Journal of Intelligent Manufacturing
https://doi.org/10.1007/s10845-024-02527-8

Deep reinforcement learning-based energy-aware disassembly planning for end-of-life products with stimuli-activated self-disassembly

Di Wang[1] · Jing Zhao[2] · Muyue Han[3] · Lin Li[1]

Received: 14 April 2024 / Accepted: 4 November 2024

Abstract
Remanufacturing stands as a cornerstone strategy for end-of-life (EOL) product management, playing a vital role in fostering a circular economy. Despite its significance, the widespread implementation remains difficult, mainly due to challenges such as labor-intensive operations, diminished quality, and time-consuming processes involved in component disassembly. A potential solution emerges in stimuli-activated self-disassembly, offering a non-destructive pathway that encourages seamless human–machine collaboration. This innovative approach facilitates the simultaneous disassembly of multiple components, reducing damage, labor costs, and energy consumption. Notably, limited studies have addressed real-time disassembly planning (DP), especially within self-disassembling workstations. Our research aims to maximize disassembly profit and energy recovery by optimizing disassembly sequences, EOL options, and a hybrid scheme that combines manual and self-disassembly operations. We propose an advanced deep reinforcement learning (DRL) algorithm that incorporates an innovative loss function, a revised training scheme, and parameter embedding to generate the Pareto frontier. Additionally, we propose a compact product representation that captures dynamics and uncertainties, such as product type variations, missing components, potential disassembly failure, and stochastic product quality. The effectiveness of our approach is demonstrated through a case study involving a TV disassembly line, benchmarked against six baselines. Furthermore, a sensitivity analysis is conducted to elucidate the impact of labor expenses and hybrid disassembly schemes on the ultimate profit recovery.

Keywords Deep reinforcement learning · Disassembly planning · End · Of · Life management · Multi · Objective optimization · Stimuli · Activated self · Disassembly

List of Symbols

x_{ij}	Whether the component i is the j_{th} disassembled component
z_{ik}	Whether component i adopts the k_{th} EOL option
d_i	Whether component i is in the stimuli-activated self-disassembly workstation
$H(i, h)$	The accessibility of component h after removing component i
$\overline{C}_1$	The minimum quality requirements threshold for the reuse
$\overline{C}_2$	The minimum quality requirements threshold for the remanufacturing
$\overline{C}_3$	The minimum quality requirements threshold for the recycling
v_j^c	The market value per unit of material j
R	The revenue of recovering a product
RC	The sum of components' EOL operation costs
C_L	The labor cost per hour
$\widetilde{T}_i$	The manual disassembly time of component i
$\widetilde{T}_S$	The shape-morphing time of the stimuli-activated self-disassembly workstation
f_1	The recovered revenue

✉ Lin Li
linli@uic.edu

[1] Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, IL 60607, USA

[2] School of Engineering, Pennsylvania State University, Erie, PA 16563, USA

[3] Department of Industrial and Systems Engineering, North Carolina A&T State University, Greensboro, NC 27411, USA

Published online: 28 November 2024

E_i	The saved energy of component i
E_{ik}^c	The energy consumption for EOL processing k
E_i^m	The embodied energy in materials of component i
$\psi_{m,i}$	The mass of component i
$c_{p,i}$	The specific heat capacity of the component i
ε_i	The heat transfer coefficient
c_i	The quality of component i
β_0	The influence of the disassembly damage loss on components' quality
r_{i1}	The values of component i being processed for the reuse
r_{i2}	The values of component i being processed for the remanufacturing
r_{i3}	The values of component i being processed for the recycling
r_{i4}	The values of component i being processed for the landfill
o_i	The component i has o_i number of materials
u_j	The units of material j
DC	The disassembly cost
C_S	The cost of the stimuli-activated self-disassembly workstation per hour
$\widetilde{T}_{ij}$	The sequence-dependent setup time between component i and component j
r_{ik}^c	The processing cost of component i with EOL option k
β_1	The EOL processing cost discount in the stimuli-activated self-disassembly workstation
f_2	The recovered energy
E_i^e	The embedded energy of component i
E_i^s	The energy consumption for thermally triggered shape morphing of component i
β_2	The discount in E_{ik}^c of the stimuli-activated self-disassembly workstation
$\psi_{a,i}$	The surface area of component i
$\Delta\tau_s$	The difference between the designated operating temperature and ambient temperature in shape morphing and recovery of component i
v_j^{e}	The embedded energy for material j per unit

Introduction

Manufacturing industry accounts for a significant 33% of global energy consumption (Yun et al., 2023). With the escalating concerns on global resource scarcity and environmental deterioration, there is a growing focus on the potential value recoverable from the vast quantities of EOL products. These products, if not properly managed, would otherwise accumulate in landfills. (Gao et al., 2021). Notably, many EOL products retain functionality and considerable salvage potential despite signs of age or wear. Common EOL options include reuse, remanufacturing, recycling, and landfill, depending highly on the returned products' quality (Qu et al., 2023). Despite its importance, widespread implementation remains difficult, mainly due to challenges such as labor-intensive operations, diminished quality, and time-consuming processes involved in component disassembly (Liu et al., 2020a, 2020b). For example, the disassembly process requires high-skill work to keep line efficiency and component integrity (Yin et al., 2023). To address such concerns, extensive research efforts are required focusing on disassembly planning from various perspectives, including process efficiency, resource allocation, environmental impacts, advanced solving methods, etc.

Recent studies have advanced the practical implementation of remanufacturing through various approaches, including the human-cyber-physical framework (Lou et al., 2024), robotic disassembly line balancing (Liu et al., 2020a, 2020b), and supply chain network optimization (Yolmeh & Saif, 2021). Despite these advancement, remanufacturing still faces challenges such as lengthy disassembly time (Cong et al., 2019; Han et al., 2023) and economic inefficiencies associated with manual disassembly (Parsa & Saadat, 2021). Additionally, mass personalization (Lou et al., 2024) and unpredictable variables often disrupt processes (Gao et al., 2021). One promising solution to these issues is stimuli-activated self-disassembly, which integrates shape memory materials into conventional additive manufacturing (AM), introducing a time dimension (Han et al., 2023). This technique, known as 4D printing, represents a transformative approach in manufacturing and demonstrates great potential for application in remanufacturing process through advancements in reversible assembly (Benyahia et al., 2023), self-assembly (Liu et al., 2021), self-evolving (Bai et al., 2022), and self-disassembly (Mazurek-Budzyńska et al., 2022).

Adopting 4D printing for stimuli-activated self-disassembly allows for concurrent disassembly of multiple sub-assemblies in response to environmental, minimizing the need for physical interaction between workers or machine tools with the product surfaces. This results in non-destructive detachment and reduces the frequency of tool changes and machine repositioning, thus streamlining the process. Consequently, a hybrid disassembly scheme that combines manual and self-disassembly workstations is expected and will be introduced in this study. Details on the stimuli-activated self-disassembly approach are discussed in Sect. "Problem formulation". In addition to the new disassembly scheme, there is a growing need for advanced approaches to address the complex DP problem associated with self-disassembly. DRL has emerged as a

highly effective method for tackling intricate NP problems, outperforming conventional DP methods that may take tens of minutes or even hours to generate satisfactory solutions through iterative searches.

This study focuses on the critical need for efficient real-time solution to DP problems within the context of stimuli-activated self-disassembly workstations. This study focuses on optimizing the profit and energy recovery by using DRL. Unlike heuristic approaches, the proposed well-trained DRL method can generate near-optimal solutions in real time, eliminating the need for repetitive problem-solving. Various uncertainties are considered in this work, including product type, missing components, disassembly failures, and product quality, to ensure robust solutions.

This study addresses the under-explored problem of self-disassembly DP problem and develops a multi-objective mathematical model that considers disassembly sequences, EOL options, and hybrid disassembly schemes. A Markov Decision Process (MDP) model is proposed with carefully designed state space, action space and reward functions. To enhance convergence, the Deep Deterministic Policy Gradient (DDPG) algorithm is modified with unique loss function and training schemes. Additionally, the performance of various neural network architectures is studied. The Pareto Frontier is further explored using a proposed parameter embedding technique. To reduce the state space and encode uncertainties, a dense grid-world product representation technique is introduced. This technique encodes geometrical relationships, sequential precedence constraints, product types, product quality, and missing components for returned EOL products. The effectiveness of the proposed approach is demonstrated through a real-world case study involving a TV disassembly line, benchmarked against six benchmark algorithms. Sensitivity analyses are conducted to evaluate the impacts of labor cost and uncertainty in product quality on profit recovery.

The significance of this work can be summarized as follows:

a) This work advances the study of real-time DP problems by exploring a hybrid disassembly scheme that integrates a self-disassembly with conventional manual disassembly workstations.
b) A multi-objective optimization approach using DRL is proposed to generate near-optimal disassembly sequences, EOL options, and hybrid disassembly schemes in real-time. This approach incorporates innovative loss functions, training schemes, and parameter embeddings.
c) Diverse uncertainties, including product type variations, missing components, potential disassembly failures, and stochastic product quality, are addressed using a compact grid-world product representation designed to reduce the state space.

The subsequent sections of this paper are structured as follows. The related work is discussed in Sect. "Related work". Sect. "Problem formulation" presents the problem definition, while Sect. "Modified DDPG-based DP with self--disassembly workstations" introduces the proposed method. The results and findings are discussed in Sect. "Numerical experiment" followed by the conclusion and future research directions highlighted in Sect. "Conclusions and future work".

Related work

In recent years, numerous meta-approaches have been proposed to address DP problems, such as Artificial Bee Colony (ABC) (Zhang et al., 2023a, 2023b, 2023c), Ant Colony Optimization (ACO) (Zha & Yu, 2014), Differential Evolution algorithm (DE) (Wang et al., 2020), and Genetic algorithm (GA) (Giudice & Fargione, 2007), etc. The focuses of these studies are not limited to optimizing recovery profit, but also areas address issues such as line balancing (Tuncel et al., 2014), energy conservation (Wang et al., 2023a, 2023b), environmental pollutant emissions (Zhang et al., 2023a, 2023b, 2023c), and resource allocation (Zhu et al., 2020).

This paper optimizes both disassembly profit and energy recovery, underscoring the significance of energy-aware remanufacturing systems, as highlighted in prior studies (Fu et al., 2024; Liang et al., 2021; Zhang et al., 2023a, 2023b, 2023c). For instance, Ren et al. investigated the trade-off between profit recovery and energy conservation for EOL products (Ren et al., 2020a, 2020b). This study extends the analysis by examining the balance between conserved energy and energy consumption across various disassembly schemes. Tian et al. investigated uncertainties in disassembly energy planning but did not consider the impact of EOL options (Tian et al., 2012). Later studies by Wang et al. incorporated EOL options but did not address real-time decision-making and cannot be directly used in the process with the self-disassembly (Wang et al., 2023a, 2023b). Wang et al. demonstrated that parallel disassembly techniques reduce energy consumption (Wang et al., 2021), which supports the hypothesis of this study that implementing self-disassembly workstations can potentially enhances energy recovery.

The disassembly process involves many uncertainties related to product quality (Bentaha et al., 2020), processing time (Slama et al., 2021), and environmental hazards (Malekkhouyan et al., 2021). Nevertheless, many of the existing methods (Fu et al., 2021) require domain knowledge. For

instance, disassembly time uncertainty for individual components is commonly represented using a normal distribution (Wang et al., 2019), while the sample average (Kim et al., 2018) provides an approximation for the variability in disassembly time.

The application of DRL has extended to conventional DP problems, with numerous studies conducted in this domain (Kong et al., 2022; Tuncel et al., 2014; Wang et al., 2022). For example, Tuncel et al. combined DQN with the Boltzmann machine to assign the probability distribution over all action spaces, aiming to minimize the number and idle time of workstations needed to operate the disassembly line (Tuncel et al., 2014). Reveliotis et al. maximized disassembly profit while considering component demand uncertainty via Q learning (Reveliotis, 2007). Wurster et al. introduced a novel DP problem, proposing a hybrid disassembly system that merges manual and autonomous workstations, solving it using a DQN approach (Wurster et al., 2022). However, these existing works primarily address conventional EOL products, excluding self-disassembly workstations. Furthermore, most existing studies focused on a restricted set of uncertainties, often neglecting energy consumption and efficiency. There is a critical demand for more comprehensive DP methods that consider a broader spectrum of uncertainty sources.

Among the early efforts on EOL product management that features stimuli-activated self-disassembly, Han et al. developed a multi-agent DRL algorithm aiming to maximize recovery profit by optimizing workstation schedules for both manual and self-disassembly tasks (Han et al., 2023). Nevertheless, their approach is confined by a predefined disassembly sequence, restricting its applicability in DP problems characterized by a high degree of product variations. Building on their findings, this study advances existing research by maximizing recovery profit and energy recovery through optimizing disassembly sequences, EOL options, and hybrid disassembly scheme, considering various types of uncertainties.

The self-disassembly workstation plays a crucial role in remanufacturing systems. One of the most significant benefits of incorporating 4D printing technology into the remanufacturing system is its self-healing capability, which can extend product lifecycles and reduce the need for frequent remanufacturing (Mishra & Behera, 2023). Furthermore, the non-destructive disassembly feature of self-disassembly workstations, as supported by (Lodha et al., 2023), enhances the remanufacturing process by minimizing material waste and preserving the integrity of components.

Problem formulation

In this section, the background is presented in Sect. "Stimuli-activated Self-disassembly" and a multi-objective optimization problem is formulated in Sect. "Problem definition" aiming to maximize the recovered energy and profit. The steps and critical definitions associated with the proposed DDPG are discussed in Sect. "Definitions of state, action, and reward function".

Stimuli-activated self-disassembly

In this study, the returned EOL products are assumed to be hybrid-manufactured, comprising both shape-morphing and conventional engineering materials. Shape-morphing sub-assemblies can be processed non-destructively at self-disassembly workstations. In contrast, for traditionally manufactured components, operators must follow sequential precedence, removing external components before internal ones. Due to the necessity of manual operation, only one component can be operated at a time, which often leads to inevitable disassembly damage. Shape-morphing structures, however, can be disassembled simultaneously using environmental triggers, without any direct interaction between workers or tools with the product surfaces, resulting in nondestructive disassembly and the acquisition of multiple sub-assemblies.

Figure 1 illustrates the differences between the two disassembly lines: manual and self-disassembly workstations. Self-disassembly shortens the process and reduces the need for additional tools, while manual disassembly is labor-intensive and requires various machines and tools. In manual workstations, work experience impacts component quality, and if adjacent components are damaged during removal, only the undamaged parts remain reusable, with defective ones relegated to recycling. In contrast, stimuli-activated self-disassembly reduces the need for skilled labor and ensures the non-destructive detachment of high-value components.

Problem definition

While self-disassembly offers several advantages, it also introduces new complexities and uncertainties into the remanufacturing system, necessitating the development of an advanced mathematical model. This model forms the basis of the MDP and supports the proposed DRL approach. The DP problems for EOL products with self-disassembly workstations are presented in Fig. 2. Unlike traditional DP problems, our model introduces an additional decision variable, the disassembly scheme. Self-disassembly allows for multiple sub-assemblies to be disassembled non-destructively, breaking the usual sequential precedence constraints and influencing the selection of EOL options. Moreover, the introduction of

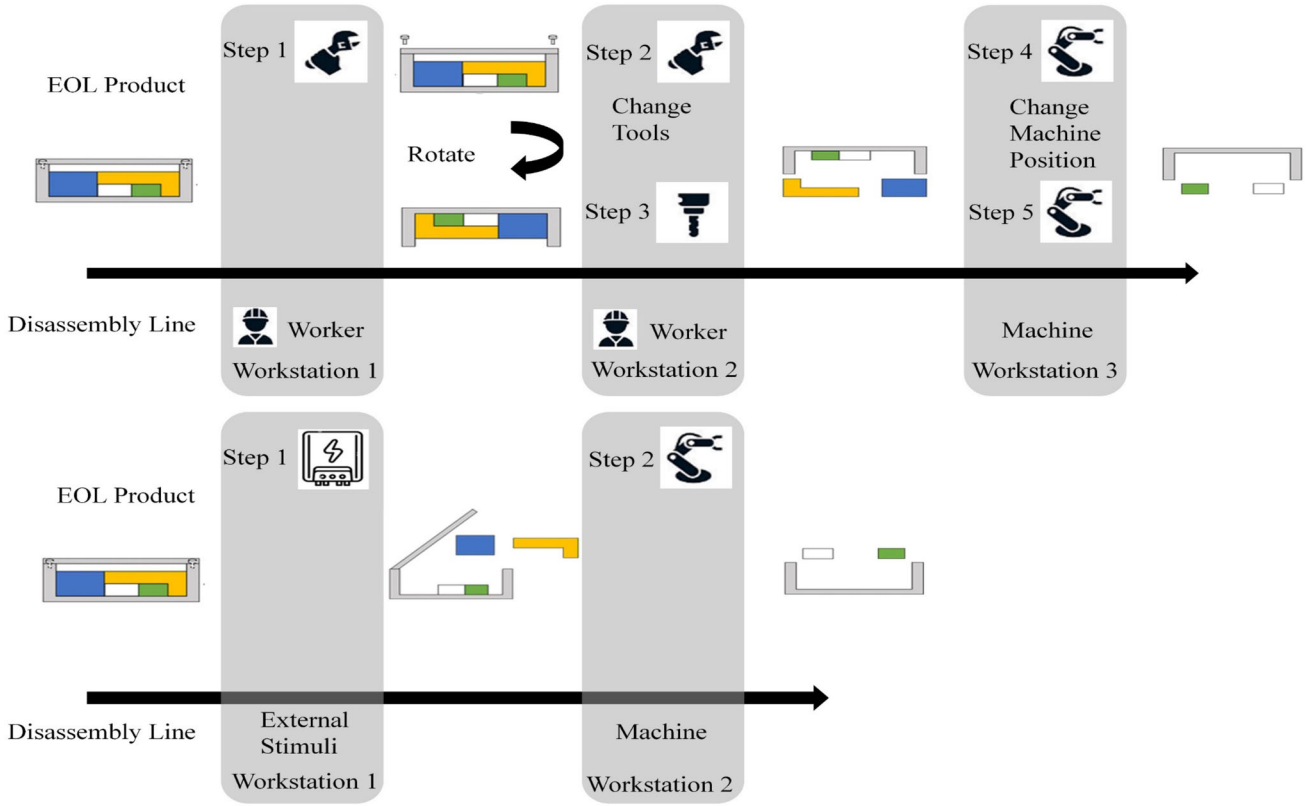


Fig. 1 Disassembly line comparisons with and without stimuli-activated self-disassembly workstations

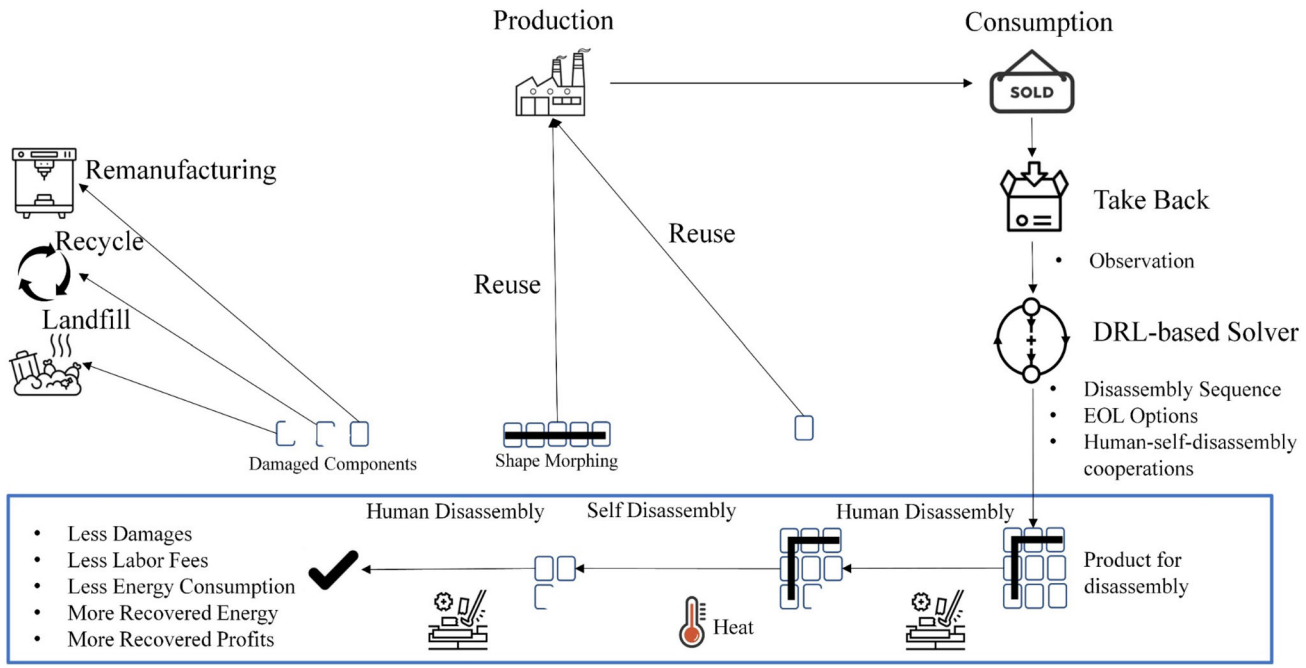


Fig. 2 Illustration of DP problems for EOL products with stimuli-activated self-disassembly workstations

self-disassembly creates new trade-offs between energy consumption, embedded energy, recovered value, disassembly cost and EOL preprocessing cost. Unlike previous studies, as discussed in Sect. "Related work", our model comprehensively addressed all mentioned factors. The impact of uncertainties will be further explored in Sect. "Definitions of state, action, and reward function".

It is assumed an EOL product consists of n components, represented by the set N. Let M and D represent the sets of components with mechanical materials and shape memory materials, respectively. Thus, $M \cup D = N$. The decision variables include disassembly sequence, EOL options, and hybrid disassembly scheme. $x_{ij} = 1$ indicates that component i is the j_{th} disassembled component; otherwise $x_{ij} = 0$. $z_{ik} = 1$ indicates that component i adopts the k_{th} EOL option, where $k = 1, 2, 3, 4$ for reuse, remanufacturing, recycling, and disposal, respectively. Furthermore, $d_i = 1$ indicates manual disassembly workstation is adopted for separating component i, while $d_i = 0$ represents the self-disassembly workstation is taken. The constraints involved are listed as follows:

$$\sum\nolimits_{j=1}^{n} x_{ij} = 1, \ \forall i \in N \tag{1}$$

$$\sum_{j=1}^{n} \frac{x_{ij}}{j} \geq \sum_{j=1}^{n} \frac{x_{hj}}{j}, \forall H(i, h) d_h d_i = 1, \ i \in N, \ h \in N \tag{2}$$

$$d_i = 1, \forall i \in M \tag{3}$$

$$\begin{cases} z_{i1} = 1, & c_i(1 - d_i\beta_0) \geq \overline{C}_1 \\ z_{i2} = 1, & \overline{C}_1 > c_i(1 - d_i\beta_0) \geq \overline{C}_2 \\ z_{i3} = 1, & \overline{C}_2 > c_i(1 - d_i\beta_0) \geq \overline{C}_3 \\ z_{i4} = 1, & \overline{C}_3 > c_i(1 - d_i\beta_0) \end{cases} \tag{4}$$

$$\sum_{j=1}^{4} z_{ij} = 1, \forall i \in N \tag{5}$$

Constraint (1) ensures that each component is disassembled only once. Constraint (2) denotes the precedence constraints for the manual disassembly workstation, and $H(i, h) = 1$ represents that component h can be taken apart after removing component i. Constraint (3) indicates that traditionally manufactured components must be processed in the manual disassembly workstation. Constraints (4) and (5) are the EOL option constraints, where c_i is the quality of component i. Components' qualities vary in different used products. $\overline{C}_1$, $\overline{C}_2$, $\overline{C}_3$ are the minimum quality requirements thresholds for EOL options and $0 < \beta_0 < 1$ measures the influence of the disassembly damage loss on components' quality.

It is assumed that the recovered values of component i for the reuse r_{i1} and the remanufacturing r_{i2} are market dependent. For the recycling option, r_{i3} equals to the market value of returned materials as shown in Eq. (6). For the landfill option, $r_{i4} = 0$.

$$r_{i3} = \sum_{j=1}^{o_i} v_j^{\text{c}} u_j \tag{6}$$

where component i has o_i number of materials; v_j^{c} is the market value per unit of material j; u_j is the units of material j. The revenue of recovering a product is shown as (7).

$$R = \sum_{i=1}^{n} \sum_{k=1}^{4} z_{ik} r_{ik} \tag{7}$$

The cost of recovery encompasses two main elements: the disassembly cost DC and the EOL processing cost RC, as detailed in Eq. (8) and (9).

$$DC = C_L\Big(\sum_{i=1}^{n} \widetilde{T}_i d_i \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{a=1}^{n-1} \widetilde{T}_{ij} d_i d_j x_{ia} x_{j(a+1)} H(a, a+1)\Big) + C_S\Big(\sum_{i=1}^{n} \widetilde{T}_S(1 - d_i)\Big) \tag{8}$$

$$RC = \sum_{i=1}^{n} \sum_{k=1}^{4} r_{ik}^{c}(1 - d_i)\beta_1 + \sum_{i=1}^{n} \sum_{k=1}^{4} r_{ik}^{c} d_i \tag{9}$$

where C_L and C_S are the labor cost per hour and the cost of self-disassembly workstation per hour; $\widetilde{T}_i$ is the manual disassembly time of component i; $\widetilde{T}_{ij}$ is the sequence-dependent setup time between component i and component j; $\widetilde{T}_S$ denotes the shape-morphing time of the 4D-printed components; r_{ik}^{c} is the processing cost of component i with EOL option k. It is assumed that the reuse option has no additional cost. In addition, β_1 ranging from 0 to 1, measures the EOL processing cost discount in the stimuli-activated self-disassembly workstation. Thus, the recovered revenue of the given product can be formulated as Eq. (10).

$$f_1 = R - DC - RC \tag{10}$$

The saved energy of component i is formulated as Eq. (11).

$$E_i = \begin{cases} E_i^e - E_i^s(1 - d_i) & z_{i1} = 1 \\ E_i^e - E_{i2}^c(1 - d_i)\beta_2 - E_{i2}^c d_i - E_i^s(1 - d_i) & z_{i2} = 1 \\ E_i^m - E_{i3}^c(1 - d_i)\beta_2 - E_{i3}^c d_i - E_i^s(1 - d_i) & z_{i3} = 1 \\ -E_i^s(1 - d_i) & z_{i4} = 1 \end{cases} \tag{11}$$

where E_i^e is the embedded energy of component i during the production, and E_{ik}^c is the energy consumption for EOL processing k. The energy consumption at the manual disassembly workstation can be calculated by different

phases, including standby, disassembly, direction changing, tool changing, and idling, throughout the disassembly process. An illustrative example is provided in Appendix A. Additionally, E_i^s is the energy consumption for thermally triggered shape morphing of component i, which can be calculated using Eq. (12) according to (Han et al., 2021); Momeni et al. calculate the theoretical limit of minimum energy consumption at the self-disassembly workstation and demonstrate that 4D printing can be the most energy-efficient method, owing to the minimal need for external inputs following initial fabrication (Momeni & Ni, 2018). E_i^m is the embodied energy in materials of component i, as shown in Eq. (13). β_2 is the discount in E_{ik}^c of the stimuli-activated self-disassembly workstation.

$$E_i^s = \psi_{m,i} c_{p,i} (\Delta\tau_s) + \psi_{a,i} \varepsilon_i (\Delta\tau_s) \widetilde{T}_s \tag{12}$$

where $\psi_{m,i}$ and $\psi_{a,i}$ represent the mass and surface area of component i, respectively; $c_{p,i}$ is the specific heat capacity of the component i; $\Delta\tau_s$ denotes the difference between the designated operating temperature and ambient temperature in shape morphing of component i; and ε_i signified the heat transfer coefficient.

$$E_i^m = \sum_{j=1}^{o_i} v_j^e u_j \tag{13}$$

where v_j^{e} is the embedded energy for material j per unit. Thus, the recovered energy of the given product can be mathematically expressed as Eq. (14).

$$f_2 = \sum_{i=1}^{n} E_i \tag{14}$$

Definitions of state, action, and reward function

MDP is particularly well-suited for handing real-time sequential decision-making tasks, such as generating optimal disassembly sequences, EOL options, and disassembly schemes at time steps to maximize the reward function. MDP is highly effective in managing uncertainties and dynamics, which is why we chose to implement this model. Furthermore, DRL enhances the efficiency of solving MDP problems by leveraging the representational power of neural networks, particularly when dealing with large state and action spaces. This integration of DRL is the core motivation behind our approach, enabling us to address complex decision-making scenarios in remanufacturing systems. In DRL, DP problem is modeled as a MDP characterized by four essential parts, the state space S, the action space A, the stationary transition function $p(s_{t+1}|s_1, a_1, s_2, a_2, \ldots, s_t, a_t) = p(s_{t+1}|s_t, a_t)$, and the reward function r.

States serve as the inputs of DRL. Given the complexity of our problem and the numerous uncertainties, such as product type variations, missing components, potential disassembly failures, and stochastic product quality, the state space can expand significantly, often exponentially with the increase in disassembly sequences. Efficiently capturing geometric structures, product variations, quality conditions, and sequential precedence relationships poses a significant challenge. In this study, the state space is $s_t = \{G_t, M_r, M_{r^c}, M_{\overline{T}}, M_{E^e}, M_{E^m}, M_{E^c}, M_{E^s}, \beta_0, \beta_1, \beta_2, C_L, C_S, \overline{C}_1, \overline{C}_2, \overline{C}_3\}$. To be more specific, $G_t \in R^{2n}$ varies at each time step, depicting grid worlds of disassembly procedures. Details of the proposed grid world representations will be discussed in Sect. "Modified DDPG-based DP with self-disassembly workstations". M_r, M_{r^c}, $M_{E^c} \in R^{4n}$ are matrices of values r_{ik}, r_{ik}^{c}, and E_{ik}^{c}, where $i \in N$, $k \in \{1, 2, 3, 4\}$. M_{E^e}, M_{E^m}, $M_{E^s} \in R^n$ denote the matrices of values E_i^{e}, $E_i^{\mathrm{m}} and E_i^s$, where $i \in N$. $M_{\widetilde{T}} \in R^{n^2+n+1}$ refers to the values of $\widetilde{T}_i$, $\widetilde{T}_{ij}$ and $\widetilde{T}_S$, where $i \in N$, $j \in N$. The quality of the component is not observable. Sect. "Product representation" introduces a compact product representation designed to reduce the dimensionality of the state space.

Actions represent decision variables, which serve as outputs of DRL. The action space can become very large, making it challenging to explore the solution space effectively. Additionally, many solutions may be infeasible. To ensure the safety of the generated policies, these infeasible solutions must be filtered out, which directly impacts the design of the action space. In this paper, our decision variables include x_{ij}, z_{ik} andd_i, where$i \in N$, $j \in N$, $k \in \{1, 2, 3, 4\}$.$A_t = [P_{s_t}, P_{EOL_t}, P_{d_t}]$, where P_{s_t}, P_{EOL_t} and$P_{d_t} \in R^n$. P_{s_t} ranges from 0 to 1, evaluating the sequential positions of components. P_{s_t} is sorted in descending order and removed nonexistent components according to constraint (1). Scan P_{s_t} from the index$i = 0$. Check$P_{d_{ti}}$, the value at index i ofP_{d_t}. If $P_{d_{ti}} > 0.5$ and this component is 4D-printed, according to constraint (3), the stimuli-activated self-disassembly workstation is taken. Otherwise, the manual disassembly workstation is taken according to constraint (2). If constraint (2) is violated, $i = i+1$. For EOL options, we check the value $P_{EOL_{ti}}$ at the index i ofP_{EOL_t}. Put $P_{EOL_{ti}}$ into one of four bins [0, 0.25), [0.25, 0.5), [0.5, 0.75), [0.75, 1] based on its value. These four bins denote four EOL options. Besides, constraints (3) and (4) should be satisfied.

Typically, DRL learns optimal policy by maximizing the collected rewards. To generate the Pareto front, the naive idea is to adjust the importance coefficients between different objectives as $F = \alpha f_1 + (1-\alpha) f_2$. Thus, the proposed DRL should have the ability to generate a series of solutions according to a list of α values as Pareto front. A list of

predefined parameter embeddings is proposed to encode different α values, and the generation of parameter embedding is discussed in Sect. "Numerical experiment". The transition function evaluates the transition probabilities among different states.

One primary limitation of the proposed MDP is the 'curse of dimensionality,' where the state and action spaces grow exponentially with the problem size. This is a key motivation for using DDPG to solve the proposed MDP problem, as DDPG is designed to handle large state and action space effectively. Another limitation is the sparse reward issue, which results in slower convergence of DRL. This further supports the choice of DDPG, which incorporates an experience replay buffer and is known for its better convergence properties. Additionally, new training loss functions and a training scheme with local search are introduced in Sect. "Modified deep deterministic policy gradient" to mitigate the issues of sparse rewards and improve convergence.

Figure 3 presents the training and real-time inference frameworks of the proposed DDPG method. The policy and the Q value are represented by the actor network with the parameter θ and the critic network with the parameter ω. The performance objective and the gradient are shown in Eq. (15) and (16). The loss function of the critic network is demonstrated in Eq. (17) and (18). Equation (15) and (16) are used for the policy improvement process while Eq. (17) and (18) are used for the policy evaluation process. Details of the DDPG training process can be found at (Wang, 2022; D. Wang & Hu, 2023)

$$J_{\rho^{\overline{\pi}}}(\mu_\theta) \approx \int_s \rho^{\overline{\pi}}(s) Q_\omega(s, \mu_\theta(s)) ds \tag{15}$$

$$\nabla_\theta J_{\rho^{\overline{\pi}}}(\mu_\theta) \approx E_{s\sim\rho^{\overline{\pi}}}[\nabla_\theta \mu_\theta(s) \nabla_a Q_\omega(s, a)|_{a=\mu_\theta(s)}] \tag{16}$$

$$L(\omega) = E_{s_t\sim\rho^{\overline{\pi}}, a_t\sim\overline{\pi}}[(y_t - Q_\omega(s_t, a_t|\omega))^2] \tag{17}$$

$$y_t = r(s_t, a_t) + \gamma Q_{\omega'}(s_{t+1}, \mu_{\theta'}(s_{t+1})|\omega') \tag{18}$$

Based on the Temporal-Difference Learning strategy, the target actor network and the critic network are taken with parameters $\omega'$and θ', respectively.

Modified DDPG-based DP with self-disassembly workstations

Product representation

The initial step in tackling DP problems is product representation. Over the past few decades, various product representations have been proposed to for different DP approaches. These include the undirected graph, directed graph, disassembly tree, AND/OR graph (AOG), and the Petri net. However, their applicability may be limited in real-world scenarios when dealing with diverse products or those with absent components. Moreover, these models aren't directly suited for stimuli-activated self-disassembly workstations. Therefore, an ideal product representation technique should effectively capture general products, different component types, their interrelationships, and precedence constraints.

A grid-world product representation is tailored for DP problems with stimuli-activated self-disassembly workstations. In our studies, the grid-world format has proven to be denser and more efficient. For instance, using our method, a product comprising 25 components—incorporating geometrical relationships matrix with dimension 25 × 25, sequential-precedence constraints matrix with dimension 25 × 25, self-disassembly constraints matrix with dimension 25 × 25, and missing-components constraints matrix with dimension 25 × 25—can be succinctly expressed within a 5 × 5 grid world. In contrast, traditional methods would necessitate at least 4 sparse matrixes with dimension 25 × 25. As the number of components grows, these conventional methods become increasingly complex.

In this grid world, priority ranks, and neighborhood relationships illustrate various constraints. The ranks delineate disassembly order. For example, red blocks symbolize components for self-disassembly workstations, while blue blocks indicate those for manual disassembly. White blocks, on the other hand, are empty, representing absent components. The numbers within these blocks indicate their priority rank. Components in outer layers possess smaller rank numbers than those within, indicating the former should be disassembled first. Furthermore, a component can only be disassembled once its neighboring components with lower ranks have been addressed.

Using the ballpoint pen depicted in Fig. 4 (left) as an example, we demonstrate the transformation of the undirected state-diagram graph from Fig. 4 (right) into a grid world. Our objective is the complete disassembly of a used pen. As per Fig. 4 (right), when the rows are scanned, the initial appearance of each component is noted. In the second row, components E and F are identified. This gives both E and F a rank of 1, indicating they can be disassembled. Meanwhile, components A, B, C, and D emerge in the fourth row, assigning them a rank of 2. This suggests that A, B, C, and D can only be disassembled following the removal of certain components.

Neighborhood relationships are utilized to establish explicit precedence constraints, as illustrated in Fig. 5 (left). For instance, component B can only be disassembled after its neighbor, F (with a smaller rank of 1), is removed. Similarly, component A can be removed based on the prior disassembly of its neighbors, E and F. At any given step, components with

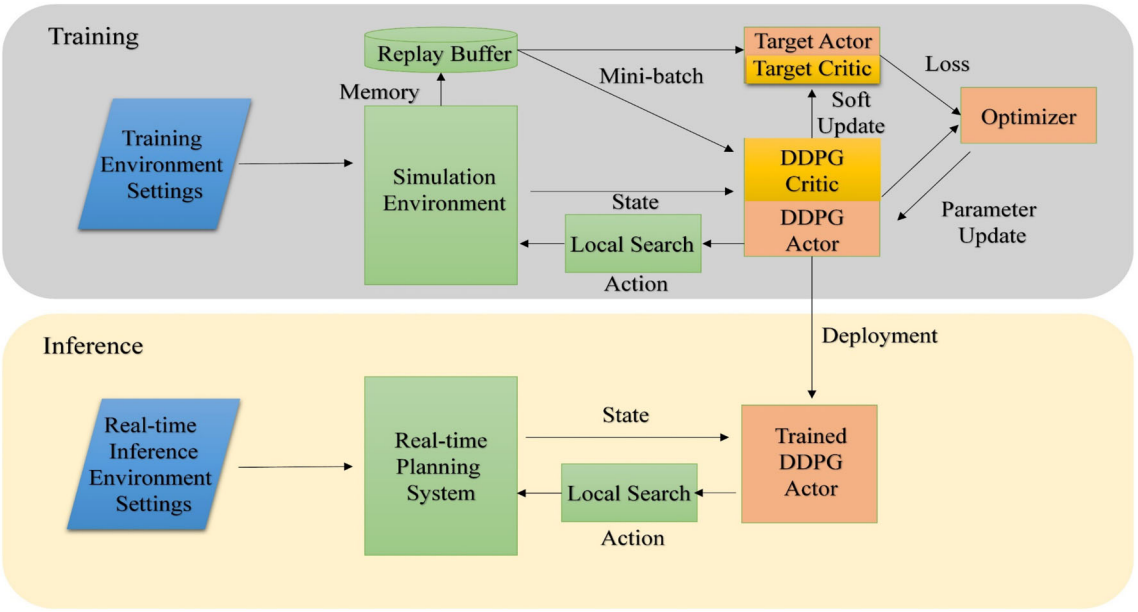
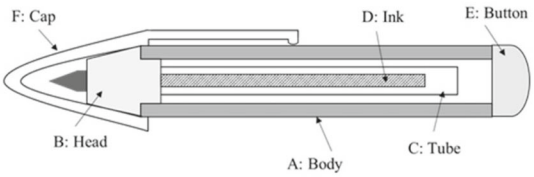


Fig. 3 The training and real-time inference frameworks of the proposed DDPG method



Row Number	Disassembly Sequence	Ranks
Row 1:	ABCDEF	[ ]
Row 2:	ABCDE,F ABCDF,E	[E^1, F^1]
Row 3:	AE,BCD,F ABCD,E,F ABF,CD,E	[E^1, F^1]
Row 4:	AE,CD,B,F A,BCD,E,F AB,CD,E,F ABF,C,D,E	[A^2, B^2, C^2, D^2, E^1, F^1]
Row 5:	AE,B,C,D,F CD,A,B,E,F AB,C,D,E,F	[A^2, B^2, C^2, D^2, E^1, F^1]
Row 6:	A,B,C,D,E,F	[A^2, B^2, C^2, D^2, E^1, F^1]

Fig. 4 Ballpoint pen example (left) and the generation of grid world (right) from the undirected state-diagram graph representation model

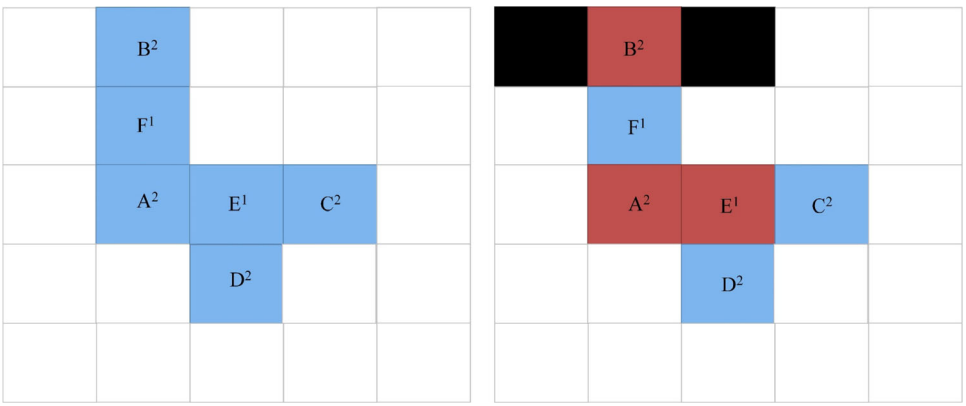


Fig. 5 Grid-world representations of a Ballpoint pen product, showing the manual disassembly scheme only (left) and the hybrid disassembly scheme (right)

a rank of 1, or those with a rank of 2 but without neighboring components of rank 1, can be disassembled.

Figure 5 (right) highlights both the advantages and challenges of stimuli-activated self-disassembly. Components A, B and E are depicted as red blocks to represent the availability of the self-disassembly scheme, where these three sub-assemblies can be non-destructively disassembled when exposed to environmental stimuli. However, black blocks enclose component B, shielding it from environmental exposure. This means that components A, B and E cannot be disassembled simultaneously unless component F, which has a lower rank of 1, is removed first. Therefore, only components A and E can be disassembled concurrently while component F remains in place.

Modified deep deterministic policy gradient

The DDPG is composed of the actor and the critic networks. The inputs of the actor network encompass both the state and predefined parameter embeddings. The predefined embedding vectors are widely used in machine learning areas (Jiang et al., 2021). The embedding vectors are randomly generated with the normal distribution of$N(\mu = 0, \sigma = \alpha)$. For simplification, a list of α values are set as[0.1, 0.2, ..., 0.9]. The heatmap of the parameter embeddings is shown in Fig. 6. For instance, the length of parameter embedding is 25, which is the same as the size of the world. The experience buffer stores the tuple (state, action, reward,α).

The architecture of the actor network is elucidated in Fig. 6. Distinct from encoders such as Multi-Layer Perceptron (MLP) (Kutschenreiter-Praszkiewicz, 2020), Convolutional Neural Network (CNN) (Lee et al., 2022), and Recurrent Neural Network (RNN) (Jeunet et al., 2021), Graph Convolutional Network (GCN) demands edge information. This not only encapsulates the connectivity details but also the specifics of the workstation type. Notably, the symbol $\oplus$ represents the concatenation operator. In the GCN encoder, the l_{th} graph convolution layer can be presented as follows: x_i^l and e_{ij}^l are the node feature vectors and edge feature vectors of the node i and edge $\widehat{ij}$. The node feature vectors, and edge feature vectors can be calculated by Eq. (19) and (20).

$$x_i^{l+1} = x_i^l + ReLU(BN(W_1^l x_i^l + \sum_{j\sim i} \varphi_{ij}^l W_2^l x_j^l)) \tag{19}$$

$$e_{ij}^{l+1} = e_{ij}^l + ReLU(BN(W_3^l e_{ij}^l + W_4^l x_i^l + W_5^l x_j^l)) \tag{20}$$

h is the size of the hidden units. ε is a small value. σ denotes the sigmoid activation function. Besides, ReLU is the rectified linear unit. BN is the batch normalization operator. Moreover, $j \sim i$ is the set of the neighborhood nodes connecting node i. More details about GCN can be found in (Ding et al., 2021). As shown in Fig. 6, the values at the first layer refer to $x^0 = x5$ and $e^0 = x6$. $x5$ and $x6$ are intermediate variables. $\varphi_{ij}^t = \sigma(e_{ij}^t)/(\sum_{j'\sim i} \sigma(e_{ij'}^t) + \varepsilon). \cdot W \in R^{h^2}$.

With local search algorithms, better policies π may be found with given states. Thus, the objective function for the actor network can be rewritten as Eq. (21).

$$\begin{aligned} J_{\rho^{\overline{\pi}}}(\mu_\theta) \approx & \int_s \rho^{\overline{\pi}}(s)\, Q_\omega(s, \alpha, \mu_\theta(s, \alpha))\, ds \\ & + E_{s_t\sim\rho^{\overline{\pi}}, a_t\sim\overline{\pi}}(I_Q\left(|M_\pi\left(\pi^l(s, \alpha)\right) - N_\pi(\mu_\theta(s, \alpha)\right)|) \end{aligned} \tag{21}$$

where $M_\pi(\pi^1(\text{s})) = [y_s, y_{EOL}, y_d]$ changes the local searched policy π^1 to target probability matrices. For example, a better policy has been found with the disassembly sequence [(3, 2), 1, 4], EOL options [1, 2, 2, 3], and disassembly options [0, 0, 1, 1]. In detail, components 3 and 2 are initially disassembled at self-disassembly workstations. This is followed by the disassembly of components 1 and 4. Furthermore, the EOL options for components 3, 2, 1, and 4 are reuse, remanufacturing, remanufacturing, and recycling. After being processed by M_π function, target probability matrices are as follows: y_s =[0.5, 0.75, 1, 0.25, 0, 0, ...], y_{EOL} =[0.125, 0.375, 0.375, 0.625, 0, 0, ...] and y_d = [1, 1, 0, 0, ...]. N_π is the max–min normalization function to make P_{s_t} compatible with y_s. Since the number of components is less than or equal to the size of the world, a matrix is adopted to mask out useless values. I_Q is the indicator function in Eq. (22). Infeasible actions will be marked out during the policy selection process to satisfy constraints.

$$I_Q = \begin{cases} 1\ r\left(s, \pi^l(s, \alpha), \alpha\right) - Q_\omega(s, \alpha, \mu_\theta(s, \alpha)) > \in \\ 0\ r\left(s, \pi^l(s, \alpha), \alpha\right) - Q_\omega(s, \alpha, \mu_\theta(s, \alpha)) \leq \in \end{cases} \tag{22}$$

where ϵ is the threshold. Here, the smooth $L1$ loss is adopted instead of the mean square error, because the Q estimation can be noisy during the training process. The smooth $L1$ loss is more robust than the $L2$ loss to abnormal values with inaccurate Q value. The loss function for the critic network can be rewritten as Eq. (23).

$$L(\omega) = E_{s_0\sim\rho^{\overline{\pi}}, a\sim\overline{\pi}}[(r(s, a, \alpha) - Q_\omega(s, \alpha, a|\omega))^2] \tag{23}$$

Leveraging the grid-world product representation, uncertainties including product types, damaged goods, and worn-out items, are integrated into the training datasets. With a well-trained DDPG actor, real-time cases or disassembly failure cases can be solved within tens of milliseconds. We select DDPG instead of DQN due to the computation complexity. For instance, a product comprising 100 components results in an action space of $100 \times 100 \times 4 \times 2$. As the number of

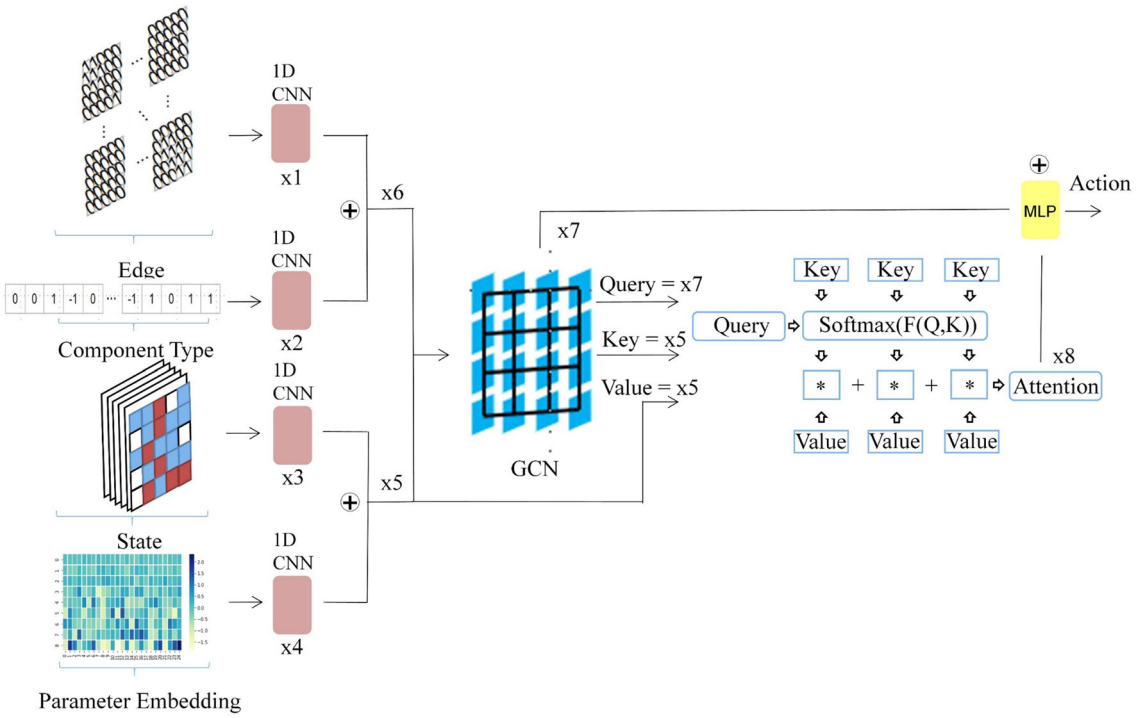


Fig. 6 The architecture of the actor network with the GCN encoder

components rises, so does the complexity. Yet, our proposed model remains less sensitive to the size of the action space. Our preference for DDPG over A2C is due to the convergence properties. A2C is an online-policy approach which requires a more extensive data pool for comprehensive training, where online data is difficult to collect.

One main limitation of the proposed DRL approach is the constrained number of non-dominated solutions on the Pareto frontier. We simplify Pareto frontier calculations through parameter embedding, enabling exploration of more non-dominated solutions with additional embedding vectors. Yang et al. integrated Pareto optimality into Q-networks, avoiding synthetic reward functions, though Q-networks are designed for limited action spaces (Yang et al., 2023). In the future, we plan to enhance our work by incorporating Pareto optimality for more accurate approximations of the Pareto frontier. The pseudocode of the proposed algorithm is shown in Algorithm 1. Wang et al. prove that the complexity of Q learning is the exponential function of the state-action space, which is intractable in the environment with large-scale state-action spaces (Wang & Hu, 2021). To have a brief view, the time complexity of the proposed method is analyzed. Assume that i episodes are needed to train the neural networks. With the replay buffer, our algorithm visits each data multiple times. Assume each data will be sampled j times at each time step. The complexity of DDPG will be $i^2 \bullet j$. Each experience trajectory will be visited a second time with an extra process/thread for our proposed method. Thus, the complexity of our method will be $2 \bullet i^2 \bullet j$.

Algorithm 1 Pseudocode of modified DDPG algorithm

Initialize critic network $Q_\omega(s, \alpha, a \mid \omega)$, target critic network $Q_{\omega'}(s, \alpha, a \mid \omega')$, actor network $\mu_\theta(s, \alpha|\theta)$, target actor network $\mu_{\theta'}(s, \alpha|\theta')$, experience reply buffer with size B, reward threshold ϵ
for epoch $i = 1$ to I **do**
  for episode $j = 1$ to J **do**
    Observe state, and generate noise-added action
    Execute action
    for α to $[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$ **do**
      Calculate rewards with $F = \alpha f_1 + (1 - \alpha) f_2$
      Add one transition $(state, action, reward, \alpha, None, None)$ to reply buffer
      Start another process to explore batches of trajectories in the experience replay buffer with a local search method. Update the corresponding tuples as (state, action, reward, α, local searched action, local searched reward)
      Sample a mini batch of trajectories from the experience replay buffer
      Update the critic network with the loss function
      Update the actor network with the performance function
    end for
  end for
end for

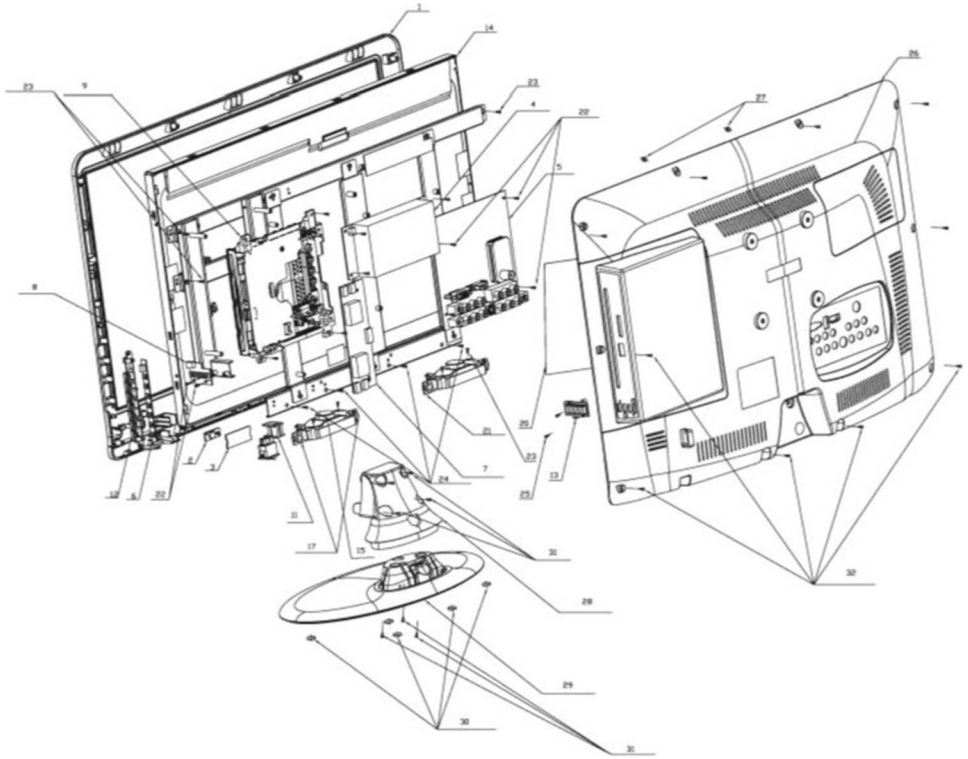


Fig. 7 Disassembly instructions for the LCD Television

Table 1 General information of the LCD TV

Components	Manual disassembly time (seconds)	Residual value (Yuan)	Mass (g)
Metal fixing plate	18.0	0.2970	15.0
Metal washer 1	2.4	0.0660	10.0
Metal washer 2	2.4	0.0660	10.0
Top metal support	21.0	0.4950	25.0
Cylindrical support 1	6.0	0.2400	30.0
Cylindrical support 2	6.0	0.1600	20.0
Toughened glass seat	13.8	0.2380	150.0
Steel plate	12.0	0.3300	50.0
Rubber gasket	6.0	0.0200	20.0
Control button	4.8	0.0100	9.2
Power switch	4.8	0.0100	5.0
Side loudspeaker	21.0	0.6000	152.0
Control receiver board	6.0	0.4000	3.0
Positive loudspeaker	15.0	0.3071	77.8
Power supply board	42.0	0.6466	118.0
Main board	42.0	0.7908	196.0
Metal board	35.4	1.2078	183.0
Metal mounting plate	109.2	4.2174	639.0
Surface frame	73.8	1.1000	270.8
LCD screen	85.2	9.6684	2900.0
Back cover	99.0	1.7904	723.8
Cover plate	1.8	0.2280	25.0
Support	2.4	0.0169	15.6

Numerical experiment

Data collection

The effectiveness of the proposed algorithm is examined through a real-world case involving the disassembly of an LCD TV, as shown in Fig. 7. The rapid technological progression in LCD TV lead to a surge in discarded LCD TV, categorizing them under WEEE (Waste Electrical and Electronic Equipment). As shown in Table 1, the reference (Ren et al., 2020a, 2020b) shares the data of an LCD TV made by China's Changhong Electronics Company Ltd.

The labor cost at the manual disassembly workstation is set as $C_L = 45$ Yuan/hour (Ren et al., 2020a, 2020b). The self-disassembly workstation requires additional device costs $C_S = 5$ Yuan/hour rather than labor costs. It is assumed that the revenue from the remanufacturing EOL option r_{i2}, $i \in N$ ranges from 60 to 100% of the residual value, which is r_{i1}, $i \in N$, ranging from 0.01 Yuan to 9.67 Yuan (Ren et al., 2020a, 2020b). The revenue from the recycling EOL option r_{i3}, $i \in N$ ranges from 40 to 80% of the residual value. Besides, the manual disassembly time $\widetilde{\widetilde{T}}_i$, $i \in N$ ranges from 1.8 s to 109.2 s. In reference to (Han et al., 2021; Zhao et al., 2020), $\widetilde{T}_S$ is assumed to be 40 s. The sequence-dependent manual disassembly time $\widetilde{T}_{ij}$, $i \in N$, $j \in N$ between two components ranges from 2 to 10 s. Moreover, the recovery cost r_{ik}^{c}, $i \in N$, $k \in \{2, 3, 4\}$ ranges from 20 to 120% of its revenue. Sutherland et al. (Sutherland et al., 2008) proved that the energy consumption in the remanufacturing process ranges from 2 to 25% of the energy used in the manufacturing process. (Ciceri et al., 2010) presented the empirical energy consumption data in the manufacturing process, which ranges from 1.3 MJ/kg to 1670 MJ/kg. E_i^{m} and E_{i3}^{c}, $i \in N$ can be found in (Ciceri et al., 2010) accordingly.

In this paper, we assume that $\beta_0 = \beta_1 = \beta_2 = 0.95$. The influence of β_0, β_1, β_2 is investigated through sensitivity analyses. Besides, the quality of each component is randomly generated with a normal distribution $N(\mu = 0.8, \sigma = 0.3)$, clipped into the range of [0.1, 1]. Moreover, $\overline{C}_1 = 0.8$, $\overline{C}_2 = 0.5$, $\overline{C}_3 = 0.3$. The size of the grid world is 5×5. The number of components is randomly generated, ranging from 15 to 25. The rank numbers are randomly assigned to each component with a maximum value of 3. The numbers of self-disassembly enabled components are randomly generated, ranging from 5 to 10.

Simulation setting

To evaluate the effectiveness of our proposed algorithm, this paper uses six heuristic methods, and three DRL methods as benchmarks. All simulated experiments in this section run on a computer with the Intel i7-6850 k CPU and the NVIDIA GeForce GTX 1080Ti. The main parameters of all methods are summarized as follows:

(1) For the DDPG method, the size of the training data set is 1,000,000. The size of the testing data set is 10,000. The size of the mini batch is 128. The default learning rate is 0.0001. Besides, the warm-up and decay strategies for the learning rate are adopted. The reward threshold $\epsilon = 0.1$.

(2) For the GCN encoder, 5 layers are deployed. Details of each layer have been illustrated above. For the RNN algorithm (Jeunet et al., 2021), two-layer Gated Recurrent Unit (GRU) units are taken. For the CNN algorithm (Lee et al., 2022), like VGG architecture, six layers are taken with varied stride sizes, including 1, 3, and 5. After each CNN layer, a batch normalization layer, a ReLU activation function, and a max-pooling layer are followed. For the MLP algorithm (Kutschenreiter-Praszkiewicz, 2020), 3 layers are taken with the ReLU activation function. The hidden unit's size is 128, and the dropout rate is 0.1.

(3) Six heuristic methods are implemented using the Python Pymoo package with the recommended parameters. For the Non-dominated Sorting Genetic Algorithm-II (NSGA2), Non-dominated Sorting Genetic Algorithm-III (NSGA3), and GA, the population size is 200. The number of offspring that are created through mating is 200. The initial population is randomly generated. The probability of crossover is 0.95, and the Eta value of the mutation is 20. For the Biased Random Key Genetic algorithm (BRKGA), the number of elites is 100. The number of offspring and the mutants are 200 and 100, respectively. The bias value is 0.7. For the DE method, the population size is 200, and the probability of the individual exchanging variable values from the donor vector is 0.3. For the Evolutionary Strategy (ES) method, the ratio of individuals surviving is 1/7, and the gamma value is 0.85. The population size is 200, and the number of offspring is 200. The default iteration value is 1000.

Simulation results

As shown in Fig. 8 (left), the Pareto frontiers of the hybrid disassembly scheme outperform those of the manual disassembly scheme. Stimuli-activated self-disassembly workstations save more energy for the same recovery profit. For example, when f_1 is around 9 Yuan, the hybrid scheme saves 3,2% more energy. Similarly, at f_2 around 2500MJ, recovery profit is 40.96% higher. Additionally, the hybrid scheme reduces disassembly time by 32.58%, from 391.96 s to 264.25 s, due to the application of self-disassembly workstations.

Figure 8 (right) illustrates the improvement ratios in recovery profit and energy achieved by the hybrid disassembly scheme. The variable f_1 represents the recovery profit obtained from the hybrid disassembly scheme, while f_{1*}

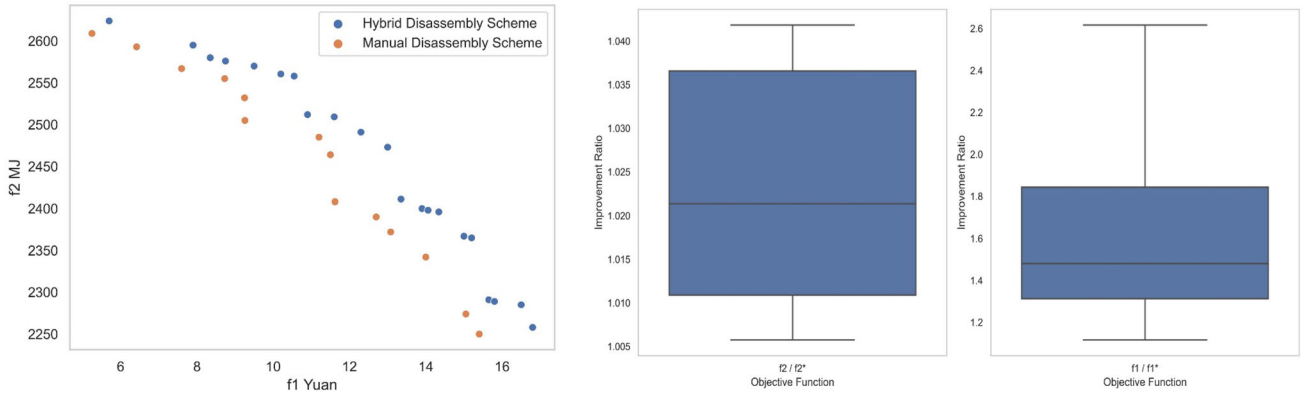


Fig. 8 Comparison of Pareto frontiers between hybrid and manual disassembly scheme (left), and the recovery profit and energy improvement ratio achieved with the hybrid disassembly scheme (right)

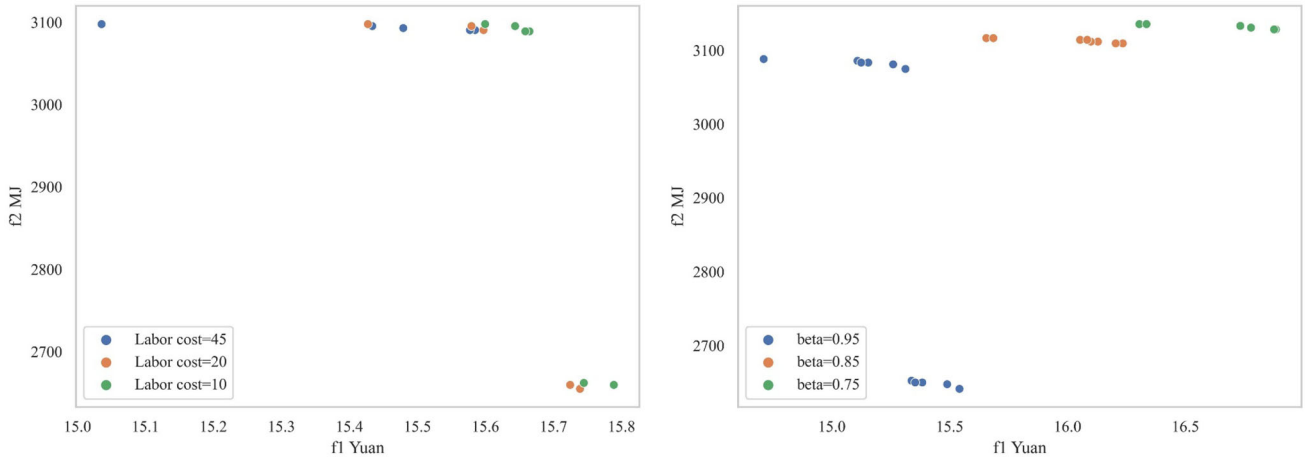


Fig. 9 Sensitivity analysis of the influence of labor cost (left) and disassembly damage (right) on Pareto frontiers

denotes the recovery profit from manual disassembly. Similarly, f_2 and f_2* correspond to the recovery energy from the hybrid and manual schemes, respectively. Interpolation is applied for a clearer depiction of the Pareto frontier. The average improvement ratios are 64.3% for recovery profit and 2.31% for recovery energy.

Figure 9 (left) studies the impacts of specific parameters like the labor cost C_L and disassembly-damage parameters β_1, β_2. As illustrated in Fig. 9 (left), expensive labor costs will cause lower recovery revenues and more applications of self-disassembly workstations. For example, when $C_L = 45$, solutions with f_2 around 2700 are nonexistent anymore. When f_2 is around 3100 MJ, f_1 with $C_L = 10$ is 1.16% more than f_1 with $C_L = 45$. As shown in Fig. 9 (right), if values of β_1, β_2 decrease, the significance of damage loss rises, concurrently boosting the likelihood of selecting self-disassembly workstations. This, in turn, results in greater recovery revenues and energy savings. For example, when $\beta_1 = \beta_2 = 0.95$, solutions with f_2 around 2650 are nonexistent anymore. Besides, f_1 with $\beta_1 = \beta_2 = 0.75$ is 8.24% more than f_1 with $\beta_1 = \beta_2 = 0.95$. f_2 with $\beta_1 = \beta_2 = 0.75$ is 1.62% more than f_2 with $\beta_1 = \beta_2 = 0.95$.

NSGA2 and NSGA3 are widely used multi-objective optimization algorithms. In the interest of fairness, the components' qualities are accessible to NSGA2 and NSGA3 methods, thereby eliminating any potential biases stemming from the stochastic and uncertainties. On the contrary, our DRL algorithm is intentionally kept uninformed about component quality. This strategic decision allows our proposed algorithm to acquire a robust understanding of quality uncertainties from its training dataset, enabling it to make informed inferences on the testing dataset. For uncertainties in product types, missing components, and disassembly failures, new cases are required to be solved repeatedly by NSGA2 and NSGA3 methods.

As shown in Table 2, the computational time for the NSGA2 method averages around 258.64 s, comparable to the disassembly time of 260 s. However, once our DRL method is trained, a new case can be solved within 0.059 s. Moreover, our algorithms provide 9 solutions at most, 6 of which are non-dominated solutions. However, NSGA2 generates 11

Table 2 Performance evaluation of algorithms on the Pareto frontiers

Algorithm	Number of nondominated solutions	CPU time (seconds)
NSGA2	11	258.64
NSGA3	12	213.75
Ours	6	0.059

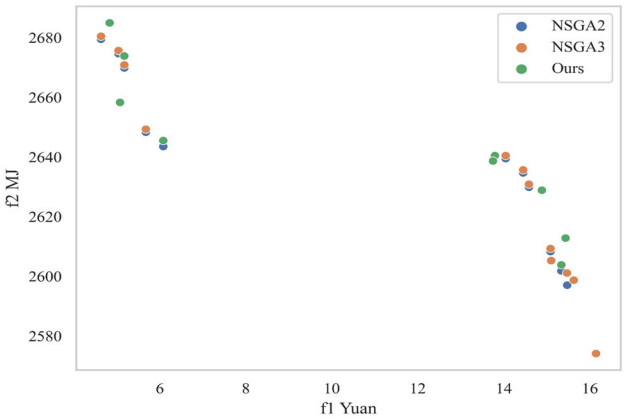


Fig. 10 Comparison of different algorithms' Pareto frontiers

nondominated solutions. Our algorithm can generate more solutions if more α values are trained in parallel. Figure 10 compares the Pareto-frontiers of our algorithm, NSGA2, and NSGA3. Unlike NSGA2 and NSGA3, solutions generated by our algorithm are not guaranteed to be in the Pareto-front line. However, compared with solutions of NSGA2, 4 better solutions are found by the proposed algorithm. For example, when f_2 is around 2610 MJ, the f_1 value of our algorithm is 2.04% more than that of NSGA2.

A comprehensive comparative analysis is conducted to examine the effectiveness of the proposed algorithm by aligning it against three DRL methods and four heuristic methods. Similar to NSGA2 and NSGA3, the quality of the components is accessible to heuristic methods, thereby eliminating any potential biases stemming from the stochastic and uncertainties. On the contrary, our DRL algorithm is intentionally kept uninformed about component quality. This strategic decision allows our proposed algorithm to acquire a robust understanding of quality uncertainties from its training dataset, enabling it to make informed inferences on the testing dataset. Table 3 provides the mean values of our methods with 10,000 testing cases and the mean values of heuristic methods with the first 500 testing cases. Besides, f_1 values are around 10 while f_2 values range around 2000. f_2 is divided by 200 to make values of f_1 and f_2 compatible. Moreover, the proposed algorithm surpasses the other methods. For instance, for $\alpha = 0.1$, our algorithm's reward is 3.51%, 1.13%, 2.33%, 10.24%, 10.49%, 14.54%, and 13.25% more than those of MLP, CNN, RNN, GA, BRKGA, DE, and ES, respectively. Higher rewards can be obtained through our proposed method compared to other benchmarks when $\alpha = 0.7$. Supplementary results for α values of 0.2, 0.4, 0.6, and 0.8 are presented in Table 4, located in the Appendix 2.

Figure 11 (left) displays rewards obtained by different algorithms across various α values. To provide a clearer visualization, we have interpolated the data to include additional α values. Our algorithm consistently achieves higher rewards than all benchmarks at every α value. Figure 11 (right) illustrates the reward improvement ratio of our algorithm over benchmarks across α values. On average, over all α values, our algorithm collects 12.08%, 17.90%, 17.13%, 13.90%, 1.15%, 2.64%, and 2.92% more rewards than the BRKGA, DE, ES, GA, CNN, RNN, and MLP benchmarks, respectively.

Heuristic methods excel in DP when exact algorithms are impractical due to the need for real-time decision-making or intricate constraints. These methods can explore a variety of solutions, providing multiple alternatives rather than a singular optimal one. Algorithms like GA and DE are known for their simplicity and extensive search capabilities. While algorithms like BRKGA can scale to accommodate larger problem sizes, scalability comes at the cost of increased computation time. Furthermore, these algorithms often necessitate tuning of parameters such as population size, mutation rate and crossover probability, a process that can be labor-intensive and requires significant domain expertise.

Deep learning, particular DRL, supports real-time decision-making in dynamic DP environments. DRL leverages extensive data for optimizing decision making process. As demonstrated, our algorithm significantly outpaces heuristic methods in solving new cases, providing solutions exponentially faster. Besides, CNN are adept at capturing spatial hierarchies and can be taken for sequential problems where input data is spatially encoded. However, CNN struggles with maintaining translation invariance, while RNN may introduce noise in the input sequences. In contrast, MLP are straightforward and ideal for scenarios with clearly defined and linear relationships, yet they may not fully harness domain knowledge in the state spaces.

As illustrated in Fig. 12, with the integration of local search, our algorithm achieves convergence more rapidly, approximately at 600,000-time steps. This local search method efficiently explores and exploits the policy space. The specialized loss function directs the algorithm to prioritize higher-potential actions, thereby reducing the computational burden of processing low-value options. Consequently, this optimized training scheme not only speeds up convergence but also enhances the overall quality of the policy learning, ensuring that the model adapts more rapidly and accurately to complex DP tasks.

Table 3 Comparison between our proposed method and state-of-art heuristic methods

α	0.1	0.3	0.5	0.7	0.9	CPU time (seconds)
Algorithm	Reward	Reward	Reward	Reward	Reward	
Ours	**17.32**	**17.00**	**16.48**	**16.10**	**15.47**	**0.059**
MLP (Kutschenreiter-Praszkiewicz, 2020)	16.74	16.42	15.92	15.24	15.23	0.055
CNN (Lee et al., 2022)	17.13	16.71	16.17	15.76	15.28	0.071
RNN (Jeunet et al., 2021)	16.93	16.31	16.09	15.79	15.16	0.088
GA (Giudice & Fargione, 2007)	15.72	14.84	14.28	13.96	13.72	1212.9
BRKGA (Biajoli et al., 2019)	15.68	15.32	14.59	14.25	13.47	1865.6
DE (Wang et al., 2020)	15.12	14.59	13.72	13.21	12.79	654.13
ES (Chien et al., 2014)	15.30	14.53	13.98	13.40	13.07	1503.5

Bold values indicate the superior results obtained by our proposed method

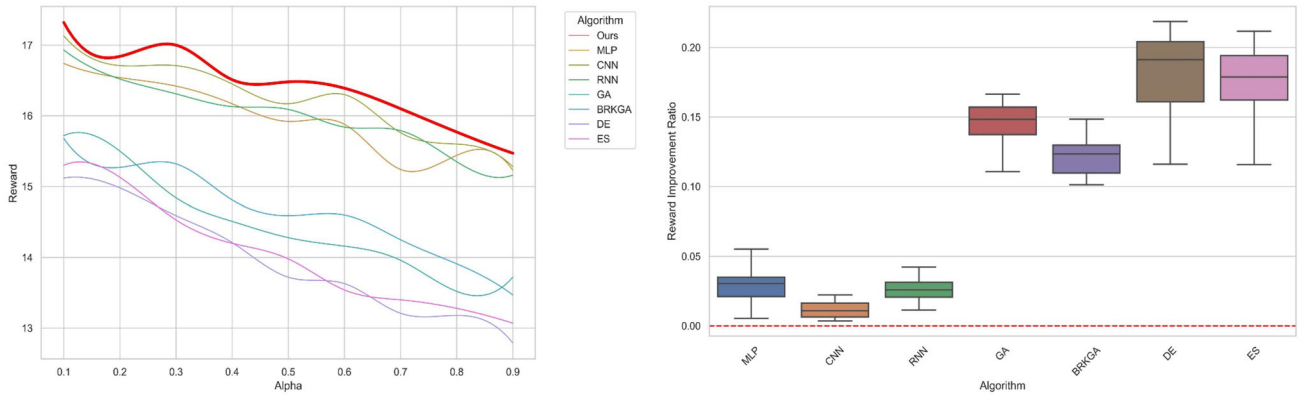


Fig. 11 Comparison of collected rewards for different algorithms across varying α values (left), and the reward improvement ratio of our algorithm over benchmarks across α values (right)

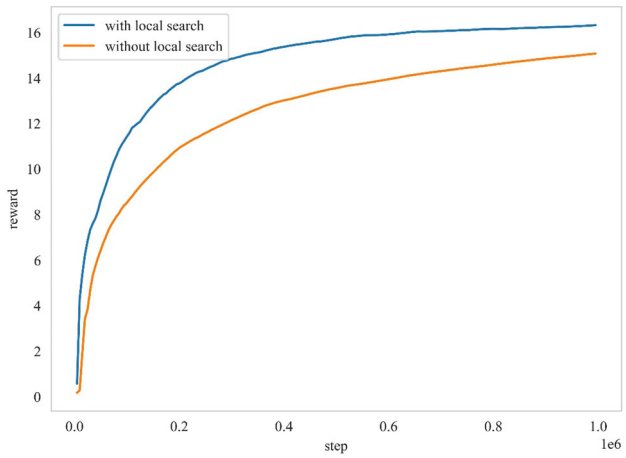


Fig. 12 Received rewards during the training with and without local search in our algorithm

This work delivers managerial insights for decision-makers on adopting the right technologies to navigate Industry 4.0/5.0 challenges. For example, in the dynamic electronics recycling industry, characterized by short product lifecycles and high-value components, integrating stimuli-activated self-disassembly technologies can markedly boost disassembly efficiency. For instance, at a facility handling diverse retired consumer electronics, implementing our work enables automatic optimization of disassembly sequences, schemes, and EOL options. This not only expedites the disassembly process at reduced costs but also lowers the risk of damaging key components like screens and batteries. Our findings highlight a route toward more sustainable manufacturing processes, recommending that decision-makers consider these innovations as they plan future technology enhancements to maintain industry leadership and environmental responsibility.

Conclusions and future work

Stimuli-activated self-disassembly facilitated by the emerging 4D printing technology provides a novel approach to sustainability and the circular economy. This approach enables simultaneous disassembly of multiple sub-assemblies,

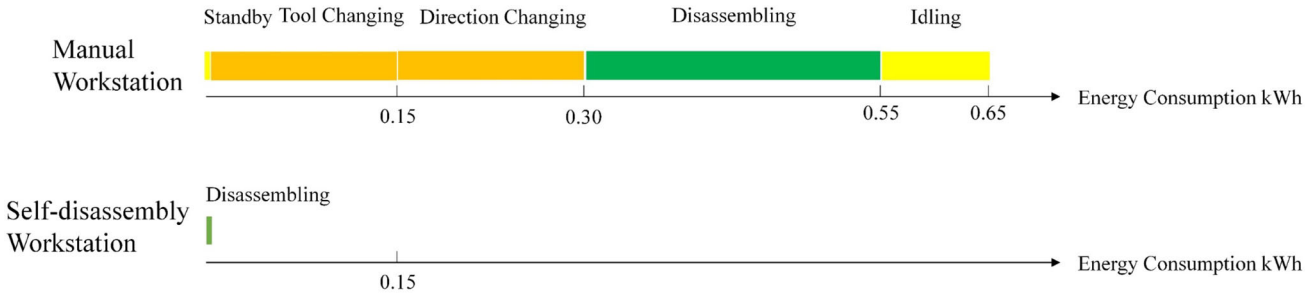


Fig. 13 Energy consumption comparison between manual and self-disassembly workstations

markedly decreasing damage, labor costs, and energy consumption. This research explores the integration of manual and self-disassembly workstations to address real-time DP challenges. The objective of this study is to maximize recovery profits and energy by optimizing disassembly sequences, EOL options, and hybrid disassembly schemes. The approach addresses several uncertainties, including variability in product types, missing components, potential disassembly failures and the quality of returned products. A comprehensive mathematical model was developed to balance critical factors including energy consumption, embedded energy, recovery value, disassembly costs and EOL preprocessing costs of self-disassembly workstations. Building on this model, a MDP is proposed with well-defined state and action space, a reward function and a compact grid-world product representation that captures the dynamics and uncertainties of the process.

To further enhance the proposed approach, the DDPG was modified with a new loss function, a revised training scheme, and parameter embedding to explore Pareto frontiers. The proposed approach was evaluated using simulation and a comprehensive case study focusing on a real LCD TV scenario. This approach was rigorously compared against six baseline methods. Additionally, sensitivity analyses were conducted to highlight the robust advantages of self-disassembly workstations. The results indicate that the hybrid disassembly scheme reduces disassembly time by 32.58% compared to manual disassembly scheme. Additionally, it achieves average improvement ratios of 64.3% in recovery profit and 2.31% in recovery energy. Furthermore, across all tested α values, the proposed algorithm outperforms the BRKGA, DE, ES, GA, CNN, RNN, and MLP benchmarks, yielding 12.08%, 17.90%, 17.13%, 13.90%, 1.15%, 2.64%, and 2.92% more rewards, respectively.

Future research will focus on the development of a multi-agent system that considers both line balance and human–robot collaboration for DP problems involving self-disassembly workstations. Additionally, the energy features of different shape-memory materials will be explored. It is also worth investigating the integration of Pareto optimization into DDPG policy generation to approximate the Pareto frontier without relying on the linear combination of objectives.

Appendix 1: An example of energy consumption at the manual and self-disassembly workstations

According to (Gao et al., 2018), the manual disassembly workstation comprises several stages: standby, disassembling, direction changing, tool changing, and idling. Here, "standby" involves loading EOL products, while "direction changing" and "tool changing" refer to adjusting the orientation of EOL products and switching out tools used for component removal. "Idling," where the workstation operates without disassembly activities, accounts for nearly 13% of total energy consumption during an 8-h work shift, as noted by (Mouzon et al., 2007). Energy usage varies across tasks. As shown in Fig. 13, Gao et al. (Gao et al., 2018) report that idle power usage is 0.10 kWh, disassembly consumes 0.25 kWh, and both direction and tool changing require 0.15 kWh each. Standby energy consumption is considered negligible.

Momeni et al. (Momeni & Ni, 2018) explore theoretical minimal energy consumption at self-disassembly workstations, highlighting significant variability based on materials and stimuli. Demoly et al. (Demoly & André, 2024) provide an simple equation for calculating energy use, exemplified by the heat capacity of a typical polymer, which, for a 100-g object heated by $40^{\circ}C$ with 80% heating efficiency, results in an energy expenditure of approximately 0.002 kWh.

Appendix 2: Supplementary Experiment Results

See Table 4.

Table 4 Supplementary results for comparison between our proposed method and state-of-art heuristic methods

α	0.2	0.4	0.6	0.8
Algorithm	Reward	Reward	Reward	Reward
Ours	**16.84**	**16.51**	**16.39**	**15.77**
MLP (Kutschenreiter-Praszkiewicz, 2020)	16.54	16.17	15.88	15.44
CNN (Lee et al., 2022)	16.71	16.45	16.30	15.60
RNN (Jeunet et al., 2021)	16.52	16.13	15.84	15.35
GA (Giudice & Fargione, 2007)	15.50	14.51	14.16	13.52
BRKGA (Biajoli et al., 2019)	15.27	14.81	14.60	13.91
DE (Wang et al., 2020)	14.98	14.21	13.63	13.18
ES (Chien et al., 2014)	15.13	14.20	13.54	13.28

Author contributions Di Wang: Conceptualization, Investigation, Writing – original draft, Software. Jing Zhao: Investigation, Visualization, Formal analysis, Writing – review & editing. Muyue Han: Conceptualization, Validation, Writing – review & editing. Lin Li: Conceptualization, Supervision, Methodology, Writing – original draft, Writing – review & editing.

Declarations

Conflict of interests The authors declare that they have no potential conflict of interest or financial conflict to disclose.

References

Bai, Y., Wang, H., Xue, Y., Pan, Y., Kim, J.-T., Ni, X., et al. (2022). A dynamically reprogrammable surface with self-evolving shape morphing. *Nature, 609*(7928), 701–708.

Bentaha, M.-L., Voisin, A., & Marangé, P. (2020). A decision tool for disassembly process planning under end-of-life product quality. *International Journal of Production Economics, 219*, 386–401.

Benyahia, K., Gomes, S., André, J.-C., Qi, H. J., & Demoly, F. (2023). Influence of interlocking blocks assembly on the actuation time, shape change, and reversibility of voxel-based multi-material 4D structures. *Smart Materials and Structures, 32*(6), 65011.

Biajoli, F. L., Chaves, A. A., & Lorena, L. A. N. (2019). A biased random-key genetic algorithm for the two-stage capacitated facility location problem. *Expert Systems with Applications, 115*, 418–426.

Chien, C.-F., Gen, M., Shi, Y., & Hsu, C.-Y. (2014). Manufacturing intelligence and innovation for digital manufacturing and operational excellence. *Journal of Intelligent Manufacturing, 25*, 845–847.

Ciceri, N. D., Gutowski, T. G., & Garetti, M. (2010). A tool to estimate materials and manufacturing energy for a product. In *Proceedings of the 2010 IEEE international symposium on sustainable systems and technology* (pp. 1–6). IEEE.

Cong, L., Zhao, F., & Sutherland, J. W. (2019). A design method to improve end-of-use product value recovery for circular economy. *Journal of Mechanical Design, 141*(4), 44502.

Demoly, F., & André, J.-C. (2024). 4D Printing: Bridging the Gap between Fundamental Research and Real-World Applications. *Applied Sciences, 14*(13), 5669.

Ding, Y., Chong, Y., Pan, S., Wang, Y., & Nie, C. (2021). Spatial-Spectral Unified Adaptive Probability Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Transactions on Neural Networks and Learning Systems*.

Fu, Y., Zhou, M., Guo, X., Qi, L., & Sedraoui, K. (2021). Multiverse Optimization Algorithm for Stochastic Biobjective Disassembly Sequence Planning Subject to Operation Failures. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Fu, Y., Zhang, Z., Liang, P., Tian, G., & Zhang, C. (2024). Integrated remanufacturing scheduling of disassembly, reprocessing and reassembly considering energy efficiency and stochasticity through group teaching optimization and simulation approaches. *Engineering Optimization*, 1–22.

Gao, Y., Lou, S., Zheng, H., & Tan, J. (2021). A data-driven method of selective disassembly planning at end-of-life under uncertainty. *Journal of Intelligent Manufacturing*, 1–21.

Gao, Y., Wang, Q., Feng, Y., Zheng, H., Zheng, B., & Tan, J. (2018). An energy-saving optimization method of dynamic scheduling for disassembly line. *Energies, 11*(5), 1261.

Giudice, F., & Fargione, G. (2007). Disassembly planning of mechanical systems for service and recovery: A genetic algorithms based approach. *Journal of Intelligent Manufacturing, 18*(3), 313–329.

Han, M., Yang, Y., & Li, L. (2021). Techno-economic modeling of 4D printing with thermo-responsive materials towards desired shape memory performance. *IISE Transactions*, 1–13.

Han, M., Yun, L., & Li, L. (2023). Deep reinforcement learning-based approach for dynamic disassembly scheduling of end-of-life products with stimuli-activated self-disassembly. Journal of Cleaner Production, 138758.

Jeunet, J., Della Croce, F., & Salassa, F. (2021). Heuristic solution methods for the selective disassembly sequencing problem under sequence-dependent costs. *Computers & Operations Research, 127*, 105151.

Jiang, T., Xie, W., Li, Y., Lei, J., & Du, Q. (2021). Weakly supervised discriminative learning with spectral constrained generative adversarial network for hyperspectral anomaly detection. *IEEE Transactions on Neural Networks and Learning Systems*.

Kim, H.-W., Park, C., & Lee, D.-H. (2018). Selective disassembly sequencing with random operation times in parallel disassembly environment. *International Journal of Production Research, 56*(24), 7243–7257.

Kong, X. T. R., Luo, H., Ballot, E., & Huang, G. Q. (2022). Driving the Physical Internet for Large-Scale Industry-Wide Deployments: A Perspective Based on Global Theoretical Frontiers. *International Journal of Production Economics*. Elsevier.

Kutschenreiter-Praszkiewicz, I. (2020). Neural network application for time standards setting in assembly and disassembly. *Journal of Machine Engineering, 20*(3), 106–116.

Lee, M.-L., Liu, W., Behdad, S., Liang, X., & Zheng, M. (2022). Robot-assisted disassembly sequence planning with real-time human motion prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 53*(1), 438–450.

Liang, P., Fu, Y., Ni, S., & Zheng, B. (2021). Modeling and optimization for noise-aversion and energy-awareness disassembly sequence planning problems in reverse supply chain. *Environmental Science and Pollution Research*, 1–13.

Liu, J., Zhou, Z., Pham, D. T., Xu, W., Cui, J., & Yang, C. (2020a). Service platform for robotic disassembly planning in remanufacturing. *Journal of Manufacturing Systems, 57*, 338–356.

Liu, J., Zhou, Z., Pham, D. T., Xu, W., Ji, C., & Liu, Q. (2020b). Collaborative optimization of robotic disassembly sequence planning and robotic disassembly line balancing problem using improved discrete Bees algorithm in remanufacturing. *Robotics and Computer-Integrated Manufacturing, 61*, 101829.

Liu, X., Wei, M., Wang, Q., Tian, Y., Han, J., Gu, H., et al. (2021). Capillary-force-driven self-assembly of 4D-printed microstructures. *Advanced Materials, 33*(22), 2100332.

Lodha, S., Song, B., Park, S.-I., Choi, H.-J., Lee, S. W., Park, H. W., & Choi, S.-K. (2023). Sustainable 3D printing with recycled materials: A review. *Journal of Mechanical Science and Technology, 37*(11), 5481–5507.

Lou, S., Zhang, Y., Tan, R., & Lv, C. (2024). A human-cyber-physical system enabled sequential disassembly planning approach for a human-robot collaboration cell in Industry 5.0. *Robotics and Computer-Integrated Manufacturing, 87*, 102706.

Malekkhouyan, S., Aghsami, A., & Rabbani, M. (2021). An integrated multi-stage vehicle routing and mixed-model job-shop-type robotic disassembly sequence scheduling problem for e-waste management system. *International Journal of Computer Integrated Manufacturing, 34*(11), 1237–1262.

Mazurek-Budzyńska, M., Behl, M., Neumann, R., & Lendlein, A. (2022). 4D-actuators by 3D-printing combined with water-based curing. *Materials Today Communications, 30*, 102966.

Mishra, A., & Behera, A. (2023). A critical review on 4D printing and their processing parameters. International Journal on Interactive Design and Manufacturing (IJIDeM), 1–31.

Momeni, F., & Ni, J. (2018). Invited Article: 4D Printing as a New Paradigm for Manufacturing with Minimum Energy Consumption. arXiv preprint arXiv:1811.12609.

Mouzon, G., Yildirim, M. B., & Twomey, J. (2007). Operational methods for minimization of energy consumption of manufacturing equipment. *International Journal of Production Research, 45*(18–19), 4247–4271.

Parsa, S., & Saadat, M. (2021). Human-robot collaboration disassembly planning for end-of-life product disassembly process. *Robotics and Computer-Integrated Manufacturing, 71*, 102170.

Qu, W., Li, J., Zhang, R., Liu, S., & Bao, J. (2023). Adaptive planning of human–robot collaborative disassembly for end-of-life lithium-ion batteries based on digital twin. *Journal of Intelligent Manufacturing*, 1–23.

Ren, Y., Jin, H., Zhao, F., Qu, T., Meng, L., Zhang, C., et al. (2020a). A multiobjective disassembly planning for value recovery and energy conservation from end-of-life products. *IEEE Transactions on Automation Science and Engineering, 18*(2), 791–803.

Ren, Y., Meng, L., Zhang, C., Zhao, F., Saif, U., Huang, A., et al. (2020b). An efficient metaheuristics for a sequence-dependent disassembly planning. *Journal of Cleaner Production, 245*, 118644.

Reveliotis, S. A. (2007). Uncertainty management in optimal disassembly planning through learning-based strategies. *Iie Transactions, 39*(6), 645–658.

Slama, I., Ben-Ammar, O., Dolgui, A., & Masmoudi, F. (2021). Genetic algorithm and Monte Carlo simulation for a stochastic capacitated disassembly lot-sizing problem under random lead times. *Computers & Industrial Engineering, 159*, 107468.

Sutherland, J. W., Adler, D. P., Haapala, K. R., & Kumar, V. (2008). A comparison of manufacturing and remanufacturing energy intensities with application to diesel engine production. *CIRP Annals, 57*(1), 5–8.

Tian, G., Liu, Y., Ke, H., & Chu, J. (2012). Energy evaluation method and its optimization models for process planning with stochastic characteristics: A case study in disassembly decision-making. *Computers & Industrial Engineering, 63*(3), 553–563.

Tuncel, E., Zeid, A., & Kamarthi, S. (2014). Solving large scale disassembly line balancing problem with uncertainty using reinforcement learning. *Journal of Intelligent Manufacturing, 25*(4), 647–659.

Wang, D., & Hu, M. (2021). Deep Deterministic Policy Gradient With Compatible Critic Network. *IEEE Transactions on Neural Networks and Learning Systems*.

Wang, D. (2022). Meta Reinforcement Learning with Hebbian Learning. In *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)* (pp. 52–58). IEEE.

Wang, D., Hu, M., & Weir, J. D. (2022). Simultaneous Task and Energy Planning using Deep Reinforcement Learning. *Information Sciences*.

Wang, D., & Hu, M. (2023). Contrastive learning methods for deep reinforcement learning. *IEEE Access*.

Wang, K., Gao, L., Li, X., & Li, P. (2021). Energy-efficient robotic parallel disassembly sequence planning for end-of-life products. *IEEE Transactions on Automation Science and Engineering, 19*(2), 1277–1285.

Wang, K., Li, X., Gao, L., & Garg, A. (2019). Partial disassembly line balancing for energy consumption and profit under uncertainty. *Robotics and Computer-Integrated Manufacturing, 59*, 235–251.

Wang, W., Tian, G., Zhang, H., Li, Z., & Zhang, L. (2023a). A hybrid genetic algorithm with multiple decoding methods for energy-aware remanufacturing system scheduling problem. *Robotics and Computer-Integrated Manufacturing, 81*, 102509.

Wang, W., Yuan, G., & Pham, D. T. (2023b). Energy-time tradeoffs for remanufacturing system scheduling using an invasive weed optimization algorithm. *Journal of Intelligent Manufacturing, 34*(3), 1065–1083.

Wang, X., Guo, S., Shen, J., & Liu, Y. (2020). Optimization of preventive maintenance for series manufacturing system by differential evolution algorithm. *Journal of Intelligent Manufacturing, 31*, 745–757.

Wurster, M., Michel, M., May, M. C., Kuhnle, A., Stricker, N., & Lanza, G. (2022). Modelling and condition-based control of a flexible and hybrid disassembly system with manual and autonomous workstations using reinforcement learning. *Journal of Intelligent Manufacturing, 33*(2), 575–591.

Yang, F., Huang, H., Shi, W., Ma, Y., Feng, Y., Cheng, G., & Liu, Z. (2023). PMDRL: Pareto-front-based multi-objective deep reinforcement learning. *Journal of Ambient Intelligence and Humanized Computing, 14*(9), 12663–12672.

Yin, T., Zhang, Z., Wu, T., Zeng, Y., Zhang, Y., & Liu, J. (2023). Multimanned partial disassembly line balancing optimization considering end-of-life states of products and skill differences of workers. *Journal of Manufacturing Systems, 66*, 107–126.

Yolmeh, A., & Saif, U. (2021). Closed-loop supply chain network design integrated with assembly and disassembly line balancing under uncertainty: An enhanced decomposition approach. *International Journal of Production Research, 59*(9), 2690–2707.

Yun, L., Wang, D., & Li, L. (2023). Explainable multi-agent deep reinforcement learning for real-time demand response towards sustainable manufacturing. *Applied Energy, 347*, 121324.

Zha, J., & Yu, J. (2014). A hybrid ant colony algorithm for U-line balancing and rebalancing in just-in-time production environment. *Journal of Manufacturing Systems, 33*(1), 93–102.

Zhang, Xin, Fang, Y., Liu, Q., & Yazdani, D. (2023). Multi-objective Robust Optimization Over Time for Dynamic Disassembly Sequence Planning. International Journal of Precision Engineering and Manufacturing, 1–20.

Zhang, Xuesong, Zhou, H., Fu, C., Mi, M., Zhan, C., Pham, D. T., & Fathollahi-Fard, A. M. (2023). Application and planning of an energy-oriented stochastic disassembly line balancing problem. Environmental Science and Pollution Research, 1–15.

Zhang, W., Zheng, Y., & Ahmad, R. (2023a). The integrated process planning and scheduling of flexible job-shop-type remanufacturing systems using improved artificial bee colony algorithm. *Journal of Intelligent Manufacturing, 34*(7), 2963–2988.

Zhao, J., Yang, Y., & Li, L. (2020). Efficiency-aware process planning for mask image projection stereolithography: Leveraging dynamic time of exposure. *Additive Manufacturing, 36*(June), 101407. https://doi.org/10.1016/j.addma.2020.101407

Zhu, L., Zhang, Z., Wang, Y., & Cai, N. (2020). On the end-of-life state oriented multi-objective disassembly line balancing problem. *Journal of Intelligent Manufacturing, 31*, 1403–1428.