An Intelligent Hybrid Framework using Deep Deterministic Policy Gradient for Semantic Communication and Computation in Digital Twin-Assisted IoD Networks

Ihab Mohamed Elsayed Koura
Science Technology and Innovation Unit,
King Saud University, P.O. Box-2454, Riyadh, 11451, Saudi Arabia
Email: ekoura@ksu.edu.sa

Abstract: Presently, there is a wide adoption of the Digital Twin (DT), Internet of Drones (IoD) and edge intelligence technology to render precision agriculture, which ensures resource optimisation, crop monitoring and intelligent decision-making. Through the Unmanned Aerial Vehicle (UAV), there is the generation of a massive volume of sensory data, which is limited by dynamic network conditions, communication bandwidth, strict energy requirements, and computational constraints. This causes major challenges for effective computation and communication in DT-assisted IoD systems. To tackle this issue, various investigations are conducted by employing standard reinforcement learning to attain the joint optimisation of resource allocation, semantic processing and task offloading with the preservation of learning privacy and reduction of communication overhead. However, the earlier framework still suffers from inadequate exploration, high decision complexity, suboptimal experience, and inefficient policy convergence, which affect the resource efficiency and adaptability in a highly dynamic environment. Therefore, this work developed an intelligent reinforcement learning framework to enhance the communication and computation efficiency in a DT-assisted IoD network. Initially, the needed environmental data from the agriculture sector are collected using IoD. Further initiates the Coordinate Attention-based Deep Deterministic Policy Gradient with Dual Prioritised Experience Replay (CA-DDPG-DPER) to jointly optimise semantic processing and task offloading under communication and computation constraints. The outcome is demonstrated through experimentation by comparing against the existing framework.

Keywords: Digital Twin; Internet of Drones; Semantic Communication and Computation; Coordinate Attention; Deep Deterministic Policy Gradient; Dual Prioritised Experience Replay

Received: June 18, 2026 Revised: June 25, 2026 Accepted: June 30, 2026

Nomenclature

Abbreviation	Description
AI	Artificial Intelligence
A-UAV	Autonomous Unmanned Aerial Vehicle
CA	Coordinate Attention
CA-DDPG-DPER	Coordinate Attention-based Deep Deterministic Policy Gradient with Dual Prioritised Experience Replay
DDPG	Deep Deterministic Policy Gradient
DNN	Deep Neural Network
DPER	Dual Prioritised Experience Replay
DT	Digital Twin
DT-IoD	Digital Twin-assisted Internet of Drones
FD-IoD	Federated Digital Internet of Drones
IoD	Internet of Drones
RL	Reinforcement Learning
UAV	Unmanned Aerial Vehicle

1. Introduction

Advanced digital technology accelerates its adoption for the emerging agriculture sector to ensure sustainable food production, efficient resource management, and data-driven farming practices [1]. A powerful technological ecosystem has been created by the combination of IoD, DT and edge intelligence for smart agriculture [2]. For physical agricultural assets, the synchronised virtual representation attained by DT enables the farmer to monitor the field and forecast future events. Multiple UAVs are deployed in IoD to recognise agricultural activity [3]. Even with the advanced features of DT-IoD, the practical implementation becomes a complex task because of the distributed computing infrastructure and wireless communication environment [4]. Multimodal sensing information acquisition for the continuous monitoring of

agriculture leads to the generation of a large amount of heterogeneous data [5]. However, there is an increasing complexity of efficient communication due to the diverse drone mobility patterns, fluctuating wireless channel quality and network congestion.

Several investigations are conducted by employing numerous optimisation strategies for improving communication and computation in DT-assisted IoD [6]. In order to enhance the operational efficiency and system adaptability, several frameworks are designed with the incorporation of reinforcement learning, edge computing, AI, and a semantic communication mechanism [7]. This system helps in transferring the task-relevant information and minimising unnecessary communication costs to enhance the scalability suitable for the real-time agricultural service [8]. Other than the advancement, the earlier frameworks encounter diverse limitations that affect efficiency in a complex IoD environment [9]. Inconsistent resource utilization resulted while processing under dynamic network conditions, which also reduces the system adaptability. In addition, the ability of the intelligent agent is restricted by the insufficient extraction of spatial information from the environmental observation [10]. To address these limitations, this work constructed an advanced framework based on hybrid DDPG to enhance the semantic communication and computation through adaptive decision-making in a DT-assisted IoD network.

1.1 Contribution and Organisation

In addressing the existing research gap, the significant contributions of this article are summarised as follows:

- To develop an advanced reinforcement learning module for semantic communication and computation resource management in a DT-assisted IoD environment. The framework integrates a feature enhancement mechanism through an attention module and an experience-based learning strategy. The presented model ensures energy-aware resource management and effective task execution, thereby providing an efficient solution in smart agriculture.
- To construct an intelligent optimisation framework, CA-DDPG-DPER for semantic communication and computation resource management in DT-assisted IoD networks. The framework is built by employing CA to the DDPG along with DPER to improve decision-making. This combinational framework enhances resource utilisation, learning stability and convergence efficiency. This ensures intelligent agricultural application in dynamic DT-IoD.

The rest of the work follows a structured organisation: Section 2 reviews the earlier literature work on energy management in DT-IoD. Section 3 describes the proposed methodology along with DT-IoD. Section 4 presents the detailed construction of the hybrid model. Section 5 provides the simulation outcome and determines the performance of the framework. Section 6 conclude the article and outlines the challenges for future research directions.

2. Existing Works

2.1 Related Works

The rising demand for real-time communication in IoD led to the development of diverse protocols to meet the system requirements. This section explains the research methodology of the previous framework in performing semantic communication and computation in a DT-assisted IoD network.

In 2025, Arsalan *et al*. [11] designed the Federated Digital (FD)-IoD module to enhance energy efficiency in task offloading. Here, the learning model was integrated with the digital twin to guarantee the IoD viability. The framework optimised mobility-aware resource management, which reduces computational overhead. Even for heavy demand, the FD-IoD attained efficient resources with less latency.

In 2023, Li *et al*. [12] presented a combinational framework that incorporated the sensing, computation and communication constraints in DT. The edge computing service was attained through the deployed UAV. The performance of the beam pattern in communication offloading was reduced through multi-objective optimisation. The flexible learning policy was achieved by the introduction of the optimisation strategy.

In 2026, Hu *et al*. [13] employed a novel hybrid framework, which incorporates the DT with sensing and communication to reduce the total service delay. Here, the multi-agent resource allocation was proposed to address the issues corresponding to the high-dimensional and

varying mobility. This effectively ensures decision-making together with UAV-vehicular user association.

In 2025, Qadir *et al*. [14] introduced an energy-efficient communication framework to reduce the energy consumption and latency of the IoD network. Here, a two-stage scheme was introduced, which was the combination of the DNN and the K-means method to address the optimisation problems. The implemented multi-layer network attained efficient optimisation results with minimum computational cost.

In 2026, Kuru *et al*. [15] developed an agent-based and scalable system for the Autonomous UAV (A-UAV) by applying the DT. Here, complex conditions were encountered using the autonomous AI agent. The effectiveness of the system was measured using the immersive tool, which incorporated a precision and force-sensitive control device. Based on this socio-cognitive interaction model, high performance was achieved in the platform, which was suitable for the critical field.

Overall, the earlier research work gives the significant contribution to enhancing semantic communication and computation in DT-IoD through different protocols. This contribution begins strong foundation for enhancing reliable and efficient DT-IoD systems for the development of an advanced framework.

2.2 Research Gaps and Challenges

Other than the substantial advancement attained in semantic communication and computation optimisation for the DT-assisted IoD networks, various research limitations remain unsolved. Despite earlier frameworks have signifies its enhancement in intelligent task execution, resource management and communication efficiency, their functionality is generally constrained by the dynamic characteristics. This section effectively determines the challenges of the literature work to construct an intelligent optimisation module.

- In general, there is complexity in learning the optimal decision policy using the standard optimisation model under a changing network environment. In order to attain a stable solution, a novel resource management framework is developed which is suitable for the large-scale DT-IoD.
- A uniform experience replay mechanism is employed by several existing models, where the interactions are served with equal importance during the learning process. The learning agent explores and restricts the policy. Hence, an intelligent decision-making process is developed, which is efficient even for dynamic network conditions.
- The existing model gives limited attention to the spatial contextual information compared to the task scheduling and resource allocation from the environmental observation. This resulted in suboptimal communication and task offloading. Hence, this framework incorporates an efficient attention mechanism to ensure accurate decision-making.

3. Brief Explanation of Proposed Approach and Systematic View of Digital Twin

3.1 Depiction of Digital Twin with IoD

For the smart agricultural application, the system designs the DT-assisted IoD network. The set of IoT devices $I = \{1,2,\ldots \mathrm{I}\}$ is comprised in the network, where the group of UAV $J = \{1,2,\ldots \Im\}$ that processed based on discrete time slots $t \in T$. The information from the heterogeneous environment is collected using the IoT device, which is deployed in the agricultural field [1]. These UAVs act as communication assistance, edge computing and aerial sensing. The virtual representation is managed using the DT through the physical agricultural environment by continuously updating the state of the environment.

The computation-intensive task is generated using the IoT device $i \in I$ with varying processing requirements and data characteristics, where the task state is expressed in Eq. (1)

$$State_i^I(t) = \left\{ t_i^{\max}, qr_i(t), size_i(t), en_i(t), c_i(t) \right\} \tag{1}$$

Here, the process of task arrival and environmental context are denoted as $qr_i(t)$ and $en_i(t)$. The required CPU and the task completion time are specified as $c_i(t)$ and $t_i^{\max}$. Similarly, based on the energy availability, channel condition, mobility, semantic quality and computational capability, UAV $j \in \Im$ is characterised. The physical parameters are synchronised by the DT to

generate the virtual state representation to offer real-time awareness of the dynamic IoD environment, and it is formed in Eq. (2)

$$S^{DT}(t) = \left\{ state_i^I(t), state_j^J(t) \right\} \tag{2}$$

Based on the continuous interaction among the virtual counterpart and physical IoD, the DT ensure adaptive decision-making under the dynamic agricultural environment, and its clear visualisation is provided in Fig. 1

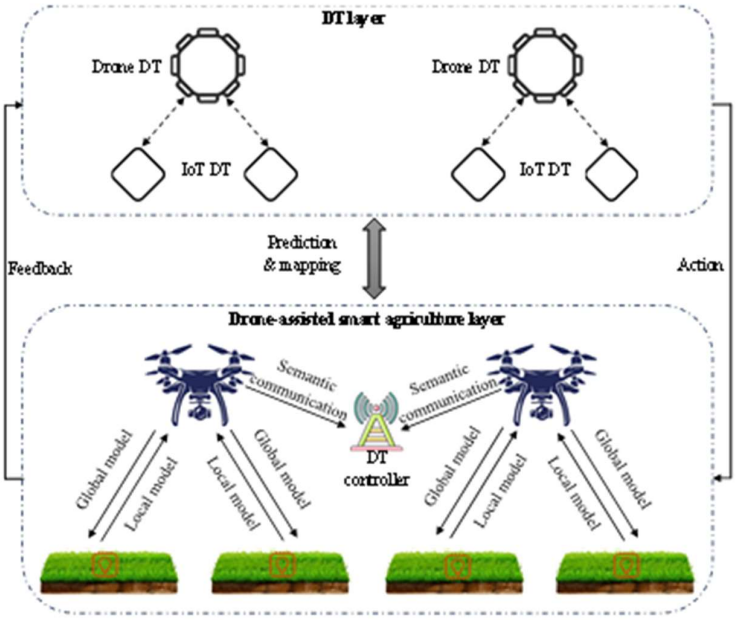


Fig. 1. System model of DT-assisted IoD network3.2 Explanation of Proposed Hybrid Framework

In general, digital agriculture highly depends on the combinational process of semantic communication, DT, and IoD to ensure intelligent data interpretation, continuous field perception and decision-making. The high-resolution sensory information is captured from the aerial platform to perform task execution, information transmission and adaptive computation. However, the conventional optimisation strategy generally faces issues in evaluating the strong coupling among the computational resource allocation, dynamic drone mobility, semantic information preservation, and communication efficiency. Computational frameworks like deep reinforcement learning models are developed to enhance intelligent resource management in the DT-IoD system. However, the existing models experience limitations like challenging contextual feature representation, inefficient exploration specially for the large action space and slow policy convergence. To address these challenges, this work developed a novel framework that jointly optimises the semantic communication and computational task for intelligent decision-making in a DT-assisted IoD network and its architectural diagram is depicted in Fig. 2

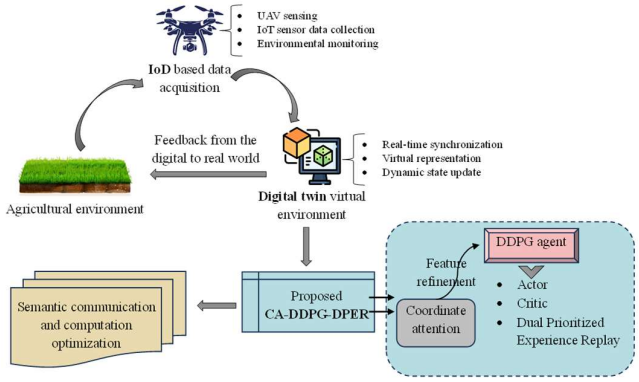


Fig. 2.Diagram of designed hybrid framework for semantic communication and computation in DT-assisted IoD

The implemented hybrid model is processed through a sequential process that incorporates intelligent decision-making, semantic communication and computational resource management in a DT-assisted IoD environment. At first, the agricultural sensory information is gathered using the IoD, which is determined in DT to manage the virtual view of the physical environment. The DDPG agent in the CA-DDPG-DPER model is provided with the updated state information. Here, the CA enhance the representation capability. Continuous control action is generated by the actor in the DDPG. The efficiency of the selected action is evaluated using the critic network based on the received reward from the environment. The replay memory is stored with the interaction experience from the DT-IoD environment. Here, the significant transition is performed using the DPER with better learning in the training phase. The implemented CA-DDPG-DPER attain adaptive computation and communication optimisation based on continuous interaction, which supports reliable and real-time agricultural applications.

4. Coordinate Attention with DDPG and Prioritised Model for Semantic Communication and Computation

4.1 DDPG

DDPG [16] is constructed based on the actor-critic model, which contains the behavioural value function and policy function. Here, the behavioural value act as a critic, which interacts with the environment and generates behaviour, whereas the policy function acts as an actor. In the learning process, the deterministic behaviour is added with a noise function to evaluate small-scale behavioural exploration. The dual parameter setting is attained by setting the critic as the target network and adding parameters to the actor network.

The present state δ initiates specific behaviour b based on the actor network. The behaviour value $K(s,b)$ is calculated by the critic network, where the subsequent state value is determined by the target actor network. In DDPG, the optimal behaviour is evaluated by the task of the actor, which also reduces the output behaviour value. In the transformation process, the samples are randomly selected to evaluate the objective function based on Eq. (3)

$$K_{T\arg et} = S_{i+1} + \gamma K'\left(s_i, B'\left(s_i \mid \theta^{B'}\right) \mid \theta^{Q'}\right) \tag{3}$$

In terms of temporal difference error, the loss of the critic network is computed using Eq. (4)

$$Loss = \frac{1}{X}\sum_{i}\left(K_{T\arg et} - K\left(s_i, b_i \mid \theta^{Q'}\right)\right)^2 \tag{4}$$

Here, the weights of the critic and actor networks are represented as θ^A and θ^Q. The term γ is the decay factor.

4.2 Coordinate Attention

CA [17] is an effective attention strategy that improves the feature representation by evaluating both the positional information and channel interdependencies. The spatial information decomposes into two-dimensional feature encoding in both the vertical and horizontal directions. This ensures the system preserves the precise location information. This enables better feature extraction without any computational overhead. The discriminative capability is enhanced through this CA based on the formulation in Eq. (5)

$$Z = Y \odot K_h \odot K_w \tag{5}$$

Here, both the attention maps that are generated in vertical and horizontal directions are denoted as K_w and K_h. The term Z is the input feature map, and $\odot$ defines the element-wise multiplication.

4.3 Developed CA-DDPG-DPER Model

To attain efficient semantic communication and computation resource management in DT-assisted IoD, an intelligent reinforcement learning framework called CA-DDPG-DPER is constructed. This framework is mainly designed to tackle the continuous decision-making problem by evaluating the optimal policy even under dynamic network conditions.

Initially, the current state information is generated from the DT-IoD environment to construct the model. This involved the semantic communication requirement, UAV resource availability, wireless channel condition and IoT devices. The state information is given to the agent in DDPG. In order to enhance the feature representation, the CA is employed, which

determines the contextual and spatial information. The control action is generated by the actor network based on the refined features from CA for managing the communication and computation.

The status is updated in DT-IoD by the generated action and gives feedback in the form of a reward. In the replay memory, the experiences are stored to enhance the learning capability of the DDPG agent. DPER [18] is the improved experience replay process that enhances reinforcement learning efficiency by providing the priority sample of the informative experiences. The importance of each transition is evaluated using DPER through its multiple priority criteria. This ensures the learning agent regains experience for effective policy enhancement. This strategy minimises the influence of redundant experiences and increases the sample utilisation of the redundant experiences, thereby improving the learning performance and stability of the DDPG model. The sampling probability is determined using Eq. (6)

$$Q(i) = \frac{q_i^{\alpha}}{\sum_k q_k^{\alpha}} \tag{6}$$

Here, the prioritisation factor is represented as α. The priority assigned to the ith transition is specified as q_i. The term $\sum_k q_k^{\alpha}$ is the normalisation. Here, the greater probability is achieved for the experiences with a higher priority value. In the training phase. This ensures the agent learns efficient and valuable interactions. The actor and critic network updated by the experience attained from DPER. The learning cycle and the repeated interaction continuously optimised the network parameters until it attains a stable policy.

The presented CA-DDPG-DPER offered intelligent decision-making for semantic communication and computation resource management in a DT-assisted IoD network. Efficient communication and computation are achieved by learning the interaction in a dynamic environment and enhancing the edge resources while satisfying the energy requirement and stringent latency. Hence, this framework is suitable for real-time applications in intelligent task execution for a dynamic network environment. The illustration of the proposed CA-DDPG-DPER is defined in Fig. 3

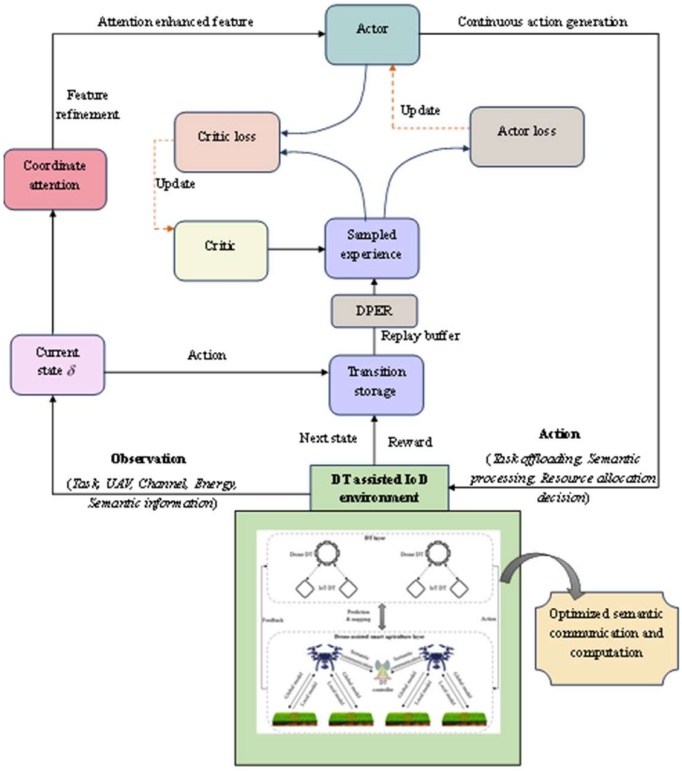


Fig. 3. The schematic diagram for the implemented CA-DDPG-DPER

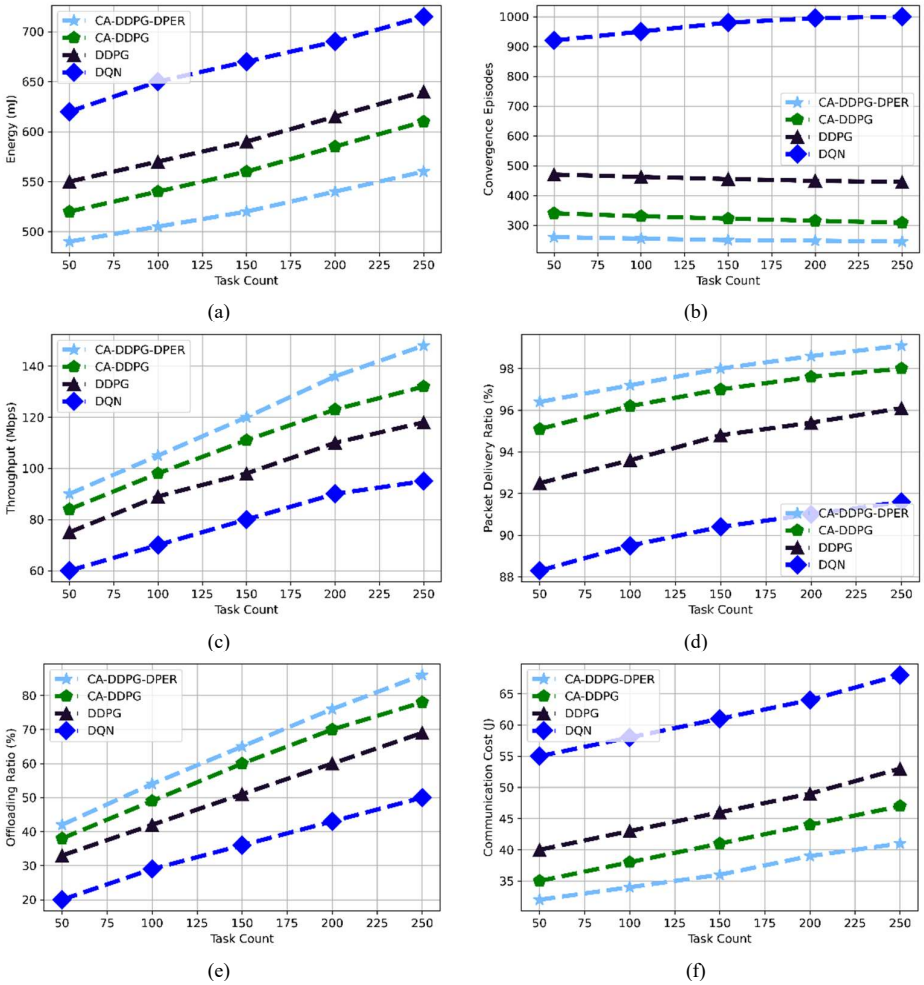
5. Results And Discussion

5.1 Experimental Setup

The presented DDPG model-based semantic communication and computation in DT-assisted IoD was evaluated in Python, and the experiments were performed. The required performance metrics were used to quantify the effectiveness of the framework. Further, the designed CA-DDPG-DPER was compared with existing models such as DQN [19], DDPG [16], and CA-DDPG [16][17] to know the efficiency of the model.

5.2 Performance analysis of CA-DDPG-DPER with existing approaches

The efficiency of the CA-DDPG-DPER is determined by differentiating its functionality from the earlier reinforcement learning models. The analysis is carried out by differentiating the task workload condition. This analysis is mainly to evaluate the functionality of the model in terms of the number of training iterations. From the obtained outcome, the CA-DDPG-DPER consistently attained a more reliable performance than 31.25% of DQN, 14.5% of DDPG, and 10.4% of CA-DDPG. From the outcome, the developed CA-DDPG-DPER has improved adaptive decision-making and learning capability. The overall outcome confirms that CA-DDPG-DPER has attained a scalable and efficient solution for intelligent communication and computation in DT-assisted IoD applications. Fig. 4(a) compares energy consumption, Fig. 4(b) illustrates convergence episodes, Fig. 4(c) presents throughput, Fig. 4(d) shows packet delivery ratio, Fig. 4(e) depicts the offloading ratio, Fig. 4(f) compares communication cost, Fig. 4(g) presents computation cost, Fig. 4(h) illustrates average queue length, Fig. 4(i) shows resource utilization, and Fig. 4(j) compares the energy efficiency of the proposed CA-DDPG-DPER framework with standard reinforcement learning methods under diverse task workloads.



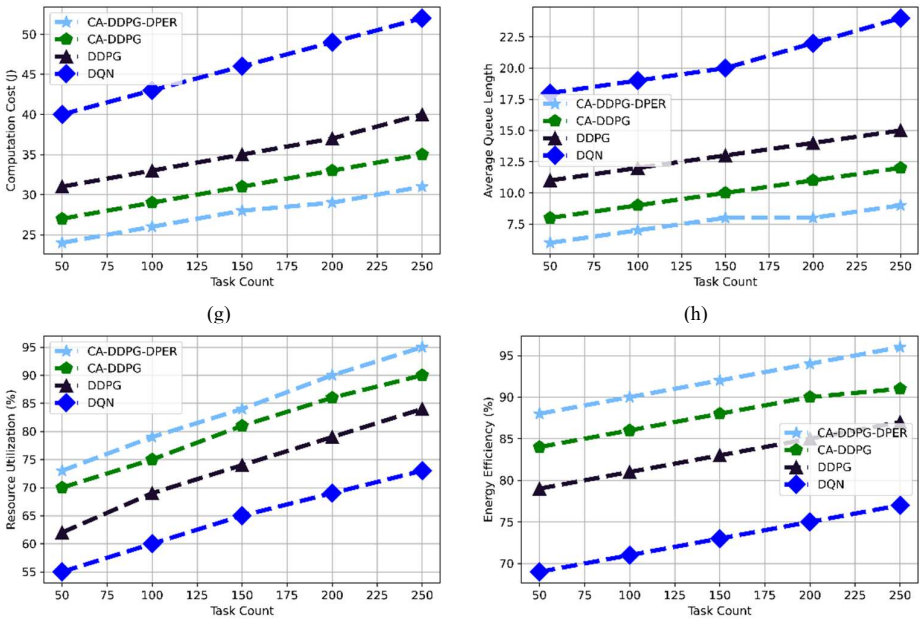


Fig. 4. Comparative assessment of CA-DDPG-DPER with standard reinforcement learning under diverse task workloads in terms of "(a) Energy, (b) Convergence episodes, (c) Throughput, (d) Packet delivery ratio, (e) offloading ratio, (f) Communication cost, (g) Computation cost, (h) Average queue length, (i) Resource utilization, (j) Energy efficiency"

6. Conclusion

This framework constructed an advanced reinforcement learning model to attain semantic communication and computation in DT-assisted IoD for smart agriculture. At first, the IoD was used to collect information from the agricultural environment through DT. The system information was optimised using the designed CA-DDPG-DPER model, which is the combination of the DDPG with the CA and DPER to enhance the feature representation and learning efficiency. The interaction between the agent and the IoD environment was effectively learned to perform the semantic processing even under different network conditions. Further, the implemented CA-DDPG-DPER was compared against previous reinforcement learning to conduct the experimental analysis. From the comparison, the CA-DDPG-DPER attained a better throughput value than 33.3% of DQN, 16.6% of DDPG, and 5.5% of CA-DDPG. From the analysis, the obtained outcome defines the superior performance of the CA-DDPG-DPER in terms of diverse performance metrics. Based on the result, the CA-DDPG-DPER provides a scalable solution for real-time decision-making in DT-assisted IoD.

Limitations and Future Direction: Even with the efficient performance of CA-DDPG-DPER, a notable limitation while training the model is the dependency on simulation-based environmental interaction, which minimises the adaptability for complex real-world agricultural scenarios. It also affects the generalisation capability of the model. Hence, in future, it will be recommended to develop a transfer learning model to ensure robustness in real-time DT-assisted IoD networks.

References

[1] Arsalan, A., Umer, T., Rehman, R.A., Kim, B.S., Murtaza, M. and Cheng, C.T., "Semantic-aware communication and computation co-optimisation in digital twin-assisted Internet of Drones networks using federated deep reinforcement learning", vol.12, pp. 517-522, 2026.

[2] Rahman, A.A., Mridul, M.C., Roy, P., Razzaque, M.A., Saleh, M.R., Hassan, M.M. and Uddin, M.Z., "A Multi-Head Attention mechanism assisted MADDPG algorithm for real-time data collection in Internet of Drones", Vehicular Communications, vol.54, pp.100944, 2025.

[3] An, D., Joo, H. and Kim, H., "Enabling low-latency digital twins for large-scale UAV networks using MQTT-based communication framework", ICT Express, vol.11(2), pp.264-269, 2025.

[4] Sundarakantham, K. and Dharani, J., "Generative AI-based blockchain framework for secure and private communication in internet of drones systems", Knowledge-Based Systems, pp.115813, 2026.

[5] Tong, H., Song, Z., Zhu, Z. and Sun, J., "Digital Twin-Driven Trajectory and Resource Optimization for UAV Swarms in Low-Altitude Urban Logistics and Communication Environments", Drones, vol.10(5), pp.376, 2026.
[6] Yuan, J., Li, L., Gong, S., Gu, B. and Li, F., "Digital Twin Enabled Simultaneous Learning and Modeling for UAV-Assisted Secure Wireless Sensing in Unknown Environment", Communications Technologies, pp. 1524-1529, 2025.
[7] Haider, Z.A., Fayaz, M., Zhang, Y. and Ali, A., "Advanced Hyperelliptic Curve-Based Authentication Protocols for Secure Internet of Drones Communication", ICCK Transactions on Advanced Computing and Systems, vol.1(3), pp.138-153, 2025.
[8] M. Chen, F. Shu, M. Zhu, D. Wu, Y. Yao and Q. Zhang, "Reinforcement-Learning-Based AAV 3-D Target Tracking and Digital-Twin-Assisted Collision Avoidance With Integrated Sensing and Communication," IEEE Internet of Things Journal, vol. 12, no. 13, pp. 24916-24928, 2025.
[9] Jianhang Tang; Jiangtian Nie; Jingpan Bai; Ji Xu; Shaobo Li; Yang Zhang, "UAV-Assisted Digital-Twin Synchronization With Tiny-Machine-Learning-Based Semantic Communications," IEEE Internet of Things Journal, vol. 11, no. 17, pp. 28437-28451, 1 Sept.1, 2024.
[10] Basharat, M. and Ferdouse, L., "Multi-Criterion Digital Twin-assisted Task Offloading with Chance Constrained Optimization in UAV Networks", IEEE Internet of Things Journal, 2025.
[11] A. Arsalan, T. Umer, R. Asif Rehman and B. -S. Kim, "Next-Gen Internet of Drones: Federated Learning and Digital Twin Synergy for Energy-Efficient Task Allocation and Seamless Service Migration," IEEE Access, vol. 13, pp. 64459-64472, 2025.
[12] Li, B., Liu, W., Xie, W., Zhang, N. and Zhang, Y., "Adaptive digital twin for UAV-assisted integrated sensing, communication, and computation networks", IEEE Transactions on Green Communications and Networking, vol.7(4), pp.1996-2009, 2023.
[13] B. Hu, H. Liu, H. Cao, S. Yuan, X. Chu and S. Xu, "Digital Twin-Assisted Large AI Task-Aware Edge Offloading and Resource Allocation for Low-Altitude Wireless Sensor Networks," in IEEE Journal of Selected Areas in Sensors, vol. 3, pp. 159-170, 2026.
[14] Qadir, M.A., Naeem, M. and Ejaz, W., "Digital twin-assisted multi-layer networks for low-latency and energy-efficient communication", Computer Communications, pp.108219, 2025.
[15] Kuru, K., Worthington, S., Ansell, D., Pinder, J.M., Sujit, A., Watkinson, B.J., Vinning, K., Moore, L., Gilbert, C., Jones, D. and Tinker-Mill, C., "Aitl-wing-hitl: Telemanipulation of autonomous drones using digital twins of aerial traffic interfaced with wing", Robotics and Autonomous Systems, pp.105486, 2026.
[16] Wu, J., Yang, Z., Liao, L., He, N., Wang, Z. and Wang, C., "A state-compensated deep deterministic policy gradient algorithm for uav trajectory tracking", Machines, vol.10(7), pp.496, 2022.
[17] Tsai, T.H. and Tseng, Y.W., "BiSeNet V3: Bilateral segmentation network with coordinate attention for real-time semantic segmentation", Neurocomputing, vol.532, pp.33-42, 2023.
[18] Li, S., Shi, P., Yang, A., Qi, H. and Dong, X., "Dual-Priority Delayed Deep Double Q-Network (DPD3QN): A Dueling Double Deep Q-Network with Dual-Priority Experience Replay for Autonomous Driving Behavior Decision-Making", Algorithms, vol.18(5), pp.291, 2025.
[19] Wang, X., He, C., Jiang, W., Wang, W., Li, L. and Xie, X., "UAV-Assisted Task Offloading and Resource Allocation in Internet of Vehicles: An Integration of Digital Twin and Generative AI", IEEE Transactions on Network Science and Engineering, vol.13, pp.5038-5055, 2025.