Quantum Machine Learning DDPG for Digital Twin Semantic Vehicular Networks

(Invited Paper)

James Adu Ansere*, Sasinda C. Prabhashana*, Octavia A. Dobre*, Trung Q. Duong*,†
* Memorial University, Canada, e-mail:{jaansere, cwelhengodag, odobre, tduong}@mun.ca
† Queen's University Belfast, UK e-mail:{trung.q.duong}@qub.ac.uk

***Abstract*—The increasing complexity of Internet-of-Vehicles (IoV) networks, driven by the need for real-time decision-making and resource management, presents significant challenges, particularly in dynamic, time-varying environments. To address these limitations, we propose a novel framework that integrates the quantum-based deep deterministic policy gradient (Q-DDPG) with digital twin networks (DTN) for distributed semantic optimization in dynamic IoV environments. The framework leverages quantum computing, such as superposition and annealing, to enhance distributed semantic decisions. DTNs provide real-time modeling for efficient task offloading and adaptive resource allocation in decentralized IoV environments under varying conditions and uncertainties. The numerical results validate the robustness of the proposed approach, significantly reduce latency, and improve energy efficiency.**

***Index Terms*—Internet of Vehicles, Deep Deterministic Policy Gradient, Digital Twin Networks, Quantum Machine Learning**

I. Introduction

Internet of Vehicles (IoV) networks have transformed transportation and mobility through the integration of vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-cloud (V2C) communications [1], [2]. These networks underpin advances such as autonomous vehicles, real-time traffic management, and intelligent transportation systems, but their expanding complexity presents challenges in resource optimization and task offloading. Efficient resource management and low-latency communication are essential for the seamless operation of decentralized IoV networks, where unpredictability arises from fluctuating traffic patterns, inconsistent user behavior, and dynamic wireless channel conditions [3]. Traditional optimization methods, often based on static assumptions, falter under such uncertainties, and scaling them to the demands of dynamic IoV networks introduces significant computational bottlenecks [4]. In distributed environments, this creates added pressure, as scalable solutions rely heavily on decentralized decision-making.

Digital twin networks (DTNs) offer a promising solution to these challenges, enabling real-time monitoring and optimization by creating virtual representations of physical IoV systems [5]. In this context, digital twins replicate and forecast network states, paving the way for proactive resource management tailored to semantic relevance. By shifting the focus from raw data throughput to the semantic meaning and relevance of information, semantic optimization refines data transmission and resource allocation strategies [6]. This prioritization of contextually significant information improves decision-making efficiency, equipping IoV networks to adapt quickly and effectively in dynamic scenarios. DT technology is introduced in [7] to enable task offloading in IoV, employing learning algorithms to predict and optimize task assignments based on real-time insights from DT models. Although effective, the framework relies heavily on the accuracy of DT representations and faces scalability challenges in large IoV networks with high volumes of vehicles and data. In [8], DT technology is integrated with intelligent reflective surfaces to optimize vehicle task absorption and resource allocation in 6G enabled IoV networks. DT facilitates real-time simulations, while IRS improves communication quality by improving signal propagation. Despite these advancements, synchronization between physical and digital twins poses significant challenges, and the complexity of efficient IRS deployment strategies limits the framework's applicability in real-time scenarios.

In [9], a framework based on a deep deterministic policy gradient (DDPG) of multiple agents is introduced to handle the task in OV, improving mobile edge computing by improving information exchange and resource coordination to ensure optimal delivery of services in vehicular networks. Scalability challenges emerge in large-scale IoV networks due to increased computational complexity, with convergence issues arising in highly dynamic environments. An adaptive joint resource allocation scheme for the IoV framework is presented, dividing resources into uplink, computing, and downlink sub-strategies [10]. Twin-delayed DDPG is used to dynamically optimize network capacity, reduce delay, and minimize energy consumption. The effectiveness of the algorithm is limited by the complexity of the model in high-dimensional problems, which can impact performance in highly unpredictable scenarios. This work employs an algorithm based on DDPG to optimize multi-user computation offloading and caching strategies in vehicular edge computing systems, aiming to minimize execution delay while improving caching and resource utilization for variability in task sequence [11]. The framework's reliance on accurate request modeling and coupled optimization raises computational demands, reducing its effectiveness in highly variable scenarios.

Driven by the above-mentioned works, we design a quantum DDPG (Q-DDPG) framework for semantic optimization in DT-enabled IoV networks. By combining Q-DDPG with DT technology, the approach minimizes latency, ensures semantic accuracy, and enhances network reliability. Quantum features like superposition and entanglement enable efficient real-time

decision-making, addressing network uncertainties and time-varying conditions. This method also overcomes the limitations of traditional machine learning algorithms, improving scalability and adaptability in complex IoV environments.

The remainder of this paper is organized as follows. Section II details the system model and the formulation of the combinatorial optimization problem. Section III introduces the Q-DDPG framework for designed dynamic IoV environment. Extensive experimental results are presented in Section IV. Finally, concluding remarks are made in Section V.

II. System Model Designed and Formulated Combinatorial Problem

We examine system models for semantic optimization in digital twin IoV and offer a mathematical formulation of the problem.

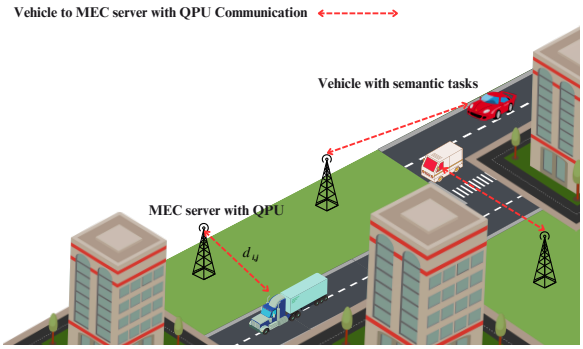

Fig. 1: IoV networks with QPU facilities

A. Semantic-Based Network Model

We consider the dynamic model of the digital twin IoV network, which consists of a set of multiple vehicles and edge servers equipped with quantum processing units (QPU) as illustrated in Fig. 1. Let the IoV network consist of vehicles $\mathcal{V} = \{i_1, i_2, \ldots, i_N\}$, and edge servers $\mathcal{E} = \{j_1, j_2, \ldots, j_M\}$.

B. Semantic-based Vehicular Mobility Model

We consider the position of the vehicle i in time frame τ, denoted as $\mathbf{p}_i(\tau) = (x_i(\tau), y_i(\tau))$, and the position of the edge server j located at $\mathbf{p}_j = (x_j, y_j)$. The channel gain between vehicle i and edge server j at time τ is given by:

$$h_{ij}(\tau) = \frac{g_0}{d_{ij}(\tau)^{\alpha}}, \tag{1}$$

where g_0 is the channel power gain at the reference distance, $d_{ij}(\tau) = \sqrt{(x_i(\tau) - x_j)^2 + (y_i(\tau) - y_j)^2}$ is the distance between vehicle i and edge server$j$, and α is the path loss exponent.

C. Stochastic Task Generation, Task Arrival, and Task Prioritization Models

1) Semantic-based Task Generation Model: Each semantic task S_k generated by vehicle i at timeframe τ is described by:

$$S_k(\tau) = \{D_i(\tau), s_c(\tau), T_{i,p}(\tau), L_D(\tau), \mathbb{P}(A_i(\tau))\}, \tag{2}$$

where $D_i(\tau)$ denotes the task data size (in bits), $s_c(\tau)$ is the semantic task complexity, indicating its contextual information, $T_{i,p}(\tau)$ is the priority of the task, $L_D(\tau)$ is the task deadline for completion, and $\mathbb{P}(A_i(\tau))$ is the stochastic arrival task process.

2) Semantic-based Stochastic Task Arrival Model: Let $\lambda_i(\tau)$ indicate the arrival rate of semantic tasks in vehicle i, and $\mu_i(\tau)$ represent the service rate of semantic tasks. The number of semantic tasks in the queue at time frame τ follows a queuing model:

$$Q_i(\tau + 1) = Q_i(\tau) + \lambda_i(\tau) - \mu_i(\tau), \tag{3}$$

where $Q_i(\tau)$ is the number of semantic tasks in the queue at time frame τ.

The semantic task generation by vehicle i is a stochastic process modeled as a Poisson process with arrival rate $\lambda_i(\tau)$, given as:

$$\mathbb{P}(A_i(\tau) = k) = \frac{(\lambda_i(\tau))^k e^{-\lambda_i(\tau)}}{k!}, \tag{4}$$

where $\mathbb{P}(A_i(\tau) = k)$ is the probability that k semantic tasks arrive at vehicle i in time interval τ.

3) Semantic Task Prioritization Model: Semantic tasks are prioritized by importance and urgency, determining their queue position:

$$S_{k,i} = \varphi \cdot s_c(\tau) + \beta \cdot (L_D(\tau) - \tau), \tag{5}$$

where φ and β are weighting factors, $L_D(\tau)$ is the task deadline for completion, and τ is the current time. Higher values of $S_{k,i}$ increase the priority of the task queue.

D. Digital Twin for Predict the Task Arrival Process

We simulate the DT layer, mirroring the IoV and edge quantum server with real-time updates of state information, historical data, and task arrival processes. The DT layer predicts the arrival rates of tasks, adjusting for uncertainties, to determine optimal offloading strategies for the physical layer. The DT at time step τ is expressed as

$$DT(\tau) = \left(DT_i^{\mathcal{V}}(\tau), DT_j^{\mathcal{E}}(\tau)\right). \tag{6}$$

Considering the uncertain and unpredictable errors between the physical entities and the DT, we model the task arrival rate $\lambda_i(\tau)$ of vehicle i with an uncertain deviation $\delta_\lambda^i(\tau)$ as

$$\lambda_i(\tau) = \hat{\lambda}_i(\tau) + \delta_\lambda^i(\tau), \quad |\delta_\lambda^i(\tau)| \leq \epsilon_\lambda, \tag{7}$$

where $\hat{\lambda}_i(\tau)$ is the estimated arrival rate predicted by the DT, and ϵ_λ is the maximum range of deviations.

E. Semantic-Based Communication Model

Each vehicle i communicates with edge server j to execute semantic tasks. The achievable transmission rate $R_{i,j}(\tau)$ between vehicle$i$ and edge server j at time frame τ is given by:

$$R_{i,j}(\tau) = B_{i,j}(\tau)\sigma_{i,j}\log_2\left(1+\frac{p_{i,j}(\tau)g_{i,j}(\tau)}{I+N_o}\right), \quad (8)$$

where $B_{i,j}(\tau)$ represents the system bandwidth at time frame τ, $p_{i,j}(\tau)$ is the transmission power, $g_{i,j}(\tau)$ is the channel gain, $\sigma_{i,j}$ is the semantic efficiency metric, N_o is the noise power, and I is the interference.

F. Semantic-Based Computation Models

Let $C_i(\tau)$ represent the local computational capacity of vehicle i, and $C_j(\tau)$ be the computational capacity of the edge server j. Let $\alpha_{i,j}(\tau)$ represent the binary offloading decision variable denoting whether semantic task S_k is processed locally or offloaded from vehicle i to edge server j at timeframe τ, defined as:

$$\alpha_{i,j}(\tau) = \begin{cases} 1, & \text{if semantic task } S_k \text{ is offloaded to server } j, \\ 0, & \text{if processed locally at vehicle } i. \end{cases} \quad (9)$$

1) Local Computation Model: The computation delay $L_i^{\text{comp}}(\tau)$ for processing the semantic task S_k in the time frame τ is:

$$L_i^{\text{comp}}(\tau) = \frac{s_c(\tau)}{C_i(\tau)} \times \sigma_i, \quad (10)$$

where σ_i is the semantic accuracy factor for local computation.

2) Offloading Computation Model: To offload the semantic task to the edge server, the computation delay is given by:

$$L_i^{\text{off}}(\tau) = \frac{D_i(\tau)}{R_{i,j}(\tau)} + L_j^{\text{comp}}(\tau), \quad (11)$$

where $L_j^{\text{comp}}(\tau) = \dfrac{s_c(\tau)}{C_j(\tau)} \times \sigma_j$ is the computation delay at the edge server, and σ_j is the semantic accuracy factor for the edge server.

The total delay for processing and offloading a semantic task is:

$$L_{\text{total}}(\tau) = \begin{cases} \frac{D_i(\tau)}{R_{i,j}(\tau)} + L_j^{\text{comp}}(\tau), & \text{if } \alpha_{i,j}(\tau) = 1, \\ L_i^{\text{comp}}(\tau), & \text{if } \alpha_{i,j}(\tau) = 0. \end{cases} \quad (12)$$

G. Semantic-based Energy Consumption Model

The energy consumption to transmit the semantic task S_k from vehicle i to server j is:

$$E_{i,j}^{\text{off}}(\tau) = p_{i,j}(\tau) \cdot \frac{D_i(\tau)}{R_{i,j}(\tau)}. \quad (13)$$

The local processing energy consumption at vehicle i is:

$$E_i^{\text{loc}}(\tau) = \kappa_i \left(C_i(\tau)\right)^2 s_c(\tau), \quad (14)$$

where κ_i is the energy efficiency coefficient of vehicle i. The total energy consumption for semantic task S_k is:

$$E_{\text{total}}(\tau) = \begin{cases} E_{i,j}^{\text{off}}(\tau), & \text{if } \alpha_{i,j}(\tau) = 1, \\ E_i^{\text{loc}}(\tau), & \text{if } \alpha_{i,j}(\tau) = 0. \end{cases} \quad (15)$$

H. Stochastic Combinatorial Offloading Problem Formulation

We aim to minimize the total cost while considering energy consumption, latency constraints, semantic accuracy, and arrival rate uncertainties. The total cost Φ at time step τ is given by:

$$\Phi(\tau) = \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{E}}\left(w_1 L_{\text{total}}^{(i,j)}(\tau) + (1-w_1)E_{\text{total}}^{(i,j)}(\tau)\right), \quad (16)$$

where w_1 represents the weight factor for latency and energy consumption. Mathematically, the combinatorial optimization problem can be formulated as follows:

$$\min_{B_{i,j}(\tau),\alpha_{i,j}(\tau),p_{i,j}(\tau)} \Phi(\tau) \quad (17a)$$

$$\text{s.t.:} \quad L_{\text{total}}^{(i)}(\tau) \le L_{\max}, \quad \forall i \in \mathcal{V}, \forall j \in \mathcal{E} \quad (17b)$$

$$E_{\text{total}}^{(i)}(\tau) \le E_i^{\max}, \quad \forall i \in \mathcal{V} \quad (17c)$$

$$s_c(\tau) \le C_i(\tau) \times \tau, \quad \text{if } \alpha_{i,j}(\tau) = 0, \quad \forall i \in \mathcal{V} \quad (17d)$$

$$\sum_{i\in\mathcal{V}} \alpha_{i,j}(\tau) s_c(\tau) \le C_j(\tau) \times \tau, \quad \forall j \in \mathcal{E} \quad (17e)$$

$$\sigma_{i,j} \ge \sigma_{\min}, \quad \sigma_{i,j}(\tau) \in \{0,1\}, \forall i \in \mathcal{V}, \forall j \in \mathcal{E} \quad (17f)$$

$$p_{i,j}(\tau) \le P_i^{\max}, \quad \forall i \in \mathcal{V}, \forall j \in \mathcal{E} \quad (17g)$$

$$\mu_i(\tau) \ge \lambda_i^{\max}(\tau), \quad \forall i \in \mathcal{V} \quad (17h)$$

$$\lambda_i^{\max}(\tau) = \hat{\lambda}_i(\tau) + \epsilon_\lambda, \quad \forall i \in \mathcal{V} \quad (17i)$$

$$\alpha_{i,j}(\tau) \in \{0,1\}, \quad \forall i \in \mathcal{V}, \forall j \in \mathcal{E} \quad (17j)$$

$$\sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}} B_{i,j}(\tau) \le B_{\max}, \quad \forall i \in \mathcal{E}, \forall j \in \mathcal{E} \quad (17k)$$

where (17b) ensures that the total latency must not exceed the task deadline for each semantic task, (17c) guarantees that the total energy consumption of a vehicle must not exceed its available energy budget, (17d) and (17e) ensure that the computational capacities of the vehicles and edge servers are not exceeded, (17f) verify that semantic tasks meet a minimum semantic accuracy threshold for local and offloaded processing, respectively, (17g) indicates the power transmission limit with the maximum transmit power $P_i^{\max}$ for vehicle i, (17i) and (17j) confirm the service rate is sufficient to handle the maximum possible arrival rate considering uncertainties, (17h)establishes the offloading decision variables, and (17k) ensures bandwidth does not surpass the system's maximum limit.

III. Proposed Quantum-based DDPG Design

The combinatorial optimization problem in (17a) for semantic optimization in digital twin IoV networks is computationally intractable and impractical with increasing vehicle numbers. We design a quantum-inspired DDPG framework for efficient exploration in high-dimensional environments, ideal

for dynamic and stochastic offloading in IoV. This section elaborates on the designed framework.

A. MDP Problem Formulation

We reformulated $P1$ as a Markov decision process (MDP) and modeled it as a tuple $(\mathcal{S}, \mathcal{A}, P, R, \gamma)$ where $\mathcal{S}$ represents the set of states, $\mathcal{A}$ is the set of continuous action space, $P(s'|s,a)$ is the probability of transition, where $P\left(s_{\tau+1} \mid s_\tau, a_\tau\right)$ represents the probability of transitioning from state s_τ to $s_{\tau+1}$ after taking action a_τ at time frame τ, $R(s,a)$ denotes the reward function and $\gamma \in [0,1)$ is the discount factor. The state, action, and reward functions are defined as follows.

1) State space: offers essential information on the current state of the environment for decision making. The state S_i for agent i (vehicle in loV) at each time step τ can be represented as:

$$s_i(\tau) = \left[L_{\text{total}}\left(\tau\right), E_{\text{total}}\left(\tau\right), S_k(\tau), R_{i,j}(\tau)\right]. \tag{18}$$

2) Action space: defines the set of possible decisions that the agent can make at each time step τ. The agent learns and observes the dynamic IoV environments as:

$$a_i\left(\tau\right) = \left[B_{i,j}\left(\tau\right), p_{i,j}\left(\tau\right), \alpha_{i,j}\left(\tau\right)\right]. \tag{19}$$

3) Reward function: provides feedback to the agent towards optimal performance. The reward function seeks to minimize a weighted combination of latency and energy as:

$$r(\tau) = -\left(w_1 L_{\text{total}}\left(\tau\right) + w_2 E_{\text{total}}\left(\tau\right)\right) - \sum_{i,j} \mathcal{U} \tag{20}$$

where $\sum_{i,j} \mathcal{U}$ represents the penalties for constraint violations. The negative sign ensures that the agent minimizes the overall cost.

B. Quantum-DDPG (Q-DDPG) Framework

Q-DDPG integrates quantum computing with DDPG by using a parametric quantum circuit (PQC) for the actor network and higher-order encoding to translate the classical state space into quantum space. The quantum state is encoded using quantum circuits with a feature map, variational ansatz, and quantum state preparation. Classical data $\mathbf{x}$ are encoded in higher-order features $\Phi(\mathbf{x})$ for the parameters of the quantum circuit. We apply $U_{\mathbf{x}} = \exp\left(i \sum_{j<k} \gamma x_j x_k Z_j Z_k\right)$. Herein, $\mathbf{x}$ is the classical input vector with elements x_j and x_k, which correspond to individual features of the data, while γ is a tunable parameter that scales the strength of these interactions. The Pauli-Z operators Z_j and Z_k act on the j-th and k-th qubits, inducing phase flips based on their states. The tensor product $Z_j Z_k$ captures feature correlations, with the exponential ensuring unitary quantum transformations.

C. PQC-based Actor Network

In our model, a hybrid action space in the Q-DDPG framework enables the agent to handle both discrete strategies and continuous control parameters. The actor network, implemented as a Parameterized Quantum Circuit (PQC), determines the optimal action $a = \mu(s \mid \theta^\mu)$, where s is the state and θ^μ the PQC parameters. The PQC encodes complex policies using parameterized single-qubit rotation gates $R_y(\theta)$, which embed trainable parameters, and controlled-Z (CZ) gates, which introduce entanglement to capture feature correlations. This combination ensures the PQC can represent expressive quantum states, optimizing the expected return as evaluated by the critic, defined as:

$$J(\theta^\mu) = \mathbb{E}_{s\sim p^\pi}[Q(s, \mu(s|\theta^\mu) \mid \theta^Q)]. \tag{21}$$

The policy gradient with respect to the actor parameters is computed as:

$$\nabla_{\theta^\mu} J(\theta^\mu) \approx \mathbb{E}_{s\sim p^\pi}\left[\nabla_a Q(s, a \mid \theta^Q) \nabla_{\theta^\mu} \mu(s \mid \theta^\mu)\right], \tag{22}$$

where $\nabla_{\theta^\mu} \mu(s \mid \theta^\mu)$ is computed based on the PQC's trainable parameters and quantum gate structure.

D. Classical based Critic Network

The critic network, implemented as a classical neural network, estimates the action-value function $Q(s, a \mid \theta^Q)$, where θ^Q are the parameters of the critic network. The critic is trained by minimizing the mean-squared error loss between the predicted and target Q-values. The loss function, $L(\theta^Q)$, is given as:

$$L(\theta^Q) = \mathbb{E}\left[(Q(s_\tau, a_\tau \mid \theta^Q) - y_\tau)^2\right], \tag{23}$$

where the target Q-value is defined as:

$$y_\tau = r + \gamma Q'(s_{\tau+1}, \mu'(s_{\tau+1} \mid \theta^{\mu'}) \mid \theta^{Q'}). \tag{24}$$

The gradient descent update for the critic parameters is:

$$\nabla_{\theta^Q} L(\theta^Q) = \mathbb{E}\left[(Q(s_\tau, a_\tau \mid \theta^Q) - y_\tau) \nabla_{\theta^Q} Q(s_\tau, a_\tau \mid \theta^Q)\right]. \tag{25}$$

E. Q-DDPG Target Networks Update

To stabilize training, target networks are used for both the actor and critic. The target network for the actor uses the PQC, while the critic's target network remains classical. These target networks are updated using Polyak averaging:

$$\begin{aligned} \theta^{\mu'} &\leftarrow \rho\theta^\mu + (1-\rho)\theta^{\mu'}, \\ \theta^{Q'} &\leftarrow \rho\theta^Q + (1-\rho)\theta^{Q'}, \end{aligned} \tag{26}$$

where $\rho \in [0,1]$ is the target update rate. The use of target networks mitigates instability caused by rapid parameter changes during training.

For the complexity analysis of the hybrid actor-critical model enhanced with the quantum-inspired algorithm, let n_s denote the state dimension, n_c as the continuous action dimension, n_q as the fixed count of qubits in the quantum network, and r as the repetitions of the variational quantum ansatz. The actor network comprises a quantum component, with a feature map and a variational ansatz scaling as $\mathcal{O}(r \cdot n_q^2)$. The Critic network, a fully connected classical network, has a forward pass complexity of $\mathcal{O}((n_s + n_c))$. The total complexity of the training per step is given by $\mathcal{O}(n_s \cdot n_q + n_c)$.

Algorithm 1: Proposed Q-DDPG Algorithm

Input: $s(\tau)$, $a(\tau)$, ψ
Output: $r(\tau)$

Initialize target network Q' and θ^Q with weights $\theta_Q^{Q'} \leftarrow \theta^Q, \theta_Q^{\mu'} \leftarrow \theta_Q^{\mu}$
Initialize replay buffer
Initialize $\tau = 0$
for $0 \leq \psi \leq \psi_{\max}$; $\psi == \psi_{\max}$; $\psi = \psi + 1$ **do**
  while $\tau > \psi$ **do**
    Select from Hilbert space $\theta_i^l(0; d)$ from $U([-\pi, \pi])$
    Determine action using quantum inspired actor network $a_\tau = \pi(s_\tau \mid \theta^\pi)$
    if *a_τ and $r(\tau)$ are obtained, note the next state $s_{\tau+1}$* **then**
      Record transition $(s_\tau, a_\tau, r_\tau, s_{\tau+1})$ store in replay buffer
    Set $y_\tau = r + \gamma Q'(s_{\tau+1}, \mu'(s_{\tau+1}|\theta^{\mu'})|\theta^{Q'})$.
    Update the critic network and minimize the loss function using equation (25).
    Update quantum-inspired actor policy via equation (24)
  else
    Update the target network weights using equation (26)
  Assign semantic offloading resource $\mathcal{L}_\tau \leftarrow r(\tau)$
  Update τ.

IV. Experimental Results and Discussions

Extensive experiments were carried out to evaluate the performance of Q-DDPG using an IBM qiskit on a 1,000 m lane road. Vehicles are uniformly deployed at RSU locations, allowing users to compute tasks locally or offload them to the quantum server. The path loss model is illustrated in detail in [12], [13]. We set the batch size to 4, actor and critic learning rates as 1×10^{-4} and 2×10^{-4}, respectively. Other simulation parameters are given in table I.

A. Results Discussion

Fig. 2 shows the total reward over training episodes for a Q-DDPG) algorithm with different learning rates. The learning rate of 1×10^{-4} offers optimal performance, ensuring stable policy optimization in dynamic IoV environments. Integrating DT and quantum-inspired features allows Q-DDPG to manage network uncertainty and varying conditions, leading to higher rewards for semantic offloading. Moreover, a learning rate of 2×10^{-4} shows moderate performance but lacks the fine-tuned optimization required for long-term stability, while a learning rate of 4×10^{-4} results in faster but less precise updates. Q-DDPG with DT and a stable learning rate ensures efficient resource management, adaptive task prioritization, and low-cost operation under dynamic IoV conditions.

TABLE I: Simulation Parameters

Parameter	Value
Path loss exponent	2
Number of vehicles	3
Number of edge servers	2
System bandwidth	20 MHz
Noise power density	-70 dBm
Vehicle semantic data size	$1 \times 10^4 - 2 \times 10^4$ bits
Semantic task complexity	$1 \times 10^6 - 2 \times 10^6$ cycles
Weighting factor	0.5
Energy coefficient of processors	1×10^{27}
Maximum latency	50 ms
Maximum transmit power per vehicle	1 W
Minimum semantic accuracy	0.5
Number of episodes	500
Buffer size	1000000

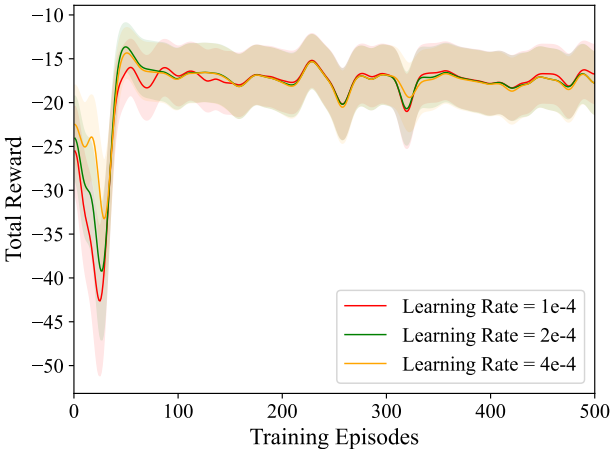

Fig. 2: Convergence performance

Fig. 3 illustrates the average cost over training episodes for the proposed Q-DDPG algorithm. The DT provides an accurate model of the IoV environment that helps Q-DDPG optimize its offloading policies, while quantum-inspired features, such as superposition and entanglement, enable faster learning and adaptability in the face of network uncertainty and time-varying conditions. Hence, Q-DDPG provides a scalable and adaptable solution to manage computational resources, data offloading, and communication in dynamic and uncertain IoV environments.

Fig.4 shows that the integration of Q-DDPG with DT significantly reduces latency in dynamic IoV environments. Quantum-inspired features enable adaptive responses to network uncertainty and time-variability, while the DT framework effectively supports strategy evaluation and improvement.

V. Concluding Remarks

This study presents a Q-DDPG-based framework for semantic optimization in digital twin-enabled IoV networks, focusing on resource allocation and semantic efficiency in dynamic vehicular settings. Utilizing quantum-inspired algorithm and

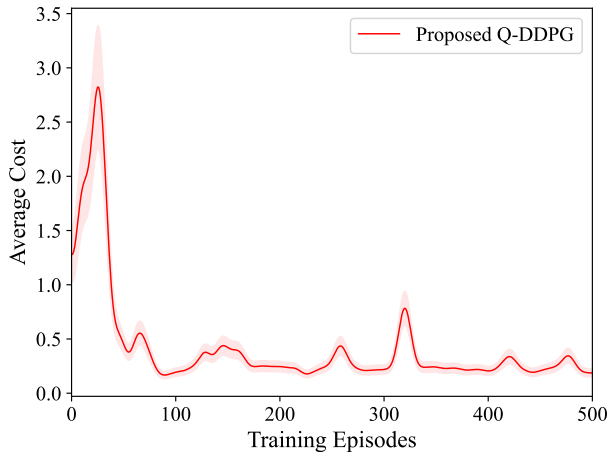

Fig. 3: Average cost performance

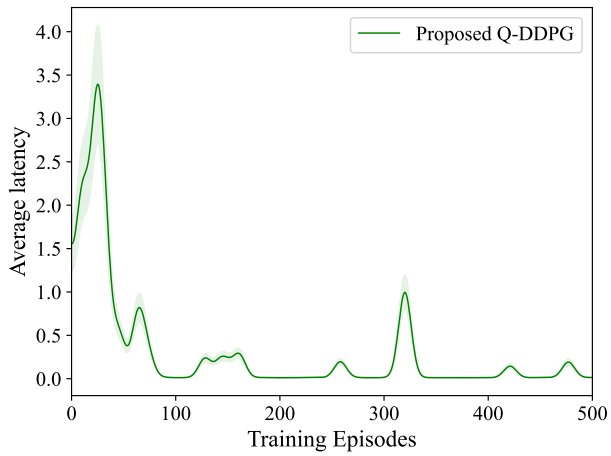

Fig. 4: Average latency performance

digital twin, it enhances semantic accuracy, energy efficiency, and latency, ensuring scalability and robustness. By combining quantum computing with semantic communication, the study paves the way for hybrid optimization, security integration, and real-world applications.

ACKNOWLEDGMENTS

This work was supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109. The work of O. A. Dobre was supported in part by the Canada Research Chair Program CRC-2022-00187.

REFERENCES

[1] K. R. Reddy and A. Muralidhar, "Machine learning-based road safety prediction strategies for internet of vehicles (IoV) enabled vehicles: A systematic literature review," *IEEE Access*, vol. 11, pp. 112 108–112 122, Sep. 2023.

[2] J. A. Ansere, G. Han, and H. Wang, "A novel reliable adaptive beacon time synchronization algorithm for large-scale vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 11 565–11 576, Dec. 2019.

[3] C.-C. Chang, Y.-M. Ooi, and B.-H. Sieh, "IoV-based collision avoidance architecture using machine learning prediction," *IEEE Access*, vol. 9, pp. 115 497–115 505, Aug. 2021.

[4] A. Waheed, M. A. Shah, S. M. Mohsin, A. Khan, C. Maple, S. Aslam, and S. Shamshirband, "A comprehensive review of computing paradigms, enabling computation offloading and task execution in vehicular networks," *IEEE Access*, vol. 10, pp. 3580–3600, Jan. 2022.

[5] Z. Liu, H. Sun, G. Marine, and H. Wu, "6G IoV networks driven by RF digital twin modeling," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 3, pp. 2976–2987, Mar 2024.

[6] Y. Yigit, L. A. Maglaras, W. J. Buchanan, B. Canberk, H. Shin, and T. Q. Duong, "AI-enhanced digital twin framework for cyber-resilient 6G internet of vehicles networks," *IEEE Internet of Things J.*, vol. 11, no. 22, pp. 36 168–36 181, Nov 2024.

[7] J. Zheng, Y. Zhang, T. H. Luan, P. K. Mu, G. Li, M. Dong, and Y. Wu, "Digital twin enabled task offloading for IoVs: A learning-based approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 1, pp. 659–670, Jan 2024.

[8] X. Yuan, J. Chen, N. Zhang, J. Ni, F. R. Yu, and V. C. M. Leung, "Digital twin-driven vehicular task offloading and IRS configuration in the internet of vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24 290–24 304, Sep. 2022.

[9] Y. Guo, D. Ma, H. She, G. Gui, C. Yuen, H. Sari, and F. Adachi, "Deep deterministic policy gradient-based intelligent task offloading for vehicular computing with priority experience playback," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 10 655–10 667, Jul 2024.

[10] J. Zhao, H. Quan, M. Xia, and D. Wang, "Adaptive resource allocation for mobile edge computing in internet of vehicles: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 73, no. 4, pp. 5834–5847, Apr 2024.

[11] L. Liu and Z. Chen, "Joint optimization of multiuser computation offloading and wireless-caching resource allocation with linearly related requests in vehicular edge computing system," *IEEE Internet of Things J.*, vol. 11, no. 1, pp. 1534–1547, Jan 2024.

[12] J. Adu Ansere, D. T. Tran, O. A. Dobre, H. Shin, G. K. Karagiannidis, and T. Q. Duong, "Energy-efficient optimization for mobile edge computing with quantum machine learning," *IEEE Wireless Commun. Lett.*, vol. 13, no. 3, pp. 661–665, Mar. 2024.

[13] D. Wang, B. Song, P. Lin, F. R. Yu, X. Du, and M. Guizani, "Resource management for edge intelligence (EI)-assisted IoV using quantum-inspired reinforcement learning," *IEEE Internet of Things J.*, vol. 9, no. 14, pp. 12 588–12 600, Jul. 2022.