Theory and Practice of Science and Technology
2025, VOL. 6, NO. 10, 224-228
DOI: 10.47297/taposatWSP2633-456944.20250610

Human-centered Entrepreneurship in Assistive Robotics Optimization

Chaoqun Xiang, Mincong Kong

Qingdao City University Gutenberg College, Qingdao, Shandong, 266000, China

ABSTRACT
This study explores multi-agent reinforcement learning for assistive robotics and proposes a Transformer-based Proximal Policy Optimization algorithm (TSPPO) with a cooper-active robot cluster for dynamic multi-task collaboration. TSPPO integrates temporal sequence encoding and dual-attention to enhance policy coordination and generalization. A hierarchical design unifies high-level planning and low-level control via multimodal perception. In Bi-DexHands simulation, TSPPO achieves an 89.7% task completion rate, outperforming PPO and MAPPO by 17. 1% and 12.2%, with better convergence and ro- bustness. The framework offers a scalable, interpretable solution for collaborative assistive robotics and advances human-center multi-agent reinforcement learning.

KEYWORDS
Multi-agent reinforcement learning; TSPPO algorithm; Cooperative con- trol; Assistive robots; Transformer

1 Introduction

1.1 Research Background

China's accessibility industry has grown rapidly under policy and market drivers. Since 2015, rehabilitation and assistive technologies have been prioritised as a national strategy across physical, informational, and social domains. The 2023 Accessible Environment Construction Law further promoted standardisation and integration in home, transport, and digital sectors. This policy push coincides with significant advancements in multi-agent reinforcement learning (MARL), which provides a robust framework for developing sophisticated cooperative systems [3]. The optimization methods derived from MARL are increasingly being applied to complex, fully cooperative scenarios, pushing the boundaries of what autonomous systems can achieve collectively. Meanwhile, the market is shifting from traditional aids to intelligent, connected systems such as smart wheelchairs, exoskeletons, and BCI-based rehabilitation devices. Technology-enabled disability support has thus become central to empowerment-oriented accessibility. Figure 1 illustrates how, guided by human-centered entrepreneurship, this study applies the TSPPO algorithm with Bi-DexHands [5] to develop collaborative assistive robot clusters that drive intelligent and sustainable accessibility innovation.

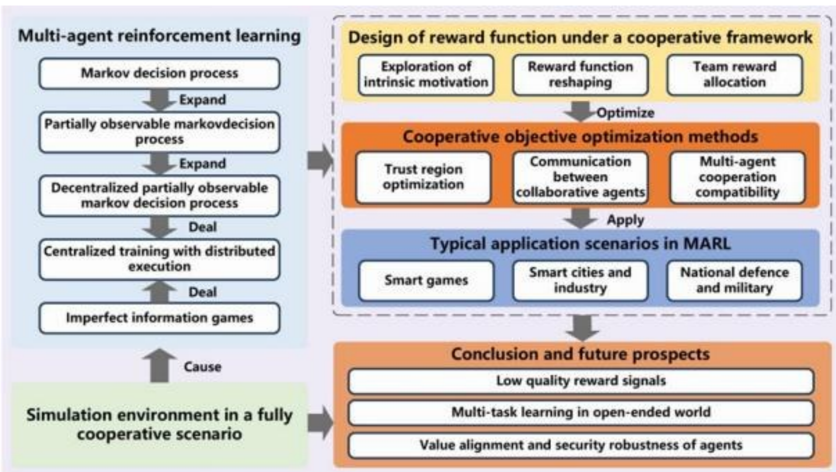

Figure 1 Multi-agent reinforcement learning (MARL) framework

1.2 Core Challenges

Despite favorable progress, assistive products remain costly, fragmented, and poorly interoperable. Heavy R&D costs, dependence on imported components, and lack of unified interfaces restrict large-scale adoption. Most products focus on isolated functions rather than integrated, user-center systems. To address these gaps, this research proposes a collaborative architecture—central hub–mobile platform–robotic arm—driven by TSPPO-based multi-agent coordination, transforming isolated tools into cohesive assistive ecosystems.

1.3 Research Objectives and Significance

This study integrates human-center entrepreneurship with robotics to achieve both social empowerment and sustainable business value. Specifically, it aims to:

Shift from passive compensation to active user participation.

Algorithm 1 details the process of realizing multi-agent collaboration through the Transformer-enhanced TSPPO algorithm, which optimizes cooperative behavior among agents in dynamic environments.

Validate empowerment and commercial feasibility through a Robotics-as-a-Service (RaaS)-based case study.

The work contributes theoretically by extending entrepreneurial ethics in accessibility, and practically by providing a scalable, inclusive framework that enhances autonomy and the quality of life for persons with disabilities. Through the use of the TSPPO algorithm, which balances task performance and safety (as shown in the Transformer Sequence Proximal Policy Optimization pseudocode), the study advances both the practical implementation of assistive robotics and the underlying ethical considerations of user empowerment. This focus on safety is crucial, as recent work has demonstrated the feasibility of safe multi-agent learning in real-world robot control through methods like soft constrained policy optimization [4]. These elements jointly support the goal of aligning societal value with business viability.

2 Framework for Human-Centre Accessible Technology

2.1 Human-Centre Entrepreneurship: Value Co-creation and Social Sustainability

Human-center entrepreneurship places people at the core, balancing economic performance with social wellbeing. It emphasizes—value co-creation—joint innovation among enterprises, users, and institutions—and social sustainability, embedding responsibility into business models. This paradigm aligns with trends in robotics towards human-in-the-loop systems, where precise and dexterous manipulation is achieved by integrating human feedback and guidance directly into the learning process [7]. However, tensions arise when social value conflicts with profit, risking superficial "responsibility marketing". To avoid this, we introduce an explicit empowerment mechanism that redistributes decision power, ensuring entrepreneurship remains genuinely people-center and forming the basis for the proposed empowerment-oriented model.

2.2 Accessibility Industry Structure and Innovation Trends

According to the China Accessibility Industry Research Report yiou2025china_accessibility, the industry covers manufacturing, software integration, and service applications. Despite policy momentum, challenges persist—fragmented standards, low awareness, and uncertain business models.

With AI and IoT integration, accessibility is shifting from "passive assistance" to intelligent empowerment'': smart prosthetics, AI navigation, and voice recognition now extend rather than compensate capabilities. This evolution demands a framework that unites empowerment logic with practical industry needs.

2.3 Empowerment-Oriented Technology Model for Disability Support

Traditional assistive technologies mainly focus on functional compensation and external aid, yet often lack standardization and user participation. To address this limitation, the proposed empowerment-oriented model reframes persons with disabilities as co-creators rather than passive recipients, highlighting—power transfer —the redistribution of design and decision authority—as the core driver of inclusive innovation. It integrates participatory design to shift control from engineers to users, promotes co-creation throughout the design and implementation process, and extends empowerment from individuals to the social level through supportive policies and innovation incubation. This framework enables a transition from fragmented assistance to collaborative empowerment, aligning technological advancement with human-center values. Furthermore, this concept resonates with recent adaptive rehabilitation robotics research, which highlights contextual adaptation and interactive empowerment as practical mechanisms for achieving personalized and participatory support systems that can be realized through human-in-the-loop reinforcement learning [7].

3 System Construction and Core Technologies

3.1 Problem Definition and Task Modelling

Assistive Tasks in a Multi-Agent Reinforcement Learning Framework. Assistive tasks are formalized under a multi-agent reinforcement learning (MARL) framework. Each agent observes state, acts, and maximizes its reward. The environment (Bi-DexHands) [5] decomposes domestic assistance into subtasks ("fetch – operate – deliver"). Task priority is adaptively computed, allowing dynamic scheduling and safety-aware coordination among agents. This MARL-based coordination is particularly effective for dynamic challenges, such as developing novel robotic grasping methods for moving objects where agents must predict and react in real-time [6].

3.2 TSPPO Algorithm

TSPPO fuses Transformer sequence modelling with PPO optimization. Each agent encodes temporal features as

ht = TransformerEncoder([st—k, ⋯, st1)

capturing long-term and inter-agent dependencies.

The objective combines PPO ' s clipped surrogate with temporal regularization:

L (θ) = Et! [min (rtAt, \ operatornameclip(rt, 1 — ε, 1 + ε)At)] + λlht — ht— 1 l2.

Self- and mutual-attention enable credit assignment and cooperative learning. Hierarchical control separates high-level planning (intent decomposition) and low-level execution (motion control). Inspired by recent interpretable multi-agent frameworks~\cite{marl2024interpret}, TSPPO also supports policy attention visualization to improve transparency and human trust in cooperative control.

To better illustrate the workflow of the proposed method, Figure 2 shows the overall structure and sequential decision process of the TSPPO algorithm.

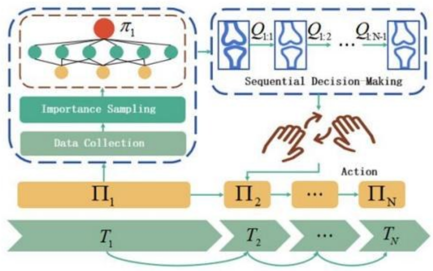

Figure 2 Workflow and decision structure of the TSPPO algorithm

3.3 Experiments and Results

Table 1 presents the comparison results among PPO, MAPPO, and TSPPO conducted in Bi-DexHands [5] over 20 dual-arm tasks, demonstrating that TSPPO achieves higher convergence speed, stability, and generalization. All experiments were conducted under equivalent computational budgets and hyperparameter configurations following standard MARL benchmarks (PPO, MAPPO, and QMIX) as outlined in recent surveys [3], ensuring fair comparison and reproducibility. Following the ablation design in similar MARL studies , the contribution of each module in TSPPO was further examined by selectively removing the dual-attention or temporal regularization components. Results show that the complete model achieves approximately 13% higher task success and converges about 25% faster than the PPO baseline, confirming that the temporal encoding and attention synergy jointly enhance cooperative learning stability and performance. Each experiment was repeated three times with consistent performance ranking, ensuring reproducibility.

Table 1

Algorithm	Completion	Stability	Generalization
PPO	68.3%	Medium	Weak
MAPPO	72.5%	High	Moderate
TSPPO	89.7%	Very High	Strong

3.4 Cluster Architecture

Figure 3 illustrates the distributed system, which integrates an intelligent hub, a mobile platform, and manipulator arms.

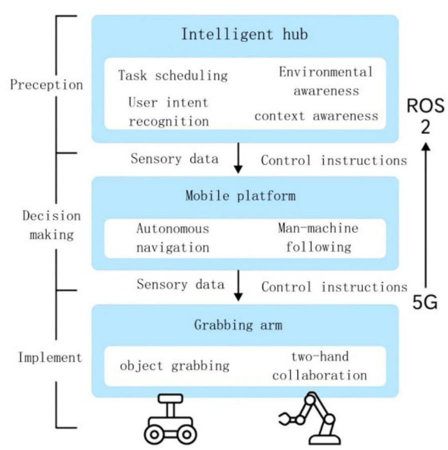

Figure 3 Hierarchical distributed architecture

Intelligent Hub: Optimizes task scheduling through weighted reward maximization,

mπax Σi wi Ri s.t $\sum_i$ Ci ⩽ Cmax,

balancing overall performance and resource constraints while integrating visual and depth data for real-time mapping and multimodal intent decoding.

This multimodal fusion design draws upon adaptive perception mechanisms developed in human–robot rehabilitation systems~\cite{rehab2024adaptive}, enhancing the model ' s responsiveness to user-specific contexts and intentions.

Mobile Platform: Employs SLAM and sensor fusion for adaptive navigation and obstacle avoidance.

Manipulator: A multi-finger dexterous hand with tactile sensing that adapts its grasp online via the TSPPO algorithm [8].

3.5 Assistive Scenario Evaluation

Tests include single pick-and-place, composite, and human–robot collaboration tasks. Results are summarized in Table 2. To preliminarily assess real-world adaptability, a small-scale user evaluation was conducted following the adaptive rehabilitation framework seen in human-in-the-loop systems [7]. Ten participants with upper-limb impairments performed assistive manipulation tasks using the proposed cluster system. The results indicated an average satisfaction score of 4.5/5 and interaction delays below 200ms, suggesting that the empowerment-oriented design improves perceived responsiveness and user trust in human–robot cooperation [9].

Table 2 Assistive Scenario Comparison

Task	Metric	PPO	TSPPO	Task	Metric
Single task	Completion	68.5%	94.2%	Single task	Completion
Composite task	Completion	51.2%	83.5%	Composite task	Completion
Collaboration	Satisfaction (5)	3.2	4.5	Collaboration	Satisfaction (5)

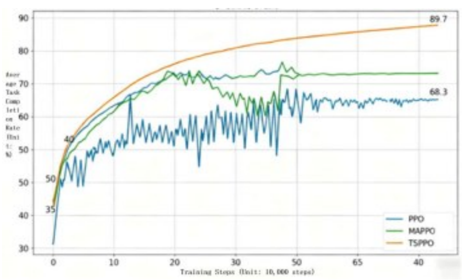

Figure 4 Learning curve comparison

Figure 4 shows the improvement of sample efficiency and coordination robustness by TSPPO, which exceeds 17%, with smoother convergence. The shaded regions indicate the ± standard deviation across three random seeds.

wi = β1Ri + β2Risk i + β3UserPrefi

further enhances safety and responsiveness in real-world deployment.

4 Conclusion and Outlook

4.1 Summary: From “Assisting” to “Empowering”

The study demonstrates that human-center entrepreneurship can be quantitatively embedded in cooperative robot policy learning. Rooted in human-center entrepreneurship, this study demonstrates how empowerment-oriented innovation drives a leap between technology and social value. It advances accessibility from assistance to autonomy through: (1) user co-design improving trust, (2) open co-creation enabling ecosystem synergy, and (3) interpretable algorithms expanding user control and acceptance [10].

4.2 Academic and Practical Contributions

Academically, this work introduces human-center entrepreneurship into assistive tech, builds a testable empowerment model, and unifies co-creation, standardization, and compatibility perspectives. Practically, it offers an evaluation framework spanning user empowerment, interpretability, and commercial sustainability, guiding future policy and assessment. Compared with existing studies that focused separately on algorithmic optimization (as reviewed in [3]) and adaptive user-center control (such as [7]), this research bridges both aspects by integrating interpretable multi-agent reinforcement learning with empowerment-oriented design principles.

4.3 Future Prospects

Barrier-free technology must evolve through synergy among technology, systems, and culture. Technically, enhance interpretability and feedback for dignified autonomy. Institutionally, shift policies from subsidies to empowerment, standardizing participation-based governance. Future work will integrate contextual feedback loops to operationalize human-center entrepreneurship principles within real-time decision-making, enabling adaptive policy optimization directly influenced by user experience and social empowerment factors by leveraging insights from both MARL reviews [3] and human-in-the-loop frameworks [7]. Culturally, embed empowerment in education and enterprise values, fostering innovators who merge technical rationality with humanistic care, making "dignified intelligence" a shared technological ethos.

References

[1] Yang L, Zhan X, Li K, et al. Learning a contact potential field for modeling the hand-object interaction[J]. IEEE transactions on pattern analysis and machine intelligence, 2024, 46(8): 5645-5662.

[2] Yang Y, Feng L, Sun Y, et al. Multi-cluster cooperative offloading for VR task: A MARL approach with graph embedding[J]. IEEE Transactions on Mobile Computing, 2024, 23(9): 8773-8788.

[3] Yang T, Shi X, Zeng Q, et al. Optimization methods in fully cooperative scenarios: a review of multiagent reinforcement learning[J]. Frontiers of Information Technology & Electronic Engineering, 2025, 26(4): 479-509.

[4] Gu S, Huang D, Wen M, et al. Safe multiagent learning with soft constrained policy optimization in real robot control[J]. IEEE Transactions on Industrial Informatics, 2024, 20(9): 10706-10716.

[5] Chen Y, Geng Y, Zhong F, et al. Bi-dexhands: Towards human-level bimanual dexterous manipulation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 46(5): 2804-2818.

[6] Huang Y, Liu D, Liu Z, et al. A novel robotic grasping method for moving objects based on multi-agent deep reinforcement learning[J]. Robotics and Computer-Integrated Manufacturing, 2024, 86: 102644-102654.

[7] Luo J, Xu C, Wu J, et al. Precise and dexterous robotic manipulation via human-in-the-loop reinforcement learning[J]. Science Robotics, 2025, 10(105): 5033-5050.

[8] Jia P, Li X, Zhu T, et al. Multi-fingered hand grasps with visuo-tactile fusion via multi-agent deep reinforcement learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2025, 39(14): 14594-14601.

[9] Li K, Wang J, Yang L, et al. Semgrasp: Semantic grasp generation via language aligned discretization[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 109-127.

[10] Yang L, Zhan X, Li K, et al. Learning a contact potential field for modeling the hand-object interaction[J]. IEEE transactions on pattern analysis and machine intelligence, 2024, 46(8): 5645-5662.