Deep Reinforcement Learning-Based Resource Allocation for Digital-Twin-Enabled Cognitive Semantic Communications in 6G Wireless Networks

Mahmoud Abir Atmaz Alsibaee
Department of Computer Science (Cybersecurity Track),
College of Technology and Business,
Riyadh ELM University, Riyadh, Saudi Arabia.
mahmoud.abir2024@student.riyadh.edu.sa

Abstract: Cognitive Semantic Communication (CSC) in 6G wireless networks represents a sophisticated paradigm that integrates semantic understanding with cognitive processes to fine-tune the data transmission. This paradigm shifts the communication focus from syntactic data delivery to intelligent interpretation of contextual meaning, enabling smarter network behavior. By targeting on the meaning and relevance of information other than raw data, this advanced approach increases efficiency, lowers bandwidth consumption, and promotes reliability, which renders it optimal for dynamic, high-necessity 6G environments with various application requirements. In this work, propose an advanced Adaptive Self-RAG with Deep Learning (ASRAG-DRL) framework is proposed for resource allocation in CSC in 6G wireless networks using a Digital Twin approach. The suggested structure incorporates real-time digital twins of communication nodes, recording their physical, cognitive, and network states, in contrast to traditional semantic systems that depend on static knowledge. This real-time mirroring ensures continuous synchronization between physical and virtual entities, improving adaptability and resilience. Predictive, context-aware decision-making enabled by the Digital Twin (DT) permits proactive modifications to network conditions. The DRL balances semantic correctness, energy efficiency, latency, and security via entropy-based reasoning and semantic knowledge graphs. To improve the performance of the proposed ASRAG-DRL, the Market Game Optimization Algorithm (MGOA) is leveraged. The method improves communication efficiency and dependability by concentrating on the delivery of valuable information rather than raw data. Simulation results show significant gains in resource utilization, semantic precision, and improved performance in dynamic network scenarios. Intelligent 6G services have a solid basis thanks to the DT-based Self-RAG framework.

Keywords: Resource Allocation; Digital-Twin; Cognitive Semantic Communications; Self-RAG with Deep Reinforcement Learning; Market Game Optimization Algorithm; 6G Wireless Networks

Received: December 11, 2025 Revised: December 19, 2025 Accepted: December 23, 2025

Nomenclature

Abbreviation	Description
CSC	Cognitive Semantic Communication
DT	Digital Twin
ASRAG-DRL	Adaptive Self-Retrieval Augmented Generation with Deep Reinforcement Learning
Self-RAG	Self-Retrieval Augmented Generation
DRL	Deep Reinforcement Learning
DQN	Deep Q-Network
DDPG	Deep Deterministic Policy Gradient
MGOA	Market Game Optimization Algorithm
NBOA	Namib Beetle Optimization Algorithm
MRA	Mud Ring Algorithm
DOA	Dream Optimization Algorithm
NOA	Narwhal Optimization Algorithm
CSCA	Cognitive Semantic Communication Agent
XAI	Explainable Artificial Intelligence
XAI-SCS	Explainable AI-based Semantic Communication System
QoE	Quality of Experience

1. Introduction

1.1 Background of Semantic Communication in 6G

CSC is a developing approach that goes beyond traditional bit-based communication by concentrating on the information transferred in 6G wireless networks [6]. It significantly reduces bandwidth requirements and energy usage by extracting, encoding, and transmitting [7] just the task-relevant elements that are essential to the user's goal rather than giving raw data indiscriminately [8]. CSC may dynamically adjust to changing network conditions and

application needs by integrating context-aware inference and knowledge graphs [9], resulting in reduced latency and greater dependability than traditional systems [10]. In addition to reducing the difficulty of intricate error-correction coding, this transition from syntactic to semantic representation allows for intelligent information prioritizing [11], guaranteeing reliable performance in extremely dynamic 6G situations [12]. In the end, CSC promises to enable real-time interactions across a variety of applications, with emerging levels of efficacy for novel 6G networks [13].

1.2 Challenges in Resource Allocation

The existing semantic communication frameworks continue to encounter a number of important obstacles [14]. Despite their promising capabilities, practical deployment challenges remain significant. First, they frequently rely on static, pre-established feature codebooks, which are insufficient to adjust semantic extraction to changing channel conditions and a variety of material [15]. Second, their reliability significantly deteriorates due to their lack of integrated systems [16] for real-time semantic distortion identification and correction, especially in noisy or unfavorable transmission circumstances [17]. Consequently, maintaining semantic fidelity becomes a critical concern in practical implementations. This flaw leads to poorer performance under difficult circumstances. The overall effectiveness and adaptability of the framework are hampered by this imbalance, particularly in highly changing network situations. In order to increase the resilience, flexibility, and dependability of semantic communication in 6G networks, it is imperative to solve these constraints.

CSC has recently surfaced as a viable paradigm with the goal of improving interpretability by imitating human cognitive processes to specifically receive, study, and encode cital semantic information. This human-inspired approach enhances the system's capability to prioritize meaningful information over redundant data. By utilizing cognitive-inspired semantic extraction, the integration of cognitive capacity with semantic communication offers a new way to overcome current constraints [18]. The majority of current image categorization techniques either fail to successfully incorporate cognitive insights into the communication process or ignore semantic-level information [19]. This limitation reduces their effectiveness in bandwidth-constrained environments. This frequently leads to less robustness and inferior efficiency, especially in situations where communication is limited [20]. As a result, creating a novel framework for cognitive semantic approaches in the context of 6G wireless networks is imperative. Such a framework should integrate adaptability, intelligence, and real-time responsiveness. Therefore, a deep learning-based resource allocation framework for DT-enabled CSC in 6G wireless networks was developed.

1.3 Contributions

This section scrutinized the major research contributions of the proposed framework. The contributions emphasize both theoretical innovation and practical applicability.

- To implement CSC in 6G wireless networks using a DT approach, the system emulates real-world environments to improve the communication efficiency. The DT model acts as a virtual replica that continuously synchronizes with physical network entities. This model allocates the resources by predicting the network behavior, streamlines data transmission, and permits real-time adaptation to unpredictable circumstances by merging intellectual abilities with semantic comprehension. Detailed modeling of network connections is facilitated by the digital twin technique, which enhances performance in intelligent, data-driven environments by decreasing latency, increasing dependability, and enabling more intelligent, context-aware decisions for 6G applications.
- To execute ASRAG-DRL based resource allocation model for DT-enabled CSC in 6G networks. The framework combines contextual knowledge and adaptive learning to handle network resources. This integration allows the system to dynamically learn optimal policies from continuously evolving environments. While DRL constantly adjusts resource allocation according to network conditions, Self-Retrieval-Augmented Generation (Self-RAG) increases semantic understanding through self-retrieval and knowledge refining. This dual mechanism enhances both learning efficiency and semantic reasoning accuracy. By replicating real-time network behavior via integration with digital twin models, the architecture supports logical, context-aware communication services in sophisticated 6G environments through providing predictive optimization, lower latency, increased dependability, and effective data transmission.

- To optimize the proposed framework for DT-enabled CSC, the MGOA is utilized. MGOA adaptively fine-tunes key hyperparameters, learning rate and number of episodes, by simulating competitive market strategies among agents. This bio-inspired optimization mechanism improves convergence speed and avoids local optima. By facilitating speedier convergence, more prediction accuracy, and effective, latency-sensitive resource allocation, this modification enhances DRL performance. As a result, in intricate 6G wireless networks, MGOA improves overall network stability, context-aware resource management, and operational efficiency.

The organization of the research work is represented here. The literature review and research gap analysis are visualized in Section 2. The proposed digital twin enabled semantic communication framework is discussed in Section 3. The Section 4 details on the deep reinforcement learning based resource allocation succeeded by the results and discussion section 5, and the conclusion in Section 6.

2. Literature Review and Research Gap Analysis

2.1 Review of Existing Digital Twin Based Communication Systems

CSC in 6G wireless networks has suggested various research works, which were a motivation for implementing this research work. These studies collectively emphasize the transition toward intelligent, context-aware communication paradigms. In 2026, Zhu *et al*. [1] explored the Next-Generation Semantic Communication (NextG-SemCom). This was investigated to fully allow the world with human-like cognitive capabilities. Such capabilities aim to replicate human reasoning in communication processes. In the year 2025, Li *et al.* [2] developed a cognitive SemComs exploit external knowledge to understand the reasoning similar to brain during semantic transmission. This work targets to survey the outcomes of the recent researches and the significant implementation of the cognitive SemComs are investigated. It also provides insights into the challenges and opportunities in deploying cognitive semantic systems. An edge Cognitive SemCom Agent (CSCA) was presented to leverage edge multimodal SemCom by Sun *et al*. [3] in late 2025. The introduction of edge intelligence reduces latency and enhances real-time decision-making capabilities. The CSCA employed an edge large model to study on the natural language intent understanding and alignment modality. Another research work was proposed in 2025 by Zhang *et al.* [4] leveraging deep generative models to generate a novel paradigm. A semantic communication framework integrated with XAI (XAI-SCS) was developed by Chen *et al*. [5] in the late 2025. This work also developed a contrastive explanation consistency metric in order to compute the key feature's sensitivity which also increases the reconstruction quality.

2.2 Research Gaps

Real-time network optimization is improved by integrating CSC utilizing a digital twin technique in 6G wireless networks. This integration enables continuous monitoring and adaptive control of network resources. It increases data transmission efficiency and adaptability to dynamic situations by mimicking real-world conditions. This technique advances 6G networks for a variety of intelligent applications with lower latency and improved dependability by forecasting network behavior, optimizing bandwidth, and enabling better, context-aware decisions. Despite these advancements, several critical challenges still hinder full-scale deployment. However, current models encountered a number of difficulties with cognitive semantic communications on 6G wireless networks, which are listed below and will be resolved in the proposed investigation.

- Prediction reliability is decreased by traditional models' inability to handle noisy inputs, incomplete data, or inadequate context. This limitation significantly affects performance in real-world, uncertain environments. Self-RAG uses self-reflection scores to assess results and dynamically collects pertinent information. For real-time applications, this adaptive retrieval guarantees context-aware decision-making, increases prediction accuracy, and strengthens robustness.
- Reliability, energy efficiency, and latency are frequently not balanced by traditional optimization techniques. This imbalance leads to inefficient utilization of limited network resources. Dynamic, multi-objective resource allocation is performed by combining reinforcement learning with Self-RAG. In real-time 6G semantic communication networks, chronic illness management, and patient monitoring, this provides reliable performance, lowers errors, and increases predictive accuracy.

- Conventional DRL models don not have transparency, which makes the decision processes complex to interpret in DT-enabled 6G networks. This lack of interpretability reduces trust in automated decision-making systems. To address this, and optimization technique is incorporated that fine-tunes model parameters, increasing the performance, enhancing interpretability, and providing reliable, context-aware, and low-latency resource allocation.

3. Proposed Digital Twin Enabled Semantic Communication Framework

With the introduction of CSC in 6G wireless networks, relevant information instead of raw data may be transmitted, increasing communication effectiveness and facilitating wise decision-making. This paradigm significantly enhances the ability of networks to interpret and act upon meaningful information rather than redundant data. However, efficient resource allocation is severely hampered by dynamic network conditions, scarce resources, and the requirement for semantic precision. Thus, an upgraded ASRAG-DRL paradigm for resource allocation in DT-enabled CSC systems is suggested in this study. The proposed framework integrates learning, reasoning, and optimization into a unified architecture.

For communication nodes, real-time DT models were initially developed. These models function as synchronized virtual counterparts of physical network entities. The physical status, network conditions, and semantic knowledge states of the nodes are continuously collected by these digital replicas. The DT allows for real-time monitoring, prediction, and adaptive resource allocation by controlling the network resources, in contrast to traditional systems that depend on static knowledge bases.

The Self-RAG mechanism in this approach enhances contextual comprehension during conversation by dynamically retrieving pertinent semantic content from knowledge networks. It further refines semantic interpretation through iterative self-evaluation and feedback mechanisms. To further improve resource allocation choices, a DRL agent is incorporated. By interacting with the environment, the DRL agent learns the best policies and assesses a variety of goals, including security, optimal resource utilization, latency, energy efficiency, and semantic accuracy. This multi-objective optimization ensures balanced performance across conflicting requirements.

The decision-making process is guided by semantic knowledge graph representations and entropy-based reasoning. These mechanisms help quantify uncertainty and improve the reliability of semantic decisions. To enhance the model performance, the MGOA is used to optimize the parameters. As a result, the suggested framework offers a effective resource allocated outcome for intelligent 6G applications including distributed AI systems, smart cities, and driverless cars. This demonstrates the scalability and applicability of the framework across diverse real-world scenarios. The pictorial view of the proposed digital twin-enabled CSC framework for resource allocation is given in Fig. 1.

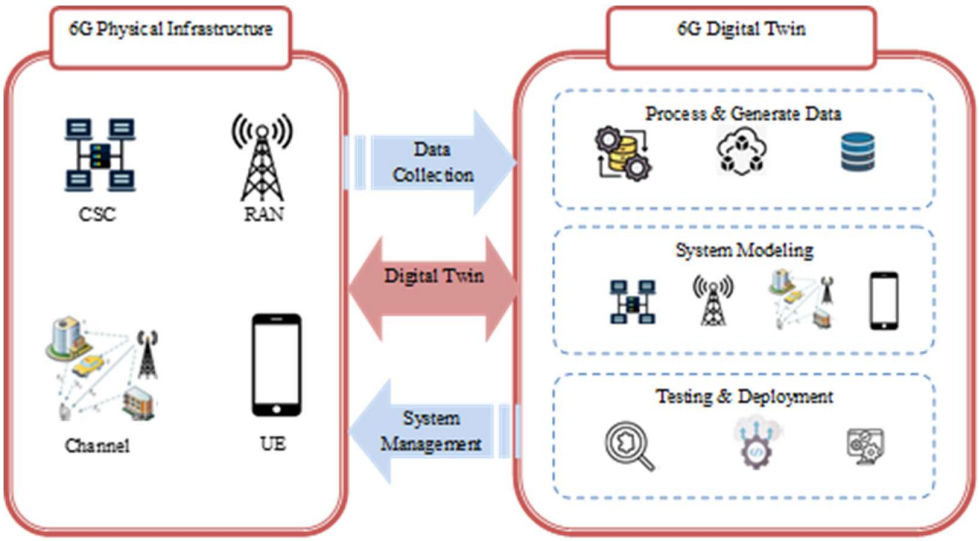


Fig. 1. Digital Twin-enabled CSC Framework for Resource Allocation Model

3.1 6G-enabled Digital Twin System Model

Every node $D_b(u \in \{1,2,...,D\})$ includes a DT representation that changes based on a set of state variable $F_b(h)$. Eq. (1) defines the physical, cognitive, and communicative states that consists the state vector. These states collectively capture both tangible and abstract aspects of the network node. The physical state in Eq. (1) is represented by $K_b(h)$, which includes factors like temperature, location, and energy level. The cognitive state is represented by $V_b(h)$, which stands for knowledge, belief, and decision-making capacity. This component enables intelligent reasoning and adaptive responses. The communication state is represented by $B_b(h)$, which describes connectivity status and communication channels. Furthermore, the historical background is represented by $X_b(h)$, which captures prior events, trends, and influences affecting the node. These digital twins provide real-time semantic communication and predictive analytics for intelligent network management with the development of 6G.

$$F_b(h) = \{K_b(h), V_b(h), B_b(h), X_b(h)\} \tag{1}$$

The state dynamics and the state dynamics of each node $F_b(h)$ is used, which is the current's derivative state $F_b(h)$ to examine the evolution of state variables and can be explained by Eq. (2), which takes the node's current condition, illustrating the intricacy and ambiguity inside the DT framework. This formulation enables modeling of nonlinear and stochastic behaviors in the system. It explains the node's state is evolvement at a specific time where W_b in Eq. (2) regarding the node J_b is the function for state transition, and it describes how the node's state varies over time, and it depends on a number of variables, including $F_b(h)$, control inputs $G_b(h)$ in relation to the node (adjustments, commands), noise $N(h)$ influencing the behavior of node, and Entropy $\partial(h)$ of the node's knowledge graph. The entropy $\partial(h)$ evaluates the information's intricacy and degree of uncertainty information and connections about node J_b in the knowledge graph.

$$F_b(h) = W_b(F_b(h), G_b(h), N(h), \partial_b) \tag{2}$$

The digital twin-enabled system model is modeled in Fig. 2.

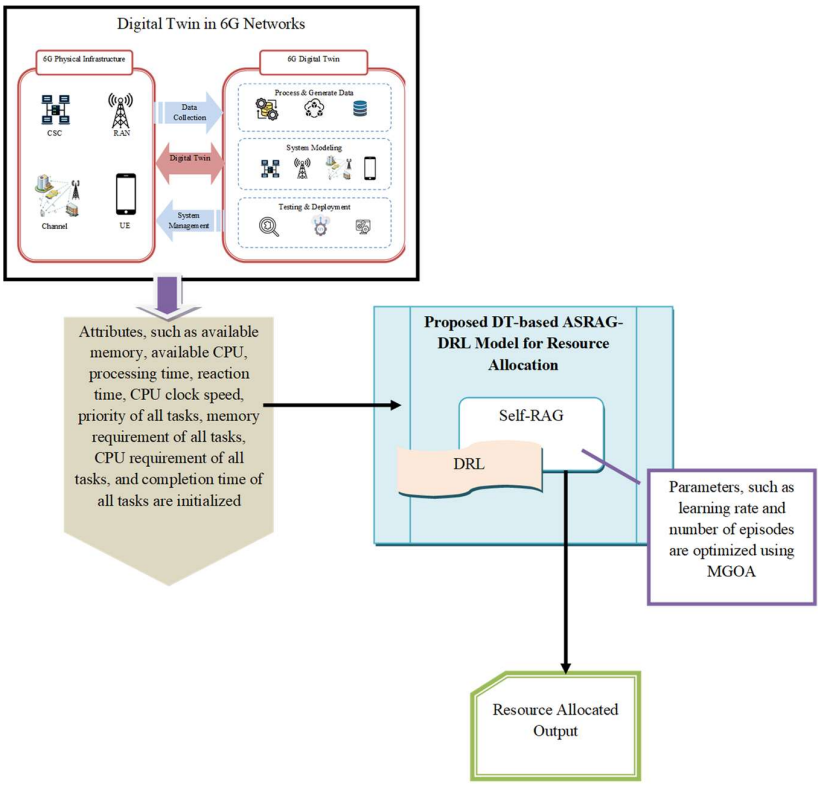


Fig. 2. Digital Twin-enabled System Model

3.2 Semantic Communication Layer

Semantic distortion, resource efficiency, latency-energy trade-off, and security are among the criteria used by the system to assess semantic interaction performance in a knowledge-graph-based 6G network. These metrics collectively define the effectiveness of semantic communication systems. The loss of meaning during transmission is measured by semantic distortion. It computes the weighted difference between received semantic information H_w and transmitted semantic information $H_w^{'}$. The knowledge graph's vertices are stated as v and edges as $M=(V,\varepsilon)$. These graph components model relationships between semantic entities. The significance of every node and relationship in the graph is represented by the weighting function $\varphi(v,q)$. Higher semantic fidelity is indicated by less distortion. Eq. (3) presents the mathematical representation of semantic distortion.

$$Z_w=\int_{v1}^{v2}\int_{\varepsilon 1}^{\varepsilon 2}\varphi(v,q)\left\|H_w^{'}(v,q)-H_w^{'}(v,q)^2\right\|d_v d_q \tag{3}$$

Here, the terms d_v and d_q are the integration over vertices and edges. Resource efficiency measures how effectively the communication resources are utilized. It contrasts the bandwidth C and transmission power Y_s with the semantic rate L_w.This comparison highlights the trade-off between resource usage and communication performance. The increased efficiency facilitates the transmission of more significant information with less resource. The equation for resource efficiency is given in Eq. (4).

$$\varepsilon_b=\frac{Y_s}{C.L_w} \tag{4}$$

The semantic rate Y_s is based on Shannon capacity and is influenced by bandwidth, transmission power, interference I, noise χ^2, channel gain $A(s)$, and channel gain H(t) .These factors collectively determine the achievable communication performance. By reducing meaning loss, optimizing resource consumption, and preserving dependable data transmission in cutting-edge 6G networks, these metrics work together to optimize semantic communication.

4. Deep Reinforcement Learning Based Resource Allocation

To enhance resource allocation in DT-enabled systems, the ASRAG-DRL method combines adaptive knowledge retrieval with reinforcement-based decision-making. This hybrid approach enables both intelligent reasoning and dynamic adaptation in complex network environments. To facilitate accurate support, the Self-RAG dynamically collects contextual data form knowledge base and evaluates through the utilization of self-reflection assessments. The DRT module then uses reward-based feedback to train optimal allocation strategies based on the retrieved contextual information. This feedback-driven mechanism allows continuous improvement of allocation policies over time. Further, the MGOA is incorporated to optimize resource allocation parameters and improve convergence efficiency. Thus, in complex DT-based systems, the ASRAG-DRL framework efficiently facilitates strategic decision-making and intelligent, real-time resource management.

4.1 MGOA

The MGOA is used to optimize the parameters of the proposed ASRAG-DRT framework. It plays a crucial role in enhancing the learning efficiency of the system. The MGOA [23] is a metaheuristic optimization method inspired by the strategic interactions between businesses and consumers in competitive markets. Candidate solutions are modeled as merchant agents that dynamically modify their strategies in response to feedback from the market. The two main stages of MGOA's operation are the (i) Collaborative Phase and (ii) Attraction Phase. These phases ensure a balance between exploitation of known solutions and exploration of new possibilities.

(i) Collaborative Phase: Collaborative Phase concentrates on local exploitation where high-performing agents interact and improve promising solutions and the update equation is presented in Eq. (5). This phase enhances convergence speed by refining near-optimal solutions.

$$Rw_{d,z}^{new}=Rw_{d,z}-r_1\cdot q_1+r_2\cdot q_2 \tag{5}$$

Here, the term $Rw_{d,j}$ denotes the current strategy vector of merchant, r_1 and r_2 are the random cooperation coefficients controlling influence of other merchants, q_1 and q_2 are the difference vectors between two randomly selected merchants and the customers.

(ii) Attraction Phase: Attraction Phase emphasizes global exploration through permitting agents to modify strategies and search various regions of the solution space and the update equation is stated in Eq. (6). This phase prevents premature convergence to local optima.

$$Rw_{d,j}^{new} = Rw_{d,j} - T \times (W_h - Rw_{d,j}) \tag{6}$$

Here, the term $Rw_{d,j}$ denotes the current strategy of merchant, W_h represent the best-known strategy, and T states the strategy adjustment strength controlling step size. The algorithm may prevent local optima and progressively converge toward globally optimal solutions by establishing a balance between exploration and exploitation. This balance is essential for achieving robust optimization performance.

Using MGOA, the hyperparameters, learning rate and number of episodes of the ASRAG-DRT framework are efficiently optimized to achieve the objectives, which focuses on reducing latency, maximizing semantic accuracy, and improving energy efficiency. This targeted optimization directly enhances the effectiveness of DRL training. This optimization ensures precise and adaptive resource allocation in the DT-based 6G CSC system. The framework improves network performance, lowers latency, and permits context-aware, low-latency resource management by offering optimum parameters.

4.2 Deep Reinforcement Learning

DRL [21] integrates deep neural networks with reinforcement learning to address challenging decision-making issues. It is particularly effective in high-dimensional and dynamic environments such as 6G networks. The DRL's objective is to increase cumulative rewards over time with an agent interacting with environment through state observation, action, and reward. The Deep Q-Network (DQN), a popular DRL technique, the action-value function and leads the agent toward the best choices. DQN approximates optimal policies using deep neural networks. Techniques like experience replay and target networks are frequently employed to increase learning stability. The model's goal is to reduce the difference between the target and forecasted Q-values, which may be stated in Eq. (5).

$$V(\partial) = R\left[(p - R(u, z|\partial))^2 \right] \tag{7}$$

Here, the term u denotes the state, z represent the action, and ∂ is the network parameters.

4.3 ASRAG-DRL: DT-based Resource Allocation Model

A novel ASRAG-DRL framework is developed to improve intelligent resource allocation in DT-based resource allocation systems. This framework integrates semantic reasoning, adaptive retrieval, and reinforcement learning into a unified model. The approach optimizes information usage, increases resource allocation efficiency, and permits quicker, context-aware judgments in dynamic contexts through using adaptive retrieval processes, self-evaluation techniques, and reinforcement-driven optimization.

4.3.1 Novelty

The proposed ASRAG-DRL combines Self-RAG with a DRL policy for DT-based resource allocation. This integration introduces a new dimension of semantic awareness in reinforcement learning. In convoluted, real-time resource allocation scenarios, this integrated system greatly enhances adaptability, decision accuracy, and robustness by simultaneously learning optimal allocation strategies and context-aware information retrieval, while leveraging optimum parameters obtained via MGOA to enhance efficiency and resource allocation performance even more.

4.3.2 Integration of DT

Real-time virtual representation of network nodes is rendered by integrating DT into the ASRAG-DRL-based resource allocation model. This enables continuous synchronization between physical and virtual environments. This facilitates adaptive control, context-aware decision-making, and predictive monitoring, guaranteeing optimal resource utilization, low latency responses, and reliable performance in dynamic 6G environments.

4.3.3 Implementation Process

The attributes, such as processing time, available memory, CPU requirement of all tasks, priority of all tasks, available CPU, reaction time, memory requirement of all tasks, CPU clock speed, and completion time of all tasks are initialized and given to the Self-RAG of the proposed ASRAG-DRL. These parameters collectively define the system state for decision-making. A methodology called Self-RAG [22] blends dynamic retrieval, self-evaluation, and iterative refinement to increase the dependability and flexibility of large language models. A generator model and a critic model are its two primary parts. Special self-reflection tokens like ISREL (relevance), ISSUP (support), and ISUSE (utility) are inserted as the generator generates responses. The critic model scores these tokens by providing normalized probabilities between 0 and 1 to assess the generated content. The generator generates a Retrieve token to retrieve pertinent documents from an external corpus if it detects inadequate information during inference. Multiple candidate responses are generated and assessed using both the generation probability and critic feedback for every document that is retrieved. This ensures selection of the most reliable response. A generated segment's total quality is calculated as stated in Eq. (6).

$$b(h_x, g) = u(h_x | s, g, h_{<x}) + J(Critique) \tag{8}$$

Here, the term $b(h_x, g)$ is the composite quality score of the generated segment with respect to the retrieved document, h_x represent the generated token, h_x states the input query provided to the model, h_x is the retrieved document, $u(h_x | s, g, h_{<x})$ denotes the conditional probability assigned by the generator model, and $J(Critique)$ is the aggregated evaluation score that evaluates the quantitative quality of the generated content. These components collectively determine the effectiveness of generated outputs. By ensuring that the final response maintains a balance between language fluency and factual dependability and utility, this scoring system enhances response, credibility, and supports accurate resource allocation decisions.

Also, Self-RAG's iterative decision-making and evaluation process reflects DRL principles. This resemblance enables seamless integration between retrieval and learning mechanisms. Like an action-selection approach, the generator model generates responses and uses the Retrieve token to dynamically determine whether to initiate document retrieval. The created segments are then assessed by the critic model using self-reflection criteria like utility, support, and relevance. The model is guided to favor higher-quality responses by these scores, which function as reward-like feedback. This creates an optimization loop that is inspired by reinforcement and enhances response quality, reliability, and resource allocated outcomes. To further enhance the performance, the MGOA is used. Thus, optimization and learning work synergistically.

The proposed ASRAG-DRL DT-based resource allocation model is designed to maximize semantic communication in DRL-based resource allocation for DT-enabled CSC in 6G wireless networks by optimizing the parameters using the MGOA. This ensures improved system performance under dynamic conditions. The MGOA optimizes key parameters, such as learning rate and number of episodes to achieve maximization of energy efficiency and security, while ensuring the minimization of latency, facilitating reliable decision-making, efficient resource utilization, and better communication performance in dynamic 6G environments. The fitness function is stated in Eq. (7).

$$FF = \underset{\left\{Lr_d^s, Ne_d^s\right\}}{\arg\min}\left(\frac{1}{Ty + \mathrm{Re} + Pi} + La\right) \tag{9}$$

Here, the terms Lr_d^s and Ne_d^s denoted the learning rate, and number of episodes in SRAG-DRL with ranges of $[0.01 - 0.99]$, and $[10000 - 100000]$. The terms Ty, Re, Pi and La are semantic correctness, energy efficiency, security, and latency, respectively, are given in Eq. (10) and Eq. (13).

$$Ty = \frac{1}{K}\sum_{o=1}^{K} sim_q(U_q, \hat{U}_q) \tag{10}$$

Here, the term K depicts the total number of transmitted semantic messages, U_q is the original semantic message, and $\hat{U}_q$ represents the received or reconstructed semantic message.

$$\mathrm{Re} = \frac{G}{Wf} \tag{11}$$

Here, the term G is the total transmitted data, and Wf is the total energy consumption.

$$Pi = 1 - Y_{attack} \tag{12}$$

Here, the term Y_{attack} is the probability of successful attack.

$$La = \frac{Bp}{Tp} \tag{13}$$

Here, the term Bp is the total transmission time and Tp denotes the number if transmitted messages. Finally, the resource allocated output is obtained. Optimal policy learning and context-aware resource management are accomplished by combining MGOA optimization with the suggested ASRAG-DRL framework for DT-based resource allocation in 6G semantic communications. This integration ensures adaptive and intelligent system behavior. This synergy guarantees quicker convergence, improved decision accuracy, and reliable allocation in dynamic, diverse network contexts by efficiently capturing semantic, temporal, and system-state relationships. In order to provide intelligent, dependable, and real-time 6G network management, adaptive Self-RAG, critic-based evaluation, DRL policy learning, and MGOA parameter optimization greatly enhance resource efficiency. Together, these components form a comprehensive solution for next-generation communication systems. The architectural framework of the introduced SRAG-DRL-based DT-enabled resource allocation model in 6G networks is displayed in Fig. 3.

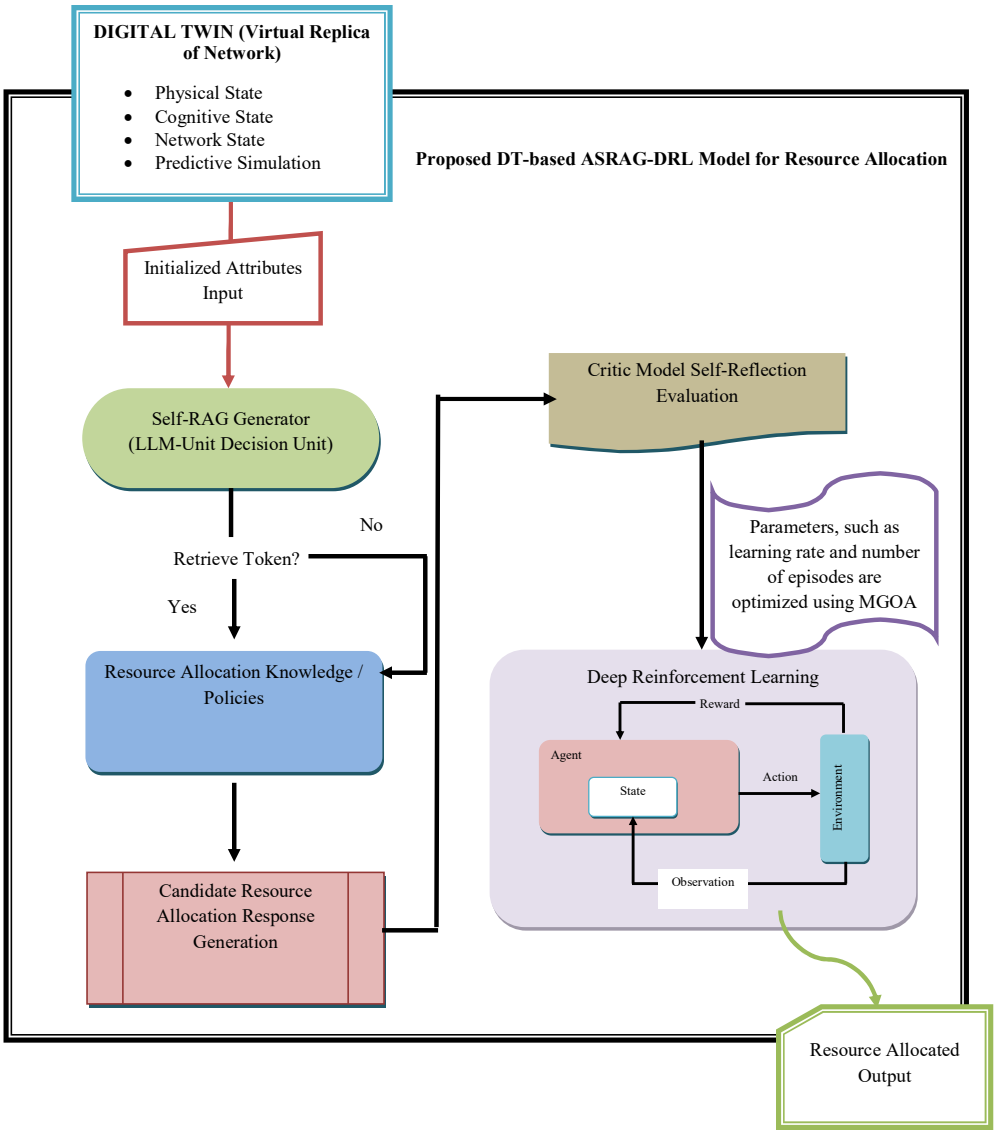


Fig. 3.Architectural Framework of the Introduced ASRAG-DRL-based DT-enabled Resource Allocation Model in 6G Networks 5. Results and Discussion

A resource allocation and decision-making model for DT-enabled 6G cognitive semantic communications was developed using a Python simulation framework. The simulation environment was designed to replicate dynamic and heterogeneous 6G network conditions. The optimization notations are modified with population size as 10, chromosome length as 2 and maximum iteration as 50. The experimental validation was conducted by comparing reinforcement learning networks, such DRL [21], DQN [24], Deep Deterministic Policy Gradient (DDPG) [25], and SRAG-DRL, respectively; and optimization algorithms, including Namib Beetle Optimization Algorithm (NBOA) [26], Mud Ring Algorithm (MRA) [27], Dream Optimization Algorithm (DOA) [28], Narwhal Optimization Algorithm (NOA) [29], and MGOA [23], to improve semantic correctness, energy efficiency, security, and latency in resource allocation. This comparative analysis highlights the strengths and weaknesses of existing approaches under identical conditions. Thus, the developed ASRAG-DRL framework effectively overcomes the limitations of classical approaches in dynamic 6G network environments.

5.1 Standard Performance Evaluation

The performance analysis of the developed ASRAG-DRL framework is estimated using the following formulae. These evaluation metrics provide a comprehensive assessment of system efficiency and responsiveness.

(a) Active Server :

$$U_s = N_{active} \tag{14}$$

Here, the term N_{active} is the count of servers, which handles the workload.

(b) Average Waiting Time (S):

$$S = \frac{1}{X}\sum_{e=1}^{X} P_e \tag{15}$$

Here, the term X denotes total number of tasks in system, and P_e is the waiting time of task before being served.

(c) Delay:

$$D = R + V_s \tag{16}$$

Here, the term V_s is service time, and R denotes waiting time in queue.

(d) Energy Consumption:

$$Y = \sum_{h=1}^{N active} F_v \cdot L_v \tag{17}$$

Here, the term F_v represents the power consumption of server and L_v represents time duration server that remains active.

5.2 Empirical Convergence Analysis

The empirical convergence analysis is represented in Fig. 4. This analysis evaluates how quickly and effectively the optimization algorithms reach optimal solutions. Based on the given input data, this resource allocation study is essential for forecasting ideal future network states. It also demonstrates the adaptability of the proposed framework under varying configurations. As illustrated in Fig. 4 (a), the generated model is assessed throughout different iterations and contrasted with current approaches, such as NBOA-ASRAG-DRL, MRA-ASRAG-DRL, NOA-ASRAG-DRL, and DOA-ASRAG-DRL. Thus, intelligent, real-time decision-making in 6G networks is rendered ideal by the suggested MGOA-ASRAG-DRL, providing effective resource usage and improved communication dependability. This confirms the robustness of the optimization strategy in dynamic environments.

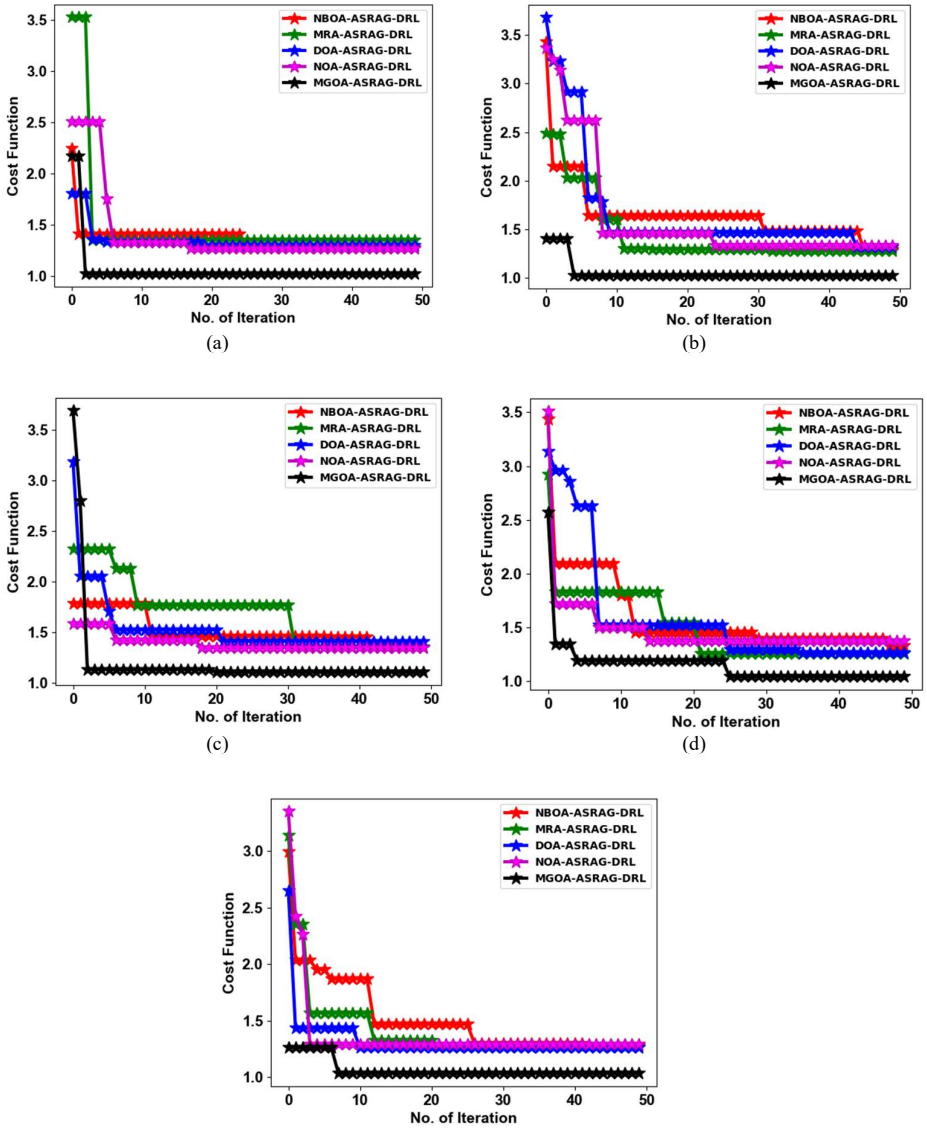


Fig. 4. Empirical Convergence Analysis regarding Configuration (a)1, (b)2, (c)3, (d)4, and (e)5

5.3 Quantitative Analysis of Results

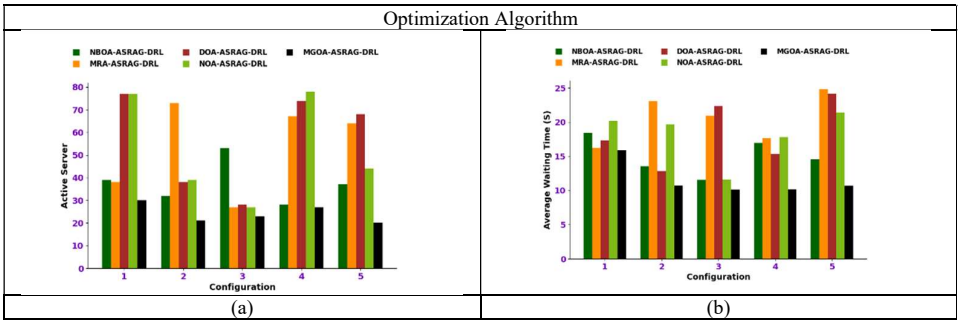
The quantitative analysis of results is represented in Table I. This statistical evaluation is crucial to evaluate the performance of the suggested model in forecasting system behavior under worst-case conditions. It highlights the robustness of the proposed approach under extreme scenarios. The worst-case for dataset-2 show that MGOA-ASRAG-DRL achieves values of approximately 11.3%, 31.9%, 50.3%, and 57.1% when compared to NBOA-ASRAG-DRL, MRA-ASRAG-DRL, DOA-ASRAG-DRL, and NOA-ASRAG-DRL, respectively. Thus, the results demonstrate that the developed framework shows robust performance with peak network loads, ensuring reliable, real-time resource allocation and enhanced efficiency in dynamic 6G semantic communication networks. The consistency across configurations further validates the effectiveness of the proposed model.

Table 1. Quantitative Analysis of Results

Optimizer	NBOA-ASRAG-DRL [26]	MRA-ASRAG-DRL [27]	DOA-ASRAG-DRL [28]	NOA-ASRAG-DRL [29]	MGOA-ASRAG-DRL
Configuration-1					
BEST	1.371	1.362	1.421	1.362	1.116
WORST	1.812	2.401	3.256	1.602	3.745
MEAN	1.536	1.723	1.552	1.407	1.209
MEDIAN	1.474	1.789	1.423	1.358	1.116
STD	0.159	0.338	0.307	0.086	0.443
Configuration-2					
BEST	1.312	1.401	1.287	1.29	1.045
WORST	2.365	3.412	1.902	2.601	2.248
MEAN	1.402	1.519	1.361	1.433	1.089
MEDIAN	1.328	1.366	1.292	1.281	1.035
STD	0.162	0.498	0.132	0.385	0.214
Configuration-3					
BEST	1.355	1.355	1.402	1.347	1.101
WORST	1.785	2.323	3.187	1.583	3.691
MEAN	1.514	1.698	1.535	1.393	1.197
MEDIAN	1.457	1.765	1.409	1.347	1.101
STD	0.147	0.324	0.293	0.077	0.428
Configuration-4					
BEST	1.287	1.271	1.269	1.304	1.047
WORST	3.072	3.201	2.715	3.405	1.278
MEAN	1.529	1.42	1.333	1.383	1.077
MEDIAN	1.479	1.271	1.269	1.304	1.047
STD	0.341	0.352	0.212	0.363	0.086
Configuration-5					
BEST	1.341	1.273	1.281	1.395	1.058
WORST	3.498	3.001	3.212	3.536	2.618
MEAN	1.613	1.508	1.598	1.497	1.171
MEDIAN	1.466	1.276	1.421	1.395	1.134
STD	0.385	0.341	0.538	0.324	0.231

5.4 Performance Assessment of the Designed Resource Allocation Approach

The performance assessment of the designed resource allocation approach is represented in Fig. 5. This assessment compares both optimization-based and learning-based approaches. In terms of configuration-4, the proposed MGOA-ASRAG-DRL model outperforms other optimization-based approaches, with NBOA-ASRAG-DRL achieving 29.000%, DOA-ASRAG-DRL reaching 69.000%, NOA-ASRAG-DRL recording 75.000%, and MRA-ASRAG-DRL attaining 80.000%, regarding active server. For network configurations based on delay, classical DRL models such as DRL, DQN, DDPG, and DRL achieve delays of 19.4s, 17.2s, 15.8s, and 14.5s respectively, whereas the MGOA-ASRAG-DRL reduces delay to 11.2s, demonstrating superior efficiency. This reduction in delay highlights the effectiveness of the integrated optimization and learning framework. These results confirm the suggested model's advantage with other conventional optimization and reinforcement learning methods for dynamic resource allocation in 6G semantic communication networks.



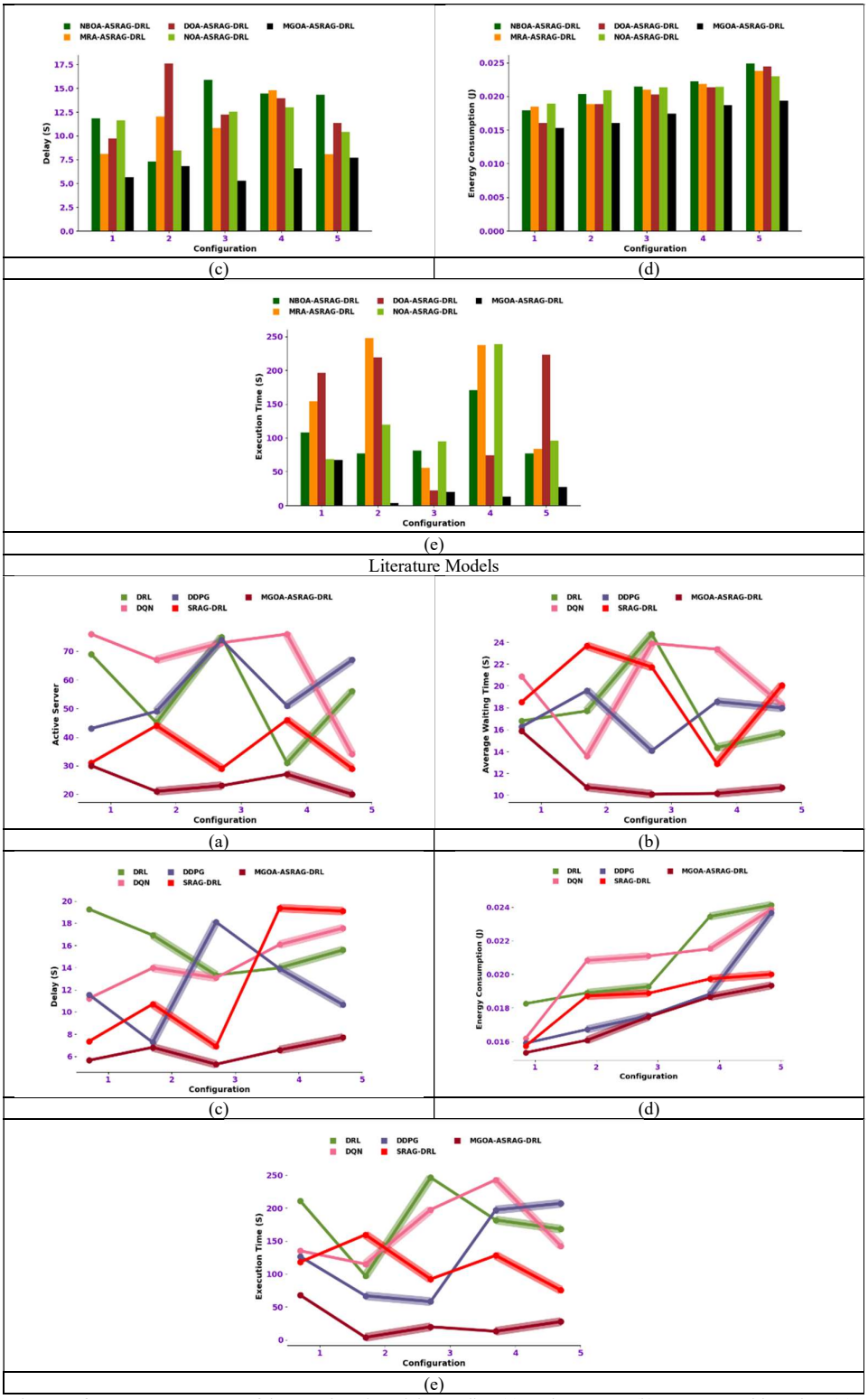


Fig. 5.Performance Assessment of the Developed Model regarding (a) Active Server, (b) Average Waiting Time (S), (c) Delay, (d) Energy Consumption, and (e) Execution Time

6. Conclusion

An innovative deep learning-based resource allocation framework for DT-enabled CSC in 6G wireless networks was developed in this work. The proposed framework integrates semantic

intelligence, adaptive learning, and optimization into a unified architecture. With improved predictive accuracy and dependable resource allocation, the model tackles important issues in real-time 6G semantic communication networks. The attributes were initialized and fed into the proposed ASRAG-DRL network for resource allocation. This guaranteed precise and context-aware forecasts, even in situations with noisy or insufficient data. Such robustness is essential for practical deployment in real-world networks. The DRL module then used reward-based feedback to maximize latency-sensitive resource allocation and decision-making. This learning mechanism enables continuous improvement of allocation strategies. To improve the suggested model's performance, the parameters, such as number of episodes and learning rate were optimized using the MGOA. Finally, the resource allocated output was obtained. The model's allocated resource showed better accuracy, robustness, and lower latency when compared to more conventional techniques like DRL, DQN, DDPG, and SRAG-DRL. These improvements highlight the effectiveness of combining retrieval, learning, and optimization strategies. The proposed ASRAG-DRL achieved values of 12.1%, 29.8%, 48.5%, and 55.6%, respectively, when compared to NBOA-ASRAG-DRL, MRA-ASRAG-DRL, DOA-ASRAG-DRL, and NOA-ASRAG-DRL, respectively, regarding execution time. When comparing with other networks, such as DRL, DQN, DDPG, and DRL achieve values of 18.7s, 16.5s, 14.9s, and 13.6s, respectively, regarding delay. Therefore, the developed DRL-based resource allocation model for DT-enabled CSC in 6G wireless networks enabled efficient, context-aware spectrum and resource management, low-latency interactions, and increased network reliability.

6.1 Limitation and Future Scope

The proposed model demand on substantial training data and computer power, which is a drawback that limits real-time implementation in dynamic 6G networks. This computational complexity may hinder deployment in resource-constrained edge devices. In order to enhance scalability, robustness, and effective resource allocation in extremely congested 6G environments, future research can concentrate on lightweight, adaptable models with transfer and online learning.

References

[1] Zhu, F., Chen, J., Wen, J., Yang, Y., Yi, C., Tie, Y., Zhang, P., Cai, J., Niyato, D. and Guizani, M., "From Data Mirror to Smart Copilot: A Survey on NextG Semantic Communication for Propelling Digital Twin World into Cognitive Stage," IEEE Communications Surveys & Tutorials, 2026.

[2] Li, Y., Zhou, F., Yuan, L., Wu, Q. and Al-Dhahir, N., "Cognitive semantic communication: A new communication paradigm for 6G," IEEE Communications Magazine, vol. 63, no. 6, pp. 122-129, 2025.

[3] Sun, Y., Liu, Y., Guo, S., Qiu, X., Chen, J., Hao, J. and Niyato, D., "Edge large AI model agent-empowered cognitive multimodal semantic communication," IEEE Transactions on Mobile Computing, 2025.

[4] Zhang, K., Li, L., Lin, W., Yan, Y., Li, R., Cheng, W. and Han, Z., "Semantic successive refinement: A generative AI-aided semantic communication framework," IEEE Transactions on Cognitive Communications and Networking, no. 11, no. 2, pp. 687-699, 2025.

[5] Chen, M., Gu, Y., He, X., Wang, S., Huang, F. and Wang, L.,"Explainable artificial intelligence enhance image semantic communication system in 6G-IoT," IEEE Internet of Things Journal, 2025.

[6] Hu, B., Ma, J., Sun, Z., Liu, J., Li, R. and Wang, L., "DRL-based intelligent resource allocation for physical layer semantic communication with IRS," Physical Communication, vol. 63, pp. 102270, 2024.

[7] Liu, M., Hou, H., Sha, J. and Garg, S., "SAFA-MMFL: A human cognition inspired unified framework of task-oriented semantic communication for wireless edge devices," Computer Communications, 2026, pp. 108468.

[8] Weng, Y., Sun, J. and Wu, J., "Cognitive semantic communications with codebook-based adaptive correction for image classification," Physical Communication, vol. 72, pp. 102764, 2025.

[9] Xie, L., Xie, W., Lu, H. and Yang, H., "Unmanned aerial vehicle-enabled mobile edge computing for semantic communications," Computer Communications, 2026, pp. 108411.

[10] Gao, P., Zhao, G., Li, Y. and Liu, R., "Towards high-reliability communication in smart grid 2.0: Semantic communications," Electric Power Systems Research, vol. 254, pp. 112639, 2026.

[11] Sagduyu, Y.E., Erpek, T., Yener, A. and Ulukus, S., "Will 6G be semantic communications? Opportunities and challenges from task oriented and secure communications to integrated sensing," IEEE Network, vol. 38, no. 6, pp. 72-80, 2024.

[12] Zhou, H., Deng, Y., Liu, X., Pappas, N. and Nallanathan, A., "Goal-oriented semantic communications for 6G networks," IEEE Internet of Things Magazine, vol. 7, no. 5, pp. 104-110, 2024.

[13] Yang, Y., Shikh-Bahaei, M., Yang, Z., Huang, C., Xu, W. and Zhang, Z., "Joint semantic communication and target sensing for 6G communication system," arXiv preprint arXiv:2401.17108, 2024.

[14] Senbagavalli, M., Debnath, S., Vaneeta, M., Arumugam, T., Beena, B.M. and Prajapat, S., "Adaptive Semantic Integrity and Privacy Shielding in Agentic AI Networks for Trustworthy 6G Cognitive Communication System," IEEE Communications Standards Magazine, 2026.

[15] Shokrnezhad, M., Mazandarani, H., Taleb, T., Song, J. and Li, R., "Semantic revolution from communications to orchestration for 6G: Challenges, enablers, and research directions," IEEE Network, vol. 38, no. 6, pp. 63-71, 2024.

[16] Jain, K., Krishnan, P., Pachiyannan, P., Jaganathan, L., Khan, M.A. and Li, Y., "Toward Smart 5G and 6G: Standardization of AI-Native Network Architectures and Semantic Communication Protocols," IEEE Communications Standards Magazine, 2025.

[17] Shi, G. and Xiao, Y., "An introduction to semantic communication and semantic-aware networking standardization for 6G," GetMobile: Mobile Computing and Communications, vol. 28, no. 3, pp. 14-19, 2024.

[18] Sugitha, G., Vasanthi, R., Solairaj, A. and Kalpana, A.V., "SeCo2: Secure Cognitive Semantic Communication in 6G-IoT Networks Using Key-Policy Attribute-Based Encryption and Elliptic Curve Cryptography," Radioengineering, vol. 34, no. 2, 2025.

[19] Chen, M., Liu, M., Wang, C., Song, X., Zhang, Z., Xie, Y. and Wang, L., "Cross-modal graph semantic communication assisted by generative AI in the metaverse for 6G," Research, vol. 7, p. 0342, 2024.

[20] Liao, H., Fan, J., Ci, H., Gu, J., Zhou, Z., Liao, B., Wang, X. and Mumtaz, S., "Electric Semantic Compression-Based 6G Wireless Sensing and Communication Integrated Resource Allocation," IEEE Internet of Things Journal, vol. 11, no. 24, pp. 39333-39345, 2024.

[21] Wang, Z., Zhang, S., Luo, W. and Xu, S., "Deep reinforcement learning with deep-Q-network based energy management for fuel cell hybrid electric truck," Energy, vol. 306, pp. 132531, 2024.

[22] Hosseini, F. and Eftekhari, M., "PFE-SELF-RAG: Balancing self-RAG evaluation metrics via Pareto efficiency," Journal of Mahani Mathematical Research, pp. 179-208, 2026.

[23] Liu, S., Xiang, Y., Guo, X., Zhao, F., Zhao, A. and Wu, W., "Market Game Optimization Algorithm: A Metaheuristic Inspired by Symmetric Competitive Behavior of Merchants and Consumers," Symmetry, vol. 17, no. 12, pp. 2118, 2025.

[24] Wang, Z., Zhang, S., Luo, W. and Xu, S., "Deep reinforcement learning with deep-Q-network based energy management for fuel cell hybrid electric truck," Energy, vol. 306, pp. 132531, 2024.

[25] Ye, J., Guo, H., Wang, B. and Zhang, X., "Deep deterministic policy gradient algorithm based reinforcement learning controller for single-inductor multiple-output DC–DC converter," IEEE Transactions on Power Electronics, vol. 39, no. 4, pp. 4078-4090, 2024.

[26] Krishnasamy, B., Muthaiah, L., Kamali Pushparaj, J.E. and Pandey, P.S., "DIWGAN optimized with Namib Beetle Optimization Algorithm for intrusion detection in mobile ad hoc networks," IETE Journal of Research, vol. 70, no. 5, pp. 4422-4441, 2024.

[27] Desuky, A.S., Hussain, S., Cifci, M.A., El Bakrawy, L.M., Mzoughi, O. and Kraiem, N., "Parameter optimization based mud ring algorithm for improving the maternal health risk prediction," IEEE Access, vol. 12, pp. 167245-167261, 2024.

[28] Lang, Y. and Gao, Y., "Dream Optimization Algorithm (DOA): A novel metaheuristic optimization algorithm inspired by human dreams and its applications to real-world engineering problems," Computer Methods in Applied Mechanics and Engineering, vol. 436, pp. 117718, 2025.

[29] Medjahed, S.A. and Boukhatem, F., "Narwhal Optimizer: A Novel Nature-Inspired Metaheuristic Algorithm," Int. Arab J. Inf. Technol., vol. 21, no. 3, pp. 418-426, 2024.