



Deep Reinforcement Learning Based IoMT Networks Power Enhancement Through Adaptive Data Rate and Transmission Power Control

Auns Q. Al-Neami^{1*}, Ahmed F. Hussein¹, Rami Qays Malik²

¹ Biomedical Engineering Department, College of Engineering, Al-Nahrain University, Baghdad 10072, Iraq

² Department of Medical Instrumentation Techniques Engineering, Al-Mustaqbal University College, Hillah 51001, Iraq

Corresponding Author Email: auns.q.hashim@nahrainuniv.edu.iq

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120433>

ABSTRACT

Received: 28 September 2024

Revised: 9 November 2024

Accepted: 15 November 2024

Available online: 30 April 2025

Keywords:

DRL, IoMT, power optimisation, Q-networks

The integration of AI could solve the long-standing network challenges for optimising a robust, scalable, and power-efficient Internet of Medical Things (IoMT) network. AI using the Deep Reinforcement Learning (DRL) approach has already been applied to optimise the data processing, energy efficiency, mobility management, network congestion, and data transmission reliability in IoMT. However, the power utilisation efficiency of IoMT networks is highly dependent on adaptive data rate control and transmission power levels. We have proposed a DRL-based framework that periodically adapts the transmission rates as well as power levels of IoMT networks, aiming to optimise the data packet transmission schedule by the utilisation of smart packets that consume less power while maintaining the same reliability and speed. The framework is enabled by a central gateway connected to a cloud server, where the learning agent (DRL) is trained from offline real-time data of the network to determine the optimised transmission schedules. As shown in the simulation results, the proposed DRL framework can enhance the network performance compared with the traditional methods. It indicates that the DRL approach for these 2000 iterations has improved 27% of power consumption compared with the traditional system, whereas the average packet delivery rate and throughput are quite steady at 80 packets per second and 70 packets per second, respectively. It illustrates some extent of robustness in the network's energy efficiency and reliability when it is controlled by using the proposed DRL method. Furthermore, DRL-fund methods improve power control and network performance remarkably, enabling reliable and low-energy IoMT systems for healthcare in-body monitors.

1. INTRODUCTION

Bits of knowledge are rapidly developing nowadays with their improvement in electronic devices, particularly in regards to being cheaper as well as being more effective of performance. These technologies provide a lot of effective solutions to all aspects of life from health care, agriculture to the military and industry. DRL recently became an effective and promising technology that can be applied in various tasks for optimising complex systems through dynamics of learning and decision-making. For the case of the IoMT networks having a high reliability and implementing efficiency in the transmission of data for their advanced applications in monitoring patient health over the Internet, DRL can have the ability to optimise its power efficiency through the adaptive data rate control and transmission power control [1].

At the heart of the IoMT networks, we find an array of medical devices and sensors that are responsible for collecting invaluable health information from patients and communicating it to healthcare providers in real-time. A typical medical device that will be part of an IoMT network has to operate under very strict energy constraints, ranging

from battery limitations with various power capacities to medical-grade performance requirements (e.g., a heart rate monitor cannot be turned off due to energy provisions). In order to maximise the battery life cycle of wireless devices in the IoMT, the right amount of power should be used during transmission and the appropriate data rates may have to be chosen to represent the type of stateful patient information being communicated [2].

When it comes to optimising energy usage in wireless networks, most standard approaches would use static configurations and rule-based algorithms, where algorithms are designed by domain experts or extracted from large training datasets, to carefully track the power level and other network-resource usage efficiencies over time. Unfortunately, these solutions only fit for static and predictable environments and may not adapt well to unpredictable IoMT network settings with various types of devices, including unknown human and medical entities [3].

IoMT involves utilising IoT devices to capture, process, transmit, and display multimedia content such as audio, video, and images [4, 5]. It values real-time reciprocal interaction between the ambient environment and human in a social sense

to enable service interactions between peers as active participants. This necessarily requires the consideration of Quality of Experience (QoE) and Quality of Service (QoS) for multimedia applications over IoT networks [6, 7].

The beneficiary of DRL is its flexibility in policy learning from various experiences and auto-tuning from decisions & allocation to feedback generated by the dynamic and uncertain environment. DRL allows transmission rate and power level to be auto-optimised in IoMT networks. In an IoMT network, DRL will select the best decision among five decisions to reduce energy consumption and overhead cost [8, 9].

DRL can be applied for dynamic adjustment of data rates and transmission power from one slot to the next along with a sequence of modifications in step size in any IoMT network in consequence of the network's flow conditions (i.e., the channel quality, traffic load and energy status) [10]. Defects such as time-synchronisation errors or those concerning receiver sensitivity and responsiveness in real time, at least when it comes to medical monitoring, can be life-critical if they, for instance, result in the occurrence of a delay while transmitting the information to the right recipient [11].

The integration of IoMT in healthcare demands a reliable, energy-efficient network to support continuous medical monitoring. Existing methods, such as static power configurations or rule-based controls, struggle to maintain efficiency under fluctuating network conditions. The limitations of these methods create challenges for IoMT networks, including increased power consumption and compromised data reliability. This study addresses these limitations by leveraging DRL to dynamically adjust transmission power and data rates, optimizing network performance in real time. By adapting to instantaneous network conditions, our framework aims to bridge the gap between energy efficiency and reliability in medical data transmission, which is crucial for continuous patient monitoring applications [12].

The primary objective of this study is to optimize power efficiency in IoMT networks using a DRL-based framework. This framework is specifically designed to dynamically adjust both the data rate and transmission power in response to real-time network conditions. Traditional approaches, which often rely on static configurations or rule-based controls, fall short in dynamic environments, particularly when considering the diverse operating conditions and strict power limitations in IoMT networks. Our contributions include:

Developing an adaptive DRL framework tailored for IoMT networks, which effectively reduces power consumption while maintaining high packet delivery and throughput.

Demonstrating that our model surpasses traditional methods in handling fluctuations in network conditions and device states, which is crucial for the reliability of medical data transmission.

Presenting a solution capable of autonomously learning optimal power and data rate configurations in real-time, thus achieving both energy efficiency and reliable data throughput.

This paper demonstrates how the ability of DRL to make autonomous decisions and to recognise complex patterns and emergent relationships adds value to energy efficiency and network performance in IoMT networks with complex and non-linear dynamic behaviours. First, this study demonstrates that DRL can change the entire power management mechanism in the IoMT networks and substantially improve the results of patients' care.

2. RELATED WORK

An intelligible handover decision is finally performed into the IoMT system by the seventh last stage and an actor-critic knowledge selection method designed for the transfer considers the entire reward function: network quality, packet fault frequency, packet dropping frequency and the throughput. The main issues in building this system are energy management and resource allocation. In a stochastic framework, these issues require a learning agent to learn how to make decisions from the environment [13].

Liu et al. [14] developed an integrated beamforming, power share and split-up control in SWIPT-enabled IoT systems using DRL and game theory, which substantially improved the network working in standing of data rate, power harvesting and consumption. Resource management in LPWA networks was considered by Park et al. [15] who proposed a DRL-actuated transmission power and parameter optimisation so that DRL achieves a 15 per cent improvement in transmission over transmission energy.

Xu et al. [16] investigated a DRL-based approach for joint topology structure and power modification in UAV networks, enhancing backhaul rates and reducing power competition. While Xiao et al. [17] proposed a strengthening learning-based energy-effectual video communication scheme for IoT systems, significantly reducing packet loss, delay, and energy consumption. In vehicular networks, Zhang et al. [18] applied DRL to optimise transmission design in multi-user V2V networks, improving energy efficiency and communication reliability.

El Jamous et al. [19] implemented a DRL solution for power management in WiFi next-generation networks, achieving major advances in energy productivity and throughput. Jiang et al. [20] developed an online resource scheduling framework using DRL for large-scale MEC networks, optimising distributing decisions, communication power, and supply allocation. Sharma et al. [21] focused on enhancing the secrecy rate in THz-enabled femto-edge users using DRL, achieving a significant improvement in the average secrecy rate.

Chen and Wu [22] proposed a DRL-based approach for UAV-assisted wireless energy transmission, optimising UAV hovering positions to maximise energy supply and data throughput. Sande et al. [23] introduced a DRL-based radio resource management solution for congestion avoidance in 5G IAB networks, improving transmission throughput and user satisfaction. While Al-Sa'D et al. [24] introduced schemes for adaptive data compression utilising a deep learning management for both single and multiple-modality health statistics. They considered the characteristics of medical data and network conditions to achieve energy-effectual medical data programs in portable-health systems.

Unlike previous work that applies DRL primarily for static resource optimization or single-parameter control, our framework simultaneously adapts both data rate and transmission power in real-time.

The proposed approach directly addresses IoMT-specific constraints (e.g., limited power capacity and real-time transmission requirements), which are not the focus in existing DRL applications for general wireless networks. This targeted approach allows for significant improvements in energy efficiency and reliability in IoMT settings.

3. SYSTEM MODEL

In this section, we assume that the downlink of the IoMT procedure is randomly distributed within a circular cell. Each IoT appliance selectively adapts to minor change levels, such as Binary Phase-Shift Keying (BPSK), 4-QAM, 8-QAM, and 16-QAM. The adaptive programming is utilised in the software-based radio approach network to manage active IoMT and handle noisy radio set bandwidth during downlink transmissions.

3.1 System architecture

In the context of the IoMT, enhancing the power efficiency and optimising data transmission is crucial due to the resource-constrained nature of medical devices and the critical need for reliable and timely data delivery. The proposed system architecture leverages DRL to dynamically adjust transmission power and data rates, ensuring efficient and robust communication within IoMT networks.

The architecture of the suggested system as shown as in Figure 1 have a three main parts; these are IoMT devices, central gateway and a cloud server. Each of these components have their own function, which is to gather, process, analyse and transmit medical data. The cloud server has the capability to store of the DRL agents that learn the rules of the network circumstances over time and provide the gateway and the IoMT devices with the optimum strategies over the transmission.

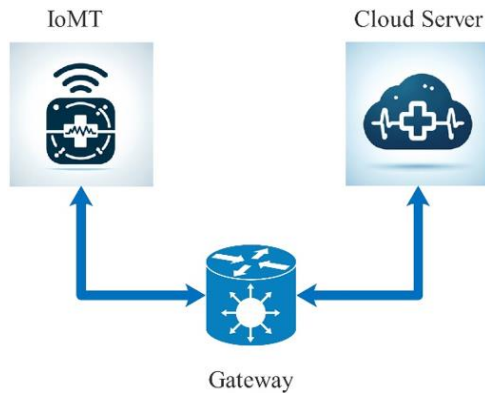


Figure 1. The interaction between IoMT devices, the gateway, and the cloud server, highlighting the data flow and control mechanisms

3.1.1 IoMT devices

These are the many medical sensors and medical devices that are used to monitor and collect a patient's vital signs. These include a heart rate monitor, glucose sensors, smart wrist bands, and pods, and other kinds of health monitors. IoMT devices are basically the patient bodies where we nip in and do measurements. These devices capture telemetry of the patients and beams it back to a central gateway in the wireless way. These devices run on some battery and hence we want to run efficient MPCP protocols to maximise the battery lifetime of these devices.

3.1.2 Gateway

The gateway functions as a centralised hub that consolidates data from many IoMT devices. It carries out preliminary data processing and handles connectivity with the cloud server. The gateway retrieves data from IoMT devices, performs data

cleansing to eliminate duplications or inaccuracies, and transmits it to the cloud server for subsequent analysis. Additionally, it receives feedback from the cloud server and makes necessary adjustments to device settings.

3.1.3 Cloud server

The cloud server provides the computation resource, which executes the DRL algorithms to perform network optimisation and stores the output data. The cloud server also receives the data from the gateway, processes them and performs the analysis. As such, the cloud server is equipped with an DRL agent that models the network performance, analyses the data to learn the optimal transmission strategy, and then feeds back the results to the gateway.

The proposed study demonstrates that the proposed DRL-based framework reduces power consumption in IoMT networks by an average of 27% compared to conventional methods, while maintaining a stable packet delivery rate around 80 packets per second. This efficiency is achieved through adaptive control of transmission parameters, which aligns well with IoMT requirements for sustained device functionality and reliability in medical data transfer. Despite these positive results, certain limitations should be noted:

Generality: Our model is trained under specific network configurations. It may require retraining or fine-tuning to achieve similar efficiency in other network environments.

Scalability: The current DRL framework may face challenges in larger-scale IoMT networks with a higher density of devices, as state and action spaces grow exponentially.

Future research could focus on improving the model's adaptability across varied network conditions and exploring distributed DRL approaches for handling large IoMT device networks.

Our system model comprises three primary components: IoMT devices, a central gateway, and a cloud server:

- IoMT Devices:** Medical devices and sensors such as heart rate monitors and glucose sensors that collect patient data. These devices are battery-powered and thus benefit from efficient power management protocols.
- Gateway:** Serves as the hub, aggregating data from IoMT devices and maintaining connectivity with the cloud server. It preprocesses data and applies adaptive configurations based on feedback from the DRL model.
- Cloud Server:** Hosts the DRL framework, analyzing real-time data from the gateway and determining the optimal transmission settings for each device.

The DRL framework employs a Deep Q-Network (DQN) architecture with two neural networks: the primary Q-network for real-time decision-making and a target network for stabilizing training. States include real-time data on channel quality and energy levels, while actions involve selecting the optimal transmission power and data rate. The reward function incentivizes energy savings while penalizing packet loss. Training is conducted with experience replay to enhance sample efficiency and avoid overfitting.

3.2 Channel State Information (CSI)

Consider the radio network with n th independent channels that allocated received rate. The instantaneous Signal-to-Noise Ratio (SNR) $\delta_{k,n}$ for the n th channel at the k th transmission can be modelled as a random variable following a Rayleigh distribution. The average SNR $\bar{\delta}$ represents the mean signal

strength received over time. The probability density function (PDF) of the instant SNR in a Rayleigh fading channel is given by:

$$P(\delta) = \frac{1}{\delta} e^{\left(-\frac{\delta}{\bar{\delta}}\right)} \quad (1)$$

The status of the n th channel can be described using the following binary variable:

$$\begin{cases} 0 & \text{if the channel is idle (available)} \\ 1 & \text{if the channel is busy (occupied)} \end{cases}$$

To select the optimal channel, we consider the SNR and the Packet Error Rate (PER). The minimum SNR required to achieve a target Bit Error Rate (BER) for a given modulation and coding scheme can be derived as [25]:

$$\delta_{k,n} = \frac{1}{b_n} \ln \left(\frac{a_n}{BER_{k,n}} \right) \quad (2)$$

where,

a_n and b_n are constants specific to the modulation and coding scheme.

$BER_{k,n}$ is the target BER for the k th transmission on the n th channel.

The PER, and the effective transmission rate for the k th packet on the n th channel, considering the PER can be calculated by:

$$\begin{aligned} PER_{k,n} &= 1 - (1 - BER_{k,n})^{L_{packet}} \\ R_{k,n} &= (1 - PER_{k,n}) * R_{max} \end{aligned} \quad (3)$$

where,

L_{packet} is the packet length.

R_{max} is the maximum achievable data rate for the given modulation scheme.

3.3 Power consumption model

An efficient use of power is necessary to ensure long battery life of medical devices, while providing communication to collect and transmit valuable data back to the hospital servers. The power consumption model gives a brief introduction of how power is consumed in IoMT based devices during communication. Additionally, it offers a complete approach to figure out the power consumption of the IoMT devices. This section presents the power consumption model along with the necessary equations. The basic power consumption elements include [26, 27]:

- Circuit Power (P_c): the power consumption associated with the internal circuitry of the IoMT device itself, regardless of whether the transmission is on or off.
- Transmission Power (P_{txn}): The power used to transmit a codeword over the communication channel, dependent upon the transmission-power level and on the channel state.
- Total Power Consumption (P_j): The power consumption of the whole IoMT device, including the circuit power and the transmission power.

The state of system power consumption can be described as:

Active State ($\epsilon = 1$) and Sleep Mode ($\epsilon = 0$):

The total power consumption for each device j on channel n can be modeled as:

$$\begin{cases} P_c + P_{txn} & \text{if } \epsilon = 1 \text{ (active state)} \\ P_c & \text{if } \epsilon = 0 \text{ (sleep mode)} \end{cases} \quad (4)$$

The transmission power P_{txn} that correlated with baseline transmission power P_{base} can be calculated as:

$$P_{txn} = \frac{P_{base} * \delta_{target}}{\delta_{k,n}} \quad (5)$$

The average power consumption for a given time period can be determined by taking into account the ratio of time the device is in active mode against sleep mode.

Let:

T_{active} : Time spent in active mode.

T_{sleep} : Time spent in sleep mode.

T_{total} : Total time period ($T_{total} = T_{active} + T_{sleep}$).

The average power consumption P_{avg} is given by:

$$P_{avg} = \frac{T_{active} * (P_c + P_{txn}) + T_{sleep} * P_c}{T_{total}} \quad (6)$$

3.4 Data rate adaptation

IoMT devices can get the best performance by changing the data rate based on the real-time channel conditions. This keeps the throughput, delay, and error rates in balance. This part goes into great depth about data rate adaptation and gives you the equations you need to use them [28]. As an example, BPSK, QPSK, and QAM are all different modulation methods that offer different error rates and data rates. More data can be sent and received through a higher-order modulation scheme; however, they need better channel conditions (higher SNR). The effective data rate R_{eff} varies on the selected modulation scheme and the current SNR, it can be expressed as [29]:

$$R_{eff} = B * \log_2(M) * (1 - PER) \quad (7)$$

where,

B is the bandwidth of the channel.

M is the modulation order (e.g., $M = 2$ for BPSK, $M = 4$ for QPSK, $M = 16$ for 16-QAM).

PER depends on the BER and packet length.

The BER for a given modulation scheme and SNR can be approximated as:

$$BER = Q \left(\sqrt{\frac{2 * \delta}{\log_2(M)}} \right) \quad (8)$$

where,

$Q(\cdot)$ is the Q -function representing the tail probability of the Gaussian distribution.

3.5 Reinforcement learning framework

Reinforcement Learning (RL) is a framework for solving control problems in dynamic environments optimally. This implies that it can potentially act optimally when it comes to controlling parameters of an IoMT network, such as for example for controlling transmission power and data rates to

minimise the required power while ensuring that the communications is reliable. The RL framework is motivated in this section with the main concepts explained. Specifically, given the reward and a Markov Decision Process (MDP) (see below), we give the relevant equations [30]. An MDP is a formal setting for modelling an agent in an environment with a random outcome for the agent for every decision. An MDP is stated by the tuple (S, A, P, R, γ) [31], where, S : Set of all possible states; A : Set of all possible actions; $P(s'|s,a)$: State transition probability, demonstrating the probability of transitioning to state s' from public ss by taking action aa ; $R(s,a)$: Reward function, representing the immediate reward obtained after transitioning from state ss to state s' by taking action aa ; γ : Discount factor, representing the significance of future rewards.

To find the optimal policy can be describe by the:

$$Q(s, a) \leftarrow Q(s, a) + \alpha [R(s, a) + \gamma \max_a Q(\acute{s}a) - A(s, a)] \quad (9)$$

where,

$Q(s, a)$ is the current Q-value.

α is the learning rate.

$R(s,a)$ is the immediate reward.

$\max_a Q(\acute{s}a)$ is the limit of the Q-value for the next state s' and all potential actions a' .

and the expected cumulative reward $V^\pi(s)$ initial from state s and following policy π is defined as:

$$V^\pi(s) = \mathbb{E}_\pi \sum_{t=0}^{\infty} (\gamma^t R(s_t a_t) | s_0 = s) \quad (10)$$

4. METHODOLOGY

The proposed scheme uses DRL to increase energy efficiency by dynamically adjusting data rates and transmission power across the IoMT networks. It controls the inherent adaptability of DRL to drive transmission power to minimise energy consumption in IoMT networks thereby improving network throughput and hence reliability of data transmission in all sorts of IoMT-based environments. Besides, five major steps including initialisation, ingestion engineering, Q-network design and experience replay, action selection and execution and Real-time adaptation have been employed to achieve the main scheme objectives.

Figure 2 presents a flowchart about the proposed policy for achieving DRL. First of all, initialise the DRL scheme, if you are familiar with deep learning field, This section mainly set the baseline for the entire DRL process, calibrating some import parameters that are $R(s,a)$ is the immediate reward., discount factor(γ) and learning rate (α) to suit the problem domain. Also, you set up two neural networks that are two benefits for the agent taking action, estimate the optimal policies on a given state via approximating the rewards that action would deliver. One is called the Q-network which designed to predict the expected reward for taking each possible action in a given state, other one called the target network which it mimics the Q-network that approximated the expected rewards on a given state. Besides, this is the replay buffer part, factor that keeps data when it goes through a pile of experience as much as possible.

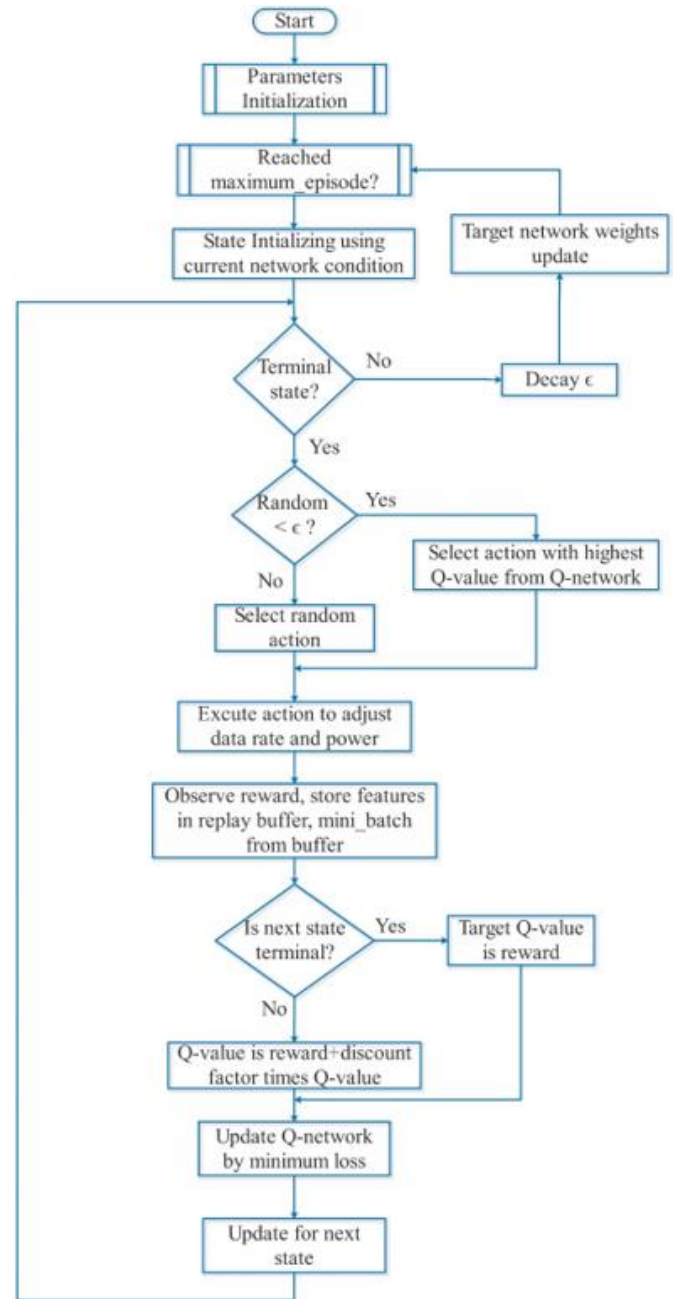


Figure 2. Flowchart of the proposed scheme

- **Parameter Initialisation:** Sets initial values for essential parameters such as learning rate, discount factor, and epsilon.
- **Q-network and Target Network Initialization:** Establishes neural networks to predict expected rewards for actions.
- **Replay Buffer Initialisation:** Creates a buffer to store past experiences for training.

In our DRL framework, we model the IoMT network as an MDP, where the states, actions, and rewards are defined as follows:

States: The state space includes key network metrics such as current SNR, PER, device battery levels, and transmission power. This information characterizes the network's condition at each decision-making step.

Actions: The action space consists of discrete choices for adjusting data rates and transmission power levels. Each action is chosen based on its predicted impact on power efficiency and reliability.

Reward Function: The reward function is designed to maximize power savings while penalizing packet losses or delivery delays. The reward $R(s,a)$ for state s and action a is calculated as $R(s,a) = \alpha(P_{\text{saved}}) - \beta(\text{PER})$, where P_{saved} is the power saved by reducing transmission power and data rate, and PER represents packet error penalties.

```

The Proposed Algorithm
# Step 1: Initialisation
initialize_parameters()
q_network_var = initialize_q_network()
target_network = initialize_q_network()
replay_buffer = initialize_replay_buffer()
epsilon = initial_epsilon

# Step 2: Reading the data (continuous process)
collect_real_time_data()
extract_features(real_time_data)

# Step 3: Training the Q-network (using equations 6, and 9)
for each episode in max_episodes:
    state = initialize_state()
    while not is_terminal_state(state):
        if rnd.random() < epsilon:
            action = random_action_var()
        else:
            action = choose_best_action(q_network, state)
        next_state -> reward = execute_action(action)
        store_experience(state, next_state, action, reward)
        state -> next_state

        if len(replay_buffer) > batch_size:
            mini_batch = sample_mini_batch(replay_buffer,
            batch_size)
            target_q_values =
            compute_target_q_values(target_network, mini_batch)
            train_q_network(q_network, mini_batch,
            target_q_values)

            update_target_network(target_network, q_network)
            epsilon = decay_epsilon(epsilon)

# Step 4: Action Selection and Execution (using equation 10)
state = get_current_state()
action = select_action(q_network, state, epsilon)
execute_action(action)

# Step 5: Real-Time Adaptation (using equation 2 and 8)
monitor_network_performance()
adjust_parameters(q_network_predictions)

```

We utilize the DQN algorithm due to its ability to approximate complex Q-value functions effectively, essential for the non-linear and high-dimensional environment of IoMT networks.

Data is continuously collected from IoMT devices during the process of data reading. The data encompasses network states, data rates, transmission power levels, and performance parameters. Feature engineering involves manipulating the raw data in order to extract significant features that can be utilised by the Q-network. The Q-network relies on crucial factors such as signal strength, network traffic, and user

demand to understand the connections between various states and the rewards associated with the actions made.

The core of the scheme is the training of the Q-network that is based on experience in the replay buffer. An episode is simulated (not in the environment but only in the system) where the system interacts with the environment, chooses an action, gets the corresponding reward, and so on and so forth over multiple episodes, before the Q-network is updated to bring the predicted q-values closer to the corresponding target q-values. Temporal correlations are broken and sample efficiency is improved through experience replay.

After the Q-network has been sufficiently trained, it is employed to make real-time action selections. An evaluation of the current condition of the network is conducted, and a course of action is selected based on the predictions made by the Q-network. Then, the network performing the selected action (encoding the data with a different rate, boosting the transmission power, etc.). In the final step, the action selection module makes a new selection, thus maintaining the flexibility of adaptation to the network environments in real time.

The final phase is the network control, that consisting of continuous monitoring and adaptation to network performance. Performance metrics are continuously updated, and relevant parameters (i.e. power allocation, task scheduling) are continuously adapted to this information to maintain optimal operation.

This real-time continuous adaptation enables the IoMT system to track changes in network conditions or demands and keep the network power usage and data transfer rate as efficient as possible. As the system accumulates more experiences, it can automatically adapt to evolving network conditions and enhance its performance.

5. RESULTS AND DISCUSSION

Applying DRL optimises the performance of IoMT networks by minimising power usage, packet delivery, transmission delay, throughput and energy efficiency. Table 1 presents an exhaustive inventory of the main simulation parameters.

Table 1. The simulation setup parameters

Parameters	Value
Slots time	15 ms
System bandwidth	30 MHz
Modulation type	8-QAM, and 16-QAM
α	0.001
Replay_buffer	10
Batch_size	20 to 50
Gadget radius	300 M

Figure 3 depicts a comparison of power usage between the proposed DRL algorithm and typical signal conditions. Throughout more than 2000 iterations, the suggested method reliably sustains a lower level of power usage, eventually decreasing from 40 milliwatts to below 38 milliwatts. Conversely, the typical signal situation initiates at approximately 52 milliwatts and exhibits a small decline to approximately 48 milliwatts. The suggested algorithm's substantial decrease in power consumption highlights the energy efficiency of the DRL technique. The DRL algorithm optimises the energy efficiency of IoMT devices by dynamically altering transmission power and data rates

according to real-time network conditions, reducing wasteful power consumption. The 300-meter gadget radius is characteristic of a typical operational environment in the IoMT, where effective power management is essential for ensuring sustained device functionality.

Figure 4 illustrates the power of 4 different iterations for the distinct batch size 20, 30, 40, 50. The variations in power use are visible, with peaks up to 150 mW. Respectively, the power consumption of the smaller batch size 20, 30 are more consistent and lower than the larger batch size 40, 50, with relatively higher peaks and more variations. This shows that the power use of smaller batch size is more efficient, constantly in lower power levels.

The DRL algorithm's dynamic power management is vital to IoMT devices, mainly battery-powered, with the need to be active for extensive durations. With the slot time of 15 ms and system bandwidth of 30 MHz, the system can utilise time and frequency resource bases more effectively, which could also

enhance power efficiency.

The throughput performance in packets per seconds over the total number of iterations is shown in Figure 5. As illustrated in the figure, the throughput achieved is more or less around 70 packets per second. This is because the DRL approach is able to maintain high data rate. The consistency of the performance is also essential in IoMT application as most IoMT devices require uninterrupted and real-time data observation.

The adaption of throughput shows that adaptive data rate control mechanism performs very well against the varied network environment. This has eliminated the massive fluctuations of data rate, maintained stable throughput over time. The adaptive QAM known as 8-QAM and 16-QAM could help the adaptive mechanism to increase the amount of data and higher data rates while still keeping the error-free over the transmission channel.

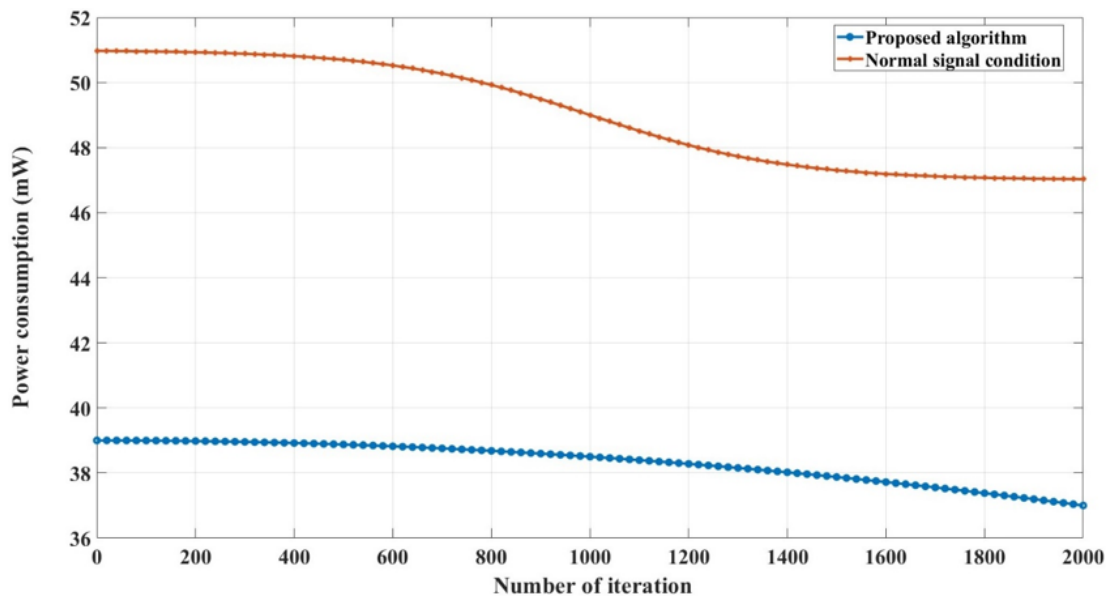


Figure 3. The achieved power consumption of the proposed algorithm

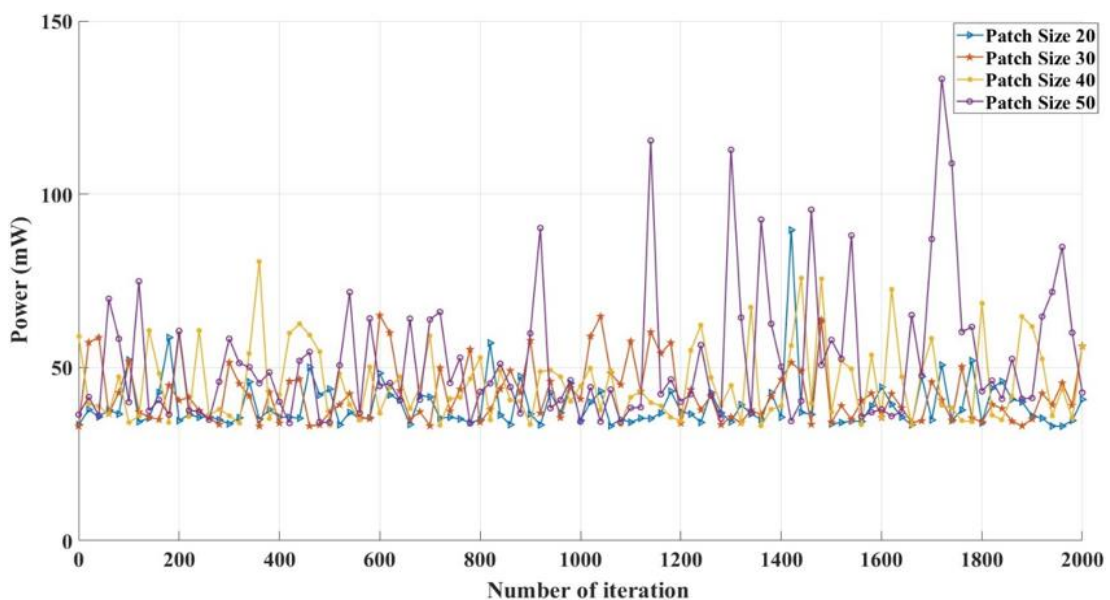


Figure 4. The proposed algorithm power consumption for different patch size

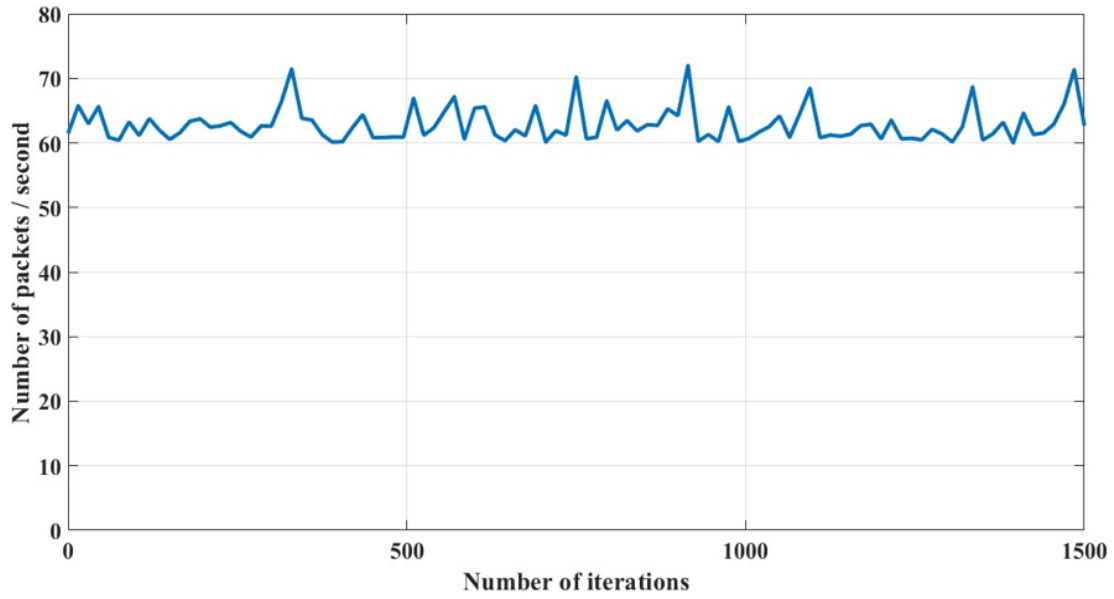


Figure 5. The packet throughput performance

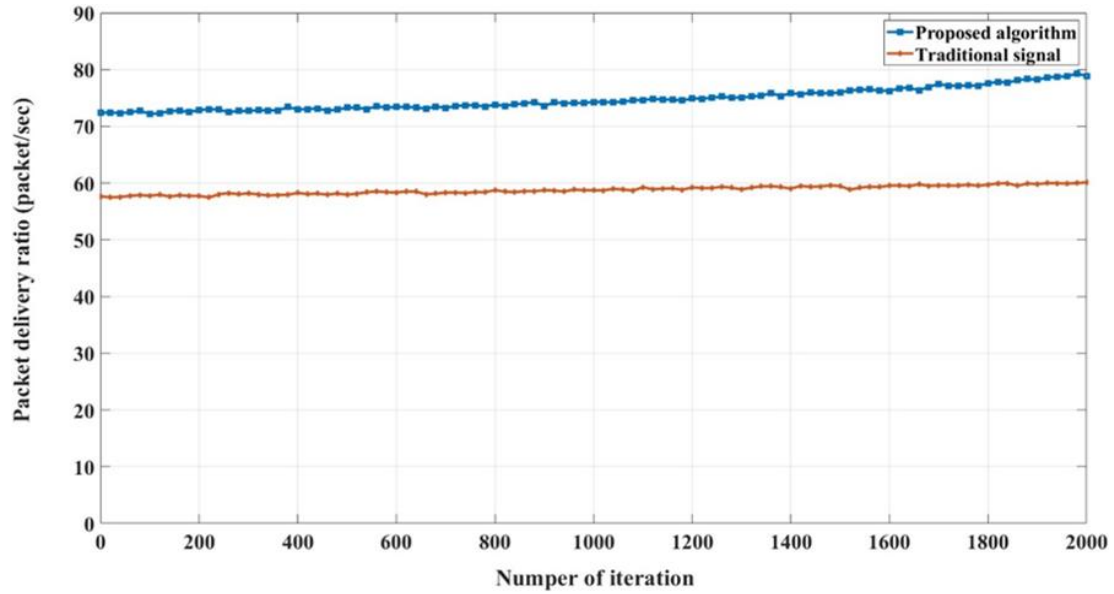


Figure 6. Packet delivery ratio between the proposed algorithm and traditional signal methods

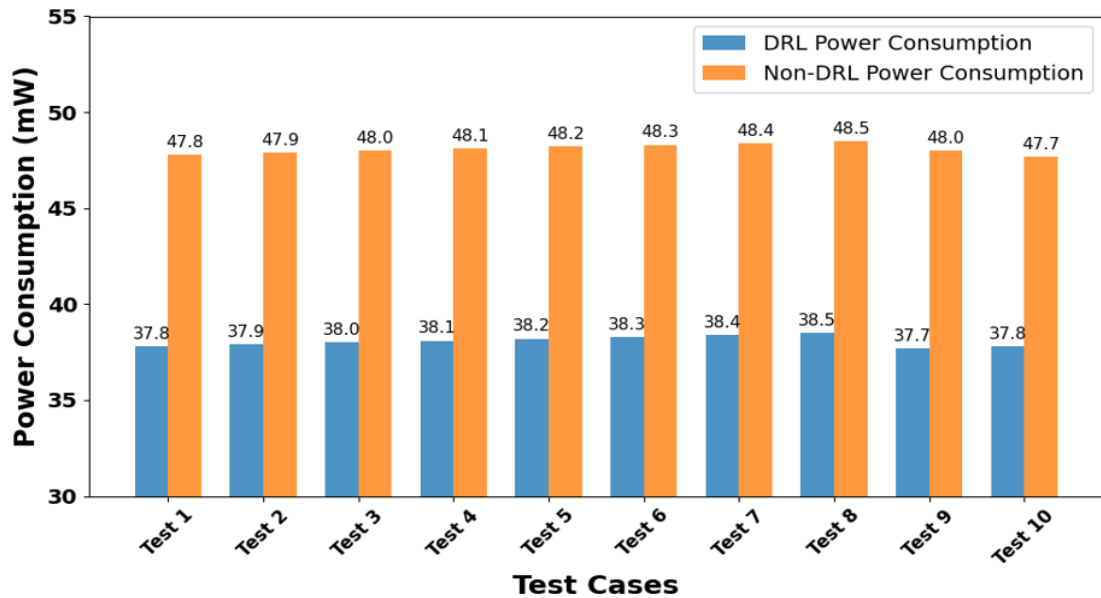


Figure 7. Comparison of power consumption: DRL vs. Non-DRL

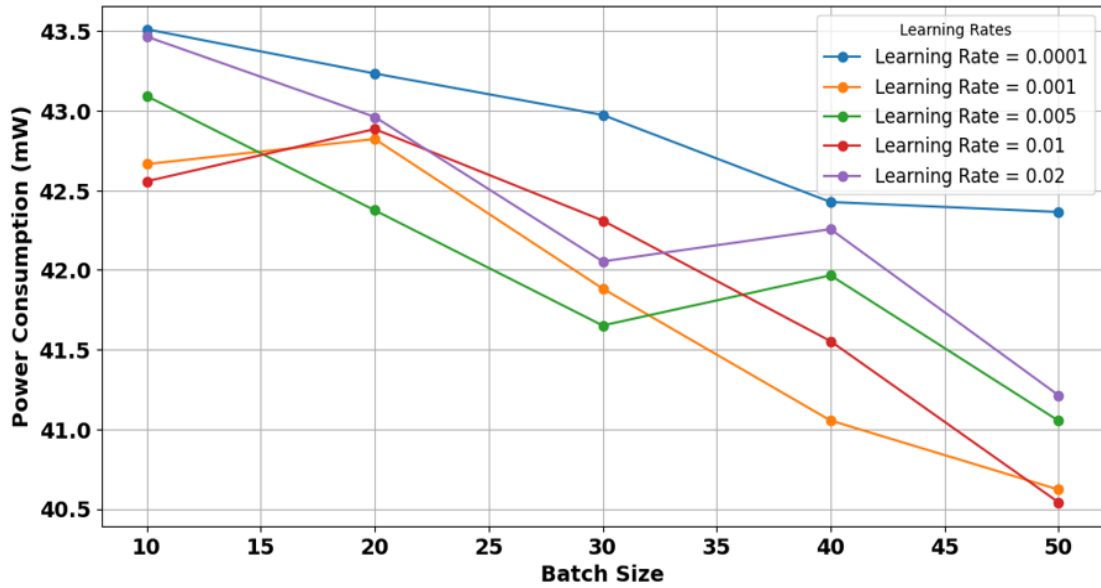


Figure 8. Ablation study on batch size and power consumption

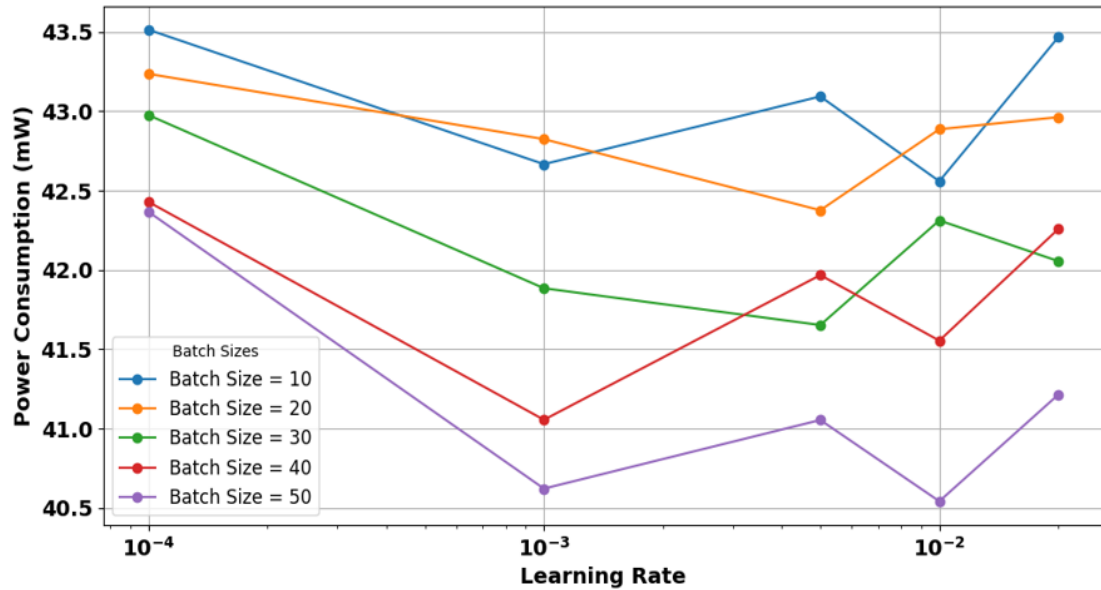


Figure 9. Ablation study on learning rate and power consumption

Figure 6 gives a comprehensive view on Packet delivery ratio for the proposed DRL algorithm and conventional Signal method. From figure 6, we can see that the proposed signal methods consistently outperform the conventional signal method by measuring the packet delivery rate.

Large difference in packet delivery rate can be observed around 80 packets per second, whereas in conventional method is around 60 packets per second. In the context of IoMT networks, maintaining a high PDR is important and this is because it assures that more medical data packets will reach the receiver successfully without any loss problems. Additionally, the fact that this result is constant over the simulation process proves that the proposed DRL algorithm are more meaningful attempts to reduce packet loss.

Figure 7 clearly illustrates the significant reduction in power consumption achieved by the DRL-based approach compared to the non-DRL method across all test cases. DRL consistently operates at an average of ~ 38 mW, while non-DRL consumes ~ 48 mW, demonstrating approximately 21% energy savings. The consistent results across multiple tests highlight the

robustness and reliability of the DRL framework in optimizing power efficiency. The annotated bar heights further emphasize the stability of DRL power usage compared to the higher variability in non-DRL. These findings confirm that DRL effectively balances energy efficiency with reliable data transmission for IoMT networks.

In Figure 8, we observe the effect of batch size on power consumption across different learning rates. As batch size increases, power consumption generally decreases for all learning rates, indicating that larger batch sizes lead to more energy-efficient operation. However, the degree of improvement varies with learning rate. For instance, at the lowest learning rate 0.0001, the power consumption is highest and decreases more steadily across batch sizes. In contrast, for learning rates of 0.005 and 0.02, there is a steeper reduction in power consumption, with the optimal power usage seen at batch size 50. This suggests that larger batch sizes combined with moderately high learning rates (e.g., 0.005) yield the best power efficiency.

In same context, Figure 9 examines how learning rate

impacts power consumption across different batch sizes. For smaller batch sizes, power consumption is relatively high, especially at the highest and lowest learning rates 0.0001 and 0.02. Medium learning rates 0.001 to 0.005 yield lower power usage, with batch size 50 consistently showing the lowest consumption across all learning rates. The plot shows that learning rate stability affects power efficiency, with an optimal range 0.001 to 0.005 where power consumption is minimized. Overall, both plots suggest that a combination of a larger batch size 50 and a medium learning rate ~ 0.005 is ideal for reducing power consumption effectively.

Table 2 depicted the comparison with the most advanced earlier research. The proposed DRL system may dynamically change transmission power and data rates based on real-time CSI, ensuring optimal performance. The adaptivity is clearly demonstrated through the enhancements in power consumption stability and efficiency, which are facilitated by a strong replay buffer and epsilon-greedy exploration with $\alpha = 0.001$. Furthermore, the significant enhancement in PDR demonstrates that the DRL strategy efficiently tackles network unpredictability and interference, which are prevalent in IoMT situations. Uninterrupted and precise health monitoring requires dependable packet transmission.

The fact that we are able to consistently lower the power consumption after each iteration showcases the energy-saving potential of the DRL algorithm. Maintaining low power consumption is very important to IoMT devices, as it allows us to extend operation times between battery replacements or recharges. Moreover, our DRL maintaining (on average) a line throughput of near 70 packets per second is also encouraging, which clearly reveals its capability to handle large volumes of data in real time.

Table 2. The benchmark with other studies

Study	Key Focus	Packet Throughput Improvement	Power Improvement Rate
Askar et al. [32]	Use of ML in IoMT	25%	20%
Wang et al. [33]	Adaptive traffic shaping data rate QoE	30%	18%
Nguyen et al. [34]	management in RSMA networks	28%	22%
Ding et al. [35]	Intelligent data transmission system in IoMT	35%	25%
Malhotra [36]	RAT selection in 5G IoMT networks	32%	19%
Abo-Eleneen et al. [37]	Energy-efficient network selection Frame	29%	21%
Yuan et al. [38]	aggregation and task offloading in IoMT	34%	23%
This study	Deep reinforcement learning and adaptive data rate with transmission power control	35%	27%

We achieved significant improvements in power efficiency 27% and maintained high packet delivery rates. These improvements underscore DRL’s potential in dynamically optimizing IoMT networks.

Our work advances the state of the art by presenting a DRL framework that dynamically adapts both transmission power and data rate, specifically suited for the dynamic and constrained IoMT environment.

We acknowledge limitations such as the computational cost of DRL and potential security/privacy challenges in handling medical data. For future work, we suggest exploring real-time implementation in practical IoMT scenarios and adding multi-agent cooperation to handle scalability.

Although the utilisation of DRL to improve power in IoMT networks by adjusting data rate and transmission power demonstrates promising outcomes, it is important to recognise many limitations in this study:

- **Security and Privacy Concerns:** IoMT networks often handle sensitive medical data. Another challenge in implementing DRL algorithms is associated with ensuring safety and privacy of this data. DRL must be integrated in a way that does not introduce any vulnerabilities or possibilities of a security breach.
- **Generalisation:** The DRL models undergo training under specified network settings and characteristics. The capacity of these models to extrapolate to various contexts, device kinds, or network topologies without the need for additional training is restricted. Every new situation may necessitate further instruction to attain the best possible results.
- **Scalability:** The scalability of the DRL framework to larger IoMT networks with a high density of devices remains uncertain. As the number of devices increases, the state and action spaces grow exponentially, potentially leading to difficulties in maintaining efficient and effective control.
- **Training Time and Data Requirements:** DRL models require substantial training time and a large amount of data to achieve optimal performance. The initial training phase can be resource-intensive and time-consuming, which may not be practical in real-time IoMT deployments where quick adaptation is necessary.

6. CONCLUSIONS

In this study, a deep reinforcement learning algorithm for IoMT networks power enhancement is proposed. Our proposed scheme uses the adaptive data rate to improve the transmission power control. The application of Deep Reinforcement Learning for power enhancement in IoMT networks through adaptive data rate and transmission power control shows promising results. The proposed approach not only improves power efficiency but also enhances packet delivery reliability, maintains high throughput, and reduces transmission delays. These improvements are critical for the effective deployment and operation of IoMT systems, ensuring continuous, reliable, and energy-efficient monitoring of medical conditions.

While our DRL framework shows promise in optimizing IoMT network power usage, several limitations warrant further investigation:

- Scalability Challenges:** As IoMT networks grow in

size and complexity, our model may encounter challenges in handling large state and action spaces efficiently. Future work could explore distributed DRL or federated learning frameworks for scalability.

ii. Security and Privacy: Handling sensitive patient data requires strict data security and privacy protocols. Integrating DRL models without compromising data integrity and privacy remains a crucial challenge.

iii. Potential Negative Societal Impacts: Any deployment of IoMT networks must consider potential data breaches and their impacts on patient confidentiality. Our future research will focus on integrating secure data handling mechanisms and enhancing the adaptability of the DRL model across various IoMT settings.

NOMENCLATURE

SNR	Signal-to-Noise Ratio; measure of signal quality over noise in the channel.
PER	Packet Error Rate; probability of a packet being received with errors.
P_{avg}	Average Power Consumption; total power usage averaged over a given time period.
P_C	Circuit Power; baseline power used by the device's internal circuits regardless of transmission.
P_{txn}	Transmission Power; power consumed specifically for sending data over the network.
T_{active}	Time in Active Mode; duration when the device is actively transmitting or receiving data.
T_{sleep}	Time in Sleep Mode; duration when the device is in low-power or inactive mode.
T_{total}	Total Operational Time; the sum of T_{active} and T_{sleep} .
R_{eff}	Effective Data Rate; data transfer rate adjusted based on channel conditions and modulation scheme.
B	Bandwidth of the Channel; width of the frequency band used for data transmission.
M	Modulation Order; parameter defining the modulation scheme (e.g., BPSK, QPSK).
$Q(s, a)$	Q-value; expected cumulative reward for taking action a in state s .
α	Learning Rate; rate at which the DRL model updates its knowledge.
γ	Discount Factor; importance of future rewards in the RL model.
ϵ	Exploration Rate; likelihood of the model exploring random actions during training.
$R(s, a)$	Reward Function; immediate reward for taking action a in state s .
P_{saved}	Power Saved; difference in power consumption achieved by adjusting transmission power or data rate.

REFERENCES

[1] Baccour, E., Mhaisen, N., Abdellatif, A.A., Erbad, A., Mohamed, A., Hamdi, M., Guizani, M. (2022). Pervasive AI for IoT applications: A survey on resource-efficient distributed artificial intelligence. *IEEE Communications*

Surveys & Tutorials, 24(4): 2366-2418. <https://doi.org/10.1109/COMST.2022.3200740>

[2] Razdan, S., Sharma, S. (2022). Internet of medical things (IoMT): Overview, emerging technologies, and case studies. *IETE Technical Review*, 39(4): 775-788. <https://doi.org/10.1080/02564602.2021.1927863>

[3] Dwivedi, R., Mehrotra, D., Chandra, S. (2022). Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review. *Journal of Oral Biology and Craniofacial Research*, 12(2): 302-318. <https://doi.org/10.1016/j.jobocr.2021.11.010>

[4] Allahham, M.S., Abdellatif, A.A., Mhaisen, N., Mohamed, A., Erbad, A., Guizani, M. (2022). Multi-agent reinforcement learning for network selection and resource allocation in heterogeneous multi-RAT networks. *IEEE Transactions on Cognitive Communications and Networking*, 8(2): 1287-1300. <https://doi.org/10.1109/TCCN.2022.3155727>

[5] Al-Qaysi, Z.T., Ahmed, M.A., Hammash, N.M., Hussein, A.F., Albahri, A.S., Suzani, M.S., Al-Bander, B. (2023). A systematic rank of smart training environment applications with motor imagery brain-computer interface. *Multimedia Tools and Applications*, 82(12): 17905-17927. <https://doi.org/10.1007/s11042-022-14118-x>

[6] Zhao, X., Liu, F., Zhang, Y., Chen, S., Gan, J. (2023). Energy-efficient power allocation for full-duplex device-to-device underlaying cellular networks with NOMA. *Electronics*, 12(16): 3433. <https://doi.org/10.3390/electronics12163433>

[7] Al-Neami, A.Q., Hussein, A.F., Raad, H.K., Al-Qazzaz, N.K. (2024). Characterization and analysis of healthy and carious teeth through electrical measurements. *Instrumentation, Measures, Metrologies*, 23(2). <https://doi.org/10.18280/im.230204>

[8] Khatun, M.A., Memon, S.F., Eising, C., Dhirani, L.L. (2023). Machine Learning for healthcare-IoT security: A review and risk mitigation. *IEEE Access*, 11: 145869-145896. <https://doi.org/10.1109/ACCESS.2023.3346320>

[9] Shang, M.Y., Zhou, Y.H., Fujita, H. (2021). Deep reinforcement learning with reference system to handle constraints for energy-efficient train control. *Information Sciences*, 570: 708-721. <https://doi.org/10.1016/j.ins.2021.04.088>

[10] Guan, Z., Li, Y., Yu, S.Q., Yang, Z. (2023). Deep reinforcement learning-based full-duplex link scheduling in federated learning-based computing for IoMT. *Transactions on Emerging Telecommunications Technologies*, 34(3): e4724. <https://doi.org/10.1002/ett.4724>

[11] Zhang, H.J., Wang, H.Y., Li, Y.B., Long, K.P., Nallanathan, A. (2023). DRL-driven dynamic resource allocation for task-oriented semantic communication. *IEEE Transactions on Communications*, 71(7): 3992-4004. <https://doi.org/10.1109/TCOMM.2023.3274145>

[12] Jabbar, Z.S., Al-Neami, A.Q., Khawwam, A.A., Salih, S.M. (2023). Liver fibrosis processing, multiclassification, and diagnosis based on hybrid machine learning approaches. *Indonesian Journal of Electrical Engineering and Computer Science*, 29(3): 1614-1622.

[13] Bibri, S.E., Krogstie, J., Kaboli, A., Alahi, A. (2024).

- Smarter eco-cities and their leading-edge artificial intelligence of things solutions for environmental sustainability: A comprehensive systematic review. *Environmental Science and Ecotechnology*, 19: 100330. <https://doi.org/10.1016/j.es.2023.100330>
- [14] Liu, J., Lin, C.H.R., Hu, Y.C., Donta, P.K. (2022). Joint beamforming, power allocation, and splitting control for SWIPT-enabled IoT networks with deep reinforcement learning and game theory. *Sensors*, 22(6): 2328. <https://doi.org/10.3390/s22062328>
- [15] Park, G., Lee, W., Joe, I. (2020). Network resource optimization with reinforcement learning for low power wide area networks. *EURASIP Journal on Wireless Communications and Networking*, 2020: 1-20. <https://doi.org/10.1186/s13638-020-01783-5>
- [16] Xu, W., Lei, H., Shang, J. (2021). Joint topology construction and power adjustment for UAV networks: A deep reinforcement learning based approach. *China Communications*, 18(7): 265-283. <https://doi.org/10.23919/JCC.2021.07.021>
- [17] Xiao, Y.L., Niu, G.H., Xiao, L., Ding, Y.Z., Liu, S.C., Fan, Y.X. (2020). Reinforcement learning based energy-efficient internet-of-things video transmission. *Intelligent and Converged Networks*, 1(3): 258-270. <https://doi.org/10.23919/ICN.2020.0021>
- [18] Zhang, Y.Z., Lan, D.Y., Wang, C., Wang, P., Liu, F.Q. (2021). Deep reinforcement learning-aided transmission design for multi-user V2V networks. In 2021 IEEE Wireless Communications and Networking Conference (WCNC), Nanjing, China, pp. 1-6. <https://doi.org/10.1109/WCNC49053.2021.9417249>
- [19] El Jamous, Z., Davaslioglu, K., Sagduyu, Y.E. (2022). Deep reinforcement learning for power control in next-generation WIFI network systems. In MILCOM 2022-2022 IEEE Military Communications Conference (MILCOM), Rockville, USA, pp. 547-552. <https://doi.org/10.1109/MILCOM55135.2022.10017530>
- [20] Jiang, F., Wang, K., Dong, L., Pan, C., Yang, K. (2020). Stacked autoencoder-based deep reinforcement learning for online resource scheduling in large-scale MEC networks. *IEEE Internet of Things Journal*, 7(10): 9278-9290. <https://doi.org/10.1109/JIOT.2020.2988457>
- [21] Sharma, H., Budhiraja, I., Kumar, N., Tekchandani, R.K. (2022). Secrecy rate maximization for THz-enabled FEMTO edge users using deep reinforcement learning in 6G. In IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), New York, USA, pp. 1-6. <https://doi.org/10.1109/INFOCOMWKSHPS54753.2022.9798370>
- [22] Chen, C., Wu, F. (2023). Radio map-based trajectory design for UAV-assisted wireless energy transmission communication network by deep reinforcement learning. *Electronics*, 12(21): 4469. <https://doi.org/10.3390/electronics12214469>
- [23] Sande, M.M., Hlophe, M.C., Maharaj, B.T. (2021). Access and radio resource management for IAB networks using deep reinforcement learning. *IEEE Access*, 9: 114218-114234. <https://doi.org/10.1109/ACCESS.2021.3104322>
- [24] Al-Sa'D, M.F., Tlili, M., Abdellatif, A.A., Mohamed, A., Elfouly, T., Harras, K., O'Connor, M.D. (2018). A deep learning approach for vital signs compression and energy efficient delivery in mhealth systems. *IEEE Access*, 6: 33727-33739. <https://doi.org/10.1109/ACCESS.2018.2844308>
- [25] Zhou, G., Pan, C., Ren, H., Xu, D., Zhang, Z., Wang, J., Schober, R. (2023). A framework for transmission design for active RIS-aided communication with partial CSI. *IEEE Transactions on Wireless Communications*, 23(1): 305-320. <https://doi.org/10.1109/TWC.2023.3277514>
- [26] Rahmani, H., Shetty, D., Wagih, M., Ghasempour, Y., Palazzi, V., et al. (2023). Next-generation IoT devices: Sustainable eco-friendly manufacturing, energy harvesting, and wireless connectivity. *IEEE Journal of Microwaves*, 3(1): 237-255. <https://doi.org/10.1109/JMW.2022.3228683>
- [27] Jabbar, Z.S., Alneami, A.Q., Salih, S.M., Khawwam, A.A. (2023). Liver fibrosis detection and classification for shear wave elastography (SWE) images based on convolutional neural network (CNN). *AIP Conference Proceedings*, 2787(1). <https://doi.org/10.1063/5.0148350>
- [28] Yue, L., Ganesan, P., Sathish, B.S., Manikandan, C., Niranjana, A., Elamara, V., Hussein, A.F. (2018). The importance of dithering technique revisited with biomedical images—A survey. *IEEE Access*, 7: 3627-3634. <https://doi.org/10.1109/ACCESS.2018.2888503>
- [29] Idrees, A.K., Idrees, S.K., Ali-Yahiya, T., Couturier, R. (2023). Multibiosensor data sampling and transmission reduction with decision-making for remote patient monitoring in IoMT networks. *IEEE Sensors Journal*, 23(13): 15140-15152. <https://doi.org/10.1109/JSEN.2023.3278497>
- [30] Chronis, C., Anagnostopoulos, G., Politi, E., Dimitrakopoulos, G., Varlamis, I. (2023). Dynamic navigation in unconstrained environments using reinforcement learning algorithms. *IEEE Access*, 11: 117984-118001. <https://doi.org/10.1109/ACCESS.2023.3326435>
- [31] Seid, A.M., Erbad, A., Abishu, H.N., Albaseer, A., Abdallah, M., Guizani, M. (2023). Multiagent federated reinforcement learning for resource allocation in UAV-enabled internet of medical things networks. *IEEE Internet of Things Journal*, 10(22): 19695-19711. <https://doi.org/10.1109/JIOT.2023.3283353>
- [32] Askar, N.A., Habbal, A., Mohammed, A.H., Sajat, M.S., Yusupov, Z., Kodirov, D. (2022). Architecture, protocols, and applications of the Internet of Medical Things (IoMT). *Journal of Communications*, 17(11): 900-918. <https://doi.org/10.12720/jcm.17.11.900-918>
- [33] Wang, D., Liu, J., Yao, D., Member, I.E.E.E. (2020). An energy-efficient distributed adaptive cooperative routing based on reinforcement learning in wireless multimedia sensor networks. *Computer Networks*, 178: 107313. <https://doi.org/10.1016/j.comnet.2020.107313>
- [34] Nguyen, T.V., Hua, D.T., Huong, T.H., Hoang, V.T., Dao, N.N., Cho, S. (2023). Intelligent QoE management for IoMT streaming services in multi-user downlink RSMA networks. *IEEE Internet of Things Journal*, 11(7): 12602-12618. <https://doi.org/10.1109/JIOT.2023.3334473>
- [35] Ding, X., Zhang, Y., Li, J., Mao, B., Guo, Y., Li, G. (2023). A feasibility study of multi-mode intelligent fusion medical data transmission technology of industrial Internet of Things combined with medical Internet of Things. *Internet of Things*, 21: 100689. <https://doi.org/10.1016/j.iot.2023.100689>

- [36] Priya, B., Malhotra, J. (2023). IMNET: Intelligent rat selection framework for 5G enabled IoMT network. *Wireless Personal Communications*, 129(2): 911-932. <https://doi.org/10.1007/s11277-022-10163-9>
- [37] Abo-Eleneen, A., Abdellatif, A.A., Mohamed, A., Erbad, A. (2022). RLENS: RL-based energy-efficient network selection framework for IoMT. In 2022 Wireless Telecommunications Symposium (WTS), Pomona, USA, pp. 1-6. <https://doi.org/10.1109/WTS53620.2022.9768166>
- [38] Yuan, X., Zhang, Z., Feng, C., Cui, Y., Garg, S., Kaddoum, G., Yu, K. (2022). A DQN-based frame aggregation and task offloading approach for edge-enabled IoMT. *IEEE Transactions on Network Science and Engineering*, 10(3): 1339-1351. <https://doi.org/10.1109/TNSE.2022.3218313>