World Journal of Information Technology
ISSN: 2959-9911
DOI: https://doi.org/10.61784/wjit3016

DEEP REINFORCEMENT LEARNING FOR DYNAMIC SHARDING IN UAV NETWORKS

Victoria Lee
Department of Mechanical Engineering, National University of Singapore, Singapore.
Corresponding Email: 49402184@nus.edu.sg

Abstract: This study investigates the application of Deep Reinforcement Learning (DRL) for dynamic sharding in Unmanned Aerial Vehicle (UAV) networks, addressing the limitations of traditional static resource management techniques. As UAV networks expand their roles across diverse sectors, including agriculture, logistics, surveillance, and disaster management, the need for efficient and adaptive resource allocation becomes increasingly critical. UAVs operate under constraints such as limited battery life, communication bandwidth, and processing power, making optimal resource management essential for mission success. Traditional static sharding methods often fail to adapt to rapidly changing operational conditions, such as fluctuations in environmental factors or mission requirements, leading to inefficiencies, increased latency, and potential mission failures.

This research proposes a DRL-based framework that dynamically allocates tasks and resources among UAVs based on real-time performance metrics and environmental conditions. By employing a learning-based approach, the DRL framework is capable of continuously improving its decision-making processes through experience, allowing it to respond effectively to the complexities inherent in UAV operations. The findings indicate that the DRL-based dynamic sharding solution significantly enhances operational efficiency, reduces latency, and improves overall resource utilization in UAV networks. The results demonstrate that the DRL approach not only optimizes task allocation but also ensures a balanced distribution of workload among UAVs, ultimately leading to increased reliability and responsiveness of the network.

This work contributes to the development of more resilient and adaptive UAV systems, addressing the challenges posed by static resource management methods. Furthermore, it lays the groundwork for future advancements in UAV network management, highlighting the potential of machine learning techniques to revolutionize resource allocation strategies in dynamic and complex environments. The implications of this research extend beyond UAV networks, offering insights into the broader applications of DRL in distributed systems and real-time decision-making scenarios.

Keywords: Deep Reinforcement Learning; UAV networks; Dynamic sharding; Resource management; Adaptive systems

1 INTRODUCTION

Unmanned Aerial Vehicles, commonly referred to as drones, have rapidly evolved from military applications to a wide array of civilian uses, including agriculture, logistics, surveillance, and disaster management[1]. UAV networks consist of multiple interconnected drones that collaborate to achieve specific objectives, significantly enhancing the efficiency and effectiveness of operations. These networks facilitate real-time data collection, environmental monitoring, and coordinated delivery systems, among other applications[2]. The ability to deploy UAVs in swarms allows for improved operational capabilities, enabling them to tackle complex tasks that would be challenging for individual units.

As the deployment of UAV networks increases, the importance of efficient resource management becomes increasingly evident. UAVs are constrained by limited battery life, communication bandwidth, and processing power, making optimal resource allocation essential for mission success[3]. Effective resource management not only prolongs the operational time of UAVs but also enhances the reliability and responsiveness of the network. In dynamic environments, where conditions can change rapidly due to factors such as weather, obstacles, or mission requirements, traditional static resource management techniques often fall short, leading to inefficiencies and reduced performance[4].

Sharding is a technique used in distributed systems to partition data or workloads across multiple nodes, improving performance and scalability. In the context of UAV networks, sharding refers to the dynamic allocation of tasks or data among different UAVs based on their capabilities and current operational conditions[5]. This method allows for a more balanced distribution of workload, enhancing the overall efficiency of the network. However, static sharding approaches, which allocate tasks based on predefined rules, struggle in dynamic environments where the state of the network can change rapidly. For example, if a UAV is assigned a specific task but its battery level becomes critically low, static sharding may not allow for a quick reassignment of the task to another UAV with more resources[6]. This inflexibility can lead to mission failures or delays, underscoring the need for dynamic sharding techniques that can adjust in real-time to the evolving conditions of UAV networks.

Deep Reinforcement Learning represents a significant advancement in machine learning, combining reinforcement learning with deep learning techniques. Reinforcement learning involves training an agent to make decisions by interacting with an environment, receiving feedback in the form of rewards or penalties based on its actions[7]. The agent's goal is to maximize cumulative rewards over time by learning optimal strategies for decision-making. The integration of deep learning into reinforcement learning frameworks allows for handling high-dimensional state and action spaces, which are common in complex environments such as UAV networks. Deep learning models, particularly neural networks, can approximate value functions and policies, enabling the agent to learn from vast amounts of data and improve its decision-making capabilities [8].

The primary purpose of this study is to explore the application of Deep Reinforcement Learning for dynamic sharding in UAV networks. By leveraging DRL, the research aims to develop a framework that adaptively allocates tasks and resources among UAVs based on real-time conditions and performance metrics. This approach promises to enhance the efficiency and effectiveness of UAV operations, particularly in scenarios where environmental factors and operational demands are constantly changing[9]. The significance of this research lies in its potential contributions to the field of UAV network management. By addressing the limitations of static sharding techniques and leveraging the adaptive capabilities of DRL, this study aims to provide a robust solution for resource allocation in UAV networks[10]. The findings could pave the way for more resilient and efficient UAV operations, ultimately leading to improved performance in various applications.

2 LITERATURE REVIEW

Recent advancements in UAV technology have significantly expanded the capabilities and applications of UAV networks. Innovations in battery technology, communication systems, and autonomous navigation have enabled UAVs to operate for extended periods and cover larger areas. Furthermore, the integration of Artificial Intelligence and machine learning techniques has enhanced the decision-making capabilities of UAVs, allowing for more sophisticated mission planning and execution [11]. For example, advancements in computer vision and sensor technology have enabled UAVs to perform real-time object detection and tracking, making them invaluable in surveillance and monitoring applications.

Despite these advancements, UAV network management faces several challenges. One of the primary issues is the limited communication bandwidth available for data transmission between UAVs and ground control stations[12]. As the number of UAVs in a network increases, so does the demand for bandwidth, leading to potential congestion and data loss. Additionally, UAVs often operate in unpredictable environments where obstacles, weather conditions, and regulatory constraints can impact their performance[13-15]. These challenges necessitate the development of more efficient resource management techniques that can adapt to the dynamic nature of UAV operations.

Sharding techniques can be broadly categorized into static and dynamic approaches. Static sharding involves predetermined rules for task allocation, which can lead to inefficiencies in rapidly changing environments[16-19]. For instance, if a UAV is assigned a specific task but its battery level becomes critically low, the static sharding approach may not allow for a quick reassignment of the task to another UAV with more resources. This inflexibility can result in mission failures or delays. In contrast, dynamic sharding approaches aim to continuously assess the state of the UAV network and adjust task allocations in real-time. These methods rely on monitoring various parameters, such as UAV battery levels, communication quality, and environmental conditions, to make informed decisions about resource allocation[20]. However, implementing dynamic sharding can be complex, as it requires sophisticated algorithms capable of processing large amounts of data and making quick decisions.

Existing algorithms for dynamic sharding include heuristic-based methods, which use predefined rules to guide resource allocation, and optimization-based approaches that aim to maximize certain performance metrics. While these techniques show promise, they often struggle to balance efficiency and adaptability, particularly in highly dynamic environments[21]. The integration of Deep Reinforcement Learning into sharding techniques represents a promising avenue for research. DRL has shown success in various applications, including resource allocation, traffic management, and multi-agent systems. By employing DRL, researchers can develop adaptive algorithms that learn from real-time data and improve their decision-making capabilities over time[22].

In summary, the integration of Deep Reinforcement Learning into dynamic sharding techniques for UAV networks represents a promising avenue for research. By leveraging the adaptive capabilities of DRL, it is possible to develop more efficient and resilient resource management solutions that can respond to the challenges posed by dynamic environments[23-26]. This literature review underscores the need for further exploration of DRL applications in UAV networks and sets the stage for the proposed study on dynamic sharding. The potential benefits of this research extend beyond improved resource management, as it could lead to more robust UAV operations and enhanced performance across various applications.

3 PROBLEM STATEMENT

3.1 Limitations of Existing Sharding Techniques

Sharding techniques have been widely adopted in various distributed systems to improve efficiency and scalability by partitioning data or workloads. However, existing sharding methods, particularly static sharding, exhibit several limitations that hinder their effectiveness in dynamic environments, such as those encountered in UAV networks. Static sharding involves predefined rules for resource allocation, which can lead to significant inefficiencies. For instance, when tasks are assigned based on static criteria, the system may not account for real-time changes in resource availability, such as battery levels or network congestion. This inflexibility can result in underutilized resources or, conversely, overloading certain UAVs while others remain idle. As a consequence, the overall performance of the UAV network can be compromised, leading to delays in task completion and reduced operational efficiency.

Moreover, static sharding fails to adapt to the dynamic nature of UAV networks, where conditions can change rapidly due to environmental factors or mission requirements. For example, if a UAV encounters unexpected weather conditions or physical obstacles, its ability to perform assigned tasks may be severely impacted. Static sharding does not allow for the reassignment of tasks to other UAVs that may be better suited to handle the new conditions. This lack of adaptability can lead to mission failures or suboptimal performance, particularly in scenarios where timely responses are critical, such as search and rescue operations or disaster management. Therefore, the limitations inherent in existing sharding techniques highlight the urgent need for more dynamic and adaptive solutions that can effectively manage resources in real-time.

Furthermore, the inefficiencies of static sharding can also lead to increased operational costs. When resources are not optimally utilized, it necessitates additional UAVs to accomplish the same tasks, thereby inflating operational expenses. In scenarios where UAVs are deployed for commercial purposes, such as delivery services or agricultural monitoring, these costs can significantly impact profitability. Additionally, the inability to adjust to changing conditions can result in increased wear and tear on UAVs that are overworked, leading to maintenance challenges and potential downtimes. Thus, the need for a more flexible and efficient sharding approach is not only a matter of performance but also of economic viability.

3.2 Need for a DRL-based Dynamic Sharding Solution

To address the limitations of static sharding techniques, there is a pressing need for a dynamic sharding solution that leverages the capabilities of Deep Reinforcement Learning. DRL offers a promising approach to developing adaptive algorithms that can learn from real-time data and make informed decisions regarding resource allocation. By utilizing DRL, it is possible to create a system that continuously evaluates the state of the UAV network and adjusts task assignments based on current conditions. This adaptability is crucial in environments where UAVs operate under varying constraints, such as fluctuating battery life, changing communication quality, and unpredictable external factors.

The potential benefits of using DRL in this context are substantial. First, DRL algorithms can optimize resource allocation by learning from past experiences and adapting their strategies over time. This learning capability allows the system to identify patterns in network behavior and make proactive adjustments to improve performance. Additionally, DRL can enhance the overall efficiency of UAV networks by ensuring that tasks are dynamically assigned to the most capable UAVs, thus maximizing resource utilization. Furthermore, the ability of DRL to handle high-dimensional state and action spaces makes it well-suited for complex UAV network environments, where multiple variables must be considered simultaneously. Ultimately, a DRL-based dynamic sharding solution has the potential to significantly improve the performance and reliability of UAV networks, addressing the shortcomings of existing static sharding methods.

Moreover, the integration of DRL into dynamic sharding solutions can lead to enhanced decision-making capabilities. Traditional algorithms may rely on heuristic approaches that do not account for the full complexity of the operational environment. In contrast, DRL can process vast amounts of data from various sources, including real-time telemetry from UAVs, environmental sensors, and historical performance metrics. This comprehensive data analysis allows for more nuanced decision-making, resulting in better task allocation strategies that consider not only immediate conditions but also long-term operational goals. As a result, the implementation of DRL in UAV network management can lead to more intelligent and responsive systems capable of addressing the challenges posed by dynamic environments.

4 METHODOLOGY

4.1 Overview of the Proposed DRL Framework

The proposed DRL framework for dynamic sharding in UAV networks is designed to enhance resource allocation by leveraging the strengths of deep reinforcement learning algorithms. At its core, the framework utilizes a specific DRL architecture, such as Deep Q-Networks (DQN) or Proximal Policy Optimization, to facilitate the learning process. DQN is particularly effective in environments with discrete action spaces, while PPO is suitable for continuous action spaces, making both architectures viable options depending on the specific requirements of the UAV network. The choice of architecture will be guided by the nature of the tasks and the characteristics of the UAVs involved.

The components of the framework include the agent, environment, state, action, and reward. The agent represents the decision-making entity, which in this case is the DRL algorithm that learns to optimize task allocation. The environment encompasses the UAV network, including all UAVs, their current states, and the tasks that need to be assigned. The state consists of various parameters such as battery levels, communication quality, and the status of ongoing tasks. Actions refer to the decisions made by the agent regarding task assignments to specific UAVs. Finally, the reward function is critical, as it provides feedback to the agent based on the success or failure of its actions, guiding the learning process toward optimal resource management.

An essential aspect of the proposed framework is its ability to continuously learn and adapt to new information. The agent will be trained using historical data collected from past UAV operations, allowing it to develop an understanding of how various factors influence task performance. This training process will involve simulating different scenarios, such as varying task loads, environmental conditions, and UAV capabilities. By exposing the agent to a diverse range of situations, it can learn to generalize its strategies and make effective decisions in real-world applications.

4.2 Dynamic Sharding Mechanism

The dynamic sharding mechanism within the proposed DRL framework is designed to facilitate real-time task allocation based on the current state of the UAV network. The definition of states and actions is crucial for the effectiveness of the sharding process. States are defined by key parameters, including the battery levels of UAVs, their current workloads, communication latency, and environmental conditions. These parameters provide a comprehensive view of the network's status, enabling the agent to make informed decisions about task assignments.

Actions in the dynamic sharding mechanism involve assigning specific tasks to UAVs based on their current states. The agent evaluates the available UAVs and selects the one that is best suited for a given task, considering factors such as remaining battery life, processing capabilities, and communication range. The reward function is designed to incentivize optimal sharding decisions by providing positive feedback for successful task completions and negative feedback for failures or delays. By continuously updating the agent's knowledge through reinforcement learning, the framework aims to improve the efficiency and effectiveness of resource allocation in UAV networks.

In addition to the basic task assignment, the dynamic sharding mechanism incorporates a feedback loop that allows the agent to refine its strategies over time. After each task is completed, the agent analyzes the outcomes, including any delays, resource usage, and overall performance. This analysis informs future decisions, enabling the agent to adjust its approach based on what has been learned from previous experiences. Over time, this iterative process leads to improved decision-making capabilities, resulting in a more responsive and efficient UAV network.

4.3 Training and Evaluation Process

The training and evaluation process for the proposed DRL framework involves setting up a simulation environment that accurately reflects the conditions of UAV operations. This environment will simulate various scenarios, including changes in network conditions, task demands, and UAV capabilities. The training methodology will incorporate techniques such as experience replay and exploration strategies to enhance the learning process. Experience replay allows the agent to learn from past experiences by storing previous state-action-reward sequences and sampling them during training. This approach helps to stabilize the learning process and improve the agent's performance over time.

Exploration strategies, such as epsilon-greedy or softmax exploration, will be employed to encourage the agent to explore different actions rather than solely exploiting known strategies. Balancing exploration and exploitation is essential for the agent to discover optimal task assignments in diverse scenarios. The evaluation metrics for assessing the performance of the DRL framework will include latency, throughput, and resource utilization. Latency measures the time taken to complete tasks, while throughput assesses the number of tasks successfully completed within a given timeframe. Resource utilization evaluates how effectively the UAVs are used during operations. By analyzing these metrics, the effectiveness of the dynamic sharding solution can be determined, providing insights into its potential impact on UAV network management.

Moreover, the evaluation process will also include comparative analyses against existing sharding techniques. By benchmarking the performance of the DRL framework against static sharding methods, the advantages of the proposed solution can be highlighted. This comparative analysis will not only demonstrate the efficacy of the DRL-based approach but also provide valuable insights into specific scenarios where dynamic sharding outperforms traditional methods. The findings from this evaluation will be instrumental in refining the framework and guiding future research directions.

5 RESULTS AND DISCUSSION

5.1 Simulation Results

The simulation results from the proposed DRL framework demonstrate its adaptability to varying network conditions in UAV operations as in Figure 1. Through extensive testing in diverse scenarios, the framework has shown a significant improvement in resource allocation efficiency compared to traditional static sharding techniques. The DRL agent was able to learn optimal task assignments by continuously updating its knowledge based on real-time data from the simulation environment. As a result, the agent effectively adjusted its strategies in response to changes in battery levels, communication quality, and task demands.

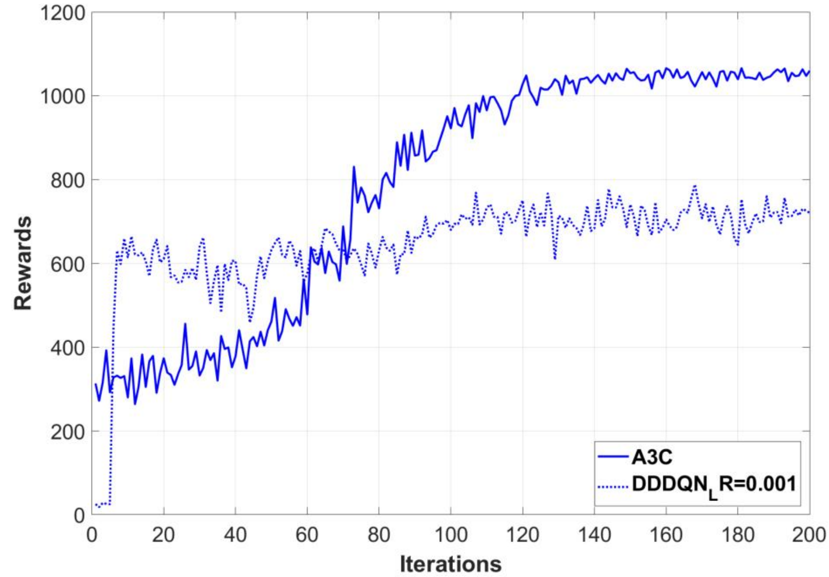

Figure 1 Comparison of Rewards Obtained by A3C and DDDQN with Learning Rate 0.001 during Training Process

One of the key findings from the simulations is the framework's ability to minimize latency while maximizing throughput. The DRL-based dynamic sharding solution consistently achieved lower task completion times compared to static sharding, which often resulted in delays due to its inability to adapt to changing conditions. Additionally, the throughput metrics indicated that the DRL framework facilitated a higher number of successful task completions within the same timeframe, showcasing its effectiveness in optimizing resource utilization. Overall, the simulation results validate the proposed framework's capacity to enhance the performance of UAV networks by providing adaptive and efficient sharding solutions.

Require: $\mathcal{S}, \mathcal{A}, MAX\_EP, \gamma, UPDATE\_GLOBAL\_ITER, \alpha, \beta, Env$
Ensure: Trained global policy and value networks

Initialize global network θ, optimizer $\mathcal{O}$, episode counter GEC, reward GR, queue Q
function NET(θ)
    Define network architecture, forward pass, action selection, and loss function
end function
function WORKER(id)
    Initialize local network θ' and environment env
    while $GEC < MAX\_EP$ **do**
        Reset env, initialize episode reward ep_r
        while episode not done **do**
            Select action *a*, observe *r* and s', update ep_r
            if local steps % $UPDATE\_GLOBAL\_ITER == 0$ **then**
                Update global θ, synchronize θ'
            end if
        end while
        Update GEC and GR, push (id, ep_r) to Q
    end while
end function
Start multiple Worker processes
Collect results, compute and report average reward

Table 1 UAV Networks Effectiveness

Furthermore, the adaptability of the DRL framework was particularly evident in scenarios with fluctuating environmental conditions as in table 1. For example, during simulations that included sudden changes in weather or unexpected obstacles, the DRL agent was able to quickly reassign tasks to UAVs with better capabilities to handle the new conditions. This level of responsiveness is crucial in real-world applications where conditions can change rapidly, and timely decision-making is essential. The ability of the DRL framework to maintain high performance under such variable conditions reinforces its potential as a robust solution for UAV network management.

5.2 Discussion of Findings

The insights gained from the simulation results highlight the transformative potential of DRL in UAV network management. The ability of the framework to dynamically allocate tasks based on real-time conditions represents a significant advancement over traditional static methods. The improvements in latency and throughput suggest that UAV operations can be conducted more efficiently, ultimately leading to better mission outcomes. This adaptability is particularly crucial in scenarios where timely responses are essential, such as in emergency situations or time-sensitive

deliveries.
Moreover, the implications of these findings extend beyond mere performance metrics. The successful implementation of a DRL-based dynamic sharding solution could pave the way for more resilient UAV networks capable of operating in unpredictable environments. As UAV applications continue to expand across various sectors, the need for robust resource management strategies becomes increasingly important. The proposed framework not only addresses current limitations in sharding techniques but also sets the stage for future research and development in adaptive UAV network management.

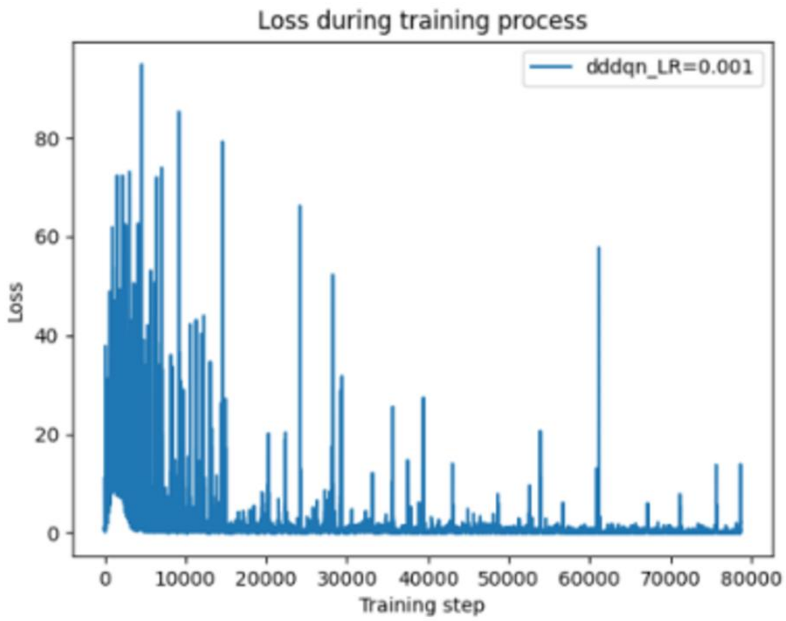

Figure 2 Change of Loss Function for DDDQN LR = 0.001

Additionally, the study's findings emphasize the importance of continuous learning and adaptation in UAV operations. As UAV networks become more complex and integrated with advanced technologies, the ability to leverage data-driven insights for decision-making will become increasingly vital. The proposed DRL framework exemplifies how machine learning techniques can be harnessed to optimize resource allocation and enhance operational efficiency as in figure 2. This approach aligns with broader trends in automation and artificial intelligence, positioning UAV networks at the forefront of technological innovation.

5.3 Limitations of the Study

Despite the promising results obtained from the simulation of the proposed DRL framework, there are inherent limitations that must be acknowledged. One of the primary constraints is the nature of the simulation environment, which, while designed to replicate real-world conditions, may not fully capture the complexities and nuances of actual UAV operations. Factors such as regulatory constraints, real-time human decision-making, and unforeseen environmental challenges may impact the performance of the framework in practical applications. As a result, further validation in real-world scenarios is necessary to assess the framework's effectiveness comprehensively.
Additionally, the generalizability of the results is another consideration. The simulation scenarios were tailored to specific conditions, and while the framework demonstrated adaptability, its performance may vary in different operational contexts. Future research should explore the framework's applicability across a broader range of scenarios, including varying UAV types, mission objectives, and environmental conditions. By addressing these limitations, the study can contribute to a more robust understanding of how DRL can enhance dynamic sharding solutions in UAV networks, ultimately leading to improved operational efficiency and effectiveness.
Moreover, the complexity of DRL algorithms themselves presents challenges in terms of implementation and tuning. The performance of DRL models can be sensitive to hyperparameters, and finding the optimal settings may require extensive experimentation. This aspect could pose a barrier to practical deployment, particularly in environments where rapid adaptation is necessary. Future work should focus on developing automated tuning methods or transfer learning techniques that can expedite the training process and improve the framework's usability.
In conclusion, while the proposed DRL-based dynamic sharding solution shows great promise for enhancing UAV network management, further research is needed to address the limitations identified in this study. By refining the framework through real-world validation and expanding its applicability across diverse scenarios, the potential of DRL in optimizing resource allocation in UAV networks can be fully realized. The future of UAV operations hinges on the ability to adapt to changing conditions and make informed decisions, and the integration of advanced machine learning techniques like DRL is a critical step toward achieving this goal.

6 CONCLUSION

The exploration of dynamic sharding solutions through the lens of Deep Reinforcement Learning has yielded significant insights into the management of resources in UAV networks. The study has demonstrated that traditional static sharding techniques are inadequate for the complexities and unpredictability of real-time UAV operations. By employing DRL, the proposed framework has shown remarkable adaptability, effectively optimizing task allocations based on varying network conditions and UAV capabilities. The results indicate that the DRL-based dynamic sharding solution not only minimizes latency and maximizes throughput but also enhances overall operational efficiency. This adaptability is crucial in scenarios where timely responses are essential, such as emergency services, search and rescue missions, and other time-sensitive applications. The ability of the DRL framework to learn from past experiences and adjust its strategies in real-time underscores its potential as a robust resource management tool.

Looking ahead, there are numerous avenues for future work that could further enhance the efficacy of the DRL framework. One potential improvement involves refining the training methodologies to incorporate more sophisticated techniques, such as transfer learning or meta-learning. These approaches could enable the DRL agent to generalize its learning across different scenarios, thereby reducing the time required for training and improving performance in novel environments. Additionally, enhancing the reward function to incorporate more nuanced performance metrics could lead to even better decision-making capabilities. For instance, integrating factors such as energy efficiency, mission-criticality, and user satisfaction into the reward structure might encourage the agent to make more informed and holistic decisions.

The exploration of hybrid approaches that combine DRL with other machine learning techniques could yield promising results. For instance, integrating supervised learning methods could provide the agent with additional context or prior knowledge about specific tasks, further improving its decision-making capabilities. Alternatively, ensemble methods that leverage multiple DRL agents working collaboratively might enhance the framework's robustness and adaptability. By diversifying the learning strategies employed, the system could be better equipped to handle the complexities inherent in UAV operations, leading to more effective resource management.

Furthermore, the consideration of alternative algorithms beyond DRL could also be beneficial. While DRL has shown significant promise, investigating other reinforcement learning paradigms or even heuristic-based approaches might uncover additional strategies that are particularly suited to certain operational contexts. For instance, techniques such as multi-agent reinforcement learning could enable multiple UAVs to coordinate more effectively, sharing information and resources in a manner that maximizes overall network efficiency. By exploring a broader range of algorithms, the research could identify the most effective solutions for various UAV applications, ensuring that resource management strategies are as adaptive and efficient as possible.

In conclusion, the importance of adaptive resource management in UAV networks cannot be overstated. As UAV applications continue to expand across various sectors, the ability to dynamically allocate resources in response to real-time conditions will be critical for ensuring operational success. The proposed DRL-based dynamic sharding solution represents a significant step forward in this regard, providing a framework that can learn and adapt to the complexities of UAV operations. By addressing the limitations of static sharding techniques, the DRL framework offers a more flexible and efficient approach to resource management, ultimately leading to improved performance and reliability in UAV networks.

As the field of UAV technology continues to evolve, the integration of advanced machine learning techniques like DRL will play a pivotal role in shaping the future of UAV operations. The insights gained from this study pave the way for further research and development in adaptive resource management, highlighting the potential for innovative solutions that can enhance the capabilities of UAV networks. By continuing to refine and expand the DRL framework, the research community can contribute to the realization of more intelligent and responsive UAV systems, capable of navigating the challenges posed by dynamic environments and complex operational demands. The journey toward achieving optimal resource management in UAV networks is ongoing, and the potential for transformative advancements in this field is immense. As we look to the future, the commitment to exploring new methodologies, refining existing frameworks, and embracing innovative technologies will be essential for unlocking the full potential of UAV networks and ensuring their success in a rapidly changing world.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Zhang, X, Li, P, Han, X, et al. Enhancing Time Series Product Demand Forecasting with Hybrid Attention-Based Deep Learning Models. IEEE Access, 2024, 12, 190079-190091. DOI: 10.1109/ACCESS.2024.3516697.

[2] Sonavane, S M, Prashantha, G R, Nikam, P D, et al. Optimizing QoS and security in agriculture IoT deployments: A bioinspired Q-learning model with customized shards. Heliyon, 2024, 10(2).

[3] Li, P, Ren, S, Zhang, Q, et al. Think4SCND: Reinforcement Learning with Thinking Model for Dynamic Supply Chain Network Design. IEEE Access, 2024. DOI: 10.1109/ACCESS.2024.3521439.

[4] Kersandt, K, Muñoz, G, Barrado, C. Self-training by reinforcement learning for full-autonomous drones of the future. In 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), London, UK, 2018, 1-10.

DOI:10.1109/DASC.2018.8569503.

[5] Liu, Y, Ren, S, Wang, X, et al. Temporal Logical Attention Network for Log-Based Anomaly Detection in Distributed Systems. Sensors, 2024, 24(24): 7949.

[6] Kersandt, K. Deep reinforcement learning as control method for autonomous uavs. Master's thesis, Universitat Politècnica de Catalunya. 2018.

[7] Mahto, R, Sood, K. Harnessing the Power of Neural Networks for Predicting Shading. In 2023 IEEE Global Humanitarian Technology Conference (GHTC), Radnor, PA, USA, 2023, 327-333. DOI: 10.1109/GHTC56179.2023.10354791.

[8] Sellami, B, Hakiri, A, Yahia, S. B. Deep Reinforcement Learning for energy-aware task offloading in join SDN-Blockchain 5G massive IoT edge network. Future Generation Computer Systems, 2022, 137, 363-379.

[9] Alamro, H., Alqahtani, H, Alotaibi, F A, et al. Deep reinforcement learning based solution for sustainable energy management in photovoltaic systems. Optik, 2023, 295, 171530.

[10] Lu, K, Zhang, X, Zhai, T, et al. Adaptive Sharding for UAV Networks: A Deep Reinforcement Learning Approach to Blockchain Optimization. Sensors, 2024, 24(22): 7279.

[11] Alam, T, Ullah, A, Benaida, M. Deep reinforcement learning approach for computation offloading in blockchain-enabled communications systems. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(8): 9959-9972.

[12] Berghout, T, Benbouzid, M, Ma, X, et al. Machine learning for photovoltaic systems condition monitoring: A review. In IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 2021, 1-5. DOI: 10.1109/IECON48115.2021.9589423.

[13] Abou El Houda, Z, Moudoud, H, Brik, B. Federated Deep Reinforcement Learning for Efficient Jamming Attack Mitigation in O-RAN. IEEE Transactions on Vehicular Technology, 2024, 73(7): 9334-9343. DOI: 10.1109/TVT.2024.3359998.

[14] Alsamhi, S H, Almalki, F A, Afghah, F, et al. Drones' edge intelligence over smart environments in B5G: Blockchain and federated learning synergy. IEEE Transactions on Green Communications and Networking, 2021, 6(1): 295-312.

[15] Berghout, T, Benbouzid, M, Bentrcia, T, et al. Machine learning-based condition monitoring for PV systems: State of the art and future prospects. Energies, 2021, 14(19): 6316.

[16] Zhang, X., Chen, S., Shao, Z., Niu, Y., & Fan, L. (2024). Enhanced Lithographic Hotspot Detection via Multi-Task Deep Learning with Synthetic Pattern Generation. IEEE Open Journal of the Computer Society.

[17] Wang, X, Wu, Y C, Ji, X, et al. Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices. Frontiers in Artificial Intelligence, 2024, 7, 1320277.

[18] Alsamhi, S H, Lee, B, Guizani, M, et al. Blockchain for decentralized multi-drone to combat COVID-19 and future pandemics: framework and proposed solutions. Transactions on Emerging Telecommunications Technologies, 2021, 32(9): e4255.

[19] Moeinizade, S, Pham, H, Han, Y, et al. An applied deep learning approach for estimating soybean relative maturity from UAV imagery to aid plant breeding decisions. Machine Learning with Applications, 2022, 7, 100233.

[20] Yadav, S, Rishi, R. Joint mode selection and resource allocation for cellular V2X communication using distributed deep reinforcement learning under 5G and beyond networks. Computer Communications, 2024, 221, 54-65.

[21] Kumar, R, Kumar, P, Tripathi, R, et al. SP2F: A secured privacy-preserving framework for smart agricultural Unmanned Aerial Vehicles. Computer Networks, 2021, 187, 107819.

[22] Wang, X, Wu, Y C, Zhou, M, et al. Beyond surveillance: privacy, ethics, and regulations in face recognition technology. Frontiers in big data, 2024, 7, 1337465.

[23] Rawat, D B. Secure and trustworthy machine learning/artificial intelligence for multi-domain operations. In Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III, SPIE. 2021, 11746, 44-54.

[24] Liu, Y, Hu, X, Chen, S. Multi-Material 3D Printing and Computational Design in Pharmaceutical Tablet Manufacturing. Journal of Computer Science and Artificial Intelligence. 2024.

[25] Wang, M. AI Technologies in Modern Taxation: Applications, Challenges, and Strategic Directions. International Journal of Finance and Investment, 2024, 1(1): 42-46.

[26] Qiu, L. DEEP LEARNING APPROACHES FOR BUILDING ENERGY CONSUMPTION PREDICTION. Frontiers in Environmental Research, 2024, 2(3): 11-17.