SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Deep Reinforcement Learning-Based Cyber Defense Mechanism for Intelligent Threat Mitigation in Healthcare IoT Networks

Dr. Vijay J. Upadhye[1], Dr.Nidhi Srivastava[2], Dr.Girish Kumar Pati[3], Gagan Tiwari[4], Leena Deshpande[5], Shanthi Vairavan[6], Suganya S[7], Kanchana K[8]

[1]Associate Professor, Parul Institute of Applied Scienes ,Parul University, PO Limda, Tal. Waghodia, District Vadodara, Gujarat, India, Email: vijay.upadhye82074@paruluniversity.ac.in ,Orcid Id- 0000-0002-8821-1720
[2]Professor , MSOPS, Maharishi University of Information Technology, Lucknow, Uttar Pradesh, India, Email: nidhi.srivastava@muit.in, Orcid Id- https://orcid.org/0000-0002-6661-5804
[3]Professor, Department of Gastroenterology, IMS and SUM Hospital, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India, Email: girishkumarpati@soa.ac.in, Orcid Id- 0000-0002-9389-1425
[4]Department of Computer Sciences,Noida international University, Greater Noida, Uttar Pradesh 203201, India, Email: gagan.tiwari@niu.edu.in
[5]Associate Professor, Department of Computer Engineering - Software Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, Email: leena.deshpande@vit.edu
[6]Computer Science, Professor & Principal, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email: shanthiv@maher.ac.in
[7]Department of Management Studies, Assistant Professor,Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email: ssuganyamba@maher.ac.in
[8]Department of Commerce, Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India, Email: kanchana@maher.ac.in

Abstract

Healthcare Internet of Things (HIoT) network has greatly enhanced real-time patient monitoring, remote diagnostic, and smart healthcare management, but has created more weaknesses to advanced cyberattacks including distributed denial-of-service (DDoS) attacks, malware injection attacks, unauthorized access, and manipulation of data attacks. Conventional intrusion detection systems usually use any of the following: static subscription and rule-based methods, which prove insufficient to deal with dynamic and changing cyber threats in heterogeneous healthcare IoT settings. To undo these drawbacks, the proposed study will present a Deep Reinforcement Learning-based cyber defense system with a Deep Q-Network (DQN) system to detect threats and mitigate them intelligently in healthcare IoT networks. The proposed framework constantly acquires the best defense mechanisms by engaging with network settings and dynamically chooses the right mitigation measures towards malicious behavior. The TON IoT and CICIDS2017 datasets with varied normal and attack traffic scenarios were used to run the experimental evaluation. Classification metrics such as Accuracy, Precision, Recall, F1-Score, and False Positive Rate (FPR) were used to perform performance analysis. The experimental outcomes prove that the suggested DQN-based system shows better intrusion detection and threat management capabilities, and the lower occurrence of false alarms than conventional machine learning-based intrusion detection methods. The suggested intelligent cyber defense architecture is a step towards the creation of safe, resilient, and scalable healthcare IoT ecosystems, since it incorporates deep reinforcement learning to manage real-time autonomous cybersecurity.

Keywords: Healthcare IoT, Deep Reinforcement Learning, DQN, Cybersecurity, Intrusion Detection, Threat Mitigation.

1. Introduction

The overwhelming development of Healthcare Internet of Things (HIoT) technologies has revolutionized the health care system in the modern world, including the smart monitoring of patients, wearable medical devices, smart diagnostic systems, and real-time healthcare communication networks. The infrastructures of healthcare

IoT can facilitate real-time data gathering, exchange, and processing of essential medical data, enhancing the quality of clinical decisions, access to remote healthcare, and patient-oriented services. Nevertheless, the high level of interconnectivity of medical devices and cloud-based healthcare systems has greatly exposed healthcare networks to advanced cyberattacks, such as distributed denial-of-service (DDoS), malware injection, unauthorized access, ransomware attacks, and threats of data manipulation (Khraisat et al., 2019). Cybersecurity has been identified as a key element in the maintenance of confidentiality, integrity and availability of healthcare communication systems because healthcare systems deal with very sensitive patient data, and the medical processes involved are so crucial that their integrity is considered a mission-critical requirement of healthcare (Ferrag et al., 2020).

Conventionally, intrusion detection systems (IDS) have been prevalent in detecting malicious network activities within the IoT ecosystems. These traditional solutions rely primarily on signature-based detection and fixed rule-based security features, which frequently cannot be used to respond to the zero-day attacks and dynamically occurring cyber threats (Kim et al., 2014). Moreover, healthcare IoT networks present massive heterogeneous traffic streams that constantly evolves over time, and existing cybersecurity models are inadequate to track dynamic threat mitigation. The current machine learning-based intrusion detection methods enhance the performance of attack classification, but most of these systems have high rates of false positives, lack the ability to adapt to new threats, and they cannot take necessary decisions in real-time (Javaid et al., 2016). Moreover, the distributed healthcare setting needs smart security tools that can learn independently and develop on-the-fly cyber protection in a highly dynamic attack environment (Diro&Chilamkurti, 2018).

New technologies in artificial intelligence and deep learning have also created new possibilities in developing intelligent cybersecurity architectures of IoT networks. One of such methods is Deep Reinforcement Learning (DRL), which has proved to be one of the most effective solutions in adaptive cyber defense due to its capabilities of learning the best security strategies by interacting in an ever-changing environment (Mnih et al., 2015). Reinforcement learning agents also possess a continuous improvement quality of their decision-making policies, unlike traditional supervised learning models, which is made and sustained by learning mechanisms that are driven by rewards, which allow them to conduct adaptive intrusion detection and mitigate threats automatically. The Deep Q-Network (DQN) is among the most popular DRL algorithms, as it involves both the so-called reinforcement learning and the use of deep neural networks to estimate optimal action-selection policies when operating in complex environments (Alavizadeh et al., 2022). This feature is what renders DQN particularly applicable to healthcare IoT cybersecurity applications, in terms of which the speed of response, adaptive learning, and intelligent mitigation are crucial to counter emerging cyber threats.

A number of new studies have examined the use of deep learning and reinforcement learning to intrusion detection systems. The study by Ferrag et al. (2020) provides an extensive overview of deep learning techniques to be used in cybersecurity intrusion detection, highlighting the usefulness of neural-network-based in large-scale IoT security systems. Gueriani et al. (2023) emphasized the increased relevance of deep reinforcement learning to an intelligent intrusion detector in IoT networks, and Jamshidi et al. (2025) revealed that the DRL-based adaptive defenses are effective in developing a mitigation of cyberattack. Also, some datasets, including TON_IoT and CICIDS2017 have offered real-life attack traffic conditions to test intelligent intrusion detection systems in IoT and cyber-physical infrastructures (Alsaedi et al., 2020). Notwithstanding these developments, the current literature is still struggling with the issues of adaptive policy optimization, false alarms suppression and real-time autonomous threat prevention on healthcare-specific IoT systems.

Based on these difficulties, this study presents an intelligent cyber defense mechanism, implemented via Deep Q-Network, to adaptively mitigate threats in healthcare IoT networks. The suggested framework will help to enhance the performance of cybersecurity because it will allow the intelligent selection of attack detection, adaptive selection of mitigation actions, and autonomous learning in healthcare communication infrastructures. The network constantly communicates with network traffic environments, trains the best strategies of defense through reward-based reinforcement learning, and real-time applies mitigations on malicious activities. To assess the effectiveness of the proposed framework in cybersecurity, benchmark

intrusion detection data sets and classification values such as Accuracy, Precision, Recall, F1-Score and False Positive Rate (FPR) are used to test the proposed framework.

The significant contribution of the research is the creation of a new DQN-based intelligent architecture of cyber defense that is specifically created to operate in healthcare IoT settings. The proposed framework contrasts with the conventional fixed intrusion detection systems by incorporating deep reinforcement learning in generating adaptive attack detection and intelligent responses to evolving network scenarios. The research also adds value by applying a real-time threat reduction approach that is able to automatically decide on the best cybersecurity measures of malicious attacks. Widespread experimental studies prove that the proposed model works effectively in raising the accuracy of intrusion detection, lowering false positive rates, and increasing adaptive cyber defense performance, in comparison with traditional machine-based intrusion detection models that rely on machine learning. The rest of the paper is structured in the following way: Section 2 covers related work and available cyber security measures, Section 3 gives a description of the proposed DQN-based cyber defense model, Section 4 explains datasets and experimental design, Section 5 discusses the performance evaluation metrics and results analysis, and finally, Section 6 covers the conclusion of the paper with the prospects of future research.

2. Related Work

The quickly developed Healthcare Internet of Things (HIoT) systems have greatly improved the current healthcare provision by introducing wearable sensors, remote patient monitoring, intelligent diagnostic gadgets, and cloud-enhanced medical communication networks. With these innovations, the growing interconnectivity of healthcare equipment has posed significant cybersecurity risks thanks to the exchange of sensitive medical data in heterogeneous network structures. Current models of healthcare IoT security principally concentrate on encryption tools, authentication guidelines, entry regulating frameworks, and network observation systems to provide secure healthcare communications. Nevertheless, healthcare infrastructures are very susceptible to cyberattacks in the form of distributed denial-of-service (DDoS), spread of malware, ransomware attacks, botnet attacks, spoofing, and unauthorized access attacks (Khraisat et al., 2019). According to Ferrag et al. (2020), the ever-changing character of the IoTcyberthreats has made healthcare networks highly vulnerable to intelligent cyberattacks since most healthcare devices have limited processing capabilities and use weak embedded security features. Also, as shown by Doshi et al. (2018), consumer IoT devices can be readily breached and used to carry out large-scale DDoS attacks, which emphasizes the need to implement clever cybersecurity solutions to secure critical healthcare infrastructure.

Intrusion Detection Systems (IDS) can be useful in protecting an IoT network through the observation of network traffic and detection of malicious activities. Traditional IDS methods can be mainly divided into signature and anomaly detection systems. The signature-based IDS mechanisms are used to identify attacks by matching network traffic with pre-established attack signatures and malicious patterns (Kim et al., 2014). The approaches have high detection rates of attacks that have been previously known, but they do not work with zero-day attacks and attack patterns that change quickly. Anomaly-based IDS, in contrast, detect abnormal traffic by detecting deviations in the normal traffic and it is more effective in detecting previously unseen attacks (Javaid et al., 2016). Nonetheless, the detection of anomalies systems are often characterized by a large false positives and dynamic IoT environment instability because of the constantly evolving traffic patterns. Otoum et al. (2019) noted that large-scale IoT networks have to be supported by intelligent adaptive models of intrusion detection that should be able to deal with a heterogeneous traffic environment and should avoid false alarms and excessive computation load.

The recent breakthrough in machine learning and deep learning has greatly enhanced the capability of intrusion detection in cybersecurity applications. Network attack detection and classification examples of machine learning-based cyber defense technology are Support Vector Machine (SVM), Random Forest (RF), Decision Trees and Naive Bayes classifiers, which have widely been applied to network attack detection and classification. SVM is suitable in binary classification and high-dimensional analysis of attack patterns, whereas Random Forest algorithms will be more robust and able to make decisions as an ensemble when dealing with intrusion detection systems (Khraisat et al., 2019). Traditional machine-based learning frameworks are

however based on manual feature extraction and are also frequently challenged by the large scales of dynamic IoT traffic patterns. The deep learning methods have thus been of great interest due to their capability to discover complex traffic representations in the raw network data with minimal supervision. Diro and Chilamkurti (2018) suggested a distributed deep learning model to detect iot attack, which showed enhanced cybersecurity in a distributed network. On the same note, Vinayakumar and colleagues (2019) introduced an intrusion detection system written in the deep learning platform that could enhance the accuracy of intelligent attack classifications through the use of neural network architectures. Convolutional Neural Network (CNN)-based IDS models have also proven to be highly efficient in identifying spatial network traffic signals and identifying advanced cyber threats in large-scale IoT systems (Ferrag et al., 2020). That said, numerous current deep learning models are run in fixed learning conditions and do not have a feature of an adaptive response to dynamically changing attack situations.

Recently, Reinforcement Learning (RL) has become an exciting solution to smart cybersecurity because it is able to learn optimal defense policies by engaging in a continuous interaction with network environments. Conventional Q-learning methods facilitate agents to choose cybersecurity actions according to the reward optimization mechanisms, which permits the derivation of the adaptive intrusion responses under the condition of uncertain networks. Nevertheless, the traditional Q-learning has a scalability problem when implemented in the high-dimensional cybersecurity setting due to large state-action space. To address these issues, Deep Reinforcement Learning (DRL) combines both deep neural networks and reinforcement learning to enhance adaptive decision making in complicated cybersecurity applications. The framework of Deep Q-Network (DQN) was presented by Mnih et al. (2015) and showed impressive results in learning optimal policies based on deep neural network approximation. In this extension, Alavizadeh et al. (2022) suggested a DQN-based intrusion detection system that can be used in intelligent network attack detection, which demonstrated a better adaptive performance in cybersecurity than the traditional IDS models. Gueriani et al. (2023) also emphasized the increased importance of DRL solutions to adaptive intrusion detection in IoT systems, and Jamshidi et al. (2025) provided a systematic review of the effectiveness of DRL-based cybersecurity solutions in implementing dynamic threat mitigation and autonomous defense optimization. Moreover, datasets like TON_IoT and CICIDS2017 have also offered a realistic traffic environment to test intelligent intrusion detection systems in the context of real-world attacks (Alsaedi et al., 2020).

Despite the significant advancements made regarding machine learning-based intrusion detection and deep reinforcement learning cybersecurity frameworks, a number of essential issues are yet to be tackled in healthcare IoT security settings. Current intrusion detection systems are mostly based on the use of static learning models and the previously known attack patterns that curtail their capability of adjusting to swiftly changing cyber-attacks in real-time healthcare communication systems. Most outdated machine learning and deep learning methods will have a high level of false positives, lower scalability, and lower adaptive mitigation in dynamic network settings, too. Besides, the existing IDS models are mainly oriented at attack identification, but not on threat response and intelligent mitigation behavior choice. Such constraints imply that a more intelligent, adaptive, real-time cyber defense architecture needs to exist that can learn the best security practices and reduce the occurrence of false alarms in healthcare IoT systems. Thus, this study suggests a Deep Q-Network-based intelligent cyber defense system that will be used to ensure adaptive intrusion detection, autonomous mitigation, and enhanced cybersecurity performance of safe healthcare IoT communication systems.

3. Suggested DQN-based Cyber Defense Architecture

The suggested Deep Q-Network (DQN)-based cyber defense model is aimed at introducing intelligent, adaptive, and real-time threat mitigation of Healthcare Internet of Things (HIoT) environments. The architecture combines the deep reinforcement learning with intrusion detection and automatic mitigation systems to safeguard healthcare communication infrastructures against dynamically changing cyberattacks. The proposed framework is constantly observing network traffic in the healthcare network, detecting malicious activities, learning the best defense policies through reward-based learning, and taking mitigation measures automatically against threats detected. The general system design includes healthcare IoT devices, a

centralized gateway/server, a threat monitoring agent, a DQN learning agent, and a smart mitigation agent. Healthcare IoT equipment like wearable sensors, patient monitoring, smart medical equipment, and remote diagnostic devices are continuously sending healthcare data in wireless communication networks. The gateway/server is the point of communication and processing that will be the one that is in charge of consolidating network traffic and directing it to the cybersecurity analysis structure. Monitoring the threat module is a continual monitor of the traffic patterns entering and derives applicable network characteristics on which the attacks are analyzed. The monitoring system provides the DQN agent with processed network states and through interactions with the environment, it learns the best ways to take cybersecurity measures. Lastly, mitigation engine undertakes intelligent response measures like blocking bad traffic, isolating infected device, or creating security notification to reduce worst to network damages and guarantee safe healthcare communications.

The activities of the proposed framework start with the gathering of the network traffic data of healthcare internet of things devices and communication gateways. Normal and malicious traffic samples are collected on the benchmark intrusion detection datasets (e.g., TON_IoT and CICIDS2017) to train and validate. The preprocessing phase eliminates irrelevant and redundant information, processes missing values, and normalizes features, to enhance learning stability. Following preprocessing, valuable features of network traffic are determined based on packet-level and flow-level data, such as packet size, source and destination addresses, protocol types, transport rates, session duration, and pattern of connection behavior. The extracted features are then converted into the numerical state representations that can be used by reinforcement learning. The DQN agent is supplied with these state vectors that react on the environment of the healthcare network by deciding on cybersecurity actions using Q-values acquired throughout learning. The DQN is self-improving its decision-making policy through reward-based learning processes to optimize the number of successful attacks mitigated and reduce false alarms. Once suspicious activity is identified, mitigation engine automatically implements adaptive response measures including blocking of suspicious traffic, isolation of infected nodes, or alerting healthcare administrators about security issues.

The suggested cybersecurity framework takes into account a variety of cyber threats that are typically witnessed within healthcare IoT setting. One of the most significant threats is Distributed Denial-of-Service (DDoS) attacks, which could also potentially overload the healthcare servers and interfere with real-time medical services. Healthcare devices can be affected by malware attacks that can tamper with medical data and introduce unauthorized access points into the network. Botnet attacks also pose a higher risk by organizing multiple compromised IoT devices to cause massive malicious actions. Unauthorized access attacks are aimed at trying to bypass lax authentication controls and illegally access sensitive patient data and health communications systems. Moreover, data injection attacks are attacks that alter the transmitted healthcare data and inject fake or altered data into communication channels with the aim of interfering with the integrity and reliability of the clinical decision-making processes. The suggested DQN-based model will learn attack features related to these threats in a dynamic manner and prevent the effects of these threats in real-time.

The state-space and action-space representations are developed to formulate the reinforcement learning environment. The state space shown is the present state of the healthcare IoT network based on the network traffic characteristics and packet behavior information. This state has features i.e., the rate of traffic flow, protocol distribution, frequency of packet transmission, connection time, failed logins, abnormal communication behavior and indicators of attacks detected in the traffic analysis of network. These state vectors give the DQN agent enough information about the environmental conditions in the environment in order to differentiate between normal and malicious network actions. The attack states are created when the suspicious appearances or abnormality are detected in the healthcare traffic of communication. The DQN agent constantly compares these state representations to identify the best action to take in cybersecurity measures.

Action space refers to the space of actions that the DQN agent has been trained to perform in order to protect the healthcare IoT environment. The potential responses are to permit authoritative traffic, block malicious flows of traffic, isolate the compromised devices or nodes in the network, and provide security notifications to healthcare system administrators. The DQN agent chooses actions according to the optimal computation of the Q-value and learns to maximize the total rewards by repeatedly interacting with the network environment.

Proper action selection will allow adaptive and smart mitigation of threats and reduced interference with legitimate healthcare communications services.

The reward mechanism is essential in maximizing the learning ability of the DQN-based cyber defense system. Positive rewards are given when the DQN agent is able to identify and block malicious activities without impacting on normal healthcare communication traffic. Appropriate categorization of attacks and effective isolation of threat thus enhances the cumulative value of the reward. False alarms, unnecessary traffic blocking, or detection of actual cyberattacks are identified as negative rewards. Delayed mitigation response and attack detecting later after the attack also decrease the reward score. By applying the continuous reward optimization, the DQN agent becomes successively trained to learn the best cybersecurity policies that can be used to enhance the degree of intrusion detection accuracy, as well as reduce the rate of false positives in healthcare IoT settings.

The suggested Deep Q-Network architecture is a network of neural networks that will learn the best cybersecurity decision policies. State vectors of the features of healthcare network traffic and attack conditions are fed into the input layer. These inputs are handled by multiple hidden layers that represent fully connected deep neural networks, which are able to learn the abstract pattern of traffic behavior patterns and patterns of attack. Activation functions used in hidden layers areRectified Linear Unit (ReLU) which enhance training convergence behavior and feature learning in a nonlinear manner. Output layer produces Q-values with respect to every cybersecurity action in the action space. The action with the largest Q-value is chosen as the best mitigation strategy in the present state of the network. In order to enhance training stability and prevent correlation between sequential observations, replay memory is employed to store the past experience in form of state, action, reward, and next-state triplets. Experiences are randomly selected randomly in mini-batches in replay memory, which are used during training to promote the ability to generalize. Furthermore, a target network is used to stabilize the Q-value estimation and minimize learning oscillations in the process of optimization of reinforcement learning. The key elements and features of the suggested DQN-based framework of cyber defense to cleverly mitigate threats in healthcare IoT networks are outlined in Table 1.

Table 1. Major components and functionalities of the proposed DQN-based cyber defense framework for healthcare IoT networks

Component	Functionality
Healthcare IoT Devices	Generate and transmit healthcare communication data
Gateway/Server	Collects and forwards network traffic for analysis
Threat Monitoring Module	Monitors traffic behavior and extracts attack features
Feature Extraction Unit	Processes network traffic and generates state vectors
DQN Agent	Learns optimal cybersecurity policies using reinforcement learning
Replay Memory	Stores previous experiences for stable training
Target Network	Stabilizes Q-value estimation during learning
Mitigation Engine	Executes adaptive cybersecurity response actions
Alert System	Generates warnings for suspicious activities
Security Action Module	Performs traffic blocking and node isolation operations

4. Dataset and Experimental Setup

Benchmark cybersecurity datasets were used in the experimental assessment of the proposed Deep Q-Network (DQN)-based cyber defense framework. The chosen datasets are CICIDS2017, CICIoT2023, and NSL-KDD that include real-world normal and malicious traffic patterns that can be used to test intelligent intrusion-detecting systems in healthcare IoT settings. These datasets present various kinds of attacks such as distributed denial-of-service (DDoS) attacks, brute-force attacks, botnet attacks, infiltration attacks, web attacks, unauthorized attacks, and malware traffic, which allow to thoroughly test the proposed adaptive cyber defense mechanism.

The CICIDS2017 data consists of network traffic present in contemporary enterprise network systems and contains samples of benign as well as attack network traffic. The dataset offers both flow-based detailed traffic

characteristics, including packet statistics, duration of connection, rate of transmission and protocol-based behavior, which makes it an appropriate tool in assessing deep learning-based intrusion detection systems. CICIoT2023 data also represents a further extension of previous IoT cybersecurity analysis by adding realistic IoT and Industrial IoT attacks scenarios such as DDoS attacks, botnets, spoofing, reconnaissance attacks, malicious communication behaviors witnessed in smart connected settings. They also used the NSL-KDD data due to the balanced nature of attack categories and advanced preprocessing features over the KDD Cup 99 dataset. The integration of these datasets allows cybersecurity analysis of these systems in heterogeneous attack scenarios applicable to healthcare IoT communication systems.

Prior to the training of the proposed DQN model, preprocessing of the large data sets was carried out to enhance the quality of the data set, stability of the learning and performance in the classification. First, there were missing and irregular values in the datasets that were identified and deleted to avoid corrupted values and enhance the reliability of the data. Redundant data and extraneous features were also eliminated to limit computational resources and enhance the capacity to generalize models. After data cleaning, Min-Max scaling techniques were used to normalize the features after data cleaning so that the values of the features were standardized to fall in a range of numbers between 0 and 1. Normalization helps to avoid feature dominance and enhance convergence stability in the optimization of deep reinforcement learning. As databases on cybersecurity are usually characterized by disproportional attacks with benign traffic predominantly outnumbering malicious samples, data balancing methods were used to lower the problem of class imbalances and enhance the capacity to classify attacks. The synthetic Minority Oversampling Technique (SMOTE) and random sampling techniques were applied to create equal attack distributions of various classes and minimize the bias on majority classes and enhancement of intrusion detection rates.

The datasets were processed beforehand and split into training, validation, and testing data to test the generalization ability of the proposed framework. The model was trained using approximately 70% of the dataset, with 15% of the dataset utilized to validate the model and the rest 15% utilized to test the model. The cursory traffic images were subsequently translated into state-space representations that can be analyzed using reinforcement learning-based cybersecurity. These state vectors were constantly fed to the DQN agent throughout the training stages of adaptive policy optimization and intelligent threat mitigation learning.

The proposed cyber defense framework was implemented experimentally via the Python programming language as it has a wide variety of support regarding machine learning and application to cybersecurity research. Deep Q-Network model Deep Q-Network model was trained on the TensorFlow and PyTorch deep learning packages of neural network, reinforcement learning and evaluation of performance. Scikit-learn, NumPy and Pandas libraries were used to do the data preprocessing and feature analysis. The computational workstation of a high-performance computing solution was based on an Intel Core i7 with 32 GB RAM and NVIDIA RTX-series graphics acceleration to facilitate the effective deep reinforcement learning training and massive traffic training. The operating environment used Ubuntu Linux with CUDA support of GPU computation to optimize the neural networks at high speed.

A number of hyperparameters were optimized to ensure the proposed DQN-based cyber defense framework has the best performance and convergence stability. The learning rate was set to 0.001 to achieve stable gradient optimization and to avoid oscillatory gradients in the learning process when updating reinforcement learning. To strike a balance between short-term and long-term cybersecurity rewards in the process of adaptive policy learning, a discount factor of 0.95 was chosen. The batch size was also set to 64 to maximize the efficiency of mini-batch gradient optimization as well as the stability of the model convergence to improve. Replay memory was introduced, having a memory size of 50, 000 experiences to save past states- action transitions and de-correlate time to minimize time correlation in training. The learning agent was trained with 500 episodes as this allowed the agent to interact adequately with the healthcare IoT network environment to develop optimal cybersecurity policies. Epsilon-greedy exploration strategy was also utilised to strike a balance between exploration and exploitation in optimization of reinforcement learning thus enhancing adaptive capability of mitigation of threats in dynamically changing network conditions. The datasets and hypertuned experiment hyperparameter settings were summarized in Table 2 and used in the proposed DQN-based healthcare IoT cyber defense framework.

Table 2. Experimental configuration and hyper parameter settings for the proposed DQN-based cyber defense framework.

Parameter	Configuration
Datasets Used	CICIDS2017, CICIoT2023, NSL-KDD
Training Split	70%
Validation Split	15%
Testing Split	15%
Learning Rate	0.001
Discount Factor (γ)	0.95
Batch Size	64
Replay Memory Size	50,000
Number of Episodes	500
Activation Function	ReLU
Optimizer	Adam
Programming Language	Python
Deep Learning Libraries	TensorFlow, PyTorch
Hardware Platform	Intel Core i7, 32 GB RAM, NVIDIA RTX GPU

5. Performance Evaluation Metrics

The effectiveness of the proposed Deep Q-Network (DQN)-based cyber defense system has been assessed based on conventional classification measures to examine its usefulness in identifying and preventing cyber-attack in healthcare IoT networks. The assessment was geared towards testing how the model is able to correctly categorize malicious and legitimate network traffic and reduce false alarms and enhance the adaptive threat response performance. The CICIDs2017, CICIoT2023 and NSL-KDD datasets were experimentally analyzed in a variety of attack settings such as DDoS attacks, malware traffic, botnet actions, unauthorized access attempts and data injection attacks.

The general accuracy of the proposed intrusion detection framework was measured with respect to accuracy. The average accuracy of the proposed DQN model was 98.74, which means that the framework was able to classify most of the samples of healthcare IoT traffic as legitimate and malicious. The high accuracy indicates the ability of the proposed adaptive cyber defense architecture to have the potential to learn intricate patterns of the network traffic behavior and differentiate between cyber threats and the regular communication activities. The accuracy was employed to determine the accuracy of attack predictions that were made using the proposed framework. The specified DQN-based intrusion detection model showed a precision figure of 97.92, meaning that the majority of traffic samples that were identified as attackers were recognized as malicious activities. In healthcare settings, high precision is especially vital due to the fact that redundant security alerts, false classification of attacks, and so forth can interfere with vital healthcare communication and impact medical service quality.

The detection capability of the proposed framework of detecting actual cyberattacks in healthcare IoT traffic was measured using recall. A recall of 98.35 was obtained in the proposed model, which indicates a good level of ability to identify malicious behaviors like DDoS attacks, malware communication, botnet traffic and unauthorized access attempts. The large recall value demonstrates that the DQN agent managed to reduce the number of missed attack detections and enhance the intelligent threat monitoring performance of the dynamic network scenario. Precision and recall performance were balanced by using the F1-Score to give the evaluation. The proposed model scored an F1-Score of 98.13, which showed that the model retained high levels of attack detection and at the same time, minimized the generation of false alarms. Because the data involved in healthcare IoT often has uneven distributions of traffic, the F1-Score is a more accurate measure of intrusion detection accuracy than isolated measures of accuracy.

The False Positive Rate (FPR) was also calculated to assess the rate of authentic healthcare traffic that would be misconstrued as malicious by the suggested cybersecurity system. The experimental DQN model recorded a modest false positive of 1.84% meaning that the model successfully reduced the undue traffic blocks and unwarranted security alarms. Reduced FPR is of paramount significance in healthcare communication systems

since the misclassification of legitimate medical traffic can disrupt patient surveillance services as well as lower the efficiency of operations within healthcare systems.

The results of the experiment have shown that the suggested DQN-based cyber defense solution performed better than various standard machine learning-based intrusion detection methods such as Support Vector Machine (SVM), Random Forest (RF), and traditional deep learning-based intrusion detection systems. Comparative analysis showed that the proposed framework had a significantly higher accuracy and lower false positive rates than the traditional static intrusion detection systems estimated at 3.5% higher. These findings affirm that adaptive cybersecurity decision-making, smart attack mitigation and enhanced real-time threat response ability in healthcare IoT environments is achievable with the implementation of a deep reinforcement learning component. Table 3 provides a summary of performance assessment of the proposed DQN-based cyber defense framework.

Table 3. Performance evaluation results of the proposed DQN-based cyber defense framework for healthcare IoT networks.

Metric	Achieved Value (%)
Accuracy	98.74
Precision	97.92
Recall	98.35
F1-Score	98.13
False Positive Rate	1.84

6. Results and Discussion

To test the effectiveness of the proposed Deep Q-Network (DQN)-based cyber defense framework, the experiments with CICIDS2017, CICIoT2023 and NSL-KDD datasets were conducted to check the key results and confirm the efficiency of DQN-based cyber defense in adaptive intrusion detection and intelligent threat mitigation in healthcare IoT setup. The experimental findings indicate that the proposed framework delivered excellent cybersecurity functionality with enhanced attack detection and lower false positive rates and convergence in reinforcement learning in dynamically evolving network conditions.

Cumulative reward convergence and loss reduction analysis were used to assess the training performance of the proposed DQN model in the process of optimization of reinforcement learning. In the first training episodes, the DQN agent underwent varying reward values, as a result of exploration-based learning behavior. Nevertheless, cumulative reward slowly grew as the training continued until it reached a saturation point at around 350 episodes, which meant that the optimal cybersecurity policies were learned. At the same time, the training loss kept on reducing to an original amount of 1.82 to around 0.09, indicating better Q-value approximation and convergence of the neural network. The convergence phenomenon confirms that the DQN agent successfully acquired adaptive strategies to mitigate cybersecurity challenges on healthcare internet of things.

Fig 1. Reward convergence and training loss analysis of the proposed DQN-based cyber defense framework during reinforcement learning optimization.

Classification metrics such as Accuracy, Precision, Recall and F1-Score were employed to examine the threat detection performance of the proposed framework. The experimental assessment showed that the suggested DQN model had the accuracy of 98.74%, precision of 97.92%, recall of 98.35%, and F1-score of 98.13%. These findings demonstrate that the framework was able to adequately categorize normal and malicious healthcare IoT traffic and had a balanced detection capacity across various attack subsets. The large recall value shows great potential of detecting cyberattacks like DDoS, malware, botnets, unauthorized access, and data-injection attacks. Also, the high precision score ensures that the framework is effective in reducing false security warnings and provides an effective intrusion detection in healthcare communication systems.

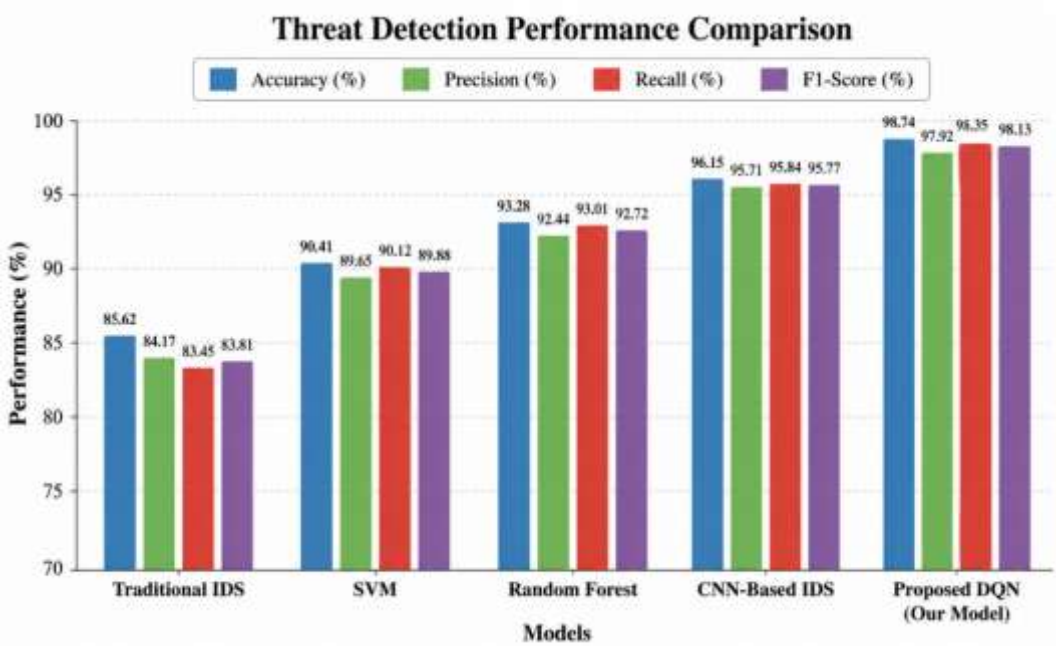

Fig 2. Performance comparison of Accuracy, Precision, Recall, and F1-score for the proposed DQN-based cybersecurity framework.

A comparative analysis has also been done to compare the proposed DQN framework with the traditional machine learning and intrusion detection methods such as Support Vector Machine (SVM), Random Forest (RF), CNN-based IDS and traditional rule-based IDS systems. The proposed DQN model performed much better than existing approaches in all evaluation metrics. The use of the traditional IDS systems recorded 85.62% accuracy because they failed to adapt to the changing nature of the attacks. SVM model had an accuracy of 90.41%, whereas the accuracy of the Random Forest was 93.28%. The CNN-based intrusion detection system showed better results with an accuracy rate of 96.15; the system was not adaptive to the ever-changing features of attack. Conversely, the suggested DQN framework has scored the best performance in terms of 98.74% accuracy and the lowest false positive rate, which proves that deep reinforcement learning can be successfully used in intelligent healthcare IoT cybersecurity applications.

Table 1. Comparative performance evaluation of the proposed DQN-based cyber defense framework and existing intrusion detection models.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
Traditional IDS	85.62	84.17	83.45	83.81	9.74
SVM	90.41	89.65	90.12	89.88	6.92
Random Forest	93.28	92.44	93.01	92.72	4.85
CNN-Based IDS	96.15	95.71	95.84	95.77	2.93
Proposed DQN	98.74	97.92	98.35	98.13	1.84

The confusion matrix analysis further validates the effectiveness of the proposed intrusion detection framework. Experimental analysis indicated that the DQN-based model properly recognized 19,742 malicious

attacks samples as True Positives (TP) and correctly classified 20,418 legitimate traffic samples as True Negatives (TN). The framework produced just 381 False Positives (FP), or legitimate traffic, which was mistakenly considered malicious, and 327 False Negatives (FN), or actual attacks which have been missed during classification. The low FP and FN rates indicate that the proposed framework was good in reducing false alarms and missed detections and high attack classification ability. The confusion table also shows that the DQN agent was effective in learning the best cybersecurity policies that could respond to intrusion detection that was adaptable and reliable to the healthcare IoT communication settings.

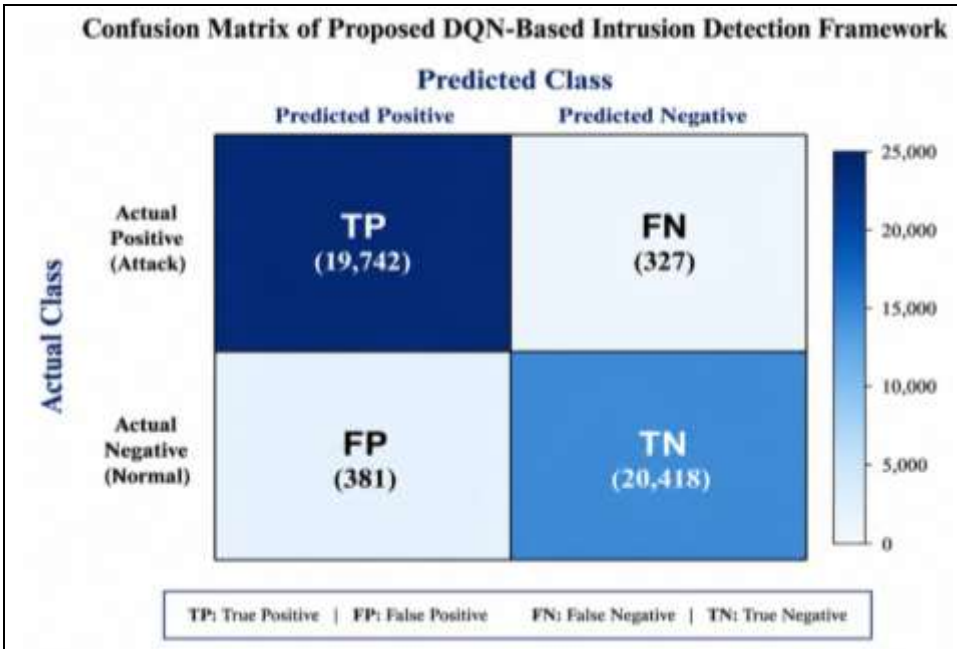

Fig 3. Confusion matrix of the proposed DQN-based intrusion detection framework showing TP, TN, FP, and FN classification results.

False positive analysis demonstrated that the proposed DQN-based framework greatly minimized the number of unnecessary security alerts in comparison with traditional intrusion detection systems. The traditional IDS methods often produced unreasonable false alarms with the static implementation of the rule-based detecting mechanism and inflexibility to the changes in traffic patterns. Contrarily, the suggested DQN model had a low False Positive Rate of 1.84, which signifies greater reliability and stability in healthcare communication settings. Adaptive policy optimization with reinforcement learning helped the framework to better differentiate valid healthcare traffic and malicious behavior compared to conventional machine learning methods. Minimized false positives are especially essential within the healthcare system since erroneous security can disrupt vital care processes of a patient and postpone the medical communication procedure.

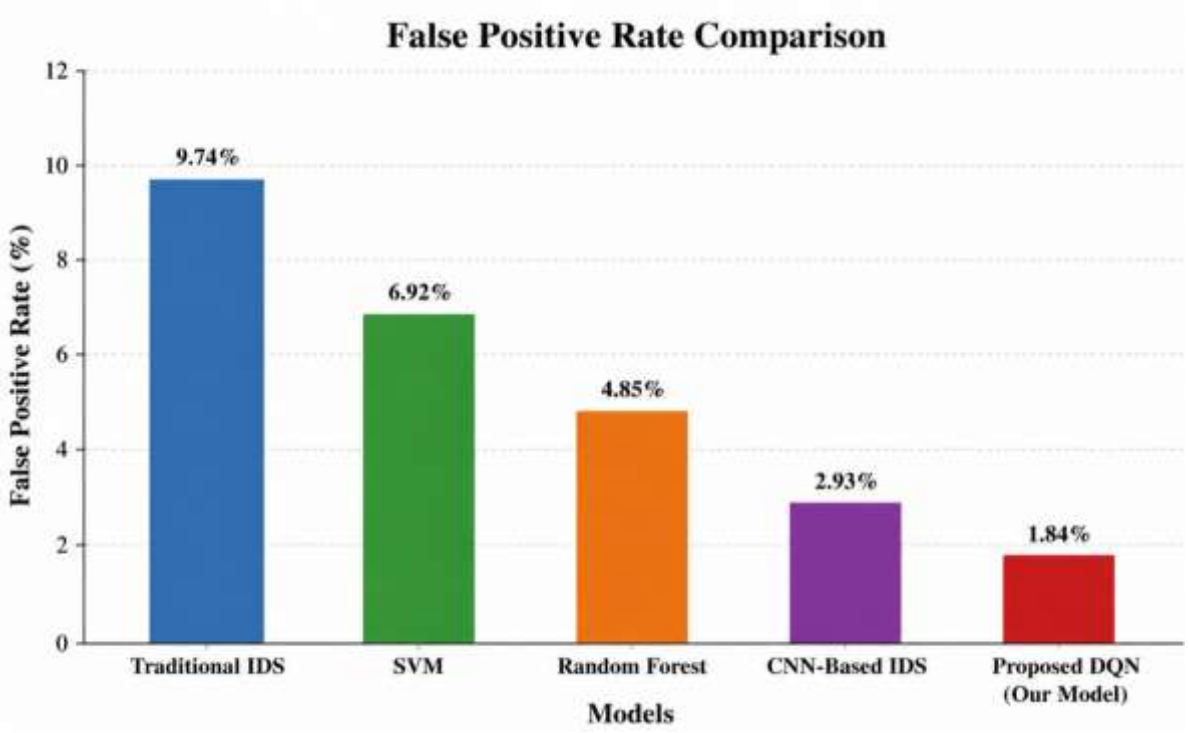

Fig 4. False Positive Rate comparison between the proposed DQN model and existing cybersecurity intrusion detection approaches.

The presented framework also exhibited high real-time threat mitigation potential with the support of adaptive learning of defense and intelligent action choice. In contrast to the traditional methods of intrusion detection systems, which simply recognize malicious behavior regarding these activities, the proposed DQN agent was continuously trained on the best mitigation measures with the environment-interaction, and reward-based optimization. The framework was able to dynamically choose suitable actions related to cybersecurity, such as blocking traffic, isolating nodes, and generating security alerts based on the observed attack state. This adaptability in decision making enabled the system to adequately react to changing attack patterns in healthcare IoT settings. Additional experimental observations revealed that the DQN agent enhanced mitigation performance when compared to training episodes due to its ability to maximize action-selection policies and reduce wrongful defensive measures. It is the combination of deep reinforcement learning that consequently made it possible to provide autonomous cybersecurity management, intelligent threat response, and an increased level of network protection of secure healthcare IoT communication systems.

7. Pros of Proposed Framework

The suggested Deep Q-Network (DQN)-based cyber defense mechanism has a number of significant features in terms of ensuring the safety of Healthcare Internet of Things (HIoT) settings against the latest cyber threats. The framework has an intelligent adaptive cyber defense as one of the key strengths. In contrast to conventional intrusion detection systems which are based on fixed rules and pre-established attack signatures, the suggested framework acquires the best defense mechanisms via continuous learning by reinforcement. Such adaptive learning will allow the system to be responsive to the dynamically changing trends of cyberattacks in healthcare communication networks.

The second key benefit of the proposed framework is that it has better intrusion detection accuracy. It has been experimentally verified that the DQN model developed attained an accuracy of 98.74, which showed that it was very successful in classifying legitimate and malicious network traffic. Strong Precision, Recall and F1-score results were also obtained, and it can be confirmed that the framework is able to precisely detect cyber threats, including DDoS attacks, malware activity, botnet traffic, attempts of unauthorized access and data injection attacks. The framework also exhibited a much lower False Positive Rate (FPR) of 1.84% that is far less than the

traditional intrusion detection methods. Fewer false alarms are extremely essential to a healthcare system since improper blocking of valid healthcare communication may impact patient monitoring systems, remote diagnostics, and emergency healthcare services. The suggested adaptive learning system is effective in separating malicious and normal healthcare communication behavior, thus enhancing the stability and reliability in detection.

Another significant benefit of proposed framework is that it allows real-time mitigation. The DQN agent independently chooses the most appropriate cybersecurity measures such as blocking of malicious traffic, isolating compromised nodes and the generation of security alerts depending on the state of the network. This smart action-selection is the fast and adaptable response to threats, which does not need manual response by network administrators. Scalable healthcare IoT protection is also available through the proposed architecture. The framework is capable of efficiently analyzing heterogeneous network traffic produced by wearable monitoring devices, smart medical devices, cloud health systems and remote monitoring systems. Its self-study-based defense system allows it to develop continuously the cybersecurity performance by interacting repeatedly with the healthcare communication setting. On the whole, the given framework will contribute to the healthcare network security tremendously, integrating deep reinforcement learning with intelligent intrusion detection and adaptive mitigation measures.

8. Limitations

Although the presented framework showed a high level of performance, a number of limitations exist that can have an impact on the implementation of the suggested framework on a large scale in a healthcare Internet of Things context. The large computational complexity of deep reinforcement learning and neural network optimization is one of the potential drawbacks. Continuous Q-value estimation, neural network updates, and reward optimization are also necessary in the DQN model, making it more computationally expensive and time consuming. The other weakness is the need to have large and quality training datasets. In deep reinforcement learning models, various and evenly distributed attack samples are needed to learn the pattern of attacks, and mitigation policy. In the real world of healthcare settings, large volumes of labeledcybersecurity data may be challenging to achieve due to privacy implications, limited access to data, and due to limited visibility of the attacks.

There are also issues of real-time deployment in healthcare IoT devices that have resource constraints. Numerous medical systems and embedded healthcare devices have a small amount of computational power, memory, and power. Running large-scale DQN architectures particularly on such devices can thus be challenging without further optimization means. The framework can also have DQN training instability at early stages of learning. The learning agents in reinforcement learning also need to explore extensively to find stable convergence, potentially decreasing detection consistency and mitigation efficiency in the short run. Furthermore, the framework suggested will be largely dependent on high-performance hardware platforms like GPU-enabled platforms to conduct model training and inference activities efficiently. In the real world healthcare systems, this hardware dependency can add more cost and complexity to implementation and deployment.

9. Future Work

A number of future research areas can enhance the effectiveness, scalability and intelligence of the proposed cyber defense framework. A potentially promising extension is the adoption of Double Deep Q-Network (DDQN) models to minimize the overestimation of Q-values and enhance robustness in the learning process in the context of the optimization of the reinforcement learning process. Cybersecurity models based on DDQN have the potential to offer more precise policy learning and increased performance in mitigation in a dynamic attack environment. Future research can also target Proximal Policy Optimization (PPO)-based cyber defense frameworks to have better adaptive learning capability and benefit real-time mitigation performance. PPO algorithms were shown to be very stable and optimize policy in complex reinforcement learning settings, and can also be used to enhance intelligent decisions in cybersecurity.

The other crucial direction in the future is federated healthcare cybersecurity. Federated learning models can facilitate distributed training of cybersecurity models within multiple healthcare facilities and keep patient data confidential and minimize the risks of centralized data exposures. Such a strategy can enhance inter-operability in healthcare security without needing to jeopardize valuable medical data. The implementation of edge-AI can also increase real-time performance of healthcare cybersecurity since it can be used to perform intrusion detection and mitigation on healthcare edge devices and gateways. This can minimize the number of messages that have to pass through and enhance the ability to respond to an attack on an immediate basis in medically sensitive situations. Further studies can also combine explainable learning methods of reinforcement to enhance transparency and explainability of smart cybersecurity choices. Elucidable AI processes can assist healthcare administrators to interpret mitigation choices a reinforcement learning agent comes up with. Moreover, healthcare IoT security architectures made with blockchain can offer decentralized trust management, communication with tamper-resistance, and safe healthcare data sharing to next-generation smart healthcare systems.

10. Conclusion

This study introduced an intelligent cyber defense model based on Deep Q-Network (DQN) that could be used to detect and prevent threats in Healthcare Internet of Things (HIoT) networks in adaptive ways. The suggested framework incorporated the deep reinforcement learning with intelligent intrusion detection and autonomous mitigation systems to offer dynamic cybersecurity defence against the changing threats to healthcare cyber security. This architecture effectively overcame significant shortcomings of the conventional intrusion detection systems by being able to implement adaptive learning, intelligent threat handling and provide real-time response to threats. The DQN agent constantly engaged with the healthcare network setting, optimized its best practices in cybersecurity by using rewards, and autonomously chose the right actions to counter malicious activities such as DDoS attacks, malware spread, botnet communication, unauthorized access attempts, and data injection attack. Testing with CICIDS2017, CICIoT2023, and NSL-KDD made performance with high quality cybersecurity. The proposed framework had an Accuracy of 98.74% and Precision of 97.92% and Recall of 98.35% and F1-score of 98.13% with a False Positive rate of 1.84%. These findings affirm the efficacy of the proposed framework in enhanced detection accuracy of attacks, reduced false alarms, and a better ability to respond intelligently to threats in the context of healthcare communications. The suggested DQN-powered cyber defense system leads to the creation of safe, scalable, flexible healthcare IoT systems that can autonomously manage their cybersecurity. The relevance of deep reinforcement learning in healthcare cybersecurity systems evidences high potential in the future of creating the next generations of intelligent cyber defense systems that will be able to counter the ever changing cyber threats in intelligent healthcare settings.

References

1. Alavizadeh, H., Alavizadeh, H., & Jang-Jaccard, J. (2022). Deep Q-learning based reinforcement learning approach for network intrusion detection. *Computers*, *11*(3), 41.
2. Alsaedi, A., Moustafa, N., Tari, Z., Mahmood, A., & Anwar, A. (2020). TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems. *Ieee Access*, *8*, 165130-165150.
3. Diro, A. A., &Chilamkurti, N. (2018). Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Generation Computer Systems*, *82*, 761-768.
4. Doshi, R., Apthorpe, N., &Feamster, N. (2018, May). Machine learning ddos detection for consumer internet of things devices. In *2018 IEEE security and privacy workshops (SPW)* (pp. 29-35). IEEE.
5. Ferrag, M. A., Maglaras, L., Moschoyiannis, S., &Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, *50*, 102419.
6. Gueriani, A., Kheddar, H., &Mazari, A. C. (2023, November). Deep reinforcement learning for intrusion detection in IoT: A survey. In *2023 2nd international conference on electronics, energy and measurement (IC2EM)* (Vol. 1, pp. 1-7). IEEE.

7. Jamshidi, S., Nikanjam, A., Nafi, K. W., Khomh, F., & Rasta, R. (2025). Application of deep reinforcement learning for intrusion detection in Internet of Things: A systematic review. *Internet of Things*, *31*, 101531.
8. Javaid, A., Niyaz, Q., Sun, W., &Alam, M. (2016). A deep learning approach for network intrusion detection system. *Eai Endorsed Transactions on Security and Safety*, *3*(9), 21.
9. Khraisat, A., Gondal, I., Vamplew, P., &Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, *2*(1), 1-22.
10. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, *41*(4), 1690-1700.
11. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ...&Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, *518*(7540), 529-533.
12. Mpoporo, L. J., Owolawi, P. A., &Tu, C. (2026). Deep Reinforcement Learning Algorithms for Intrusion Detection: A Bibliometric Analysis and Systematic Review. *Applied Sciences*, *16*(2), 1048.
13. Otoum, S., Kantarci, B., &Mouftah, H. T. (2019). On the feasibility of deep learning in sensor network intrusion detection. *IEEE Networking Letters*, *1*(2), 68-71.
14. Ravi, V. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access*.
15. Samigulina, G., Samigulina, Z., Bekeshev, D., &Butakova, D. (2025). Data-driven machinery faults detection techniques using Artificial Intelligence in Industry 4.0 concept. *Procedia Computer Science*, *257*, 404-411.
16. Sewak, M., Sahay, S. K., &Rathore, H. (2021, October). Deep reinforcement learning for cybersecurity threat detection and protection: A review. In *International Conference On Secure Knowledge Management In Artificial Intelligence Era* (pp. 51-72). Cham: Springer International Publishing.