IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 1 March 2024, accepted 31 March 2024, date of publication 5 April 2024, date of current version 16 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3385426

TOPICAL REVIEW

Advancements in Deep Reinforcement Learning and Inverse Reinforcement Learning for Robotic Manipulation: Toward Trustworthy, Interpretable, and Explainable Artificial Intelligence

RECEP OZALP[1], AYSEGUL UCAR[1], (Senior Member, IEEE), AND CUNEYT GUZELIS[2]
[1]Engineering Faculty, Mechatronics Engineering Department, Firat University, 23119 Elâziğ, Turkey
[2]Engineering Faculty, Electrical and Electronics Engineering, Yaşar University, 35100 İzmir, Turkey

Corresponding author: Recep Ozalp (rozalp@firat.edu.tr)

This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 117E589.

ABSTRACT This article presents a literature review of the past five years of studies using Deep Reinforcement Learning (DRL) and Inverse Reinforcement Learning (IRL) in robotic manipulation tasks. The reviewed articles are examined in various categories, including DRL and IRL for perception, assembly, manipulation with uncertain rewards, multitasking, transfer learning, multimodal, and Human-Robot Interaction (HRI). The articles are summarized in terms of the main contributions, methods, challenges, and highlights of the latest and relevant studies using DRL and IRL for robotic manipulation. Additionally, summary tables regarding the problem and solution are presented. The literature review then focuses on the concepts of trustworthy AI, interpretable AI, and explainable AI (XAI) in the context of robotic manipulation. Moreover, this review provides a resource for future research on DRL/IRL in trustworthy robotic manipulation.

INDEX TERMS Deep reinforcement learning, inverse reinforcement learning, robotic manipulation, artificial intelligence, trustworthy AI, interpretable AI, eXplainable AI.

I. INTRODUCTION

Robots are devices that are produced for different tasks and environments and are developed to meet the needs of people in almost every field. They are designed and programmed to perform specific tasks [1]. The robots are employed in industry and manufacturing, agriculture, health and medicine, space, exploration, military and defense and service sectors [2], [3], [4], [5], [6], [7]. Robotic manipulation applications are among the most used applications in robotics [8], [9], [10], [11].

Deep Reinforcement Learning (DRL) is a frequently used machine learning technique in robotic manipulation applications [12], [13]. The DRL algorithms attempt to find the most appropriate policy for the problems by trial and error [14]. The algorithms have been processed faster with the development of computing tools [15]. They are used to train robots in many manipulation tasks, such as robotic grasping [16], robotic hand manipulation [17], and object manipulation [18]. The algorithms face specific challenges, such as the need for expert knowledge to determine the appropriate reward function [19], to model the complex environment [20], and to construct the proper algorithm for complex tasks [21].

Inverse reinforcement learning (IRL) is a machine learning approach based on obtaining the reward function by observing a policy [22]. The policy can be an expert representation or a working policy [23]. In recent years, IRL has been

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

applied to a wide variety of robotic manipulation tasks such as grasping [24], combining [25], and manipulating objects [26], [27], [28]. The IRL algorithms use observation to learn from a given policy, which makes them too sensitive to noisy observation data [29], [30]. Furthermore, as they also account for the noise within the data, the trained policy is likely incorrect. Moreover, deriving possible multiple reward functions from a single policy causes uncertainty in the solution.

The recent advances in trustworthy Artificial Intelligence (AI) have led to the development of trustworthy robotic technologies [31]. Trustworthy AI emphasizes that the outcomes of AI actions should be explained, and the outputs need to be interpreted [32]. Researchers have discussed it with the concepts of explainable AI (XAI)/ interpretable AI. The concept of trustworthy AI has become vital, especially for critical and sensitive [33]. The XAI algorithms address the need for AI to explain the reasons of the actions it takes [34]. Interpretable AI is defined as understanding the output of the algorithm for the end user [35].

Thanks to the developments in Artificial Intelligence (AI), concepts such as Trustworthy AI, eXplainable AI (XAI), and Interpretable AI which are closely interrelated concepts focusing on the reliability and understandability of AI systems have become increasingly important in instilling confidence in these systems among humans [31]. This confidence is related to the algorithms' accuracy, reliability, and fairness [32]. Trustworthy AI develops systems capable of making correct and fair decisions while preventing misinterpretation or misleading use of data. Trustworthiness does not rely solely on algorithm performance but also depends on the comprehensibility of how these algorithms operate [33]. At that point, the concepts of XAI and Interpretable AI became a current issue. XAI concentrates on the ability of AI models to explain their decision-making processes and outcomes. Humans can easily understand the XAI models and trust them [34]. Interpretable AI emphasizes the understandability of the internal workings of AI models, which is necessary for explaining why a particular decision was made or a specific outcome was reached [35].

Section II presents the classification of DRL considering AI algorithms, robotic manipulation applications, and studies in this field. The problems encountered in robotic manipulation and proposed solutions are addressed there. Section III provides the IRL classification and the studies on robotic manipulation of IRL. The problems being solved by IRL are highlighted in the section. Section IV presents the concepts of trustworthiness/ explainability/ interpretation in robotic manipulation, with articles written about them and DRL and IRL. In Section V, the possible future works are provided.

II. DEEP REINFORCEMENT LEARNING FOR ROBOTIC MANIPULATION

The DRL methods are among the most effective deep learning methods for performing various robotic manipulation tasks. They train an agent in an environment to maximize the reward function [36]. The DRL agent learns through trial and error with a principal learning objective that maps the states of the environment to actions.

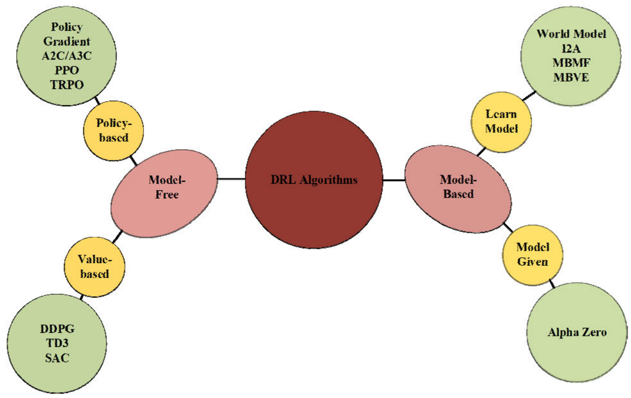


FIGURE 1. DRL classification.

There are several different ways to classify the DRL algorithms. One approach is to categorize DRL according to the particular algorithm or approach used [37]. As shown in Figure 1, DRL is classified according to algorithms [38].

1- Model-free DRL includes algorithms that an agent learns to make decisions and take actions directly from interaction with its environment without explicitly modeling the dynamics of the environment. In model-free DRL, the agent learns a policy or value function directly from experience, typically through trial and error, without requiring a model of the environment's transition dynamics [39]. The algorithms are separated into two branches.

a) Value-based methods are the RL approaches where the agent learns to make decisions and take actions based on estimating the value of different actions or states in the environment. In value-based DRL, the agent typically learns a value function, which assigns a value to each possible action or state. The value is the expected cumulative reward the agent can achieve by taking that action or being in that state and following a particular policy. [40]. Deep Deterministic Policy Gradient (DDPG) [41], Twin Delayed DDPG (TD3) [42], and Soft Actor-Critic (SAC) [43] are value-based algorithms.

b) Policy-based methods are algorithms that directly optimize policy without estimating the value of states or state-action pairs by determining the actions taken by the agent as a function of the agent's state and environment [27]. Policy Gradient [44], Advantage Actor-Critic (A2C) [45], Asynchronous Advantage Actor-Critic (A3C) [46], Proximal Policy Optimization (PPO) [47], and Trust Region Policy Optimization (TRPO) [48] are policy-based algorithms.

2- Model-based DRL includes algorithms that follow the framework of an agent that interacts with an environment, learns a model of that environment, and then uses the model to make decisions [49]. The algorithms are divided into two parts.

a) Learning the model includes algorithms such as The World Model, Imagination-Augmented Agents (I2A) [50], Model-Based RL with Model-Free FineTuning (MBMF) [51], and Model-Based Value Expansion [52].

b) Working on the given model includes algorithms such as the AlphaZero algorithm [53].

In robotic manipulation the classification of DRL is given according to the specific tasks and problem-solving approaches:

Grasping and manipulation: Applications include using DRL for robots to perform ambiguous manipulation tasks such as grasping and hand manipulation [12].

Navigation and localization: Applications where DRL is not used for robots to navigate and localize different environments [54], [55].

Multi-agent systems: Applications where DRL is used to train multiple agents to interact and coordinate with each other [56].

Figure 2 shows some examples from the studies on manipulation tasks such as robotic grasping [57], robotic hand manipulation [58], and object manipulation [59] using DRL algorithms. There are notable studies in the field of robot manipulation with DRL [16]. In the study conducted by OpenAI [60], a 24-degree-of-freedom robotic hand was successfully trained in short training times using learning from human demonstrations for applying complex manipulation tasks. The Robotics Institute at Carnegie Mellon University presented that model-based DRL has higher performance than model-free DRL in object manipulation [61].

FIGURE 2. Robotic manipulation applications by using DRL. (a) Baxter robot grasping [57]. (b) Robot hand block grip [58]. (c) SoftGym robot [59].

In addition to these articles, many other studies have used DRL to train robots to perform manipulation tasks such as sorting objects [62], assembling parts [63], and manipulating flexible objects [64], [65]. As DRL evolves more robots will likely be trained to perform increasingly complex manipulation tasks. Figure 3 shows some studies related to these tasks.

The growing utilization of DRL algorithms has led to its continuous improvement. The industrial applications of DRL in robotic manipulation were explored in [66] and [67] as well as in industrial automation applications [68], path planning [69], electronic circuit production [70], and assembly tasks [71]. Figure 4 presents visuals of the applications.

In training a humanoid robot with DRL for gripper object manipulation in an environment with obstacles [72],the robot

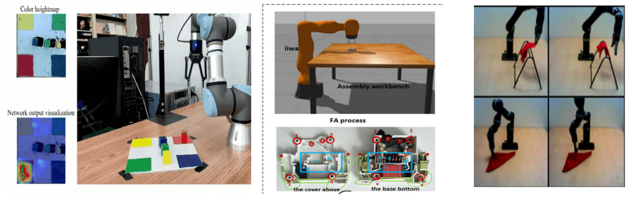


FIGURE 3. Robotic manipulation applications by using DRL. (a) Sequencing four blocks in the real world [62]. (b) Assembling parts [63]. (c) Manipulating flexible objects [64].

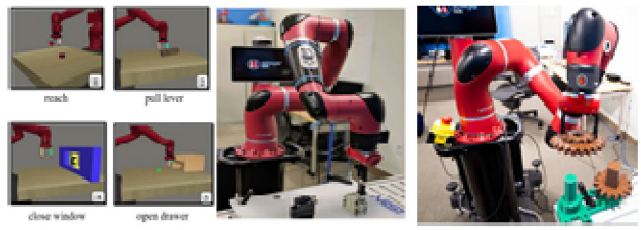


FIGURE 4. Robots are trained by using DRL for robotic manipulation. (a) Adaptable automation task [68]. (b) The electronic circuit production [70]. (c) An assembly task [71].

successfully detected obstacles without extracting features from the image data. An approach to grasping based on DRL haptic feedback to improve grip performance was presented, outperforming non-tactile rewards with tactile reward equations [73].

A study in [74] presented a DRL-based approach to grasp and manipulating unknown objects in real life demonstrating that object manipulation can be performed with DRL using 3D image data without segmentation and image enhancement methods.

A self-monitoring model-based approach was proposed in [75]. The approach learns to predict the future directly from raw sensory readings such as camera images. It was shown that the obtained model works well with previously unseen objects.

A task-oriented comprehension network (TOG-Net) was proposed to jointly optimize the task-oriented comprehension of the tool and manipulation policy [76].

An approach presented in [77] was implemented in dataset collection for robot insight to enhance the efficiency of learning the principles of deep-dip gripping using 3D object CAD models. This approach demonstrated training policies for lifting and moving new objects with complex geometry from a desktop or a box.

The essential structural features of a Markov decision process from offline data were discussed in [78]. This discussion included the performance of surgical robot control, and the creation of efficient execution plans.

Visual perception-based RL was combined with low-level reactive control based on tactile perception to prevent

TABLE 1. The problems in robotic manipulation and the solutions searched using deep reinforcement learning algorithms in the articles between 2018 and 2023.

Article	Task/Problem	Solution Method
Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation [80]	Problem of learning vision-based dynamic manipulation skills.	Closed-loop vision-based control, in which the robot constantly updates its grip strategy based on the latest observations.
Composable Deep Reinforcement Learning for Robotic Manipulation [81]	Poor performance in real-world tasks with limited interaction with the environment.	Soft Q-learning to learn and optimize multimodal exploration strategies.
Goal Representations for Instruction Following: A Semi-Supervised Language Interface to Control [82]	Dompelling data containing representations of tasks tagged with language instruction.	A method fusing language and a common image and introducing target conditional policies using only a small amount of language data.
Training Diffusion Models with Reinforcement Learning [83]	Direct optimization of diffusion models.	The DRL algorithm.
Automatic Parameter Optimization Using Genetic Algorithm in Deep Reinforcement Learning for Robotic Manipulation Tasks [84]	Effect of the learning process on selecting of the values of the hyperparameters used in the learning algorithm.	Genetic algorithm (GA) to fine-tune the values of hyperparameters.
Provably Safe Deep Reinforcement Learning for Robotic Manipulation [85]	No method in RL-based manipulator control guarantees the safety of highly dynamic obstacles.	A rapid accessibility analysis of humans and manipulators to ensure that the manipulator comes to a complete stop before a human comes within range.
Bi-Touch: Bimanual Tactile Manipulation with Sim-to-Real Deep Reinforcement Learning [86]	Due to the complexity of designing effective controllers, bilateral manipulation with haptic feedback is less explored than single-handed manipulation.	A dual-arm tactile robotic system (Bi-Touch) based on the Tactile Gym 2.0 setup, which combines two affordable industrial-level robot arms with low-cost, high-resolution tactile sensors (TacTips).
SURREAL: Open-Source Reinforcement Learning Framework and Robot Manipulation Benchmark [87]	Reproducibility.	SURREAL, an open-source, scalable framework that supports cutting-edge distributed RL algorithms.
Swarm Deep Reinforcement Learning for Robotic Manipulation [88]	Insufficient data sharing among robots and data privacy and security issues.	A swarm RL method, a decentralized DRL technology based on blockchain, has been proposed.
Hierarchical Reinforcement Learning Integrating With Human Knowledge for Practical Robot Skill Learning in Complex Multi-Stage Manipulation [89]	Performing complex manipulation tasks.	The Hierarchical Reinforcement Learning (HRL) framework.
Sim-to-Real Model-Based and Model-Free Deep Reinforcement Learning for Tactile Pushing [90]	Lack of tactile sensing.	A DRL approach to object pushing using haptic sensing without visual input.
Motion Planner Augmented Reinforcement Learning for Robot Manipulation in Obstructed Environments [91]	The RL algorithms requires a great deal of experience in environments with many obstacles that make exploration difficult, and movement planners give erroneous results in tasks that require contact with the environment.	The motion planner-augmented RL (MoPA-RL) to increase an RL agent's field of action with the long-range planning capabilities of motion planners.
Deep Reinforcement Learning Using Genetic Algorithm for Parameter Optimization [92]	Selection of values for the learning algorithm parameters.	GA combined with Hindsight Experience Replay (HER) was used to find the values of parameters used in the Deep DDPG to speed up the learning agent.

TABLE 1. ***(Continued.)*** **The problems in robotic manipulation and the solutions searched using deep reinforcement learning algorithms in the articles between 2018 and 2023.**

SoftGym: Benchmarking Deep Reinforcement Learning for Deformable Object Manipulation [93]	Deformable objects are difficult to manipulate because of their high-dimensional state representation and complex dynamics.	SoftGym, a set of open-source simulated benchmarks for manipulating deformable objects, with a standard OpenAI Gym API and a Python interface to create new environments.
Comparison of End-to-End and Hybrid Deep Reinforcement Learning Strategies for Controlling Cable-Driven Parallel Robots [94]	Effects of integrating DRL with non-learning-based approaches on learning rate and discussion of DRL's robustness for modeling uncertainties.	An end-to-end DRL strategy and a hybrid DRL strategy to control a wired parallel robot.
Deep Multi-Agent Reinforcement Learning for Decentralized Continuous Cooperative Control [95]	Increasing the use of decentralized collaborative robotic control.	Multi-Agent Mujoco, an easily expandable multi-agent benchmark package for robotic the control in continuous motion fields.
A Framework for Efficient Robotic Manipulation [96]	Real robot learning of RL principles has yet to increase as much as in simulation.	A method using data augmentation and unsupervised learning to obtain efficient training of real robot arm policies from sparse rewards.
Dynamic Cloth Manipulation with Deep Reinforcement Learning [97]	Complex reward functions.	A sparse reward approach to improve control policy learning.
A Reinforcement Learning Neural Network for Robotic Manipulator Control [98]	Unknown parameters and dead zones in dynamic fabric processing tasks.	Deep reinforcement learning.
Self-Supervised Sim-to-Real Adaptation for Visual Robotic Manipulation [99]	Difficulty collecting and automatically obtaining reward signals from real robotic visual data for training RL algorithms.	The latent state representation learned indirectly and then adapted to real space using real unlabeled robot data with DRL in simulation.
Learning Mobile Manipulation through Deep Reinforcement Learning [100]	Mobile manipulation is difficult because of the complex coordination of the mobile base and manipulator.	A new mobile manipulation system
Deep reinforcement learning with smooth policy update: Application to robotic cloth manipulation [101]	Massive number of training samples for learning.	Combined policy updates with automatic feature extraction in deep neural networks to improve sample efficiency and learning stability with fewer samples.
Visual Foresight: Model-Based Deep Reinforcement Learning for Vision-Based Robotic Control [75]	Absence of core reality reward signals in real-world robotic tasks.	A self-monitoring model-based approach in which a predictive model learns to predict the future directly from raw sensory readings such as camera images.
Deep Reinforcement Learning for Collision Avoidance of Robotic Manipulators [102]	Collision avoidance problem.	A normalized advantage function (NAF) model-free algorithm.
Learning Synergies Between Pushing and Grasping with Self-Supervised Deep Reinforcement Learning [103]	Skilled robot manipulation, performing non-grasping (e.g., pushing) and grasping (e.g. grasping) actions.	To discover and learn non-apprehensive (e.g., pushing) and grasping (e.g., grasping) actions from scratch through model-free DRL.
Dexterous Manipulation with Deep Reinforcement Learning: Efficient, General, and Low-Cost [104]	Controlling dexterous multi-fingered robotic hands poses a significant challenge because of the high dimensionality of configuration spaces and complex intermittent contact interactions.	An end-to-end approach to map sensor readings to actions with DRL directly.
Reinforcement Learning Tracking Control for Robotic Manipulator With Kernel-Based Dynamic Model [105]	The challenge RL has to perform continuous control tasks.	A reward function according to the characteristics of the tracking control and a kernel-based transition dynamic model RL tracking controller to accelerate the learning process.
Self-Configuring Robot Path Planning With Obstacle Avoidance via Deep Reinforcement Learning [106]	Avoid full-body collision.	The DRL approach to avoid obstacles while performing tasks in the operational space.

TABLE 1. *(Continued.)* **The problems in robotic manipulation and the solutions searched using deep reinforcement learning algorithms in the articles between 2018 and 2023.**

Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey [107]	Degraded performance of policies after models are transferred to real robots due to the gap between the simulated and real worlds.	The essential background behind sim-to-real transfer in DRL and the current methods.
Variable Compliance Control for Robotic Peg-in-Hole Assembly: A Deep-Reinforcement-Learning Approach [108]	The problem of peg-in-hole mounting safely solving complex, high-precision assembly in an unstructured environment.	An out-of-policy, model-independent RL method based on learning to solve peg-in-hole tasks with hole location uncertainty.
IRIS: Implicit Reinforcement without Interaction at Scale for Learning Control from Offline Robot Manipulation Data [109]	Problem learning from offline task demonstrations.	Implicit reinforcement without interaction at scale (IRIS), a new framework for learning from large-scale demonstration datasets.
Deep Reinforcement Learning with Optimized Reward Functions for Robotic Trajectory Planning [110]	To improve the efficiency of DRL-based methods for robotic trajectory planning in an unstructured working environment with obstacles.	An azimuth reward function to accelerate the learning process by modeling position and orientation constraints that can reduce exploratory blindness and a subtask-level reward function to further increase efficiency.
Learning Robotic Manipulation through Visual Planning and Acting [111]	Difficulty in analytically modeling soft or deformable objects.	Robot's self-supervised interaction with the object to perform purposeful object manipulation directly from the raw image data using the Causal InfoGAN generative model.
Learning to Combine Primitive Skills: A Step Towards Versatile Robotic Manipulation [112]	Full status observability in tasks such as preparing food, assembling furniture, and difficulty adapting to dynamic scene changes.	An RL approach to task planning that learns to incorporate primitive skills.
Complicated Robot Activity Recognition by Quality-Aware Deep Reinforcement Learning [113]	Existing manipulator control methods, such as position control and vision-based control methods, do not meet the requirements of autonomous learning.	A DRL scheme quality model was proposed to achieve end-to-end manipulator control.
A Deep Reinforcement-Learning Approach for Inverse Kinematics Solution of a High Degree of Freedom Robotic Manipulator [114]	Computational difficulty of inverse kinematics problems.	A DRL approach to solving the kinematics problem of a 7-degree of freedom robotic manipulator using the Product of Exponentials and Deep Q-Network as an advanced kinematics (FK) computing tool.
Reinforcement Learning for Facilitating Human-Robot-Interaction in Manufacturing [115]	Developing the ability of robotic operators to adapt their behavior to changes in human task performance is a significant challenge to overcome to ensure that many ideas in the larger smart manufacturing paradigm can be realized.	A methodology to effectively model robotic operators and an RL agent capable of autonomous decision-making.
Demonstration-Guided Deep Reinforcement Learning of Control Policies for Dexterous Human-Robot Interaction [116]	Human-robot interactions such as handshaking or clapping.	The DRL-based control policies.
Towards Safe Human-Robot Collaboration Using Deep Reinforcement Learning [117]	Safety in Human-Robot Collaboration (HRC).	The DRL algorithm using in HRC scenarios to increase the intelligence and safety of robots and thus reduce the dangers posed by robots.
An AR-Assisted Deep Reinforcement Learning-Based Approach Towards Mutual-Cognitive Safe Human-Robot Interaction [118]	Due to the increasing individualized demand for production tasks, traditional rule-based secure HRI measures cannot meet security requirements well due to a lack of flexibility and synergy.	A mutual cognitive safe HRI approach, which includes visual augmentation, robot speed control, motion preview and collision detection with Digital Twin, and collision avoidance and motion planning based on DRL in AR.
Prehensile and Non-Prehensile Robotic Pick-and-Place of Objects in Clutter Using Deep Reinforcement Learning [119]	Intelligent and self-controlled industrial pick and place operation for complex environments.	A DRL temporal difference learning model.

TABLE 1. *(Continued.)* **The problems in robotic manipulation and the solutions searched using deep reinforcement learning algorithms in the articles between 2018 and 2023.**

Facilitating Sim-to-real by Intrinsic Stochasticity of Real-Time Simulation in Reinforcement Learning for Robot Manipulation [120]	The RL agents are sensitive to discrepancies between the simulation and the real world.	The properties of the intrinsic stochasticity of real-time simulation (RT-IS) of off-the-shelf simulation software and its potential to improve RL performance.

slippage in [79]. This study aims to fulfill the target of manipulation and minimize the interference of tactile control.

The reviewed articles and studies have revealed DRL's widespread and vital use in robotic manipulation. In addition, they have demonstrated that a wide variety of tasks can be performed using DRL.

Table 1 lists the articles published between 2018 and 2023 that offer solutions to common problems in robotic manipulation tasks. In general, the articles focus on a few difficulties. These are commonly shown as difficulties related to the environment [117], the structure of the robot [89], transferring from simulation to real environment [95], security [88], algorithm speed [112], and parameter selection [92]. Problems arising from the environment are seen as the inability to model complex environments well enough [117], the inability to model the interacting objects in the environment [91], the poor observability of the environment [110], or the high-dimensional dataset problems encountered [100] in modeling the environment. The difficulties arising from the robot's structure include calculating the multi-joint robot's kinematic equations [112] or its movement in complex environments due to the limited freedom of movement [103]. In the problems experienced in the transfer from the simulation to the real environment, the simulation environment dynamics [105] and the visuals cannot match the real environment [98]. When we look at the security problems, doubts about ensuring the safety of people in environments where human-robot interaction mostly come to the fore [115].

The flexible structure of the DRL has solved the problems mentioned above. Methods such as task fragmentation have been used in complex environment tasks. Some studies combined DRL with different machine-learning algorithms in tasks requiring parameter optimization to increase the algorithm speed with multi-agent systems. The difficulties in modeling the environment and transferring from simulation to reality have been accomplished with DRL.

III. INVERSE REINFORCEMENT LEARNING FOR ROBOTIC MANIPULATION

In recent years, with the development of RL algorithms, robots can learn from experience [29]. In addition, with the developments in simulation technology, the further development of training algorithms has enabled them to improve their performance over time, and the area and number of uses have increased accordingly. Developments in simulation environments such as ROS [119], Gazebo [93], Webots [120], and V-REP [121] and advances in simulation physics engines have accelerated the development of DRL.

This section focuses on recent developments in IRL for robotic manipulation, particularly over the past five years (2018-2023). Various problems in this area are discussed, and articles on their solutions are reviewed and summarized. Thus, it has been aimed at contributing to the solution of similar problems.

The following are classified IRL according to specific tasks or applications of robotic manipulation:

1-Grasping: Robotic grasping tasks using IRL aims to learn a grip policy that a robot executes to grasp an object [122], [123], [124], [125], [126], [127].

2-Assembly: Robotic assembly tasks using IRL aim to learn a policy being executed by a robot to assemble a product [25], [128], [129], [130].

3-Manipulation with indefinite rewards: The use of IRL arises in robotic manipulation tasks where it may be difficult or impossible to give the robot explicit rewards [30], [131], [132].

4-Multitasking and transfer learning: Transfer learning is used to increase the sample efficiency of IRL by allowing a robot to transfer information between tasks [126], [133], [134], [135], [136].

5-Multimodal IRL: Learning from different feedback forms, such as visual, tactile, and verbal feedback is facilitated by fusing IRL [137], [138], [139], [140].

6- Active IRL: The robots utilize IRL to actively solicit feedback from humans, which increases sample efficiency and robustness of learned policies [141], [142], [143], [144].

7- HRI: Incorporating human feedback and preferences, IRL improves the performance and robustness of learned policies [145], [146], [147], [148], [149].

In recent years, a wide variety of robotic manipulation tasks such as grasping, combining, and manipulating objects have been addressed by using IRL [24], [150], [151]. Figure 5 provides a visual representation of these studies.

When the articles including IRL have been examined, the following articles have come to the fore: a study [126] aims to contribute to developing multitasking IRL in the computationally more efficient maximum causal entropy (MCE) IRL framework. A Bayesian Inverse Reinforcement Learning Fail (BIRLF) algorithm allows the agent to use successful and unsuccessful observations by taking advantage of failed demonstrations [122]. In [152], the DDPG and Principal Component Analysis (PCA) methods have been used to show

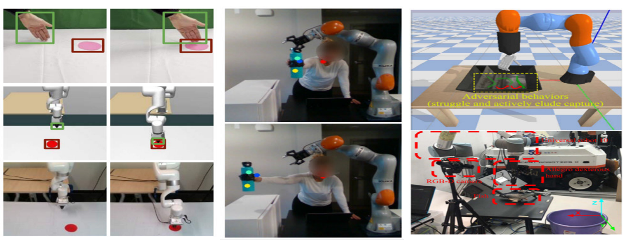
TABLE 2. The problems in robotic manipulation and the suggested solutions using IRL algorithms in the articles between 2018 and 2023.

Article	Task/Problem	Solution Method
Robot Hand-Eye Cooperation Based on Improved Inverse Reinforcement Learning [165]	A highly optimized hand-eye coordination model of the robot was designed to improve the robot's on-site decision-making ability.	Combining an IRL algorithm and a generative adversarial network.
Meta-Inverse Reinforcement Learning with Probabilistic Context Variables [166]	The problem with IRL is that the agent requires multiple representations to correctly understand the reward for each task.	A model extracting rewards for new, structurally similar tasks from a single demonstration.
A Context-Based Multi-task Hierarchical Inverse Reinforcement Learning Algorithm [167]	Low data efficiency and poor performance of existing MIL algorithms for complex long tasks.	Multitasking Hierarchical Contention IRL (MH-AIRL) to learn the principles of hierarchically structured multitasking.
Deep Imitation Learning for Bimanual Robotic Manipulation [168]	Causes difficulty generalizing manipulation skills to objects in different locations.	Modelling relational information in the environment to significantly improve generalization.
Learning Forceful Manipulation Skills from Multi-modal Human Demonstrations [169]	Limiting task-parameter representations to pose representations only, and thus only to skills with spatial and temporal properties.	The DRL and demonstration learning (LfD) framework to expand with multimodal demonstrations, including robot end-effector poses, force and torque readings, and operation scenes to address powerful manipulation skills such as assembly.
Robust Multi-Modal Policies for Industrial Assembly via Reinforcement Learning and Demonstrations: A Large-Scale Study [170]	DRL lacks prohibitively large design space for industrial assembly.	Defining industry-focused DRL criteria and comparing a family of learning approaches (DRL from demonstration) according to these criteria with a professional industrial integrator in the recently established NIST assembly benchmark.
Inferring Task Goals and Constraints using Bayesian Nonparametric Inverse Reinforcement Learning [25]	Failure of standard IRL approaches to model the existence of locally consistent constraints that may only be active on a portion of an impression.	Constraint-Based Bayesian Non-Parametric IRL (CBN-IRL), models the observed behavior as a set of subtasks, each consisting of a target and a set of locally effective constraints.
Deep Adaptive Multi-Intention Inverse Reinforcement Learning [171]	A DRL framework can learn a previously unknown number of nonlinear reward functions from the demonstrations of unlabeled experts.	Tools from Dirichlet processes, and an adaptive approach to simultaneously account for both a complex and an unknown number of reward functions.
Adversarial Inverse Reinforcement Learning With Self-Attention Dynamics Model [172]	Due to the stochastic policy of contentious IRL (AIRL), the current computational graph cannot be differentiated end-to-end.	Model-Based Arbitrary IRL (MAIRL), an end-to-end model-based policy optimization method.
Graph Inverse Reinforcement Learning from Diverse Videos [127]	Much of the previous work is still limited to training from a relatively limited field of video.	The true potential of a third-party URL for increasing the diversity of videos for better scaling.
SWIRL: A Sequential Windowed Inverse Reinforcement Learning Algorithm for Robot Tasks With Delayed Rewards [126]	Delayed rewards and overfitting of training data cause difficulties in robot learning.	The principles of sequential robot tasks using a specific demonstration sequence.
Reward Identification in Inverse Reinforcement Learning [30]	Problem of reward identifiability.	The formulation of the reward identification problem in IRL and identifiability related to the features of the MDP model.
Learning Soft Robotic Assembly Strategies from Successful and Failed Demonstrations [132]	New soft robotic assembly strategies.	Formulation of the problem as an RL task and the reward function obtained from human demonstrations.
Better-than-Demonstrator Imitation Learning via Automatically-Ranked	Difficulty in deriving preferences over representations in imitation learning and unclear when it can be successfully predicted	Disturbance Reward Extrapolation (D-REX), a sequencing imitation learning method that adds noise to a policy learned through behavioral

TABLE 2. ***(Continued.)*** **The problems in robotic manipulation and the suggested solutions using IRL algorithms in the articles between 2018 and 2023.**

Demonstrations [133]	beyond the demonstrator's performance.	cloning to automatically generate ranked displays.
Learning Variable Impedance Control via Inverse Reinforcement Learning for Force-Related Tasks [131]	The DRL and LfD-based approaches are typically task-specific and can be sensitive to task-setting changes.	An IRL-based approach to recover the variable impedance policy and the reward function from expert demonstrations. In addition, different action domains of reward functions have been explored to obtain a more general representation of expert variable impedance skills.
Online Inverse Reinforcement Learning with Learned Observation Model [134]	Extending I2RL to real-world robotics applications using noisy observations and an unknown observation model.	A model to determine the maximum entropy distribution through the observation features that govern the perception process and then uses the inferential observation model to learn the reward function.
Efficient Insertion Control for Precision Assembly Based on Demonstration Learning and Reinforcement Learning [130]	Multiple peg-in-hole insertion control.	A state-to-action policy mapping model based on the Gaussian Mixing Model (GMM) and a policy learning process of insertion with GMR to generalize policy reuse.
Learning Virtual Grasp with Failed Demonstrations via Bayesian Inverse Reinforcement Learning [124]	Use of failed demonstrations.	Bayesian IRL with Failure (BIRLF).

how IRL can transfer task knowledge from a human expert to a robot in a dynamic environment. A method for applying demonstration learning using IRL has been presented in [129]. Figure 5 shows an application of this study.



(a) (b) (c)

FIGURE 5. Examples of IRL. (a). A real environment where a dexterous hand is used to grasp toy fish [24]. (b). Human Notation [150] used for the IRL algorithm to extract cost functions. (c). Image-based policies are trained in simulation with learned reward functions and performed on a real robot [151].

Another area of research has been on the use of IRL for robotic assembly; here, the aim is to learn a policy that a robot can execute to assemble a product [128], [153], [154]. There has also been research on using IRL for robotic manipulation with ambiguous rewards, where it is difficult or impossible to give the robot explicit rewards [155], [156], [157]. Researchers have proposed methods for using IRL in such scenarios by incorporating uncertainty in rewards or using other feedback forms such as sensor data or human demonstrations [158].

In addition to the applications for this particular task, the researchers have proposed new IRL algorithms that can improve performance, robustness, and sample efficiency. These include methods such as maximum entropy IRL [159], inverse optimal control [160], and IRL with expert demonstrations [161].

The performance of IRL has steadily improved as IRL allows robots to learn from human demonstrations without defining a reward function, which creates a promising situation for using IRL in robotic manipulation in future studies. Table 2 shows some distinguished articles on the use of IRL in robotic manipulation tasks between 2018 and 2023, the problems discussed in the articles, and their solutions. These articles discuss various problems in using DRL and IRL in robotic manipulation tasks and the approaches to solving these problems. These problems include difficulty in obtaining the appropriate reward function [30], [163], inability to model the environment well enough [25], sensitivity to noise [132], task-specificity and generalizability [129], [165], excessive dependence on the quality of representations [131], scalability problems [169], and data overload for complex tasks [164].

Despite these difficulties, several IRL methods have been applied to various robotic manipulation tasks, and different types of IRL methods have been introduced to eliminate these difficulties. Due to their strengths and difficulties, the IRL algorithms will continue to evolve. Moreover, the requirement for IRL usage in real-world robotic systems is expected to increase to solve issues such as robustness, and scalability.

TABLE 3. Examples of articles in the field of trustworthy/interpretable/explainable AI for robotic manipulation.

Article	Task/Problem	Solution Method
An AR-assisted Deep Reinforcement Learning-based approach towards mutual-cognitive safe human-robot interaction [118]	Traditional rule-based secure HRI measures fail to meet security requirements well due to a lack of flexibility and synergy.	A mutual cognitive safe HRI approach that includes visual augmentation, robot speed control, motion preview and collision detection with Digital Twin, and collision avoidance and motion planning based on DRL in AR.
Provably Safe Deep Reinforcement Learning for Robotic Manipulation [85]	No method in RL-based manipulator control guarantees the safety of highly dynamic obstacles.	A rapid accessibility analysis of humans and manipulators to ensure that the manipulator comes to a complete stop before a human comes within range.
Data-Efficient, Explainable and Safe Payload Manipulation: An Illustration of the Advantages of Physical Priors in Model-Predictive Control [182]	Decisions made by learned policies or predictions made by learned dynamic models have several disadvantages, such as inability to be easily interpreted by a human user without using XAI techniques and increased difficulty in debugging and integrating into security-critical systems.	Incorporating prior knowledge of environmental kinematics and dynamics into the environmental model or decision process.
Safe Reinforcement Learning using Formal Verification for Tissue Retraction in Autonomous Robotic-Assisted Surgery [183]	Current DRL methods do not guarantee any safety criteria because they maximize cumulative rewards without considering the risks associated with the actions taken.	A Safe-DRL framework with security constraints for the automation of surgical subtasks through DRL training.
Formal Verification for Safe Deep Reinforcement Learning in Trajectory Generation [184]	Secure the DRL problem using formal validation in a trajectory generation task.	An approach to verify whether a trained model can generate trajectories guaranteed to meet safety specifications (for example, operating in a limited workspace).
Deep Learning Technology for Construction Machinery and Robotics [185]	Dataset limitation in deep learning for construction machinery and robotics applied in autonomous construction: lack of interpretability and insufficient autonomous intelligence.	Datasets with expert knowledge, trustworthy AI, productive deep learning, and extraterrestrial structure have been used to compensate for this deficiency.
A Formal Methods Approach to Interpretable Reinforcement Learning for Robotic Planning [178]	This raises security concerns because reward functions for complex tasks are difficult to define formally, and erroneous rewards tend to be exploited by the learner agent.	A method to integrate task specifications with domain-specific knowledge, make the reward creation process easily interpretable, guide policy-making according to the specification, and ensure critical security components.
Interpretable Machine Learning: Definitions, Methods, and Applications [186]	It is unclear how the variety of suggested interpretation methods in machine learning models relate to each other and what common concepts can be used to evaluate them.	The predictive, descriptive, relevant (PDR) framework to define interpretability and discuss interpretations in machine learning.
Constrained Reinforcement Learning for Robotics via Scenario-Based Programming [187]	Adopting security and optimizing performance for safety-critical missions where human safety and expensive equipment may be a concern.	Incorporating domain expert knowledge into a restricted DRL training cycle.
Robotic Lever Manipulation using Hindsight Experience Replay and Shapley Additive Explanations [189]	Problem of interpretability or human explainability of robot decision-making processes.	A hierarchical DRL system consisting of a low-level agent that manages significant actions.
Learning Deep Energy Shaping Policies for Stability-Guaranteed Manipulation [190]	The traditional stability analysis of DRL becomes difficult due to the uninterpretable nature of neural network policies and unknown system dynamics.	Achieving stability deriving a deep interpretable policy structure based on the energy shaping control of Lagrange systems.

TABLE 3. ***(Continued.)*** **Examples of articles in the field of trustworthy/interpretable/explainable AI for robotic manipulation.**

Explainability in Deep Reinforcement Learning, a Review into Current Methods and Applications [191]	Lack of interpretability.	Explainable AI
Causal versus Marginal Shapley Values for Robotic Lever Manipulation Controlled using Deep Reinforcement Learning [192]	The effect of including application knowledge about the causal relationships of states of a robotic system when generating descriptions of deep neural network policies has been explored.	Two methods from the XAI KernelSHAP and causal SHAP were compared on a deep neural network trained using DRL to control a lever using a robotic manipulator.
Explainability in Deep Reinforcement Learning [181]	How XAI techniques can help us understand models beyond classification tasks for RL.	Recent studies on XAI intended for use in public applications with diverse audiences that require ethical, responsible, and trustworthy algorithms.

TABLE 4. **Advantages and disadvantages of DRL and IRL.**

Deep Reinforcement Learning		Inverse Reinforcement Learning	
Advantage	**Disadvantage**	**Advantage**	**Disadvantage**
Performing complex manipulation tasks [89].	Effect of the learning process on selection of the values of the hyperparameters used in the learning algorithm [94].	An IRL framework can learn a previously unknown number of nonlinear reward functions from the demonstrations of unlabeled experts [171].	The problem with IRL is that the agent requires multiple representations to understand the reward for each task correctly [166].
Overcoming the challenge of analytically modeling soft or deformable objects [111].	No method in RL-based manipulator control guarantees the safety of highly dynamic obstacles [85].	New soft robotic assembly strategies [132].	Low data efficiency and poor performance of existing MIL algorithms for complex long tasks [167].
Full status observability in tasks such as preparing food, assembling furniture, and adapting to dynamic scene changes [112].	Reinforcement learning algorithms requires a great deal of experience in environments with many obstacles that make exploration difficult, and movement planners give erroneous results in tasks that require contact with the environment [91].		Failure of standard IRL approaches to model the existence of locally consistent constraints that may only be active on a portion of an impression [25].
Overcoming the computational difficulty of inverse kinematics problems [114].	Effects of integrating DRL with non-learning-based approaches on learning rate and discussion of DRL's robustness for modeling uncertainties [94].		Due to the stochastic policy of contentious IRL (AIRL), the current computational graph cannot be differentiated end-to-end [125].
Developing the ability of robotic operators to adapt their behaviour to changes in human task performance ensures that many ideas in the larger smart manufacturing paradigm can be realized [115].	The challenge RL has to perform continuous control tasks [105].		Problem of reward identifiability [30].
DRL in HRC scenarios to increase the intelligence and safety of robots and thus reduce the dangers posed by robots [117].	Agents of RL are sensitive to discrepancies between the simulation and the real world [193].		
Intelligent and self-controlled industrial pick and place operation for complex environments [119].	DRL lacks prohibitively large design space for industrial assembly [170].		

IV. TRUSTWORTHY/INTERPRETABLE/EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR ROBOTIC MANIPULATION

The High-Level Expert Group on AI has announced that trustworthy AI should have three vital components. They are: i) lawful, ii) ethical, and iii) robust [170]. The reality of controlling robots most of our lives today makes these concepts essential. Only the applications of AI that comply with these principles can be considered trustworthy. Moreover, an AI algorithm should also explain the reasons for this, such as making a decision [171]. In addition, the algorithm's output should be interpretable; that is, the algorithm's outputs should be understood [172]. In this way, scientific progress can be made by interpreting the wrong output, even in false outputs.

Trustworthy/ interpretable/ explainable AI concepts are gaining importance as robotic manipulation is used in important areas such as manufacturing, logistics, and healthcare [173]. With the development of Industry 5.0, the need for autonomous systems that can perform complex robotic manipulation tasks in smart factories and cyber-physical systems has increased. The trustworthiness of robots has proven to be even more important in critical manipulation applications such as medicine, space, nuclear fields, and HRI [174]. At the same time, it is expected that the explainability and interpretability of robots' actions and decision-making mechanisms will be high to continue developments in these areas and maintain trust in robots [175].

In recent years, researchers have used DRL and IRL in robotic manipulation. They have attempted to make the decision-making process interpretable and explainable by learning control policies and ensuring the safety and robustness of the learned policies. However, the topic is challenging for DRL and IRL [176].

In the case of trustworthy AI, various approaches of DRL have been presented to learn robust and secure principles that can operate in uncertain and dynamic environments. In the proposed study [177], a safe system has been developed to grasp and lift objects connected by a human operator in a power-assisted robotic system. The algorithms relating to DRL and IRL are applied to interpretable and explainable AI to understand the decision-making step of learned policies [178]. Currently, the methods have different challenges, which are helpful for debugging, monitoring, and explainability.

Table 3 presents the review of articles written in the field of Trustworthy/ Interpretable/ explainable AI for robotic manipulation between 2018 and 2023. The problems addressed in these articles and their solution methods are indicated. The articles have obtained the policies people can understand and showed how DRL can be used to understand learned policy decision-making. In this way, they have created solutions for the difficulties of DRL in the context of trustworthy/ explainable/ interpretable AI in robotic manipulation applications.

In addition, the articles have discussed problems such as the lack of flexibility in rule-based security restrictions in HRI applications [116], security problems against dynamic obstacles [85], policy-making deficiencies in DRL and IRL techniques without addressing security [179], optimizing policy speed to ensure security [184], and the lack of interpretability and explainability of robot decision-making processes [185]. As problem-solving is discussed in the articles, algorithms and methods have been proposed as trustworthy, explainable, and interpretable [178], [183]. Solutions, such as the use of digital twin methods, greater inclusion of environmental dynamics in the decision-making process, the use of the XAI algorithm for ethical issues and areas requiring security, and the inclusion of expert knowledge in the DRL training process have been adopted. Moreover, because DRL tries to maximize the given reward function, the authors have suggested adding security restrictions to this process by creating security vulnerabilities [180]. In addition, faulty reward functions also damage trustworthiness. Therefore, the authors proposed more interpretable and explainable algorithms.

V. FUTURE WORKS

In this section, future works are given on using DRL and IRL in robotic manipulation and then on integrating trustworthy/interpretable/explainable AI into them.

This article suggests some future work as below:

1. Using the concepts of trustworthiness, interpretability, and explainability with DRL and IRL methods in robotic manipulation. Examples include determining the metrics and benchmarks for the concepts and considering human feedback to incorporate these insights in the learning process.

2. Providing new debugging, monitoring, and explainability tools to demonstrate how the learned policies carry out decision-making. An example includes developing real-time visualization tools for monitoring policy execution.

3. Performing DRL and IRL in multi-robot systems and swarm robots for collaboration aim in manipulation tasks. Examples cover defining collaboration metrics and protocols for multi-robot manipulation and investigating communication schemes among swarm robots for efficient collaboration.

4. Investigating efficient use of DRL and IRL in the case of continuous motion and high-dimensional observations in robotic manipulation tasks. Examples include the optimization of continuous action spaces and the reduction of handling high-dimensional observations.

5. Developing new methods to evaluate the performance of DRL and IRL methods for robotic manipulation tasks by considering deployment and security challenges in the real world.

6. Developing the applications of DRL and IRL in manipulation tasks using mobile robots, aerial robots, and service robots in different industrial fields such as automation.

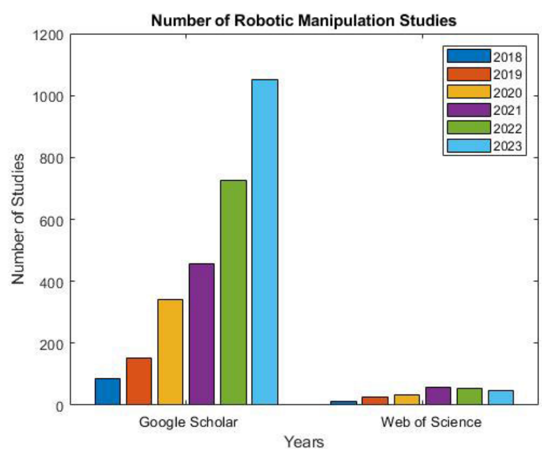


FIGURE 6. Number of robotic manipulation studies using DRL (Google scholar and web of science in 2018-2023) [190], [191].

7. Using DRL and IRL with other machine learning techniques and computer vision to increase performance in robotic manipulation tasks.

8. Developing new transfer learning methods between various robotic platforms and manipulation tasks to reduce learning and deployment costs.

9. Fusing Natural Language Processing, visual sensors, and other sensors in robotics manipulation tasks and integrating digital twins into new methods. Examples encompass the design of new interfaces for natural language communication with robots and the integration of digital twins to improve simulation-to-real-world transfer.

10. Developing the transfer from simulation to real-world applications in robotic manipulation tasks.

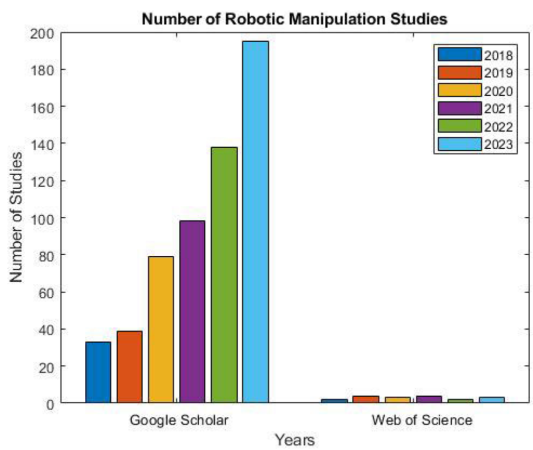


FIGURE 7. **Number of robotic manipulation studies using IRL (Google scholar and web of science in 2018-2023) [192], [193].**

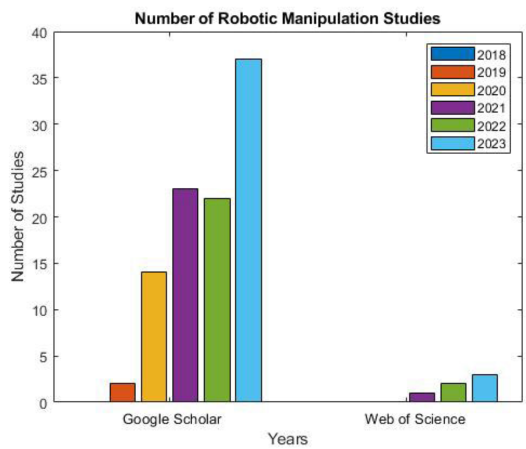


FIGURE 8. **Number of robotic manipulation studies using Trustworthy/Interpretable/ eXplainable AI (Google scholar and web of science in 2018-2023) [194], [195], [196], [197], [198], [199].**

VI. CONCLUSION

This review aims to provide an overview of the last five years of research on DRL and IRL for robotic manipulation. First, the DRL algorithms are introduced, and studies in the field of robotic manipulation are examined. Then, the problems and solution methods that the prominent articles using IRL in robotic manipulation are given. Finally, trustworthy/ interpretable/ eXplainable AI concepts are given in the context of robotic manipulation and articles have been reviewed in this field. These articles provide the main problems and solutions in this field. The difficulties arising from robotic manipulation, the applications of DRL and IRL to robotic manipulation in simulation or real environments and the proposed solutions are provided in these articles. Our literature review concludes that i) both DRL and IRL are commonly used to train robots to perform a wide variety of manipulation tasks, ii) both DRL and IRL have their advantages and disadvantages and can be used to achieve different goals in robotic manipulation, iii) in the context of trustworthy AI, interpretable AI and XAI, DRL and IRL can be used to train robots to perform tasks more efficiently and trustworthy. Table 4 shows the reviewed studies' advantages and disadvantages of DRL and IRL.

The number of articles reviewed in this article for 2018 and 2023 are shown in Figure 6-7-8. Figure 6 shows the number of articles written in the field of robot manipulation using the DRL algorithm in Google Scholar and Web of Science. Figure 7 shows the number of articles written in the field of robot manipulation using the IRL algorithm in Google Scholar and Web of Science. Figure 8 shows the number of articles written robot manipulation using of Trustworthy/ Interpretable/ Explainable AI in Google Scholar. As seen in the graphics, the number of articles in the examined areas has increased from year to year. This increase shows that DRL and IRL are suitable algorithms for robotic manipulation, and their use in solving the difficulties in robotic manipulation applications is increasing daily. It can be predicted that its use will further increase in the coming years.

Studies in this area should attempt to develop new methods to train robots to perform complex manipulation tasks in more complex and dynamic environments that align with current developments. While performing these tasks, the robots are expected to be interpretable, explainable, and trustworthy so they can be used more safely and transparently. Researchers have conducted many studies on these issues. More research is needed to make DRL and IRL more interpretable, explainable, and trustworthy so that robots can be used more safely and transparently while performing these tasks. It has been observed that the use of DRL in robotic manipulation has increased over the years since 2018. Currently, ongoing studies show that the use of DRL in robotic manipulation will increase. This article provides a quick literature review for new researchers working in this field.

REFERENCES

[1] S. B. Niku, *Introduction To Robotics: Analysis, Control, Applications*. Wiley, 2011. [Online]. Available: http://ci.nii.ac.jp/ncid/BB04086836

[2] R. R. Murphy, "Introduction to AI robotics," *Ind. Robot: Int. J.*, vol. 28, no. 3, pp. 266–267, Jun. 2001, doi: 10.1108/ir.2001.28.3.266.1.

[3] M. Ben-Ari and F. Mondada, "Robots and their applications," in *Elements of Robotics*, 2018, pp. 1–20, doi: 10.1007/978-3-319-62533-1_1.

[4] I. Tsitsimpelis, C. J. Taylor, B. Lennox, and M. J. Joyce, "A review of ground-based robotic systems for the characterization of nuclear environments," *Prog. Nucl. Energy*, vol. 111, pp. 109–124, 2019, doi: 10.1016/j.pnucene.2018.10.023.
[5] D. Patil, M. Ansari, D. Tendulkar, R. Bhatlekar, V. N. Pawar, and S. Aswale, "A survey on autonomous military service robot," in *Proc. Int. Conf. Emerg. Trends Inf. Technol. Eng. (ic-ETITE)*, Feb. 2020, pp. 1–7, doi: 10.1109/ic-ETITE47903.2020.78.
[6] T. Duckett, "Agricultural robotics: The future of robotic agriculture," 2018, *arxiv.1806.06762*.
[7] J. Luo, E. Solowjow, C. Wen, J. A. Ojea, A. M. Agogino, A. Tamar, and P. Abbeel, "Reinforcement learning on variable impedance controller for high-precision robotic assembly," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 3080–3087, doi: 10.1109/ICRA.2019.8793506.
[8] M. Shridhar, L. Manuelli, and D. Fox, "CLIPORT: What and where Pathways for robotic manipulation," in *Proc. 5th Annu. Conf. Robot Learn.*, Jun. 2021, pp. 1–13. [Online]. Available: https://openreview.net/pdf?id=9uFiX_HRsIL
[9] S. Nair, A. Rajeswaran, V. Kumar, C. Finn, and A. Gupta, "R3M: A universal visual representation for robot manipulation," 2022, *arXiv:2203.12601*.
[10] Z. Feng, G. Hu, Y. Sun, and J. Soon, "An overview of collaborative robotic manipulation in multi-robot systems," *Annu. Rev. Control*, vol. 49, pp. 113–127, Jan. 2020, doi: 10.1016/j.arcontrol.2020.02.002.
[11] E. Papadopoulos, F. Aghili, O. Ma, and R. Lampariello, "Robotic manipulation and capture in space: A survey," *Frontiers Robot. AI*, vol. 8, Jul. 2021, Art. no. 686723, doi: 10.3389/frobt.2021.686723.
[12] R. Liu, F. Nageotte, P. Zanne, M. de Mathelin, and B. Dresp-Langley, "Deep reinforcement learning for the control of robotic manipulation: A focussed mini-review," *Robotics*, vol. 10, no. 1, p. 22, Jan. 2021, doi: 10.3390/robotics10010022.
[13] H. Nguyen and H. La, "Review of deep reinforcement learning for robot manipulation," in *Proc. 3rd IEEE Int. Conf. Robotic Comput. (IRC)*, Feb. 2019, pp. 590–595, doi: 10.1109/IRC.2019.00120.
[14] A. S. Morgan, D. Nandha, G. Chalvatzaki, C. D'Eramo, A. M. Dollar, and J. Peters, "Model predictive actor-critic: Accelerating robot skill acquisition with deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2021, pp. 6672–6678, doi: 10.1109/ICRA48506.2021.9561298.
[15] J. Hua, L. Zeng, G. Li, and Z. Ju, "Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning," *Sensors*, vol. 21, no. 4, p. 1278, Feb. 2021, doi: 10.3390/s21041278.
[16] M. Q. Mohammed, K. L. Chung, and C. S. Chyi, "Review of deep reinforcement learning-based object grasping: Techniques, open challenges, and recommendations," *IEEE Access*, vol. 8, pp. 178450–178481, 2020, doi: 10.1109/ACCESS.2020.3027923.
[17] H. Zhu, J. Yu, A. Gupta, D. Shah, K. Hartikainen, A. Singh, V. Kumar, and S. Levine, "The ingredients of real-world robotic reinforcement learning," 2020, *arXiv:2004.12570*.
[18] M. Vecerik, O. Sushkov, D. Barker, T. Rothorl, T. Hester, and J. Scholz, "A practical approach to insertion with variable socket position using deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot Autom.*, May 2019, pp. 754–760, doi: 10.1109/ICRA.2019.8794074.
[19] P. Ladosz, L. Weng, M. Kim, and H. Oh, "Exploration in deep reinforcement learning: A survey," *Inf. Fusion*, vol. 85, pp. 1–22, Sep. 2022, doi: 10.1016/j.inffus.2022.03.003.
[20] G. Kahn, A. Villaflor, B. Ding, P. Abbeel, and S. Levine, "Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 5129–5136, doi: 10.1109/ICRA.2018.8460655.
[21] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3826–3839, Sep. 2020, doi: 10.1109/TCYB.2020.2977374.
[22] J. Jara-Ettinger, "Theory of mind as inverse reinforcement learning," *Current Opinion Behav. Sci.*, vol. 29, pp. 105–110, Oct. 2019, doi: 10.1016/j.cobeha.2019.04.010.
[23] S. N. Aslan, R. Ozalp, A. Uçar, and C. Güzelis, "End-to-end learning from demonstation for object manipulation of robotis-Op3 humanoid robot," in *Proc. Int. Conf. Innov. Intell. Syst. Appl. (INISTA)*, Aug. 2020, pp. 1–6, doi: 10.1109/INISTA49547.2020.9194630.
[24] Z. Hu, Y. Zheng, and J. Pan, "Grasping living objects with adversarial behaviors using inverse reinforcement learning," *IEEE Trans. Robot.*, vol. 39, no. 2, pp. 1151–1163, Apr. 2023, doi: 10.1109/TRO.2022.3226108.
[25] D. Park, M. Noseworthy, R. Paul, S. Roy, and N. Roy, "Inferring task goals and constraints using Bayesian nonparametric inverse reinforcement learning," in *Proc. PMLR*, May 2020, pp. 1005–1014, Accessed: May 21, 2023. [Online]. Available: https://proceedings.mlr.press/v100/park20a.html
[26] B. Fang, S. Jia, D. Guo, M. Xu, S. Wen, and F. Sun, "Survey of imitation learning for robotic manipulation," *Int. J. Intell. Robot. Appl.*, vol. 3, pp. 362–369, Sep. 2019, doi: 10.1007/S41315-019-00103-5.
[27] Z. Xie, Q. Zhang, Z. Jiang, and H. Liu, "Robot learning from demonstration for path planning: A review," *Sci. China Technological Sci.*, vol. 63, no. 8, pp. 1325–1334, Aug. 2020, doi: 10.1007/S11431-020-1648-4.
[28] Y. Liu, A. Gupta, P. Abbeel, and S. Levine, "Imitation from observation: Learning to imitate behaviors from raw video via context translation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1118–1125, doi: 10.1109/ICRA.2018.8462901.
[29] S. Arora and P. Doshi, "A survey of inverse reinforcement learning: Challenges, methods and progress," *Artif. Intell.*, vol. 297, Aug. 2021, Art. no. 103500, doi: 10.1016/j.artint.2021.103500.
[30] K. Kim, S. Garg, K. Shiragur, and S. Ermon, "Reward identification in inverse reinforcement learning," in *Proc. PMLR*, Jul. 2021, pp. 5496–5505, Accessed: Jun. 4, 2023. [Online]. Available: https://proceedings.mlr.press/v139/kim21c.html
[31] J.-P.-A. Yaacoub, H. N. Noura, O. Salman, and A. Chehab, "Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations," *Int. J. Inf. Secur.*, vol. 21, no. 1, pp. 115–158, Mar. 2021, doi: 10.1007/s10207-021-00545-8.
[32] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "XAI—Explainable artificial intelligence," *Sci. Robot.*, vol. 4, no. 37, Dec. 2019, Art. no. eaay7120, doi: 10.1126/scirobotics.aay7120.
[33] M. Ryan, "In AI we trust: Ethics, artificial intelligence, and reliability," *Sci. Eng. Ethics*, vol. 26, no. 5, pp. 2749–2767, Oct. 2020, doi: 10.1007/S11948-020-00228-Y.
[34] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. Zhu, "Explainable AI: A brief survey on history, research areas, approaches and challenges," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11839, 2019, pp. 563–574, doi: 10.1007/978-3-030-32236-6_51.
[35] J. D. Fuhrman, N. Gorre, Q. Hu, H. Li, I. El Naqa, and M. L. Giger, "A review of explainable and interpretable AI with applications in COVID-19 imaging," *Med. Phys.*, vol. 49, no. 1, pp. 1–14, Jan. 2022, doi: 10.1002/mp.15359.
[36] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 285–286, Jan. 2005, doi: 10.1109/tnn.2004.842673.
[37] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, and Q. Miao, "Deep reinforcement learning: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 28, 2022, doi: 10.1109/TNNLS.2022.3207346.
[38] S. E. Li, *Reinforcement Learning for Sequential Decision and Optimal Control*. Singapore: Springer, 2023, doi: 10.1007/978-981-19-7784-8.
[39] H. Dong, Z. Ding, S. Zhang, *Deep Reinforcement Learning Fundamentals, Research and Applications: Fundamentals, Research and Applications*. New York, NY, USA: Springer Nature, 2020.
[40] C. Wan and M. Hwang, "Value-based deep reinforcement learning for adaptive isolated intersection signal control," *IET Intell. Transp. Syst.*, vol. 12, no. 9, pp. 1005–1010, Nov. 2018, doi: 10.1049/iet-its.2018.5170.
[41] C. Do, C. Gordillo, and W. Burgard, "Learning to pour using deep deterministic policy gradients," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 3074–3079, doi: 10.1109/IROS.2018.8593654.
[42] M. Kim, D.-K. Han, J.-H. Park, and J.-S. Kim, "Motion planning of robot manipulators for a smoother path using a twin delayed deep deterministic policy gradient with hindsight experience replay," *Appl. Sci.*, vol. 10, no. 2, p. 575, Jan. 2020, doi: 10.3390/app10020575.
[43] I. Nematollahi, E. Rosete-Beas, A. Röfer, T. Welschehold, A. Valada, and W. Burgard, "Robot skill adaptation via soft actor-critic Gaussian mixture models," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 8651–8657, doi: 10.1109/ICRA46639.2022.9811770.
[44] H. Zhang, F. Wang, J. Wang, and B. Cui, "Robot grasping method optimization using improved deep deterministic policy gradient algorithm of deep reinforcement learning," *Rev. Sci. Instrum.*, vol. 92, no. 2, Feb. 2021, Art. no. 025114, doi: 10.1063/5.0034101/369268.

[45] P. Shukla, M. Pegu, and G. C. Nandi, "Development of behavior based robot manipulation using actor-critic architecture," in *Proc. 8th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Aug. 2021, pp. 469–474, doi: 10.1109/SPIN52536.2021.9566102.
[46] T. Chen, J.-Q. Liu, H. Li, S.-R. Wang, and W.-J. Niu, "Robustness assessment of asynchronous advantage actor-critic based on dynamic skewness and sparseness computation: A parallel computing view," *J. Comput. Sci. Technol.*, vol. 36, no. 5, pp. 1002–1021, Oct. 2021, doi: 10.1007/S11390-021-1217-Z.
[47] F. Ye, X. Cheng, P. Wang, C.-Y. Chan, and J. Zhang, "Automated lane change strategy using proximal policy optimization-based deep reinforcement learning," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Oct. 2020, pp. 1746–1752, doi: 10.1109/IV47402.2020.9304668.
[48] T. Kurutach, I. Clavera, Y. Duan, A. Tamar, and P. Abbeel, "Model-ensemble trust-region policy optimization," 2018, *arXiv:1802.10592*.
[49] Z. Huang, W. Heng, and S. Zhou, "Learning to paint with model-based deep reinforcement learning," 2019, *arXiv:1903.04411*.
[50] M. Thabet, "Imagination-augmented deep reinforcement learning for robotic applications," A thesis, Dept. Doctor Philosophy, Univ. Manchesterr, Manchester, U.K., 2022.
[51] A. Nagabandi, G. Kahn, R. S. Fearing, and S. Levine, "Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 7559–7566, doi: 10.1109/ICRA.2018.8463189.
[52] M. Janner, J. Fu, M. Zhang, and S. Levine, "When to trust your model: Model-based policy optimization," in *Proc. Neural Inf. Process. Syst.*, vol. 32, Jun. 2019, pp. 12498–12509. [Online]. Available: https://papers.nips.cc/paper/9416-when-to-trust-your-model-model-based-policy-optimization.pdf
[53] G. Marcus, "Innateness, AlphaZero, and artificial intelligence," 2018, *arXiv:1801.05667*.
[54] K. Zhu and T. Zhang, "Deep reinforcement learning based mobile robot navigation: A review," *Tsinghua Sci. Technol.*, vol. 26, no. 5, pp. 674–691, Oct. 2021, doi: 10.26599/TST.2021.9010012.
[55] N. F. Bar, H. Yetis, and M. Karakose, "Deep reinforcement learning approach with adaptive reward system for robot navigation in dynamic environments," in *Interdisciplinary Research in Technology and Management*, Sep. 2021, pp. 349–355, doi: 10.1201/9781003202240-55.
[56] W. Du and S. Ding, "A survey on multi-agent deep reinforcement learning: from the perspective of challenges and applications," *Artif. Intell. Rev.*, vol. 54, no. 5, pp. 3215–3238, Jun. 2021, doi: 10.1007/S10462-020-09938-Y.
[57] S. Joshi, S. Kumra, and F. Sahin, "Robotic grasping using deep reinforcement learning," in *Proc. IEEE 16th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2020, pp. 1461–1466, doi: 10.1109/CASE48305.2020.9216986.
[58] M. Saeed, M. Nagdi, B. Rosman, and H. H. S. M. Ali, "Deep reinforcement learning for robotic hand manipulation," in *Proc. Int. Conf. Comput., Control, Electr., Electron. Eng. (ICCCEEE)*, Feb. 2021, pp. 1–5, doi: 10.1109/ICCCEEE49695.2021.9429619.
[59] X. Lin, Y. Wang, J. Olkin, and D. Held, "SoftGym: Benchmarking deep reinforcement learning for deformable object manipulation," in *Proc. PMLR*, Oct. 2021, pp. 432–448, Accessed: Jun. 2, 2023. [Online]. Available: https://proceedings.mlr.press/v155/lin21a.html
[60] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," 2017, *arXiv:1709.10087*.
[61] L. Manuelli, L. Li, P. Florence, and R. Tedrake, "Keypoints into the future: Self-supervised correspondence in model-based reinforcement learning," in *Proc. Conf. Robot Learn.*, Jan. 2020, pp. 693–710. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr2009.html#abs-2009-05085
[62] J. Bao, G. Zhang, Y. Peng, Z. Shao, and A. Song, "Learn multi-step object sorting tasks through deep reinforcement learning," *Robotica*, vol. 40, no. 11, pp. 3878–3894, Nov. 2022, doi: 10.1017/s0263574722000650.
[63] F. Li, Q. Jiang, S. Zhang, M. Wei, and R. Song, "Robot skill acquisition in assembly process using deep reinforcement learning," *Neurocomputing*, vol. 345, pp. 92–102, Jun. 2019, doi: 10.1016/j.neucom.2019.01.087.
[64] J. Matas, S. James, and A. J. Davison, "Sim-to-real reinforcement learning for deformable object manipulation," *PMLR*, pp. 734–743, Oct. 2018, Accessed: Jun. 2, 2023. https://proceedings.mlr.press/v87/matas18a.html
[65] A. Singh, L. Yang, C. Finn, and S. Levine, "End-to-end robotic reinforcement learning without reward engineering," *Robot., Sci. Syst.*, vol. 15, p. 73, Jun. 2019, doi: 10.15607/RSS.2019.XV.073.
[66] Ì. Elguea-Aguinaco, A. Serrano-Muñoz, D. Chrysostomou, I. Inziarte-Hidalgo, S. Bøgh, and N. Arana-Arexolaleiba, "A review on reinforcement learning for contact-rich robotic manipulation tasks," *Robot. Comput.-Integr. Manuf.*, vol. 81, Jun. 2023, Art. no. 102517, doi: 10.1016/j.rcim.2022.102517.
[67] A. Acuto, P. Barillà, L. Bozzolo, M. Conterno, M. Pavese, and A. Policicchio, "Variational quantum soft actor-critic for robotic arm control," 2022, *arXiv:2212.11681*.
[68] Z. Raziei and M. Moghaddam, "Adaptable automation with modular deep reinforcement learning and policy transfer," *Eng. Appl. Artif. Intell.*, vol. 103, Aug. 2021, Art. no. 104296, doi: 10.1016/j.engappai.2021.104296.
[69] B. Sangiovanni, G. P. Incremona, M. Piastra, and A. Ferrara, "Self-configuring robot path planning with obstacle avoidance via deep reinforcement learning," *IEEE Control Syst. Lett.*, vol. 5, no. 2, pp. 397–402, Apr. 2021, doi: 10.1109/LCSYS.2020.3002852.
[70] G. Schoettler, A. Nair, J. Luo, S. Bahl, J. Aparicio Ojea, E. Solowjow, and S. Levine, "Deep reinforcement learning for industrial insertion tasks with visual inputs and natural rewards," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 5548–5555, doi: 10.1109/IROS45743.2020.9341714.
[71] J. Luo, E. Solowjow, C. Wen, J. A. Ojea, and A. M. Agogino, "Deep reinforcement learning for robotic assembly of mixed deformable and rigid objects," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, Dec. 2018, pp. 2062–2069, doi: 10.1109/IROS.2018.8594353.
[72] W. Yuan, K. Hang, D. Kragic, M. Y. Wang, and J. A. Stork, "End-to-end nonprehensile rearrangement with deep reinforcement learning and simulation-to-reality transfer," *Robot. Auto. Syst.*, vol. 119, pp. 119–134, Sep. 2019, doi: 10.1016/j.robot.2019.06.007.
[73] A. Koenig, Z. Liu, L. Janson, and R. Howe, "The role of tactile sensing in learning and deploying grasp refinement algorithms," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2022, pp. 7766–7772, doi: 10.1109/IROS47612.2022.9981915.
[74] M. H. Sayour, S. E. Kozhaya, and S. S. Saab, "Autonomous robotic manipulation: Real-time, deep-learning approach for grasping of unknown objects," *J. Robot.*, vol. 2022, pp. 1–14, Jun. 2022, doi: 10.1155/2022/2585656.
[75] F. Ebert, C. Finn, S. Dasari, A. Xie, A. Lee, and S. Levine, "Visual foresight: Model-based deep reinforcement learning for vision-based robotic control," 2018, *arXiv:1812.00568*.
[76] K. Fang, Y. Zhu, A. Garg, A. Kurenkov, V. Mehta, L. Fei-Fei, and S. Savarese, "Learning task-oriented grasping for tool manipulation from simulated self-supervision," *Int. J. Robot. Res.*, vol. 39, nos. 2–3, pp. 202–216, Mar. 2020, doi: 10.1177/0278364919872545.
[77] J. Mahler. (2018). *Efficient Policy Learning for Robust Robot Grasping*. [Online]. Available: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-120.pdf
[78] S. Krishnan. (2018). *Hierarchical Deep Reinforcement Learning For Robotics and Data Science*. [Online]. Available: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-101.pdf
[79] P. Falco, A. Attawia, M. Saveriano, and D. Lee, "On policy learning robust to irreversible events: An application to robotic in-hand manipulation," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1482–1489, Jul. 2018, doi: 10.1109/LRA.2018.2800110.
[80] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, and S. Levine, "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Proc. PMLR*, Oct. 2018, pp. 651–673, Accessed: Jun. 20, 2023. [Online]. Available: https://proceedings.mlr.press/v87/kalashnikov18a.html
[81] T. Haarnoja, V. Pong, A. Zhou, M. Dalal, P. Abbeel, and S. Levine, "Composable deep reinforcement learning for robotic manipulation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 6244–6251, doi: 10.1109/ICRA.2018.8460756.
[82] V. Myers, A. He, K. Fang, H. Walke, P. Hansen-Estruch, C.-A. Cheng, M. Jalobeanu, A. Kolobov, A. Dragan, and S. Levine, "Goal representations for instruction following: A semi-supervised language interface to control," 2023, *arXiv:2307.00117v1*.
[83] K. Black, M. Janner, Y. Du, I. Kostrikov, and S. Levine, "Training diffusion models with reinforcement learning," 2023, *arXiv:2305.13301v2*.
[84] A. Sehgal, N. Ward, H. La, and S. Louis, "Automatic parameter optimization using genetic algorithm in deep reinforcement learning for robotic manipulation tasks," 2022, *arXiv:2204.03656v2*.

[85] J. Thumm and M. Althoff, "Provably safe deep reinforcement learning for robotic manipulation in human environments," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2022, pp. 6344–6350, doi: 10.1109/ICRA46639.2022.9811698.

[86] Y. Lin, A. Church, M. Yang, H. Li, J. Lloyd, D. Zhang, and N. F. Lepora, "Bi-touch: Bimanual tactile manipulation with sim-to-real deep reinforcement learning," 2023, *arXiv:2307.06423*.

[87] L. Fan, "SURREAL: Open-source reinforcement learning framework and robot manipulation benchmark," in *Proc. PMLR*, Oct. 2018, pp. 767–782, Accessed: Jun. 20, 2023. [Online]. Available: https://proceedings.mlr.press/v87/fan18a.html

[88] X. Zhu, F. Zhang, and H. Li, "Swarm deep reinforcement learning for robotic manipulation," *Proc. Comput. Sci.*, vol. 198, pp. 472–479, Jan. 2022, doi: 10.1016/j.procs.2021.12.272.

[89] X. Liu, G. Wang, Z. Liu, Y. Liu, Z. Liu, and P. Huang, "Hierarchical reinforcement learning integrating with human knowledge for practical robot skill learning in complex multi-stage manipulation," *IEEE Trans. Autom. Sci. Eng.*, early access, Jul. 17, 2004, doi: 10.1109/TASE.2023.3288037.

[90] M. Yang, Y. Lin, A. Church, J. Lloyd, D. Zhang, D. A. W. Barton, and N. F. Lepora, "Sim-to-real model-based and model-free deep reinforcement learning for tactile pushing," *IEEE Robot. Autom. Lett.*, vol. 8, no. 9, pp. 5480–5487, Sep. 2023, doi: 10.1109/LRA.2023.3295236.

[91] J. Yamada, "Motion planner augmented reinforcement learning for robot manipulation in obstructed environments," in *Proc. Conf. Robot. Learn*, 2021, pp. 589–603, Accessed: Jun. 20, 2023. [Online]. Available: https://proceedings.mlr.press/v155/yamada21a.html

[92] A. Sehgal, H. La, S. Louis, and H. Nguyen, "Deep reinforcement learning using genetic algorithm for parameter optimization," in *Proc. 3rd IEEE Int. Conf. Robotic Comput. (IRC)*, Feb. 2019, pp. 596–601, doi: 10.1109/IRC.2019.00121.

[93] H. Xiong, T. Ma, L. Zhang, and X. Diao, "Comparison of end-to-end and hybrid deep reinforcement learning strategies for controlling cable-driven parallel robots," *Neurocomputing*, vol. 377, pp. 73–84, Feb. 2020, doi: 10.1016/j.neucom.2019.10.020.

[94] B. Peng, T. Rashid, C. A. Schroeder de Witt, P.-A. Kamienny, P. H. S. Torr, W. Böhmer, and S. Whiteson, "FACMAC: Factored multi-agent centralised policy gradients," 2020, *arXiv:2003.06709*.

[95] A. Zhan, R. Zhao, L. Pinto, P. Abbeel, and M. Laskin, "A framework for efficient robotic manipulation," in *Proc. NeurIPS*, Dec. 2021, pp. 1–15.

[96] R. Jangir, G. Alenya, and C. Torras, "Dynamic cloth manipulation with deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 4630–4636, doi: 10.1109/ICRA40945.2020.9196659.

[97] Y. Hu and B. Si, "A reinforcement learning neural network for robotic manipulator control," *Neural Comput.*, vol. 30, no. 7, pp. 1983–2004, Jul. 2018, doi: 10.1162/neco_a_01079.

[98] R. Jeong, Y. Aytar, D. Khosid, Y. Zhou, J. Kay, T. Lampe, K. Bousmalis, and F. Nori, "Self-supervised sim-to-real adaptation for visual robotic manipulation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 2718–2724, doi: 10.1109/ICRA40945.2020.9197326.

[99] C. Wang, Q. Zhang, Q. Tian, S. Li, X. Wang, D. Lane, Y. Petillot, and S. Wang, "Learning mobile manipulation through deep reinforcement learning," *Sensors*, vol. 20, no. 3, p. 939, Feb. 2020, doi: 10.3390/s20030939.

[100] Y. Tsurumine, Y. Cui, E. Uchibe, and T. Matsubara, "Deep reinforcement learning with smooth policy update: Application to robotic cloth manipulation," *Robot. Auto. Syst.*, vol. 112, pp. 72–83, Feb. 2019, doi: 10.1016/j.robot.2018.11.004.

[101] B. Sangiovanni, A. Rendiniello, G. P. Incremona, A. Ferrara, and M. Piastra, "Deep reinforcement learning for collision avoidance of robotic manipulators," in *Proc. Eur. Control Conf. (ECC)*, Jun. 2018, pp. 2063–2068, doi: 10.23919/ECC.2018.8550363.

[102] A. Zeng, S. Song, S. Welker, J. Lee, A. Rodriguez, and T. Funkhouser, "Learning synergies between pushing and grasping with self-supervised deep reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 4238–4245, doi: 10.1109/IROS.2018.8593986.

[103] H. Zhu, A. Gupta, A. Rajeswaran, S. Levine, and V. Kumar, "Dexterous manipulation with deep reinforcement learning: Efficient, general, and low-cost," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 3651–3657, doi: 10.1109/ICRA.2019.8794102.

[104] Y. Hu, W. Wang, H. Liu, and L. Liu, "Reinforcement learning tracking control for robotic manipulator with kernel-based dynamic model," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3570–3578, Sep. 2020, doi: 10.1109/TNNLS.2019.2945019.

[105] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: A survey," in *Proc. IEEE Symp. Series Comput. Intell. (SSCI)*, 2020. pp. 737–744, doi: 10.1109/SSCI47803.2020.9308468.

[106] C. C. Beltran-Hernandez, D. Petit, I. G. Ramirez-Alpizar, and K. Harada, "Variable compliance control for robotic peg-in-hole assembly: A deep-reinforcement-learning approach," *Appl. Sci.*, vol. 10, no. 19, p. 6923, Oct. 2020, doi: 10.3390/app10196923.

[107] A. Mandlekar, F. Ramos, B. Boots, S. Savarese, L. Fei-Fei, A. Garg, and D. Fox, "IRIS: Implicit reinforcement without interaction at scale for learning control from offline robot manipulation data," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 4414–4420, doi: 10.1109/ICRA40945.2020.9196935.

[108] J. Xie, Z. Shao, Y. Li, Y. Guan, and J. Tan, "Deep reinforcement learning with optimized reward functions for robotic trajectory planning," *IEEE Access*, vol. 7, pp. 105669–105679, 2019, doi: 10.1109/ACCESS.2019.2932257.

[109] A. Wang, T. Kurutach, K. Liu, P. Abbeel, and A. Tamar, "Learning robotic manipulation through visual planning and acting," *Robot., Sci. Syst.*, pp. 74–86, May 2019, doi: 10.15607/RSS.2019.XV.074.

[110] R. Strudel, A. Pashevich, I. Kalevatykh, I. Laptev, J. Sivic, and C. Schmid, "Learning to combine primitive skills: A step towards versatile robotic manipulation," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2020, pp. 4637–4643, doi: 10.1109/ICRA40945.2020.9196619.

[111] X. Li, J. Zhong, and M. M. Kamruzzaman, "Complicated robot activity recognition by quality-aware deep reinforcement learning," *Future Gener. Comput. Syst.*, vol. 117, pp. 480–485, Apr. 2021, doi: 10.1016/j.future.2020.11.017.

[112] A. Malik, Y. Lischuk, T. Henderson, and R. Prazenica, "A deep reinforcement-learning approach for inverse kinematics solution of a high degree of freedom robotic manipulator," *Robotics*, vol. 11, no. 2, p. 44, Apr. 2022, doi: 10.3390/robotics11020044.

[113] H. Oliff, Y. Liu, M. Kumar, M. Williams, and M. Ryan, "Reinforcement learning for facilitating human-robot-interaction in manufacturing," *J. Manuf. Syst.*, vol. 56, pp. 326–340, Jul. 2020, doi: 10.1016/j.jmsy.2020.06.018.

[114] S. Christen, S. Stevšic, and O. Hilliges, "Demonstration-guided deep reinforcement learning of control policies for dexterous human-robot interaction," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 2161–2167, doi: 10.1109/ICRA.2019.8794065.

[115] M. El-Shamouty, X. Wu, S. Yang, M. Albus, and M. F. Huber, "Towards safe human–robot collaboration using deep reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 4899–4905, doi: 10.1109/ICRA40945.2020.9196924.

[116] C. Li, P. Zheng, Y. Yin, Y. M. Pang, and S. Huo, "An AR-assisted deep reinforcement learning-based approach towards mutual-cognitive safe human–robot interaction," *Robot. Comput.-Integr. Manuf.*, vol. 80, Apr. 2023, Art. no. 102471, doi: 10.1016/j.rcim.2022.102471.

[117] M. B. Imtiaz, Y. Qiao, and B. Lee, "Prehensile and non-prehensile robotic pick-and-place of objects in clutter using deep reinforcement learning," *Sensors*, vol. 23, no. 3, p. 1513, Jan. 2023, doi: 10.3390/s23031513.

[118] R. Dershan, A. M. Soufi Enayati, Z. Zhang, D. Richert, and H. Najjaran, "Facilitating sim-to-real by intrinsic stochasticity of real-time simulation in reinforcement learning for robot manipulation," 2023, *arXiv:2304.06056*.

[119] N. G. Lopez, Y. L. E. Nuin, E. B. Moral, L. U. S. Juan, A. S. Rueda, V. M. Vilches, and R. Kojcev, "Gym-gazebo2, a toolkit for reinforcement learning using ROS 2 and gazebo," 2019, *arXiv:1903.06278*.

[120] S. N. Aslan, B. Taşçi, A. Uçar, and C. Güzeliş, "Learning to move an object by the humanoid robots by using deep reinforcement learning," in *Proc. Intell. Environments Workshop 17th Int. Conf. Intell. Environments*, vol. 29, Jul. 2021, pp. 143–155, doi: 10.3233/AISE210092.

[121] C. Chen, H.-Y. Li, X. Zhang, X. Liu, and U.-X. Tan, "Towards robotic picking of targets with background distractors using deep reinforcement learning," in *Proc. WRC Symp. Adv. Robot. Autom. (WRC SARA)*, Aug. 2019, pp. 166–171, doi: 10.1109/WRC-SARA.2019.8931932.

[122] X. Xie, C. Li, C. Zhang, Y. Zhu, and S.-C. Zhu, "Learning virtual grasp with failed demonstrations via Bayesian inverse reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 1812–1817, doi: 10.1109/IROS40897.2019.8968063.
[123] J. Sun, L. Yu, P. Dong, B. Lu, and B. Zhou, "Adversarial inverse reinforcement learning with self-attention dynamics model," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 1880–1886, Apr. 2021, doi: 10.1109/LRA.2021.3061397.
[124] S. Krishnan, A. Garg, R. Liaw, B. Thananjeyan, L. Miller, F. T. Pokorny, and K. Goldberg, "SWIRL: A sequential windowed inverse reinforcement learning algorithm for robot tasks with delayed rewards," *Int. J. Robot. Res.*, vol. 38, nos. 2–3, pp. 126–145, Mar. 2019, doi: 10.1177/0278364918784350.
[125] S. Kumar, J. Zamora, N. Hansen, R. Jangir, and X. Wang, "Graph inverse reinforcement learning from diverse videos," in *Proc. PMLR*, Mar. 2023, pp. 55–66, Accessed: Jun. 4, 2023. [Online]. Available: https://proceedings.mlr.press/v205/kumar23a.html
[126] A. Gleave and O. Habryka, "Multi-task maximum entropy inverse reinforcement learning," 2018, *arXiv:1805.08882*.
[127] I. Batzianoulis, F. Iwane, S. Wei, C. G. P. R. Correia, R. Chavarriaga, J. D. R. Millán, and A. Billard, "Customizing skills for assistive robotic manipulators, an inverse reinforcement learning approach with error-related potentials," *Commun. Biol.*, vol. 4, no. 1, p. 1406, Dec. 2021, doi: 10.1038/s42003-021-02891-8.
[128] Y. Ma, D. Xu, and F. Qin, "Efficient insertion control for precision assembly based on demonstration learning and reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4492–4502, Jul. 2021, doi: 10.1109/TII.2020.3020065.
[129] X. Zhang, L. Sun, Z. Kuang, and M. Tomizuka, "Learning variable impedance control via inverse reinforcement learning for force-related tasks," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2225–2232, Apr. 2021, doi: 10.1109/LRA.2021.3061374.
[130] M. Hamaya, F. von Drigalski, T. Matsubara, K. Tanaka, R. Lee, C. Nakashima, Y. Shibata, and Y. Ijiri, "Learning soft robotic assembly strategies from successful and failed demonstrations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 8309–8315, doi: 10.1109/IROS45743.2020.9341504.
[131] D. S. Brown, W. Goo, and S. Niekum, "Better-than-demonstrator imitation learning via automatically-ranked demonstrations," in *Proc. PMLR*, May 2020, pp. 330–359, Accessed: Jun. 22, 2023. [Online]. Available: https://proceedings.mlr.press/v100/brown20a.html
[132] S. Arora, P. Doshi, and B. Banerjee, "Online inverse reinforcement learning with learned observation model," in *Proc. PMLR*, Mar. 2023, pp. 1468–1477, Accessed: Jun. 4, 2023. [Online]. Available: https://proceedings.mlr.press/v205/arora23a.html
[133] K. Nishi and M. Shimosaka, "Fine-grained driving behavior prediction via context-aware multi-task inverse reinforcement learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 2281–2287, doi: 10.1109/ICRA40945.2020.9197126.
[134] S.-W. Yoo and S.-W. Seo, "Learning multi-task transferable rewards via variational inverse reinforcement learning," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 434–440, doi: 10.1109/ICRA46639.2022.9811697.
[135] S. Arora, P. Doshi, and B. Banerjee, "Min-max entropy inverse RL of multiple tasks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 12639–12645, doi: 10.1109/ICRA48506.2021.9561771.
[136] K. Kobayashi, T. Horii, R. Iwaki, Y. Nagai, and M. Asada, "Situated GAIL: Multitask imitation using task-conditioned adversarial inverse reinforcement learning," 2019, *arXiv:1911.00238*.
[137] K. Hausman, Y. Chebotar, S. Schaal, G. Sukhatme, and J. J. Lim, "Multi-modal imitation learning from unstructured demonstrations using generative adversarial nets," in *Proc. Adv Neural Inf. Process Syst.*, May 2017, pp. 1236–1246, Accessed: Jun. 5, 2023.
[138] S. Piao, Y. Huang, and H. Liu, "Online multi-modal imitation learning via lifelong intention encoding," in *Proc. IEEE 4th Int. Conf. Adv. Robot. Mechatronics (ICARM)*, Jul. 2019, pp. 786–792, doi: 10.1109/ICARM.2019.8833960.
[139] R. H. Kaiser, M. T. Treadway, D. W. Wooten, P. Kumar, F. Goer, L. Murray, M. Beltzer, P. Pechtel, A. Whitton, A. L. Cohen, N. M. Alpert, G. El Fakhri, M. D. Normandin, and D. A. Pizzagalli, "Frontostriatal and dopamine markers of individual differences in reinforcement learning: A multi-modal investigation," *Cerebral Cortex*, vol. 28, no. 12, pp. 4281–4290, Dec. 2018, doi: 10.1093/cercor/bhx281.
[140] K. Stanojević, S. P. Samuel, K. Advisor, and P.-A. Murena. (2021). *Non-Sequential Bayesian Multi-Modal Inverse Reinforcement Learning*, Accessed: Jul. 2, 2023. [Online]. Available: https://aaltodoc.aalto.fi:443/handle/123456789/111789
[141] D. S. Brown, Y. Cui, and S. Niekum, "Risk-aware active inverse reinforcement learning," in *Proc. PMLR*, Oct. 2018, pp. 362–372, Accessed: Jun. 5, 2023. [Online]. Available: https://proceedings.mlr.press/v87/brown18a.html
[142] M. Lopes, F. Melo, and L. Montesano, "Active learning for reward estimation in inverse reinforcement learning," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5782, 2009, pp. 31–46, doi: 10.1007/978-3-642-04174-7_3.
[143] F. Memarian, Z. Xu, B. Wu, M. Wen, and U. Topcu, "Active task-inference-guided deep inverse reinforcement learning," in *Proc. 59th IEEE Conf. Decis. Control (CDC)*, Dec. 2020, pp. 1932–1938, doi: 10.1109/CDC42340.2020.9304190.
[144] D. Lindner, A. Krause, and G. Ramponi, "Active exploration for inverse reinforcement learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, Dec. 2022, pp. 5843–5853, Accessed: Jul. 2, 2023. [Online]. Available: https://github.com/lasgroup/aceirl
[145] M. Fahad, Z. Chen, and Y. Guo, "Learning how pedestrians navigate: A deep inverse reinforcement learning approach," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 819–826, doi: 10.1109/IROS.2018.8593438.
[146] D. Mukherjee, K. Gupta, L. H. Chang, and H. Najjaran, "A survey of robot learning strategies for human-robot collaboration in industrial settings," *Robot. Comput.-Integr. Manuf.*, vol. 73, Feb. 2022, Art. no. 102231, doi: 10.1016/j.rcim.2021.102231.
[147] W. Wang, R. Li, Y. Chen, Z. M. Diekel, and Y. Jia, "Facilitating human–robot collaborative tasks by teaching-learning-collaboration from human demonstrations," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 2, pp. 640–653, Apr. 2019, doi: 10.1109/TASE.2018.2840345.
[148] B. Woodworth, F. Ferrari, T. E. Zosa, and L. D. Riek, "Preference learning in assistive robotics: Observational repeated inverse reinforcement learning," in *Proc. Mach. Learn. Res.*, vol. 85, Nov. 2018, pp. 420–439, Accessed: May 21, 2023. [Online]. Available: https://proceedings.mlr.press/v85/woodworth18a.html
[149] M. Kollmitz, T. Koller, J. Boedecker, and W. Burgard, "Learning human-aware robot navigation from physical interaction via inverse reinforcement learning," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, Oct. 2020, pp. 11025–11031, doi: 10.1109/IROS45743.2020.9340865.
[150] N. Das, S. Bechtle, T. Davchev, D. Jayaraman, A. Rai, and F. Meier, "Model-based inverse reinforcement learning from visual demonstrations," in *Proc. PMLR*, Oct. 2021, pp. 1930–1942, Accessed: Jun. 4, 2023. [Online]. Available: https://proceedings.mlr.press/v155/das21a.html
[151] W. Xue, P. Kolaric, J. Fan, B. Lian, T. Chai, and F. L. Lewis, "Inverse reinforcement learning in tracking control based on inverse optimal control," *IEEE Trans. Cybern.*, vol. 52, no. 10, pp. 10570–10581, Oct. 2022, doi: 10.1109/TCYB.2021.3062856.
[152] E. B. Hansen, R. E. Andersen, S. Madsen, and S. Bøgh, "Transferring human manipulation knowledge to robots with inverse reinforcement learning," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2020, pp. 933–937, doi: 10.1109/SII46433.2020.9025873.
[153] O. M. Manyar, Z. McNulty, S. Nikolaidis, and S. K. Gupta, "Inverse reinforcement learning framework for transferring task sequencing policies from humans to robots in manufacturing applications," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2023, pp. 849–856, doi: 10.1109/ICRA48891.2023.10160687.
[154] W. Luo, J. Zhang, P. Feng, D. Yu, and Z. Wu, "A deep transfer-learning-based dynamic reinforcement learning for intelligent tightening system," *Int. J. Intell. Syst.*, vol. 36, no. 3, pp. 1345–1365, Mar. 2021, doi: 10.1002/int.22345.
[155] Q. Wang, F. R. Sanchez, R. McCarthy, D. C. Bulens, K. McGuinness, N. O'Connor, M. Wüthrich, F. Widmaier, S. Bauer, and S. J. Redmond, "Dexterous robotic manipulation using deep reinforcement learning and knowledge transfer for complex sparse reward-based tasks," *Expert Syst.*, vol. 40, no. 6, p. 13205, Nov. 2022, doi: 10.1111/exsy.13205.
[156] A. Bobu, M. Wiggert, C. Tomlin, and A. D. Dragan, "Feature expansive reward learning: Rethinking human input," in *Proc. 16th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2021, pp. 216–224, doi: 10.1145/3434073.3444667.

[157] A. Bobu, M. Wiggert, C. Tomlin, and A. D. Dragan, "Inducing structure in reward learning by learning features," *Int. J. Robot. Res.*, vol. 41, no. 5, pp. 497–518, Apr. 2022, doi: 10.1177/02783649221078031.

[158] D. S. Brown, W. Goo, P. Nagarajan, and S. Niekum, "Extrapolating beyond suboptimal demonstrations via inverse reinforcement learning from observations," in *Proc. PMLR*, May 2019, pp. 783–792, Accessed: Jun. 4, 2023. [Online]. Available: https://proceedings.mlr.press/v97/brown19a.html

[159] Z. Wu, L. Sun, W. Zhan, C. Yang, and M. Tomizuka, "Efficient sampling-based maximum entropy inverse reinforcement learning with application to autonomous driving," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5355–5362, Oct. 2020, doi: 10.1109/LRA.2020.3005126.

[160] C. You, J. Lu, D. Filev, and P. Tsiotras, "Advanced planning for autonomous vehicles using reinforcement learning and deep inverse reinforcement learning," *Robot. Auto. Syst.*, vol. 114, pp. 1–18, Apr. 2019, doi: 10.1016/j.robot.2019.01.003.

[161] A. Tucker, A. Gleave, and S. Russell, "Inverse reinforcement learning for video games," 2018, *arXiv:1810.10593*.

[162] N. Yu, L. Nan, and T. Ku, "Robot hand-eye cooperation based on improved inverse reinforcement learning," *Ind. Robot: Int. J. Robot. Res. Appl.*, vol. 49, no. 5, pp. 877–884, Jun. 2022, doi: 10.1108/ir-09-2021-0208.

[163] L. Yu, T. Yu, C. Finn, and S. Ermon, "Meta-inverse reinforcement learning with probabilistic context variables," 2019, *arXiv:1909.09314*.

[164] J. Chen, T. Lan, and V. Aggarwal, "Option-aware adversarial inverse reinforcement learning for robotic control," 2022, *arXiv:2210.01969*.

[165] F. Xie, A. Chowdhury, M. C. De Paolis Kaluza, L. Zhao, L. L. S. Wong, and R. Yu, "Deep imitation learning for bimanual robotic manipulation," 2020, *arXiv:2010.05134v2*.

[166] A. T. Le, M. Guo, N. v. Duijkeren, L. Rozo, R. Krug, A. G. Kupcsik, and M. Bürger, "Learning forceful manipulation skills from multi-modal human demonstrations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2021, pp. 7770–7777, doi: 10.1109/IROS51168.2021.9636828.

[167] J. Luo, O. Sushkov, R. Pevceviciute, W. Lian, C. Su, M. Vecerik, N. Ye, S. Schaal, and J. Scholz "Robust multi-modal policies for industrial assembly via reinforcement learning and demonstrations: A large-scale study," *Robot., Sci. Syst.*, vol. 17, pp. 88–97, Mar. 2021, doi: 10.15607/RSS.2021.XVII.088.

[168] A. Bighashdel, P. Meletis, P. Jancura, and G. Dubbelman, "Deep adaptive multi-intention inverse reinforcement learning," in *Proc. Mach. Learn. Knowl. Discovery Databases. Res. Track*, in Lecture Notes in Computer Science, 2021, pp. 206–221, doi: 10.1007/978-3-030-86486-6_13.

[169] M. Imani and S. F. Ghoreishi, "Scalable inverse reinforcement learning through multifidelity Bayesian optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 4125–4132, Aug. 2022, doi: 10.1109/TNNLS.2021.3051012.

[170] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Mach. Intell.*, vol. 1, no. 9, pp. 389–399, Sep. 2019, doi: 10.1038/s42256-019-0088-2.

[171] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, and W. Samek, "Explainable AI methods—A brief overview," in *XxAI—Beyond Explainable AI* (Lecture Notes in Computer Science), 2022, pp. 13–38, doi: 10.1007/978-3-031-04083-2_2.

[172] A. Ghosh and D. Kandasamy, "Interpretable artificial intelligence: Why and when," *Amer. J. Roentgenology*, vol. 214, no. 5, pp. 1137–1138, May 2020, doi: 10.2214/ajr.19.22145.

[173] M. T. Mason, "Toward robotic manipulation," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 1, no. 1, pp. 1–28, May 2018, doi: 10.1146/annurev-control-060117-104848.

[174] A. Akundi, D. Euresti, S. Luna, W. Ankobiah, A. Lopes, and I. Edinbarough, "State of industry 5.0—Analysis and identification of current research trends," *Appl. Syst. Innov.*, vol. 5, no. 1, p. 27, Feb. 2022, doi: 10.3390/asi5010027.

[175] X. Li, Z. Serlin, G. Yang, and C. Belta, "A formal methods approach to interpretable reinforcement learning for robotic planning," *Sci. Robot.*, vol. 4, no. 37, Dec. 2019, Art. no. aay6276, doi: 10.1126/scirobotics.aay6276.

[176] M. Z. Naser, "An engineer's guide to eXplainable artificial intelligence and interpretable machine learning: Navigating causality, forced goodness, and the false perception of inference," *Autom. Construction*, vol. 129, Sep. 2021, Art. no. 103821, doi: 10.1016/j.autcon.2021.103821.

[177] S. M. Mizanoor Rahman, "Trustworthy power assistance in object manipulation with a power assist robotic system," in *Proc. SoutheastCon*, Apr. 2019, pp. 1–8, doi: 10.1109/southeastcon42311.2019.9020523.

[178] A. Heuillet, F. Couthouis, and N. Díaz-Rodríguez, "Explainability in deep reinforcement learning," *Knowledge-Based Syst.*, vol. 214, Feb. 2021, Art. no. 106685, doi: 10.1016/j.knosys.2020.106685.

[179] A. Salehi and S. Doncieux, "Data-efficient, explainable and safe box manipulation: Illustrating the advantages of physical priors in model-predictive control," 2023, *arXiv:2303.01563*.

[180] A. Pore, D. Corsi, E. Marchesini, D. Dall'Alba, A. Casals, A. Farinelli, and P. Fiorini, "Safe reinforcement learning using formal verification for tissue retraction in autonomous robotic-assisted surgery," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2021, pp. 4025–4031, doi: 10.1109/IROS51168.2021.9636175.

[181] D. Corsi, E. Marchesini, A. Farinelli, and P. Fiorini, "Formal verification for safe deep reinforcement learning in trajectory generation," in *Proc. 4th IEEE Int. Conf. Robotic Comput. (IRC)*, Nov. 2020, pp. 352–359, doi: 10.1109/IRC.2020.00062.

[182] K. You, C. Zhou, and L. Ding, "Deep learning technology for construction machinery and robotics," *Autom. Construction*, vol. 150, Jun. 2023, Art. no. 104852, doi: 10.1016/j.autcon.2023.104852.

[183] W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu, "Definitions, methods, and applications in interpretable machine learning," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 44, pp. 22071–22080, Oct. 2019, doi: 10.1073/pnas.1900654116.

[184] D. Corsi, R. Yerushalmi, G. Amir, A. Farinelli, D. Harel, and G. Katz, "Constrained reinforcement learning for robotics via scenario-based programming," 2022, *arXiv:2206.09603*.

[185] B. Beyret, A. Shafti, and A. A. Faisal, "Dot-to-dot: Explainable hierarchical reinforcement learning for robotic manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 5014–5019, doi: 10.1109/IROS40897.2019.8968488.

[186] S. B. Remman and A. M. Lekkas, "Robotic lever manipulation using hindsight experience replay and Shapley additive explanations," in *Proc. Eur. Control Conf. (ECC)*, Jun. 2021, pp. 586–593, doi: 10.23919/ECC54610.2021.9654850.

[187] S. A. Khader, H. Yin, P. Falco, and D. Kragic, "Learning deep energy shaping policies for stability-guaranteed manipulation," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 8583–8590, Oct. 2021, doi: 10.1109/LRA.2021.3111962.

[188] T. Hickling, A. Zenati, N. Aouf, and P. Spencer, "Explainability in deep reinforcement learning, a review into current methods and applications," 2022, *arXiv:2207.01911*.

[189] S. B. Remman, I. Strümke, and A. M. Lekkas, "Causal versus marginal Shapley values for robotic lever manipulation controlled using deep reinforcement learning," in *Proc. Amer. Control Conf. (ACC)*, Jun. 2022, pp. 2683–2690, doi: 10.23919/ACC53348.2022.9867807.

[190] (2018). *Number of Robotic Manipulation Studies Using DRL in Google Scholar From 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://scholar.google.com/scholar?q=%22deep+reinforcement+learning%22+and+%22robot+manipulation%22&hl=tr&as_sdt=0%2C5&as_ylo=2018&as_yhi=2023

[191] *Number of Robotic Manipulation Studies Using DRL in Web of Science from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.webofscience.com/wos/woscc/summary/9e95bb47-f6ad-42ef-b4a0-53c5734c42e7-cf8587d7/relevance/1

[192] *Number of Robotic Manipulation Studies Using IRL in Google Scholar from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://scholar.google.com/scholar?q=%22inverse+reinforcement+learning%22+and+%22robot+manipulation%22&hl=tr&as_sdt=0%2C5&as_ylo=2018&as_yhi=2023

[193] *Number of Robotic Manipulation Studies Using DRL in Web of Science from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.webofscience.com/wos/woscc/summary/7b18565c-d2ea-4604-bef1-f04062fb5054-cf8a6 ff3/relevance/1

[194] *Number of Robotic Manipulation Studies Using Explainable AI in Google Scholar from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://scholar.google.com/scholar?q=%22explainable+ai%22+and+%22robot+manipulation%22&hl=tr&as_sdt=0%2C5&as_ylo=2018&as_yhi=2023

[195] *Number of Robotic Manipulation Studies Using Explainable AI in Web of Science from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.webofscience.com/wos/woscc/summary/3df660b8-412a-4abf-bf8b-f0da5551d751-cf8 a9774/relevance/1

[196] *Number of Robotic Manipulation Studies Using Trustworthy AI in Google Scholar from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://scholar.google.com/scholar?q=%22trustworthy+ai%22+and+%22robot+manipulation%22&hl=tr&as_sdt=0%2C5&as_ylo=2018&as_yhi=2023

[197] *Number of Robotic Manipulation Studies Using Trustworthy AI in Web of Science from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.webofscience.com/wos/woscc/summary/a8e5f037-8538-4f04-ad69-174f5dcac5d6-cf8aa795/relevance/1

[198] *Number of Robotic Manipulation Studies Using Interpretable AI in Google Scholar from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://scholar.google.com/scholar?q=%22interpretable+ai%22+%22robot+manipulation%22&hl=tr&as_sdt=0%2C5&as_ylo=2018&as_yhi=2023

[199] *Number of Robotic Manipulation Studies Using Interpretable AI in Web of Science from 2018 to 2023*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.webofscience.com/wos/woscc/summary/4bcdc18c-d624-4556-bc6e-8155875f10eb-cf8 acb75/relevance/1

RECEP OZALP received the B.S. and M.S. degrees from the Mechatronics Engineering Department, Firat University, Turkey, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree in education. He has been a Lecturer with the Baskil Vocational School, Firat University, since 2019. His research interests include humanoid robots and artificial intelligence, engineering applications, and robotic vision.

AYSEGUL UCAR (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the Department of Electrical and Electronics Engineering, Firat University, Turkey, in 1998, 2000, and 2006, respectively. In 2013, she was a Visiting Professor with the Division of Computer Science and Engineering, Louisiana State University, USA. Since 2020, she has been a Professor with the Department of Mechatronics Engineering, Firat University. She has more than 24 years of background in autonomous technologies and artificial intelligence, its engineering applications, robotics vision, teaching, and research. She is active in several professional bodies, particularly as a European Artificial Intelligence Alliance Committee Member and an Associate Editor of IEEE Access.

CUNEYT GUZELIS received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from İstanbul Technical University, İstanbul, Turkey, in 1981, 1984, and 1988, respectively. He was a Visiting Researcher and a Lecturer with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA, USA, from 1989 to 1991. He was a full-time Faculty Member with Istanbul Technical University, from 1991 to 2000, where he became a Full Professor, in 1998. He was a Professor of electrical and electronics engineering with Dokuz Eylül University, İzmir, Turkey, from 2000 to 2011, where he was the Dean of the Faculty of Engineering and İzmir University of Economics, İzmir, from 2011 to 2015, where he was the Director of the Graduate School of Natural and Applied Science. Since 2015, he has been a Professor of electrical and electronics engineering with Yaşar University, İzmir, where he was the Director of the Graduate School. He has supervised 17 M.S. and 14 Ph.D. students and published over 50 SCI-indexed journal articles, six peer-reviewed book chapters, and more than 80 peer-reviewed conference papers. He has participated in over 20 scientific research projects funded by national and international institutions, such as the British Council and the French National Council for Scientific Research. His research interests include artificial neural networks, biomedical signal and image processing, nonlinear circuits-systems and control, and educational systems.