to tackle decision-making in complex, sequential, and uncertain scenarios.

Policy. Policy defines the agent's strategy by mapping perceived environmental states to corresponding actions. Analogous to stimulus–response mechanisms in psychology, policies range from simple lookup tables to sophisticated mappings requiring extensive computation. In more complex environments, policies often rely on models to optimize action sequences for maximum cumulative rewards. Policies can be deterministic (a specific action for each state) or stochastic (probabilistic distributions over actions).

Reward Function. The reward function is the objective feedback of the environment to the agent, which sets the RL objective. At each time step, the environment assigns an integer or scalar value to the state–action pairs, known as a reward. The long-term goal of the agent is to maximize the cumulative reward, as the reward gauges the quality of behavior. While the agent cannot directly control how rewards are assigned, it can influence reward outcomes indirectly by altering its actions and the subsequent state transitions.

Value Function. If the reward function evaluates short-term benefits, the value function encapsulates long-term advantages. The value function assigns an estimate to a state, representing the cumulative future rewards achievable from that state onward. Thus, even states with low immediate rewards may hold significant value if they lead to high-reward states in subsequent steps.

Model of Environment. The presence of the model of environment is optional. If the model of environment is available, the RL algorithm is termed model-based. In this method, predictions about the upcoming state and reward are made based on the current state and action, often used for planning purposes.

Table 3
Imitation learning for anthropomorphic five-fingered hands.

Category	Ref.	Year	Demonstration	Robot hand	Task	Experimental validation
BC	[128]	2013	Graspl! & Gradient-free optimization	iCub Hand	Grasp	Sim
	[139]	2016	Teleoperation	SCHUNK S5FH	Grasp	Real
	[26]	2017	Human annotation	Human Hand	Grasp	Real
	[108]	2018	Kinesthetic teaching	Shadow	Grasp	Real & Sim
	[34]	2019	Human annotation	Shadow	Grasp	Sim
	[143]	2019	Human annotation	Pisa/IIT SoftHand	Grasp	Real
	[43]	2019	Markerless	Adroit	Ham a nail, open a door, relocation, reposition a pen	Sim
	[152]	2019	Kinesthetic teaching	DLR/HIT II, Human Hand	Grasp	Real & Sim
	[39]	2020	Human annotation	AR10	Grasp	Real & Sim
	[142]	2020	Indirect	AR10	Grasp	Real
	[107]	2020	Kinesthetic teaching	Pisa/IIT SoftHand	Grasp	Real
	[146]	2020	Mocap	Shadow	open a bottle, pick and place, wipe	Real
	[60]	2020	Teleoperation	Dexmart Hand	Pour coffee, open toolbox latches, grasp and manipulate, play the board game	Real & Sim
	[40]	2021	Human annotation	Shadow	Grasp	Real & Sim
	[122]	2022	Simulator-based synthesis	DLR/HIT II	Grasp	Real & Sim
	[121]	2022	Simulator-based synthesis	DLR/HIT II	Grasp	Real & Sim
	[145]	2023	Simulator-based synthesis	DLR/HIT II	Handover	Real & Sim
	[144]	2023	Human annotation	Shadow	Grasp	Real
	[56]	2023	Human annotation	Shadow	Nonprehensile pregrasp	Sim
	[147]	2023	Human annotation & gradient-based optimization	Shadow	Functional grasp	Real & Sim
[33]	2024	Human annotation & gradient-based optimization	Shadow, DLR/HIT II, SCHUNK S5FH	Functional grasp	Real & Sim	
[153]	2024	Teleoperation	Shadow	Continuous grasp	Real & Sim	
[126]	2024	Gradient-free optimization	DLR/HIT II	Grasp	Real & Sim	
[127]	2024	Gradient-free optimization	Shadow	Bimanual grasp	Sim	
[148]	2024	Gradient-based optimization	Shadow	Grasp	Sim	
IRL	[150]	2021	Markerless	Adroit	Ham a nail, open a door, relocation, reposition a pen	Sim
GAIL	[137]	2018	Markerless	Modular Prosthetic Limb	Grasp	Sim

4.2.1. Model-based methods

Model-based methods rely on either an explicit or a learned representation of the environment, which serves as a foundation for predicting future states and rewards. By leveraging this predictive capability, agents can strategically plan their actions, considering not just immediate outcomes but also long-term consequences. This foresight allows for efficient exploration and optimization of strategies, significantly reducing the need for direct interaction with the real environment. Simulations within the environment model enable agents to evaluate potential strategies and refine their decisions, leading to faster convergence and improved performance. Furthermore, model-based approaches offer enhanced interpretability, as the environment model provides insights into the system's dynamics and underlying structure, making them particularly valuable for solving complex, high-dimensional tasks.

When the environment model is available, optimal strategies can be derived using model-based methods such as value iteration and policy iteration. These algorithms work by evaluating and refining policies based on predictions of future states and rewards. In value iteration, the algorithm iteratively updates the value of each state by considering the expected rewards from all possible actions, eventually converging on the optimal value function. The optimal policy is then derived from this value function by choosing the action that leads to the highest value in each state. Policy iteration, on the other hand, alternates between evaluating the current policy and improving it. In the evaluation step, the algorithm computes the value function for the current policy. In the improvement step, it updates the policy by choosing the action that maximizes the expected value at each state, repeating these steps until convergence to an optimal policy.

The development of model-based RL for anthropomorphic five-fingered hands has been significantly advanced through various studies, each contributing to the refinement and application

of these techniques. Gupta et al. [154] integrate Guided Policy Search (GPS) with RL to optimize policies for a soft robotic hand, employing a model-based approach to optimize control policies. The GPS framework is used to train a neural network policy on expert demonstrations, enabling the hand to replicate complex motions like valve turning and abacus manipulation. By fitting time-varying linear models to the system's dynamics, the method improves task execution efficiency. The GPS algorithm is extended to not only learn from multiple demonstrations but also to generalize these skills across various initial states, showcasing an improvement over traditional model-based methods by incorporating deep learning for policy optimization and generalization. Kumar et al. [155] utilize RL to control a pneumatically-actuated hand by applying time-varying linear models. This approach highlights how model-based techniques can effectively manage dynamic, real-time constraints in robotic hands. By incorporating a system's dynamic model, the method optimizes hand positioning and object manipulation while minimizing the need for excessive trial-and-error interactions with the environment. Such a strategy proves particularly advantageous in systems where accurately modeling physical dynamics is essential for precise manipulation. Nagabandi et al. [61] present online planning with deep dynamics models (PDDM), a novel RL approach that leverages deep dynamics models to enhance dexterous manipulation with multi-fingered hands. This method employs a policy-based framework that directly links data to control strategies, using predictions from the model to facilitate real-time planning and decision-making. One of the key innovations is the use of deep neural networks to model complex system dynamics, significantly improving the model's ability to handle intricate interactions. Additionally, PDDM integrates model-predictive control (MPC), a technique that optimizes control actions based on predictions of future system states, enabling efficient trajectory optimization at each step. The approach

demonstrates remarkable sample efficiency, learning advanced manipulation skills with minimal real-world data, and shows strong generalization capabilities, allowing the learned policy to be applied to a variety of tasks in different settings. Similarly, Omer et al. [156] integrate MPC with the Actor–Critic framework, a RL framework that simultaneously evaluates policies (the Actor) and estimates value functions (the Critic). This integration further enhances the real-time planning and adaptation of control policies, combining the strengths of both MPC's optimal control and the learning capability of the Actor–Critic model. By evaluating both actions and policies in parallel, this approach allows for more effective handling of complex, dynamic tasks such as valve rotation and object manipulation, offering improved flexibility and performance.

Model-based RL methods, despite their effectiveness, face significant challenges, primarily due to their heavy dependence on the accuracy of the environment model. In complex and dynamic environments where explicit representations are lacking, obtaining a precise model can be exceedingly difficult. These methods are not only computationally intensive but also sensitive to errors in the model, which can result in suboptimal policies. Furthermore, the policies derived may have limited generalization capabilities and require continuous updates to adapt to environmental changes, highlighting the need for approaches that can handle uncertainties and inaccuracies inherent in real-world applications.

4.2.2. Model-free methods

Model-free methods refer to a class of RL algorithms that do not require a model of the environment to make decisions. In contrast to model-based methods, which rely on an explicit model to predict future states and rewards, model-free methods directly learn optimal policies or value functions through interactions with the environment, typically by trial and error.

Model-free policy-based methods have played a crucial role in advancing dexterous robotic manipulation, particularly for high-dimensional action spaces. Trust Region Policy Optimization (TRPO) [157], a foundational approach, optimizes policies without requiring an explicit environment model. It introduces trust region updates to ensure stable policy changes and uses natural gradient optimization for more efficient learning in complex control tasks. In [158], TRPO was used to train an anthropomorphic robotic hand for grasping tasks, demonstrating its ability to generalize across diverse object geometries and material properties. The learned policy was not object-specific and showed the ability to generalize to novel objects with distinct characteristics. When augmented with touch sensing, the policy exhibited improved robustness, maintaining performance across varying levels of environmental noise. Proximal Policy Optimization (PPO) [159], a successor to TRPO, simplified the trust region mechanism by introducing clipped updates, maintaining training stability while improving computational efficiency. For instance, in [46], PPO was applied to train the Shadow Dexterous Hand for vision-based object reorientation tasks. The approach utilized a distributed RL system to enhance data collection efficiency and employed extensive simulation randomization to achieve successful sim-to-real transfer. This method demonstrated robustness to variations in physical properties, friction coefficients, and object appearances, marking a significant advancement in dexterous manipulation for robotic hands. Building upon PPO, Soft PPO integrated attention mechanisms to enhance task-specific performance, particularly for complex manipulation scenarios [160]. By focusing on relevant regions of the environment, the approach optimized the exploration–exploitation balance and achieved over 90% success rates in complex grasping tasks. Similarly, domain randomization, a key technique for improving robustness in real-world

environments, was effectively combined with PPO in [52] to enable zero-shot transfer for ball rotation tasks. Further extending PPO's capabilities, hierarchical curriculum designs and dense reward functions have been employed to accelerate learning in functional grasping tasks. For instance, [54] used a dense multi-component reward function to reduce reliance on expert demonstrations and enable efficient learning of functional grasps on novel objects, such as drills, spray bottles, and mugs. The approach achieved a 94.1% success rate and utilized a two-stage curriculum to improve convergence stability. Meanwhile, the PGDM framework introduced in [161] leveraged pre-grasp postures—specific hand configurations that prepare for object interaction—to reduce the complexity of exploration in dexterous manipulation tasks. This innovation enabled the framework to effectively learn complex manipulation behaviors with minimal task-specific engineering or hyperparameter tuning, allowing PPO to achieve performance comparable to methods requiring more extensive supervision. In a recent advancement, PPO has been effectively applied to human-assisting dexterous grasping tasks [162]. The article presents a novel policy-based RL approach for human-assisting dexterous grasping, focusing primarily on the development and implementation of a RL method to control a robotic hand's fingers for assisting users in grasping objects. The RL component, which is the core of the method, is designed to address the challenges of adapting to diverse user intentions and object geometries in real-time. Experiments conducted in the paper demonstrate the superiority of the proposed method over baselines in terms of success rate, posture similarity to human grasps, and stability of the grasped object. The method significantly outperforms baselines across various metrics, including a success rate of 66% in real-world human-assisted grasping scenarios. PPO has also proven versatile in dynamic manipulation of deformable objects, as demonstrated by the DexDLO framework [163], which employs end-to-end strategies for tasks such as grabbing, pulling, stretching, and bending deformable linear objects (DLOs). The framework successfully trains a policy to dynamically minimize the distance between a selected point on the DLO and a goal position. Experimental results show high success rates across various tasks, including over 60% success on DLO grabbing right, over 80% on DLO grabbing left, and nearly 100% success on DLO pulling. Moreover, frameworks like DiscoSyn [164] and SURRL [165] have incorporated PPO into multi-task learning setups, leveraging low-dimensional latent spaces and structural priors. These methods significantly reduced task-specific data requirements and facilitated effective transfer across diverse tasks. Policy-based methods are advantageous because they directly optimize the policy, making them effective for high-dimensional action spaces and continuous control tasks. However, they often require large amounts of interaction data, and the training process can be slow, with a risk of getting stuck in local optima in complex environments.

Value-based methods are a class of algorithms that focus on learning the value function, which estimates the expected return or cumulative reward for each state–action pair. The policy, which dictates the agent's actions, is then derived from these value estimates, often using a greedy strategy or an epsilon-greedy strategy that balances exploration and exploitation. Relative Entropy Q-Learning (REQ), as introduced in [64], is a model-free RL method designed for complex robotic tasks like bimanual LEGO stacking. It combines policy iteration with importance sampling to leverage both off-policy data from suboptimal experts and on-policy exploration. REQ's key feature is its temperature parameter, which balances exploitation of the learned Q-function with regularization, crucial for learning from partial solutions provided by suboptimal experts. This approach is adept at handling high-dimensional state–action spaces and sparse rewards, common in

Table 4
Reinforcement learning for anthropomorphic five-fingered hands.

Category	Ref.	Year	Algorithm	Value/Policy	Robot hand	Task	Experimental validation
Model-based	[155]	2016	RL with linear-Gaussian controllers	Policy-based	Adroit	Hand positioning and object manipulation	Real & Sim
	[61]	2019	PDDM	Policy-based	Shadow	Valve rotation, in-hand reorientation, handwriting, manipulating baoding balls	Real & Sim
Model-free	[156]	2021	MPC-SAC	Actor-critic	Shadow	Valve rotation, manipulate cube	Sim
	[158]	2018	TRPO	Policy-based	MPL	Grasp	Sim
	[167]	2019	D4PG	Actor-critic	Shadow	Gentle contact-rich manipulation	Real & Sim
	[168]	2019	DDPG+HER	Actor-critic	Shadow	Solve a rubik's cube	Sim
	[46]	2020	PPO	Policy-based	Shadow	In-hand manipulation	Real & Sim
	[160]	2020	Soft PPO	Policy-based	Seed	Grasp	Sim
	[169]	2020	DDPG++	Actor-critic	Adroit	Open door	Sim
	[170]	2021	DDPG+HER	Actor-critic	Shadow	Multi-task	Sim
	[164]	2021	DisoSyn	Policy-based	Shadow	Multi-task	Sim
	[165]	2022	PPO	Policy-based	Adroit	Multi-task	Sim
	[54]	2023	PPO	Policy-based	SCHUNK SIH	Pregrasp	Sim
	[52]	2023	PPO	Actor-critic	Faive hand	Ball rolling	Real
	[161]	2023	PPO	Policy-based	Shadow	Pregrasp	Real & Sim
	[162]	2024	PPO	Policy-based	Shadow	Grasp	Real & Sim
	[163]	2024	PPO	Policy-based	Shadow	Manipulate deformable linear Object	Sim
[166]	2024	DQN	Value-based	Shadow	Throw ball, guess fist	Sim	

dexterous manipulation. REQ also uses KL divergence to constrain policy updates, ensuring a balance between aggressive improvement and conservative behavior retention, which is essential for intricate environmental interactions in tasks like LEGO stacking and insertion. Deep Q-Networks (DQN) are a groundbreaking approach in RL that combines the Q-learning algorithm with deep neural networks to approximate the action-value function, which estimates the expected return for each state-action pair. DQN's key innovations include the use of experience replay and target networks. Experience replay involves storing experiences in a buffer and sampling them randomly for training, which helps to break the correlation between consecutive experiences and stabilizes the learning process. Target networks are used to provide a stable target for the network to learn from, reducing the variance in the learning process and improving convergence. In the context of multi-finger dexterous hands, the article [166] introduces a novel application of DQN for interactive learning. This study leverages DQN's capabilities to address the intricate challenges of coordinating finger movements and controlling interactions with the environment in tasks requiring precision and rapid response, such as ball throwing and playing rock-paper-scissors. The approach uses a hierarchical deep reinforcement learning framework that simplifies complex tasks into more manageable sub-tasks, enhancing motion planning and control. The integration of experience replay and target networks within the DQN framework contributes to the stability and efficiency of the learning process. The study demonstrates the effectiveness of this DQN-based approach, achieving a high normalized reward of 0.83 and an 84% success rate in simulated tasks. These results highlight the potential of DQN in advancing the capabilities of multi-finger dexterous hands for interactive applications and human-robot interaction. Value-based methods excel in learning state-action value functions (Q-values), enabling efficient policy updates and generally offering more stable training compared to policy-based methods. However, they struggle with continuous action spaces and may experience low learning efficiency in high-dimensional environments due to challenges in estimating the value function.

Actor-Critic methods combine the strengths of both policy-based and value-based approaches, where the actor learns the policy and the critic evaluates it by estimating the value function. One prominent algorithm in this category is Deep Deterministic Policy Gradient (DDPG), which is tailored for continuous action spaces, combining elements of Deterministic Policy Gradients (DPG) and Q-learning. Unlike traditional policy-gradient

methods, DDPG employs a deterministic policy, where the actor directly maps states to actions, eliminating the randomness seen in stochastic methods like REINFORCE. The critic, on the other hand, estimates the value of state-action pairs using a Q-function. In [168], a novel approach for solving the Rubik's Cube using a multi-fingered robotic hand, the Shadow Hand, was introduced. This method employed a hierarchical reinforcement learning framework, splitting the task into two components: planning and manipulation. A model-based solver was used for the planning phase to compute an optimal move sequence, while the manipulation phase relied on a model-free cube operator, guided by the DDPG algorithm. This separation of tasks allowed the system to efficiently handle complex operations and ensure precise control. The integration of Hindsight Experience Replay (HER) played a crucial role in enhancing learning efficiency. HER enables the agent to learn from failed attempts by adjusting the goals to states visited during those failures, effectively turning failures into learning opportunities. This method achieved impressive results, solving 1400 randomly scrambled Rubik's Cubes with a success rate of 90.3%, highlighting the combined strengths of DDPG and HER in tackling complex manipulation tasks, especially in high-dimensional spaces with sparse rewards. Similarly, [170] applied DDPG and HER in robotic manipulation tasks, focusing on improving generalization to unseen objects. The authors used point cloud encoding to represent objects geometrically, enhancing the robot's ability to interact with novel objects in a variety of scenarios. By incorporating multi-task learning, the model was able to adapt to different objects and environments, further boosting its generalization capabilities. HER allowed the agent to modify its goals during failed episodes, enabling the system to learn from a wider range of experiences. This approach improved performance in handling diverse tasks, demonstrating the power of combining DDPG with HER for robust generalization, even in scenarios with noisy data or unfamiliar objects. Building on these advancements, [169] introduced DDPG++, a refined version of DDPG designed to improve stability and efficiency in continuous action spaces. The key innovation of DDPG++ was the incorporation of advantage estimation and multi-Q function min-value updates, which addressed the instability often seen in standard DDPG. These updates helped to stabilize training and improve sample efficiency, making DDPG++ more robust and less sensitive to hyperparameters. The enhanced stability of DDPG++ made it particularly effective for high-dimensional tasks, further demonstrating the value of optimizing DDPG for complex, real-world

robotic applications. Distributed Distributional Deterministic Policy Gradient (D4PG) is an enhancement over DDPG, addressing key challenges such as sample inefficiency and uncertainty in continuous action spaces. While DDPG uses a deterministic policy and Q-function to estimate state–action values, D4PG improves upon this by adopting distributional value learning, which models the full distribution of returns rather than just their expected values. This allows D4PG to more effectively manage uncertainty and improve stability during the learning process. In [167], D4PG is applied to a curiosity-driven approach for gentle object manipulation, where the goal is to train policies that promote safe and delicate interaction with objects. The method integrates an impact penalty and a surprise-based intrinsic reward, encouraging the agent to explore in a manner that minimizes environmental disruption. The distributed nature of D4PG enables more efficient use of data, making it particularly effective for tasks requiring precise, gentle manipulation. When compared to earlier model-free methods like DDPG, D4PG shows improved efficiency and greater resilience in the face of environmental uncertainties, demonstrating its suitability for real-world robotic applications where safety and finesse are paramount. Actor-critic methods combine policy optimization and value function estimation, making them efficient in continuous action spaces. However, their training process is more complex and can be unstable, as the updates to the policy and value function may not always align properly, leading to inconsistent learning behavior.

Model-free methods, compared to model-based approaches, offer notable advantages in certain robotic manipulation tasks. These methods do not rely on an explicit model of the environment, simplifying their application in real-world scenarios where building an accurate model can be challenging or impractical. By learning optimal policies or value functions through interaction with the environment, they can handle highly dynamic or uncertain environments where model-based predictions may fail or require constant recalibration. However, model-free methods typically require large amounts of interaction data and can experience slower convergence, especially in complex, high-dimensional environments. Despite these challenges, they provide a significant advantage in environments where real-time adaptability is crucial. Table 4 provides a summary of key studies on RL for anthropomorphic five-fingered hands, detailing the methods and technical aspects explored in this field. Given the challenges inherent in both model-based and model-free RL, combining IL with RL presents a promising approach to mitigate these issues.

4.3. Integrating imitation learning with reinforcement learning for anthropomorphic five-fingered hands

Researchers have been exploring the integration of RL and IL to leverage the strengths of both approaches. While RL excels in exploration, it often struggles with efficiency, especially in complex manipulation tasks where exploration can be costly and time-consuming. In contrast, IL can provide valuable guidance by leveraging expert demonstrations but is limited by the quality of available samples and the lack of interactive optimization. Combining RL and IL helps address these challenges by using IL for policy initialization, data augmentation, guided exploration, reward learning, and policy distillation. This fusion enhances learning efficiency and accelerates adaptation to complex tasks, with each paradigm playing a key role in optimizing the agent's learning process.

Policy initialization through IL provides RL agents with an effective starting point, reducing the need for extensive trial-and-error exploration by leveraging expert demonstrations. In [45], the Demonstration Augmented Policy Gradient (DAPG) method

combines BC with RL for dexterous manipulation tasks. BC is used to initialize the policy by mimicking expert demonstrations, providing a starting point for the Natural Policy Gradient (NPG) algorithm. This initialization sidesteps complex reward engineering and accelerates the learning process. Experiments demonstrated that DAPG outperformed NPG with shaped rewards and DDPG with demonstrations, achieving policies with human-like behaviors and requiring significantly fewer samples for convergence. In [140], IL is employed to initialize policies for vision-based grasping tasks. Human motion capture data is used to pretrain neural networks that emulate grasping motions, providing an initial policy for Policy Improvement with Path Integrals (PI2), a model-based RL method. Experiments on an SVH hand-arm system demonstrated that the initialized policy improved the number of successful grasps, with iterative updates reducing the force-closure cost and resulting in stable grasps for novel objects. [171] introduced a similar integration of IL and RL for dexterous manipulation tasks. BC was used to train an initial policy on expert demonstrations, capturing key visual and proprioceptive features. This policy served as the basis for the Distributional Maximum a-posteriori Policy Optimization (DMPO) algorithm. Quantitative results showed that a BC-initialized policy trained on 30 demonstrations achieved a success rate of around 30%, which improved to nearly 90% when combined with residual RL. In contrast, training DMPO from scratch achieved only 59.2% success on the Pick task after 8 million frames, whereas the IL-initialized approach reached 98.8%, highlighting the substantial benefits of policy initialization. [172] further explored this paradigm for anthropomorphic robotic hands. The DAPG method was again employed, where BC pretrained policies using human demonstrations. This provided a robust starting point for PPO, which refined the manipulation skills. Experimental results showed that policies trained for six different tools achieved nearly 100% success rates, demonstrating resilience across varying tool scales. The initialization via IL significantly reduced the exploration burden and accelerated skill acquisition.

Data augmentation, by incorporating expert demonstrations into the training process, enriches the RL agent's experience. This approach helps accelerate learning and improve performance by enabling the agent to leverage high-quality demonstrations, reducing the need for extensive self-generated interactions. Rajeswaran et al. [45] introduced the Demonstration-Augmented Policy Gradient (DAPG) method, which incorporates expert demonstrations into the RL process through an augmented policy gradient term. This approach allows the model to not only learn from its own interactions but also to benefit from the expert demonstrations, effectively augmenting the policy's learning process with external knowledge. Experimental results show that DAPG outperforms traditional RL methods, such as NPG and DDPG, by improving sample efficiency and significantly reducing the training time required to achieve robust, human-like policies. This highlights the effectiveness of demonstration data as a powerful tool for enhancing RL, especially in complex, high-dimensional tasks. Radosavovic et al. [44] proposed State-Only Imitation Learning (SOIL), which uses state-only demonstrations to generate state–action pairs for RL training. This approach allows RL agents to learn from demonstrations that only provide partial information, such as states, without the need for full trajectory data or action annotations. The SOIL framework demonstrates that even limited demonstration data, when used in conjunction with RL, can significantly improve an agent's ability to generalize to new, unseen environments. Their results indicate that the integration of state-only demonstrations led to faster convergence and better overall performance, particularly in tasks where full trajectory data was unavailable. Similarly, Qin et al. [59] introduced DexMV, a framework that integrates IL with

RL to improve the sample efficiency of RL agents performing dexterous manipulation tasks. By incorporating demonstrations into the RL process, DexMV enables the model to learn faster from fewer interactions, effectively enhancing the policy's ability to handle intricate object manipulation tasks. Their experimental results show that leveraging demonstration data significantly reduced the number of required training episodes to achieve a high level of performance, outperforming baseline RL methods that relied solely on exploration.

Guided exploration techniques enhance the exploration process by focusing the agent's attention on high-value areas of the state space, prioritizing actions that are more likely to lead to successful outcomes. For instance, Gupta et al. [154] introduced Guided Policy Search (GPS), which integrates IL with RL to optimize control policies for complex robotic tasks. By leveraging expert demonstrations, GPS helps to steer the agent's exploration towards more effective trajectories, enabling it to replicate complex behaviors such as valve turning and abacus manipulation. This approach not only facilitates faster learning but also reduces the exploration burden by aligning the agent's policy with expert knowledge. Jeong et al. [64] similarly used expert demonstrations to enhance the sample efficiency of RL, although they did not specifically use GPS. Instead, their approach focused on augmenting RL with IL to guide the agent's exploration and improve its ability to handle fine-grained manipulation tasks. By incorporating demonstrations directly into the RL process, they were able to accelerate learning and avoid inefficient exploration paths. Wu et al. [173] also utilized expert demonstrations, but their method specifically employed a ranking mechanism to prioritize difficult demonstrations. By focusing on the most challenging examples, this ranking mechanism enabled more efficient exploration, helping the agent to learn faster and more effectively in highly complex environments. These works highlight the importance of expert demonstrations in guiding exploration, showcasing how expert knowledge can significantly improve sample efficiency and accelerate the learning process for complex robotic manipulation tasks.

Reward learning adapts the reward function based on expert demonstrations or feedback, enabling the RL agent to refine its understanding of desirable outcomes without relying on hand-crafted reward signals. [174] introduced a Deep Grasping Probability Network (DGPN) that predicts grasp success probabilities based on human-like hand poses and touch areas, providing a reward signal to guide RL towards natural and functional grasping behaviors. This approach achieved an average grasping success rate of 89.40%, a substantial improvement over standard RL. Building on this, [175] leveraged human hand poses extracted from Internet videos to generate an auxiliary reward, Rpose, which encourages human-like grasp poses during manipulation. The method accelerated training by 20% and significantly improved grasp stability and naturalness. [176] proposed a task-agnostic reward model derived from human preferences over robot trajectories to fine-tune RL policies, enhancing human-likeness in robotic behaviors. After four iterations of reward model training, the approach increased human-likeness by 22.3% while maintaining high task success rates, demonstrating strong generalization across tasks. Lastly, [177] combined IL-derived reference trajectories with trajectory-guided rewards to encourage pose similarity while solving tasks. The reward design not only aligned policies with human demonstrations but also improved generalization to unseen object poses through trajectory augmentation. Together, these studies highlight the evolving role of IL in shaping reward functions to guide RL, achieving more human-like and effective robotic manipulation.

Policy distillation condenses the learned behaviors from a complex policy into a simpler, more efficient one, enabling faster

deployment and improving the agent's performance in resource-constrained environments. [178] utilizes DAgger, an IL algorithm, to distill the expertise of a teacher policy trained with PPO under privileged oracle information into a student policy that operates on realistic sensory inputs. This approach achieves over 70% success rates in simulation for the teacher model and demonstrates robust generalization capabilities for the student model across unseen objects. Similarly, [179] applies policy distillation via DAgger to a multi-object grasping framework, where the teacher policy, trained with PPO in simulation, achieves a 44.13% success rate in dual-object grasping, with further validation of adaptability and tolerance in real-world tests. Moving beyond direct policy distillation, [180] introduces the Teacher-Augmented Policy Gradient (TAPG) framework, which combines IL methods such as Behavior Cloning and DAgger with RL to accelerate learning. TAPG not only initializes student policies through distillation but also leverages teacher demonstrations to augment training data, significantly improving sample efficiency and stability during training. This method enables superior performance in visual-based grasping tasks, demonstrating robust zero-shot transfer capabilities to real-world scenarios. Expanding this concept further, [55] employs a diffusion-based policy distillation mechanism to aggregate knowledge from multiple teacher policies into a single student policy, targeting functional pre-grasp manipulations. The method achieves a 72.6% success rate across over 30 object categories and over 10,000 goal poses, with notable robustness under noisy object pose observations. Collectively, these studies illustrate the versatility and effectiveness of policy distillation in enabling robust, generalizable, and high-performing dexterous manipulation policies.

The integration of IL with RL optimizes the learning process, especially for complex tasks. IL provides an effective foundation by utilizing expert demonstrations, reducing the need for extensive trial-and-error exploration and speeding up learning. It also enhances the training data through augmentation, offering the RL agent a more diverse set of experiences without extra self-interactions. Additionally, combining IL and RL facilitates guided exploration, directing the agent toward promising regions of the state space and ensuring more efficient learning. Reward learning within this framework adapts the reward function based on expert input, helping the agent refine its understanding of desired outcomes, particularly in tasks with challenging reward definitions. Policy distillation further simplifies the learned behaviors into a more efficient form, improving deployment and performance in resource-constrained environments. In summary, the combination of IL and RL leverages their strengths, resulting in a more effective and adaptable learning process. **Table 5** provides a summary of key studies that demonstrate the successful integration of IL with RL for anthropomorphic five-fingered hands, highlighting advancements and applications across various manipulation tasks.

Indeed, the five-fingered hand's structural resemblance to the human hand provides a natural foundation for replicating human-like behavior and coincides with the intuitive understanding of human-like dexterity. However, the applicability of the aforementioned learning methods extends beyond five-fingered hands; they also hold potential applicability to four-fingered and three-fingered robots for achieving human-like dexterity. This is in line with the prevailing trend of developing generalized learning algorithms, as exemplified by the paper [27,182]. It is essential to underscore that designing robot learning algorithms for five-fingered hands might entail greater complexity compared to algorithms tailored exclusively for four or three fingers. Several reasons contribute to this complexity:

Table 5
Integrating imitation learning with reinforcement learning for anthropomorphic five-fingered hands.

Category	Ref.	Year	Algorithm	Robot hand	Task	Experimental validation
Policy initialization	[45]	2017	BC & DAPG & NPG	Adroit	Relocate objects, in-hand manipulation, use tools, open doors	Sim
	[140]	2019	BC & PI2	SCHUNK S5FH	Grasp	Real
	[171]	2021	BC & DMPO	Shadow	Relocate objects, in-hand manipulation, use tools, open doors	Sim
Data augmentation	[181]	2023	BC & DAPG	Adroit	Grasp	Sim
	[172]	2023	DAPG & PPO	Adroit	Grasp	Sim
	[45]	2017	DAPG & NPG	Adroit	Relocate objects, in-hand manipulation, use tools, open doors	Sim
	[44]	2021	SOIL & DAPG & NPG	Adroit	Ham a nail, open a door, relocation, reposition a pen	Sim
Guided exploration	[59]	2022	GAIL & DAPG & SOIL & TRPO	Adroit	Grasp	Sim
	[154]	2016	Demonstrations & GPS	RBO Hand 2	Rotate valve, push abacus, grasp bottle	Real
	[64]	2020	BC & REQ	Shadow	Bimanual LEGO stacking	Sim
Reward learning	[173]	2023	BC & DAPG & TRPO	Adroit	Relocate	Sim
	[174]	2021	BC & DGPN & NPG	Adroit	Grasp	Sim
	[175]	2022	Demonstrations & PPO	Adroit	Grasp	Sim
	[176]	2023	Bradley–Terry–Luce & PPO	Shadow	Binamual manipulation	Real & Sim
	[177]	2024	DAPG & PPO	Adroit	Grasp	Sim
Policy distillation	[178]	2023	Dagger & PPO	Shadow	Grasp	Sim
	[55]	2024	DDPM & PPO	Shadow	Grasp	Sim
	[179]	2024	Dagger & PPO	Shadow	Grasp multi-object	Real & Sim
	[180]	2024	BC & Dagger & TAPG	Schunk SIH	Grasp in clutter	Real & Sim

- **Complexity of action space:** Robots equipped five-fingered hands confront a higher-dimensional action space due to the presence of multiple joints in each finger. Consequently, the potential combinations of actions increase significantly. Coping with this complexity demands more sophisticated learning techniques for efficient exploration and optimization within this expanded action space. In contrast, robots with three or four fingers may have a relatively lower-dimensional action space due to fewer fingers.
- **Diversity of task demands:** Five-fingered robotic hands excel at executing a wide range of tasks, manipulating objects of varying shapes, sizes, and stiffness levels, spanning from dexterous manipulation to robust grasp. Consequently, robot learning algorithms for five-fingered hands must prioritize enhanced generalizability and adaptability to address the spectrum of operational tasks.
- **Richness of tactile information:** Designing learning algorithms for these hands necessitates the ability to capture and process this extensive tactile information accurately. This processing capability contributes to guiding operations with higher precision.

5. Challenges and future directions

Recent advances in the design of anthropomorphic five-fingered hands have laid a robust hardware foundation for achieving dexterous manipulation, complemented by the rapid development of robot learning algorithms that promise to significantly enhance manipulation capabilities. However, bridging the gap between current achievements and the demands of complex real-world tasks remains challenging. This section delves into the key open challenges and outlines potential future research directions to address these gaps, fostering innovation in this critical field.

5.1. Technological challenges and innovations

5.1.1. Structure and mechanism

Designing the structure and drive system of a dexterous robotic hand presents a key challenge in achieving sufficient DoFs and flexibility, while ensuring reliability and durability. A primary issue is the design of joints and actuators that can adapt

to a variety of tasks, without introducing excessive complexity that may lead to high costs and reduced efficiency. Traditional mechanical systems, while reliable, can sometimes become bulky or inefficient when scaling to the required dexterity for complex manipulation tasks. To address these challenges, one potential direction is the exploration of lightweight and high-performance structural designs. Incorporating innovative materials, such as soft materials and shape-memory alloys, can enhance flexibility and allow for more adaptable configurations. These materials offer advantages in terms of flexibility, compactness, and energy efficiency. Additionally, improving drive systems through more efficient actuators, such as high-torque motors, pneumatic, or hydraulic actuators, could offer a solution to achieving human-like dexterity while maintaining system reliability and durability. The integration of under-actuated mechanisms and tendon-based actuation could further reduce mechanical complexity while still enabling versatile and precise movements.

5.1.2. Perception and feedback

Achieving high-precision perception and feedback systems is crucial for dexterous robotic manipulation, particularly when it comes to providing tactile and force feedback. The human hand's soft, flexible skin acts as a high-frequency, high-resolution sensor, offering detailed feedback about touch, pressure, and texture. This tactile feedback is essential for adjusting grip strength, adapting to object shape, and ensuring stable manipulation. However, current robotic hands typically rely on rigid components and fingertip sensors, which lack the dexterity and sensitivity found in human skin. These limitations restrict their ability to perform complex tasks that require fine manipulation or adaptive responses to varying forces. To address these limitations, advancements in materials science and 3D manufacturing are critical for integrating human-like soft materials into robotic hands. By incorporating flexible and lightweight artificial skin, robotic hands could achieve enhanced tactile perception and more accurate feedback during manipulation. These materials could enable continuous force sensing across the hand, facilitating better interaction with objects and a higher degree of precision. Moreover, the integration of multi-modal sensors—combining touch, force, and vision—could further enhance real-time feedback, enabling more adaptive and stable manipulation in diverse environments.

5.1.3. Balancing anthropomorphism and functionality

Designing anthropomorphic five-fingered hands is inspired by the natural structure of the human hand, representing the ultimate goal in robotic hand development. Structurally, this involves a comprehensive replication of human hand anatomy, including size, weight, phalangeal ratios, joint-to-bone connections, and particularly the thumb, which requires careful consideration. Functionally, achieving human-like kinematics and dynamics necessitates powerful actuators, sensitive sensors, efficient transmission mechanisms, and robust control systems. However, balancing anthropomorphism with functionality remains a key challenge. While replicating the human hand's anatomy is important for natural movement, maintaining stability and precision in complex tasks is equally crucial. To optimize this balance, the design can focus on improving the trade-off between flexibility and stability. For example, incorporating anthropomorphic joint designs can enhance flexibility without compromising grasping stability. Additionally, dynamic joint configurations can be adapted according to task-specific requirements, allowing the robotic hand to maintain both dexterity and task performance across different applications.

5.1.4. Real-time performance and computational resources

Achieving real-time performance in robotic hands is critical for ensuring responsive and adaptive manipulation, especially during complex tasks. The ability to process sensory data, control actuators, and adjust movements in real-time requires significant computational resources. However, the complexity of high-fidelity perception, actuation control, and feedback loops demands both fast processing and efficient use of resources. Current challenges include balancing the need for real-time decision-making with the computational power required for precise control and feedback. The large volumes of data generated by sensors, particularly in high-resolution force and tactile feedback, add additional strain on processing capabilities. Moreover, achieving low-latency control while maintaining system stability, especially in dynamic or unpredictable environments, remains a challenge. To address these issues, potential solutions include optimizing algorithms for faster computation, improving hardware accelerators such as GPUs and FPGAs, and utilizing efficient sensor fusion techniques to reduce computational load. Additionally, integrating edge computing could help offload intensive tasks and enable faster decision-making, improving overall system responsiveness.

5.2. Algorithmic advancements for enhanced dexterity

5.2.1. Multi-task learning

Multi-task learning draws inspiration from humans' ability to master multiple skills simultaneously, using shared representations to accelerate learning across tasks. In robotics, especially with anthropomorphic five-fingered hands, multi-task learning is key to enabling robots to handle a range of real-world tasks, such as pouring liquids, tidying spaces, opening doors, or folding clothes. These tasks vary in complexity and require adaptability to different physical conditions and environments. While basic multi-task learning implementations are seen in simpler systems—like robots grasping and pushing with parallel grippers or manipulating objects using different grasp types—there is increasing demand for multi-task learning to support more complex interactions. For instance, anthropomorphic robots could use multi-task learning to perform fine-grained manipulations, such as unscrewing a lid while maintaining a steady grip on a cup, or folding a shirt while avoiding nearby obstacles. Future advancements could involve the development of more sophisticated models, like multi-agent RL or hierarchical approaches, to better integrate task-specific knowledge across domains. These enhancements would help robots perform diverse tasks in dynamic environments with increased efficiency and adaptability.

5.2.2. Multi-modal learning

Humans integrate multiple sensory inputs—such as vision, touch, and proprioception—to perceive and interact with the world, enabling accurate and rich information gathering for tasks like shaping clay or assembling objects. In robotics, multi-modal learning is essential for replicating this capability, especially when handling deformable or non-rigid objects that cannot be manipulated by vision alone. In robotic systems, multi-modal learning integrates sensor modalities (vision, tactile feedback, and force sensing) to create a unified representation of the environment. This is crucial when interacting with soft or irregular objects, where visual cues alone are insufficient. For instance, when manipulating clay, a robot combines visual and tactile feedback to adjust force and understand the material's deformation. Future advancements could focus on developing deep learning models that process and fuse multi-sensor data in real-time, improving sensor fusion techniques. Additionally, innovations in soft sensors and tactile skins may enhance robots' tactile sensitivity, enabling more precise manipulation of delicate objects. Integrating multi-modal learning with RL could also allow robots to adapt actions based on real-time multi-sensory feedback.

5.2.3. Sim-to-real transfer

In robotics, sim-to-real transfer refers to the challenge of transferring policies or models learned in simulation environments to real-world robots, overcoming the "reality gap" that often arises due to differences in physics, sensor noise, and unmodeled real-world phenomena. Simulation environments provide a controlled and safe space for testing robotic systems and training complex behaviors without the constraints and risks of physical robots. However, the discrepancies between simulated environments and reality—such as sensor inaccuracies, unmodeled friction, and imperfect dynamics—often cause policies trained in simulations to underperform or fail when transferred to real-world scenarios. To address the reality gap, advancements in sim-to-real transfer methods are critical. Techniques such as domain randomization, where simulation parameters are randomized to cover a wide range of real-world variations, have shown promise in improving the generalization of learned behaviors. Additionally, adversarial domain adaptation methods, which aim to minimize the domain shift between simulation and reality, can be incorporated into training pipelines. Research could also explore hybrid simulation-real-world training, where robots interact with physical environments in conjunction with simulated environments to refine learned behaviors. Combining simulation with real-world feedback loops will enhance the performance and robustness of robotic systems when deployed in complex, dynamic environments.

5.2.4. Lifelong learning

Lifelong learning, also known as continuous learning, is the ability of humans to acquire, retain, and refine knowledge and skills throughout their lives, adapting to new challenges while retaining previously learned information. In robotics, lifelong learning involves designing algorithms that allow robots to continually improve their performance over time without forgetting prior experiences. This contrasts with traditional machine learning, where models are trained on static datasets and often suffer from catastrophic forgetting when exposed to new information. One promising direction is the development of algorithms that support incremental learning, allowing robots to integrate new knowledge without losing the ability to perform previously learned tasks. Techniques such as elastic weight consolidation, progressive neural networks, or memory-based approaches can help mitigate catastrophic forgetting. Additionally, lifelong learning for robots could be enhanced through meta-learning, where

robots learn how to learn more efficiently by generalizing across tasks and environments. Research could also investigate the application of transfer learning, where knowledge acquired from one task is adapted to different but related tasks, enabling robots to handle diverse scenarios with minimal retraining.

5.2.5. Lack of unified evaluation framework

The absence of a unified evaluation framework poses a significant challenge in the field of anthropomorphic robotic hands, as it hampers the ability to accurately assess and compare the performance of different systems across various tasks and environments. The current fragmented approach, with its diverse benchmarks and metrics, complicates the evaluation of generalizability and adaptability, particularly when translating performance from simulated environments to real-world conditions. To address this, the development of standardized benchmarks that can encompass a broad range of dexterous manipulation tasks is essential. Future research should focus on creating adaptable and dynamic evaluation protocols that can evolve alongside advancing technology, allowing for meaningful comparisons and guiding the development of more capable robotic hands. Integrating real-world data into these frameworks will also be crucial for reflecting the complexities of practical applications and ensuring that robotic hands can seamlessly transition from simulation to reality. Establishing a cohesive evaluation methodology is imperative for driving progress in the field, facilitating the development of robotic hands that are not only versatile but also practical for real-world deployment.

6. Conclusion

This paper comprehensively reviews human-like dexterous manipulation for anthropomorphic five-fingered hands. Key contributions of this research include:

- An initial definition of human-like dexterity and an overview of typical human-like dexterous manipulation tasks.
- In-depth exploration of anthropomorphism and the anthropomorphic five-fingered hand, covering definition, design, and evaluation.
- A survey of learning methods, with a focus on imitation learning, reinforcement learning, and their integration, enabling robots to acquire human-like dexterity.

Despite notable progress in the field, significant challenges remain in both the structural design and functional capabilities of anthropomorphic robotic hands. Structurally, the focus must be on refining the mechanics and materials to better replicate the human hand's intricate anatomy and adaptability, ensuring that the hands can withstand the demands of complex tasks without sacrificing precision or durability. Functionally, advancing the algorithms that enable these hands to learn, adapt, and perform a wide range of manipulations with human-like finesse is crucial. This includes developing more sophisticated learning models capable of handling multi-tasking, processing multi-modal sensory inputs, and effectively transferring skills from simulated environments to real-world scenarios. Furthermore, establishing a unified evaluation framework is essential to benchmark and compare the performance of different robotic hands across standardized tasks, providing a clear direction for further research and development. Through continued advancements in both structural and functional aspects, the full potential of robotic hands can be realized, bridging the gap between current capabilities and the dexterity and versatility of the human hand. This potential spans across various sectors, where anthropomorphic robotic hands can significantly enhance both precision and adaptability

in complex tasks.

In the industrial sector, anthropomorphic robotic hands offer transformative potential for tasks such as pick-and-place, assembly, and precision tool use. These robots can handle intricate assembly tasks, like fitting small, delicate components in electronics or assembling automotive parts that require precise alignment. With enhanced dexterity, force feedback, and versatility, they can perform delicate operations such as polishing, grinding, or welding, while adapting to diverse environments. Their fine motor control also enables more complex manufacturing processes, such as assembling micro-components or performing precision tasks in dynamic settings. Furthermore, these robotic hands can assist in quality control by detecting subtle product defects through tactile sensing, ensuring consistently high production standards.

In the service sector, anthropomorphic robotic hands can revolutionize human-robot interaction by performing tasks that require dexterity and human-like engagement. They can assist with household chores like organizing groceries, folding laundry, and handling fragile objects. Additionally, these robots can provide support in eldercare by helping with daily tasks, such as opening medication containers, preparing meals, and guiding physical activities. In interactive entertainment, they can engage in games, simulate musical performances, or guide creative activities such as painting or sculpting, all while responding sensitively to human cues. These robots not only enhance everyday life but also offer meaningful companionship, expanding their role in educational and recreational activities.

In the medical sector, anthropomorphic robotic hands could play a pivotal role in both surgical and rehabilitation contexts. They could assist in minimally invasive surgeries, where precision is critical, by manipulating surgical instruments with delicate accuracy in confined spaces. This would reduce patient trauma and recovery time. Additionally, robotic hands could support remote surgeries, especially in regions where expert human surgeons are unavailable, offering a lifeline to patients in underserved areas. In rehabilitation, these robots could tailor physical therapy to individual patients, adapt exercises in real-time, monitor progress, and provide valuable feedback, ultimately improving the recovery process.

In specialized sectors such as space exploration, military operations, and hazardous environments, anthropomorphic robotic hands offer substantial potential. These robots can perform precise tasks in settings where human involvement is limited due to risk, such as assisting in disaster recovery by manipulating debris or performing repairs in hazardous locations. Their ability to operate in extreme environments with fine-tuned precision opens up new possibilities for high-stakes applications. In space or underwater missions, these hands could perform repairs on fragile equipment, assemble structures in dangerous or remote locations, or collect samples without risking contamination. As these systems become more robust and adaptable, they will play an increasingly critical role in scenarios that demand both precision and autonomy, offering solutions in environments that are otherwise too dangerous for human intervention.

As these capabilities continue to advance, human-like dexterous manipulation with anthropomorphic five-fingered robotic hands has the potential to transition from laboratory research to impactful real-world applications, enhancing human efforts across diverse and challenging domains.

CRedit authorship contribution statement

Yayu Huang: Writing – original draft, Visualization, Methodology, Conceptualization. **Dongxuan Fan:** Resources, Investigation, Data curation. **Haonan Duan:** Writing – review & editing. **Dashun Yan:** Writing – review & editing. **Wen Qi:** Writing – review & editing. **Jia Sun:** Visualization, Validation. **Qian Liu:** Visualization, Validation. **Peng Wang:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (91748131, 62006229, and 61771471), in part by Young Scientists Fund of the National Natural Science Foundation of China (62303454), in part by the Strategic Priority Research Program of Chinese Academy of Science (XDB32050106), and in part by the InnoHK Project.

References

- [1] A. Baratta, A. Cimino, M.G. Gnoni, F. Longo, Human robot collaboration in industry 4.0: a literature review, *Procedia Comput. Sci.* 217 (2023) 1887–1895.
- [2] L. Yu, Y. Wang, X. Wei, C. Zeng, Towards low-carbon development: The role of industrial robots in decarbonization in Chinese cities, *J. Environ. Manag.* 330 (2023) 117216.
- [3] B. Liu, X. Yang, J. Zhang, Nonlinear effect of industrial robot applications on carbon emissions: evidence from China, *Environ. Impact Assess. Rev.* 104 (2024) 107297.
- [4] M. Zhang, G. Tian, Y. Cui, Y. Zhang, Z. Xia, Hierarchical semantic knowledge-based object search method for household robots, *IEEE Trans. Emerg. Top. Comput. Intell.* (2023).
- [5] H. Xiong, H. Fu, J. Zhang, C. Bao, Q. Zhang, Y. Huang, W. Xu, A. Garg, C. Lu, Robotube: Learning household manipulation from human videos with simulated twin environments, in: *Conference on Robot Learning*, PMLR, 2023, pp. 1–10.
- [6] A. Soni, S. Alla, S. Dodda, H. Volikatla, Advancing household robotics: Deep interactive reinforcement learning for efficient training and enhanced performance, 2024, arXiv preprint arXiv:2405.18687.
- [7] S. Poorghasem, Y. Bao, Review of robot-based automated measurement of vibration for civil engineering structures, *Measurement* 207 (2023) 112382.
- [8] B. Ma, Z. Jiang, Y. Liu, Z. Xie, Advances in space robots for on-orbit servicing: A comprehensive review, *Adv. Intell. Syst.* 5 (8) (2023) 2200397.
- [9] E. Tokgöz, M.A. Carro, Robotics applications in facial plastic surgeries, in: *Cosmetic and Reconstructive Facial Plastic Surgery: A Review of Medical and Biomedical Engineering and Science Concepts*, Springer, 2023, pp. 307–341.
- [10] S. Yang, J. Chen, A. Li, K. Deng, P. Li, S. Xu, Accuracy of autonomous robotic surgery for single-tooth implant placement: A case series, *J. Dent.* 132 (2023) 104451.
- [11] J.E. Knudsen, U. Ghaffar, R. Ma, A.J. Hung, Clinical applications of artificial intelligence in robotic surgery, *J. Robot. Surg.* 18 (1) (2024) 102.
- [12] A.M. Okamura, N. Smaby, M.R. Cutkosky, An overview of dexterous manipulation, in: *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, vol. 1, IEEE, 2000, pp. 255–262.
- [13] E. Mattar, A survey of bio-inspired robotics hands implementation: New directions in dexterous manipulation, *Robot. Auton. Syst.* 61 (5) (2013) 517–544.
- [14] E.N. Gama Melo, O.F. Aviles Sanchez, D. Amaya Hurtado, Anthropomorphic robotic hands: a review, *Ing. y Desarro.* 32 (2) (2014) 279–313.
- [15] C. Melchiorri, M. Kaneko, Robot hands, in: *Springer Handbook of Robotics*, Springer, 2016, pp. 463–480.
- [16] C. Piazza, G. Grioli, M. Catalano, A. Bicchi, A century of robotic hands, *Annu. Rev. Control., Robot., Auton. Syst.* 2 (2019) 1–32.
- [17] Y. Liu, Z. Li, H. Liu, Z. Kan, B. Xu, Bioinspired embodiment for intelligent sensing and dexterity in fine manipulation: A survey, *IEEE Trans. Ind. Inform.* 16 (7) (2020) 4308–4321.
- [18] O. Kroemer, S. Niekum, G. Konidaris, A review of robot learning for manipulation: Challenges, representations, and algorithms, *J. Mach. Learn. Res.* 22 (2021) 30–31.
- [19] J. Hua, L. Zeng, G. Li, Z. Ju, Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning, *Sensors* 21 (4) (2021) 1278.
- [20] H. Duan, P. Wang, Y. Huang, G. Xu, W. Wei, X. Shen, Robotics dexterous grasping: The methods based on point cloud and deep learning, *Front. Neurobotics* (2021).
- [21] A.R. Sobinov, S.J. Bensmaïa, The neural mechanisms of manual dexterity, *Nature Rev. Neurosci.* 22 (12) (2021) 741–757.
- [22] C. Yu, P. Wang, Dexterous manipulation for multi-fingered robotic hands with reinforcement learning: a review, *Front. Neurobotics* 16 (2022) 861825.
- [23] Z. Xia, Z. Deng, B. Fang, Y. Yang, F. Sun, A review on sensory perception for dexterous robotic manipulation, *Int. J. Adv. Robot. Syst.* 19 (2) (2022) <http://dx.doi.org/10.1177/17298806221095974>.
- [24] S. Kadalagere Sampath, N. Wang, H. Wu, C. Yang, Review on human-like robot manipulation using dexterous hands, *Cogn. Comput. Syst.* 5 (1) (2023) 14–29.
- [25] J.R. Napier, The prehensile movements of the human hand, *J. Bone Jt. Surg. Br. Vol.* 38 (4) (1956) 902–913.
- [26] H. Yuan, D. Li, J. Wu, Efficient learning of grasp selection for five-finger dexterous hand, in: *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, CYBER, IEEE, 2017*, pp. 1101–1106.
- [27] S. Brahmabhatt, A. Handa, J. Hays, D. Fox, Contactgrasp: Functional multi-finger grasp synthesis from contact, in: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2019*, pp. 2386–2393.
- [28] Y.-R. Liu, M.-B. Huang, H.-P. Huang, Automated grasp planning and path planning for a robot hand-arm system, in: *2019 IEEE/SICE International Symposium on System Integration, SII, IEEE, 2019*, pp. 92–97.
- [29] P. Song, Z. Fu, L. Liu, Grasp planning via hand-object geometric fitting, *Vis. Comput.* 34 (2) (2018) 257–270.
- [30] H. Jiang, S. Liu, J. Wang, X. Wang, Hand-object contact consistency reasoning for human grasps generation, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021*, pp. 11107–11116.
- [31] P. Grady, C. Tang, C.D. Twigg, M. Vo, S. Brahmabhatt, C.C. Kemp, ContactOpt: Optimizing contact to improve grasps, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021*, pp. 1471–1481.
- [32] P. Mandikal, K. Grauman, Learning dexterous grasping with object-centric visual affordances, in: *2021 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2021*, pp. 6169–6176.
- [33] W. Wei, P. Wang, S. Wang, Y. Luo, W. Li, D. Li, Y. Huang, H. Duan, Learning human-like functional grasping for multi-finger hands from few demonstrations, *IEEE Trans. Robot.* (2024).
- [34] Z. Deng, G. Gao, S. Frintrop, F. Sun, C. Zhang, J. Zhang, Attention based visual analysis for fast grasp planning with a multi-fingered robotic hand, *Front. Neurobotics* 13 (2019) 60.
- [35] F. Ficuciello, Hand-arm autonomous grasping: Synergistic motions to enhance the learning process, *Intell. Serv. Robot.* 12 (1) (2019) 17–25.
- [36] T. Osa, J. Peters, G. Neumann, Hierarchical reinforcement learning of multiple grasping strategies with human instructions, *Adv. Robot.* 32 (18) (2018) 955–968.
- [37] K. Higashi, K. Koyama, R. Ozawa, K. Nagata, W. Wan, K. Harada, Functionally divided manipulation synergy for controlling multi-fingered hands, in: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2020*, pp. 9190–9197.
- [38] R. Chauhan, B. Sebastian, P. Ben-Tzvi, Grasp prediction toward naturalistic exoskeleton glove control, *IEEE Trans. Hum.-Mach. Syst.* 50 (1) (2019) 22–31.
- [39] B. Rao, H. Li, K. Krishnan, E. Boldsai Khan, H. He, Knowledge-augmented dexterous grasping with incomplete sensing, 2020, arXiv preprint arXiv:2011.08361.
- [40] Z. Deng, B. Fang, B. He, J. Zhang, An adaptive planning framework for dexterous robotic grasping with grasp type detection, *Robot. Auton. Syst.* 140 (2021) 103727.
- [41] Z. Deng, Y. Jonetzko, L. Zhang, J. Zhang, Grasping force control of multi-fingered robotic hands through tactile sensing for object stabilization, *Sensors (Switzerland)* 20 (2020) <http://dx.doi.org/10.3390/s20041050>.
- [42] F. Veiga, B.B. Edin, J. Peters, In-hand object stabilization by independent finger control, 2018, arXiv preprint arXiv:1806.05031.
- [43] D. Jain, A. Li, S. Singhal, A. Rajeswaran, V. Kumar, E. Todorov, Learning deep visuomotor policies for dexterous hand manipulation, in: *2019 International Conference on Robotics and Automation, ICRA, IEEE, 2019*, pp. 3636–3643.
- [44] I. Radosavovic, X. Wang, L. Pinto, J. Malik, State-only imitation learning for dexterous manipulation, in: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2021*, pp. 7865–7871.
- [45] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, S. Levine, Learning complex dexterous manipulation with deep reinforcement learning and demonstrations, 2017, arXiv preprint arXiv:1709.10087.
- [46] O.A.M. Andrychowicz, B. Baker, M. Chociej, R. Józefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, W. Zaremba, Learning dexterous in-hand manipulation, *Int. J. Robot. Res.* 39 (2020) 3–20, <http://dx.doi.org/10.1177/0278364919887447>.

- [47] Z. Deng, J.W. Zhang, Learning synergies based in-hand manipulation with reward shaping, *CAAI Trans. Intell. Technol.* 5 (2020) 141–149, <http://dx.doi.org/10.1049/trit.2019.0094>.
- [48] A. Nagabandi, K. Konolige, S. Levine, V. Kumar, Deep dynamics models for learning dexterous manipulation, in: *Conference on Robot Learning*, PMLR, 2020, pp. 1101–1112.
- [49] Y. Xue, L. Tang, Y.-B. Jia, Dynamic finger gaits via pivoting and adapting contact forces, in: *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2023*, pp. 8784–8791, <http://dx.doi.org/10.1109/IROS55552.2023.10342156>.
- [50] G. Khandate, M. Haas-Heger, M. Ciocarlie, On the feasibility of learning finger-gaiting in-hand manipulation with intrinsic sensing, in: *2022 International Conference on Robotics and Automation, ICRA, 2022*, pp. 2752–2758, <http://dx.doi.org/10.1109/ICRA46639.2022.9812212>.
- [51] A.S. Morgan, K. Hang, B. Wen, K. Bekris, A.M. Dollar, Complex in-hand manipulation via compliance-enabled finger gaiting and multi-modal planning, *IEEE Robot. Autom. Lett.* 7 (2) (2022) 4821–4828, <http://dx.doi.org/10.1109/LRA.2022.3145961>.
- [52] Y. Toshimitsu, B. Forrai, B.G. Cangan, U. Steger, M. Knecht, S. Weirich, R.K. Katzschmann, Getting the ball rolling: Learning a dexterous policy for a biomimetic tendon-driven hand with rolling contact joints, in: *2023 IEEE-RAS 22nd International Conference on Humanoid Robots, Humanoids, 2023*, pp. 1–7, <http://dx.doi.org/10.1109/Humanoids57100.2023.10375231>.
- [53] J. Shi, J.Z. Woodruff, P.B. Umbanhowar, K.M. Lynch, Dynamic in-hand sliding manipulation, *IEEE Trans. Robot.* 33 (4) (2017) 778–795.
- [54] D. Pavlichenko, S. Behnke, Deep reinforcement learning of dexterous pre-grasp manipulation for human-like functional categorical grasping, in: *2023 IEEE 19th International Conference on Automation Science and Engineering, CASE, IEEE, 2023*, pp. 1–8.
- [55] T. Wu, Y. Gan, M. Wu, J. Cheng, Y. Yang, Y. Zhu, H. Dong, Unidexfpm: Universal dexterous functional pre-grasp manipulation via diffusion policy, 2024, arXiv preprint [arXiv:2403.12421](https://arxiv.org/abs/2403.12421).
- [56] S. Chen, A. Wu, C.K. Liu, Synthesizing dexterous nonprehensile pregrasp for ungraspable objects, in: *ACM SIGGRAPH 2023 Conference Proceedings, 2023*, pp. 1–10.
- [57] M. Qin, J. Brawer, B. Scassellati, Rapidly learning generalizable and robot-agnostic tool-use skills for a wide range of tasks, *Front. Robot. AI* 8 (2021) <http://dx.doi.org/10.3389/frobt.2021.726463>, URL <https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2021.726463>.
- [58] Y. Zhou, P. Zhou, S. Wang, Y. She, In-hand singulation and scooping manipulation with a 5 DOF tactile gripper, 2024, URL <https://arxiv.org/abs/2408.00610>. arXiv:2408.00610.
- [59] Y. Qin, Y.-H. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, X. Wang, Dexmv: Imitation learning for dexterous manipulation from human videos, in: *European Conference on Computer Vision, Springer, 2022*, pp. 570–587.
- [60] S. Katyara, F. Ficuciello, D.G. Caldwell, B. Siciliano, F. Chen, Leveraging kernelized synergies on shared subspace for precision grasping and dexterous manipulation, *IEEE Trans. Cogn. Dev. Syst.* (2021) <http://dx.doi.org/10.1109/TCDS.2021.3110406>, 1–1.
- [61] A. Nagabandi, K. Konolige, S. Levine, V. Kumar, Deep dynamics models for learning dexterous manipulation, 2019, pp. 1–12, arXiv.
- [62] L. Tang, Y.-B. Jia, Y. Xue, Robotic manipulation of hand tools: The case of screwdriving, in: *2024 IEEE International Conference on Robotics and Automation, ICRA, 2024*, pp. 13883–13890, <http://dx.doi.org/10.1109/ICRA57147.2024.10610831>.
- [63] F. Krebs, T. Asfour, A bimanual manipulation taxonomy, *IEEE Robot. Autom. Lett.* 7 (4) (2022) 11031–11038.
- [64] R. Jeong, J.T. Springenberg, J. Kay, D. Zheng, Y. Zhou, A. Galashov, N. Heess, F. Nori, Learning dexterous manipulation from suboptimal experts, 2020, arXiv preprint [arXiv:2010.08587](https://arxiv.org/abs/2010.08587).
- [65] M. Laghi, L. Raiano, F. Amadio, F. Rollo, A. Zunino, A. Ajoudani, A target-guided telemanipulation architecture for assisted grasping, *IEEE Robot. Autom. Lett.* 7 (4) (2022) 8759–8766.
- [66] K. Shaw, Y. Li, J. Yang, M.K. Srirama, R. Liu, H. Xiong, R. Mendonca, D. Pathak, Bimanual dexterity for complex tasks, 2024, arXiv preprint [arXiv:2411.13677](https://arxiv.org/abs/2411.13677).
- [67] Z. Jiang, Y. Xie, K. Lin, Z. Xu, W. Wan, A. Mandlekar, L. Fan, Y. Zhu, DexMimicGen: Automated data generation for bimanual dexterous manipulation via imitation learning, 2024, arXiv preprint [arXiv:2410.24185](https://arxiv.org/abs/2410.24185).
- [68] Y. Chen, T. Wu, S. Wang, X. Feng, J. Jiang, Z. Lu, S. McAleer, H. Dong, S.-C. Zhu, Y. Yang, Towards human-level bimanual dexterous manipulation with reinforcement learning, *Adv. Neural Inf. Process. Syst.* 35 (2022) 5150–5163.
- [69] Y. Chen, Y. Geng, F. Zhong, J. Ji, J. Jiang, Z. Lu, H. Dong, Y. Yang, Bidexhands: Towards human-level bimanual dexterous manipulation, *IEEE Trans. Pattern Anal. Mach. Intell.* (2023).
- [70] J. Grannen, Y. Wu, B. Vu, D. Sadigh, Stabilize to act: Learning to coordinate for bimanual manipulation, in: *Conference on Robot Learning, PMLR, 2023*, pp. 563–576.
- [71] Y. Huang, Z. Wang, X. Shen, Q. Liu, P. Wang, Human-like dexterous manipulation for the anthropomorphic hand-arm robotic system via teleoperation, in: *International Conference on Intelligent Robotics and Applications, Springer, 2023*, pp. 309–321.
- [72] S. Li, X. Ma, H. Liang, M. Görner, P. Ruppel, B. Fang, F. Sun, J. Zhang, Vision-based teleoperation of shadow dexterous hand using end-to-end deep neural network, 2018, arXiv preprint [arXiv:1809.06268](https://arxiv.org/abs/1809.06268). URL <http://arxiv.org/abs/1809.06268>.
- [73] Q. Gao, Z. Deng, Z. Ju, T. Zhang, Dual-hand motion capture by using biological inspiration for bionic bimanual robot teleoperation, *Cyborg Bionic Syst.* 4 (2023) 0052, <http://dx.doi.org/10.34133/cbsystems.0052>, URL <https://spj.science.org/doi/abs/10.34133/cbsystems.0052>. arXiv:https://spj.science.org/doi/pdf/10.34133/cbsystems.0052.
- [74] X. Cheng, J. Li, S. Yang, G. Yang, X. Wang, Open-television: Teleoperation with immersive active visual feedback, 2024, arXiv preprint [arXiv:2407.01512](https://arxiv.org/abs/2407.01512).
- [75] R. Ding, Y. Qin, J. Zhu, C. Jia, S. Yang, R. Yang, X. Qi, X. Wang, Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning, 2024, arXiv preprint [arXiv:2407.03162](https://arxiv.org/abs/2407.03162).
- [76] L. Biagiotti, F. Lotti, C. Melchiorri, G. Vassura, How Far Is the Human Hand? A Review on Anthropomorphic Robotic End-Effectors, Tech. Rep., University of Bologna, 2004.
- [77] S. Puhlmann, J. Harris, O. Brock, RBO hand 3: A platform for soft dexterous manipulation, *IEEE Trans. Robot.* 38 (6) (2022) 3434–3449.
- [78] M.G. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, A. Bicchi, Adaptive synergies for the design and control of the Pisa/IIT SoftHand, *Int. J. Robot. Res.* 33 (5) (2014) 768–782.
- [79] J.T. Belter, M.T. Leddy, K.D. Gemmell, A.M. Dollar, Comparative clinical evaluation of the yale multigrasp hand, in: *2016 6th IEEE International Conference on Biomedical Robotics and Biomechanics, BioRob, IEEE, 2016*, pp. 528–535.
- [80] S. Takamuku, A. Fukuda, K. Hosoda, Repetitive grasping with anthropomorphic skin-covered hand enables robust haptic recognition, in: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2008*, pp. 3212–3217.
- [81] A.M. Dollar, R.D. Howe, The highly adaptive SDM hand: Design and performance evaluation, *Int. J. Robot. Res.* 29 (5) (2010) 585–597.
- [82] C. Melchiorri, G. Palli, G. Berselli, G. Vassura, Development of the ub hand iv: Overview of design solutions and enabling technologies, *IEEE Robot. Autom. Mag.* 20 (3) (2013) 72–81.
- [83] O. Shorthose, A. Albin, L. He, P. Maiolino, Design of a 3D-printed soft robotic hand with integrated distributed tactile sensing, *IEEE Robot. Autom. Lett.* 7 (2) (2022) 3945–3952.
- [84] T. Feix, J. Romero, H.-B. Schmedtmayer, A.M. Dollar, D. Kragic, The grasp taxonomy of human grasp types, *IEEE Trans. Hum.-Mach. Syst.* 46 (1) (2015) 66–77.
- [85] H. Liu, K. Wu, P. Meusel, N. Seitz, G. Hirzinger, M. Jin, Y. Liu, S. Fan, T. Lan, Z. Chen, Multisensory five-finger dexterous hand: The DLR/HIT hand II, in: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2008*, pp. 3692–3697.
- [86] L.B. Bridgwater, C.A. Ihrke, M.A. Diftler, M.E. Abdallah, N.A. Radford, J.M. Rogers, S. Yayathi, R.S. Askew, D.M. Linn, The robonaut 2 hand - designed to do work with tools, in: *2012 IEEE International Conference on Robotics and Automation, 2012*, pp. 3425–3430, <http://dx.doi.org/10.1109/ICRA.2012.6224772>.
- [87] SCHUNK 5-finger hand, 2012, https://schunk.com/de_en/gripping-systems/highlights/svh/.
- [88] U. Kim, D. Jung, H. Jeong, J. Park, H.-M. Jung, J. Cheong, H.R. Choi, H. Do, C. Park, Integrated linkage-driven dexterous anthropomorphic robotic hand, *Nat. Commun.* 12 (1) (2021) 1–13.
- [89] Shadow dexterous hand e1 series, 2013, <http://www.shadowrobot.com/wpcontent/uploads/shadowdexteroushandtechnicalspecificationE120130101.pdf>.
- [90] H. Yang, G. Wei, L. Ren, Z. Qian, K. Wang, H. Xiu, W. Liang, A low-cost linkage-spring-tendon-integrated compliant anthropomorphic robotic hand: MCR-hand III, *Mech. Mach. Theory* 158 (2021) 104210.
- [91] W. Ryu, Y. Choi, Y.J. Choi, Y.G. Lee, S. Lee, Development of an anthropomorphic prosthetic hand with underactuated mechanism, *Appl. Sci.* 10 (12) (2020) 4384.
- [92] L. Biagiotti, Advanced robotic hands: Design and control aspects, *Dep. Electron., Comput. Sci. Syst.* (2002).
- [93] M.V. Liarokapis, P.K. Artemiadis, K.J. Kyriakopoulos, Quantifying anthropomorphism of robot hands, in: *2013 IEEE International Conference on Robotics and Automation, IEEE, 2013*, pp. 2041–2046.
- [94] Y. Liu, D. Yang, L. Jiang, H. Liu, A synthetic framework for evaluating the anthropomorphic characteristics of prosthetic hands, in: *2015 IEEE International Conference on Advanced Intelligent Mechatronics, AIM, IEEE, 2015*, pp. 877–884.

- [95] J.T. Belter, A.M. Dollar, Performance characteristics of anthropomorphic prosthetic hands, in: 2011 IEEE International Conference on Rehabilitation Robotics, IEEE, 2011, pp. 1–7.
- [96] J. Zhou, Y. Chen, D.C.F. Li, Y. Gao, Y. Li, S.S. Cheng, F. Chen, Y. Liu, 50 benchmarks for anthropomorphic hand function-based dexterity classification and kinematics-based hand design, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2020, pp. 9159–9165, <http://dx.doi.org/10.1109/IROS45743.2020.9340982>.
- [97] I. Llop-Harillo, A. Pérez-González, J. Starke, T. Asfour, The anthropomorphic hand assessment protocol (AHAP), *Robot. Auton. Syst.* 121 (2019) 103259.
- [98] I. Llop-Harillo, J.L. Iserte, A. Pérez-González, Benchmarking anthropomorphic hands through grasping simulations, *J. Comput. Des. Eng.* 9 (2) (2022) 330–342.
- [99] T. Feix, J. Romero, C.H. Ek, H.-B. Schmiedmayer, D. Kragic, A metric for comparing the anthropomorphic motion capability of artificial hands, *IEEE Trans. Robot.* 29 (1) (2012) 82–93.
- [100] M.Q. Mohammed, L.C. Kwek, S.C. Chua, A. Al-Dhaqm, S. Nahavandi, T.A.E. Eisa, M.F. Miskon, M.N. Al-Mhiqani, A. Ali, M. Abaker, et al., Review of learning-based robotic manipulation in cluttered environments, *Sensors* 22 (20) (2022) 7938.
- [101] V. Caggiano, S. Dasari, V. Kumar, Myodex: a generalizable prior for dexterous manipulation, in: International Conference on Machine Learning, PMLR, 2023, pp. 3327–3346.
- [102] W. Li, W. Wei, P. Wang, Continual learning for anthropomorphic hand grasping, *IEEE Trans. Cogn. Dev. Syst.* 16 (2) (2023) 559–569.
- [103] A. Sahbani, S. El-Khoury, P. Bidaud, An overview of 3D object grasp synthesis algorithms, *Robot. Auton. Syst.* 60 (3) (2012) 326–336.
- [104] A.T. Miller, P.K. Allen, Graspit! a versatile simulator for robotic grasping, *IEEE Robot. Autom. Mag.* 11 (4) (2004) 110–122.
- [105] M. Malvezzi, G. Gioioso, G. Salvietti, D. Prattichizzo, A. Bicchi, SynGrasp: A MATLAB toolbox for grasp analysis of human and robotic hands, in: 2013 IEEE International Conference on Robotics and Automation, 2013, pp. 1088–1093, <http://dx.doi.org/10.1109/ICRA.2013.6630708>.
- [106] B. Fang, S. Jia, D. Guo, M. Xu, S. Wen, F. Sun, Survey of imitation learning for robotic manipulation, *Int. J. Intell. Robot. Appl.* 3 (4) (2019) 362–369.
- [107] C. Gabellieri, M. Garabini, F. Angelini, V. Arapi, A. Palleschi, M.G. Catalano, G. Grioli, L. Pallottino, A. Bicchi, M. Bianchi, Grasp it like a pro: Grasp of unknown objects with robotic hands based on skilled human expertise, *IEEE Robot. Autom. Lett.* 5 (2020) 2808–2815, <http://dx.doi.org/10.1109/LRA.2020.2974391>.
- [108] F. Song, Z. Zhao, W. Ge, W. Shang, S. Cong, Learning optimal grasping posture of multi-fingered dexterous hands for unknown objects, in: 2018 IEEE International Conference on Robotics and Biomimetics, ROBIO, IEEE, 2018, pp. 2310–2315.
- [109] V. Arapi, Y. Zhang, G. Averta, M.G. Catalano, D. Rus, C. Della Santina, M. Bianchi, To grasp or not to grasp: an end-to-end deep-learning approach for predicting grasping failures in soft hands, in: 2020 3rd IEEE International Conference on Soft Robotics (RoboSoft), IEEE, 2020, pp. 653–660.
- [110] V. Kumar, A. Gupta, E. Todorov, S. Levine, Learning dexterous manipulation policies from experience and imitation, 2016, arXiv preprint [arXiv:1611.05095](https://arxiv.org/abs/1611.05095).
- [111] H. Liang, L. Cong, N. Hendrich, S. Li, F. Sun, J. Zhang, Multifingered grasping based on multimodal reinforcement learning, *IEEE Robot. Autom. Lett.* (2021).
- [112] S. Brahmabhatt, C. Tang, C.D. Twigg, C.C. Kemp, J. Hays, ContactPose: A dataset of grasps with object contact and hand pose, in: Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16, Springer, 2020, pp. 361–378.
- [113] O. Taheri, N. Ghorbani, M.J. Black, D. Tzionas, GRAB: A dataset of whole-body human grasping of objects, in: Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16, Springer, 2020, pp. 581–600.
- [114] F. Krebs, A. Meixner, I. Patzer, T. Asfour, The kit bimanual manipulation dataset, in: 2020 IEEE-RAS 20th International Conference on Humanoid Robots, Humanoids, IEEE, 2021, pp. 499–506.
- [115] Z. Fan, O. Taheri, D. Tzionas, M. Kocabas, M. Kaufmann, M.J. Black, O. Hilliges, ARCTIC: A dataset for dexterous bimanual hand-object manipulation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 12943–12954.
- [116] Y. Liu, H. Yang, X. Si, L. Liu, Z. Li, Y. Zhang, Y. Liu, L. Yi, Taco: Benchmarking generalizable bimanual tool-action-object understanding, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 21740–21751.
- [117] Y.-W. Chao, W. Yang, Y. Xiang, P. Molchanov, A. Handa, J. Tremblay, Y.S. Narang, K. Van Wyk, U. Iqbal, S. Birchfield, et al., DexYCB: A benchmark for capturing hand grasping of objects, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 9044–9053.
- [118] J. Jian, X. Liu, M. Li, R. Hu, J. Liu, Affordpose: A large-scale dataset of hand-object interactions with affordance-driven hand pose, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 14713–14724.
- [119] S. Hampali, M. Rad, M. Oberweger, V. Lepetit, Honnotate: A method for 3d annotation of hand and object poses, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3196–3206.
- [120] T. Kwon, B. Tekin, J. Stühmer, F. Bogo, M. Pollefeys, H2O: Two hands manipulating objects for first person interaction recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10138–10148.
- [121] Y. Li, W. Wei, D. Li, P. Wang, W. Li, J. Zhong, HGC-Net: Deep anthropomorphic hand grasping in clutter, *IEEE Int. Conf. Robot. Autom. (ICRA)* (2022).
- [122] W. Wei, D. Li, P. Wang, Y. Li, W. Li, Y. Luo, J. Zhong, DVGG: Deep variational grasp generation for dextrous manipulation, *IEEE Robot. Autom. Lett.* (2022).
- [123] Y. Hasson, G. Varol, D. Tzionas, I. Kalevatykh, M.J. Black, I. Laptev, C. Schmid, Learning joint reconstruction of hands and manipulated objects, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 11807–11816.
- [124] M. Liu, Z. Pan, K. Xu, K. Ganguly, D. Manocha, Deep differentiable grasp planner for high-dof grippers, 2020, arXiv preprint [arXiv:2002.01530](https://arxiv.org/abs/2002.01530).
- [125] L.F.C. Murrilo, N. Khargonkar, B. Prabhakaran, Y. Xiang, MultiGripper-Grasp: A dataset for robotic grasping from parallel jaw grippers to dexterous hands, 2024, arXiv preprint [arXiv:2403.09841](https://arxiv.org/abs/2403.09841).
- [126] L. Zhang, K. Bai, G. Huang, Z. Bing, Z. Chen, A. Knoll, J. Zhang, Multi-fingered robotic hand grasping in cluttered environments through hand-object contact semantic mapping, 2024, arXiv preprint [arXiv:2404.08844](https://arxiv.org/abs/2404.08844).
- [127] Y. Shao, C. Xiao, Bimanual grasp synthesis for dexterous robot hands, *IEEE Robot. Autom. Lett.* (2024).
- [128] B. Huang, S. El-Khoury, M. Li, J.J. Bryson, A. Billard, Learning a real time grasping strategy, in: Proceedings - IEEE International Conference on Robotics and Automation, 2013, pp. 593–600, <http://dx.doi.org/10.1109/ICRA.2013.6630634>.
- [129] T. Liu, Z. Liu, Z. Jiao, Y. Zhu, S.-C. Zhu, Synthesizing diverse and physically stable grasps with arbitrary hand structures using differentiable force closure estimator, *IEEE Robot. Autom. Lett.* 7 (1) (2021) 470–477.
- [130] R. Wang, J. Zhang, J. Chen, Y. Xu, P. Li, T. Liu, H. Wang, Dexgraspnet: A large-scale robotic dexterous grasp dataset for general objects based on simulation, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2023, pp. 11359–11366.
- [131] D. Turpin, T. Zhong, S. Zhang, G. Zhu, E. Heiden, M. Macklin, S. Tsogkas, S. Dickinson, A. Garg, Fast-grasp'D: Dexterous multi-finger grasp generation through differentiable simulation, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2023, pp. 8082–8089.
- [132] L. Yang, K. Li, X. Zhan, F. Wu, A. Xu, L. Liu, C. Lu, Oakink: A large-scale knowledge repository for understanding hand-object interaction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 20953–20962.
- [133] Y. Liu, Y. Liu, C. Jiang, K. Lyu, W. Wan, H. Shen, B. Liang, Z. Fu, H. Wang, L. Yi, Hoi4d: A 4d egocentric dataset for category-level human-object interaction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 21013–21022.
- [134] W. Xie, Z. Yu, Z. Zhao, B. Zuo, Y. Wang, Hmdo: Markerless multi-view hand manipulation capture with deformable objects, *Graph. Models* 127 (2023) 101178.
- [135] P. Li, T. Liu, Y. Li, Y. Geng, Y. Zhu, Y. Yang, S. Huang, GenDexGrasp: Generalizable dexterous grasping, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, 2023, pp. 8068–8074, <http://dx.doi.org/10.1109/ICRA48891.2023.10160667>.
- [136] X. Zhan, L. Yang, Y. Zhao, K. Mao, H. Xu, Z. Lin, K. Li, C. Lu, OAKINK2: A dataset of bimanual hands-object manipulation in complex task completion, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2024, pp. 445–456.
- [137] D. Antotsiou, G. Garcia-Hernando, T.-K. Kim, Task-oriented hand motion retargeting for dexterous manipulation imitation, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018.
- [138] S. Ottenhaus, D. Renninghoff, R. Grimm, F. Ferreira, T. Asfour, Visuo-haptic grasping of unknown objects based on Gaussian process implicit surfaces and deep learning, in: 2019 IEEE-RAS 19th International Conference on Humanoid Robots, Humanoids, IEEE, 2019, pp. 402–409.
- [139] F. Ficuciello, D. Zaccara, B. Siciliano, Learning grasps in a synergy-based framework, in: International Symposium on Experimental Robotics, Springer, 2016, pp. 125–135.
- [140] F. Ficuciello, A. Migliozi, G. Laudante, P. Falco, B. Siciliano, Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework, *Science Robotics* 4 (26) (2019) ea04900.

- [141] B.D. Argall, S. Chernova, M. Veloso, B. Browning, A survey of robot learning from demonstration, *Robot. Auton. Syst.* 57 (5) (2009) 469–483.
- [142] H. Li, J. Tan, H. He, MagicHand: Context-aware dexterous grasping using an anthropomorphic robotic hand, in: 2020 IEEE International Conference on Robotics and Automation, ICRA, 2020, pp. 9895–9901, <http://dx.doi.org/10.1109/ICRA40945.2020.9196538>.
- [143] C.D. Santana, V. Arapi, G. Averta, F. Damiani, G. Fiore, A. Settimi, M.G. Catalano, D. Bacciu, A. Bicchì, M. Bianchi, Learning from humans how to grasp: A data-driven architecture for autonomous grasping with anthropomorphic soft hands, *IEEE Robot. Autom. Lett.* 4 (2019) 1533–1540, <http://dx.doi.org/10.1109/LRA.2019.2896485>.
- [144] D. Hidalgo-Carvajal, H. Chen, G.C. Bettelani, J. Jung, M. Zavaglia, L. Busse, A. Naceri, S. Leutenegger, S. Haddadin, Anthropomorphic grasping with neural object shape completion, *IEEE Robot. Autom. Lett.* (2023).
- [145] H. Duan, P. Wang, Y. Li, D. Li, W. Wei, Learning human-to-robot dexterous handovers for anthropomorphic hand, *IEEE Trans. Cogn. Dev. Syst.* 15 (3) (2023) 1224–1238, <http://dx.doi.org/10.1109/TCDS.2022.3203025>.
- [146] P. Ruppel, J. Zhang, Learning object manipulation with dexterous hand-arm systems from human demonstration, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2020, pp. 5417–5424.
- [147] R. Wu, T. Zhu, W. Peng, J. Hang, Y. Sun, Functional grasp transfer across a category of objects from only one labeled instance, *IEEE Robot. Autom. Lett.* 8 (5) (2023) 2748–2755.
- [148] G.-H. Xu, Y.-L. Wei, D. Zheng, X.-M. Wu, W.-S. Zheng, Dexterous grasp transformer, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 17933–17942.
- [149] S. Zhifei, E.M. Joo, A survey of inverse reinforcement learning techniques, *Int. J. Intell. Comput. Cybern.* (2012).
- [150] J. Orbik, A. Agostini, D. Lee, Inverse reinforcement learning for dexterous hand manipulation, in: 2021 IEEE International Conference on Development and Learning, ICDL, IEEE, 2021, pp. 1–7.
- [151] J. Ho, S. Ermon, Generative adversarial imitation learning, *Adv. Neural Inf. Process. Syst.* 29 (2016) 4565–4573.
- [152] M.S. Kopicki, D. Belter, J.L. Wyatt, Learning better generative models for dexterous, single-view grasping of novel objects, *Int. J. Robot. Res.* 38 (10–11) (2019) 1246–1267.
- [153] Y. Liu, Y. Yang, Y. Wang, X. Wu, J. Wang, Y. Yao, S. Schwertfeger, S. Yang, W. Wang, J. Yu, et al., Realdex: Towards human-like grasping for robotic dexterous hand, 2024, arXiv preprint [arXiv:2402.13853](https://arxiv.org/abs/2402.13853).
- [154] A. Gupta, C. Eppner, S. Levine, P. Abbeel, Learning dexterous manipulation for a soft robotic hand from human demonstrations, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2016, pp. 3786–3793.
- [155] V. Kumar, E. Todorov, S. Levine, Optimal control with learned local models: Application to dexterous manipulation, in: 2016 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2016, pp. 378–383.
- [156] M. Omer, R. Ahmed, B. Rosman, S.F. Babikir, Model predictive-actor critic reinforcement learning for dexterous manipulation, in: 2020 International Conference on Computer, Control, Electrical, and Electronics Engineering, ICCCEE, IEEE, 2021, pp. 1–6.
- [157] J. Schulman, S. Levine, P. Abbeel, M. Jordan, P. Moritz, Trust region policy optimization, in: International Conference on Machine Learning, PMLR, 2015, pp. 1889–1897.
- [158] M. Mudigonda, P. Agrawal, M. Dewese, J. Malik, Investigating deep reinforcement learning for grasping objects with an anthropomorphic hand, 2018.
- [159] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, 2017, arXiv preprint [arXiv:1707.06347](https://arxiv.org/abs/1707.06347).
- [160] B. Wu, I. Akinola, A. Gupta, F. Xu, J. Varley, D. Watkins-Valls, P.K. Allen, Generative attention learning: a “GenerAL” framework for high-performance multi-fingered grasping in clutter, *Auton. Robots* 44 (6) (2020) 971–990.
- [161] S. Dasari, A. Gupta, V. Kumar, Learning dexterous manipulation from exemplar object trajectories and pre-grasps, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2023, pp. 3889–3896.
- [162] T. Wu, M. Wu, J. Zhang, Y. Gan, H. Dong, Learning score-based grasping primitive for human-assisting dexterous grasping, *Adv. Neural Inf. Process. Syst.* 36 (2024).
- [163] S. Zhaole, J. Zhu, R.B. Fisher, Dexdlo: Learning goal-conditioned dexterous policy for dynamic manipulation of deformable linear objects, in: 2024 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2024, pp. 16009–16015.
- [164] Z. He, M. Cioarlie, Discovering synergies for robot manipulation with multi-task reinforcement learning, 2021, arXiv preprint [arXiv:2110.01530](https://arxiv.org/abs/2110.01530).
- [165] F. Zhang, Y. Chen, H. Qiao, Z. Liu, SURRL: Structural unsupervised representations for robot learning, *IEEE Trans. Cogn. Dev. Syst.* (2022).
- [166] B. Li, S. Qiu, J. Bai, B. Wang, Z. Zhang, L. Li, H. Wang, X. Wang, Interactive learning for multi-finger dexterous hand: A model-free hierarchical deep reinforcement learning approach, *Knowl.-Based Syst.* 295 (2024) 111847.
- [167] S.H. Huang, M. Zambelli, J. Kay, M.F. Martins, Y. Tassa, P.M. Pilarski, R. Hadsell, Learning gentle object manipulation with curiosity-driven deep reinforcement learning, 2019, ArXiv.
- [168] T. Li, W. Xi, M. Fang, J. Xu, M.Q.-H. Meng, Learning to solve a rubik’s cube with a dexterous hand, in: 2019 IEEE International Conference on Robotics and Biomimetics, ROBIO, IEEE, 2019, pp. 1387–1393.
- [169] R. Fakoor, P. Chaudhari, A.J. Smola, DDPG++: striving for simplicity in continuous-control off-policy reinforcement learning, 2020, arXiv preprint [arXiv:2006.15199](https://arxiv.org/abs/2006.15199).
- [170] W. Huang, I. Mordatch, P. Abbeel, D. Pathak, Generalization in dexterous manipulation via geometry-aware multi-task learning, 2021, arXiv preprint [arXiv:2111.03062](https://arxiv.org/abs/2111.03062).
- [171] M. Alakuijala, G. Dulac-Arnold, J. Mairal, J. Ponce, C. Schmid, Residual reinforcement learning from demonstrations, 2021, arXiv preprint [arXiv:2106.08050](https://arxiv.org/abs/2106.08050).
- [172] L. Huang, W. Cai, Z. Zhu, Z. Zou, Dexterous manipulation of construction tools using anthropomorphic robotic hand, *Autom. Constr.* 156 (2023) 105133.
- [173] Y.-H. Wu, J. Wang, X. Wang, Learning generalizable dexterous manipulation from human grasp affordance, in: Conference on Robot Learning, PMLR, 2023, pp. 618–629.
- [174] E. Valarezo Anazco, P. Rivera Lopez, N. Park, J. Oh, G. Ryu, M.A. Al-antari, T.-S. Kim, Natural object manipulation using anthropomorphic robotic hand through deep reinforcement learning and deep grasping probability network, *Appl. Intell.* 51 (2021) 1041–1055.
- [175] P. Mandikal, K. Grauman, Dexvip: Learning dexterous grasping with human hand pose priors from video, in: Conference on Robot Learning, PMLR, 2022, pp. 651–661.
- [176] Z. Ding, Y. Chen, A.Z. Ren, S.S. Gu, Q. Wang, H. Dong, C. Jin, Learning a universal human prior for dexterous manipulation from human preference, 2023, arXiv preprint [arXiv:2304.04602](https://arxiv.org/abs/2304.04602).
- [177] Z. Chen, S. Chen, C. Schmid, I. Laptev, ViViDex: Learning vision-based dexterous manipulation from human videos, 2024, arXiv preprint [arXiv:2404.15709](https://arxiv.org/abs/2404.15709).
- [178] Y. Xu, W. Wan, J. Zhang, H. Liu, Z. Shan, H. Shen, R. Wang, H. Geng, Y. Weng, J. Chen, et al., Unidexgrasp: Universal robotic dexterous grasping via learning diverse proposal generation and goal-conditioned policy, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 4737–4746.
- [179] Y. Li, B. Liu, Y. Geng, P. Li, Y. Yang, Y. Zhu, T. Liu, S. Huang, Grasp multiple objects with one hand, *IEEE Robot. Autom. Lett.* (2024).
- [180] M. Mosbach, S. Behnke, Grasp anything: Combining teacher-augmented policy gradient learning with instance segmentation to grasp arbitrary objects, 2024, arXiv preprint [arXiv:2403.10187](https://arxiv.org/abs/2403.10187).
- [181] Q. Liu, Y. Cui, Q. Ye, Z. Sun, H. Li, G. Li, L. Shao, J. Chen, DexRepNet: Learning dexterous robotic grasping network with geometric and spatial hand-object representations, in: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2023, pp. 3153–3160.
- [182] Y. Qin, H. Su, X. Wang, From one hand to multiple hands: Imitation learning for dexterous manipulation from single-camera teleoperation, 2022, arXiv:2204.12490.