Multi-Agent Reinforcement Learning Method for Disassembly Sequential Task Optimization Based on Human–Robot Collaborative Disassembly in Electric Vehicle Battery Recycling

Jinhua Xiao[1]
School of Transportation and Logistics Engineering,
Wuhan University of Technology,
Wuhan 430063, China
e-mail: xiaojinh@whut.edu.cn

Jiaxu Gao
School of Transportation and Logistics Engineering,
Wuhan University of Technology,
Wuhan 430063, China
e-mail: gaojiaxu@whut.edu.cn

Nabil Anwer
École Normale Supérieure Paris Saclay, LURPA,
Gif-sur-Yvette F-91190, France
e-mail: nabil.anwer@ens-paris-saclay.fr

Benoit Eynard
Department of Mechanical Engineering,
University of Technology of Compiè,
Roberval Laboratory, CS 60319,
Compiègne Cedex 60203, France
e-mail: benoit.eynard@utc.fr

With the wide application of new Electric Vehicle (EV) batteries in various industrial fields, it is important to establish a systematic intelligent battery recycling system that can be used to find out the resource wastes and environmental impacts of the retired EV battery. By combining the uncertain and dynamic disassembly and echelon utilization of EV battery recycling in the remanufacturing fields, human–robot collaboration (HRC) disassembly method can be used to solve huge challenges about the efficiency of retired EV battery recycling. In order to find out the disassembly task planning based on HRC disassembly process for retired EV battery recycling, a dynamic disassembly sequential task optimization method algorithm is proposed by Multi-Agent Reinforcement Learning (MARL). Furthermore, it is necessary to disassemble the retired EV battery disassembly trajectory based on the HRC disassembly task in 2D planar, which can be used to acquire the optimal disassembly paths in the same disassembly planar by combining the Q-learning algorithm. The disassembly task sequence can be completed through standard trajectory matching. Finally, the feasibility of the proposed method is verified by disassembly operations for a specific battery module case. [DOI: 10.1115/1.4062235]

Keywords: electric vehicle battery, human–robot collaboration, disassembly operation, multi-agent reinforcement learning, disassembly task optimization, sustainable manufacturing

1 Introduction

With the wide application of electric vehicle (EV) batteries in various industrial fields, highly intelligent EV battery recycling is gradually being urgent to alleviate the resource waste and environmental impacts, which will promote the automatic manufacturing of EV battery disassembly technology [1]. Due to the complexity and uncertainty of EV battery disassembly, it is necessary to develop an automatic disassembly process based on robot manufacturing. However, the main gap for the automatic EV battery disassembly is to find a better way to balance the automatic disassembly production line and human involvement by optimizing the low cost and disassembly capacity. However, the optimal disassembly sequence of human–robot collaboration (HRC) can provide the disassembly orders of retired EV battery structure to save the overall disassembly operation time than other orders. The efficiency of disassembly operations will be improved through the optimal disassembly sequence to assign the disassembly task to humans or robots for the specific disassembly operations. Actually, optimal disassembly task planning can reduce the switch of disassembly tools and the unnecessary disassembly operations that might cause the safety problems, such as collisions, mistaking behaviors, etc. Therefore, it becomes increasingly popular to combine HRC manufacturing to improve the efficiency and stability of disassembly and recycling in the flexible production and manufacturing process. HRC disassembly for EV batteries still has many key issues to realize the industrial manufacturing automation to make the machining robots better interact with human intentions that can be used to cooperate with human behaviors to complete the highly efficient disassembly operation and manufacturing tasks. The HRC disassembly needs to find out the disassembly operator switching from human to robot or robot to human, which can better complete the complex disassembly tasks to ensure the collaboration of humans and robots with more interaction and flexibility [2]. With the development of more advanced sensing technology and interaction technology, human and robot disassembly has achieved interactive and collaborative manufacturing, which not only reduces the labor cost for workers but also expands the efficiency of flexible disassembly operations [3].

With the requirements of green energy technology, EV battery recycling for EV batteries is becoming increasingly a huge challenge. If EV batteries are not effectively recycled, they might cause environmental pollutions and even threaten the health and safety of human beings. Therefore, it is important to establish a scientific and complete automated EV battery recycling and disassembly process [4]. Because there are still certain shortcomings for robots to autonomously complete tedious disassembly tasks in a relatively complex working environment, an efficient and safe HRC disassembly work can effectively solve the problems in the actual disassembly process for EV battery recycling. The disassembly tasks for EV battery recycling are repetitive and labor-intensive, and the robot can significantly improve work efficiency and reduce labor costs to assist in the accomplishment of disassembly tasks [5].

Compared with the traditional mode of EV battery recycling, the new disassembly method has the characteristics of a high dangerous degree based on HRC workings with high recycling requirements.

[1]Corresponding author.

Manuscript received November 16, 2022; final manuscript received March 27, 2023; published online July 21, 2023. Assoc. Editor: Pai Zheng.

However, a highly flexible remanufacturing disassembly production line needs to be executed in the actual disassembly process to acquire the desired disassembling sub-products in a single disassembly line [6]. The HRC disassembly operations need to accomplish the allocation of disassembly tasks that considers the optimization decision-making between human and robot. However, a simple disassembly process is not suitable enough to adapt to the uncertain disassembly objects. We know that the EV battery structure is relatively complex and it is difficult to unify the different battery types from various enterprises for their various layouts and modules [7]. Although the disassembly operation based on robot-assisted manufacturing can better cope with the uncertain and flexible battery products, human operations mainly focus on the intelligent decision-making of disassembly tasks that will promote the robot to accomplish a stable disassembly efficiency of EV battery. If there are no involvements about humans in the disassembly process, it is difficult to complete the complex disassembly tasks independently. By combining humans and robots in the specific disassembly process, high-efficient disassembly tasks can be better accomplished to establish reasonable disassembly sequences and disassembly task assignments.

The paper mainly proposes multi-agent reinforcement learning for HRC disassembly optimization in retired EV battery recycling. In Sec. 2, literature reviews will be presented to compare the various optimization methods and disassembly technology. The optimization of the disassembly process based on reinforcement learning will be proposed to discuss the optimal disassembly sequence and disassembly tasks in Sec. 3. Section 4 gives a case study to demonstrate the disassembly optimization process for certain an EV battery module. We will discuss the future challenges and possible research points for EV battery disassembly in Sec. 5. Section 6 gives a conclusion.

2 Related Work

2.1 Electric Vehicle Battery Disassembly Structure. By reviewing many related literature studies, most automotive EV batteries are composed of the following parts: battery housing, battery module, cooling system, BMU, and electronic devices such as the wiring harness. The structure of the battery module is further divided into module shell, insulation board, battery cell, and control circuit board and related wiring harness. The battery housing accounts for about 20–30% of the total weight of the battery, and the material is mostly aluminum alloy [8]. As known, the retired EV battery recycling needs to consider more possible applications for echelon utilization and disassembly [9]. However, there are many evaluation methods to judge the specific application scenarios by analyzing the specific characteristics of the specific EV battery pack. The entire battery pack needs to be discharged to a safe voltage before it can be disassembled, and the entire disassembly process can be divided into several large stages: disassembly of the battery package, disassembly of the battery module, and final disassembly of the battery cell. In the disassembly of the whole package of batteries, the current technical means are still disassembled by simple HRC or pure manual ways with inefficient and labor-intensive operations [10]. In order to facilitate the subsequent disassembly work, the operator first needs to remove the upper cover of the battery shell, which usually has two connection methods between the upper cover plate and the housing, the fastener connection and the sealant connection.

According to the complex product structure and uncertain product types, the size and position of the connection in the disassembly process are different to remove different fastener positions. Therefore, the design needs to optimize the disassembly sequence and to conform to the actual operation of the disassembly process. However, there are many engineering applications in visual identification parts and removal of product connections, which can be well solved in dealing with some traditional problems, and the connection position recognition of the cover plate on the battery shell is a relatively simple situation based on machine vision recognition [11]. After the removal of the battery cover, the disassembly of the wiring harness and electronic devices and the entire cooling system is relatively complex, and the connection between various battery modules and the battery shell is also important to the disassembly task. The connection method used in the battery module is basically connected by fasteners, many designers will apply fixing glue at the bottom of the module [12]. The traditional method is still relying on manpower to use special tools to violently remove the connection glue, which may cause the module to be damaged with greater safety hazards and low efficiency [13]. Because the wiring harness and cooling system in the battery pack have a certain degree-of-freedom, HRC disassembly operation requires a human to give necessary positioning assistance. The disassembly of the battery module requires the removal of the upper end cover and the internal wiring harness, and the connection removal of the module. However, the structure of the module for EV battery disassembly is more suitable for HRC disassembly operations to better accomplish disassembly tasks [14].

2.2 Human–Robot Collaboration Disassembly Methods. In the past research on HRC manufacturing, human–robot interactions are also becoming more and more abundant. However, many literature studies suggest that the repetitive operation task of the robot is easier and more accurate than that of the human action. Many scholars mentioned that the HRC methods to solve specific manufacturing problems toward assembly or disassembly, which can support the making decisions of the specific manufacturing operations between humans and robots [15,16]. In the HRC disassembly process, robots are usually required to update their disassembly actions dynamically with human operators in a shared workspace according to disassembly task planning. A reliable HRC workstation needs to develop its execution platform that can provide real-time updated models and data in a dynamic disassembly environment [17]. Furthermore, pattern recognition under high-speed cameras from traditional sensing technology further improves the possibility of HRC [18]. With the development of intelligent sensors and safety disassembly, the brainwave-driven HRC disassembly is gradually becoming a possibility to better accomplish the disassembly task executions and real-time data feedback [19]. However, human intention and action prediction can be achieved by recognizing the trajectory of human joint action, which can effectively improve the overall disassembly efficiency and ensure safety problems such as collisions that may cause wrong operations with intelligent robots [20]. In addition to human intention recognition, other methods also can be used to predict the disassembly tasks and related operations based on HRC by considering the multi-agents actions and intelligent algorithms [21,22]. As known, the intelligent models based on neural network algorithms can be also used to combine learning predictions of the disassembly process, which can predict simple human actions and determine the robot to accomplish the collaborative commands.

The EV retired batteries need to be recycled to ensure the recycling requirements that the disassembly process in the entire remanufacturing process should consider non-destructive disassembly. However, it is difficult to develop the use of a fully automatic disassembly mode based on HRC disassembly [23]. In terms of the optimization of the HRC disassembly sequence as shown in Table 1, the optimization process of the disassembly task is to generate the related basic disassembly sequence according to the actual task requirements based on the specific algorithm to optimize the final multi-objective optimization combined with multiple conditions [24]. The specific disassembly representation can be described by the disassembly graph model, which disassembles the original product into the subcomponents to form a tree graph structure. The selection of a disassembly graph model can be usually accompanied by multiple constraints to determine disassembly objects by optimizing the disassembly cost and times [25]. As known, the traditional optimization methods mainly include genetic algorithms

Table 1 Research works on disassembly sequence issues

Ref.	Disassembly optimization method	Research descriptions and contents
Xiao et al. [30]	DBN	• To propose a dynamic disassembly Bayesian network approach • To deduce the optimal disassembly sequences of EVB model using the forward–backward algorithm and the Viterbi decoding algorithm
McGovern et al. [36]	ACO(DLBP)	• To use ant colony algorithms to solve the balance problem between disassembly lines • To combine reasonable plans by the number of disassembly workstations • To improve the efficiency of disassembly line
Tripathi et al. [37]	ASGA	• To propose a fuzzy disassembly optimization model and Self-Guided Ants algorithms • To calculate the optimal disassembly strategy and optimal disassembly depth
Xia et al. [38]	STLBO	• To proposes a STLBO algorithm (teaching and learning) to optimize DSP issues • To ensure the characteristics of fast convergence speed and strong adaptability
Tseng et al. [39]	PSO	• To create a closed-loop model to optimize the disassembly sequences with PSO method • To reduce the lowest cost of disassembly
Zhao et al. [31]	DRL	• To propose the DQN algorithm based on the RL framework to solve the SDSP problem • To obtain the optimal disassembly sequence with the shortest disassembly time
Kheder et al. [40]	GA	• To propose a genetic algorithm to solve DSP problem • To optimize disassembly tools and disassembly process of disassembled parts
Liu et al. [41]	DRL	• To use DQN method to optimize the Human-machine collaboration operation and task allocation in real industrial production
Tian et al. [42]	DDA	• To propose a dual-objective discrete artificial bee colony to solve the dual-objective DSP problem (economical and energy) • To propose the disassembly uncertainties that need to be presented in practice

Note: DBN, Dynamic Bayesian network; ACO, Ant colony optimization; DLBP, disassembly line-balancing problem; ASGA, Algorithm of Self-Guided Ants; STLBO, Simplified Teaching-learning-based optimization; PSO, Particle swarm optimization; GA, Genetic Algorithm; DRL, Deep Reinforcement learning; DDA, Dual-objective discrete artificial bee colony.

based on bionic phenomena [26], ant colony algorithms [27], and other heuristic algorithms. Genetic algorithms are one of the commonly used methods to solve the optimal programming problems for EV battery disassembly optimization, which can be used to generate the optimal disassembly sequences for the entire disassembly process [28]. The ant colony algorithm can be used to find the possible optimal disassembly sequences for complex disassembled products, which finally achieves the selection of optimal disassembly sequences [29]. Recently, some new ideas are also used to find out the optimization of disassembly sequences based on dynamic Bayesian networks that can achieve the optimization of dynamic disassembly operations with uncertain conditions [30]. However, it should be a good method to consider reinforcement learning (RL) models that can be used to determine the disassembly sequence of complex retired products [31]. It is possible to achieve the decision-making of disassembly sequences based on deep reinforcement learning (DRL) architecture that can obtain the optimal disassembly sequence by considering the HRC disassembly tasks [32].

For the disassembly of retired EV batteries, the basic disassembly process can be used to acquire the stereoscopic visual positioning technology by machine vision to obtain the 3D positions of each disassembly component in the physical coordinate. The target recognition algorithm can complete the extraction of disassembled object features, which can finally synthesize the identification results to obtain the disassembly task allocation for each component of the battery [33]. At the module disassembly stage, because the average quality of the module after disassembly into a single body is higher, the disassembly object recognition is better than the battery pack. Therefore, by analyzing the disassembly module, it is found that the robot might better accomplish the high-speed cutting and other dangerous work, while a human can cooperate with to disassemble based on a HRC algorithm [34]. EV battery disassembly can be established by optimizing the working mode of the HRC disassembly platform, which can select the robot fully as an assistant that can be positioned by machine vision with efficient and non-destructive disassembly that can further select the robot disassembly tools [35].

However, there are many research points for EV battery disassembly based on the HRC disassembly process, including disassembly tool selection, disassembly task assignment, path optimization, and disassembly safety evaluations. Disassembly safety is a vital part of massive EV battery recycling that robot-assisted disassembly should change its machining speed with the disassembly operation action under the safety of HRC disassembly. The workstation of disassembly operation should have the ability to ensure the trajectory of the robot and the position of the human under a safe distance [34]. However, the disassembly operation environment through traditional industrial robots can be realized by security measures strictly to protect human actions. The safety of HRC disassembly can be used to consider the common data through cameras or other sensors, which should also be combined with hardware improvements. In summary, the optimization of HRC disassembly operations can be decomposed into multiple components in the real-time monitoring and control, disassembly working force, and specific disassembly operations based on disassembly robots.

By combining the current disassembly problem for automotive retired EV batteries, there are two main challenges for HRC disassembly methods as follows:

- Due to the complexity of the battery pack structure, various EV battery types and the various disassembly tools need to be determined for the variety of retired EV batteries. Therefore, it is necessary to find out the disassembly strategy by optimizing the appropriate disassembly sequence and tool selection according to their disassembly conditions such as collision protection, path optimization and human–robot making decisions.
- The recycling of EV battery is different from general retired products that EV battery pack may cause uncertain structure deformation due to impact, wear and tear. The uncertain structure might cause the complex disassembly process and unfixed disassembly sequence for each disassembled product. Therefore, it is difficult to meet the actual disassembly requirements to determine the entire disassembly strategy that can optimize the disassembly process based on HRC disassembly for massive EV battery.

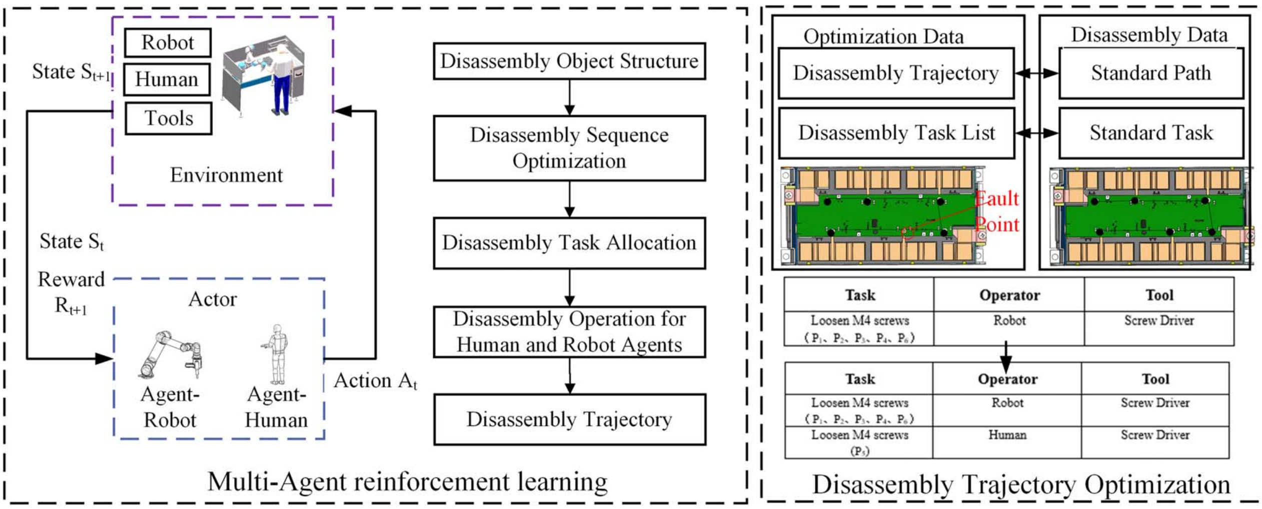

Fig. 1 The framework of disassembly task planning based on HRC disassembly

3 Dynamic Human–Robot Collaboration Disassembly Model

As discussed earlier, the dynamic HRC disassembly model needs a certain learning ability and efficient reasoning ability for disassembly tool selection, disassembly path optimization, and human–robot decision-making in the entire disassembly process, which should take into account the disassembly sequence optimization and disassembly tasks allocation in HRC disassembly process. After optimizing the disassembly sequence of the retired EV battery, a dynamic disassembly task planning model can be proposed to determine the disassembly task assignment and the disassembly operations according to disassembly features and retired product structure. However, an optimal disassembly sequence can be generated in real-time uncertain disassembly process according to a specific optimization algorithm for the retired disassembly object structure as shown in Fig. 1. The left part of Fig. 1 shows the procedure of the disassembly process for retired EV batteries based on multi-agent reinforcement learning for humans and robots, including disassembly object structure, disassembly sequence optimization, disassembly task assignment, multi-agent disassembly operations, and disassembly trajectory. The right part of Fig. 1 presents multi-agent disassembly path optimization through disassembly tasks in 2D disassembly planar. The disassembly operation and optimization trajectory for dynamic disassembly sequence can be accomplished by a deep reinforcement learning approach to match the optimal disassembly trajectory by combining the optimal disassembly sequence and disassembly tasks based on HRC disassembly. It is necessary to monitor the implementation of the disassembly system according to the specific disassembly process and disassembly paths to ensure disassembly operations. The novel multi-agent reinforcement learning approach based on HRC disassembly will be used to assign specific disassembly tasks based on the optimal disassembly sequence for the specific EV battery.

3.1 Disassembly Structure Graph Representation. Because the internal structure of a retired EV battery is relatively complex and there are more constraints between various components, it is more appropriate to represent the product disassembly graph that combines the directed graph and the undirected graph. As shown in Fig. 2, a retired product can be decomposed into component *a* and subassembly *bcd*, which indicates that the two subassemblies do not have priority in the actual disassembly process through direct contact and interference operations. The different disassembly operations have various disassembly priorities in the actual disassembly process. The composition of a disassembled hybrid diagram consists of three elements, represented by the diagram $G = \{V, E, DE\}$, where G represents the hybrid diagram, V refers to the node in the diagram, in the disassembly scenario refers to the individual parts that need to be disassembled, and E represents the undirected edge, which represents the contact constraint for the disassembled product. DE represents the direction edge, which

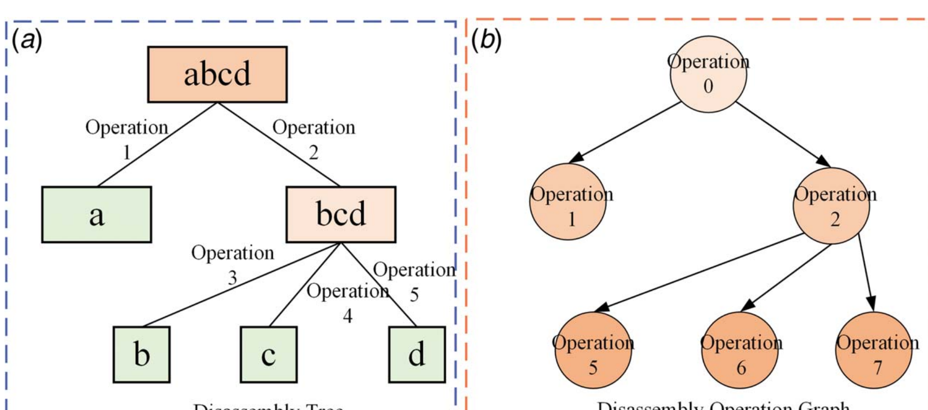

Fig. 2 (*a*) Disassembly Tree and (*b*) Disassembly operation graph

Table 2 Related symbol abbreviation definitions

Symbol abbreviations	Definition
PM	Priority constraint matrix
CM	Association constraint matrix
G	Total system benefit of the continuous process
S	State space of the environment
A	Action space of the agent
P	Agent action probability
R	Reward space
γ	Attenuation rate
S_H	State of the Agent-Human
S_R	State of the Agent-Robot
a_H	Action of the Agent-Human
a_R	Action of the Agent-Robot
O_A	The total number of operations
O_H	The number of operations completed by human
O_T	The number of operations completed by robot
D_H	The difficulty of operations completed by human
D_R	The difficulty of operations completed by robot
T_R	The time of completion by robot
π	The policy of the agent
π^*	Optimal policy
$\varepsilon - greedy$	Greedy probability
V	State value function
Q	Action value function
N	The number of disassembly targets

represents the priority constraint between the parts. The related mathematical symbol abbreviation has also been defined as listed in Table 2.

According to the representation of the disassembly graph model, we can generate a priority constraint matrix and an association constraint matrix, which can be used to visually represent the priority constraint between any two components, let i and j be one of the components as shown in the following:

$$PM_{ij} = \begin{cases} 1 & \text{Part } i \text{ has a priority constraint with part } j \\ 0 & \text{Otherwise} \end{cases} \tag{1}$$

The association matrix visually represents the association constraints between any two components and can be generated as follows:

$$CM_{ij} = \begin{cases} 1 & \text{Part } i \text{ is connected with part } j \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

The goal of the disassembly sequence can be used to optimize the disassembly time and costs in the complex disassembly process. The target of EV battery disassembly should be disassembled for some fasteners, which can be regarded as various disassembly tasks and operations to separate the assembly into subassembly or components. According to the connection types in the EV battery assembly structure, the specific connection for retired EV battery needs to be accomplished by the specific operations, which can define the total number of operations as O_A, the sum of the number of operations O_H completed by human, and the sum of the number of operations O_T completed by the robot. However, we define the disassembly operations for humans as D_H and the operation for the robot as D_R that can be used to analyze the disassembly operations for different disassembly tasks. According to the difficulty of disassembly tasks between the operation D_H completed by a human alone and the operation D_R completed by the robot alone, the time of independent disassembly operation by the robot can be defined as the time of disassembly task T_R. The disassembly time as an optimization goal can be used to acquire the optimal disassembly structure graph. Therefore, we can know that the optimal disassembly sequence can be calculated by the related constraints and disassembly requirements according to the specific disassembly structure for certain retired EV batteries.

3.2 Multi-Agent Reinforcement Learning. Reinforcement learning can accomplish the interaction between the individual and the environment to continuously acquire feedback data, which can finally obtain the desired results based on the input data. The method can be used to find out the optimal disassembly paths for the disassembly EV battery structure. The whole learning process can be discretized for each different moment, and the individual will interact with the environment that will trigger the environment a certain action. The environment will feedback to the individual state parameters under the activity and the reward generated by the previous activity in the discrete moment T. Each moment will allow the individual to receive the state S_t from the environment. The individual will choose an action A_t and the environment receives this action signal. The environment will feedback on the state S_{t+1} at the next moment t+1 and reward R_{t+1}. The framework of multi-agent reinforcement learning for HRC can be used to describe two agents' application optimization to represent the specific operations of humans and robots. For the total benefit of the continuous disassembly process, G_t should present an attenuation rate γ $(0<\gamma<1))$ calculated as follows:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{3}$$

The final result of reinforcement learning provides an optimal strategy for action that can guide the agent to take the next action. The probability of making the action at state S_t described in $\pi(a|s)$. Through continuous iteration, the total benefit will continue to increase and the strategy will become better and better to achieve the final optimization effect. The reinforcement learning model will focus on the Markov Decision Model (MDP) to determine the optimal strategy, which includes five parameters in the optimization process, namely (S, A, P, R, γ). The parameter S represents the state space, because the model in this scenario belongs to the finite Markov model, so the parameter in the state space collection is also limited. The parameter A represents the feasible space of action in the S state. The parameter $P(s'|s, a)$ indicates the probability that the agent will take the action a to the next action s' in the S state

$$P(s'|s, a) = \Pr[S_{t+1} = s' | S_t = s, A_t = a] = \sum_{r \in \mathbb{R}} p(s', r | s, a) \tag{4}$$

where $R(s', s, a)$ represents the real-time reward that is immediately received when the agent takes action to move to status s' under the state S. In this model, the agent can choose different actions according to different states, and the probability of state transfer can be also changed. There are two functions that follow the strategy to obtain the expected benefit based on reinforcement learning method, involving the state value function $V_\pi(s)$ and the action value function $Q_\pi(s, a)$. $V_\pi(s)$ indicates that the strategy does not change to harvest the expected return in the state S, and $Q_\pi(s, a)$ indicates that the desired return is harvested according to the strategic action after adopting action a in state S as follows:

$$V_\pi(s) = \mathbb{E}_\pi[G_0 | S_0 = s] = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R_{t+1} | S_0 = s\right] \tag{5}$$

$$Q_\pi(s, a) = \mathbb{E}_\pi[G_0 | S_0 = s, A_0 = a] = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R_{t+1} | S_0 = s, A_0 = a\right] \tag{6}$$

The state value function $V_\pi(s)$ and the action value function $Q_\pi(s, a)$ are recursively expanded to obtain the Bellman equation, which represents the recursive relationship between the function in the current state and the function in the subsequent state.

$$V_\pi(s) = \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a)[r + \gamma V_\pi(s')] \tag{7}$$

$$Q_\pi(s, a) = \sum_{s',r} p(s', r|s, a)\left[r + \gamma \sum_{a'} \pi(a'|s')Q_\pi(s', a')\right] \tag{8}$$

The optimal strategy is to find the maximum value of the two functions when the strategy π, and the optimal strategy is expressed as follows:

$$V^*(s) = \max_\pi V_\pi(s) \tag{9}$$

$$Q^*(s, a) = \max_\pi Q_\pi(s, a) \tag{10}$$

$$V^*(s) = \max_a Q^*(s, a) \tag{11}$$

The best strategy requires finding the strategy used by the optimal value function. You can take the way of maximizing $Q^*(s, a)$ to find the optimal strategy π

$$\pi^*(a|s) = \begin{cases} 1 & \text{if a = arg max Q*(s, a)} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

However, we can get the Bellman equation in the optimal case:

$$V^*(s) = \max_{a \in A} \sum_{s',r} p(s', r|s, a)[r + \gamma V^*(s')] \tag{13}$$

$$Q^*(s, a) = \sum_{s',r} p(s', r|s, a)\left[r + \gamma \max_{a'} Q^*(s', a')\right] \tag{14}$$

However, there are many ways to solve Bellman's equations: Value Iteration and Policy Iteration. According to the proposed Multi-Agent reinforcement learning framework, the traditional two-dimensional grid world is used as the environment to establish the disassembly task sequence representation different from the traditional optimization method to calculate the disassembly task time. The comprehensive negative return value in this framework evaluates the actions completed by different roles, which can be expressed as the statistics of three factors, the calculation formula is as follows:

$$Reward = \alpha^* operating\ time + \beta^* difficulty + \gamma^* quality \text{ impact} \tag{15}$$

However, the negative return of tasks that are difficult for robots to complete independently is defined as −99, and the return value in other cases is calculated according to the comprehensive negative return. Specifically, the current proposed reinforcement learning method for multi-agent human and robot operations can be used to determine the optimal disassembly path, which optimizes the maximum reward of the multi-agent disassembly task to specify the disassembly paths. The disassembly task planning based on multi-agent reinforcement learning can be divided into the following two disassembly ways. The first one is to optimize the specific disassembly sequential tasks based on the disassembly graph model for the specific retired EV battery, which will be completed by one operator or multiple operators cooperatively. As known, the next disassembly tasks cannot be completed before the task has not been completed. The other one is to execute a parallel disassembly task, which means that the same level of disassembly tasks can be carried out by two operators at the same time without interfering with each other in the process of parallel disassembly tasks. By optimizing the disassembly task planning for the retired EC battery, the value of the combined negative return will be reduced by 40% due to the reduction of disassembly operation time. The same level of disassembly tasks can only be completed by one or two operators, which need to cooperatively complete the specific operation step of the disassembly sequential task as shown in Fig. 3. The entire reinforcement learning environment is composed of $5 \times N$ grids that can be determined by the disassembly task goal. According to different disassembly object structures and disassembly task lists, the reinforcement environment has also huge differences. The first row of white grids is the start state where the agents start a single round of search tasks from the start point. The rightmost red grid row is the endpoint, when the agent reaches the endpoint, it can be considered to have completed the current round of search tasks. The bottom is to achieve the proper disassembly depth. The same pair of white grids in the middle and the starting point is the adjustment

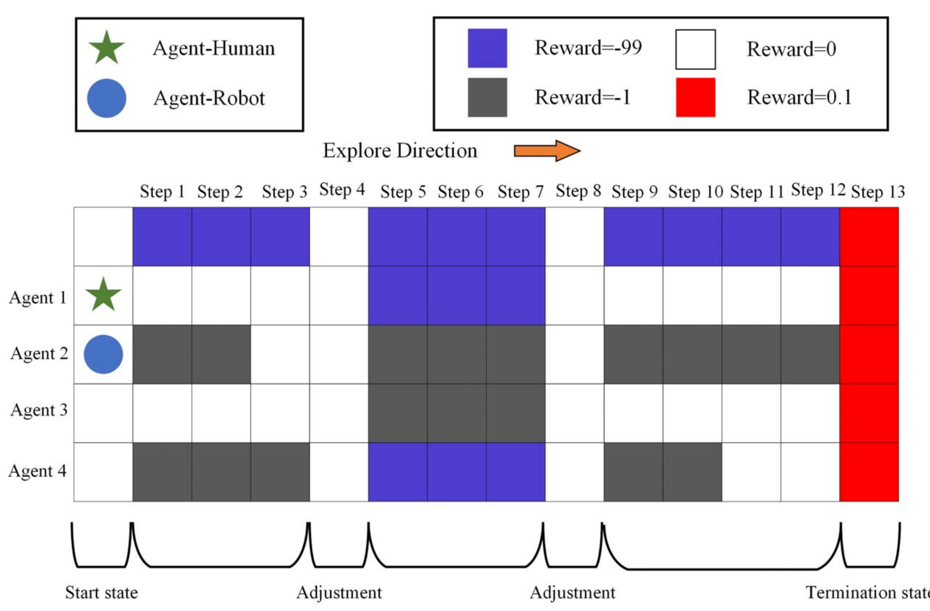

Fig. 3 Disassembly sequential operations and parallel operations based on multiple-agent reinforcement learning algorithm

selection area, and the rest of the areas are task areas. In the disassembly task area, each row represents a task selection. Two agents cannot appear in the same row, and the agents cannot change ranks in the task area. The grids of different colors represent different negative feedback values. The dark blue grids represent that the task cannot be completed independently by the robot–agent. The white grids indicate that the agent waits for another task to be completed and does not take any action itself. The total value of the black grids in the environment represents the cost required by the agent to complete the entire task. In this environment, the agent's exploration direction can only be right, up, or down.

In the disassembly task searching area, the exploration direction of the agent can only go to the right from the starting point to the adjustment selection area, and the exploration direction of the agent can only be up, down, or right. For the HRC disassembly environment, dual agents can select at most two tasks to accomplish simultaneously according to the order of disassembly priorities. The disassembly sequential task can be only completed by a single agent, and the other agent needs to wait for the completion of the disassembly task to quit the sequential task area after the disassembly task accomplishment. The adjustment selection area will reselect the agent for the next task. The execution of the parallel task is completed by two agents at the same time. The dual-agent decides who will complete the two parallel tasks (task A and task B): human for task A or robot for task B. The robot for task A can be accomplished by continuous exploration and matching of agents that the optimal task matching can be found based on a reinforcement learning environment. Even if the total negative income is the smallest, it is considered that the total cost (including disassembly time and disassembly difficulty) is the least in the real environment to reach the minimum value of disassembly time and disassembly difficulty.

3.3 Disassembly Task Planning and Tool Selection. In order to acquire the optimization of disassembly task planning for the real-time disassembly sequence, the entire disassembly operations need to be assigned to different operators, such as human or robot agents. After obtaining the related disassembly tasks for various agents, the specific operator can select the specific disassembly tool according to the disassembly task assignment. The disassembly process will consider two ways to improve the intelligent disassembly of complex EV batteries. The first is to visually identify the disassembly tools required for the related disassembly task, such as selecting the related series of sleeves according to various screw sizes. The second is to consider the related disassembly operations through the disassembly sequential task lists in the disassembly system. Because the flexibility of the disassembly robot is difficult to select suitable disassembly tools and the switching of the disassembly tools are different according to the complex disassembly objects. However, the frequent switching process will also lower the disassembly efficiency for multi-agent disassembly operations in the more accurate control system. In the previous disassembly sequence optimization, the switching process of disassembly tools can be optimized in actual production to determine the disassembly tool selection that can be used to change the entire disassembly process.

In order to implement the disassembly tasks, the initial assignments should be prioritized by operation types. For example, the robot is more efficient than the human to complete some operations such as loosening. The specific type of disassembly operations involving the loosening operation can consider the robot operation to complete the disassembly task. The efficiency of disassembling a wire harness by manual removal is much higher than that of robots that will consider more manual disassembly tasks involving flexible disassembly. However, in the actual disassembly process, it is possible to consider humans to accomplish some loosening disassembly tasks because of the damaged disassembly objects and the efficiency problem of robot switching disassembly tools. Therefore, it is necessary to combine the disassembly sequence planning to optimize the actual disassembly tool and the final disassembly task allocation. Owing to the uncertainty of the internal and external complex conditions of the retired EV battery pack, the actual disassembly task may dynamically optimize the final assignment of entire disassembly tasks.

By considering a retired EV battery module as an example, it is necessary to make a detailed representation of the basic disassembly tasks for an automotive EV battery module. Although the disassembly task for a retired EV battery is complex to determine the disassembly operations, the disassembly structure can be represented by separating the general structures, including decomposing the battery internally into a single module and the module into a single cell. According to the actual disassembly situation, the disassembly task can be specifically decomposed into the robot operations and the human operations, which can be mathematically optimized to determine the robot and the human according to the different difficulty coefficients based on various agents for the appropriate allocation of disassembly tasks. However, it is important to generate the disassembly path that can be used to accomplish the disassembly task from three stages. The first stage is to identify the disassembly target by the machine vision method, which can be usually achieved using the target recognition algorithm in the vision system. The second stage is that a related disassembly path can be generated according to the disassembly target in the disassembly process. The third stage is to optimize the disassembly path by perceiving the surrounding environment to achieve the optimization of the disassembly paths and allocate the specific disassembly tasks. In general, the selection of the optimal disassembly path mainly considers the comprehensive disassembly time and the complexity of the disassembly operations based on various disassembly tasks.

By considering the Q-learning algorithm based on reinforcement learning, we can also solve the disassembly path optimization problem for complex disassembly operations. Because retired EV batteries have the characteristics of multi-operation targets in the same plane, the coordinated action of the entire disassembly robot can be discretized in a two-dimensional plane. The area needs to be first meshed that can be used to capture the disassembly features and entire images can be also divided into feature and non-feature areas. The target grid to be disassembled is dyed black, while the network to be disassembly in other processes will be dyed gray, which can simplify the robot disassembly trajectory and its disassembly strategy is similar to the artificial intelligence decision to generate the retired EV battery disassembly path. We can assume that the robot starts from any black square, and the next step can fall on the other 8 squares adjacent to it. The probability of each square is 0.125, and the return of the white square is −0.01, the income of the black square is 100. We define n as the number of disassembly targets, and the optimization goal is to stop when the total return reaches $(n-1) * 100$ or more. In order to prevent the same black grid from being accessed repeatedly, the robot will reset the grid to a white grid with a return of −0.01 every time it passes through the grid. The optimal strategy will be selected for the maximum total return. As the number of iterations increases, the reward value will converge more and more to the maximum reward value based on the Q-learning algorithm as shown in Fig. 4.

In the actual battery recycling process, because the battery pack has generally huge differences for retired EV battery quality, it is difficult to identify the battery shapes and structures such as deformation and rust that might affect the variations of the entire disassembly process. Based on the template matching method, a new recognition can be proposed to compare the existing recognition results according to disassembly sequences, which can be used to complete more efficient disassembly task assignments. Yolo and other pattern recognition algorithms can perceive the disassembly depth for the specific EV battery structure by comparing the difference between screws and their holes. But it is difficult to accomplish the disassembly task planning that considers the real-time generation of the disassembly sequence for each disassembly operation. According to all kinds of battery packs in the retired EV battery recycling, a fundamental EV battery structure and its graph model

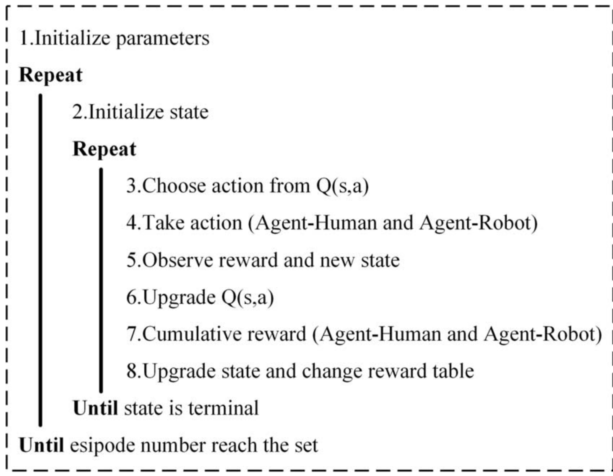

Fig. 4 The flowchart of Q-learning reinforcement learning for disassembly process

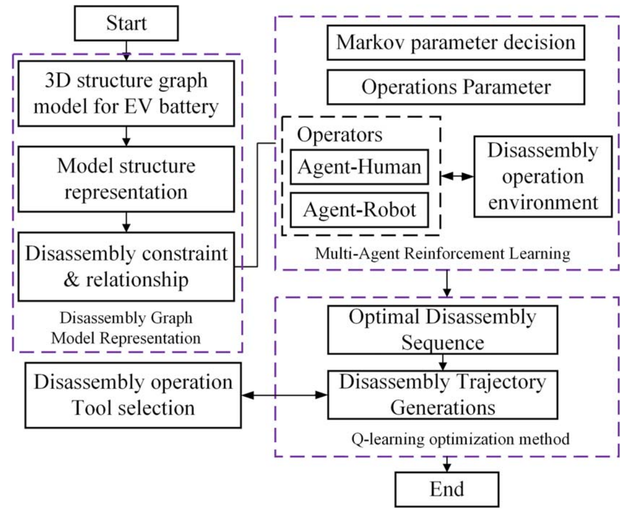

Fig. 5 Disassembly optimization based on multi-agent reinforcement learning

should be considered to determine the final disassembly strategy and operations. By combining the image capture of 3D camera in the actual disassembly process, the similarity of the existing structure should be first matched based on the external structure features. The related battery pack can be screened to match the disassembly planning with the optimal disassembly path. However, it is necessary to solve the problems of the actual disassembly tasks and give some specific marks for all the problem points that can be used to assign to disassembly operations for humans in the disassembly task. The human can use special disassembly tools to manually disassembly EV battery structure and finally interact with the system to determine the final disassembly strategy.

However, the implementation of retired EV battery disassembly mainly relies on the initial matching of the outer profile or other features to obtain a rough battery pack structure model and to generate the subsequent disassembly tasks and disassembly operations as shown in Fig. 5 The EV battery disassembly needs to capture as many characteristics as possible to distinguish different retired EV batteries in the disassembly operations, which also need to rely on the operator to determine the specific EV battery disassembly strategy that helps HRC disassembly achieve the disassembly flexibility in complex retired EV battery recycling. Due to the uncertainty of battery structure and quality in the actual disassembly identification process, the disassembly trajectory for retired EV batteries is also crucial to determine the optimal disassembly strategy based on the basic structure of the battery pack. Due to the uncertainty of battery pack quality, the battery pack size and shape are also difficult to guarantee consistency with the normal battery pack structure with any possible damages in the actual identification process. Because humans and robots share a common disassembly operation space in the actual disassembly workstation for the retired EV battery pack, it is important to ensure the disassembly safety by considering more human–robot interactions in the actual disassembly process. The robot-assisted disassembly operation will present an anti-collision design method to protect the robot and human from accident events based on HRC disassembly by the robot sensors.

4 Case Study

By considering the MS372P5s Samsung battery module as an example, it can be roughly divided into the following subassembly or components: (1) Upper End Cover; (2) Module Circuit Board; (3) Structure Frame; (4) Pole Head; (5) Pole Piece; (6) Insulating Plates; (7) Battery Cell; (8) Thick Side Shell. As known, it is necessary to optimize the disassembly process that accomplishes optimal disassembly planning by determining the disassembly sequences and the specific disassembly operations for each disassembly task. However, various subassemblies or components have some complex relationships between their connections that might make the difficulty of disassembly operations without any disassembly tools or assistant devices. The battery module can be finally divided into eight components that can be used to effectively recycle the components by HRC disassembly. The disassembly sequences can be represented by the disassembly graph model as shown in Fig. 6, which can be used to optimize the disassembly operations and disassembly sequences based on various disassembly parameters for EV battery structure.

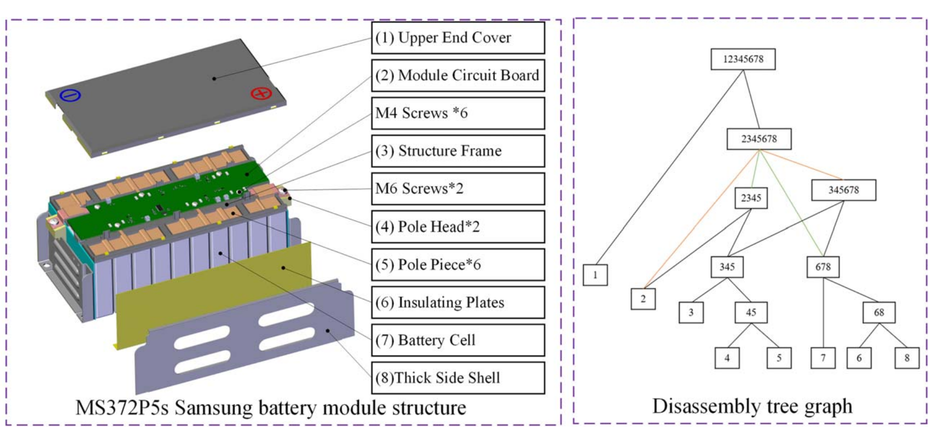

Fig. 6 Disassembly object 3D model and disassembly graph structure

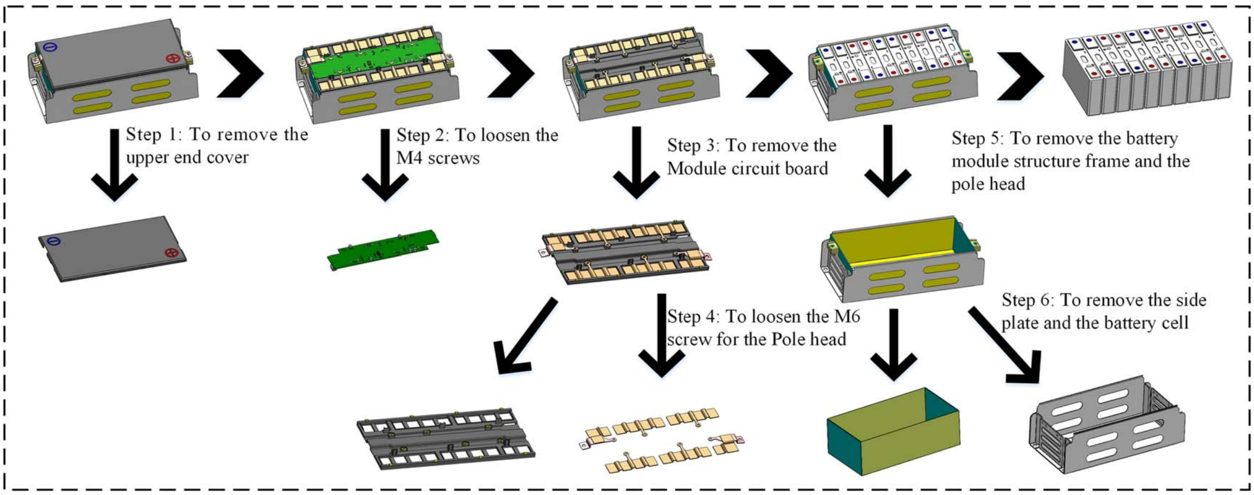

Fig. 7 The procedure of battery module separation into battery cells

Furthermore, EV battery disassembly is to acquire the fundamental components or the valuable subassembly to re-utilize its values or its functions through EV battery recycling. The case study can be used to demonstrate the entire disassembly process of the EV battery module from the original structure to each component, which can be described by a series of disassembly operations based on the optimal disassembly sequences in the actual disassembly process. However, it is necessary to discuss the disassembly sequences based on HRC disassembly that needs to specify the disassembly steps. The original manual disassembly sequence is to remove the upper end cover first, and the step has no fasteners to remove because the traditional manual operation is less difficult to accomplish the removal without any robot operations. The second step requires loosening the M4 Phillips screws that fix the Module circuit board in six installation places, which is difficult to operate by traditional manual operation. Therefore, the robot-assisted disassembly operation should be considered to accomplish repeatable disassembly tasks. The third step is to remove the Module circuit board from the EV battery module by separating the fastener connectors with manual disassembly operations. The fourth step requires loosening the M6 Phillips screw that holds the Pole head in a connection place that is generally difficult to manually disassembly operations with huge repeatable actions. The operations might cause low disassembly efficiency when facing a load of EV battery recycling. The fifth step is to remove the battery module structure frame and the pole head in the actual disassembly operations. However, the disassembly process mainly includes the specific flexible disassembly operations that will be accomplished by manually disassemble rather than robot actions. The sixth step is to separate the EV battery pole pieces from the battery module by manual disassembly operations. The seventh step is to remove the side plate and the battery cell, which normally requires the cutting disassembly operations to remove the outer shell or side plates as shown in Fig. 7.

However, the disassembly sequential task can be represented by a disassembly graph structure with various disassembly constraints in the actual disassembly process from the EV battery module to battery cells. The disassembly graph for the EV battery module case can be represented by related association and priority relationships for the actual disassembly sequences for each disassembly subassembly or component. As shown in Fig. 7, MS372P5s Samsung battery module can be represented by the disassembly graph structure into eight components, including (1) upper end cover, (2) module circuit board, (3) structure frame, (4) pole head, (5) pole pieces, (6) insulation plates, (7) battery cells, and (8) thick side shell. As known, component (1) has the priority for the specific disassembly operations, which directly links to component (8). The component (2) has more priority than components (4) and (5). The disassembly sequence for the disassembly case for component (3), component (2), component (4), and components (5) has a direct association relationship in the actual disassembly process. Component (3) has more priority than component (6) and component (7). And component (6) has more priority than component (7). Therefore, the association and precedent relationship between disassembly processes can be described as shown in Fig. 8.

In order to accomplish the optimal disassembly sequence and disassembly tasks based on HRC disassembly, it is necessary to determine the disassembly operations for the specific EV battery. The specific disassembly operations have been discussed for related steps for one human as an agent or one robot as another agent in the actual disassembly process. However, all disassembly operations are scored on the difficulty of the disassembly task, and any of the disassembly operations are scored according to the different working characteristics of humans and robots. We define the score of easy completion of characters as 0.25, the score of disassembly operations that are not easy to complete is 0.75, and the average degree of disassembly operation is 0.5. Therefore, we can obtain the evaluation scores of EV battery disassembly tasks based on HRC as shown in Table 3.

The multi-agent reinforcement learning can be used to optimize the disassembly sequences and operations for EV batteries in the actual disassembly process. By disassembling the specific EV battery structure based on human and robot agents as disassembly operators, the optimization of the disassembly process can be obtained by defining many parameters for the specific algorithm. The total number of disassembly operations (O_A) can define as 8. The total disassembly target is eight components and two fasteners to build the grid world based on the disassembly graph structure. We define the experience rate as 0.2, learn rate as 0.8, and the reward discount rate as 0.98. However, the optimization results of the disassembly process for the different teardown paths can be

$$\mathrm{CM}_{ij}=\begin{bmatrix} 0&0&0&0&0&0&0&1\\ 0&0&1&0&0&0&0&0\\ 0&1&0&1&1&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{bmatrix}$$

(1) Priority Constraint Matrix

$$\mathrm{PM}_{ij}=\begin{bmatrix} 0&1&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0 \end{bmatrix}$$

(2) Association Matrix

Fig. 8 Constraint matrix (1) and association matrix (2) for EV battery module

Table 3 The evaluation of EV battery disassembly task based on reinforcement learning for human–robot agents

Specific disassembly operation step	Operator	Tool	Difficulty of operation	Comprehensive negative reward
To remove the upper end cover	Human	Hand	0.25	−2
	Robot	Fixture	0.5	−4
To loosen the M4 screws	Human	Screw driver	0.5	−5
	Robot	Screw driver	0.25	−3
To remove the module circuit board	Human	Hand	0.25	−2
	Robot	Fixture	0.75	−4
To loosen the M6 screws	Human	Screw driver	0.5	−5
	Robot	Screw driver	0.25	−3
To remove the structure frame and pole pieces	Human	Hand	0.25	−3
	Robot	Fixture	1.0	−99
To separate pole pieces from the structure frame	Human	Hand	0.25	−3
	Robot	Fixture	1.0	−99
To fetch the battery cell	Human	Hand or Special tools	0.25	−3
	Robot	Fixture	1.0	−99
To separate insulating plate	Human	Hand	0.25	−3
	Robot	Fixture	1.0	−99

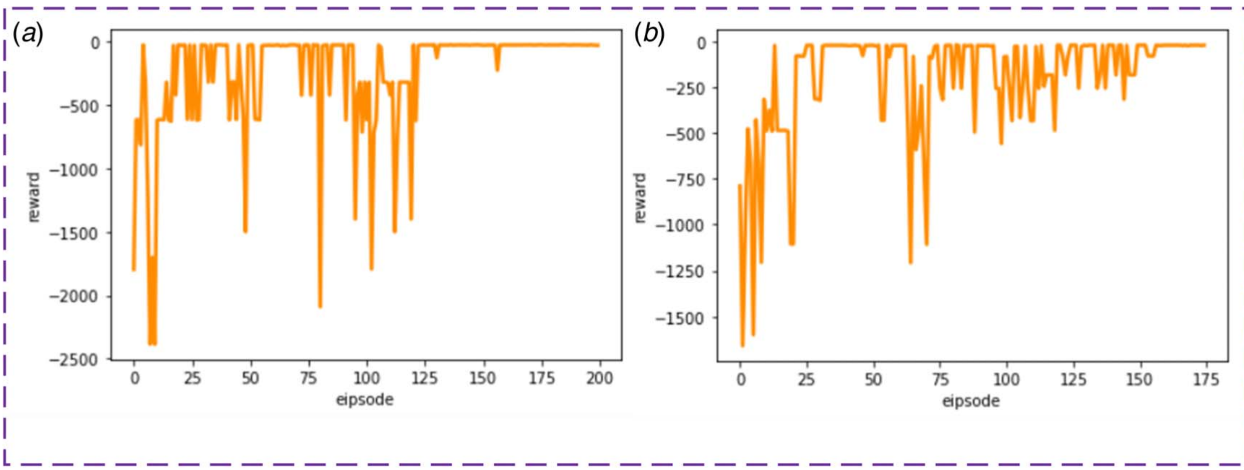

Fig. 9 The optimization results for two different disassembly tree graphs (*a* and *b*) based on the MS372P5s Samsung battery module (see Fig. 7).

obtained as follows in Fig. 9. We know that the disassembly tree graph can be divided into two kinds of disassembly sequences for the specific EV battery. The optimization results for the two disassembly sequences can be used to explain the final convergence after the iterations of about 150 times or over 160 times, respectively. Finally, the optimal disassembly sequence can be determined by the minimum rewards based on the optimization algorithm.

By optimizing the disassembly sequential tasks for the specific EV battery, the multi-agent disassembly operations can be accomplished to acquire the optimal disassembly strategy for the specific disassembly battery. The robot–agent disassembly process mainly focuses on repeatable and simple disassembly operations that can be accomplished by the specific disassembly step with the disassembly tools. The human–agent disassembly process might solve the flexible disassembly tasks with uncertain disassembly situations that the process is difficult to be accomplished by robot operations. Therefore, it is necessary to adopt the HRC disassembly to allocate the specific disassembly operations for the disassembly tasks. As shown in Fig. 10, each disassembly task can allocate a specific disassembly operation by determining the specific human–agent or robot–agent operators to accomplish the disassembly process based on the optimization method. According to the actual disassembly requirements and constraints, the disassembly steps can be regularized by defining the disassembly waiting steps and replacing the disassembly tool steps in the actual disassembly process. The entire disassembly process will be executed by the disassembly operations to acquire the final disassembly components or subassembly for retired EV battery.

In order to further discuss the disassembly trajectory based on robot agents, it is important to find out the optimal disassembly path for the disassembly tasks in the disassembly process. It is necessary to select the robot–agent disassembly task as an optimization example to demonstrate the disassembly path optimization for the retired EV battery module based on the reinforcement learning method. The disassembly task for loosening the M4 screws can be used to optimize the robot–agent disassembly trajectory for the final disassembly strategy. The disassembly planar can be divided into discrete grids that provide the disassembly optimization environment with respect to the optimization conditions. The recognition of disassembly screws can be accomplished by image vision as the initial optimization conditions, which can acquire the optimal disassembly paths by the reinforcement learning method. As shown in Fig. 11, by defining the second step of robot disassembly recognition for the M4 screw as the disassembly path on the two-dimensional planar, it can be set the target reward value to 500, and the robot traverses six points as an iteration. The experience rate can be defined as 0.2, learn rate as 0.01, and the reward discount rate as 0.95. However, the final optimization results can be demonstrated as the disassembly paths for the specific disassembly task, such as the 4M screws disassembly task.

5 Discussion

The optimization of disassembly task planning based on reinforcement learning is proposed in this paper to combine the

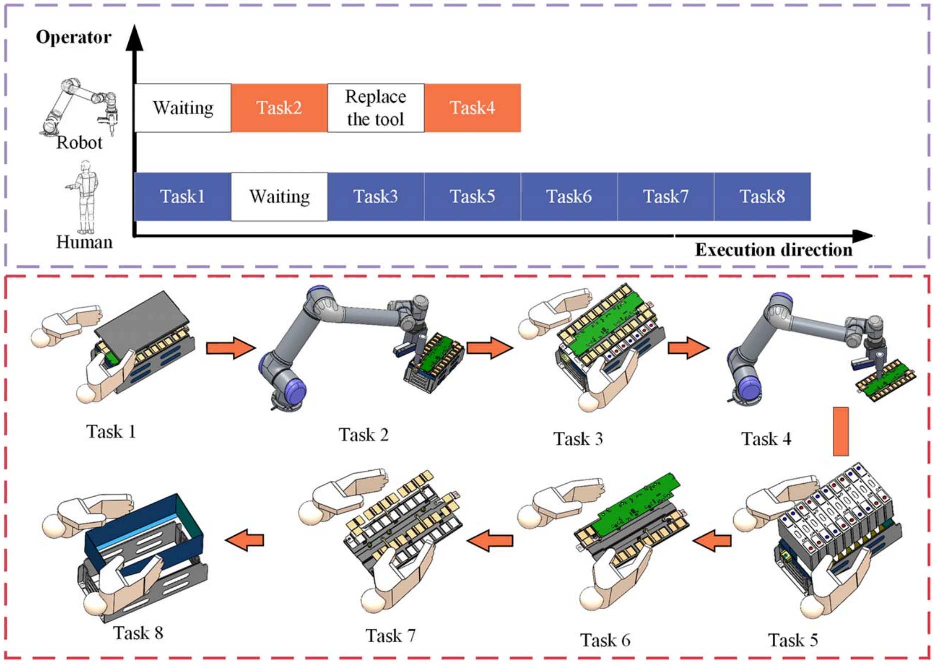

Fig. 10 The optimal disassembly sequence for human–robot collaboration disassembly

necessary disassembly conditions for HRC disassembly such as disassembly trajectory, disassembly task assignment, and disassembly tool selection, which can be used to solve the specific two problems for retired EV battery:

- Owing to the complexity and uncertainty of EB battery disassembly, the optimization of disassembly sequences is still a huge challenge for better solving the problem. However, it is difficult to acquire the optimal disassembly sequences based on the dynamic disassembly process, which might cause changes of the optimal disassembly sequence owing to the real-time disassembly status and condition alterations [43]. However, dynamic disassembly sequence optimization is necessary to solve the real-time optimization of the disassembly process for the EV battery packs or modules in a complex and changeable environment. The intelligent optimization

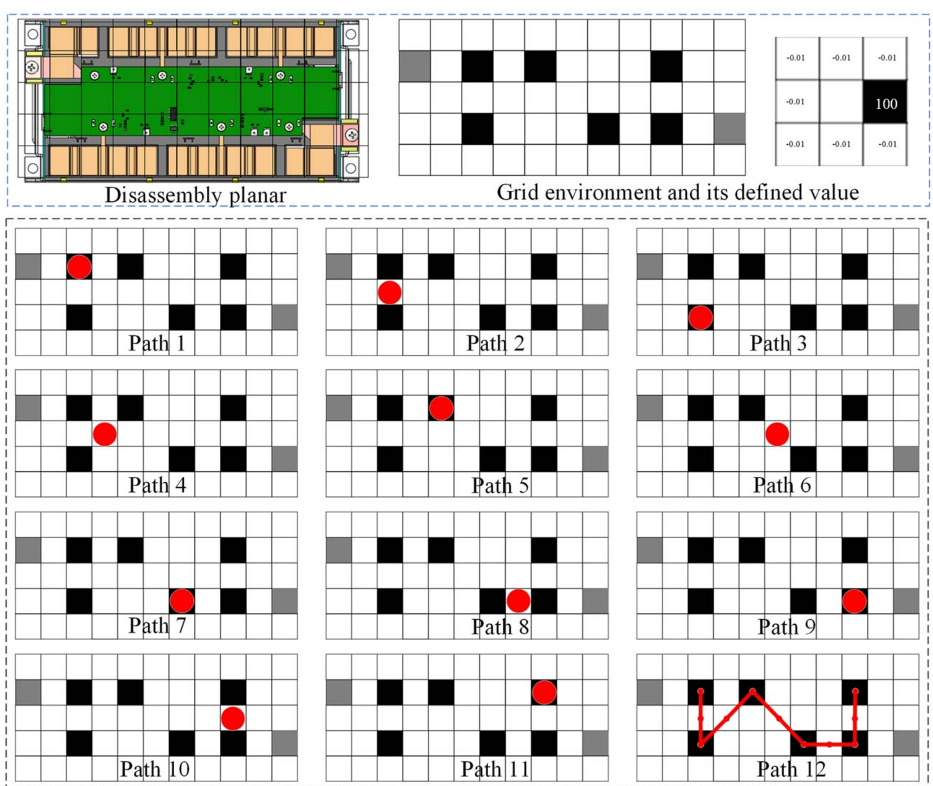

Fig. 11 The optimization procedure of disassembly paths for 4M screws

algorithms provide a possible way to solve the complex and uncertain disassembly process. Multi-agent deep reinforcement learning can dynamically optimize the entire disassembly sequential tasks in a complex and changeable environment. By defining human–robot agent disassembly tasks, EV battery disassembly interaction with the complex disassembly environment can be accomplished to combine with dual network parameter optimization by Actor-based and Critic-based reinforcement learning methods, which plays a more effective role in solving disassembly sequence optimization than traditional heuristics algorithms.

- Owing to the complexity of the dynamic disassembly trajectory for the specific disassembly task in the actual EV battery disassembly, it is necessary to explore the optimal disassembly paths based on intelligent optimization algorithms. The disassembly task will involve various disassembly operations to determine the human or robot disassembly execution in the actual disassembly process [44]. The disassembly optimization process will generate the suspected fault point after the analysis of the disassembly trajectory or the disassembly obstacle caused by the uncertain EB battery quality. However, it is important to solve the problems of the re-optimization based on the obstacle point. At present, in the entire disassembly process, human disassembly operations have more flexibility to solve the fault points in a more reliable way. However, it is possible to combine the new digital twin technology with efficient data mining and optimization identification that can cope with special situations to further reduce human fatigue.

By reviewing many recent literature studies, the disassembly process optimization based on deep learning is gradually replacing many traditional optimization methods to provide a greater potential for disassembly tasks in complex scenarios. However, there are many perspectives in future research as follows:

- However, it is possible to explore the further disassembly optimization of EV battery packs, the standard structure of the battery should be further classified to reduce the difficulty of the recycling and disassembly process by HRC disassembly. The robot-assisted disassembly process has higher efficiency for repeatable disassembly operations. Similarly, the selection of disassembly tools can also be more concise to accomplish the specific disassembly tasks, which not only helps to further improve the existing disassembly speed but also makes the entire disassembly process more secure and stable.
- In terms of disassembly safety, because the EV battery is a retired product with certain dangers, it cannot be directly separated without the full release of its internal power in the entire disassembly process. However, infrared and other temperature sensors can use neural network algorithms to predict the battery health states to avoid the occurrence of undesired accidents. With the evolution of disassembly optimization for EV batteries, the battery disassembly path will be smarter to reduce the possibility of disassembly risks. In addition, the predictive algorithm can also identify the human action intention in time before the collision and dynamically adjust the disassembly trajectory of the robot.

6 Summary

In this paper, an HRC disassembly optimization method for the recycling of EV batteries has been proposed based on multi-agent reinforcement learning. The disassembly structure of an EV battery has been represented by a disassembly graph model and by combining reinforcement learning algorithms to specify the disassembly tasks for certain EV batteries. For the specific disassembly task assignment, disassembly operations based on the HRC disassembly method have been demonstrated to dynamically optimize 2D disassembly operation paths on the disassembly planar. Then, the disassembly trajectory has provided a dynamic planning method to achieve the adjustment of the optimal disassembly paths. Finally, we discussed the current disassembly of EV battery and its potential research directions by combining with intelligent algorithms to achieve a more efficient and safer way in the entire disassembly process.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The data sets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Zhu, Y., Ding, Q., Zhao, Y., Ai, J., Li, Y., and Cao, Y., 2020, "Study on the Process of Harmless Treatment of Residual Electrolyte in Battery Disassembly," Waste Manage. Res., **38**(11), pp. 1295–1300.
[2] Bauer, A., Wollherr, D., and Buss, M., 2008, "Human-Robot Collaboration: A Survey," Int. J. Humanoid Rob., **5**(1), pp. 47–66.
[3] Wu, T., Zhang, Z., Yin, T., and Zhang, Y., 2022, "Multi-Objective Optimisation for Cell-Level Disassembly of Waste Power Battery Modules in Human-Machine Hybrid Mode," Waste Manage., **144**(4), pp. 513–526.
[4] Xiao, J., Li, W., Lv, Y., and Du, G., 2021, "Disassembly Information Interoperability for Electric Vehicle Battery in Re-Manufacturing Based on STEP Standards," Procedia CIRP, **104**(1), pp. 1873–1877.
[5] Lu, Q., Ren, Y., Jin, H., Meng, L., Li, L., Zhang, C., and Sutherland, J. W., 2020, "A Hybrid Metaheuristic Algorithm for a Profit-Oriented and Energy-Efficient Disassembly Sequencing Problem," Rob. Comput. Integr. Manuf., **61**(1), p. 101828.
[6] Ilgin, M. A., and Taşoğlu, G. T., 2016, "Simultaneous Determination of Disassembly Sequence and Disassembly-to-Order Decisions Using Simulation Optimization," ASME J. Manuf. Sci. Eng., **138**(10), p. 101012.
[7] Kay, I., Farhad, S., Mahajan, A., Esmaeeli, R., and Hashemi, S. R., 2022, "Robotic Disassembly of Electric Vehicles' Battery Modules for Recycling," Energies, **15**(13), p. 4856.
[8] Yang, Y., Okonkwo, E. G., Huang, G., Xu, S., Sun, W., and He, Y., 2021, "On the Sustainability of Lithium Ion Battery Industry—A Review and Perspective," Energy Storage Mater., **36**(1), pp. 186–212.
[9] Xiao, J., Jiang, C., and Wang, B., 2023, "A Review on Dynamic Recycling of Electric Vehicle Battery: Disassembly and Echelon Utilization," Batteries, **9**(1), p. 57.
[10] Bentaha, M. L., Battaïa, O., Dolgui, A., and Hu, S. J., 2014, "Dealing With Uncertainty in Disassembly Line Design," CIRP Ann., **63**(1), pp. 21–24.
[11] Liu, Z., Liu, Q., Xu, W., Liu, Z., Zhou, Z., and Chen, J., 2019, "Deep Learning-Based Human Motion Prediction Considering Context Awareness for Human-Robot Collaboration in Manufacturing," Procedia CIRP, **83**(1), pp. 272–278.
[12] Marconi, M., Palmieri, G., Callegari, M., and Germani, M., 2019, "Feasibility Study and Design of an Automatic System for Electronic Components Disassembly," ASME J. Manuf. Sci. Eng., **141**(2), p. 021011.
[13] Marshall, J., Gastol, D., Sommerville, R., Middleton, B., Goodship, V., and Kendrick, E., 2020, "Disassembly of Li Ion Cells—Characterization and Safety Considerations of a Recycling Scheme," Metals (Basel), **10**(6), pp. 1–22.
[14] Jiang, H., Yi, J., Zhu, X., and Li, Z., 2018, "Generating Disassembly Tasks for Selective Disassembly Using Ontology-Based Disassembly Knowledge Representation," Assem. Autom., **38**(2), pp. 113–124.
[15] Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X. V., Makris, S., and Chryssolouris, G., 2019, "Symbiotic Human-Robot Collaborative Assembly," CIRP Ann., **68**(2), pp. 701–726.
[16] Liu, S., Wang, L., and Wang, X. V., 2021, "Sensorless Haptic Control for Human-Robot Collaborative Assembly," CIRP J. Manuf. Sci. Technol., **32**(1), pp. 132–144.
[17] Liu, S., Wang, X. V., and Wang, L., 2022, "Digital Twin-Enabled Advance Execution for Human-Robot Collaborative Assembly," CIRP Ann., **71**(1), pp. 25–28.
[18] Liu, H., and Wang, L., 2018, "Gesture Recognition for Human-Robot Collaboration: A Review," Int. J. Ind. Ergon., **68**(11), pp. 355–367.
[19] Wang, L., Liu, S., Cooper, C., Wang, X. V., and Gao, R. X., 2021, "Function Block-Based Human-Robot Collaborative Assembly Driven by Brainwaves," CIRP Ann., **70**(1), pp. 5–8.
[20] Butepage, J., Kjellstrom, H., and Kragic, D., 2018, "Anticipating Many Futures: Online Human Motion Prediction and Generation for Human-Robot Interaction," Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Vol. 1, No. 1, pp. 4563–4570.
[21] Li, S., Wang, R., Zheng, P., and Wang, L., 2021, "Towards Proactive Human–Robot Collaboration: A Foreseeable Cognitive Manufacturing Paradigm," J. Manuf. Syst., **60**(1), pp. 547–552.

[22] Li, S., Zheng, P., Fan, J., and Wang, L., 2022, "Toward Proactive Human—Robot Collaborative Assembly: A Multimodal Transfer-Learning-Enabled Action Prediction Approach," IEEE Trans. Ind. Electron., **69**(8), pp. 8579–8588.
[23] Hjorth, S., and Chrysostomou, D., 2022, "Human–Robot Collaboration in Industrial Environments: A Literature Review on Non-Destructive Disassembly," Rob. Comput. Integr. Manuf., **73**(1), p. 102208.
[24] Zhou, Z., Liu, J., Pham, D. T., Xu, W., Ramirez, F. J., Ji, C., and Liu, Q., 2019, "Disassembly Sequence Planning: Recent Developments and Future Trends," Proc. Inst. Mech. Eng. Part B J. Eng. Manuf., **233**(5), pp. 1450–1471.
[25] Hu, Y., and Ameta, G., 2019, "A Charts-Based Approach to Estimate Disassembly Time: Hypothesis, Model and Validation," ASME J. Manuf. Sci. Eng., **141**(2), p. 021009.
[26] Kongar, E., and Gupta, S. M., 2006, "Disassembly Sequencing Using Genetic Algorithm," Int. J. Adv. Manuf. Technol., **30**(5–6), pp. 497–506.
[27] Wang, J. F., Liu, J. H., Li, S. Q., and Zhong, Y. F., 2003, "Intelligent Selective Disassembly Using the Ant Colony Algorithm," Artif. Intell. Eng. Des. Anal. Manuf., **17**(4–5), pp. 325–333.
[28] Tao, F., Bi, L., Zuo, Y., and Nee, A. Y. C., 2018, "Partial/Parallel Disassembly Sequence Planning for Complex Products," ASME J. Manuf. Sci. Eng., **140**(1), p. 011016.
[29] Ilgin, M. A., 2019, "A DEMATEL-Based Disassembly Line Balancing Heuristic," ASME J. Manuf. Sci. Eng., **141**(2), p. 021002.
[30] Xiao, J., Anwer, N., Li, W., Eynard, B., and Zheng, C., 2022, "Dynamic Bayesian Network-Based Disassembly Sequencing Optimization for Electric Vehicle Battery," CIRP J. Manuf. Sci. Technol., **38**(1), pp. 824–835.
[31] Zhao, X., Li, C., Tang, Y., and Cui, J., 2021, "Reinforcement Learning-Based Selective Disassembly Sequence Planning for the End-of-Life," IEEE Rob. Autom. Lett., **6**(4), pp. 7807–7814.
[32] Xu, W., Tang, Q., Liu, J., Liu, Z., Zhou, Z., and Pham, D. T., 2020, "Disassembly Sequence Planning Using Discrete Bees Algorithm for Human-Robot Collaboration in Remanufacturing," Rob. Comput. Integr. Manuf., **62**(1), p. 101860.
[33] Choux, M., Bigorra, E. M., and Tyapin, I., 2021, "Task Planner for Robotic Disassembly of Electric Vehicle Battery Pack," Metals (Basel, **11**(3), pp. 1–18.
[34] Fan, J., Member, S., Zheng, P., Member, S., Li, S., and Member, S., 2022, "An Integrated Hand-Object Dense Pose Estimation Approach With Explicit Occlusion Awareness for Human-Robot Collaborative Disassembly," IEEE Trans. Autom. Sci. Eng., **10**(4), pp. 1–10.
[35] Osti, F., Ceruti, A., Liverani, A., and Caligiana, G., 2017, "Semi-Automatic Design for Disassembly Strategy Planning: An Augmented Reality Approach," Procedia Manuf., **11**(5), pp. 1481–1488.
[36] McGovern, S. M., and Gupta, S. M., 2006, "Ant Colony Optimization for Disassembly Sequencing With Multiple Objectives," Int. J. Adv. Manuf. Technol., **30**(5), pp. 481–496.
[37] Tripathi, M., Agrawal, S., Pandey, M. K., Shankar, R., and Tiwari, M. K., 2009, "Real World Disassembly Modeling and Sequencing Problem: Optimization by Algorithm of Self-Guided Ants (ASGA)," Rob. Comput. Integr. Manuf., **25**(3), pp. 483–496.
[38] Xia, K., Gao, L., Li, W., and Chao, K. M., 2014, "Disassembly Sequence Planning Using a Simplified Teaching-Learning-Based Optimization Algorithm," Adv. Eng. Inf., **28**(4), pp. 518–527.
[39] Tseng, Y. J., Yu, F. Y., and Huang, F. Y., 2011, "A Green Assembly Sequence Planning Model With a Closed-Loop Assembly and Disassembly Sequence Planning Using a Particle Swarm Optimization Method," Int. J. Adv. Manuf. Technol., **57**(9), pp. 1183–1197.
[40] Kheder, M., Trigui, M., and Aifaoui, N., 2015, "Disassembly Sequence Planning Based on a Genetic Algorithm," Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci., **229**(12), pp. 2281–2290.
[41] Liu, Z., Liu, Q., Wang, L., Xu, W., and Zhou, Z., 2021, "Task-Level Decision-Making for Dynamic and Stochastic Human-Robot Collaboration Based on Dual Agents Deep Reinforcement Learning," Int. J. Adv. Manuf. Technol., **115**(11), pp. 3533–3552.
[42] Chen, B., Xu, W., Liu, J., Ji, Z., and Zhou, Z., 2020, "Robotic Disassembly Sequence Planning Considering Robotic Collision Avoidance Trajectory in Remanufacturing," Proceedings of the 2020 IEEE 18th International Conference on Industrial Informatics (INDIN), Warwick, UK, July 20–23, pp. 494–501.
[43] Liu, Z., Liu, Q., Xu, W., Wang, L., and Zhou, Z., 2022, "Robot Learning Towards Smart Robotic Manufacturing: A Review," Rob. Comput. Integr. Manuf., **77**(1), p. 102360.
[44] Poschmann, H., Brüggemann, H., and Goldmann, D., 2020, "Disassembly 4.0: A Review on Using Robotics in Disassembly Tasks as a Way of Automation," Chem.-Ing.-Tech., **92**(4), pp. 341–359.