GSC Advanced Research and Reviews

eISSN: 2582-4597 CODEN (USA): GARRC2

Cross Ref DOI: 10.30574/gscarr

Journal homepage: https://gsconlinepress.com/journals/gscarr/

(REVIEW ARTICLE)

Check for updates

Decentralized reinforcement learning collectives advancing autonomous automation strategies for dynamic, scalable and secure operations under adversarial environmental uncertainties

Olabode Soetan [1,*] and Juwon Kehinde Olowonigba [2]

[1] *Retail Operations Coordination and Analytics, Access Bank Plc, Nigeria.*

[2] *Department of Business Administration, Nexford University, USA.*

GSC Advanced Research and Reviews, 2021, 09(03), 164-183

Publication history: Received on 10 November 2021; revised on 24 December 2021; accepted on 30 December 2021

Article DOI: https://doi.org/10.30574/gscarr.2021.9.3.0294

Abstract

The increasing complexity of modern industrial, financial, and socio-technical systems has outpaced the capabilities of traditional automation and centralized decision-making frameworks. As infrastructures become more distributed and interconnected, they face heightened vulnerabilities from adversarial environments, ranging from cyberattacks to unexpected disruptions in physical operations. Reinforcement learning (RL) has emerged as a promising paradigm for adaptive decision-making, yet single-agent or centrally coordinated RL models often struggle with scalability, robustness, and trust in hostile conditions. This study proposes a paradigm of decentralized reinforcement learning collectives, where multiple agents operate autonomously while coordinating decisions through distributed consensus mechanisms. By leveraging decentralized architectures, these collectives avoid single points of failure, enhance resilience against adversarial manipulations, and support dynamic adaptation in uncertain environments. The integration of cryptographic coordination and distributed consensus ensures secure interactions between agents, enabling collaborative optimization without reliance on centralized controllers. Moreover, the collective learning process allows agents to share policy insights while maintaining independence, fostering scalability in large and heterogeneous networks. Case illustrations demonstrate how decentralized RL collectives can advance autonomous automation strategies across sectors such as industrial control, smart grids, logistics, and defense systems, where adversarial uncertainty is persistent. The framework highlights pathways for embedding interpretability, accountability, and operational fairness, ensuring that scalability does not compromise trust. By bridging reinforcement learning with decentralized governance, this approach contributes a foundational model for secure, dynamic, and resilient automation in next-generation critical infrastructures.

Keywords: Decentralized Reinforcement Learning; Autonomous Automation; Adversarial Environments; Collective Intelligence; Secure Operations; Scalable Decision-Making

1. Introduction

1.1. Context: Automation in Adversarial and Uncertain Environments

Automation is increasingly deployed in environments characterized by high uncertainty, adversarial dynamics, and rapid change. Examples include autonomous vehicles navigating unpredictable traffic, unmanned aerial systems operating in contested airspace, and robotic swarms conducting disaster response missions. These settings demand adaptive decision-making under incomplete or noisy information, where traditional rule-based control is insufficient [1]. Reinforcement learning (RL) has become a central framework for enabling adaptive policies, as agents learn optimal behaviors through trial and error.

* Corresponding author: Olabode Soetan

Adversarial conditions, however, amplify complexity. In cyber-defense, automated systems must adapt to evolving threats, while in field robotics, changing terrain and unforeseen hazards continually test resilience [2]. These environments require not just efficiency but robustness agents must withstand perturbations, deliberate interference, and shifting dynamics.

Uncertainty also arises from multi-agent interactions. When multiple autonomous entities operate simultaneously, competition, cooperation, and emergent behaviors introduce new unpredictability's [3]. In adversarial domains such as defense or security, agents face deliberate opposition from intelligent adversaries, making the challenge more akin to dynamic games than static optimization.

Thus, automation in adversarial and uncertain contexts requires learning frameworks that are scalable, robust, and capable of distributed coordination. The shortcomings of centralized methods in such environments highlight the need for decentralized learning paradigms [4].

1.2. Limits of Centralized RL Approaches in Scalability and Robustness

Centralized reinforcement learning approaches have historically dominated applications where a single agent or a central controller dictates system policy. While effective in controlled settings, these approaches encounter severe limitations in adversarial and uncertain environments. The computational burden of maintaining global state information grows exponentially with the number of agents, creating scalability bottlenecks [5]. In large-scale domains such as swarm robotics or networked sensors, central learners become communication and processing choke points.

Robustness is another limitation. Centralized systems introduce single points of failure: if the central controller is compromised, the system collapses. This vulnerability is especially problematic in adversarial contexts, where attackers may deliberately target central nodes [6]. Latency also reduces responsiveness, as centralized systems must aggregate dispersed state information before adapting policies, leading to sluggish reactions in dynamic environments.

Centralized RL also struggles with partial observability. Agents often operate with only local data, while centralized models depend on global states. This disconnect undermines both efficiency and generalization across heterogeneous environments [7]. The curse of dimensionality further complicates matters, making large-scale centralized RL models difficult to train and prone to overfitting.

These limitations have motivated the exploration of decentralized alternatives, where learning and decision-making are distributed among multiple agents. Such approaches avoid the bottlenecks and vulnerabilities inherent in centralized systems while better aligning with real-world constraints [1].

1.3. Decentralized Reinforcement Learning Collectives as a Solution

Decentralized reinforcement learning (DRL) shifts the focus from a single central learner to distributed collectives of agents that learn and adapt locally. Each agent makes decisions using partial observations and limited communication, while collective intelligence emerges through interaction and coordination [8]. This architecture directly addresses scalability by distributing computation across agents and reduces dependence on centralized control.

Robustness is enhanced because decentralized systems avoid single points of failure. If one agent is lost or compromised, the collective continues to function, a property especially valuable in adversarial or uncertain environments [3]. Local adaptation further enables agents to respond quickly to disturbances without waiting for centralized coordination.

Applications already demonstrate this promise: multi-drone navigation, distributed traffic routing, and large-scale sensor networks have all benefited from decentralized learning structures. By fostering emergent coordination, DRL mirrors natural systems, such as insect colonies or neural networks, where simple local rules give rise to robust global behaviors [6].

Thus, DRL offers a compelling alternative to centralized RL by combining scalability, resilience, and adaptability qualities essential for automation in adversarial and uncertain contexts.

Objectives and Scope

This article examines the limitations of centralized reinforcement learning and explores decentralized reinforcement learning collectives as a solution to automation challenges in adversarial and uncertain environments. The objectives are to: (1) outline the contextual pressures driving resilient automation; (2) evaluate the scalability and robustness limits of centralized RL; (3) analyze DRL frameworks and their benefits; and (4) situate DRL within broader theoretical and applied research. The scope spans robotics, cyber-physical systems, and multi-agent domains. By doing so, the article establishes a foundation for understanding decentralized learning architectures and their role in next-generation automation [5].

2. Foundations of reinforcement learning and decentralization

2.1. Fundamentals of Reinforcement Learning

Reinforcement learning (RL) is a computational framework where agents learn to make sequential decisions by interacting with an environment. The central concept involves an agent receiving observations about the environment, selecting actions, and obtaining feedback in the form of rewards. Over time, the agent seeks to maximize cumulative reward by refining its policy a mapping from states to actions [6]. This iterative trial-and-error process differentiates RL from supervised learning, which relies on labeled data, and from unsupervised learning, which focuses on pattern recognition.

The mathematical foundations of RL are grounded in Markov decision processes (MDPs). An MDP formalizes an environment with states, actions, transition probabilities, and rewards. The agent's objective is to find a policy that maximizes the expected long-term return. Algorithms such as Q-learning and policy gradient methods are widely used to approximate optimal policies in environments where the full state space is large or unknown [7].

Exploration versus exploitation is a fundamental tension in RL. Agents must balance trying new actions to discover potentially higher rewards with exploiting known strategies for consistent returns. This balance is especially critical in dynamic or adversarial environments, where the optimal policy may shift over time [8].

Function approximation techniques, particularly those involving neural networks, have enabled RL to scale to high-dimensional problems. Deep reinforcement learning (DRL) integrates RL with deep learning architectures, enabling breakthroughs in domains such as robotics, games, and autonomous driving. However, these advances come at the cost of increased computational demands and vulnerability to instability during training [9].

While RL offers powerful mechanisms for adaptive decision-making, it assumes environments are either stationary or follow known stochastic processes. In practice, adversarial and uncertain contexts violate these assumptions. As a result, RL must be extended into multi-agent and decentralized frameworks to address the complexity of real-world systems [10].

2.2. Multi-Agent Systems and Coordination Challenges

As automation expands into complex domains, reinforcement learning increasingly involves not one agent but many. Multi-agent reinforcement learning (MARL) generalizes the single-agent paradigm to systems where multiple autonomous entities interact. Each agent pursues its own objective, which may align, conflict, or partially overlap with others. This creates a dynamic and interdependent environment in which coordination is both a necessity and a challenge [11].

In cooperative settings, agents must coordinate to achieve shared goals. Examples include robotic swarms exploring hazardous terrain or fleets of autonomous vehicles navigating urban networks. Coordination requires agents to share information, align policies, and avoid conflicts. However, communication limitations, partial observability, and resource constraints complicate the achievement of efficient cooperation. In adversarial environments, additional challenges arise, as some agents may behave competitively or even maliciously [12].

Credit assignment is another critical issue in MARL. When outcomes depend on the combined actions of multiple agents, it becomes difficult to attribute success or failure to specific individuals. Without proper mechanisms for credit assignment, agents may learn suboptimal or unstable strategies. Similarly, non-stationarity emerges because the learning process of one agent alters the environment for others, complicating convergence [9].

Scalability further complicates MARL. As the number of agents increases, the state-action space expands combinatorially, making it computationally infeasible for centralized approaches to manage. This complexity demands new learning structures that reduce dependence on central controllers while maintaining coordination.

Despite these challenges, MARL frameworks are central to emerging applications in logistics, defense, and distributed sensing. They capture the interactive dynamics of real-world systems better than isolated single-agent models. Still, scaling coordination under adversarial uncertainty remains unresolved, motivating research into decentralized and collective learning approaches [6].

2.3. Decentralization and Collective Intelligence Frameworks

Decentralization in reinforcement learning shifts decision-making away from a central controller and distributes it across autonomous agents. Each agent operates using local observations and limited communication, while coordination emerges through distributed interactions. This approach mitigates scalability challenges and enhances robustness by eliminating single points of failure [13].

Collective intelligence frameworks draw inspiration from natural systems such as insect colonies, bird flocks, and neural networks, where simple local interactions generate robust global behaviors. In decentralized RL, these frameworks are applied to design learning rules that balance autonomy with coordination. By relying on local information and adaptive interactions, decentralized collectives achieve scalable learning even under adversarial disruptions [7].

Key to these frameworks is the principle of emergent behavior. Instead of explicitly programming coordination, agents adopt decentralized policies that evolve through repeated interaction. Over time, stable patterns of cooperation or competition arise, enabling the collective to adapt to environmental uncertainty. For example, decentralized drone swarms can maintain formations or adjust to environmental hazards without requiring centralized commands [8].

Resilience is another advantage of collective frameworks. Since agents adapt locally, disruptions such as the loss of individuals or localized attacks do not cripple the entire system. This property makes decentralized RL particularly suitable for adversarial contexts where robustness is paramount [10].

As illustrated in Figure 1, the evolution from single-agent reinforcement learning to multi-agent systems and then to decentralized collectives demonstrates the trajectory of increasing complexity and adaptability. The figure highlights how decentralization builds on multi-agent foundations by addressing scalability, coordination, and resilience in uncertain environments.

By embracing decentralized and collective intelligence approaches, reinforcement learning expands from solving isolated tasks to managing complex, interactive systems. This theoretical progression sets the stage for examining how vulnerabilities manifest in adversarial settings, where the strengths of decentralized RL are most critically tested [11].

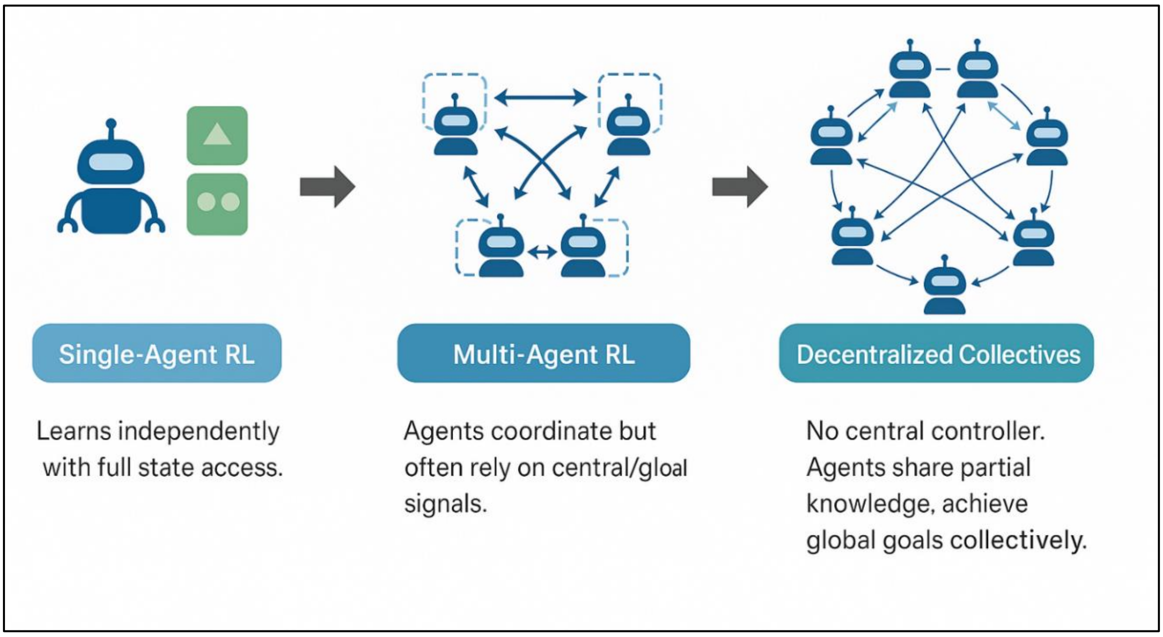

Figure 1 Evolution from single-agent RL → multi-agent RL → decentralized collectives

3. Adversarial environmental uncertainties in automation

3.1. Cybersecurity Threats in Distributed Automation

Cybersecurity has emerged as one of the most critical vulnerabilities in distributed automation systems. As industrial control, logistics, and robotic networks increasingly rely on interconnected digital infrastructures, they become targets for sophisticated cyberattacks. Threat actors exploit weak authentication, unpatched software, and insecure communication protocols to gain access to automation networks [12]. These vulnerabilities enable attacks ranging from data manipulation to system hijacking.

Distributed automation systems, such as fleets of autonomous vehicles or smart grid controllers, present unique attack surfaces. Each node is a potential entry point, and compromise of even a single agent can cascade through the network. In multi-agent reinforcement learning systems, adversaries may inject false data into sensor streams or disrupt policy updates, leading to systemic performance degradation [13]. Such poisoning attacks are particularly damaging in adversarial settings, where trust in information exchange is vital for coordination.

Denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks also threaten real-time performance. Automation requires low-latency communication, but overwhelming system bandwidth with malicious traffic can paralyze decision-making. In logistics networks, such disruptions can halt supply chains, leading to significant financial and social costs [15]. Additionally, ransomware attacks targeting industrial automation systems have grown in frequency, locking operators out of critical processes until payments are made.

The distributed nature of automation compounds detection difficulties. Centralized monitoring tools often struggle to identify localized anomalies, and adversaries exploit this by remaining stealthy while progressively corrupting nodes. Cybersecurity research has emphasized the importance of anomaly detection algorithms that learn normal traffic patterns and identify deviations. However, adversaries increasingly employ adaptive strategies that mimic legitimate behaviors, complicating detection [16].

Ultimately, cybersecurity threats expose the fragility of distributed automation when confronted with intelligent adversaries. The challenge lies not only in preventing intrusions but in building resilient architectures that can adaptively respond to compromised conditions without collapsing entirely [14].

3.2. Physical Disruptions in Industrial and Logistics Environments

Beyond cyberspace, distributed automation systems face vulnerabilities from physical disruptions. In industrial settings, automated production lines rely on interconnected robotic arms, conveyors, and sensors. Any physical interference whether accidental or deliberate can cascade into system-wide breakdowns. For instance, sensor tampering or the physical obstruction of robotic paths can trigger production delays or even safety hazards [12].

In logistics, automated warehouses and transport fleets encounter risks from equipment failures, environmental hazards, and intentional sabotage. For example, disruptions to automated guided vehicles (AGVs) or drones through physical blockages can cause significant delays. Similarly, attacks on supply nodes, such as tampering with cargo scanners or disabling loading equipment, create vulnerabilities in distributed logistics operations [17].

Environmental unpredictability also plays a role. Harsh weather, floods, or seismic events can disable infrastructure supporting automation, such as communication towers or power supply systems. Unlike centralized systems, distributed automation is expected to adapt locally, but physical disruptions often exceed the capacity of individual nodes. For instance, a drone swarm may adapt to the loss of a few units, but large-scale environmental interference such as dust storms can degrade sensors across the entire system [13].

Adversaries can exploit these vulnerabilities by targeting physical weak points. Disabling sensors or jamming communication links in industrial plants can induce cascading effects, forcing shutdowns or emergency halts. Logistics networks, which depend on just-in-time supply chains, are especially vulnerable. Even short delays amplify downstream inefficiencies and financial losses [15].

Mitigation strategies involve redundancy, fault-tolerant design, and predictive maintenance. Embedding multiple sensors for the same function can reduce single points of failure, while predictive analytics can anticipate equipment breakdowns. Yet, these solutions increase system costs and may still fail under coordinated adversarial attacks.

The growing reliance on automation in physical infrastructure highlights the urgent need to design systems resilient not only to cyber manipulation but also to tangible disruptions in industrial and logistics environments [18].

3.3. Complex Adversarial Dynamics in Socio-Technical Systems

Automation increasingly operates within socio-technical systems where technology interacts with human decision-makers, regulatory structures, and social dynamics. In such contexts, adversarial uncertainties extend beyond technical attacks to include human manipulation, misinformation, and institutional failures [14]. For example, autonomous vehicles depend not only on onboard algorithms but also on interactions with human drivers and pedestrians, who may act unpredictably. Similarly, industrial automation must coexist with labor forces, where strikes, sabotage, or resistance to technology adoption can undermine performance [16].

In socio-technical domains, adversaries exploit psychological and institutional vulnerabilities. Disinformation campaigns can disrupt decision-making by feeding false signals into distributed systems. In critical infrastructure, such as smart grids, adversaries may manipulate public perception to provoke regulatory overreactions that destabilize operations [12]. The challenge lies in integrating human trust and institutional accountability into the design of resilient automation.

Coordination complexity further increases in multi-agent socio-technical environments. Agents must adapt not only to technical uncertainties but also to shifting policies, evolving norms, and adversarial manipulation of human behavior. This creates a hybrid adversarial landscape that is far less predictable than purely cyber or physical domains. Unlike conventional attacks, socio-technical disruptions exploit ambiguity, making detection and response particularly difficult [13].

Table 1 presents a typology of adversarial uncertainties cyber, physical, and socio-technical and their impact on distributed automation. The table illustrates how each category introduces distinct vulnerabilities: cyberattacks exploit digital weaknesses, physical disruptions target hardware and environments, and socio-technical threats leverage human and institutional fragilities. By categorizing these uncertainties, the typology highlights the multi-layered complexity that automation systems must endure.

Addressing socio-technical adversarial dynamics requires interdisciplinary approaches that blend engineering, social science, and governance mechanisms. Designing robust automation is not only a technical challenge but a societal one, requiring integration of ethical considerations, institutional trust, and adaptive regulation [17].

In this way, socio-technical adversarial dynamics reveal the broadest layer of uncertainty. Unlike cyber or physical disruptions, which can be mitigated through technical hardening, socio-technical vulnerabilities demand holistic frameworks that account for human and institutional variability [18].

Table 1 Typology of adversarial uncertainties (cyber, physical, socio-technical) and their impact on automation

Category	Description	Examples	Impact on Automation
Cyber Uncertainties	Attacks targeting digital communication, data integrity, or system logic.	Data poisoning, ransomware, denial-of-service (DoS), adversarial inputs to RL policies.	Disruption of decision-making; manipulation of agent behavior; paralysis of distributed coordination.
Physical Uncertainties	Disruptions originating in hardware, infrastructure, or environmental systems.	Equipment failures, sensor tampering, weather hazards (storms, floods), intentional sabotage.	Cascading breakdowns of industrial and logistics systems; loss of redundancy; degradation of safety margins.
Socio-Technical Uncertainties	Threats emerging from human, institutional, or governance factors.	Misinformation campaigns, labor strikes, regulatory manipulation, social resistance to automation.	Erosion of trust; fragmented governance; inequitable or unpredictable system performance.

4. Architecture of decentralized RL collectives

4.1. Agent Design and Local Policy Learning

The foundation of decentralized reinforcement learning (DRL) collectives lies in the design of individual agents and their ability to learn local policies. Unlike centralized reinforcement learning systems, where a central controller dictates behavior, DRL agents operate autonomously, guided by localized data and adaptive strategies. Each agent perceives only a subset of the global state space, relying on its own observations and limited interactions with peers. This localized design reduces dependence on global information, enabling scalability and robustness [17].

Local policy learning is critical to ensuring that agents can adapt to dynamic environments. Policies are typically optimized through reinforcement signals derived from agent-environment interactions. For instance, drones in a swarm may each receive rewards for maintaining safe distances, conserving energy, and covering search areas efficiently. By training locally, the collective becomes resilient to failures of individual units, as other agents can compensate for missing peers [18].

To address the challenge of partial observability, agents often use function approximation methods, such as neural networks, to generalize from limited experiences. This allows them to adapt to unseen conditions without requiring a complete environmental model [19]. Importantly, decentralized design enhances fault tolerance: the system does not collapse if one agent is disabled, a critical property in adversarial or uncertain contexts.

Moreover, agent heterogeneity plays an important role. In many applications, agents are not identical; they differ in capabilities, sensors, or computational resources. Local policy learning frameworks must therefore accommodate heterogeneity, ensuring that specialized roles emerge naturally. For example, in industrial automation, some robots may specialize in inspection while others focus on manipulation tasks, yet all contribute to collective objectives [20].

Overall, the success of DRL collectives begins with the careful design of agents capable of independent, adaptive learning while contributing to broader collective goals. This foundation allows higher layers of communication, coordination, and consensus to build on robust individual policies.

4.2. Communication and Coordination Protocols

While agent design ensures autonomy, the power of decentralized RL lies in communication and coordination protocols that enable collective intelligence. Communication allows agents to share local information—such as observations, policy updates, or reward signals—with neighbors or the wider network. Coordination protocols then use this information to align agent behaviors and achieve global objectives [21].

Communication can be explicit or implicit. Explicit protocols involve structured message exchanges, such as position updates in drone swarms or transaction records in digital ledgers. Implicit communication relies on observable actions, where agents infer peer strategies by monitoring behavior. Both forms are vital for maintaining flexibility and resilience in uncertain environments.

The structure of communication networks also shapes coordination. Fully connected systems maximize information flow but suffer from scalability issues, while sparse networks reduce overhead at the cost of slower convergence. Hybrid models balance efficiency and robustness by connecting agents in modular clusters, which can coordinate locally and interact globally. Consensus algorithms often rely on such hybrid structures to reduce delays [22].

Coordination protocols must also address adversarial disruptions. Malicious or compromised agents can inject false information, undermining collective decision-making. Secure communication layers, such as encrypted exchanges or blockchain-inspired ledgers, are increasingly used to protect message integrity. These mechanisms ensure that agents trust shared data, which is essential in adversarial contexts.

Figure 2 illustrates a conceptual model of decentralized RL collectives with layered structures: local policy learning at the agent level, communication protocols linking agents into networks, and consensus layers ensuring alignment. This layered perspective highlights how communication acts as the bridge between local autonomy and global coherence.

In practice, effective coordination emerges from balancing communication overhead with decision accuracy. Too much communication increases delays and energy costs, while too little risks fragmentation. Adaptive communication protocols, where agents exchange information only when uncertainty is high, offer promising solutions to this challenge [23].

Thus, communication and coordination form the backbone of decentralized RL collectives, enabling autonomous agents to transform local learning into coherent, system-wide behavior.

Table 2 Comparison of centralized RL, federated RL, and decentralized RL collectives

Dimension	Centralized RL	Federated RL	Decentralized RL Collectives
Control Structure	Single central controller manages learning and policy updates.	Multiple local agents train models locally, aggregated periodically by a central server.	Fully distributed; each agent learns locally with coordination through peer-to-peer communication.
Scalability	Limited; state-action space grows exponentially with more agents.	Moderate; reduces communication cost but still constrained by central aggregation.	High; distributed learning eliminates single bottlenecks, enabling large-scale multi-agent systems.
Fault Tolerance	Low; single point of failure can collapse the system.	Medium; partial resilience, but central aggregator remains a vulnerability.	High; no central dependency, resilience maintained even if some agents fail.
Communication Overhead	High; requires full state aggregation and centralized updates.	Moderate; periodic aggregation reduces load but requires stable connectivity.	Adaptive; agents communicate selectively, optimizing bandwidth and energy use.
Security	Vulnerable; central node is a primary attack target.	Improved; data stays local but central aggregator can still be compromised.	Strong; consensus protocols and redundancy mitigate adversarial manipulation.
Adaptability	Slow; updates constrained by central processing speed and latency.	Faster than centralized but adaptation tied to aggregation intervals.	Rapid; local policy learning enables real-time adaptation to adversarial and uncertain conditions.
Example Applications	Small-scale robotics, simple simulations.	Distributed mobile applications, health data systems.	Drone swarms, smart grids, industrial automation, defense logistics.

4.3. Consensus-Driven Policy Updates for Secure Adaptation

Consensus mechanisms play a critical role in ensuring that decentralized RL collectives adapt securely and consistently. While agents learn local policies, collective effectiveness requires agreement on certain shared rules or updates. Without consensus, systems risk fragmentation, where individual agents pursue divergent goals that undermine collective performance [24].

Consensus can be achieved through voting protocols, averaging schemes, or blockchain-inspired consensus layers. In voting protocols, agents collectively decide on updates to shared parameters, ensuring that outliers do not dominate. Averaging schemes, such as federated averaging, allow distributed agents to synchronize their policies periodically while retaining local autonomy. Blockchain-style consensus adds a layer of security by creating immutable records of updates, protecting against adversarial tampering [18].

Consensus is particularly important in adversarial environments. If some agents are compromised, consensus protocols can prevent them from corrupting collective learning. Byzantine fault tolerance techniques, for example, allow systems to continue functioning even when a fraction of agents behaves maliciously. These approaches enhance resilience by ensuring that trust is distributed across the collective rather than concentrated in a single authority [19].

Table 2 compares centralized RL, federated RL, and decentralized RL collectives across several dimensions: scalability, fault tolerance, communication overhead, and security. The table highlights how decentralized collectives outperform centralized systems in resilience and adaptability, while offering improvements over federated models in robustness to adversarial interference.

Consensus mechanisms also support incremental adaptation. By periodically realigning policies, collectives can adapt to new threats, environmental changes, or evolving objectives without losing coherence. Importantly, consensus does not imply homogeneity. Agents may still specialize in local tasks while adhering to shared global objectives, allowing diversity and coordination to coexist [20].

In summary, consensus-driven updates provide the structural glue that binds decentralized RL collectives, enabling them to adapt securely in uncertain and adversarial conditions. They ensure that local autonomy is balanced with global coherence, making collective learning both resilient and scalable.

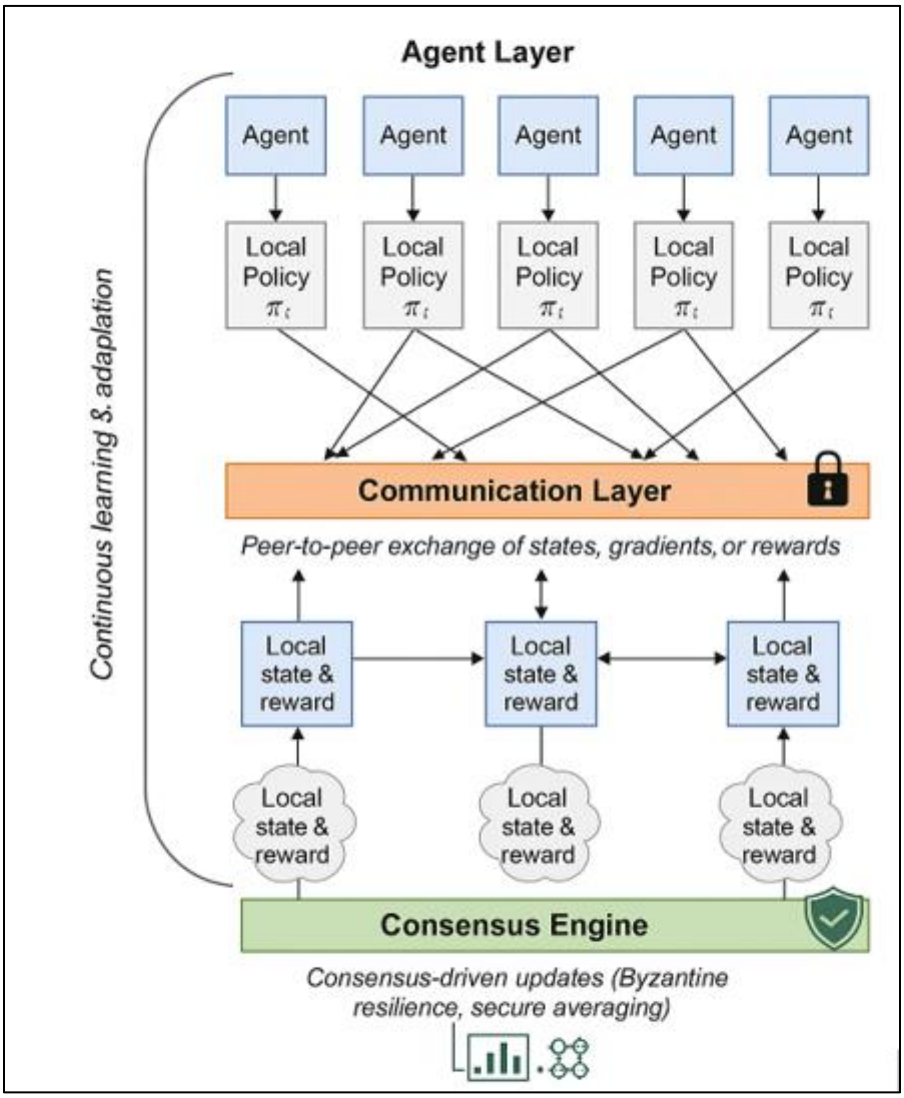

Figure 2 Conceptual model of decentralized RL collectives with communication and consensus layers

4.4. Resource Efficiency and Scalability

A persistent challenge in decentralized reinforcement learning is managing resource efficiency while scaling to large numbers of agents. Communication, computation, and energy all represent scarce resources in distributed automation. Designing protocols that minimize these costs without sacrificing performance is central to the scalability of DRL systems [17].

Resource efficiency is often achieved through selective communication. Agents may only share information when confidence levels fall below thresholds, thereby reducing bandwidth usage. Similarly, lightweight policy models allow agents with limited computational capacity to participate effectively in collective learning [21]. Energy-aware protocols, particularly for mobile systems such as drones or ground robots, further extend operational lifespans in the field.

Scalability is supported by modular architectures, where subgroups of agents coordinate locally before integrating into global collectives. This reduces communication overhead and enables hierarchical organization. Importantly, decentralized frameworks naturally scale better than centralized ones, as the burden of computation and decision-making is distributed [22].

By emphasizing efficiency and scalability, DRL collectives position themselves as viable solutions for real-world applications where large agent populations must function reliably under resource and adversarial constraints.

5. Applications across domains

5.1. Industrial Automation and Manufacturing

Industrial automation and manufacturing are among the most prominent beneficiaries of decentralized reinforcement learning (DRL) collectives. Traditional centralized automation systems, while effective for structured assembly lines, often falter in dynamic or high-variability environments. By contrast, DRL frameworks allow machines and robots to learn locally, adapt to uncertainties, and coordinate actions across distributed manufacturing plants [23].

One major application lies in predictive maintenance and adaptive scheduling. Agents embedded in machines can monitor local states such as vibration, temperature, or energy consumption and autonomously update maintenance policies. Instead of waiting for centralized instructions, each agent responds proactively to prevent breakdowns. This distributed approach reduces downtime and increases throughput, ensuring resilience against local disruptions [24].

Manufacturing environments also demand flexibility. Product customization, small-batch production, and fluctuating supply chains require machines to adjust rapidly. DRL collectives enable robots and conveyors to adapt by learning policies that balance throughput with quality control. Coordination among multiple agents ensures synchronization across different production lines, reducing bottlenecks [25].

Another benefit is fault tolerance. If one robotic unit fails, neighboring agents can redistribute tasks, maintaining operational flow without requiring human intervention. This redundancy makes production systems resilient to both technical faults and adversarial disruptions. Additionally, DRL supports safe human-robot collaboration, as agents learn to adaptively adjust force, speed, and trajectories in response to human co-workers [26].

Industrial plants adopting DRL frameworks also gain efficiency in resource allocation. Agents can dynamically optimize energy consumption, raw material use, and machine allocation in real time. This not only reduces costs but also contributes to sustainability goals by minimizing waste.

By embedding learning agents across manufacturing systems, DRL collectives transform rigid automation into adaptive, resilient, and resource-efficient ecosystems. These capabilities are particularly critical in Industry 4.0 contexts, where interconnected machines must continuously adjust to uncertainty and complexity [27].

5.2. Critical Infrastructure and Energy Systems

Critical infrastructure including power grids, water systems, and communication networks represents another sector where DRL collectives provide transformative advantages. These systems are highly interdependent and vulnerable to adversarial disruptions, making resilience a top priority. Decentralization allows each node whether a power substation, water pump, or communication relay to learn policies locally while coordinating with neighbors [23].

In energy systems, DRL has particular relevance for smart grids. Power distribution requires balancing demand and supply under fluctuating conditions. Agents embedded in substations and renewable energy sources, such as wind and solar farms, can autonomously regulate load balancing, frequency stabilization, and energy storage deployment. Local learning reduces latency, allowing real-time adaptation to variable generation and sudden demand spikes [24].

DRL also enhances resilience against cyber-physical attacks. For instance, if adversaries attempt to destabilize grid nodes, decentralized agents can collectively re-route power flows, preventing cascading blackouts. Unlike centralized controllers, which may present single points of failure, DRL collectives distribute decision-making, making systems more robust to targeted disruptions [26].

Water distribution networks face similar challenges. Pumps, valves, and treatment units often operate across wide geographic areas. DRL agents can locally monitor water pressure, flow, and quality, optimizing performance while coordinating with other units. This ensures efficient resource use, reduced leakage, and improved service delivery, even in the face of environmental uncertainties or deliberate tampering [28].

Telecommunication infrastructure also benefits from DRL. Distributed agents in base stations can dynamically allocate bandwidth, manage congestion, and adapt to interference. Local learning, combined with collective consensus protocols, ensures that networks remain functional during crises when centralized command may be disrupted.

Overall, DRL collectives transform critical infrastructure from fragile, centralized systems into adaptive, resilient networks. By distributing intelligence across nodes, they provide continuity of essential services under both adversarial and environmental uncertainties, reinforcing national security and societal stability [25].

5.3. Defense, Logistics, and Mobility Networks

Defense, logistics, and mobility networks operate in adversarial and uncertain environments where DRL collectives provide decisive advantages. In defense contexts, autonomous systems such as drone swarms or robotic ground vehicles must operate with high resilience and adaptability. Decentralized learning allows each unit to learn local policies while contributing to broader mission objectives, avoiding vulnerabilities associated with central command structures [24].

In logistics, DRL collectives support adaptive supply chains. Agents embedded in warehouses, transport fleets, and distribution centers can autonomously adjust routing, scheduling, and inventory levels in response to real-time disruptions. This flexibility ensures continuity even when networks face shocks such as port closures, equipment breakdowns, or targeted adversarial attacks [23].

Mobility networks, including autonomous vehicles and public transport systems, also benefit from DRL. Local policy learning enables vehicles to adapt to unpredictable traffic conditions while coordination protocols ensure collective flow optimization. This distributed control prevents congestion, reduces emissions, and enhances safety [27].

Figure 3 maps DRL collective applications across defense, logistics, and mobility sectors, illustrating their role in ensuring resilience. The figure highlights how decentralized architectures reduce bottlenecks, enable redundancy, and provide secure adaptation mechanisms in contested or uncertain conditions.

A unique advantage of DRL in defense and logistics lies in resilience to adversarial manipulation. If adversaries compromise a subset of agents, consensus protocols prevent systemic collapse, allowing remaining agents to continue operating effectively. This property is particularly critical in contested environments, where adaptability and redundancy directly determine mission success [26].

Mobility networks in smart transportation corridors similarly rely on decentralized learning for scalability. Agents embedded in vehicles, traffic lights, and routing systems collectively adapt to shifting flows, ensuring equitable distribution of resources such as road space and energy [28].

By embedding learning and coordination directly into nodes, DRL collectives transform defense and logistics systems into secure, adaptive infrastructures. They not only withstand uncertainty but also exploit it, turning variability into opportunities for optimized performance and resilience [25].

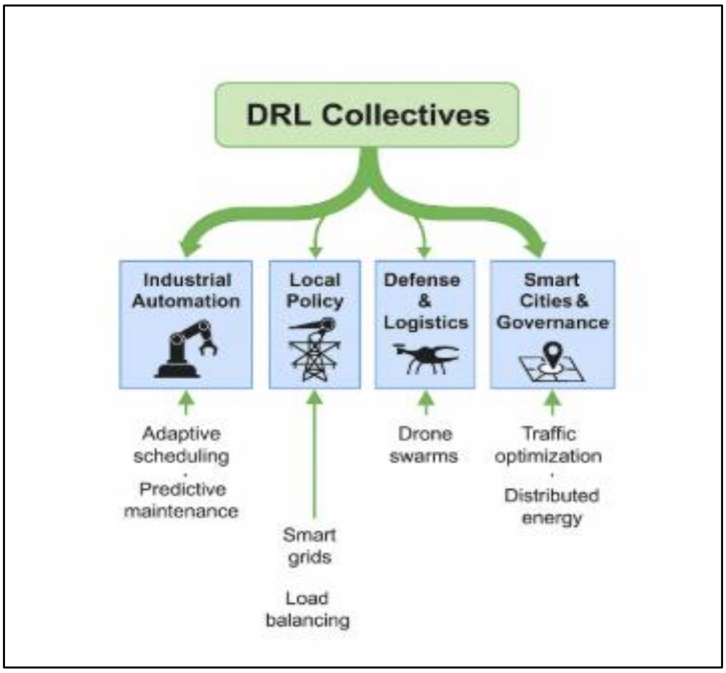

Figure 3 Sectoral map showing DRL collective applications

5.4. Smart Cities and Socio-Technical Governance

Smart cities embody the intersection of technology, governance, and society, making them ideal candidates for DRL collectives. In urban environments, decentralized learning agents embedded in transportation, energy, and waste systems adaptively optimize service delivery while balancing competing demands. For instance, agents in traffic networks can reduce congestion by locally adjusting signal timings, while energy nodes manage distributed renewables in real time [23].

Socio-technical governance also benefits from DRL transparency. Digital trails created by decentralized transactions reduce opportunities for corruption and foster citizen trust. By embedding consensus-driven protocols, city systems maintain accountability even when operating across multiple stakeholders [27].

Resilience is further enhanced when urban infrastructures integrate adaptive DRL frameworks. Local disruptions whether technical faults, cyber intrusions, or natural hazards are mitigated as agents autonomously compensate, preventing systemic collapse. This redundancy ensures that essential services such as water, transport, and electricity remain available.

Ultimately, DRL collectives provide cities with the adaptability and resilience necessary to govern complex socio-technical systems under uncertainty. By distributing intelligence across infrastructures, they empower cities to sustain efficiency, equity, and transparency in the face of adversarial and environmental challenges [28].

6. Performance evaluation and governance

6.1. Performance Metrics: Adaptability, Resilience, and Security

Evaluating decentralized reinforcement learning (DRL) collectives requires performance metrics that extend beyond accuracy or efficiency. Three key metrics adaptability, resilience, and security define their effectiveness in uncertain and adversarial environments.

Adaptability measures how quickly and effectively agents respond to changes in their environment. Unlike static models, DRL collectives continuously update policies in response to novel conditions. In industrial automation, for instance, adaptability is reflected in the ability of robotic swarms to adjust to fluctuating workloads or unforeseen machine failures [27]. Metrics here often involve convergence speed, learning stability, and the capacity to maintain performance under new inputs.

Resilience evaluates the ability of collectives to maintain function under disruptions. Whether disruptions arise from hardware failures, cyber intrusions, or environmental hazards, resilience is tested by how gracefully systems degrade and recover. Distributed architectures naturally enhance resilience by avoiding single points of failure, but empirical evaluation requires stress-testing systems under adversarial and stochastic perturbations [28].

Security is central in adversarial domains. Metrics for secure learning include robustness against data poisoning, resistance to model manipulation, and detection rates of malicious behavior. Simulation-based adversarial attacks are often employed to measure collective security properties [30]. By benchmarking adaptability, resilience, and security together, DRL evaluations provide holistic assessments that align with real-world operational demands.

6.2. Governance through Transparency, Explainability, and Accountability

Governance is essential for embedding DRL collectives within socio-technical systems. Unlike purely technical metrics, governance emphasizes transparency, explainability, and accountability. These dimensions ensure that DRL systems remain trustworthy and ethically aligned with societal expectations [29].

Transparency relates to the visibility of decision-making processes. In decentralized collectives, this is achieved by recording and sharing digital trails of agent interactions. Transparent records reduce opportunities for corruption, prevent manipulation, and allow stakeholders to audit performance. For example, decentralized ledgers applied in critical infrastructure create verifiable transaction histories that strengthen trust between operators and regulators [31].

Explainability focuses on making agent behaviors interpretable to human stakeholders. Without explainability, decisions made by distributed agents risk appearing opaque or arbitrary. Visualization tools, interpretable policy models, and counterfactual explanations are increasingly integrated to allow stakeholders to understand why collectives make certain decisions. This is particularly vital in defense or healthcare domains, where accountability for errors carries ethical and legal consequences [27].

Accountability ensures that actors deploying DRL collectives remain responsible for outcomes. Distributed design risks diluting responsibility by diffusing decision-making across agents. Mechanisms such as consensus logs and hierarchical oversight ensure that accountability is preserved even when learning is decentralized [32]. Together, transparency, explainability, and accountability form the governance framework necessary for public acceptance and regulatory compliance.

6.3. Policy Implications for Fairness and Equitable Automation

Fairness and equity are increasingly central to the evaluation of DRL collectives. Decentralized systems influence access to resources, opportunities, and services, raising policy concerns about inclusivity and bias. Policies must therefore ensure that automation benefits are equitably distributed and that marginalized groups are not systematically disadvantaged [28].

Bias in DRL systems can arise from uneven access to data or structural inequities embedded in reward functions. For example, if autonomous traffic systems prioritize affluent neighborhoods with better infrastructure, lower-income communities may face persistent service gaps. Policy interventions must mandate fairness audits to ensure equitable performance across demographics and geographies [30].

Fairness also extends to labor and employment impacts. In industrial automation, DRL collectives may displace certain jobs while creating new roles in oversight and system design. Equitable transition policies, such as retraining programs and inclusive workforce participation, are essential to avoid reinforcing economic inequality [29].

Equity in access to services is another concern. For example, in smart city governance, decentralized energy distribution must ensure that underserved communities benefit equally from efficiency gains. Otherwise, DRL systems risk widening existing gaps in resource availability [33]. Policymakers must establish regulatory frameworks that require distributive equity across all applications.

Table 3 presents an evaluation framework for DRL collectives across three dimensions: performance (adaptability, resilience, security), governance (transparency, explainability, accountability), and fairness (equity, inclusivity, accessibility). The table demonstrates how these dimensions intersect, highlighting the necessity of integrated evaluation. It underscores that performance cannot be considered in isolation from governance and fairness, as long-term sustainability depends on balancing all three.

Finally, equity policies must recognize global disparities. Developing regions may lack the infrastructure to implement DRL at the same scale as advanced economies. International cooperation and inclusive technology transfer mechanisms can help ensure that benefits of decentralized automation are shared more broadly [31].

Table 3 Evaluation framework for decentralized RL collectives across performance, governance, and fairness

Dimension	Evaluation Criteria	Description	Indicative Metrics
Performance	Adaptability	Ability of agents to update policies in response to environmental changes.	Convergence speed, learning stability, task completion under novel conditions.
	Resilience	System's capacity to function despite disruptions or agent failures.	Recovery time, degradation rates under stress, fault-tolerance thresholds.
	Security	Resistance to adversarial manipulation and malicious agents.	Attack detection rate, robustness to data poisoning, percentage of successful defenses.
Governance	Transparency	Visibility of decision-making and data exchanges within the collective.	Availability of audit logs, proportion of verifiable transactions.
	Explainability	Interpretability of agent policies and collective outcomes.	Model interpretability scores, stakeholder comprehension ratings.
	Accountability	Clear allocation of responsibility for outcomes.	Existence of consensus logs, oversight mechanisms, traceability indices.
Fairness	Equity	Distribution of resources and services across populations or agents.	Equity indices, service-level differences across demographic or geographic groups.
	Inclusivity	Degree of participation of diverse stakeholders or agents.	Representation levels, inclusion of marginalized groups, system accessibility scores.
	Accessibility	Extent to which benefits of DRL systems are widely available.	Coverage rates, affordability metrics, availability in underserved regions.

7. Risks, limitations, and challenges

7.1. Technical Challenges: Latency, Bandwidth, and Interoperability

Despite their potential, decentralized reinforcement learning (DRL) collectives face significant technical challenges that limit real-world deployment. Among the most pressing are latency, bandwidth constraints, and interoperability across heterogeneous systems.

Latency refers to the delay between action and system response. In distributed systems, agents often require rapid information exchange to coordinate effectively. However, communication delays caused by physical distance, network congestion, or processing limitations can reduce performance. For instance, in autonomous vehicle collectives, latency can lead to unsafe delays in collision avoidance maneuvers [32]. Addressing latency requires lightweight communication protocols and adaptive algorithms that minimize reliance on constant updates.

Bandwidth is another critical bottleneck. DRL collectives often involve hundreds or thousands of agents exchanging messages simultaneously. Excessive communication can overwhelm networks, leading to congestion and degraded performance [33]. Selective communication strategies, where agents share information only when uncertainty thresholds are crossed, have been proposed, but balancing efficiency with robustness remains unresolved.

Interoperability compounds these issues. In industrial, defense, or urban infrastructures, DRL agents must interact with legacy systems, sensors, and protocols. Mismatches in communication standards, data formats, or control architectures hinder seamless integration [35]. Without standardized interoperability frameworks, collectives risk fragmentation, where subsystems operate in isolation rather than as cohesive wholes.

Thus, technical challenges highlight the gap between theoretical DRL architectures and their practical scalability in complex, real-world environments.

7.2. Ethical Challenges: Unintended Bias and Accountability Gaps

Beyond technical hurdles, DRL collectives raise pressing ethical concerns. Two key challenges are unintended bias in decision-making and accountability gaps in distributed governance.

Bias in DRL systems often originates from uneven data representation or misaligned reward structures. For example, traffic management collectives trained on data from affluent neighborhoods may prioritize service efficiency in those areas, while systematically under-serving marginalized communities [34]. Such biases can inadvertently reinforce inequality, especially in socio-technical contexts where DRL is embedded in public infrastructure. Addressing bias requires fairness audits, diverse training datasets, and continuous monitoring to detect disparities as they emerge [32].

Accountability presents a different challenge. Distributed decision-making dilutes responsibility, making it unclear who is to blame when failures occur. In centralized systems, accountability is typically assigned to operators or designers. In DRL collectives, however, autonomous agents make local decisions that collectively produce outcomes. If a swarm of delivery drones causes property damage, questions arise over whether liability lies with the software developer, the system operator, or the autonomous agents themselves [36].

These accountability gaps undermine public trust and create regulatory uncertainty. Mechanisms such as consensus logs, explainable policy models, and hierarchical oversight frameworks are being explored to restore accountability. However, striking a balance between autonomy and responsibility remains an ongoing ethical dilemma [35].

Ethical challenges thus emphasize that the success of DRL collectives depends not only on technical efficiency but also on their alignment with societal values and norms.

7.3. Adoption Barriers: Cost, Trust, and Institutional Resistance

Even when technical and ethical challenges are addressed, adoption barriers hinder widespread deployment of DRL collectives. High costs represent one major obstacle. Implementing decentralized architectures requires investments in advanced hardware, software, and communication infrastructure. For many organizations, especially in resource-constrained regions, the financial burden is prohibitive [33].

Trust is equally critical. Stakeholders may be reluctant to adopt DRL systems if they perceive them as opaque, unreliable, or vulnerable to manipulation. Building trust requires not only demonstrable performance but also transparency in governance and decision-making [37]. Without this, organizations risk public backlash or lack of user acceptance.

Institutional resistance further slows adoption. Established industries and public institutions often favor familiar centralized models, viewing decentralized systems as disruptive and risky. Regulatory bodies may also struggle to adapt existing frameworks to accommodate distributed decision-making [34]. Overcoming resistance requires education, policy alignment, and pilot projects that demonstrate value without undermining stability.

Figure 4 presents a risk matrix of DRL collective deployment, categorizing adoption barriers such as cost, trust, and institutional inertia alongside technical and ethical risks. This framework underscores the multi-dimensional challenges that must be resolved before DRL collectives achieve mainstream adoption.

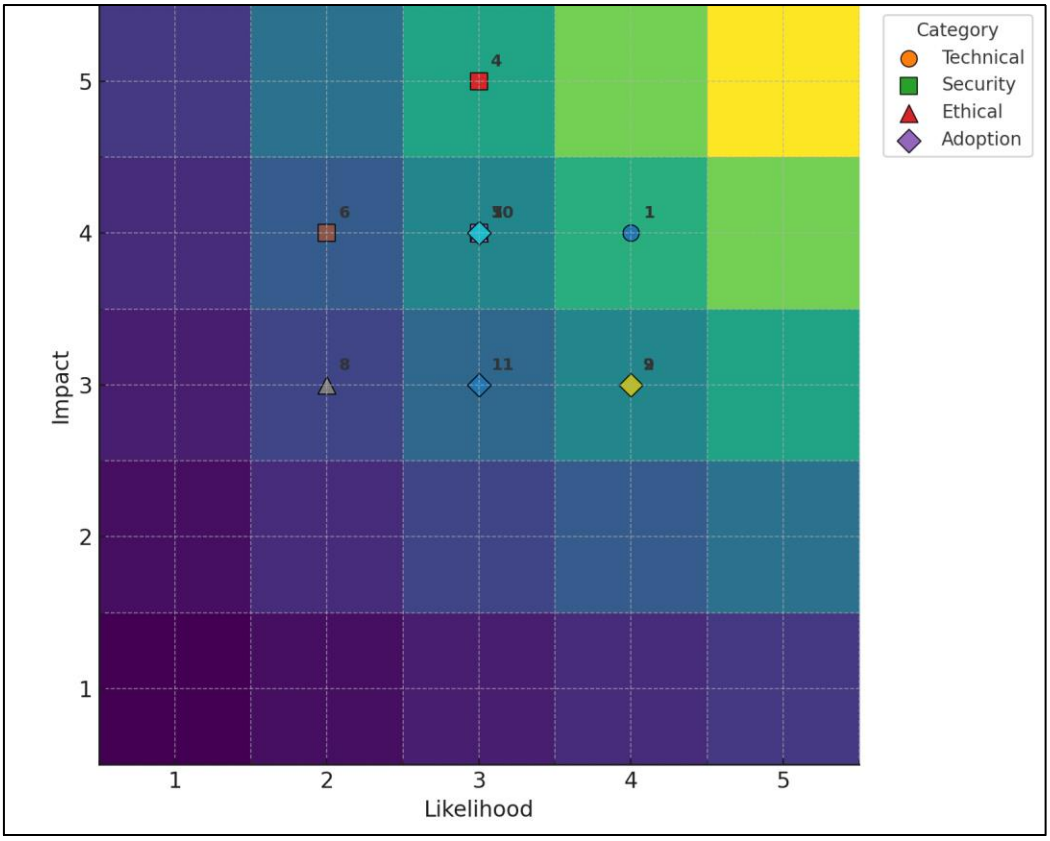

Figure 4 Risk matrix of DRL collective deployment

8. Future directions and research opportunities

8.1. Integrating Blockchain and Secure Consensus into DRL Collectives

One promising pathway for advancing decentralized reinforcement learning (DRL) collectives is the integration of blockchain and secure consensus protocols. Decentralized systems inherently face challenges of trust, especially when adversaries attempt to manipulate communication channels or inject false information. Blockchain technologies offer a potential solution by providing immutable, verifiable ledgers that record agent interactions and policy updates [36].

Consensus protocols built on blockchain principles allow agents to synchronize without relying on central authorities. For example, Byzantine fault tolerance (BFT) mechanisms ensure that the collective continues to function even when a subset of agents behaves maliciously. Such resilience is essential in adversarial environments like defense or critical infrastructure, where tampered updates could have catastrophic consequences [38].

Beyond security, blockchain integration strengthens accountability. Immutable transaction records enable stakeholders to trace back decisions, clarifying responsibility in multi-agent contexts. This is particularly important for governance frameworks where transparency and auditability are required by regulators [39].

Challenges remain, particularly in scalability. Standard blockchain systems suffer from latency and high energy costs, making them difficult to deploy in real-time DRL applications. Research into lightweight, application-specific blockchains and hybrid consensus models is therefore necessary to align blockchain's security benefits with DRL's performance needs [37].

By embedding secure consensus mechanisms into DRL collectives, systems gain both resilience against adversarial manipulation and institutional trustworthiness.

8.2. Lightweight AI and Energy-Aware Decentralized Learning

Another critical direction for DRL innovation is the development of lightweight AI models and energy-aware learning strategies. As collectives often operate in resource-constrained environments such as drone fleets, sensor networks, or industrial robots balancing computational demand with energy efficiency becomes central to sustainability [40].

Lightweight AI involves reducing model size and complexity without sacrificing performance. Techniques such as model pruning, knowledge distillation, and quantization allow DRL agents to run efficiently on low-power hardware. This ensures that even devices with limited resources can participate meaningfully in collective decision-making [36].

Energy-aware protocols complement these efforts by aligning learning processes with resource availability. For instance, agents may adjust communication frequency based on battery levels, or dynamically scale computation during peak demand periods. These strategies extend the lifespan of mobile systems and reduce operational costs [42].

Figure 5 presents a roadmap for future DRL collective research, highlighting how lightweight AI and energy-aware protocols form a key pillar. The figure situates energy optimization alongside blockchain integration and global scalability, emphasizing the interdependence of these research streams.

Importantly, energy-aware learning also has environmental implications. As automation expands, ensuring that DRL systems minimize carbon footprints becomes essential for aligning innovation with sustainability goals [38].

Thus, lightweight and energy-efficient decentralized learning provides a pathway to scalable, environmentally responsible DRL deployment.

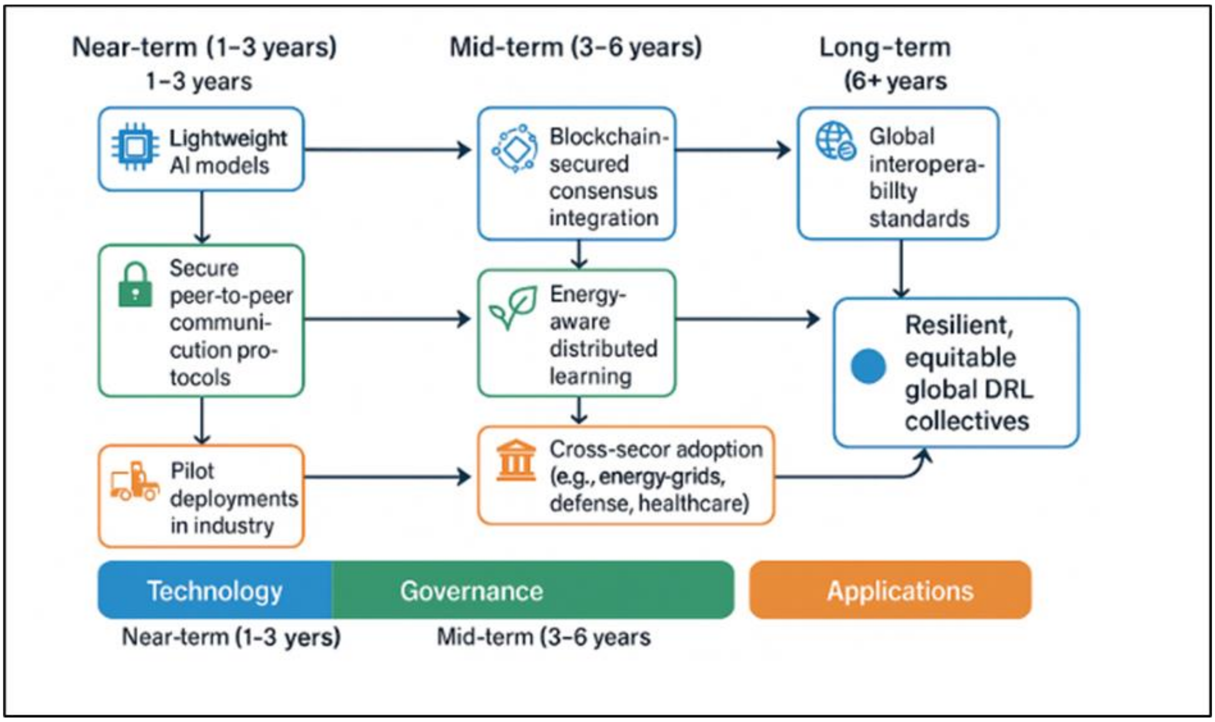

Figure 5 Roadmap for future decentralized RL collective research and deployment

8.3. Scaling Across Global Infrastructures

The final frontier for DRL collectives is scaling across global infrastructures. While current deployments often focus on localized systems such as factories, logistics hubs, or city networks future applications must extend across national and transnational infrastructures. This includes energy grids, global supply chains, and international communication systems [41].

Scaling introduces unique challenges. Cultural, regulatory, and technical heterogeneity complicate integration, while adversarial risks increase when systems span geopolitical boundaries. Standardization of communication protocols, interoperability frameworks, and ethical guidelines will therefore be essential. Without coordinated policies, fragmentation may prevent DRL from realizing its full potential [37].

At the same time, global scaling offers opportunities. DRL collectives could enable resilient coordination in climate adaptation, disaster response, or pandemic management. Distributed agents embedded worldwide could autonomously coordinate resources, making global infrastructures more adaptive and inclusive [40].

In sum, scaling DRL across global infrastructures represents both the greatest challenge and the greatest opportunity transforming localized innovations into global systems capable of supporting resilience, sustainability, and equity at planetary scale [39].

9. Conclusion

Summary of Contributions

This article has traced the evolution of decentralized reinforcement learning (DRL) collectives from their theoretical foundations to their applications, challenges, and future directions. It began by highlighting the limitations of centralized reinforcement learning in adversarial and uncertain environments, establishing the case for decentralized architectures. The discussion then examined how DRL collectives address these challenges through agent-level policy learning, secure communication protocols, consensus-driven updates, and scalable resource management. Sectoral applications were explored in depth, ranging from industrial automation and critical infrastructure to defense systems, logistics, and smart city governance. Evaluation frameworks were introduced to capture performance, governance, and fairness, while risks such as latency, bias, and adoption barriers were assessed. Finally, pathways for future innovation blockchain integration, energy-aware learning, and global scaling were outlined. Collectively, these contributions establish DRL collectives as a robust paradigm for building adaptive, transparent, and equitable automation systems.

Implications for Automation and Governance

The implications of DRL collectives extend across both technical and governance dimensions. On the technical side, decentralization offers scalable resilience: systems can adapt dynamically to uncertainty without reliance on fragile centralized nodes. This makes DRL architectures especially valuable for critical infrastructures and defense contexts, where robustness is essential. On the governance side, decentralization challenges existing regulatory and accountability frameworks. With decision-making dispersed across multiple agents, new mechanisms of oversight, explainability, and transparency must be developed. The potential for unintended bias or inequity further underscores the importance of policy alignment. Governance structures must therefore evolve in parallel with technical systems to ensure that decentralization does not undermine trust or fairness. The integration of accountability logs, fairness audits, and ethical safeguards will be pivotal. Ultimately, DRL collectives offer not only a technological advance but also a governance shift demanding interdisciplinary collaboration across engineering, law, and policy.

Final Reflections on Resilience and Equity

Resilience and equity emerge as the twin pillars of decentralized reinforcement learning collectives. Resilience ensures that systems withstand disruption, adapt to uncertainty, and continue functioning in adversarial environments. Equity ensures that the benefits of automation are shared inclusively, avoiding systemic biases or exclusion. Together, these dimensions define the sustainability of DRL. As societies face increasingly complex challenges ranging from cyber threats to climate disruptions the ability to design systems that are both robust and just will determine the transformative potential of automation. DRL collectives, if developed responsibly, embody this vision of adaptive, equitable, and trustworthy automation.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Weerakkody S, Ozel O, Mo Y, Sinopoli B. Resilient control in cyber-physical systems: Countering uncertainty, constraints, and adversarial behavior. Foundations and Trends® in Systems and Control. 2019 Dec 1;7(1-2):1-252.

[2] Nkrumah MA. Forecasting pension fund liabilities through multivariate time series models with structural breaks and demographic statistical trend analysis. World J Adv Res Rev. 2020;5(3):219-38. doi: https://doi.org/10.30574/wjarr.2020.5.3.0058

[3] Bodkhe U, Mehta D, Tanwar S, Bhattacharya P, Singh PK, Hong WC. A survey on decentralized consensus mechanisms for cyber physical systems. Ieee Access. 2020 Mar 17;8:54371-401.

[4] Chukwunweike J. Design and optimization of energy-efficient electric machines for industrial automation and renewable power conversion applications. Int J Comput Appl Technol Res. 2019;8(12):548–560. doi: 10.7753/IJCATR0812.1011.

[5] Olowononi FO, Rawat DB, Liu C. Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for CPS. IEEE Communications Surveys and Tutorials. 2020 Nov 9;23(1):524-52.

[6] Aradi S. Survey of deep reinforcement learning for motion planning of autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems. 2020 Sep 30;23(2):740-59.

[7] Nkrumah MA. Actuarial risk evaluation of health insurance portfolios using copula-based time series and Bayesian statistical learning approaches. Int J Comput Appl Technol Res. 2020;9(12):394-407.

[8] Liu Y, Yu FR, Li X, Ji H, Leung VC. Blockchain and machine learning for communications and networking systems. IEEE communications surveys and tutorials. 2020 Feb 24;22(2):1392-431.

[9] Kingma DP. Reinforcement Learning for Multi-Agent Systems. International Journal of Artificial Intelligence and Machine Learning. 2018 Sep 18;1(2).

[10] D'Angelo M, Gerasimou S, Ghahremani S, Grohmann J, Nunes I, Pournaras E, Tomforde S. On learning in collective self-adaptive systems: State of practice and a 3d framework. In2019 IEEE/ACM 14th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS) 2019 May 25 (pp. 13-24). IEEE.

[11] Sifakis J. Autonomous systems–an architectural characterization. InModels, Languages, and Tools for Concurrent and Distributed Programming: Essays Dedicated to Rocco De Nicola on the Occasion of His 65th Birthday 2019 Jul 3 (pp. 388-410). Cham: Springer International Publishing.

[12] Sun Y, Liu J, Wang J, Cao Y, Kato N. When machine learning meets privacy in 6G: A survey. IEEE Communications Surveys and Tutorials. 2020 Jul 23;22(4):2694-724.

[13] Liang F, Hatcher WG, Liao W, Gao W, Yu W. Machine learning for security and the internet of things: the good, the bad, and the ugly. Ieee Access. 2019 Oct 22;7:158126-47.

[14] Verbraeken J, Wolting M, Katzy J, Kloppenburg J, Verbelen T, Rellermeyer JS. A survey on distributed machine learning. Acm computing surveys (csur). 2020 Mar 13;53(2):1-33.

[15] Abdelzaher T, Taliaferro A, Sullivan P, Russell S. The multi-domain operations effect loop: from future concepts to research challenges. InArtificial Intelligence and Machine Learning for Multi-Domain Operations Applications II 2020 May 14 (Vol. 11413, p. 1141304). SPIE.

[16] Feron E, Abed E, Balcan N, Baras J, Young P, Kaebling L, Martins N, Ozdaglar A, Shamma J, Frazzoli E. Distributed learning and information dynamics in networked autonomous systems. 2015 Nov 20.

[17] Chung SJ, Paranjape AA, Dames P, Shen S, Kumar V. A survey on aerial swarm robotics. IEEE Transactions on robotics. 2018 Aug 3;34(4):837-55.

[18] Kim JH, Vadakkepat P. Multi-agent systems: a survey from the robot-soccer perspective. Intelligent Automation and Soft Computing. 2000 Jan 1;6(1):3-17.

[19] Chhetri MB, Uzunov A, Vo B, Nepal S, Kowalczyk R. Self-improving autonomic systems for antifragile cyber defence: Challenges and opportunities. In2019 IEEE international conference on autonomic computing (ICAC) 2019 Jun 16 (pp. 18-23). IEEE.

[20] Bayrak AE, Gorsich D, Epureanu BI. Future of Autonomous High-Mobility Military Systems. SAE International Journal of Connected and Automated Vehicles. 2020 Oct 19;3(12-03-03-0016):205-15.

[21] Jahan F, Sun W, Niyaz Q, Alam M. Security modeling of autonomous systems: A survey. ACM Computing Surveys (CSUR). 2019 Sep 13;52(5):1-34.

[22] Huang L, Zhou M, Hao K, Hou E. A survey of multi-robot regular and adversarial patrolling. IEEE/CAA Journal of Automatica Sinica. 2019 Jun 19;6(4):894-903.

[23] Talwandi NS, Khare S, Thakur P, Kumar R. Network security and data privacy in the 6G environment: Impacts and challenges. InNetwork Security and Data Privacy in 6G Communication (pp. 211-233). CRC Press.

[24] Finn A, Kabacinski K, Drake SP. Design challenges for an autonomous cooperative of UAVs. IEEE; 2007 Feb 12.

[25] Bithas PS, Michailidis ET, Nomikos N, Vouyioukas D, Kanatas AG. A survey on machine-learning techniques for UAV-based communications. Sensors. 2019 Nov 26;19(23):5170.

[26] Theron P, Kott A, Drašar M, Rzadca K, LeBlanc B, Pihelgas M, Mancini L, De Gaspari F. Reference architecture of an autonomous agent for cyber defense of complex military systems. InAdaptive autonomous secure cyber systems 2020 Feb 5 (pp. 1-21). Cham: Springer International Publishing.

[27] Steinberg M. Biologically-inspired approaches for self-organization, adaptation, and collaboration of heterogeneous autonomous systems. InDefense Transformation and Net-Centric Systems 2011 2011 May 23 (Vol. 8062, pp. 127-139). SPIE.

[28] Chaterji S, Naghizadeh P, Alam MA, Bagchi S, Chiang M, Corman D, Henz B, Jana S, Li N, Mou S, Oishi M. Resilient cyberphysical systems and their application drivers: A technology roadmap. arXiv preprint arXiv:2001.00090. 2019 Dec 20.

[29] Uzunov AV, Nepal S, Chhetri MB. Proactive antifragility: A new paradigm for next-generation cyber defence at the edge. In2019 IEEE 5th international conference on collaboration and internet computing (CIC) 2019 Dec 12 (pp. 246-255). IEEE.

[30] Yang R, Yu FR, Si P, Yang Z, Zhang Y. Integrated blockchain and edge computing systems: A survey, some research issues and challenges. IEEE Communications Surveys and Tutorials. 2019 Jan 23;21(2):1508-32.

[31] Gerasimou S, D'Angelo M, Ghahremani S, Grohmann J, Nunes I, Pournaras E, Tomforde S. On Learning in Collective Self-adaptive Systems: State of Practice and a 3D Framework. In14th International Symposium on Software Engineering for Adaptive and Self-Managing Systems 2019 Mar 22.

[32] Mittal A, Ouaissa M, Ouaissa M. Metaverse security architecture: Designing a secure environment. InDefending the Metaverse (pp. 29-46). CRC Press.

[33] Nawari NO, Ravindran S. Blockchain and the built environment: Potentials and limitations. Journal of Building Engineering. 2019 Sep 1;25:100832.

[34] Hoang TN, Xiao Y, Sivakumar K, Amato C, Howl JP. Near-optimal adversarial policy switching for decentralized asynchronous multi-agent systems. In2018 IEEE International Conference on Robotics and Automation (ICRA) 2018 May 21 (pp. 6373-6380). IEEE.

[35] Liu J, Wu J. Multiagent robotic systems. CRC press; 2018 Oct 8.

[36] Yang GZ, Bellingham J, Dupont PE, Fischer P, Floridi L, Full R, Jacobstein N, Kumar V, McNutt M, Merrifield R, Nelson BJ. The grand challenges of science robotics. Science robotics. 2018 Jan 31;3(14):eaar7650.

[37] Zhou Y, Yu FR, Chen J, Kuo Y. Cyber-physical-social systems: A state-of-the-art survey, challenges and opportunities. IEEE Communications Surveys and Tutorials. 2019 Dec 12;22(1):389-425.

[38] Queralta JP, Taipalmaa J, Pullinen BC, Sarker VK, Gia TN, Tenhunen H, Gabbouj M, Raitoharju J, Westerlund T. Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. Ieee Access. 2020 Oct 12;8:191617-43.

[39] Li, Y., Zhou, W., Wang, H., Ding, B. and Xu, K., 2019, August. Improving fast adaptation for newcomers in multi-robot reinforcement learning system. In 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 753-760). IEEE.

[40] Tavčar J, Horvath I. A review of the principles of designing smart cyber-physical systems for run-time adaptation: Learned lessons and open issues. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2018 Mar 30;49(1):145-58.