Contents lists available at *Dergipark*

Journal of Scientific Reports-A

journal homepage: *https://dergipark.org.tr/tr/pub/jsr-a*

E-ISSN: 2687-6167 ***Number 59, December 2024***

RESEARCH ARTICLE

Receive Date: 20.11.2024 *Accepted Date: 27.12.2024*

Reinforcement learning for energy optimization in IoT based landslide early warning systems

Seyfullah Arslan[a,]*, Safa Dörterler[b], Fırat Aydemir[c]

[a]*Kutahya Dumlupınar University, Faculty of Engineering, Computer Engineering Dept., 43100, Kütahya, Türkiye*
ORCID: 0000-0002-2573-273X
[b]*Kutahya Dumlupınar University, Faculty of Engineering, Computer Engineering Dept., 43100, Kütahya, Türkiye*
ORCID: 0000-0001-8778-081X
[c]*Kutahya Dumlupınar University, Faculty of Engineering, Computer Engineering Dept., 43100, Kütahya, Türkiye*
ORCID: 0000-0002-8965-1429

Abstract

This study introduces a novel energy management model based on Deep Reinforcement Learning for IoT-based landslide early warning systems, aiming to achieve energy neutrality and enhance system resilience, efficiency, and sustainability. Unlike traditional energy optimization methods, the proposed model employs a Deep Q-Network (DQN) to dynamically optimize the duty cycle of sensor nodes by leveraging real-time energy availability. By adaptively balancing energy harvesting and consumption, sensor nodes can maintain continuous operation even under highly variable environmental conditions, maximizing their performance during high-energy periods while preserving battery life when energy is limited. Extensive simulations using real-world solar radiation data demonstrate the model's superior capability in extending system longevity and operational stability compared to existing approaches. Addressing critical energy management challenges in landslide monitoring systems, this work enhances system reliability, scalability, and adaptability, offering a robust foundation for broader IoT applications deployed in energy-limited and dynamic environments. The proposed method represents a significant improvement over conventional techniques, as it autonomously optimizes energy resources to ensure the continuous and sustainable operation of IoT ecosystems.

* Corresponding author. Tel.: +90 (274) 443 42 68; fax: +90 (274) 265 20 13 - 14.
E-mail address: seyfullah.arslan@dpu.edu.tr
http://dx.doi.org/10.1016/j.cviu.2017.00.000

.

Keywords: Reinforcement learning; internet of things; energy management optimization; landslide early warning systems

1. Introduction

Landslides are one of the serious natural disasters that cause significant loss of life and property worldwide. The 2020 Gjerdrum landslide in Norway led to fatalities, the displacement of dozens of families, and millions of dollars in economic loss [1]. In regions such as South Asia, landslides frequently result in severe casualties and infrastructure destruction, particularly during heavy monsoon rains, further compounding the economic and social impact [2]. These disasters occur in mountainous and sloping areas and are triggered by sudden rainfall, groundwater movements, and weak soil structures [3]. Dense settlements and infrastructure projects in the regions where these disasters occur further increase the impact of landslides. Early warning systems have been developed to prevent loss of life and economic damages caused by landslides [4]. These systems identify the risk of landslides in advance, providing the opportunity to evacuate people living in affected areas to safe areas. However, the reliability of these systems depends on sensor networks that can provide accurate and continuous data [5]. Sensors used in landslide early warning systems must be capable of long-term operation, especially under harsh environmental conditions and energy constraints [6]. Therefore, energy management of sensors poses a significant challenge.

In recent years, Internet of Things (IoT)-based sensor networks have been employed extensively in landslide early warning systems [7]. However, most of the sensors in these networks have limited battery life, which has increased the importance of energy harvesting techniques [8]. Energy harvesting sensors can sustain long-term operation by collecting energy from solar, wind, and environmental vibrations [9]. However, the irregular and unpredictable dynamics of energy harvesting are one of the main factors limiting the effectiveness of such systems. Energy-harvesting sensors must efficiently utilize the harvested energy and sustain their duty cycles without power outages [10]. However, intelligent energy management algorithms are needed to optimize energy consumption.

Reinforcement Learning (RL) has significant potential for application in the field of energy management and duty cycle optimization. The RL algorithms can potentially enable energy-neutral operation (ENO) through autonomous optimization of sensor nodes' energy harvesting and consumption processes [11]. In particular, Deep Reinforcement Learning (DRL) algorithms facilitate the expeditious adaptation of sensors to environmental alterations and enhance energy efficiency [12]. The RL algorithms utilized in these studies guarantee uninterrupted system operation by establishing a balance between the energy collection and consumption processes [13].

The objective of this paper is to present a novel energy management model for IoT-based landslide early warning systems. The proposed model uses a reinforcement learning algorithm to optimize the battery levels of solar-powered sensor nodes and increase their continuous operation capability. This study introduces a system to enhance the energy efficiency of IoT nodes in landslide early warning systems. Furthermore, a Deep Q-Networks (DQN) based reinforcement learning algorithm is developed to optimize the energy consumption of nodes. This algorithm aims to achieve energy-neutral operation by balancing the energy collection and consumption of nodes.

Real-world solar radiation data is used to evaluate the performance of the developed system, and the energy harvesting process is simulated. Furthermore, the impact of the system on extending the battery life and improving the reliability of landslide early warning systems is analyzed. Simulation results show that the proposed model improves energy efficiency and ensures the uninterrupted operation of sensors.

The main contributions of this study are as follows:

- In early warning systems of natural disasters such as landslides, IoT nodes are provided to operate continuously without energy interruption.

- The proposed model addresses a significant gap in achieving energy-neutral operation under dynamic and unpredictable environmental conditions, which has been a critical limitation in prior studies.

- Energy consumption and harvesting processes are dynamically balanced using a reinforcement learning algorithm, specifically a Deep Q-Network (DQN), to adaptively respond to real-time variations in energy availability.

- Unlike conventional approaches, this model ensures uninterrupted operation of IoT nodes by preventing overcharging or depleting battery levels, which enhances both system longevity and reliability.

- Simulations with real-world solar radiation data validate the model's ability to self-adapt, demonstrating superior performance in extending battery life and maintaining operational stability compared to existing techniques.

- The system introduces an autonomous and scalable solution to energy management challenges in harsh and energy-limited environments, providing a robust framework for broader IoT applications beyond landslide monitoring.

- Using reinforcement learning, the model allows IoT nodes to intelligently optimize their duty cycles, enhancing energy efficiency without manual intervention.

- The proposed approach lays a foundation for future advancements in reinforcement learning-based energy management systems, addressing key challenges such as energy variability, scalability, and operational sustainability.

In conclusion, this study addresses energy management challenges in energy-harvesting IoT-based landslide early warning systems and introduces an approach to enhance their effectiveness. By optimizing energy harvesting and consumption processes, the model ensures uninterrupted sensor operation and supports the development of reliable early warning systems for natural disasters like landslides.

This paper consists of 5 sections in total. The second section, "Related Work," provides a comprehensive review of previous work on landslide early warning systems and energy harvesting technologies. The third section, "Materials and Methods," presents details of the proposed reinforcement learning-based energy management model. This section describes the methods developed to improve energy efficiency for IoT nodes and the model's training process. The fourth section, "Results and Discussion," discusses the simulation results and the performance of the model. The last section, "Conclusion," contains the study's overall conclusions and recommendations for future research.

2. Related work

Landslide early warning systems are highly essential in curbing devastating effects of natural disasters on human lives. These systems give early warning of future events of landslides, hence prompting the relocation of people in landslide-prone areas. Events of landslides are caused by various natural factors: excessive rainfall, soil moisture, fault movement, and rockfalls [14]. Accurate forecasting of these events will involve selecting the right sensors, efficient data processing, and optimization of decision-making processes. The research area of the landslide forecasting has witnessed a remarkable development in a number of techniques and methodologies over the last ten years. Some of these techniques include machine learning, artificial neural networks, data mining, geographic information systems, numerical modeling, and hybrid approaches.

Machine learning has been successfully used to solve complex problems in areas such as optimization [15-17], data classification [18-20], object recognition [21], [22], medical diagnosis [23], [24], educational technologies [22], voice analysis [25], [26], image analysis [27], [28], and decision support systems [29-31]. Recently, energy optimization and efficient resource management in wireless sensor networks (WSNs) and Internet-of-Things (IoT) systems have gained increasing attention to enhance network lifetime and reduce energy consumption. Ali et al. propose a novel ARSH-FATI-based Cluster Head Selection (ARSH-FATI-CHS) algorithm integrated with ranked-based clustering to address the energy efficiency challenges in WSNs [32]. Their approach dynamically balances exploration and exploitation during the cluster head selection process, taking into account residual energy, communication distance, and workload parameters. Simulation results demonstrate that the ARSH-FATI-CHS algorithm outperforms traditional techniques, such as particle swarm optimization (PSO), achieving approximately 25% improvement in network lifetime. This study highlights the importance of advanced metaheuristic algorithms in reducing communication energy and prolonging the lifespan of sensor nodes in WSNs, which are fundamental components in IoT ecosystems. In addition, Ali et al. provide a comprehensive survey on energy optimization techniques for Multiprocessor System-on-Chip (MPSoC) architectures in IoT and consumer electronics systems [33]. They emphasize the critical role of workload mapping and scheduling approaches, such as Dynamic Voltage and Frequency

Scaling (DVFS), Dynamic Power Management (DPM), and inter-processor communication reduction, in achieving substantial energy savings. The study also underscores the significance of combining coarse-grained software pipelining techniques with DVFS to improve performance and energy efficiency. By evaluating the integration of MPSoCs in IoT systems, the authors offer insightful directions for future research, particularly in balancing computational performance and energy consumption.

Researchers have used various machine learning techniques, including supervised and unsupervised learning, to analyze extensive datasets from areas with high landslide risk. Artificial neural networks have achieved success in landslide risk prediction by using parameters such as soil moisture, rainfall, slope, and soil type. Thirugnanam et al. developed a landslide early warning system [34]. They divided early warning systems into six phases which are data collection, transmission, modeling, analysis, prediction, and warning. In the data collection and transmission phase, they used various rain, humidity, and slope meter sensors to collect their data in a center with the developed IoT system. These data from the field are analyzed in real time, and landslides are successfully predicted with the developed machine learning algorithm. In addition, the researchers took into account that the data transmission stops and the system does not work correctly in case of various IoT sensors not working, malfunctioning, battery depletion, etc. In such cases, they predicted the data that could not come from the sensors with machine learning algorithms and ensured the security of the system. In another study, Hemalatha et al. developed a landslide early warning system by collecting the parameters affecting landslides in real time with rainfall, humidity, water pressure, and motion sensors [35]. They trained their model with current-PWP and 24-PWP algorithms based on the Support Vector Regression method. The continuity of the warning system depends on the parameters coming from the sensors and the system sensors operate 24/7 depending on electricity. In case of a power outage, sensors powered by solar energy ensure the continuity of the system. They also used machine learning algorithms to predict sensor values as a second measure. Collini et al. used various machine learning methods by analyzing 341 landslide events in Florence, Italy, and obtained the most successful results with the XGBoost model [36]. In another study, Ng et al. developed landslide prediction models using meteorological and topographic data collected from landslide-free and prone areas [37]. Their results turned out to be more successful when using the random forest model. Data collection and analysis are one of the key components in this developed system that uses machine learning methods. A commodity of a large amount of historical and real-time data is being modeled and made meaningful in this process to produce forecasts and determine early warning systems. These data have been used for the development of various models of forecasting using different machine learning and deep learning algorithms, such as Logistic regression, Support Vector Machine, Random Forest, and Convolutional Neural Network.

Another important technique that is normally used in the prediction of landslide risks includes data mining. Different machine learning algorithms analyze sensors or geographic information system data, such as soil type, measurement of moisture, GPS data, condition of weather, and variation in position for the prediction of risk from landslides. Franceschini et al. have generated a landslide inventory by applying methods of data mining and open-source websites with diverse geological and hydrological data in Italy [38]. Based on this aspect of the study, web pages revealed the use of predefined descriptive keywords in crawling web pages related to landslide events between 2010 and 2019, whereas those showing relevance were collected in an inventory. Subsequently, this was linked to several geographical databases that showed factors affecting landslides, such as population density, slope, rainfall, and geological structure. The relationships between these yielded a risk map related to areas prone to landslides and hence predicted future landslide events. As a result, the authors were then able to portray the usability of automated data mining methods in developing a landslide inventory and the usefulness of associating those methods with various geographical factors when conducting landslide risk analysis. Qian collected and analyzed data from many different sources and created a disaster emergency management system based on data mining [39]. They showed that this big data platform collected from various sources was successfully used as a disaster emergency management system and was more successful than traditional methods. Another study obtained images of 26 deep mountains and valleys in China. There were 146 landslide disasters in the locations where these images were obtained. Various machine learning algorithms classified these landslides according to landslide types and the results were compared with each other. As a result, solutions are presented to prevent landslide disasters and ensure the sustainability of rural industries by using remote sensing technologies and IoT sensors [40]. Pennington et al. introduced a global landslide event reporting tool using photo data from social media and various sources and artificial intelligence algorithms [41]. This tool detects landslide events by collecting data from social media platforms such as Twitter and using natural language processing and machine learning techniques. With the system developed by the authors, landslide events were monitored with an accuracy rate of 76%, and useful and valuable information was provided to emergency teams and decision-makers. Many countries around the world have created national and regional landslide databases by collecting data from various social media and news sources to investigate the impact of landslide disasters, document damage and loss of life, identify dangerous areas, etc. [42-47]. These databases are then processed and analyzed with various artificial intelligence applications to mitigate the effects of the disaster or to develop early warning systems to warn decision-makers and the public before the disaster occurs [41].

Geographic Information Systems (GIS) are an important tool for real-time monitoring of natural disasters locally and globally and for emergency response. GISs collect data through aerial and satellite images [48-50], radars [48], [51], [52] and various sensors [53-55]. At the same time, GIS technology has been successfully used in the early detection and prevention of landslides [51]. In recent years, many studies have been conducted on the use of GIS and have shown the effectiveness of GIS in detecting landslides. Can et al. created a geographic information system into which citizens can voluntarily upload photos of landslides [56]. They further proposed the CNN architecture to automate the detection of landslide photos from image data in the system and evaluate the accuracy of the landslide events. The CNN architecture proposed detects landslide photos automatically, which will help the rapid assessment of the quality of data when a manual evaluation of data is time-consuming and costly. It therefore means that the CNN architecture could identify photos of landslides with 94% accuracy, which can be applied to community-based emergency response systems. Goniewicz et al. presented how geographic information systems technologies can be used in disaster risk analysis, emergency management, and post-disaster recovery and reconstruction processes [51]. As a result, they exemplified several illustrations of geographic information systems technology for disaster and emergency management and went forward to show that the use of such technologies was an important tool in disaster and emergency management. Yang et al. developed a web-based emergency response and visualization system using Cesium Digital Earth technology [49]. The system will be designed in such a way that it will be able to implement a geographic information system that can help the emergency responders take appropriate actions a lot faster and more potentially during natural disasters. Moreover, a landslide scenario example was used to design how the systems will be employed and the effectiveness they will have.

Yet another technique is the numerical modeling technique. In this technique, landslide risk is estimated by simulating factors such as soil moisture, precipitation, slope, and soil type [57]. The numerical modeling techniques can also be combined with GIS data and machine learning techniques [58]. Similarly, Park et al. proposed a combination of statistical and physically based approaches for the development of a landslide early warning system [59]. It was a study within some regions in South Korea that needed to estimate landslide risk and develop early warning systems based on two different landslide modeling approaches. The first model represents a heuristic approach based on the statistical analysis of past landslides, while the second one is a physically based model that considers precipitation, soil moisture content, and other factors in predicting occurrences of landslide events. Whereas the statistical model offers information related to the analysis of past events, the physical model provides the estimate of landslide risk considering current conditions. Thus, it brings about a more efficient landslide early warning system by combining these two models into one space for both approaches to landslide early warning systems.

Meanwhile, in another research finding, Harilal et al. have also proposed a system for early warning against landslides by issuing warnings for the onset of rainfall conditions that exceed a certain threshold [57]. Landslide probabilities could then be determined based on the geographical characteristics and daily rainfall data latitude provided by the Indian meteorological service. It was then compared with real-time rainfall data, and automatically, in the case of increased risk of landslide, a warning message was generated. The performance of the system was also evaluated. Since then, it found that the rainfall threshold determination for developing a real-time landslide early warning system yielded successful results. Salvatici et al. discussed in their paper [60] the application of physically based models for regional-scale predictions of rainfall-induced shallow landslides. Causes, mechanisms of occurrence, characteristics of shallow landslides, and estimation of shallow landslide probabilities in light of precipitation, moisture, slope, soil properties, and vegetation cover were studied. The model performance is assessed in the paper by comparisons to measurements made at a site known for its accuracy. The results proved that the model makes good predictions of shallow landslide probability and can be used as an uncomplicated early warning system.

Another methodology that could be used either for the detection of areas where landslides occur or to forecast landslide disasters is hybrid approaches. The data obtained by image processing, satellite images, aerial photographs, drone images, or various sensors are rendered meaningful through different techniques. Using artificial neural networks, possible scenarios of natural disasters that may happen in the future can be created by analyzing variously obtained data. These high-resolution satellite images identify the landslides. Characteristics of landslide features can be identified by applying techniques of image processing to these images. The occurrence of the landslide can also be identified by image processing methods, which in turn identifies the factors that cause the occurrence of it to help determine damages caused by landslides and take measures for the safety of the local people. Interferometric synthetic aperture radar and Google Earth imagery were used to develop a landslide early warning system by Nhu et al. in their study [61]. The data they obtained was used to develop a geographic information system, in which those data were classified using machine learning algorithms. For the learning process of landslide susceptibility mapping, they used AdaBoost, Decision Tree, and hybridized versions of these two machine learning algorithms and obtained the most successful results from the AdaBoost model. Thus, they have shown that data obtained with image processing techniques and machine learning algorithms can be used efficiently for the detection of landslides.

A literature synthesis reveals that the principal structure of landslide disaster early warning systems is as follows: data collection, data transmission, modeling, analysis, prediction, and warning. Data is collected through various sensors such as rain gauges, moisture meters, slope meters, groundwater level meters, strain gauges, and geophones in the data collection phase. Data transmission is provided using GPRS, internet, telephone connections, telemetry, Wi-Fi, and various IoT systems. The collected data is then analyzed using various methods to make landslide predictions, and the warning system is activated. Data collection and transmission are among the most critical components of this system. However, if we consider the system as a chain in this way, the system will not work if there is a problem with any link in the chain. Sensors, which are one of the critical components of the system, may be damaged in harsh environmental conditions, or suitable conditions may not be provided for the battery conditions of the sensors, power outages may occur, and the data collection/transmission component will be adversely affected, and the modeling, analysis, and prediction component will not work. Therefore, without data, no output will be produced, and the system will become unusable. Researchers have taken various measures and developed alternative solutions to keep the system running in these and similar situations. Thirugnanam et al. proposed a solution with machine learning algorithms [34]. With the proposed solution, they predicted the data that could not be received from the sensors. They used the previous values of these parameters to predict the changes in slope, which are frequently seen in landslide events and sent them as input to other system components. Harilal et al. [57], Hariharan et al. [62], Guntha et al. [63], and Guntha and Ramesh [63] used multiple meteorological and meteorological data sensors to maintain the system. Huang et al. [64], Fathani et al. [65], Hemalatha et al. [66], and Harilal et al. [57] used multiple data transmission links to overcome data transmission problems. Fathani et al. [65], Hemalatha et al. [66], and Harilal et al. [57] used multiple thresholds in case of data flow problems.

Studies on energy management and energy harvesting of sensors offer essential solutions for wireless sensor networks and IoT applications. Researchers have developed different methods to ensure sensors' energy independence and guarantee the continuity of data collection/transmission processes, especially in harsh environmental conditions. In addition to these methods, reinforcement learning and energy harvesting strategies stand out. Zhang and Lin used RL algorithms to optimize energy-neutral operations in hybrid energy-harvesting wireless body area networks, TBANs [67]. In the work here, energy efficiency is improved by using DQN. Murad et al. optimized the automatic management of energy harvesting IoT nodes with the Proximal Policy Optimization algorithm and gave more autonomous energy management [68]. Omidkar et al. used Q-learning in optimizing spectrum and power allocation in energy harvesting-enabled device-to-device (D2D) communications [69]. Chu et al. optimized power control in multiple access systems with reinforcement learning and improved the system's energy efficiency [70]. Finally, Aoudia et al. developed a reinforcement learning-based energy management algorithm called RLMan [71]. This algorithm optimizes the energy consumption of nodes in wireless sensor networks and dynamically adjusts the power usage based on the state of charge of the node's energy storage device. The algorithm achieved high performance under time-varying energy resource conditions and increased packet transmission rates by up to 70%.

The literature presents significant progress in energy management and harvesting techniques, with various studies proposing innovative methods to improve energy efficiency and ensure the continuous operation of IoT systems. Techniques such as reinforcement learning, dynamic duty cycling, and adaptive power optimization have shown promise in enhancing system performance. However, their original contributions to balancing energy harvesting and consumption remain limited. Many existing approaches focus on optimizing either energy collection or energy consumption independently rather than integrating both processes into a unified and adaptive framework. Additionally, while reinforcement learning algorithms such as Q-learning, DQN, and Proximal Policy Optimization have been applied, their scalability, real-time adaptability, and performance under dynamic and unpredictable environmental conditions are not consistently validated with real-world data. The effectiveness of current methods for achieving energy-neutral operations (ENO) in IoT devices, especially under tough conditions, remains uncertain.

The proposed model aims to fill the critical gaps identified in the literature by introducing a reinforcement learning-based energy management approach that is designed for IoT-based landslide early warning systems. A Deep Q-Network (DQN) is exploited to address the key limitation of the previous methods in achieving an energy-neutral operation under dynamic and unpredictable environmental conditions. Model can be used for dynamically balancing energy-harvesting and power-consumption processes at IoT nodes, adapting to any real-time variations in energy availability. Thus, model guarantees faultless operation and prolongs the operating life of a system by avoiding overcharging or deep depletions of the batteries. Besides, in-depth simulations for real-world solar radiation have also confirmed the good self-adaptability features of the model, showing superior performances compared to traditional approaches. Proposed model enhances the dependability and scalability of the landslide monitoring system and gives the basic foundation for general IoT applications in energy-limited environments.

This paper discusses the most important challenge of any landslide early warning system, how to ensure the long-term energy independency of IoT sensors with continuing their operation. The strategy proposed in this paper seeks to optimize a combined reinforcement learning-energy harvesting model in a coherent, self-adjusting framework that furnishes an integrated solution to a problem that was separately solved by earlier works dealing specifically with either energy harvesting or consumption policies, respectively. With dynamic optimization of energy harvesting and consumption, continuity for data collection and transmission has been provided in changing environmental conditions or in case of a power outage. In this way, model ensures enhancement in the reliability, scalability, and adaptability features with significant contributions to IoT-based natural disaster management systems and beyond.

3. Material and methods

This section describes the material used and methods adopted to develop and validate the reinforcement learning-based approach for optimizing energy management in IoT nodes for the proposed landslide early warning system.

3.1. Reinforcement learning framework and Deep Q-Network (DQN)

This work implements reinforcement learning to enhance energy efficiency in IoT-based landslide early warning systems. RL is a learning process whereby an agent develops a strategy that maximizes its cumulative rewards through interactions with the environment [72]. Accordingly, with respect to each instance, the agent perceives its current state, performs an appropriate action, and obtains a reward as feedback. These rewards contribute to the agent's long-term strategy that involves maximizing the total expected rewards. This total discounted reward, which the agent will try to maximize, is described in Equation 1.

$$R_t = \sum_{t'=t}^{T} \gamma^{t'-t} r_{t'} \tag{1}$$

In this equation, R_t represents the cumulative discounted reward at time step t, with γ serving as the discount factor that reduces the influence of future rewards on current decision-making, with $0 \leq \gamma \leq 1$. The term $r_{t'}$ denotes the instantaneous reward received at each time step, and T is the time step at which the task concludes. By adjusting γ, the agent can balance the trade-off between immediate and future rewards, optimizing its long-term strategy. The agent's objective is to learn a policy that selects actions yielding the highest expected rewards in each state.

To achieve optimal action selection, the Q-learning algorithm estimates an optimal action-value function $Q^*(s,a)$, defined through the Bellman equation. This Equation allows the agent to determine the most suitable action in a given state by considering the maximum expected reward obtainable in the subsequent state. The Bellman equation is formulated in Equation 2 [73].

$$Q^*(s,a) = E_{s' \sim E}[r + \gamma max_{a'} Q^*(s',a')|s,a] \tag{2}$$

In Equation 2, $Q^*(s,a)$ represents the optimal action-value function for a given state-action pair (s,a). The expectation operator E represents the expected value over all possible outcomes. The variable s' denotes the subsequent state, and E represents the environment governing state transitions. The term r signifies the immediate reward, while γ is the discount factor that scales down the importance of future rewards. Finally, $max_{a'} Q^*(s',a')$ identifies the maximum Q-value for the next state s' across all possible actions a'. This equation enables the agent to leverage information about future states to make optimal decisions at each step.

Applying Q-learning in large and complex state spaces presents significant challenges [74]. To address this, Deep Q Networks (DQN) are utilized in this study to approximate the Q-function in high-dimensional state spaces using neural networks [75]. DQN leverages deep neural networks to enhance the learning process, allowing the agent to make efficient decisions in complex environments [76]. In DQN, a loss function $L_i(\theta_i)$ is defined and iteratively optimized, as shown in Equation 3.

$$L_i(\theta_i) = \mathbb{E}_{s,a \sim \rho(\cdot)}[(y_i - Q(s,a;\theta_i))^2] \tag{3}$$

In Equation 3, $L_i(\theta_i)$ denotes the loss function at iteration i, where θ_i represents the network's weight parameters for that iteration. The behavior distribution $p(s,a)$ indicates the probability distribution over the agent's experiences in state s and action a. The target value y_i at each iteration i is given by $y_i = E_{s' \sim E}[r + \gamma max_{a'} Q^*(s',a')|s,a]$, which depends on the expected reward and the maximum estimated Q-value from the previous iteration's parameters θ_{i-1}. This loss function guides weight parameter updates, enhancing the Q-function approximation's accuracy in each iteration.

The gradient of the loss function concerning the weights is computed to update the network parameters, as defined in Equation 4.

$$\nabla_{\theta_i} L_i(\theta_i) = \mathbb{E}_{s,a \sim \rho(\cdot);s' \sim E}[(r + \gamma max_{a'} Q(s',a';\theta_{i-1}) - Q(s,a;\theta_i))\nabla_{\theta_i} Q(s,a;\theta_i)] \tag{4}$$

In Equation 4, $\nabla_{\theta_i} L_i(\theta_i)$ represents the gradient of the loss function at iteration i concerning the weight parameters θ_i. The terms s and a indicate the current state-action pair, while $s' \sim E$ specifies that the next state is sampled from the environment E. The expression r denotes the immediate reward obtained, and γ is the discount factor emphasizing present over future rewards. The term $max_{a'}Q(s', a'; \theta_{i-1})$ selects the a value that gives the highest Q-value for the next state s', based on the parameters from the previous iteration. Finally, $(s, a; \theta_i)$ denotes the predicted Q-value for the current state-action pair with the current parameters θ_i.

This gradient optimizes the network's weight parameters by minimizing the difference between predicted and target Q-values. By leveraging this structure, DQN enables the agent to efficiently learn optimal actions in high-dimensional environments, improving its decision-making process and leading to more accurate policy learning.

3.2. Dataset

The present work proposes a model that uses instantaneous battery level, air temperature, and energy harvesting amount to maintain the battery at an optimum level. To avoid randomness, energy harvesting is simulated. A solar radiation dataset, containing hourly radiation data, is used to calculate the energy harvested by a specified solar panel.

The dataset used is taken from the European Commission Photovoltaic Geographic Information System, and it includes measurements taken in Istanbul-Turkey between 2019 and 2020 [77]. There are 17544 data in the dataset used in training. The first 100 hours of the same dataset were used in the testing phase.

The data used has six columns, which include time, G(i), H_sun, T2m, WS10m, and Int. The time column gives the date and time of measurement. G(i): the amount of global radiation on the inclined plane, a unit of W/m2. H_sun: solar altitude, unite of degree. T2m: air temperature 2 meters above the ground, its unit is degree Celsius. WS10m total wind speed 10 meters above the ground, its unit is m/s. "Int" takes values 0 or 1; 1 is the value when the solar radiation value is reconstructed. During training, G and values are being used for the state space.

3.3. Proposed model

This study proposes a reinforcement learning model to improve energy efficiency and optimize the battery life of IoT nodes used in landslide early warning systems. The proposed model aims to balance IoT nodes' energy consumption and energy harvesting to prevent batteries from being completely depleted or overcharged. In this way, the reliability of landslide early warning systems will be increased by ensuring the uninterrupted and efficient operation of IoT nodes in the field.

It is not feasible to sustain the continuous energy demands of IoT nodes in regions prone to landslides with conventional energy sources. Therefore, the energy needs of IoT nodes should be met by renewable energy sources such as solar energy. In the proposed model, each IoT node is equipped with a solar panel and a rechargeable battery. The energy harvested through the solar panel charges the battery and meets the energy needs of the node. The architecture of this reinforcement learning-based energy management system for IoT nodes is illustrated in Figure 1.

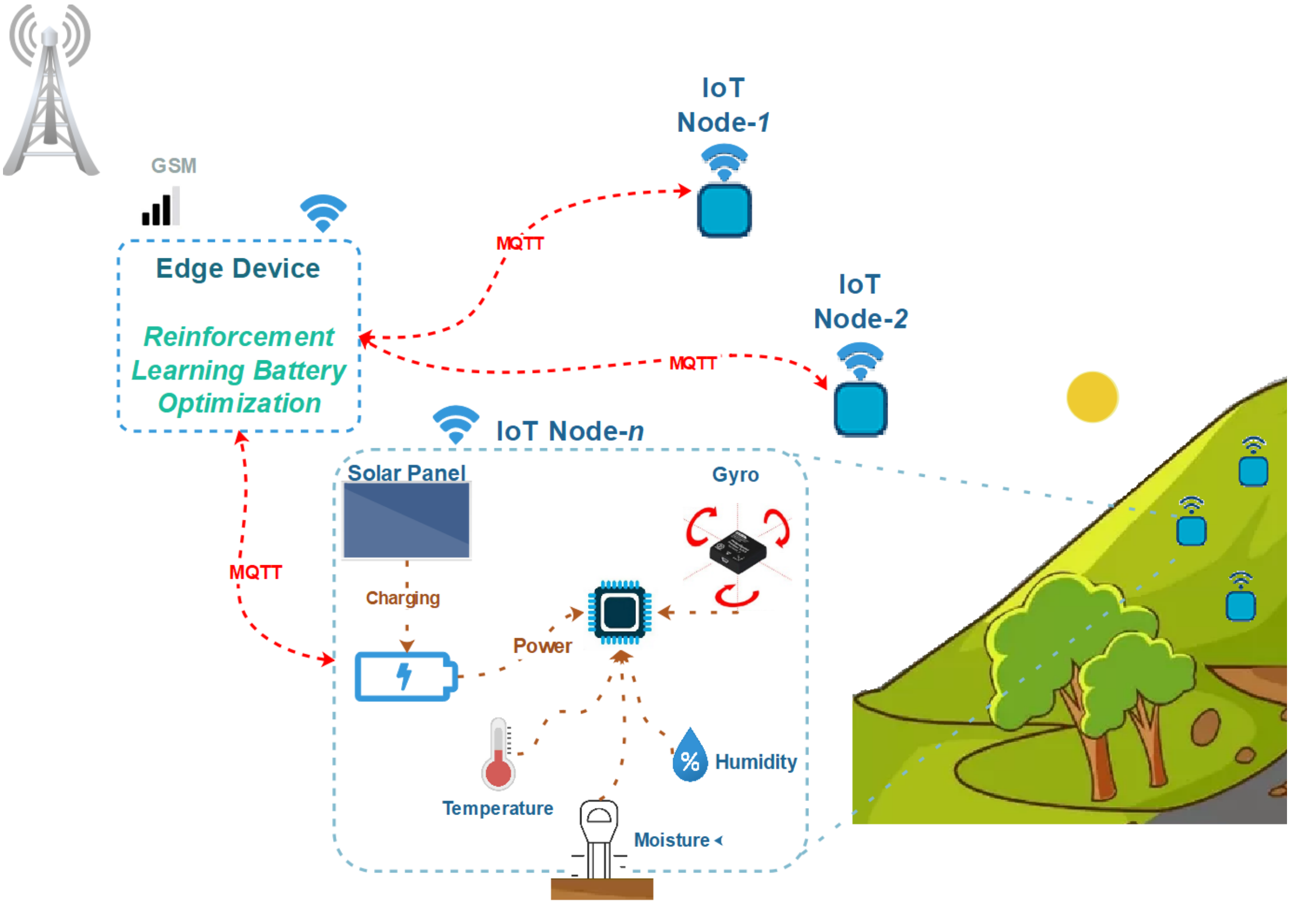

Fig. 1. Architecture of the reinforcement learning-based energy management system for landslide early warning

Figure 1 illustrates the architecture of the IoT-based landslide early warning system that is powered by reinforcement learning-based energy optimization. The formed system includes several IoT nodes equipped with different sensors for monitoring environmental conditions. Each node gathers environmental data such as soil moisture, temperature, and humidity, and sends it wirelessly to a central edge device using the MQTT protocol. The IoT nodes harvest energy through solar panels to meet their energy requirements, with this harvested energy stored in the nodes' batteries.

The reinforcement learning model will analyze data collected to come up with optimum energy consumption by the IoT nodes. The model dynamically adjusts the duty cycle of nodes with consideration of the present level of the battery and how much energy has been harvested. In this way, the system can improve energy efficiency for non-interrupted operation. This architecture is designed to improve energy sustainability in IoT nodes and enhance the reliability of the landslide early warning system.

3.3.1. Reinforcement learning model

The proposed approach is modeled as a Markov Decision Process (MDP) within the reinforcement learning framework. An MDP is defined by the tuple $\langle S, A, P, R, \gamma \rangle$, consisting of four main components: S represents the state space, A represents the action space, P denotes the state transition probabilities, R denotes the reward function, and γ is the discount factor.

In this study, the state space is represented by a three-dimensional vector:

$$s_t = [B_t, T_t, E_{h,t}] \tag{5}$$

$$B_t = \left(\frac{B_{\mathrm{current}}}{B_{\max}}\right) \times 100 \tag{6}$$

The vector s_t, which is provided as input to the trained DQN model, represents the state at time t. Here, B_t indicates the battery level at time t (expressed as a percentage of mAh), which represents the current state of the energy storage in the IoT node and serves as a critical parameter for determining the appropriate duty cycle based on energy availability. T_t represents the ambient temperature at time t (in °C), and $E_{h,t}$ is the amount of harvested energy at time t (in mAh).

The action space encompasses the potential decisions the trained agent can make. In this study, the action space consists of five distinct duty cycle levels.

$$A = \{a_1, a_2, a_3, a_4, a_5\} \tag{7}$$

In the above representation, A defines the state space, and a_1, a_2, a_3, a_4, and a_5 correspond to 20%, 40%, 60%, 80%, and 100% duty cycles, respectively. By simulating energy consumption under these duty cycles, the battery level is updated according to Equation 8.

$$E_{c,t} = P_{\min} + (P_{\max} - P_{\min}) \times \left(\frac{D(a_t) - D_{\min}}{D_{\max} - D_{\min}}\right) \times \Delta t \tag{8}$$

where $E_{c,t}$ denotes the amount of energy consumed over a given time interval. $P_{\min}$ and $P_{\max}$ represent the minimum and maximum energy consumption levels of the IoT node, respectively. $D(a_t)$ is the duty cycle action selected for time t, while $D_{\min}$ and $D_{\max}$ define the minimum and maximum selectable duty cycles, respectively. Following this, the amount of energy harvested from the solar panel during the specified time interval is calculated using Equation 9 to determine the increase in battery level.

$$E_{h,t} = \frac{G_t \times A_p \times \eta_p}{V_s} \times 1000 \times \Delta t \tag{9}$$

In Equation (9), $E_{h,t}$ represents the energy harvested during the time interval t (in mAh). G_t is the solar irradiance at time t (W/m²), A_p is the area of the solar panel (m²), η_p is the efficiency percentage of the solar panel, and V_s is the system voltage (V). The battery level update is performed based on Equations 10 and 11.

$$B_s = B_t - E_{c,t} + E_{h,t} \tag{10}$$

$$B_{t+1} = min(max(B_s, 0), B_{\max}) \tag{11}$$

In these equations, B_s represents the total energy available in the battery after accounting for both the energy consumed by the IoT node and the energy harvested through the solar panel. B_{t+1} denotes the battery level in the next time step; if the total harvested energy causes the battery to exceed its maximum capacity, the level is capped at $B_{\max}$.

The goal of the Reinforcement Learning agent is to maintain the battery level at an optimal level continuously and to be rewarded for the duration it remains at this level. Therefore, the agent is penalized during training for overcharging or discharging. The reward function defined below is used to determine the reward-penalty amounts.

$$d_n = |B_{t+1} - (B_{\max} \times 0.8)| \tag{12}$$

$$O_t = \begin{cases} \left(\frac{B_s - B_{\max}}{B_{\max}}\right) \times \frac{4-k}{4}, & B_s > B_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

$$R_t = \begin{cases} 5, & d_n < 500 \\ 5 - 50 \times \left(\frac{d_n}{B_{\max}} + O_t\right), & 500 \leq d_n < 3000 \\ -50, & d_n \geq 3000 \end{cases} \tag{14}$$

Here, d_n represents the energy neutrality distance, which increases as the battery level deviates from the optimal level of 80%. O_t denotes the amount of overcharge. If the battery level at a given time and the total harvested energy in that interval exceed the maximum battery capacity, and the duty cycle selected is less than 100%, the penalty rate increases. In Equation 13, the value k represents the coefficient for the overcharge penalty, indicating the index of the chosen action in the action space, ranging from 0 to 4. The highest penalty is given when the duty cycle is set to 20%. Finally, using the values of d_n and O_t , the total reward R_t is calculated. R_t reflects the balance between energy efficiency and operational performance, penalizing deviations from the optimal battery level and rewarding actions that maintain energy neutrality

3.3.2. Deep Q Network architecture and model training

The proposed model employs the Deep Q-learning (DQN) algorithm as the reinforcement learning method. DQN is an extension of the traditional Q-learning algorithm using deep neural networks, enabling it to operate effectively in high-dimensional state spaces. The DQN network here consists of a three-dimensional input layer that takes the state space as input, three hidden layers of 64, 32, and 8 dimensions, respectively, and a five-dimensional output layer that provides an action from the action space as output. The architecture of the network used in training is shown in Figure 2.

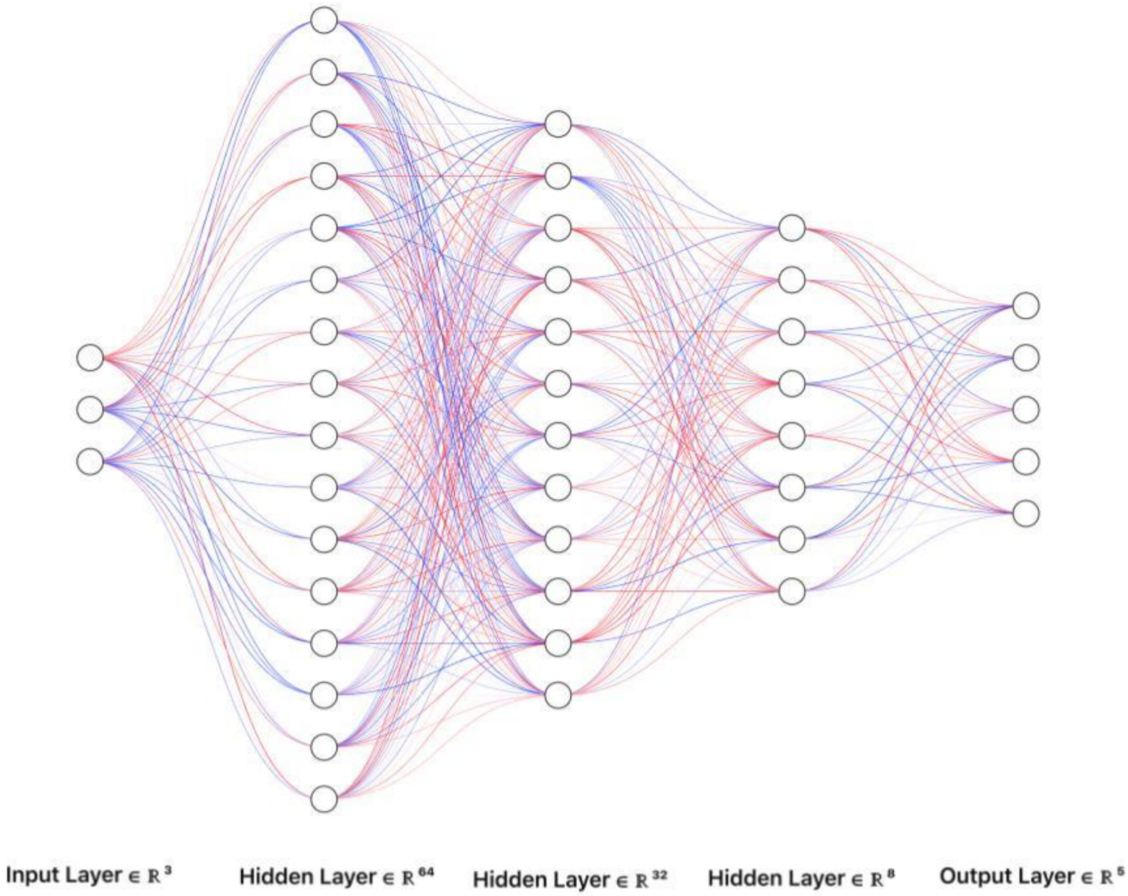

Fig. 2. Architecture of the DQN used in training.

This neural network is used to approximate the $Q(s, a; \theta)$ state-action value function, where θ represents the neural network's weights.

During the training process, the neural network parameters θ are initialized randomly. The Epsilon-Greedy strategy is applied in training, with the epsilon (ε) value initially set to 1,0 [78]. A random number between 0 and 1 is then selected and compared with the ε value. If the randomly chosen number is less than ε, an action is chosen randomly from the action space. Otherwise, the action with the highest Q value, as predicted by the neural network, is selected. At the end of each time step, the value of ε is reduced according to Equation 15.

$$\varepsilon = max(\varepsilon_{min}, \varepsilon \times \varepsilon_{\text{decay}}) \tag{15}$$

where ε_{min} is the minimum allowable epsilon value, and $\varepsilon_{\text{decay}}$ is the decay rate per time step. After these operations, the action a_t is used to select a duty cycle, resulting in the next state s_{t+1} and the reward R_t. The information from this experience is then stored in the experience replay memory. An example of the format for stored data in the replay memory is shown below:

$$(s_t, a_t, R_t, s_{t+1}, \partial) \quad (16)$$

Here, ∂ is a boolean variable indicating whether the training episode ends after this step. Experience data is stored in memory until a specified memory size is reached.

There are two main components in the application of the DQN model. One is the primary model used while the agent continuously interacts with the environment. The other is the target model, which is used to prevent instability caused by rapid changes in the weights of the primary model. The weights of the target model are gradually transferred from the primary model to the target model using a method known as soft update [79].

$$y_j = \begin{cases} R_j, & \text{if } \partial_j = \text{True} \\ R_j + \gamma \times max_{a'} Q(s_{j+1}, a'; \theta^-), & \text{otherwise} \end{cases} \quad (17)$$

where y_j represents the target Q values, θ^- represents the parameters of the target network and γ represents the the discount factor, which determines the importance of future rewards compared to immediate rewards. A value of γ close to 1 prioritizes long-term rewards, while a lower value focuses on short-term gains, allowing the model to balance immediate performance with sustainable energy neutrality over time. The weights of the target network are adjusted to approach the weights of the main network according to the following assignment:

$$\theta^- \leftarrow \tau \times \theta + (1 - \tau) \times \theta^- \quad (18)$$

Here, τ is the soft update rate, typically a small value like 0,001. The loss function used in the trained model is the Mean Squared Error (MSE), calculated for the specified model as follows:

$$L(\theta) = \frac{1}{N} \sum_i (y_j - Q(s_j, a_j; \theta))^2 \quad (19)$$

Figure 3 presents the pseudocode of the proposed model, which outlines the step-by-step process of the reinforcement learning-based energy optimization framework, including initialization, action selection, energy calculations, reward assignment, and model updates.

```
  - Initialize the DQN parameters θ
  - Clone θ to θ^- (target network)
  - Initialize replay memory with capacity defined in Replay_buffer_size
  - Set ε = Exploration_rate
for_each_episode:
    - Set B_current = B_max
    - Set done = false
    - Set total_reward = 0
    - Set time_step t = 0
    - Initialize state s_0 = [Irradiation_0, Temperature_0, (B_current/B_max)x100]
    while_not_done:
        - Select action a_t randomly with probability ε
        - Otherwise, a_t = argmax_a Q(s_t; θ)
        - duty_cycle = (a_t + 1) × (incremental percentage)
        - E_c = Base_Consumption + (Additional_Consumption × a_t)
        - E_h = ((Irradiation_t × A_p × η_p) / V_s) × 1000 (mAh)
        - B_temp = B_current - E_c + E_h
        - if B_temp > B_max:
            - overcharge_penalty_rate = (B_temp - B_max) / B_max
            - overcharge_penalty_rate = overcharge_penalty_rate × ((4 - a_t) / 4)
            - B_current = B_max
          else:
            - overcharge_penalty_rate = 0
            - B_current = max(B_temp, 0)
            - if B_current == 0:
                - done = true
        - s_(t+1) = [Irradiation_(t+1), Temperature_(t+1), (B_current/B_max)x100]
        - neutrality_dist = |B_current - B_opt|
        - if neutrality_dist < Threshold_1:
            - r_t = High_Value
        - else_if (neutrality_dist >= Threshold_1 and neutrality_dist < Threshold_2):
            - r_t = Medium_Base_Value - (Additional_Penalty_Factor ×
                  ((neutrality_dist / B_max) + overcharge_penalty_rate))
        - else:
            - r_t = Heavy_Penalty
        - Store (s_t, a_t, r_t, s_(t+1), done) in replay memory
        - if replay_memory_size > Batch_size:
            - Sample a minibatch of transitions
            - for each sampled transition:
                if done:
                  - y_j = r_j
                else:
                  - Compute target y_j = r_j + Discount_factor × max_a Q(s_(j+1); θ^-)
            - Update θ by minimizing the loss between Q(s_j, a_j; θ) and y_j
            - Soft update target network: θ^- ← τxθ + (1 - τ)xθ^-
        - s_t = s_(t+1)
        - total_reward += r_t
        - Increment t by 1
        - If done or t reaches the end of the data sequence, break the loop
    - ε = max(epsilon_min, ε × Decay_Rate)
    - Optionally save model parameters (θ and θ^-) at scheduled intervals
```

Fig. 3. Pseudocode of the proposed model.

4. Results and discussion

This section covers the results from our study on how the RL-based model works in optimizing energy management in IoT nodes to allow for early detection of natural hazards like landslides and outlines some benchmark parameters that had been used for training the RL model and the agent and a simulation environment which was set up with these parameters. Details of the parameters and the values used for training are listed in Table 1.

Table 1. Parameters and configuration for agent training and simulation environment.

Parameter Type	Parameter Name	Parameter Value
Agent	State Space Size	3
Agent	Action Space Size	5
Agent	Exp. Replay Memory Size	1000
Agent	Gamma (γ)	0.95
Agent	Epsilon Decay ($\varepsilon_{\text{decay}}$)	0.9995
Agent	Tau (τ)	0.001
Model	Hidden Layer Activation	ReLU
Model	Output Layer Activation	Linear
Model	Optimizer	Adam
Model	Optimizer Learning Rate	0.001
Model	Loss Function	MSE
Environment	Panel Area (A_p)	0.5 m2
Environment	Max Battery Level (B_{max})	5000 mAh
Environment	Panel Efficiency (η_p)	0.15
Environment	Optimum Battery Level	4000 mAh
Environment	System Voltage	5 Volt
Training	Batch Size	32
Training	Episode	1000

The variables in Table 1 describe the parameters and configurations used for the training and simulation environment of the reinforcement learning model proposed in the article. State Space Size refers to the three-dimensional data vector that the model receives as input, while Action Space Size defines five different duty cycles that the model can choose from. Exp. Replay Memory Size indicates the amount of experiences retained during the learning process. Gamma value is the discount factor that determines the effect of future rewards on today's rewards. Epsilon Decay regulates how the selection rate of random actions in the epsilon-change strategy decreases over time. Tau defines the soft update of the weights of the target model to the main model. The Hidden Layer Activation and Output Layer Activation functions used in the training of the model specify how the activation between layers will occur during the learning process of the neural network. Optimizer and Optimizer Learning Rate determine the weight update strategy and speed of the model, while Loss Function is an error measure that evaluates the performance of the model. Parameters such as Panel Area, Panel Efficiency, System Voltage, Max Battery Level and Optimum Battery Level used in the simulation environment describe the physical properties used in energy collection and consumption calculations. Finally, Batch Size and Episode define the amount of data processing in the training process of the model and the total number of training cycles. These parameters allow the proposed model to be configured for energy management optimization.

Training and simulations were performed using the computer hardware specification: AMD Ryzen 5600X, NVIDIA RTX 3060 with CUDA for training acceleration, and RAM 16 GB. The results demonstrated a balance between energy harvesting and consumption, enabling the IoT node to operate continuously within the specified duty cycles. It is observable from Figure 4 that during highly energy-available periods of the year, the battery level remains full. During winter, when energy becomes really scarce, the level of the battery decreases but never to zero. This means the model maintains the IoT node always on by keeping the battery at an optimal level even when energy harvesting is low. In this way, the IoT node can operate continuously without interruption and never allow the battery level to drop to zero.

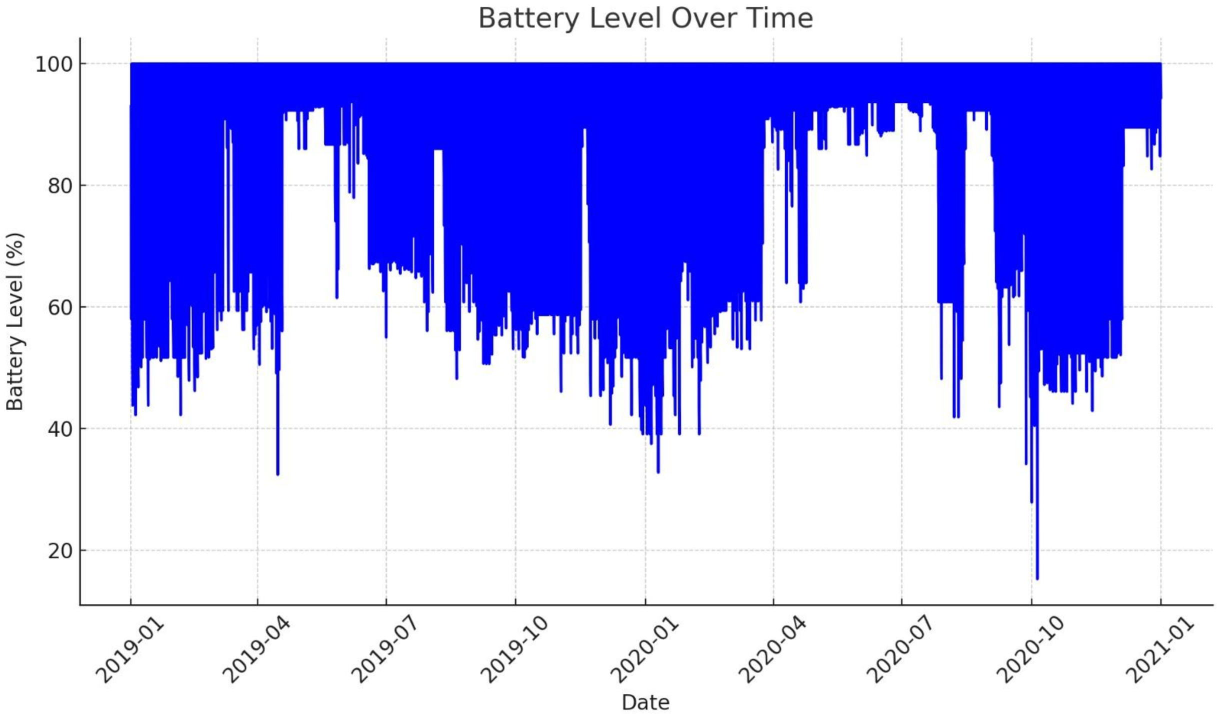

Fig. 4. Variation of battery level over time.

As shown in Figure 5, the energy harvesting data reveal that the model dynamically adjusts the IoT node's operational cycle by considering seasonal variations throughout the year. This allows for maximum energy harvesting during summer months with high solar irradiance. Therefore, the model is able to maintain a higher level of battery and increase its IoT node activity cycle. The model improves system performance by allowing it to operate at a high-duty cycle when the battery level and energy harvesting are high. Conversely, if the energy harvesting during winter is low, the proposed model optimizes the energy consumption by operating an IoT node on a low-duty cycle. It is an effective adaptiveness for energy neutrality, enabling the system to utilize energy efficiently.

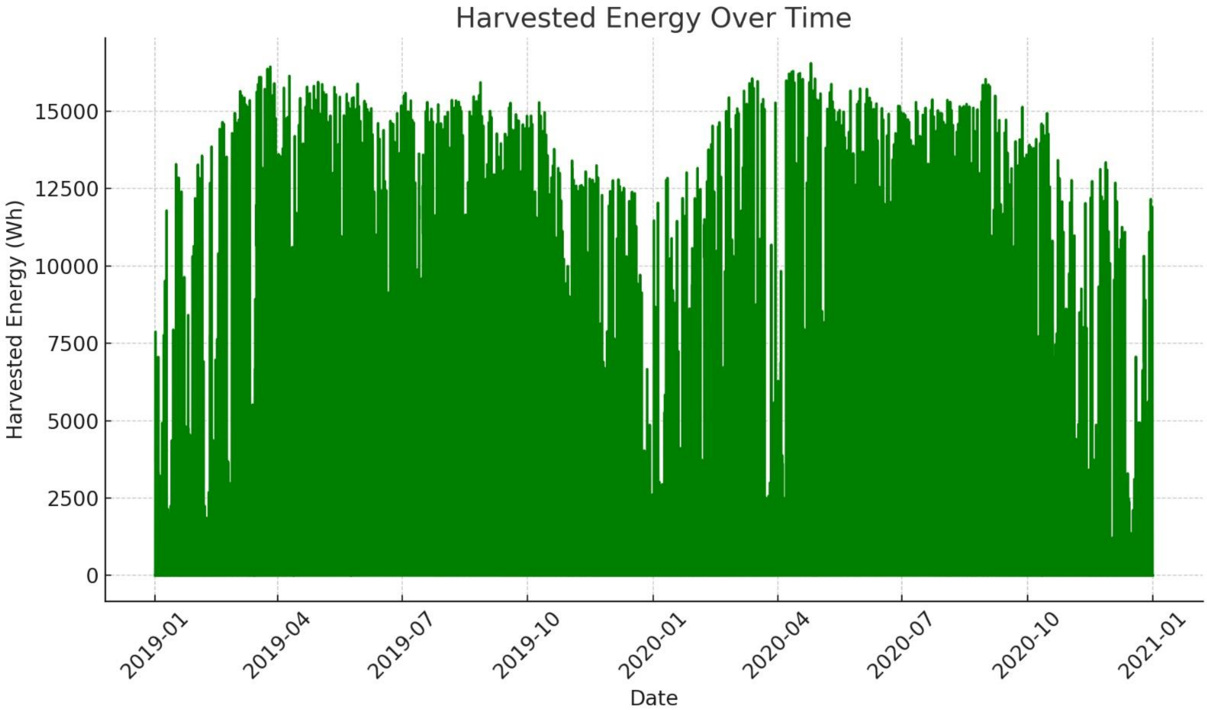

Fig. 5. Variation of harvested energy over time.

Figure 6 illustrates how dynamic and adaptive structure of the model adjusts the duty cycle in response to varying battery levels, ensuring energy neutrality. At lower battery levels (e.g., 20%-40%), the model automatically switches to a reduced duty cycle (20%-40%) to minimize energy consumption and prevent over-discharging. On the other hand, when the battery level increase (e.g., 80%-100%), the model progressively selects higher duty cycles (80%-100%), maximizing performance and data collection while maintaining energy sustainability.

This adaptive behavior highlights the ability to balance energy harvesting and consumption, ensuring uninterrupted operation even during periods of limited energy availability. The figure clearly demonstrates this relationship, with annotations providing additional clarity on the specific duty cycles selected at various battery levels. This feature ensures the IoT node operates efficiently and remains energy-neutral, adapting seamlessly to fluctuations in energy availability

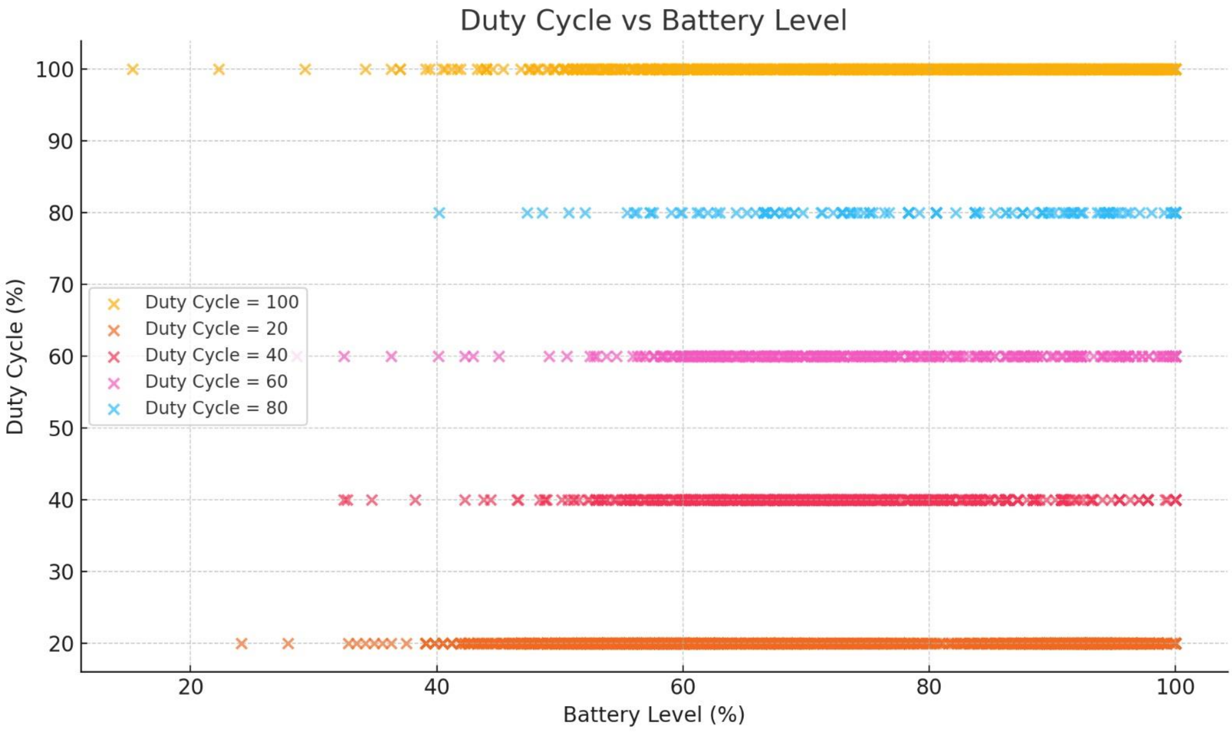

Fig. 6. Relationship between battery level and selected duty cycle.

This adaptive operational strategy is further corroborated in the relationship between total reward and duty cycle as depicted in Figure 7, whereby actions by the agent return higher rewards than random actions. It changes the duty cycle, depending on the battery level, for doing some energy efficiency enhancement. In this way, the IoT node gets operated uninterruptedly. This signifies that decisions made by the model are optimum concerning power management and thereby highest efficiency in the operation of the IoT node.

Fig. 7. Relationship between selected duty cycles and obtained rewards.

In Figure 8, which illustrates the relationship between action types and total reward, the difference between random actions (denoted as Action Type 0) and the actions selected by the trained RL agent (denoted as Action Type 1) is evident. The trained RL agent has achieved significantly higher rewards compared to random actions. This demonstrates that the RL agent makes more informed and optimized decisions regarding the energy management of the IoT node. The model selects appropriate duty cycles at every given time, taking into consideration the battery level and energy harvesting, hence increasing energy efficiency and feeding the continuous operation of the IoT node. It prefers low-duty cycles mainly during the periods of low-energy harvesting to save energy, while during high-energy harvesting periods, it operates the system more efficiently with high-duty cycles. These prove that the model is more successful than random actions and it can select the optimal action for continuous and efficient operation of IoT node.

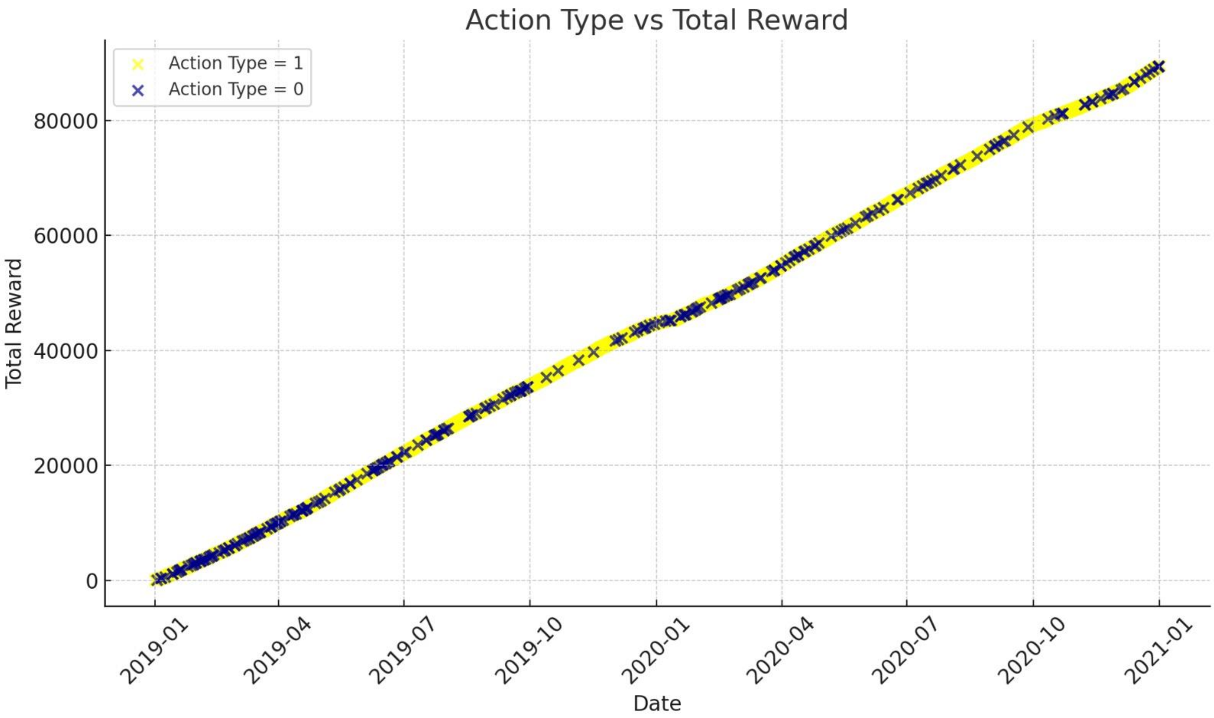

Fig. 8. Total rewards obtained by random actions and those selected by the agent.

Finally, it is observed from Figure 9 that the line that shows the relation between the epsilon value and total reward is such that while the epsilon value decreases, the total reward increases. This indicates that the total reward obtained with randomly determined actions was substantially lower, but higher rewards were achieved when the model chose the actions to be performed. This reflects the effectiveness that the model gained during the learning process.

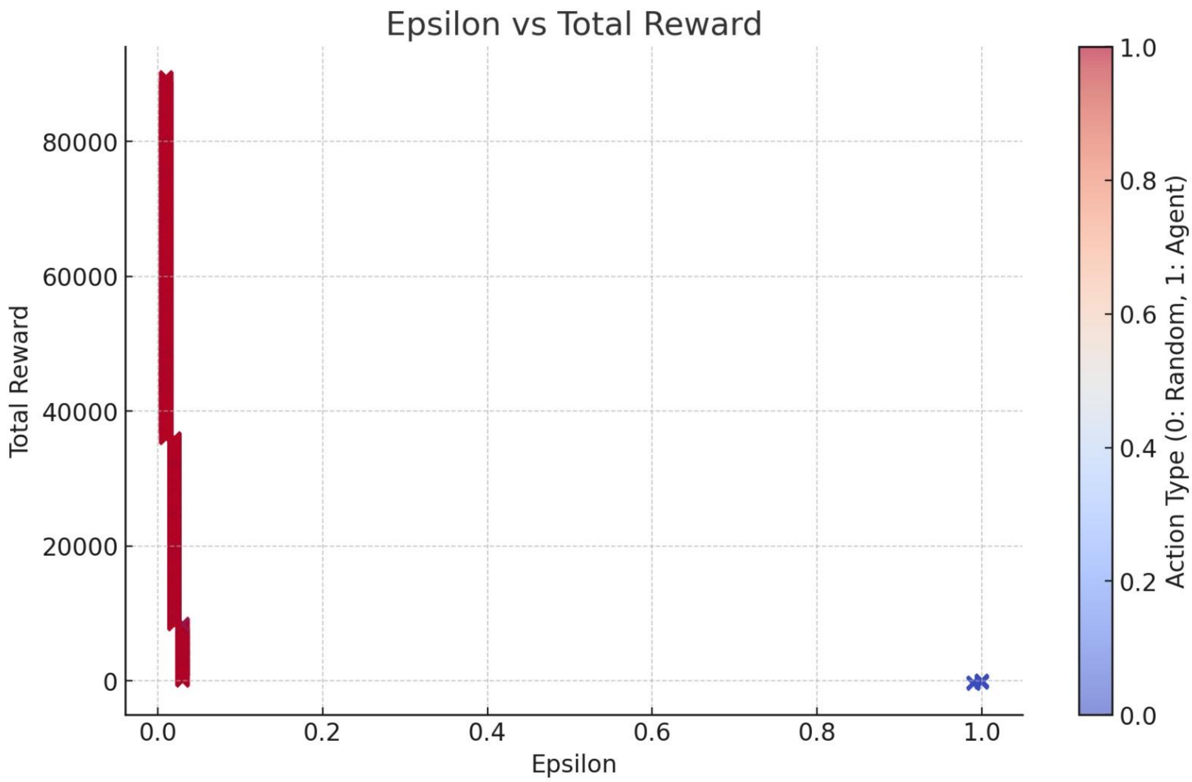

Fig. 9. Relationship between epsilon value and obtained reward.

These results confirm the efficiency of the proposed model based on reinforcement learning towards improving energy efficiency in IoT-based landslide early warning systems and enabling continuous operation. The model helps the IoT node use energy efficiently by keeping the battery level in an optimal state and enhances system reliability. These reflect that a dynamic balance exists between energy harvesting and the duty cycle and that the model effectively presents a sustainable solution for energy management in IoT systems.

Table 2. Uptime comparison with studies in the literature.

Study	Uptime (Seconds)	Uptime (Hours)	Uptime (Days)
Ait Aoudia et al. [80]	23,328,000	6,480	270
Murad et al. [81]	604,800	168	7
Charef et al. [82]	500	0.14	~0.006
Abadi et al. [83]	3,000	0.83	~0.035
Proposed Method	**63,158,400**	**17,544**	**731**

The comparison in the table clearly shows that the proposed model provides a significant superiority over the existing studies in the literature in terms of the working time of the IoT node. The proposed method worked continuously for 17544 hours (731 days, 2 years), exceeding the highest time of 270 days in the table by approximately 2.7 times. These results show that the proposed model offers a more efficient strategy in terms of duty cycle optimization and battery management, and ensures that the IoT node remains operational for a long time. This is a significant advantage, especially in remote and maintenance-requiring environments where energy collection is limited.

5. Conclusion

This work presents a novel RL-based model to improve energy efficiency and operational sustainability in IoT-based landslide early warning systems. This study will employ the DQN algorithm to achieve energy neutrality by effectively optimizing the balance between energy harvesting and consumption. Our strategy is centered on an optimum level of the battery of IoT sensor nodes in such a way that continuous operational capability would be ensured for those nodes installed or deployed in remote disaster areas where the probability of frequent maintenance is not feasible. Unlike conventional energy management models, our model dynamically adapts to variations in environmental energy availability—solar radiation specifically—through adjustment in the duty cycle of sensor nodes. Thus, when energy harvest is high, the model increases the duty cycle to capture and utilize more data for better system reliability and response. At low energies, the model reduces the duty cycle—a reason being to save the battery life without completely compromising any core functionalities.

However, the study has some limitations that should be addressed in future work. Firstly, while the model was validated using real-world solar radiation data, its performance under other environmental conditions (e.g., wind energy, vibration-based energy harvesting) was not considered. Future research could expand the scope by incorporating multi-source energy harvesting and investigating the impact of environmental variability on system performance. Future studies will also focus on validating the proposed model using different datasets collected from various geographic locations and environmental conditions to further enhance its robustness and applicability

These results from the simulation verify that the model is effective in ensuring continuity during prolonged operations. We have used real-world solar radiation data to demonstrate that the model's adaptive management can enable IoT nodes to attain high-duty cycles during energy-rich conditions and low-duty cycles during energy-starved conditions. This kind of strategy extends not only the battery lifetime of sensor nodes but also enhances landslide early warning system resilience to keep it operational through seasonal cycles and inclement weather.

As the limitation, this study relied on simulations for validation. While the results are promising, deploying the model in real-world landslide-prone areas and monitoring its long-term performance would provide more robust insights. This can include assessing the model's ability to adapt to hardware failures, unpredictable weather patterns, and sensor inaccuracies.

These features make the proposed RL-based model quite suitable for IoT applications in unpredictable and energy-limited environments. The work also contributes to a wide variety of energy management issues in IoT by showing how reinforcement learning can manage energy resources autonomously in this study. Additionally, further work could focus on the following directions: integrating advanced machine learning techniques, such as hybrid reinforcement learning or transfer learning, to enhance the system's adaptability and decision-making in highly dynamic environments. Developing more energy-efficient hardware for IoT nodes to complement the software-based optimization approaches. Exploring the scalability of the proposed model in larger IoT networks with heterogeneous energy sources and varying duty cycle requirements.

Our findings have proved that the RL approach can be expanded and tailored to accommodate complex demands for energy in IoT systems other than landslide monitoring, providing a foundation for future research focused on RL-based energy management in diverse IoT applications.

Author contributions

S.A., S.D., and F.A. actively participated in conducting the experimental studies and writing the manuscript.

Acknowledgements

This study did not receive any specific funding or financial assistance from governmental, commercial, or nonprofit organizations.

References

[1] R. Steen, E. Roud, T. M. Torp, and T.-A. Hansen, "The impact of interorganizational collaboration on the viability of disaster response operations: The Gjerdrum landslide in Norway," *Saf Sci*, vol. 173, p. 106459, 2024.
[2] D. Petley, "Global patterns of loss of life from landslides," *Geology*, vol. 40, no. 10, pp. 927–930, 2012.
[3] M. T. Chaudhary and A. Piracha, "Natural disasters—origins, impacts, management," *Encyclopedia*, vol. 1, no. 4, pp. 1101–1131, 2021.
[4] F. Guzzetti *et al.*, "Geographical landslide early warning systems," *Earth Sci Rev*, vol. 200, p. 102973, 2020.
[5] M. A. Jamshed, K. Ali, Q. H. Abbasi, M. A. Imran, and M. Ur-Rehman, "Challenges, applications, and future of wireless sensors in Internet of Things: A review," *IEEE Sens J*, vol. 22, no. 6, pp. 5482–5494, 2022.
[6] N. Casagli, E. Intrieri, V. Tofani, G. Gigli, and F. Raspini, "Landslide detection, monitoring and prediction with remote-sensing techniques," *Nat Rev Earth Environ*, vol. 4, no. 1, pp. 51–64, 2023.
[7] M. Esposito, L. Palma, A. Belli, L. Sabbatini, and P. Pierleoni, "Recent advances in internet of things solutions for early warning systems: A review," *Sensors*, vol. 22, no. 6, p. 2124, 2022.
[8] J. Singh, R. Kaur, and D. Singh, "Energy harvesting in wireless sensor networks: A taxonomic survey," *Int J Energy Res*, vol. 45, no. 1, pp. 118–140, 2021.
[9] A. J. Williams, M. F. Torquato, I. M. Cameron, A. A. Fahmy, and J. Sienz, "Survey of energy harvesting technologies for wireless sensor networks," *IEEE Access*, vol. 9, pp. 77493–77510, 2021.
[10] T. Sanislav, G. D. Mois, S. Zeadally, and S. C. Folea, "Energy harvesting techniques for internet of things (IoT)," *IEEE access*, vol. 9, pp. 39530–39549, 2021.
[11] S. Shresthamali, M. Kondo, and H. Nakamura, "Multi-objective reinforcement learning for energy harvesting wireless sensor nodes," in *2021 IEEE 14th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, IEEE, 2021, pp. 98–105.
[12] A. Karras, C. Karras, I. Karydis, M. Avlonitis, and S. Sioutas, "An Adaptive, Energy-Efficient DRL-Based and MCMC-Based Caching Strategy for IoT Systems," in *International Symposium on Algorithmic Aspects of Cloud Computing*, Springer, 2023, pp. 66–85.
[13] E. O. Arwa and K. A. Folly, "Reinforcement learning techniques for optimal power control in grid-connected microgrids: A comprehensive review," *Ieee Access*, vol. 8, pp. 208992–209007, 2020.
[14] L. M. Highland and P. Bobrowsky, The landslide handbook-A guide to understanding landslides, no. 1325. *US Geological Survey*, 2008.
[15] A. Tekerek and M. Dörterler, "The adaptation of gray wolf optimizer to data clustering," *Politeknik Dergisi*, p. 1, 2020.
[16] S. Dörterler, "Kanser Hastalığı Teşhisinde Ölüm Oyunu Optimizasyon Algoritmasının Etkisi," *Mühendislik Alanında Uluslararası Araştırmalar VIII*, p. 15, 2023.
[17] İ. Akgül and V. Kaya, "Learning performance of optimization algorithms in convolutional neural networks: An application," in *INSAC Scientific Researches in Natural and Engineering Sciences*, 1st ed., vol. 1, Duvar, 2022, ch. 11, pp. 215–236.
[18] N. Yağmur, H. Temurtaş, and İ. Dağ, "Anemi hastalığının yapay sinir ağları yöntemleri kullanılarak sınıflandırılması," *Journal of Scientific Reports-B*, no. 008, pp. 20–34, 2023.
[19] N. N. Arslan, E. Şahin, and M. Akçay, "Deep learning-based isolated sign language recognition: a novel approach to tackling communication barriers for individuals with hearing impairments," *Journal of Scientific Reports-A*, no. 055, pp. 50–59, 2023.
[20] V. Kaya, I. Akgül, and Ö. Tanır, "A novel hybrid model based on machine and deep learning techniques for the classification of microalgae.," *Phyton-International Journal of Experimental Botany*, vol. 92, no. 9, pp. 2519-2534, 2023.
[21] G. Arslan, F. Aydemir, and S. Arslan, "Enhanced license plate recognition using deep learning and block-based approach," *Journal of Scientific Reports-A*, no. 058, pp. 57–82, 2023.
[22] F. Aydemir and S. Arslan, "A System Design With Deep Learning and IoT to Ensure Education Continuity for Post-COVID," *IEEE Transactions on Consumer Electronics*, 2023.
[23] S. Dörterler, H. Dumlu, D. Özdemir, and H. Temurtaş, "Hybridization of Meta-heuristic Algorithms with K-Means for Clustering Analysis: Case of Medical Datasets," *Gazi Mühendislik Bilimleri Dergisi*, pp. 1–23.
[24] E. Şahin, D. Özdemir, and H. Temurtaş, "Multi-objective optimization of ViT architecture for efficient brain tumor classification," *Biomed Signal Process Control*, vol. 91, p. 105938, 2024.
[25] F. Aydemir and S. Arslan, "Covid-19 pandemi sürecinde çocukların el yıkama alışkanlığının nesnelerin interneti tabanlı sistem ile izlenmesi," *Mühendislik Bilimleri ve Araştırmaları Dergisi*, vol. 3, no. 2, pp. 161–168, 2021.
[26] M. C. Özbalcı, H. Şahin, and T. T. Bilgin, "Makine Öğrenmesi Yöntemleri ile GTZAN Veri Kümesine Ait Müzik Türlerinin Sınıflandırılması," *Mühendislik Bilimleri ve Araştırmaları Dergisi*, vol. 6, no. 1, pp. 77–87.
[27] E. Şahin and M. F. Talu, "WY-NET: A NEW APPROACH to IMAGE SYNTHESIS with GENERATIVE ADVERSARIAL NETWORKS," *Journal of Scientific Reports-A*, no. 050, pp. 270–290, 2022.
[28] E. Şahin and M. F. Talu, "Automatic Mustache Pattern Production on Denim Fabric with Generative Adversarial Networks," *Computer Science*, vol. 7, no. 1, pp. 1–9, 2022.
[29] N. Yagmur, İ. Dag, and H. Temurtas, "Classification of anemia using Harris hawks optimization method and multivariate adaptive regression

spline," *Neural Comput Appl*, pp. 1–20, 2024.
[30] N. Yagmur, I. Dag, and H. Temurtas, "A new computer-aided diagnostic method for classifying anaemia disease: Hybrid use of Tree Bagger and metaheuristics," *Expert Syst*, p. e13528, 2023.
[31] İ. Akgül, V. Kaya, E. Karavaş, S. Aydın, and A. Baran, "A Novel Artificial Intelligence-Based Hybrid System to Improve Breast Cancer DetectionUsing DCE-MRI," *Bulletin Of The Polish Academy Of Sciences. Technical Sciences*, vol. 72, no. 3, 2024.
[32] H. Ali, U. U. Tariq, M. Hussain, L. Lu, J. Panneerselvam, and X. Zhai, "ARSH-FATI: A novel metaheuristic for cluster head selection in wireless sensor networks," *IEEE Syst J*, vol. 15, no. 2, pp. 2386–2397, 2020.
[33] H. Ali *et al.*, "A survey on system level energy optimisation for MPSoCs in IoT and consumer electronics," *Comput Sci Rev*, vol. 41, p. 100416, 2021.
[34] H. Thirugnanam, M. V. Ramesh, and V. P. Rangan, "Enhancing the reliability of landslide early warning systems by machine learning," *Landslides*, vol. 17, pp. 2231–2246, 2020.
[35] T. Hemalatha, M. V. Ramesh, and V. P. Rangan, "Effective and accelerated forewarning of landslides using wireless sensor networks and machine learning," *IEEE Sens J*, vol. 19, no. 21, pp. 9964–9975, 2019.
[36] E. Collini, L. A. I. Palesi, P. Nesi, G. Pantaleo, N. Nocentini, and A. Rosi, "Predicting and understanding landslide events with explainable AI," *IEEE Access*, vol. 10, pp. 31175–31189, 2022.
[37] C. W. W. Ng, B. Yang, Z. Q. Liu, J. S. H. Kwan, and L. Chen, "Spatiotemporal modelling of rainfall-induced landslides using machine learning," *Landslides*, vol. 18, pp. 2499–2514, 2021.
[38] R. Franceschini, A. Rosi, F. Catani, and N. Casagli, "Exploring a landslide inventory created by automated web data mining: the case of Italy," *Landslides*, vol. 19, no. 4, pp. 841–853, 2022.
[39] X. Qian, "Regional Geological Disasters Emergency Management System Monitored by Big Data Platform," *Processes*, vol. 10, no. 12, p. 2741, 2022.
[40] X. Chen, "Early Warning of Regional Landslide Disaster and Development of Rural Ecological Industrialization Based on IoT Sensor," *Sci Program*, vol. 2022, 2022.
[41] C. V. L. Pennington *et al.*, "A near-real-time global landslide incident reporting tool demonstrator using social media and artificial intelligence," *International Journal of Disaster Risk Reduction*, vol. 77, p. 103089, 2022.
[42] G. Herrera *et al.*, "Landslide databases in the Geological Surveys of Europe," *Landslides*, vol. 15, pp. 359–379, 2018.
[43] K. Sassa and P. Canuti, *Landslides-disaster risk reduction*. Springer Science & Business Media, 2008.
[44] B. Rosser, S. Dellow, S. Haubrock, and P. Glassey, "New Zealand's national landslide database," *Landslides*, vol. 14, pp. 1949–1959, 2017.
[45] C. S. Juang, T. A. Stanley, and D. B. Kirschbaum, "Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR)," *PLoS One*, vol. 14, no. 7, p. e0218657, 2019.
[46] M. J. Froude and D. N. Petley, "Global fatal landslide occurrence from 2004 to 2016," *Natural Hazards and Earth System Sciences*, vol. 18, no. 8, pp. 2161–2181, 2018.
[47] Q. Han, K. Sassa, and M. Mikoš, "International programme on landslides (ipl): a programme of the icl for landslide disaster risk reduction," *Understanding and Reducing Landslide Disaster Risk: Volume 1 Sendai Landslide Partnerships and Kyoto Landslide Commitment 5th*, pp. 187–203, 2021.
[48] A. C. Teodoro and L. Duarte, "The synergy of remote sensing and geographical information systems in the management of natural disasters," in *Nanotechnology-Based Smart Remote Sensing Networks for Disaster Prevention*, Elsevier, 2022, pp. 217–230.
[49] Z. Yang, J. Li, J. Hyyppä, J. Gong, J. Liu, and B. Yang, "A comprehensive and up-to-date web-based interactive 3D emergency response and visualization system using Cesium Digital Earth: taking landslide disaster as an example," *Big Earth Data*, pp. 1–23, 2023.
[50] M. N. I. Sarker, Y. Peng, C. Yiran, and R. C. Shouse, "Disaster resilience through big data: Way to environmental sustainability," *International Journal of Disaster Risk Reduction*, vol. 51, p. 101769, 2020.
[51] K. Goniewicz *et al.*, "Geographic information system technology: review of the challenges for its establishment as a major asset for disaster and emergency management in Poland," *Disaster Med Public Health Prep*, vol. 15, no. 5, pp. 573–578, 2021.
[52] C. Nefros, G. Kitsara, and C. Loupasakis, "Geographical Information Systems and Remote Sensing Techniques to Reduce the Impact of Natural Disasters in Smart Cities," *IFAC-PapersOnLine*, vol. 55, no. 11, pp. 72–77, 2022.
[53] P. V Gorsevski, M. K. Brown, K. Panter, C. M. Onasch, A. Simic, and J. Snyder, "Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park, Ohio," *Landslides*, vol. 13, pp. 467–484, 2016.
[54] N. M. Abdullah, N. Sulaiman, U. Nazir, M. Ismail, S. K. K. A. Latib, and N. P. N. Mahmud, "Geographical Information System (GIS) and Remote Sensing (RS) Applications in Disaster Risk Reduction (DRR) in Malaysia," *International Journal of Integrated Engineering*, vol. 14, no. 5, pp. 25–37, 2022.
[55] I. Alcántara-Ayala and R. J. Garnica-Peña, "Landslide Warning Systems in Upper Middle-Income Countries: Current Insights and New Perspectives," in *Progress in Landslide Research and Technology, Volume 1 Issue 2, 2022*, Springer, 2023, pp. 159–168.
[56] R. Can, S. Kocaman, and C. Gokceoglu, "A convolutional neural network architecture for auto-detection of landslide photographs to assess citizen science and volunteered geographic information data quality," *ISPRS Int J Geoinf*, vol. 8, no. 7, p. 300, 2019.
[57] G. T. Harilal, D. Madhu, M. V. Ramesh, and D. Pullarkatt, "Towards establishing rainfall thresholds for a real-time landslide early warning system in Sikkim, India," *Landslides*, vol. 16, no. 12, pp. 2395–2408, 2019.
[58] X. Fan *et al.*, "Prediction of a multi-hazard chain by an integrated numerical simulation approach: the Baige landslide, Jinsha River, China," *Landslides*, vol. 17, pp. 147–164, 2020.
[59] J.-Y. Park, S.-R. Lee, D.-H. Lee, Y.-T. Kim, and J.-S. Lee, "A regional-scale landslide early warning methodology applying statistical and physically based approaches in sequence," *Eng Geol*, vol. 260, p. 105193, 2019.

[60] T. Salvatici *et al.*, "Application of a physically based model to forecast shallow landslides at a regional scale," *Natural Hazards and Earth System Sciences*, vol. 18, no. 7, pp. 1919–1935, 2018.
[61] V.-H. Nhu *et al.*, "Landslide susceptibility mapping using machine learning algorithms and remote sensing data in a tropical environment," *Int J Environ Res Public Health*, vol. 17, no. 14, p. 4933, 2020.
[62] B. Hariharan and R. Guntha, "Crowdsourced landslide tracking-lessons from field experiences of landslide tracker mobile app," in *EGU General Assembly Conference Abstracts*, 2021, pp. EGU21-12711.
[63] R. Guntha, S. N. Rao, and M. V Ramesh, "Deployment experiences with Amrita Kripa: a user-friendly feature rich crowdsourced humanitarian application," *Procedia Comput Sci*, vol. 171, pp. 302–311, 2020.
[64] C. He, N. Ju, Q. Xu, and J. Huang, "Automated data processing and integration of large multiple data sources in geohazards monitoring," *International Journal of Georesources and Environment-IJGE (formerly Int'l J of Geohazards and Environment)*, vol. 3, no. 1–2, pp. 9–21, 2017.
[65] T. F. Fathani, D. Karnawati, and W. Wilopo, "An adaptive and sustained landslide monitoring and early warning system," in *Landslide Science for a Safer Geoenvironment: Volume 2: Methods of Landslide Studies*, Springer, 2014, pp. 563–567.
[66] T. Hemalatha, M. V. Ramesh, and V. P. Rangan, "Adaptive learning techniques for landslide forecasting and the validation in a real world deployment," in *Advancing Culture of Living with Landslides: Volume 5 Landslides in Different Environments*, Springer, 2017, pp. 439–447.
[67] L. Zhang and P. Lin, "Reinforcement learning based energy-neutral operation for hybrid EH powered TBAN," *Future Generation Computer Systems*, vol. 140, pp. 311–320, 2023.
[68] A. Murad, F. A. Kraemer, K. Bach, and G. Taylor, "Autonomous management of energy-harvesting iot nodes using deep reinforcement learning," in *2019 IEEE 13th International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, IEEE, 2019, pp. 43–51.
[69] A. Omidkar, A. Khalili, H. H. Nguyen, and H. Shafiei, "Reinforcement-learning-based resource allocation for energy-harvesting-aided D2D communications in IoT networks," *IEEE Internet Things J*, vol. 9, no. 17, pp. 16521–16531, 2022.
[70] M. Chu, X. Liao, H. Li, and S. Cui, "Power control in energy harvesting multiple access system with reinforcement learning," *IEEE Internet Things J*, vol. 6, no. 5, pp. 9175–9186, 2019.
[71] F. A. Aoudia, M. Gautier, and O. Berder, "RLMan: An energy manager based on reinforcement learning for energy harvesting wireless sensor networks," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 408–417, 2018.
[72] Y. Li, "Deep reinforcement learning: An overview," *arXiv preprint arXiv:1701.07274*, 2017.
[73] Y. Fei, Z. Yang, Y. Chen, and Z. Wang, "Exponential bellman equation and improved regret bounds for risk-sensitive reinforcement learning," *Adv Neural Inf Process Syst*, vol. 34, pp. 20436–20446, 2021.
[74] J. Clifton and E. Laber, "Q-learning: Theory and applications," *Annu Rev Stat Appl*, vol. 7, no. 1, pp. 279–301, 2020.
[75] F. Liu, L. Viano, and V. Cevher, "Understanding deep neural function approximation in reinforcement learning via $\epsilon $-greedy exploration," *Adv Neural Inf Process Syst*, vol. 35, pp. 5093–5108, 2022.
[76] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE Trans Cybern*, vol. 50, no. 9, pp. 3826–3839, 2020.
[77] European Commission, "European Commission Photovoltaic Geographic Information System." Accessed: Nov. 09, 2024. [Online]. Available: https://re.jrc.ec.europa.eu/pvg_tools/en/
[78] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
[79] T. P. Lillicrap, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
[80] F. Ait Aoudia, M. Gautier, and O. Berder, "RLMan: An Energy Manager Based on Reinforcement Learning for Energy Harvesting Wireless Sensor Networks," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 408–417, Jun. 2018, doi: 10.1109/TGCN.2018.2801725.
[81] A. Murad, F. A. Kraemer, K. Bach, and G. Taylor, "Autonomous management of energy-harvesting iot nodes using deep reinforcement learning," in *2019 IEEE 13th International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, IEEE, 2019, pp. 43–51.
[82] N. Charef, M. Abdelhafidh, A. Ben Mnaouer, K. Andersson, and S. Cherkaoui, "RL-Based Adaptive Duty Cycle Scheduling in WSN-Based IoT Nets," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*, IEEE, 2023, pp. 3777–3782.
[83] A. F. E. Abadi, S. A. Asghari, M. B. Marvasti, G. Abaei, M. Nabavi, and Y. Savaria, "RLBEEP: Reinforcement-Learning-Based Energy Efficient Control and Routing Protocol for Wireless Sensor Networks," *IEEE Access*, vol. 10, pp. 44123–44135, 2022, doi: 10.1109/ACCESS.2022.3167058.