IoT-Enabled Deep Reinforcement Learning for Adaptive Waste Management in Hospital Environments

M. M. U. R. Usmani, R. Rafiq, and M. R. Khan,
"IoT-Enabled Deep Reinforcement Learning for Adaptive Waste Management in Hospital Environments",
Latin-American Journal of Computing (LAJC), vol. 13, no. 2, 2026.

ARTICLE HISTORY
Received 24 December 2025
Accepted 23 June 2026
Published 7 July 2026

Muhammad Masood Ul Rahman Usmani
Department of Computer Science
COMSATS University Islamabad, Sahiwal Campus
Sahiwal, Pakistan
muhammadmasoodulrahmanusmani@gmail.com
ORCID: 0009-0001-9948-111X

Rimsha Rafiq
Department of Computer Science
COMSATS University Islamabad, Sahiwal Campus
Sahiwal, Pakistan
rimshach026@gmail.com
ORCID: 0009-0009-4002-0798

Makki Riaz Khan
Department of Information Technology
Bahauddin Zakariya University, Lodhran Campus
Lodhran, Pakistan
makkiriazkhanwapda@gmail.com
ORCID: 0009-0009-0698-3857

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

IoT-Enabled Deep Reinforcement Learning for Adaptive Waste Management in Hospital Environments

Muhammad Masood Ul Rahman Usmani
COMSATS University Islamabad, Sahiwal Campus
Department of Computer Science
Sahiwal, Pakistan
muhammadmasoodulrahmanusmani@gmail.com

Rimsha Rafiq
COMSATS University Islamabad, Sahiwal Campus
Department of Computer Science
Sahiwal, Pakistan
rimshach026@gmail.com

Makki Riaz Khan
Bahauddin Zakariya University, Lodhran Campus
Department of Information Technology
Lodhran, Pakistan
makkiriazkhanwapda@gmail.com

***Abstract*—Healthcare industry produces considerable amount of biomedical waste, which needs efficient and safe ways of handling. The research paper puts forward a scheme of IoT-enabled DRL for hospital waste management. Unlike previously known schemes, the suggested scheme incorporates IoT-based sensing capabilities with a Proximal Policy Optimization agent. The performance of the framework is analyzed by developing a custom simulation environment based on the OpenAI Gym library, wherein the process of waste production is modeled as stochastic. According to the experimental results, the suggested scheme of PPO surpasses its competitors in all key criteria. Statistical validation using ANOVA and t-tests confirms that the improvements are significant. The findings highlight the potential of IoT-DRL integration for intelligent, adaptive, and efficient hospital waste management systems.**

***Keywords*—Biomedical Waste Management, Internet of Things (IoT), Deep Reinforcement Learning (DRL), Healthcare Systems, Intelligent Optimization, Hospital Safety, Sustainability.**

I. Introduction

The rising number of healthcare institutions has worsened the situation regarding biomedical and hospital waste management. Inadequate methods of separating and disposing of hospital waste can pose severe threats to the environment and raise the risks of disease transmission to healthcare providers and communities nearby [1], [2]. Conventional separation by healthcare providers can be laborious, prone to mistakes, and dangerous to healthcare workers [3], [4].

Recent developments in IoT and AI technologies have opened avenues towards new paradigms in smart waste management. Low-cost technologies in IoT (e.g., ultrasonic level sensors, RFID tags, and environment sensors) enable waste level detection at all times and provide real-time input to central processing units [5]–[7]. The application of intelligent systems to already existing infrastructure in IoT makes it possible to manage waste proactively and decrease the chances of overflow and pollution in sensitive sectors like hospitals [7].

Machine learning algorithms, specifically deep learning networks such as EfficientNet, ResNeXt, MobileNet, and YOLO, have achieved significant success in classifying different types of healthcare waste accurately with a classification accuracy of more than 95% [1], [4], [8]. Such systems bring the process of proper healthcare waste segregation at par with the national healthcare waste management guidelines, thus preventing the risks of exposure to harmful healthcare waste [8], [22].

However, despite such advancements, the majority of existing systems are still either reactive or limited to a labscale environment. For instance, even though IoT-enabled systems can send notifications once the bin is full, they do not have the predictive models required for optimal collection route planning [7], [19]. Similarly, vision-based sorting models, although data-intensive, may fail to effectively adapt to a healthcare setting [9], [10]. Thus, a critical need emerges for adaptive and smart systems to adapt to the ever-changing hospital waste production dynamics.

Deep Reinforcement Learning (DRL) is also an emerging promising technology for adaptive decision-making in challenging and uncertain environments. In contrast to static ML models, DRL agents are capable of learning optimal waste management techniques such as dynamic routing, vehicle allocation, and prioritization of bins through trial and error interactions in the environment [11], [12]. Multi-agent DRL techniques are also further extended and employed for simultaneous allocation of waste collection vehicles that aim to reduce energy costs, minimize delays in waste collection, and avoid the formation of dangerous wastes [11], [13].

In a hospital setting, the adoption of IoT-capable DRL models will be beneficial in making hospital waste management activities more predictive and adaptive. Real-time data inputs from sensors, along with models for reinforcement learning, such as Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN), will be useful in making hospital waste management more adaptive by optimizing routes, minimizing overflow levels, and maintaining biomedical waste regulations [7], [11].

This study introduces an IoT-assisted Deep Reinforcement

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

Learning framework for adaptive hospital waste management. The framework is founded on the recent advancements in the classification of healthcare waste in [8], proactive IoT assistance in hospital waste management in [7], and multiagent reinforcement learning with the goal of smart cities in [11].

The contributions of this paper are:

- An IoT-based sensor system capable of continuously tracking waste generation, availability of bins, and other risk factors in hospital environments.
- An accurate model for classification of different types of medical waste (hazardous, infectious, general, sharps, etc.) using deep learning techniques on images and sensor data.
- A DRL agent responsible for learning adaptive policies regarding waste collection, bin services, routing, and disinfection or storage considering safety and cost aspects.
- Evaluation on hospital data to assess improved classification accuracy, lower cost and risk, and ability to adapt to variable amounts of waste.

The second part highlights related literature on IoT, machine learning, and RL on Waste Management. The third section reviews the proposed IoT-enabled DRL framework. The fourth part highlights the experiment setup. The fifth section presents results discussion and comparison. The sixth part provides conclusions and future directions.

II. Related Work

Hospital and biomedical waste management has been an increasingly growing problem has attracted a lot of interest from the scientific community lately. This is because of the associated effects of infection control, sustainability, and efficiency. Traditional methods involving manual segregation and collection at pre-documented times have been found to be increasingly ineffective in dealing with the dynamic and dangerous nature of hospital waste [1] [2]. As such, there has been considerable interest in the use of emerging technologies such as IoT, ML, and DRL for smart and automated hospital waste management models.

A. Deep Learning Applied to Healthcare Waste Classification

A recent area that has shown great promise for automatic classification of healthcare waste is based on the principles of deep learning. Various research articles that applied transfer learning with ResNeXt, EfficientNet, and MobileNet models have demonstrated classification achievements above 90% for different categories of healthcare wastes [1] [2]. Specifically, YOLO models have also demonstrated their efficiency with respect to real-time processing, with YOLOv5 and YOLOv8 models achieving above 95% classification accuracy for different categories of wastes with lower inference times, making these models suitable for hospital settings [3] [4] [8]. Nevertheless, these models also rely heavily upon large-scale healthcare datasets, and their applicability for classification with hospital settings has shown some challenges [9] [10].

B. IoT-Enabled Smart Waste Systems

IoT-Smart garbage management systems in the IoT environment are also explored in terms of urban waste collection points (WCC) and smart cities. The usage of low-cost sensor nodes and cloud computing contributes to realtime monitoring of the fill level, temperature, humidity, and environmental aspects of garbage bins in the IoT-based system [5]– [7]. ProWaste technology, which involves the application of machine learning algorithms and IoT sensors, demonstrates prediction rates of above 99% in forecasting situations of overflow at WCCs. The conceptual structures of smart garbage bins in the IoT environment are also explored, incorporating the use of sensors, compression systems, and solar-based circuits to maximize effectiveness and minimize costs [7] [11] [20] [21].

C. Reinforcement Learning and Optimization

Routing optimization for waste collection is an important applications area in the healthcare sector, considering the consequences of inefficient waste collection, which can result in accelerated health risk escalation. Conventional methods for optimization, also originating in the Vehicle Routing Problem, have been adapted for superior effectiveness, thereby incorporating stochastic optimization techniques [12] [19]. In the more recent past, reinforcement learning has been employed for effective dynamic optimization in uncertain settings for waste routing, thereby facilitating adaptability according to real-time waste generation levels [11] [13]. Furthermore, multi-agent deep reinforcement learning setups ensure greater flexibility, incorporating vehicle dispatch and management for multiple zones in an upscale healthcare facility, thereby increasing energy efficiency and overall [11].

D. Integration in Healthcare Contexts

However, their application to a hospital environment is still very limited despite significant advances in general smart waste management. Biomedical waste is unique in hazardous material classification, strict regulatory compliance, and infection risks for healthcare staff [8] [22]. IoT-enabled sensing with DRL can help build adaptive, predictive frameworks that manage the complexity of the waste system of a hospital. In addition, such integration fits into the sustainability and public health goals while ensuring compliance with the national and international biomedical waste standards [2], [7], [11].

Summarily, from the literature, there are promising fronts but also some identified gaps. The existing systems based on IoT are mostly on notification, and classifications based on deep learning focused on correctness in lieu of optimized routes. The application of reinforcement learning, although a robust method, is less considered in a healthcare setting scenario. There consequently arises a need for a unified framework of DRL using IoT-enabled adaptive hospital waste management.

Table I summarizes recent approaches in IoT, machine learning, and DRL-based waste management systems. The comparison shows the limitations of the existing methods, motivating the need for an integrated IoT-DRL framework.

III. Methodology

The proposed framework utilizes a combination of IoT sensing and DRL optimization techniques to provide an adaptive biomedical waste management system within a hospital setting. The methodology is comprised of three fundamental

TABLE I. Comparative Summary of Recent IoT, ML/DL, and DRL Approaches for Waste Management (2021–2025)

Authors / Year	Method / Model	Application / Dataset	Key Results / Findings
Zhou et al. (2022) [1]	ResNeXt-50 (DL)	Private 8-class medical waste dataset	Achieved 97.2% accuracy in healthcare waste classification.
Kumar et al. (2021) [2]	EfficientNet-B7 (TL, DL)	COVID-related biomedical waste streams	Reported 99% accuracy; highlighted AI for circular economy in healthcare waste.
Kunwar & Rai (2025) [8]	YOLOv5-s, YOLOv8, EfficientNet-B0	Medical Waste Dataset 4.0 + Pharma-biomedical dataset (Nepal)	YOLOv5-s achieved 95.06% accuracy; deployed with bin-color mapping to Nepal's HCW standards.
Mok (2024) [4]	YOLO + IoT Integration	67,860 images of HCW detection	Achieved mAP of 98%; demonstrated IoT-enhanced automated sorting.
Lahoti et al. (2024) [3]	Computer Vision + Robotic Arm	Multi-class waste segregation prototype	Enabled real-time robotic segregation of hospital waste.
Stephan et al. (2025) [6]	ProWaste (IoT + ML, Decision Tree + BPSO)	Urban WCCs, 6,954 daily records (Bengaluru)	Reduced missed pickups; >99.8% macro-F1 with only 3 predictive features; deployed via mobile app.
Patil et al. (2021) [5]	Ultrasonic Sensors + IoT	E-waste and bin monitoring prototype	IoT-based notification system reduced manual inspections.
Shanthini et al. (2021) [7]	IoT, RFID, WSN	Smart City Waste Collection	Found LoRaWAN-based IoT systems outperform others in efficiency.
Rajani et al. (2022) [11]	Multi-Agent Deep RL (DRL)	IoT-driven smart waste management (simulation)	Proposed platform-agnostic DRL framework; optimized routing, reduced overflow.
Mishra et al. (2022) [12]	Route Optimization (VRP + IoT)	Bhubaneswar City MSW data	Reduced vehicle distance by 30.28%, OpEx by 29.07%.
Abuga et al. (2022) [13]	IoT + Fuzzy Logic	Real-time smart garbage bins	Achieved high reliability in dynamic bin monitoring and waste-level prediction.
Gondal et al. (2023) [9]	Hybrid Deep Learning Model	Real-time garbage classification	Achieved 99% training/validation accuracy with automated bin sorting.
Zhang et al. (2022) [10]	Cascade R-CNN (enhanced with dilated convolutions)	Garbage detection for small objects	Improved precision in detecting small waste objects in cluttered environments.

components, namely: (i) IoT Layer, (ii) DRL Layer, and (iii) System Workflow. Figure 1 shows the overall architecture.

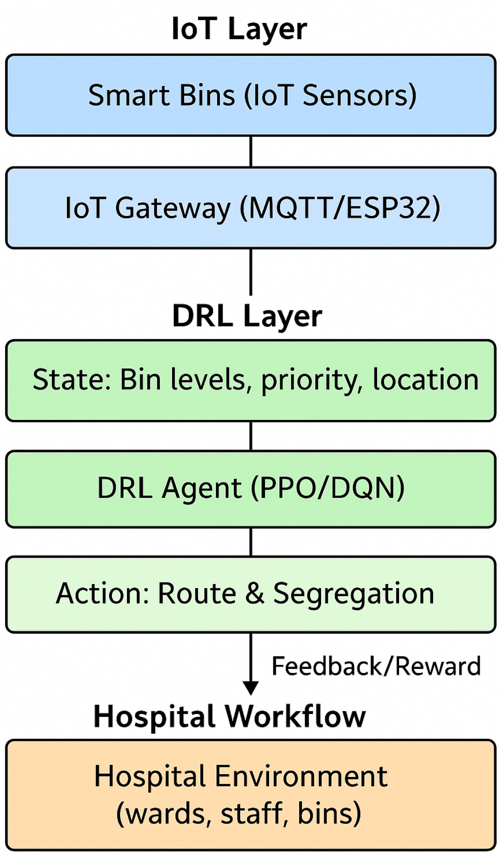

Fig. 1: Proposed IoT-Enabled DRL Framework for Adaptive Hospital Waste Management

A. IoT Layer

This layer contains a network of intelligent biomedical waste dumpsters with ultrasonic sensors for measuring the fill levels of the dumpsters, an RFID for identifying biowaste types, and temperature sensors for the detection of infectious diseases. Data from the dumpers is relayed through lightweight messaging transports such as MQTT to a cloud server. This enables real-time mapping of biowaste generation behavior in hospital wards, operation theatres, and labs. Figure 2 illustrates the PPO architecture used in this study.

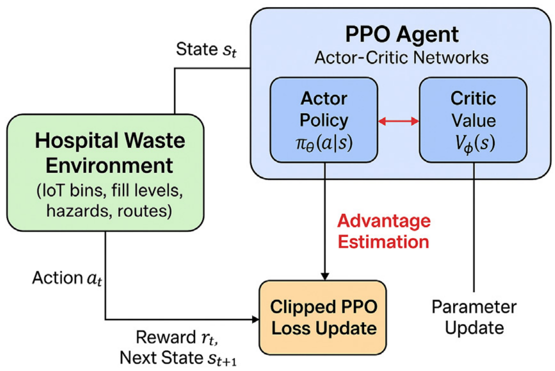

Fig. 2: Proposed PPO Model Architecture

B. DRL Layer

The DRL layer formulates waste management as a sequential decision-making problem. Every individual in charge of collecting waste material from different hospital locations (for instance, a robot cart or human personnel) is considered an agent functioning within the hospital premises. The state space will consist of bin levels, location, waste importance

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

Algorithm 1 Training Procedure of the Proposed PPO Agent

Randomly initialize the policy parameters θ and value network parameters ϕ
for each training episode **do**
  Reset the environment and obtain the initial state s
  for each interaction step until horizon T **do**
    Sample an action a according to the current policy $\pi_\theta(\cdot|s)$
    Apply a to the environment
    Receive reward r and the subsequent state s'
    Save the transition (s, a, r, s') in the rollout buffer
    Set $s \leftarrow s'$
  end for
  Estimate discounted returns and compute advantages using GAE
  Evaluate the clipped objective function:

$$L_{\text{PPO}} = \mathbb{E}\Big[\min\Big(r(\theta)\hat{A}, \text{clip}(r(\theta), 1-\epsilon, 1+\epsilon)\hat{A}\Big)\Big]$$

  Optimize the policy network using gradient ascent
  Update the value network by minimizing the value prediction error
end for

(hazardous or non-hazardous), and availability of employees. Actions will comprise routing actions (next bin to collect, segregation of waste material, or inaction). Reward functions will aim at (i) reducing the routing distance and duration of waste collection; (ii) avoiding overflowing of bins containing hazardous waste materials; and (iii) compliance with the biomedical safety standards. PPO will be chosen as the algorithm to use based on DRL since it is stable and suited for continuous state spaces.

C. *System Workflow*

The overall system workflow is as follows:

- IoT sensors collect real-time data from hospital waste bins.
- Data is transmitted to a cloud server and preprocessed (normalization, bin classification).
- The DRL agent receives the current hospital state and selects an optimal action (e.g., which bin to service).
- The action is executed by the waste collection system (robot/human).
- The environment returns feedback (updated state, reward signal).
- The DRL model updates its policy iteratively through training episodes.

This adaptive loop allows the system to dynamically optimize collection routes and schedules while minimizing risks of contamination and operational inefficiencies.

1) PPO Training Algorithm: Use Proximal Policy Optimization (PPO) to train the DRL agent [14]. Algorithm 1 presents the pseudocode of the training process.

D. *Simulation Environment*

For assessing the proposed framework, a simulation environment was designed with the help of the OpenAI Gym API. The environment considers hospital waste bins, collectors,

Algorithm 2 Custom OpenAI Gym Environment for Hospital Waste Management

Require: Number of bins B, Hospital layout L, Maximum steps T
Ensure: State transitions, reward signals
Initialize: Bin fill levels $f_b = 0$, Waste types $w_b \in \{hazardous, non\text{-}hazardous\}$, Agent position p_0
for Epoches $= 1$ to N **do**
  Reset: $f_b \sim U(0, 0.3)$, $p_0 =$ nurse station
  for $t = 0$ to $T-1$ **do**
    State: $s_t = \{f_b, w_b, p_t\}$
    Agent selects action $a_t \in$ {move to bin, collect, idle}
    if action == collect **then**
      Empty selected bin, update $f_b \leftarrow 0$
      Reward $r_t = +\alpha$ (hazardous) or $+\beta$ (non-hazardous)
    end if
    Update fill levels: $f_b \leftarrow f_b + \delta$
    if $f_b > 1.0$ **then**
      Penalize: $r_t = -\gamma$ (overflow)
    end if
    Update agent position p_{t+1} and hospital state
  end for
  if all bins empty or $t = T$ **then**
    Terminate episode
  end if
  Return trajectory $\{s_t, a_t, r_t, s_{t+1}\}$ for PPO training
end for

and IoT sensors dynamics. In each episode, the simulation reflects the dynamics of a hospital shift, where the agent tries to find an optimal path and waste segregation strategy. The algorithm for the environment is presented in Algorithm 2.

IV. Results & Experiments

This section outlines the experimental configuration, explains the evaluation criteria, and analyzes the performance of IoT-based DRL framework for adaptive hospital waste management system. The experiments have been performed using the Python environment that uses the custom OpenAI Gym environment described in Algorithm 2 and TensorFlow for DRL. Experiments were conducted on a workstation with Intel i7-12700 CPU, 32GB RAM, and NVIDIA RTX 3080 GPU.

A. *Experimental Setup*

The hospital environment was simulated with $B = 50$ bins distributed across wards, laboratories, and operation theaters. Each bin was randomly assigned as hazardous (30%) or non-hazardous (70%), and waste generation rates followed a Poisson distribution. It is important to note that the experimental evaluation is conducted using a simulated hospital environment. IoT data streams are synthetically generated to emulate real-world waste generation patterns. The DRL agent (PPO) was compared against baseline methods:

- Rule-Based Scheduling (RBS): Fixed-time collection intervals.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

TABLE II. Simulation and PPO Hyperparameters

Parameter	Value
α (Hazardous reward)	10
β (Non-hazardous reward)	5
γ (Overflow penalty)	15
λ (Poisson rate)	0.2–0.5
Learning Rate	3×10^{-4}
Discount Factor (γ)	0.99
PPO Clipping (ϵ)	0.2
Batch Size	64
Epochs	10

TABLE III. Performance Comparison of Waste Management Approaches

Method	Overflow (%)	Coll. Time (min)	Haz. Score	Energy (kWh)
RBS	18.2	12.5	0.62	4.8
SPH	12.9	9.4	0.71	4.1
DQN	9.7	8.1	0.76	3.7
Proposed PPO	**5.4**	**6.9**	**0.89**	**3.2**

- Shortest Path Heuristic (SPH): Greedy routing to the nearest non-empty bin.
- Deep Q-Network (DQN): Model-free baseline with discrete actions.

B. Hyperparameter Settings

To ensure reproducibility, the following hyperparameters were used:

C. Metrics for Evaluation

The system was evaluated using the following metrics:

- **Overflow Rate (%)**: Percentage of bins that exceeded capacity.
- **Average Collection Time (min)**: Mean time per bin service.
- **Hazardous Waste Priority Score**: Ratio of hazardous waste collected before overflow.
- **Energy Efficiency (kWh)**: Energy consumed by collection robots/vehicles.

D. Results and Discussion

Table III shows the comparison across baseline methods. The proposed PPO-based DRL outperformed all baselines, achieving the lowest overflow rate and highest hazardous waste priority compliance. Fig. 3 and Fig. 4 illustrate the overflow reduction and training convergence, respectively.

The findings validate that DRL-based adaptive scheduling helps minimize the risks of overflow as well as enhances the efficiency of hazardous waste management. Additionally, the reduced energy consumption indicates the sustainability advantages of such an approach in practical settings of hospitals as well.

E. Statistical Significance Analysis

We further conducted statistical significance tests to assess the robustness of the outcomes of the proposed method and the baselines. We executed 30 independent episodes for each method based on randomly generated patterns of waste production. A one-way ANOVA test is followed by pairwise two-tailed t-tests using Bonferroni correction for overflow rate.

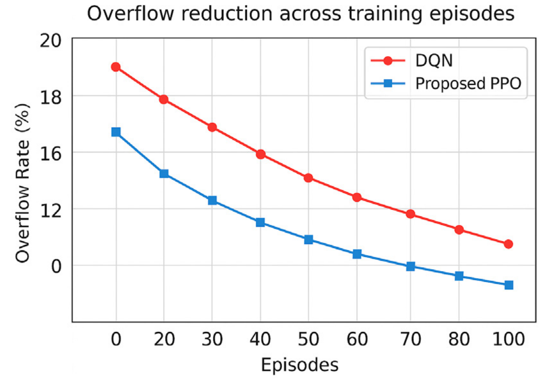

Fig. 3: Overflow reduction across training episodes

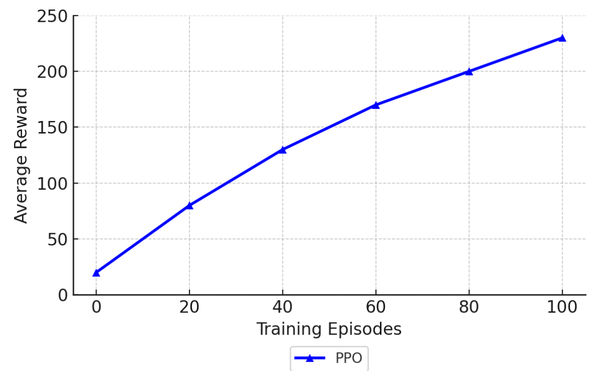

Fig. 4: Training convergence of PPO agent

Table IV shows that the ANOVA test yielded a statistically significant difference among methods ($p < 0.001$). Post-hoc pairwise t-tests (Table V) confirmed that the proposed PPO significantly outperformed RBS, SPH, and DQN.

TABLE IV. One-Way ANOVA Results for Overflow Rate Across Methods

Source	DF	F-Value	p-Value
Between Groups	3	24.56	< 0.001
Within Groups	116	–	–
Total	119	–	–

TABLE V. Pairwise t-Test Results (Overflow Rate, Bonferroni Corrected)

Comparison	t-Statistic	p-Value
PPO vs. RBS	7.82	< 0.001
PPO vs. SPH	6.11	< 0.001
PPO vs. DQN	3.45	0.002

The statistical analysis confirms that the improvements of PPO are not due to random chance, but are significant at the 95% confidence level. Thus, the proposed DRL-based approach provides a robust and reliable optimization strategy for hospital waste management.

F. Complexity Analysis

To assess the computational feasibility of the proposed framework, we analyzed the time and space complexity of both the IoT data acquisition and the DRL training phases.

1) IoT Communication Overhead: The IoT layer relies on MQTT-based communication for transmitting sensor data from B bins. Each message has an average payload size of $O(1)$ (bin fill level, type, timestamp). Thus, the per-step communication complexity is

$$O(B) \tag{1}$$

For a hospital with $B = 100$ bins, the per-second communication load remains within 50–100 KB, which is well within the capabilities of low-cost WiFi/LoRa gateways.

2) PPO Training Complexity: The PPO algorithm involves two key components: trajectory collection and policy updates. Let N be the number of steps per episode, E the number of episodes, and M the number of policy update iterations.

- **Trajectory Collection**: Each step requires state evaluation and action selection, with complexity $O(d)$, where d is the dimensionality of the state vector. Thus, trajectory collection over N steps per episode costs

$$O(Nd) \tag{2}$$

- **Policy Update**: PPO uses stochastic gradient descent on batches of size b over M iterations. Each forward-backward pass has complexity $O(\theta)$, where θ is the number of neural network parameters. Thus, update complexity is

$$O(Mb\theta) \tag{3}$$

Overall training complexity is therefore

$$O(END + EMb\theta) \tag{4}$$

3) Space Complexity: The memory footprint consists of: replay buffer $O(Nd)$, neural network parameters $O(\theta)$, and IoT data queue $O(B)$. Total space complexity is

$$O(Nd + \theta + B) \tag{5}$$

In practice, with $N = 1000$ steps, $E = 500$ episodes, $M = 10$ iterations, $b = 64$, and $\theta \approx 10^5$, training can be completed using a single NVIDIA RTX 3060 GPU in under 4 hours. The IoT communication overhead is negligible relative to network capacity. Thus, the framework is computationally efficient and feasible for real-world hospital environments.

V. Discussion

The obtained results confirm that the proposed IoT-enabled PPO framework provides a significant improvement in hospital waste management efficiency compared to baseline approaches, including Rule-Based Scheduling (RBS), Shortest Path Heuristic (SPH), and Deep Q-Network (DQN). Specifically, the proposed model achieves lower overflow rates, reduced collection times, and improved prioritization of hazardous waste. These improvements are statistically validated using ANOVA and post-hoc t-tests, confirming that the observed performance gains are not due to random variation.

A. Interpretation of Results

These performance gains can be explained by two major reasons. Firstly, the IoT-based data acquisition layer allows for the adaptive tracking of waste production dynamics and thus helps make the decisions in a timely fashion depending on current conditions in the hospital, which is more efficient than static or heuristic methods.

Secondly, the PPO method offers stable learning in stochastic environments thanks to the clipped surrogate objective function that does not allow for very large updates of the policy function during learning. This feature is especially important in hospitals due to the uncertainty in waste production and other constraints.

B. Limitations

However, despite all the positive results achieved, some limitations should be mentioned. First of all, the evaluation of the proposed solution takes place within a simulation setting wherein the patterns of waste generation are artificially generated. Even though these patterns are created with an intention to mirror realistic conditions at the hospital, they might not be sufficiently diverse enough.

Waste generation in real life may be subject to some unexpected factors like outbreaks of infectious diseases, changes in seasons, and the increase in patient flow, among others. In particular, during a large-scale outbreak, the volume and content of biomedical waste may change substantially. Hence, the evaluation of the system that takes place within a simulation is an idealistic case study.

C. Ethical and Privacy Considerations

This approach uses information from hospital waste, some of which might be confidential metadata, like timestamps that may inadvertently correspond to patient behavior. Compliance with data protection laws, HIPAA and GDPR, is a prerequisite before implementation of such an approach.

In addition, although automation decreases the exposure to harmful waste for people, it is vital to retain proper human control. System malfunctions, wrong classifications or any other circumstances can be dangerous. Thus, the ethical implementation of the system requires proper mechanisms of monitoring and fail-safes.

D. Generalizability and Future Deployment

Though the model has been designed with hospital waste management in mind, the suggested IoT-DRL model can be easily customized for any other type of waste management operations, such as intelligent cities and industries. The implementation of multi-agent reinforcement learning will help scale the model even more, as it will allow coordination among several collectors.

Further research should be directed at applying the model in practice through collaboration with the healthcare sector. Besides, introducing uncertainty analysis, anomaly detection, and edge computing is a good idea as well.

VI. Conclusion and Future Work

This paper provided a novel IoT based deep reinforcement learning (DRL) approach to adaptive hospital waste management. The proposed method consists of using an IoT based sensor along with proximal policy optimization (PPO) for dynamic optimization of waste collection, route planning, and prioritization of hazardous wastes.

From the results of our experiments performed in the custom simulation environment built on top of the OpenAI Gym, we can see that the behavior of the proposed DRL method is significantly better than those of the baselines methods RBS, SPH, and DQN. Our model provides the results with lower overflow rate, lesser collection time, and better prioritization of hazardous wastes.

The primary contributions of this study are summarized as follows:

- Development of an integrated IoT-DRL framework for adaptive hospital waste management.
- Design of a custom simulation environment to model realistic waste generation and collection dynamics.
- Comprehensive performance evaluation demonstrating the effectiveness of PPO in complex and stochastic environments.
- Analysis of computational complexity and consideration of ethical and deployment-related aspects.

However, there are some limitations of this study. The assessment was conducted based on artificial datasets in a simulation setting, which does not completely account for real-life randomness in hospital operations. Hence, the obtained results correspond to an ideal case scenario.

The future works will be directed towards practical implementation of the proposed approach together with hospitals. Furthermore, the use of the multi-agent reinforcement learning may make the system capable of collecting waste cooperatively via several autonomous agents in hospital areas. Some additional improvements might include robustness enhancement via modeling of uncertainties, anomalies detection and failure-proofed IoT communications.

Additionally, the employment of energy-efficient DRL models and edge computing methods will enhance the scalability and application of the framework in limited-resource settings. Moreover, apart from being applied to hospitals, the proposed framework can potentially be used for other applications such as smart city waste management systems, industrial waste handling systems and other public health infrastructure.

Thus, the combination of IoT and DRL may lead to the development of an advanced waste management framework.

Funding Statement

There was no outside support for this study.

Acknowledgment

The Department of Computer Science at COMSATS University Islamabad's Sahiwal Campus is acknowledged by the authors for providing the academic resources and assistance required to complete this study.

Author Contributions

The authors' contributions follow the CRediT (Contributor Roles Taxonomy) as follows:

- **Muhammad Masood Ul Rahman Usmani**: Conceptualization, Methodology, Writing Original Draft & Editing
- **Rimsha Rafiq**: Data Curation, Visualization, Review & Editing
- **Makki Riaz Khan**: Conceptualization, Writing – Review & Editing

References

[1] H. Zhou, X. Yu, A. Alhaskawi, Y. Dong, Z. Wang, Q. Jin, X. Hu, Z. Liu, V. G. Kota, M. H. Abdulla, S. H. A. Ezzi, B. Qi, J. Li, B. Wang, J. Fang, and H. Lu, "A deep learning approach for medical waste classification," *Scientific Reports*, vol. 12, no. 1, p. 2159, 2022.

[2] N. M. Kumar, M. A. Mohammed, K. H. Abdulkareem, R. Damasevicius, S. A. Mostafa, M. S. Maashi, and S. S. Chopra, "Artificial intelligence-based solution for sorting COVID related medical waste streams and supporting data-driven decisions for smart circular economy practice," *Process Safety and Environmental Protection*, vol. 152, pp. 482–494, 2021.

[3] J. Lahoti, J. Sn, M. V. Krishna, M. Prasad, R. Bs, N. Mysore, and J. S. Nayak, "Multi-class waste segregation using computer vision and robotic arm," *PeerJ Computer Science*, vol. 10, p. e1957, 2024.

[4] M. H. Mok, "YOLO combined with IoT for detection of healthcare waste," *Applied Sciences*, vol. 14, no. 3, p. 1167, 2024.

[5] U. Patil *et al.*, "IoT based smart waste management system," in *Proc. IEEE Int. Conf. Inventive Research in Computing Applications (ICIRCA)*, 2021, pp. 1–6.

[6] T. Stephan, S. M. Hari Krishna, C.-C. Lin, U. Sumesh, S. Agarwal, and H. Kim, "ProWaste for proactive urban waste management using IoT and machine learning," *Scientific Reports*, vol. 15, no. 1, p. 27790, 2025.

[7] S. Vishnu, S. R. J. Ramson, S. Senith, T. Anagnostopoulos, A. M. AbuMahfouz, X. Fan, S. Srinivasan, and A. A. Kirubaraj, "IoT-enabled solid waste management in smart cities," *Smart Cities*, vol. 4, no. 3, pp. 1004–1017, 2021.

[8] S. Kunwar and P. Rai, "Healthcare waste classification using deep learning aligned with Nepal's bin color guidelines," *arXiv preprint arXiv:2508.07450*, 2025.

[9] A. U. Gondal, M. I. Sadiq, T. Ali, M. Irfan, A. Shaf, M. Aamir, M. Shoaib, A. Glowacz, R. Tadeusiewicz, and E. Kantoch, "Real time multipurpose smart waste classification model for efficient recycling in smart cities using multilayer convolutional neural network and perceptron," *Sensors*, vol. 21, no. 14, p. 4916, 2021.

[10] C. Zhang, X. Zhang, D. Tu, and Y. Wang, "Small object detection using deep convolutional networks: Applied to garbage detection system," *Journal of Electronic Imaging*, vol. 30, no. 4, p. 043013, 2021.

[11] K. R. Rajani, A. Gaddam, and J. Gaddam, "IoT-based smart waste management using deep reinforcement learning," *Preprints*, 2022.

[12] A. Mishra, S. Ghosh, and D. P. Jena, "Internet of Things based waste management system for smart cities: A real time route optimization for waste collection vehicles," *International Journal of Scientific Research in Computer Science and Engineering*, vol. 6, no. 1, pp. 37–42, 2019.

[13] D. Abuga and N. S. Raghava, "Real-time smart garbage bin mechanism for solid waste management in smart cities," *Sustainable Cities and Society*, vol. 75, p. 103347, 2021.

[14] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," *arXiv:1707.06347*, 2017.

[15] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, 2015.

[16] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press, 2018.

[17] Y. Li, "Deep reinforcement learning: An overview," *arXiv preprint arXiv:1701.07274*, 2017.

[18] L. D. Xu, W. He, and S. Li, "Internet of Things in industries: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[19] A. Martikkala, B. Mayanti, P. Helo, A. Lobov, and I. F. Ituarte, "Smart textile waste collection system—Dynamic route optimization with IoT," *Journal of Environmental Management*, vol. 335, p. 117548, 2023.

[20] S. R. J. Ramson, D. J. Moni, S. Vishnu, T. Anagnostopoulos, A. A. Kirubaraj, and X. Fan, "An IoT-based bin level monitoring system for solid waste management," *Journal of Material Cycles and Waste Management*, vol. 23, no. 2, pp. 516–525, 2021.

[21] D. Baldo, A. Mecocci, S. Parrino, G. Peruzzi, and A. Pozzebon, "A multi-layer LoRaWAN infrastructure for smart waste management," *Sensors*, vol. 21, no. 8, p. 2600, 2021.

[22] S. Karki, S. R. Niraula, and S. Karki, "Perceived risk and associated factors of healthcare waste in selected hospitals of Kathmandu, Nepal," *PLOS ONE*, vol. 15, no. 7, p. e0235982, 2020.

[23] H. Wu, F. Tao, and B. Yang, "Optimization of vehicle routing for waste collection and transportation," *International Journal of Environmental Research and Public Health*, vol. 17, no. 14, p. 4963, 2020.

AUTHORS

Muhammad Masood Usmani

Muhammad Masood ul Rahman Usmani received his M.S. degree in Computer Science from COMSATS University Islamabad, Sahiwal Campus, Pakistan, in 2024, following his Master of Computer Science degree from The Islamia University of Bahawalpur, Pakistan. He is currently serving as a Faculty Member at Bahauddin Zakariya University, Lodhran Sub Campus, Pakistan, where he teaches undergraduate courses in Computer Science. His research interests include Machine Learning, Deep Learning, Reinforcement Learning, the Internet of Things (IoT), Computer Vision, and Artificial Intelligence for healthcare. He has authored and coauthored several research articles published in or submitted to international journals and conferences. His current research focuses on intelligent IoT systems, medical image analysis, and deep reinforcement learning based optimization for smart environments.

Rimsha Rafiq

Rimsha Rafiq received her M.S. degree in Computer Science from COMSATS University Islamabad, Sahiwal Campus, Pakistan, in 2024, following her Bachelor degree in Computer Science from Bahauddin Zakariya University, Multan, Pakistan. She is currently serving as a Teaching Faculty at COMSATS University Islamabad, Sahiwal Campus, Pakistan, where she teaches undergraduate courses in Computer Science. Her research interests include Machine Learning, Deep Learning, Reinforcement Learning and Artificial Intelligence for healthcare. Her current research focuses on medical image analysis and deep reinforcement learning based optimization for smart environments.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026

AUTHORS

Makki Riaz Khan

Makki Riaz Khan is a BS Information Technology student at Bahauddin Zakariya University, Sub Campus Lodhran, Pakistan. His research interests include Artificial Intelligence, Machine Learning, Deep Learning, and Computer Vision. He is a co-author of the paper "Multi-Class Classification of Alzheimer's Impairment using EfficientNet-B0". His technical expertise includes Python, TensorFlow, Keras, Scikit-learn, and OpenCV. He has developed several AI-based applications in healthcare and computer vision and is passionate about applying intelligent technologies to solve real-world problems.

M. M. U. R. Usmani, R. Rafiq, and M. R. Khan,
"IoT-Enabled Deep Reinforcement Learning for Adaptive Waste Management in Hospital Environments",
Latin-American Journal of Computing (LAJC), vol. 13, no. 2, 2026.