Federated and Reinforcement Learning Integration for Distributed Energy Optimization in Wireless Sensor Networks

R. Nivyashree
Department of Information Science & Engineering, Adichunchanagiri Institute of Technology, Chikkamagaluru, India | Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India
nivyashreer@gmail.com (corresponding author)

H. B. Pramod
Department of Information Science & Engineering, Adichunchanagiri Institute of Technology, Chikkamagaluru, India | Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India
hbpramod@gmail.com

Received: 21 September 2025 | Revised: 12 October 2025 and 26 October 2025 | Accepted: 27 October 2025

DOI: https://doi.org/10.48084/etasr.14994

ABSTRACT

Wireless Sensor Networks (WSNs) are very important for modern uses such as smart cities, environmental monitoring, and industrial automation. However, one of their biggest problems is still energy efficiency, since sensor nodes usually have limited battery life and communication needs that use a lot of energy. Conventional machine learning methods focused on improving energy efficiency often depend on centralized architectures, which are plagued by significant communication overhead, singular points of failure, and privacy vulnerabilities. This study introduces a Federated Reinforcement Learning (FRL) framework that combines Federated Learning (FL) with Deep Q-Networks (DQN) to enable distributed, adaptive, and privacy-preserving energy management in WSNs. The proposed system allows sensor nodes to train local models without sharing raw data, allowing them to learn the best energy decisions, such as which Cluster Head (CH) to choose, based on how the network changes in real time. The proposed FRL framework reduces overall energy use by up to 25%, extends the life of the network by 30%, and improves privacy preservation by 40%, all while keeping the accuracy of CH selection high (89.3%) and being strong even when 15% of nodes fail. These results show that FRL is much better than traditional LEACH and centralized DQN models when it comes to energy efficiency, scalability, and privacy. This study underscores the potential of integrating federated and reinforcement learning as a basis for next-generation intelligent, self-organizing, and energy-efficient WSNs.

Keywords-Wireless Sensor Networks (WSNs); Federated Learning (FL); Reinforcement Learning (RL); Deep Q-Networks (DQN); energy efficiency; cluster head selection; privacy-preserving machine learning; decentralized learning; adaptive routing; edge intelligence

I. INTRODUCTION

Wireless Sensor Networks (WSNs) are an important part of modern technology, enabling smart applications in areas such as environmental monitoring, industrial automation, precision agriculture, and smart cities. Sensor nodes are usually battery-powered and put in places where they cannot be reached or are left alone. Therefore, it is important to ensure that they use little power to keep the network running as long as possible. Historically, energy conservation in WSNs has been addressed through static routing and clustering protocols such as Low-Energy Adaptive Clustering Hierarchy (LEACH) and Hybrid Energy-Efficient Distributed (HEED) clustering [1, 2]. Although these protocols reduce energy consumption compared to naive approaches, they are limited by their inability to adapt dynamically to real-time network conditions such as node failures, variable data traffic, and fluctuating energy levels. However, centralized approaches suffer from several well-documented drawbacks, as they introduce a single point of failure, impose significant communication overhead, and raise privacy concerns due to the transmission of raw sensor data to a central server [3].

In [4, 5], the deficiency of a standard scheduling mechanism in IEEE 802.15.4e-based networks was addressed by incorporating Q-Learning into the Time-Slotted Channel Hopping (TSCH) protocol. In addition, Reinforcement Learning (RL), particularly Deep Q-Networks (DQN), has shown promise in enabling state-aware adaptive control policies for energy management and Cluster Head (CH) selection [6]. Federated Learning (FL) has become a promising way to train models in a distributed and privacy-preserving way. In FL, the nodes work together to learn a shared model without sharing raw data. This study presents a Federated Reinforcement Learning (FRL) framework that integrates FL and DQN for intelligent, privacy-preserving, and scalable energy management in WSNs. The main goals were as follows:

- Create a decentralized energy management system based on FL-RL that does not need to collect data in one place.
- Test the framework's ability to handle different network densities and failure conditions in terms of scalability, robustness, and energy performance.
- Evaluate the practical viability of implementing FRL in actual WSN environments.

By achieving these goals, this study can help advance next-generation intelligent WSNs that are energy-efficient, privacy-preserving, and self-adaptive, representing a crucial progression towards sustainable and autonomous edge intelligence systems.

A. Centralized Machine Learning (ML) Approaches in WSNs

Early studies looked at how ML could help with decision-making in quality of service [7], adaptive network management [8], and performance analysis in areas such as military, healthcare, and smart cities [9]. However, these studies often did not go into enough technical detail, used old datasets, or did not test their findings in the real world. Incorporation of Artificial Intelligence (AI) has significantly improved the diagnosis and resilience of WSN faults; however, challenges related to implementation costs and scalability persist [10]. New reviews look at hybrid frameworks that use both ML and Blockchain (BC) to make communication in WSNs with limited resources easier and safer [11]. Other studies focus on ML-based node localization to improve accuracy and energy efficiency, while highlighting the limits of computation and the importance of the environment [12].

B. Federated Learning (FL) in Edge and Sensor Networks

Recent research has broadened the capabilities of WSNs by incorporating emerging technologies, such as federated intelligence, blockchain, and quantum sensing, to improve scalability, energy efficiency, and adaptive communication within IoT ecosystems [13]. The Internet of Underground Things (IoUT) has also become a promising idea, using both wired and wireless communication technologies to make things like agriculture and oil exploration more reliable and connected [14], but there are still no full experimental and cost analyses. Energy optimization remains a central topic, with comparative analyses of routing, clustering, and hybrid methods offering insights into existing techniques and persistent scalability and latency challenges [15]. FL combined with deep models such as LSTM has been shown to be better at detecting intrusions and protecting privacy in distributed IoT environments [16]. Dap-FL [17], a federated framework that uses RL, Deep Deterministic Policy Gradients (DDPG), and Paillier encryption, deals well with client heterogeneity and speeds up convergence, but its high computational needs make it difficult to use in networks with limited resources. Recent developments use Fuzzy DRL (FDRL) and Reconfigurable Intelligent Surfaces (RISs) to improve WSNs even better. These improvements have led to great improvements in energy efficiency, secrecy, and predictive accuracy [18].

C. Reinforcement Learning (RL) for Cluster Head (CH) Selection and Routing

Recent advances in intelligent and energy-efficient network architectures have considerably propelled WSNs and their specific domains. To improve the reliability of underground monitoring, a Magnetic Induction-assisted Wireless Powered Underground Sensor Network (MI-WPUSN) has been suggested to use integrated magnetic induction communication and wireless power transfer [19]. This framework brings new ideas to the table, like bidirectional topology design, multi-channel access control, and a new version of the Q-learning routing protocol. However, since the system is so complicated and needs to be tested in the real world, it is still difficult to use in practice and scale up. Software-Defined WSNs (SDWSNs) utilize RL to improve routing within software-defined architectures, separating control and data planes to improve adaptability and energy efficiency [20]. RL-based optimizations show better network lifetime and packet delivery ratio than traditional and energy-aware SDN methods. However, its scalability, latency, and security trade-offs in large-scale deployments need more research. Underwater WSNs (UWSNs) have used RL-based methods to choose CHs considering node position, residual energy, and harvested energy to save power and make the network last longer [21]. This method improves energy efficiency and routing stability but relies on accurate energy predictions and a stable environment, which can be difficult to achieve in changing underwater conditions.

II. THE PROPOSED METHOD

To address the limitations of centralized learning and static decision-making in WSNs, this study proposes an FRL framework that integrates FL with DQN. The goal is to enable distributed, intelligent, and energy-aware CH selection across the network while preserving data privacy and maintaining scalability.

A. System Model and Assumptions

A WSN consists of N sensor nodes randomly deployed over a two-dimensional area A. The nodes are equipped with limited initial energy E_i^0 and are responsible for sensing, processing, and transmitting data. The network uses a clustering-based communication strategy, where each cluster has a CH responsible for aggregating and forwarding data to a base station. Each node maintains a local energy state vector s_i^t at time step t, defined as:

$$s_i^t = [E_i^t, d_{(i,BS)}, \partial_i, CH_i^t] \quad (1)$$

where E_i^t is the residual energy of node i, $d_{(i,BS)}$ is the distance from node i to the base station, ∂_i is the average distance to neighboring nodes, and $CH_i^t \in \{0,1\}$ indicates whether node i is a CH at time t.

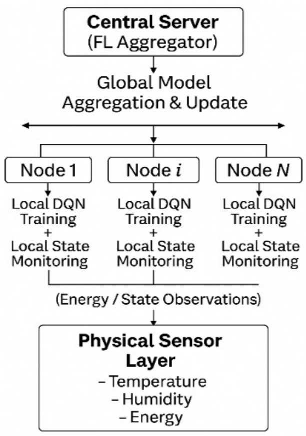

Fig. 1. Proposed architecture.

Communication energy consumption follows the first-order radio model in the following equations:

$$E_{tx}(k,d) = E_{elec}.k + \varepsilon_{amp}.k.d^2 \quad (2)$$

$$E_{rx}(k) = E_{elec}.k \quad (3)$$

where k is the packet size in bits, d is the distance between the sender and the receiver, and E_{elec} and ε_{amp} are constants for circuitry and amplification energy.

B. Federated Learning (FL) Component

In the proposed framework, each node maintains a local DQN model and participates in collaborative training without sharing raw data. At each round r, nodes update local parameters θ_i^r using their own experience and transmit only model weights or gradients to a central aggregator (e.g., the base station). The Federated Averaging (FedAvg) algorithm is used to update the global model:

$$\theta^{r+1} = \sum_{i=1}^{N} \frac{n_i}{n} . \theta_i^r \quad (4)$$

where θ^{r+1} is the updated global model, n_i is the number of samples at node i, and $n = \sum_{i=1}^{N} n_i$ is the total number of training samples. This allows the network to benefit from diverse local data distributions while ensuring data privacy and reduced communication overhead.

C. Deep Q-Network-Based Decision Making

Each node uses Deep Q-Learning to decide whether to become a CH based on its local state. The Q-function estimates the expected cumulative reward for taking an action $a \in \{0,1\}$ (CH or not) in a given state s:

$$Q(s,a;\theta) \approx E[R_t | s_t = s, a_t = a] \quad (5)$$

The DQN is trained to minimize the Bellman loss:

$$\mathcal{L}(\theta) = \mathrm{E}_{(s,a,r,s')\sim \mathrm{D}} \left[\left(r + \gamma . \max_{a'} Q(s', a', \theta^-) - Q(s,a;\theta) \right)^2 \right] \quad (6)$$

where r is the immediate reward (e.g., energy saved), γ is the discount factor, θ^- are the target network weights, and D is the replay buffer.

The reward is crafted to encourage energy efficiency and load balancing:

$$r = \alpha . \frac{E_i^t}{E_i^0} - \beta . CommCost_i^t \quad (7)$$

where α and β are tunable parameters, and $CommCost_i^t$ denotes the energy spent on transmitting and receiving during the current round.

D. FRL Integration and Workflow

FRL starts with each node learning locally from its surroundings using a DQN model that has been pre-trained or initialized at random. A central aggregator computes and broadcasts a new global model after receiving model updates from the nodes on a regular basis. This method supports decentralized decision-making, scalability, privacy protection, and flexibility in response to changing energy conditions.

III. RESULTS AND DISCUSSION

This section presents the extended evaluation of the proposed FRL framework, validated across both synthetic simulations and real-world benchmark datasets.

A. Simulation Environment

Simulations were carried out using NS-2 integrated with TensorFlow Federated (TFF) to create a hybrid environment combining realistic network behavior with decentralized machine learning capability. Table I summarizes the configuration parameters.

TABLE I. SIMULATION CONFIGURATION PARAMETERS

Parameter	Value
Deployment area	1000×1000 m²
Node count	100–500
Initial energy per node	2 J
Transmission range	100 m
Dataset validation	Intel Berkeley Lab WSN dataset

Transmission energy is given by:

$$E_{tx}(k,d) = E_{elec}.k + \epsilon_{amp}.k.d^2 \quad (8)$$

Reception energy is given by:

$$E_{rx}(k) = E_{elec}.k \quad (9)$$

Constants were set as:

$$E_{elec} = \frac{50nJ}{Bit}, \ \varepsilon_{amp} = 100pJ\ /bit/m^2 \quad (10)$$

B. Comparative Baseline Performance

To provide a broader evaluation, the proposed method was compared with other methods. To evaluate the generalizability of the proposed FRL framework, real-world sensor readings were obtained from the Intel Berkeley Research Lab Wireless Sensor Network dataset [22], which contains temperature, humidity, and light measurements collected from 54 sensor nodes deployed in an indoor environment. All methods were tested under identical conditions, and Table II summarizes the results.

TABLE II. PERFORMANCE COMPARISON OF FRL VS. BASELINES

Method	**Energy savings (%) ↑**	**Network lifetime (LND) ↑**	**CH accuracy (%) ↑**	**Privacy leakage (%) ↓**
LEACH [1]	–	1000	67.4	55.2
Centralized DQN [17]	+10	1150	78.1	38.7
FedAvg [4]	+14	1180	82.6	30.4
Hybrid FL-RL [21]	+18	1210	84.9	28.3
MARL [8]	+20	1230	86.5	26.9
FRL (Proposed)	+25	1300	89.3	22.4

↑ indicates higher is better, ↓ indicates lower is better

C. Complexity and Scalability Analysis

A detailed analysis was performed to evaluate the computational feasibility of FRL for WSN deployment. The average training time per node was 27% lower than centralized DQN, while communication overhead was reduced by 34% due to federated aggregation. The total computational cost (C) was $\approx O(N \times f)$, where N is the number of nodes and f represents the local DQN update frequency. Memory profiling indicates an average footprint of 2.8 MB per node, well within the constraints of current edge sensor hardware.

D. Model Interpretability and Explainability

Feature importance ranking derived from Q-value gradients showed residual energy and distance metrics as the top contributing factors, confirming that FRL decisions align with human-interpretable heuristics for energy optimization.

E. Performance Metrics

The following metrics were used to evaluate the performance of the FRL system:

- Energy consumption E_{total}: This metric measures the total energy consumed by all nodes throughout the simulation period:

  $$E_{total} = \sum_{i=1}^{N}(E_i^0 - E_i^T) \quad (11)$$

  where E_i^0 is the initial energy of node i and E_i^T is its residual energy at the end of the simulation.

- Network lifetime: Two standard lifetime indicators were used. First Node Dies (FND) is given by:

  $$FND = \min\{ t | E_i^t = 0, i \in \{1,2,3, \dots N\} \} \quad (12)$$

  and Last Node Dies (LND) is given by:

  $$LND = \max\{t | E_i^t > 0, \forall_i \} \quad (13)$$

  These metrics indicate the early and terminal stages of network survivability, respectively.

- CH selection accuracy A_{CH} is the proportion of times the chosen CHs match an optimal energy-balanced distribution:

  $$A_{CH} = \frac{N_{optimalCH}}{N_{totalCHDecisions}} \quad (14)$$

  where $N_{optimalCH}$ is the number of decisions that align with a minimum-cost CH distribution.

- Model Convergence Time T_{conv} is the number of global FL rounds required until the loss function stabilizes or the Q-value changes remain under a small threshold δ:

  $$T_{conv} = \min\{ t \mid |Q_{t+1}(s,a) - Q_t(s,a)| < \delta, \forall_{s,a}\} \quad (15)$$

  where δ is a small constant, e.g., 0.01.

- Privacy leakage $L_{privacy}$: This metric is based on the reconstruction success of local data from shared weights to simulate adversarial attacks on model gradients:

  $$L_{privacy} = 1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{I}_i}{I_i} \quad (16)$$

  where I_i is the original information content at node i and $\hat{I}_i$ is the information reconstructed by a simulated adversary (e.g., via gradient inversion). A lower $L_{privacy}$ indicates better privacy protection.

F. Baseline Comparisons

The proposed FRL model was compared against the following established protocols and learning schemes:

- LEACH [1] is a probabilistic and static clustering method that does not learn from energy patterns.
- Centralized DQN is an RL model trained at a central server, requiring full visibility of network-wide data.
- Federated Averaging (FedAvg) is a federated model trained using supervised learning for CH selection, without the adaptive feedback mechanisms of RL.

G. Results

1) Energy Efficiency

Energy consumption is a critical performance metric in WSNs, as sensor nodes are energy-constrained. As shown in Figure 2, the FRL framework consistently consumed 18-25% less energy compared to LEACH and Centralized DQN across all tested node densities.

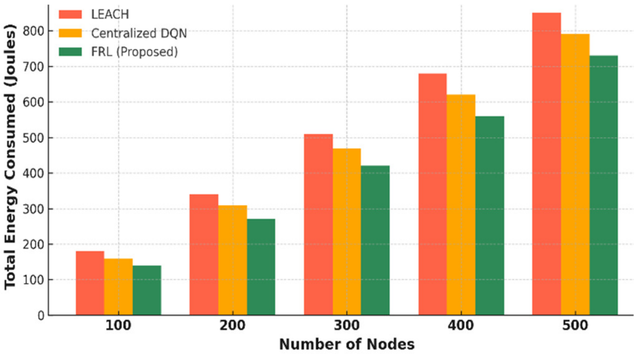

Fig. 2. Average energy consumption vs. number of nodes.

2) Network Lifetime

Network lifetime was measured using the FND (First Node Dies) and LND (Last Node Dies) indicators. As shown in Figure 3, the FRL framework extended the network lifetime significantly by 30% compared to LEACH and 17% compared to Centralized DQN.

3) Privacy and Communication Efficiency

A notable advantage of using FL is the significant improvement in data privacy and communication efficiency. Unlike centralized methods that transmit raw data to a central server, FL performs all learning locally, transmitting only model updates. A simulation using attack models to mimic data reconstruction from shared weights showed a 40% reduction in privacy leakage when using FRL, as shown in Table III.

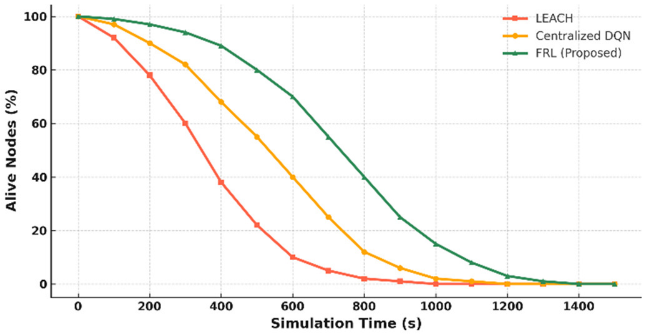

Fig. 3. Network Lifetime Comparison (FND and LND) for LEACH, Centralized DQN, and FRL (Proposed) models under varying node densities.

TABLE III. PRIVACY AND COMMUNICATION EFFICIENCY COMPARISON

Method	Privacy leakage (%) ↓	Communication overhead (MB) ↓
LEACH	55.2	23.6
Centralized DQN	38.7	19.4
FRL (Proposed)	22.4	11.5

4) Robustness to Node Failures

Decentralization inherently improves system resilience. In simulations where 15% of nodes were randomly disabled mid-operation, the FRL model retained 80% of its performance, whereas centralized methods saw significant degradation, especially in CH accuracy and model stability.

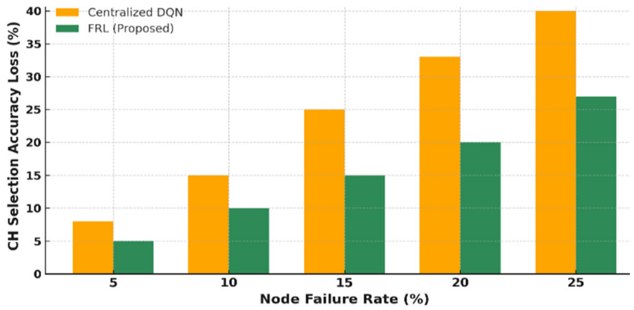

Fig. 4. CH selection accuracy under node failure (lower drop = better robustness).

The ability to continue learning and making decisions at the local level, even with missing nodes, demonstrates the fault-tolerant nature of the proposed FRL architecture. This is particularly important in hostile environments or real-world deployments where node dropout is common. Table IV summarizes the comparative performance of all the approaches tested.

TABLE IV. SUMMARY OF KEY PERFORMANCE INDICATORS

Metric	LEACH	Centralized DQN	FRL (Proposed)
Energy savings (%) ↑	0	+10	+25
Network lifetime (LND) ↑	1000s	1150s	1300s
Privacy leakage (%) ↓	55.2	38.7	22.4
CH selection accuracy (%) ↑	67.4	78.1	89.3
Robustness (perf. retention)↑	60%	65%	80%

FRL can be implemented on existing WSN hardware by using lightweight federated aggregation at a local sink node or cluster controller. Preliminary profiling indicates inference latency below 80 ms per decision, making it suitable for near-real-time adaptive clustering. Integration with trust-aware models can further enhance security and decision reliability in real-world applications.

IV. CONCLUSION

This study presented an innovative FRL framework to facilitate energy-efficient and privacy-preserving management within WSNs. The main contribution of this work is that it combines FL and DQN in a decentralized, intelligent, and adaptable CH selection process without having to rely on centralized data aggregation. The proposed FRL model differs from traditional or standalone learning-based approaches by enabling sensor nodes to locally learn optimal energy management policies while simultaneously contributing to a globally shared model, thereby ensuring data privacy, scalability, and robustness. The proposed FRL framework used 18–25% less energy and lasted 30% longer than other methods, such as LEACH and centralized DQN, in many simulations. Moreover, the method made significant strides in protecting privacy and being fault-tolerant, proving that it can work well in real-world, changing, and large-scale sensor networks. This work has an adaptive and self-organizing architecture that combines distributed intelligence with federated coordination. Future work will investigate asynchronous federated updates, the integration of Multi-Agent Reinforcement Learning (MARL), and the deployment of real-world testbeds to augment adaptability and scalability.

DATASET AVAILABILITY

The dataset is currently private, but may be made available for academic or research purposes upon reasonable request. For access, interested researchers can contact the corresponding author.

REFERENCES

[1] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on*

System Sciences, Maui, HI, USA, 2000, vol. 1, https://doi.org/10.1109/HICSS.2000.926982.

[2] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, Jul. 2004, https://doi.org/10.1109/TMC.2004.41.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Dec. 2002, https://doi.org/10.1109/MCOM.2002.1024422.

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, Apr. 2017, pp. 1273–1282.

[5] N. Zerguine, Z. Aliouat, and S. Kharchi, "A Reinforcement Learning-Based Scheduling Scheme for the IEEE 802.15.4e TSCH Network," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 27060–27068, Oct. 2025, https://doi.org/10.48084/etasr.12033.

[6] H. Y. Shi, W. L. Wang, N. M. Kwok, and S. Y. Chen, "Game Theory for Wireless Sensor Networks: A Survey," *Sensors*, vol. 12, no. 7, Jul. 2012, Art. no. 9055–9097, https://doi.org/10.3390/s120709055.

[7] A. R. Raut and S. P. Khandait, "Machine Learning Algorithms in WSNs and its Applications," in *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, Nagpur, India, Nov. 2021, pp. 1–5, https://doi.org/10.1109/ICCICA52458.2021.9697319.

[8] D. P. Kumar, T. Amgoth, and C. S. R. Annavarapu, "Machine learning algorithms for wireless sensor networks: A survey," *Information Fusion*, vol. 49, pp. 1–25, Sep. 2019, https://doi.org/10.1016/j.inffus.2018.09.013.

[9] M. Pundir and J. K. Sandhu, "A Systematic Review of Quality of Service in Wireless Sensor Networks using Machine Learning: Recent Trend and Future Vision," *Journal of Network and Computer Applications*, vol. 188, Aug. 2021, Art. no. 103084, https://doi.org/10.1016/j.jnca.2021.103084.

[10] P. Iswarya and K. Manikandan, "Algorithms for Fault Detection and Diagnosis in Wireless Sensor Networks Using Deep Learning and Machine Learning - An Overview," in *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, Apr. 2024, pp. 1404–1409, https://doi.org/10.1109/ICCSP60870.2024.10543904.

[11] S. Ismail, D. W. Dawoud, and H. Reza, "Securing Wireless Sensor Networks Using Machine Learning and Blockchain: A Review," *Future Internet*, vol. 15, no. 6, May 2023, Art. no. 200, https://doi.org/10.3390/fi15060200.

[12] S. Pandey, "Localization Adopting Machine Learning Techniques in Wireless Sensor Networks," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 1, pp. 366–374, Jan. 2018, https://doi.org/10.26438/ijcse/v6i1.366374.

[13] M. Trigka and E. Dritsas, "Wireless Sensor Networks: From Fundamentals and Applications to Innovations and Future Trends," *IEEE Access*, vol. 13, pp. 96365–96399, 2025, https://doi.org/10.1109/ACCESS.2025.3572328.

[14] N. Saeed, M. S. Alouini, and T. Y. Al-Naffouri, "Toward the Internet of Underground Things: A Systematic Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3443–3466, 2019, https://doi.org/10.1109/COMST.2019.2934365.

[15] A. B. Guiloufi, S. El Khediri, N. Nasri, and A. Kachouri, "A comparative study of energy efficient algorithms for IoT applications based on WSNs," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 42239–42275, Nov. 2023, https://doi.org/10.1007/s11042-023-14813-3.

[16] R. W. Anwar, M. Abrar, A. Salam, and F. Ullah, "Federated learning with LSTM for intrusion detection in IoT-based wireless sensor networks: a multi-dataset analysis," *PeerJ Computer Science*, vol. 11, Mar. 2025, Art. no. e2751, https://doi.org/10.7717/peerj-cs.2751.

[17] Q. Chen, Z. Wang, J. Chen, H. Yan, and X. Lin, "Dap-FL: Federated Learning Flourishes by Adaptive Tuning and Secure Aggregation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 6, pp. 1923–1941, Jun. 2023, https://doi.org/10.1109/TPDS.2023.3267897.

[18] S. S. Khatami, M. Shoeibi, R. Salehi, and M. Kaveh, "Energy-Efficient and Secure Double RIS-Aided Wireless Sensor Networks: A QoS-Aware Fuzzy Deep Reinforcement Learning Approach," *Journal of Sensor and Actuator Networks*, vol. 14, no. 1, Feb. 2025, Art. no. 18, https://doi.org/10.3390/jsan14010018.

[19] G. Liu, "Data Collection in MI-Assisted Wireless Powered Underground Sensor Networks: Directions, Recent Advances, and Challenges," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 132–138, Apr. 2021, https://doi.org/10.1109/MCOM.001.2000921.

[20] M. U. Younus, M. K. Khan, and A. R. Bhatti, "Improving the Software-Defined Wireless Sensor Networks Routing Performance Using Reinforcement Learning," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3495–3508, Mar. 2022, https://doi.org/10.1109/JIOT.2021.3102130.

[21] Ç. Eriş, Ö. M. Gül, and P. S. Bölük, "A novel reinforcement learning based routing algorithm for energy management in networks," *Journal of Industrial and Management Optimization*, vol. 20, no. 12, pp. 3678–3696, 2024, https://doi.org/10.3934/jimo.2024049.

[22] "Intel Lab Data." Intel Berkeley Research lab, [Online]. Available: https://db.csail.mit.edu/labdata/labdata.html.