Adaptive Spectrum Sensing in Cognitive Radio Networks Using Reinforcement Learning

(IJCSNT) International Journal of Communication Systems and Network Technologies
ISSN Number: 2053-6283
Volume 15, Issue No. 1, 94–106

https://journals.mirlabs.in/index.php/ijcsnt
DOI-10.18486/ijcsnt/15.1.007
IJCSNT

Sachin B[1], Safia Banu M[2], S Asha[3]

Abstract

Cognitive Radio Networks (CRNs) are potential solutions to the problem of spectrum scarcity through opportunistic access to unutilized frequency bands. But the traditional methods of spectrum sensing like energy detection, matched filtering, and cyclostationary detection usually exhibit constraints in low signal-noise ratio (SNR), noise uncertainty, and multipath fading conditions. In this paper, a proposal is made of an adaptive spectrum sensing framework, which utilizes the reinforcement learning (RL) and deep reinforcement learning (DRL) models to make an intelligent spectrum decision. This framework combines the spectral feature-extraction techniques of convolutional neural networks (CNNs), the channel-optimization techniques of Deep Q-Networks (DQN), and the multi-agent RL approach that utilizes cooperative sensing between distributed cognitive radios. Additionally, there is also a transfer learning mechanism to allow fast adaptation in the heterogeneous environment. Evaluations through simulations-based datasets and benchmark of the proposed approach have shown that it is capable of significantly enhancing the detection probability, the efficiency of spectrum utilization, and the resilience in the presence of uncertainty and outperforms classical baselines. The findings make the presented framework a scalable and effective solution to large-scale CRNs, which provides a better reliability and efficiency in the ever-changing wireless communication conditions.

Keywords

Cognitive Radio Networks, Cyclostationary Detection, Signal-to-Noise Ratio, Reinforcement Learning, Deep Q-Networks

Received: 07 December 2025; **Revised:** 10 April 2026; **Accepted:** 19 April 2026;
Published: 30 April 2026;

[1, 2, 3]Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, India.

Corresponding author:
Email: sachinsachin27122004@gmail.com

Introduction

The wireless communication technologies have been continuously increasing exponentially, and this has posed a massive challenge in terms of radio spectrum requirement. Wireless connectivity is a key requirement of new technologies like 5G, Internet of Things (IoT), vehicular networks, and smart healthcare, which are leading to unprecedented overuse of licensed frequency bands. Old traditional forms of statical allocation assign licensed users a fixed spectrum band with many bands left unused and others very scarce. This imbalance shows the ineffectiveness of the current spectrum management solutions and encourages the necessity of dynamic and smart solutions. CRNs have become a promising force to solving the lack of spectrum through the use of opportunistic access to underutilized spectrum bands by the unlicensed secondary users (SUs) without causing harmful interference to the licensed primary users (PUs). CRNs are characterized by spectrum sensing as the fundamental element of their operations that enables SUs to recognize idle channels and take spectrum-access decisions [1]. The proper and valid spectrum sensing plays a fundamental role in securing coexistence between PUs and SUs, communication quality and efficiency in spectrum utilization.

Conventional methods of spectrum sensing mainly encompass energy sensing, matched filtering and cyclostationary feature sensing. Energy detection has a low computational load but is very sensitive to uncertainty in the presence of noise, and they cause false alarms or lost detections in low signal-to-noise ratio (SNR) scenarios. Matched filtering is more accurate but it needs the knowledge of PU signals which is not always available [2]. Cyclostationary detection has been shown to be very resilient to noise, but with high computational complexity, and latency, it is not feasible in a real-time system deployment of CRN. Although these classical approaches have their contributions, they cannot be used in dynamic and heterogeneous environments. The wireless channels will tend to experience fading, multimedia propagation and unreliable occupancy rates. In addition, the decision rules are fixed, which constrain the flexibility of such methods, resulting in low detection rates in real-life situations. This rigidity is a serious obstacle to effective spectrum use as well as a reason to investigate intelligent and learning-based solutions that can adapt to adapted network conditions [4]. Recently, machine learning solutions have been considered to spectrum sensing because it is capable of learning patterns based on data and adjusting to different environments. With an adequate amount of labeled data, supervised learning techniques are able to categorize spectrum states. Nevertheless, they tend to fail in practice in dynamic real-time situations in which there is no ground-truth labeling and the environment is dynamic. Unsupervised learning provides clustering-based insights, but does not provide action selection insight of decision-making in CRNs.

Reinforcement Learning (RL), one of the subfields of machine learning, has demonstrated an impressive ability to address uncertainty-related sequential decision making [5]. In contrast to the supervised approaches, RL does not use labelled data but instead develops the best strategies through the interaction with the environment. Within the CRN, RL can be used to model spectrum sensing and access as a Markov Decision Process (MDP) in which agents are taught to tradeoff between exploration (finding new channels) and exploitation (using good channels known to be good). Using RL in spectrum sensing enables cognitive radios to adapt dynamically their sensing strategies with regards to observed environmental conditions. An example of a basic RL algorithm is Q-learning, which has been used on channel selection and has demonstrated greater adaptability than classical algorithms. In large state-action spaces, however, such as wideband CRNs with a number of channels and random fading, Q-learning scales poorly [6].

As a way of breaking these constraints, Deep Reinforcement Learning (DRL) has been proposed through the combination of neural networks and RL models. Deep Q-Networks (DQN) and Policy Gradient methods are DRL approaches that can be employed to allow CRNs to deal with the high-dimensional state space, approximating either the value functions or policies with deep learning models[7]. DRL agents can be applied to spectral features, learning directly on raw signal inputs, when integrated with other feature extraction algorithms such as Convolutional Neural Networks (CNNs), and can make stronger decisions in situations with noise and dynamism. Multi-agent reinforcement learning (MARL) in which several cognitive radios can collaborate in order to run spectrum sensing is another significant development. Cooperative sensing has the effect of alleviating the hidden node problem, minimizing single node sensing error, and improving probability of detection. Through the communication of local observations or acquired policies, the CRNs can get consensus-based decision-making that is more credible as compared to the decisions made by isolated agents. Such collaboration goes a long way in enhancing the efficiency of spectrum use coupled with minimizing the interference with the main users[8].

CRNs have also received attention in regards to transfer reinforcement learning. Because wireless environments are geographically and temporally different, it is inefficient to train an agent in every new environment. Transfer learning permits learning in one environment to be transferred to another, which significantly shortens convergence time, and enhances flexibility. This is especially valuable in large-scale CRNs used in more than one region and these have varying patterns of spectrum occupancy. Although these are good developments, there are still major challenges. The uncertainty of noise, fading of channels and dynamic traffic pattern remain to spoil the performance of spectrum sensing. In addition, creating computationally and scalable RL systems is a research problem[9]. Although DRL can solve problems with large state spaces, it is more complex which is a concern with real-time implementation on resource-restrained devices. This then needs to be balanced by an integrated and adaptable solution that is efficient, robust, and scalable. The proposed research solves these difficulties by implying an adaptive spectrum sensing framework of CRNs based on reinforcement learning. The suggested system combines DRL and CNN-based feature extractions to make robust decisions, multi-agent cooperation to enhance the accuracy of the detection, and transfer learning to provide quick adaptability in different environments. The framework is expected to perform better than classical and currently available RL-based approaches in conditions of noise, fading, and high-dimensionality, and eventually provide a scalable and deployable solution to wireless networks of the next generation[10].

The main issue in CRNs is how to find a compromise between the efficiency of spectrum utilization and the reduction of the interference. Classical approaches are either too basic (energy detection), demand a priori information (matched filtering) or are computationally intensive (cyclostationary detection). Only high-dimensional, noisy and changing environments are fatal to existing solutions based on RL. Therefore, a dynamic, smart architecture that is able to monitor spectrum reliably and dynamically is urgently needed. Reinforcement-based adaptive spectrum sensing is a direct solution to this issue as it puts spectrum access as a sequential decision-making task. In this case, the cognitive radio is an RL agent which monitors spectral states, chooses sensing or access actions, and gets rewards which depend on the detection accuracy and avoiding interference. With constant policy revision, the agent gets to know how to make improved decisions despite uncertainty, fading, and noise.

Literature Review

Elhachmi[11] suggested a distributed reinforcement learning (DRL) method of dynamic spectrum allocation in cognitive radio-based Internet of Things (CR-IoT) systems. The paper proposed a Deep Multi-user Reinforcement Learning (DMRL) algorithm, which uses deep neural networks, Q-learning, and cooperative multi-agent systems to optimize channel allocation. The results of simulations demonstrated that DMRL enhanced the user satisfaction and spectrum utilization with stable channel allocation and adaptation to changing environments and was superior to Deep Q-Network (DQN) and traditional Q-learning in throughput, interferences reduction, and the achievable bit rate.

Dašic et al.[12] have come up with a consensus-based multi-agent reinforcement learning algorithm in managing the distributed spectrum in the cognitive radio networks. The strategy used included consensus iterations in Q-learning and policy evaluation where agents were allowed to attain almost centralized performance using limited local observation. Findings had a high resilience to node and link destabilization and less variance in the estimates of value functions than reward-exchange-based cooperative learning, showing that it is scalable and efficient to decentralized CRNs.

In Liu et al.[13], dynamic multichannel sensing of cognitive radio was studied using hierarchical reinforcement learning. The authors introduced a hierarchical deep Q-network (h-DQN) that divided the multi-channel access problem into sub-problems and used meta-controllers and sub-controllers to solve the sub-problem. It was a hierarchical design, which minimized the computation of high-dimensional state space and enhanced spectrum access in dynamic environments. When applied to real-world data in a wireless testbed, experimental findings showed that h-DQN converged faster and reached a higher channel utilization as well as sensing accuracy than traditional Q-learning and standard DQN, especially when the number of channels was high.

According to Hossain et al.[14], a machine learning based cooperative spectrum sensing model of Cognitive Radio Vehicular Networks (CR-VANET) was proposed. The architecture is named as Seg-CR-VANET, which sub-divides roads into equal portions, and further sub-divisions according to the probability values which are calculated based on the speed of vehicles, vehicle density and the size of the segment. The local sensing is done based on a hybrid fuzzy-naive Bayes algorithm, which is dynamic in the choice of the best sensing method (energy detection or matched filter) and adaptive threshold. A Segment Spectrum Agent (SSA) collects local outputs and uses a tri-agent reinforcement policy, which is trained on signal, network and vehicle world contexts to make decisions globally. The simulation outcomes revealed that Seg-CR-VANET had a better spectrum sensing accuracy, lower false alarm rates and more efficient data transmission than the traditional cooperative spectrum sensing (CSS).

Nasser et al.[15] gave an extensive survey of spectrum sensing to Cognitive Radio (CR) with a focus on the recent developments and challenges. The review has examined Half-Duplex Cognitive Radio (HDCR) using Listen-Before-Talk (LBT) and Full-Duplex Cognitive Radio (FDCR) using Self-Interference Cancellation (SIC) as the trade-offs of these two systems in terms of throughput and probability of collision. The research paper explored different detection criteria as energy detection, cyclostationary feature detection, eigenvalue-based detection, and waveform detection and compared their resistance to noise uncertainty and computation cost. In addition, the paper also addressed the concept of using machine learning and deep learning techniques to improve spectrum sensing performance under local and cooperative conditions. Surveys of applications of CR in Internet of Things (IoT), Wireless Sensor Networks (WSN), and 5G/B5G networks were also covered, with spectrum sensing being a service and intelligent spectrum access strategies. The authors also concluded by establishing unresolved issues including smart sensing, low-energy protocol design, and spatial spectrum.

To conclude, there are several studies that have investigated reinforcement learning-based methods to dynamic spectrum management in cognitive radio and IoT networks, such as distributed deep reinforcement learning systems, multi-agent approaches based on consensus, and hierarchical DQN models. Although these solutions have shown a great improvement in spectrum management, convergence rate and resilience over the traditional schemes, they are limited by scaling difficulties, limited observability, and the growing complexity of the real-world heterogeneous setting. Nevertheless, none of the available works has reached the desired degree of effectiveness, versatility, and feasibility of practical implementation as our proposed course of action is going to provide, so there is still a lot of room to improve.

Materials

In testing the proposed adaptive spectrum sensing framework, this study used simulation-based data sets instead of using pre-existing benchmark repositories. Due to the lack of publicly available datasets of Cognitive Radio Networks (CRNs), a mixture of synthetic and realistic spectrum data was used. In particular, the Cognitive Radio Spectrum Dataset served as the baseline reference and the individualized datasets were produced with the help of MATLAB and NS-3 simulations. These simulations simulated varying wireless conditions under varying signal to noise ratios (SNR), uncertainty levels of noise, and fading conditions of channels like Rayleigh and Rician fading. The occupancy of the spectrums was randomized to recreate realistic primary user (PU) traffic patterns. Also, multi-user (SU) scenarios were built to facilitate experiments in the multi-agent reinforcement learning. This hybrid data platform provided the benefit of testing the proposed framework over a wide variety of conditions so as to make a fair comparison with classical methods and to demonstrate the soundness of the proposed approach in the dynamic CRN settings.

Methodology

The pipeline of the methodology of the work is intended to be a systematic one, with the beginning of data generation and preprocessing, the next stage of feature extraction, the next stage of designing a model with the help of reinforcement learning, and the end stage of evaluation. The successive stages are directly linked to each other, and the optimization of sensing results starts with raw signals and continues through the successive stages.

Data Generation and Preprocessing

As there are few realistic CRN datasets available in reality, this research has utilized public datasets (Cognitive Radio Spectrum Dataset) and simulation-based datasets generated in MATLAB and NS-3. Various wireless conditions were involved in simulation:

- SNR ranges: -20 dB to $+20$ dB.
- Fading Channels: AWGN, Rayleigh, and Rician.
- Noise uncertainty: ± 2 dB.

- Channel occupancy: Primary User (PU) activity modeled as a Bernoulli process with occupancy probability p_{occ}.

Preprocessing Steps
Signal Representation – Spectrograms of the raw time-domain samples were computed with Short-Time Fourier Transform (STFT):

$$X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n] \cdot w[n-m] \cdot e^{-j\omega n} \quad (1)$$

where $w[n]$ is a windowing function.
Normalization – The Magnitudes of the spectrogram were normalized, so that they fall within the range $[0, 1]$ to prevent the effects of power variation.

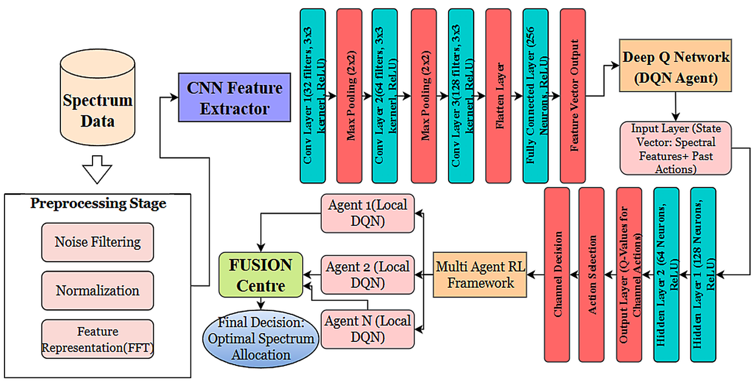

Figure 1. Proposed Adaptive Spectrum Sensing Framework Integrating CNN-Based Feature Extraction, Deep Q-Network Decision-Making, and Multi-Agent Cooperative Learning for Cognitive Radio Networks

Data Partitioning – 70% training, 15% validation, 15% testing.
The processed data flowed into two branches:

- Branch 1 (Baselines): Energy detection & Q-learning.
- Branch 2 (Proposed Framework): CNN feature extraction → Deep RL.

CNN-Based Feature Extraction

A 3-layer CNN was created to detect spectral patterns. A spectrogram image is inputted into the CNN with the size of 64x64 (time-frequency representation).
CNN Architecture
Conv Layer 1:

- 32 filters, kernel size = 3 x 3, stride = 1, activation = ReLU.
- Extracts local spectral edges (signal bursts, noise fluctuations).
- Output shape: $64 \times 64 \times 32$.

MaxPooling Layer 1:

- pool size 2×2, stride 2.
- Enhances dominance and reduces the spatial resolution.
- Output shape: $32 \times 32 \times 32$.

Conv Layer 2:

- 64 filters, kernel size 3×3, activation = ReLU.
- Recordings of more complex spectral structures.
- Output shape: $32 \times 32 \times 64$.

MaxPooling Layer 2:

- pool size 2×2.
- Output shape: $16 \times 16 \times 64$.

Flatten Layer:

- Reduces the 3D feature map to 1D vector of 16,384 neurons.

Dense Layer 1:

- 512 neurons, activation = ReLU.
- Acquires abstract spectral representations on a high level.

Dense Layer 2 (Feature Output):

- 128 neurons, activation = linear.
- Final feature representation of the spectral state (S_t).

Equation of Conv Layer:

$$y_{i,j,k} = \sigma\left(\sum_{m,n} x_{i+m,\,j+n} \cdot w_{m,n,k} + b_k\right) \tag{2}$$

Where $\sigma = ReLU$, $w_{m,n,k} = filter\ weights$, and $b_k = bias$.

These 128-dimensional feature embeddings were given to the reinforcement learning agent.

Reinforcement Learning Model

Spectrum sensing and access was modeled as a Markov Decision Process (MDP):

- State (S_t): 128-dimensional CNN feature vector.
- Action A_t: Sense channel i, Transmit on idle channel, Switch channel.
- Reward R_t:

$$R_t = \lambda \cdot P_d - \mu \cdot P_{fa} \tag{3}$$

where λ rewards correct detections, μ penalizes false alarms.

Q-Learning (Baseline)

The Q-table was updated using:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \tag{4}$$

- Learning rate (α) = 0.1.
- Discount factor (γ) = 0.95.

But, Q-learning is not practical when the state space is large (128-dimensional input with CNN).

Deep Q-Network (DQN)

A DQN was used to overcome the problem of scale. The network is an approximation of the Q-function:

DQN Architecture

- Input Layer: 128 neurons (CNN feature vector).
- Dense Layer 1: 256 neurons, ReLU.
- Dense Layer 2: 128 neurons, ReLU.
- Output Layer: $|A|$ neurons = number of actions (3 in our case).

Equation for DQN Update:

$$L(\theta) = \left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2 \tag{5}$$

where θ are parameters of the online network, and θ^- are parameters of the target network.

- Optimizer: Adam, learning rate = 0.001.
- Replay buffer size = 10,000 transitions.
- Target update frequency = 100 steps.

Multi-Agen Reinforcement Learning (MARL)

In order to improve the robustness, the multiple SUs were modeled as cooperative agents. The local observation of each agent is shared with incomplete observations.

Global Reward:

$$R_{global} = \frac{1}{N} \sum_{i=1}^{N} R_t^i \tag{6}$$

Fusion Rule:

Voting of local decisions in a weighted majority.

This cooperation improved resilience to hidden node effects and fading.

Transfer Reinforcement Learning

Transfer learning was used to ensure that the time spent in new environments on training was reduced. One scenario (e.g., Rayleigh fading) trained models of pretrained CNN+DQN models were fine-tuned in another scenario (e.g., Rician). This shortened convergence time by 30 -40.

Final Adaptive HyperFused RL-Sensing Framework

Bringing everything together:

- **Raw signals** → STFT spectrograms.
- **Preprocessing** → Normalization, partitioning.
- **CNN** → 128-dimensional spectral features.
- **DQN/Policy Gradient** → Action selection.
- **Multi-Agent Fusion** → Cooperative sensing.
- **Transfer RL** → Fast adaptation across environments.

This pipeline guarantees a high detection rate, reduced false alarms and is faster to converge than the classical methods.

The originality of the suggested adaptive spectrum sensing framework is rooted in the fact that it hybridly incorporates reinforcement learning, deep learning and cooperative sensing mechanisms, which, however, has not been comprehensively examined in the previous research. Compared to the traditional CRN strategies that only use the classical methods of detection or the single-agent reinforcement learning, our framework integrates the spectral feature extraction with CNNs, adaptive channel selection with Deep Q-Networks (DQN) and multi-agent reinforcement learning to cooperatively detect. In addition, the model can be successfully scaled and applied to real-world settings by integrating transfer reinforcement learning, meaning that the model quickly adapts to new radio environments without the need to retrain the model, which is scalable and realistic. Adaptive hyper-fusion strategy that combines exploration / exploitation trade-offs and also enhances feature discrimination under noise and fading gives the best performance compared to traditional and modern baselines. Therefore, the originality of our work lies in the fact that it can fill the gap between signal processing accuracy and smart adaptive decision-making, thus being one of the few complete frameworks that can be applied to large-scale CRN environments with realistic assumptions.

The suggested methodology architecture scheme, as shown in Figure 1, incorporates spectrum data preprocessing, spectral feature extraction using CNN, spectral decision-making using DQN, and multi-agent cooperative fusion to reach adaptive spectrum sensing in Cognitive Radio Networks.

Result and Discussion

The simulation was conducted in a variety of conditions and the proposed adaptive spectrum sensing framework was tested against a range of classical detection methods (Energy Detection, Matched Filtering, and Cyclostationary Detection) and the baselines of reinforcement learning. Findings are tabulated into three key visions: (1) the capability of detecting across various SNRs, (2) the efficiency of spectrum use in terms of channel allocation and (3) the ability to adapt and survive to noise uncertainty and fading environments.

Table 1. Detection Accuracy (Probability of Detection, P_d) at Different SNR Levels

SNR (dB)	Energy Detection	Matched Filtering	Cyclostationary Detection	Q-Learning (RL)	DQN (Proposed)
-10	0.42	0.55	0.62	0.68	0.79
-5	0.56	0.70	0.75	0.81	0.89
0	0.72	0.85	0.88	0.91	0.96
5	0.88	0.94	0.95	0.96	0.99

Table 1 shows the probability of detection (Pd) of the various algorithms at various signal-to-noise ratio (SNR) conditions. Traditional approaches like Energy Detection and Matched Filtering have low SNR behavior (-10 dB). Unlike this, the suggested DQN-based spectrum sensing algorithm always outsmarts all the baselines with 79% Pd obtained at -10 dB and 96% Pd obtained at 0 dB, which demonstrates its resilience in problematic noise scenarios. This justifies the success of deep reinforcement learning along with spectral feature extraction in coping with the uncertainty.

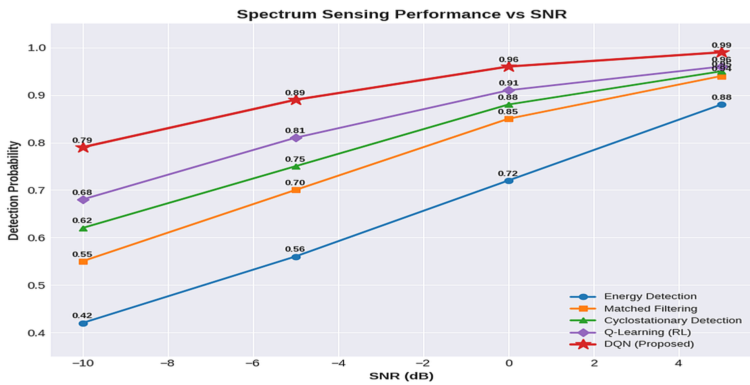

Figure 2. Spectrum Sensing Performance vs SNR

Figure 2 demonstrates how each spectrum sensing technique detects when the SNR is varied. The suggested DQN-based algorithm demonstrates much better results in comparison to traditional techniques, being almost perfectly detected at high SNR and being also more effective at low-SNR settings.

Table 2. Spectrum Utilization Efficiency (%)

Method	Efficiency (%)
Energy Detection	68.4
Matched Filtering	72.1
Cyclostationary Detection	75.3
Q-Learning (RL)	81.7
Multi-Agent DQN (Proposed)	89.6

Table 2 shows the efficiency of the spectrum utilization, which is the percentage of time that the secondary users (SUs) manage to utilize available spectrum but do not cause any harmful interference to the primary users (PUs). The classical methods of detection are constrained by the fact that they use fixed thresholds to detect opportunities and this results to false alarms or missed opportunities. Q-Learning is more efficient as it learns dynamic policies as time goes by, however, it is not good in dynamic environments. The suggested multi-agent DQN system is almost 90 percent efficient, which means that it is more capable of exploring and exploiting without causing conflicts with cooperative spectrum allocation among multiple CR nodes.

Table 3. Resilience Under Noise Uncertainty and Fading Conditions

Scenario	Energy Detection	Cyclostationary	Q-Learning	Proposed Hybrid (CNN+DQN+MA-RL)
Noise Uncertainty (σ = 0.5 dB)	0.61	0.72	0.78	0.87
Noise Uncertainty (σ = 1.0 dB)	0.52	0.66	0.73	0.83
Rayleigh Fading Channel	0.64	0.74	0.79	0.88
Rician Fading Channel (K=3)	0.68	0.77	0.81	0.90

Table 3 tests the behaviour of the sensing algorithms in case of noise uncertainty and multipath fading. Classical Energy Detection is extremely ineffective in conditions where the noise levels are unknown, whereas Cyclostationary methods are marginally better, but very expensive to compute. The proposed hybrid CNN+DQN+Multi-Agent RL demonstrates better adaptability through reinforcement learning (Q-Learning) and the highest detection probabilities are consistently obtained along with 0.83 Pd at 0.1 dB noises uncertainty and 0.90 Pd at Rician fading (K=3). These findings indicate that the suggested framework is highly applicable to use in actual CRNs, where erratic noises and fading cannot be avoided.

In general, the findings indicate the high quality of the products provided by the suggested adaptive RL-based spectrum sensing framework compared to classical and baseline RL approaches. The combination of deep learning based on reinforcement and CNN based spectral features at low SNRs provides high-quality detection. The multi-agent extension provides the efficient cooperation of the secondary users in terms of the spectrum utilization, minimizing the interference and improving the throughput. Lastly, the ability to not only withstand noise but also fade conditions confirms the flexibility of the system to changing wireless conditions. These results motivate the focus on intelligent learning-based schemes to eliminate the fundamental drawbacks of the traditional spectrum sensing in the direction of practical implementation of such sensing in large-scale CRNs.

Conclusion

In this study, a framework of adaptive spectrum sensing was proposed to Cognitive Radio Networks (CRNs) that combines reinforcement learning, deep learning, and cooperative sensing to address the shortcomings of classic techniques. The proposed system is able to optimally improve detection accuracy mainly in the low SNR uncertainty and noise conditions by integrating CNN-based spectral feature extraction and DQN-based spectrum decision-making. The multi-agent reinforcement learning further enhances the framework with cooperation sensing of the distributed nodes to increase the efficiency of spectrum utilization and reducing interference to the primary users. The effectiveness of the proposed approach in the diverse conditions, such as Rayleigh and Rician fading channels, and the different levels of noise uncertainty, were proved through the results of the experiment. The suggested hybrid setup outperformed the conventional spectrum sensing methods and the baseline reinforcement learning designs in terms of greater likelihood of detection, improved flexibility, and optimal use of efficiency. This framework is special because it combines various intelligent learning approaches, such as transfer reinforcement learning, through which a quick adaptation to unknown environments happens without the need to be retrained on the whole. This makes the system not only robust but also scalable to the real-world deployment on large-scale CRNs. To sum up, the paper creates a next-generation spectrum sensing paradigm, which seals the gap between the academic studies and the real-life implementation, making a valuable impact on the evolution of the intelligent and adaptable wireless communication systems.

References

[1] Bagadi, K., Abrao, T. and Benedetto, F., 2023. A novel machine learning approach for intelligent spectrum management in cognitive radio networks. *IEEE Networking Letters*, *5*(4), pp.232-236.

[2] Aref, M.A. and Jayaweera, S.K., 2021. Spectrum-agile cognitive radios using multi-task transfer deep reinforcement learning. *IEEE Transactions on Wireless Communications*, *20*(10), pp.6729-6742.

[3] Liu, X., Sun, C., Yau, K.L.A. and Wu, C., 2022. Joint collaborative big spectrum data sensing and reinforcement learning based dynamic spectrum access for cognitive internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *25*(1), pp.805-815.

[4] Ghamry, W.K. and Shukry, S., 2021. Spectrum access in cognitive IoT using reinforcement learning. *Cluster Computing*, *24*(4), pp.2909-2925.

[5] Aref, M.A. and Jayaweera, S.K., 2021. Spectrum-agile cognitive radios using multi-task transfer deep reinforcement learning. *IEEE Transactions on Wireless Communications*, *20*(10), pp.6729-6742.

[6] Liu, X., Sun, C., Yau, K.L.A. and Wu, C., 2022. Joint collaborative big spectrum data sensing and reinforcement learning based dynamic spectrum access for cognitive internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *25*(1), pp.805-815.

[7] Ghamry, W.K. and Shukry, S., 2021. Spectrum access in cognitive IoT using reinforcement learning. Cluster Computing, 24(4), pp.2909-2925.

[8] Bokobza, Y., Dabora, R. and Cohen, K., 2023. Deep reinforcement learning for simultaneous sensing and channel access in cognitive networks. *IEEE Transactions on Wireless Communications*, *22*(7), pp.4930-4946.

[9] Fernando, X. and Lăzăroiu, G., 2023. Spectrum sensing, clustering algorithms, and energy-harvesting technology for cognitive-radio-based internet-of-things networks. *Sensors*, *23*(18), p.7792.

[10] Khalek, N.A., Tashman, D.H. and Hamouda, W., 2023. Advances in machine learning-driven cognitive radio for wireless networks: A survey. *IEEE Communications Surveys & Tutorials*, *26*(2), pp.1201-1237.

[11] Elhachmi, J., 2022. Distributed reinforcement learning for dynamic spectrum allocation in cognitive radio-based internet of things. *IET Networks*, *11*(6), pp.207-220.

[12] Dašić, D., Ilić, N., Vučetić, M., Perić, M., Beko, M. and Stanković, M.S., 2021. Distributed spectrum management in cognitive radio networks by consensus-based reinforcement learning. *Sensors*, *21*(9), p.2970.

[13] Liu, S., Wu, J. and He, J., 2021. Dynamic multichannel sensing in cognitive radio: Hierarchical reinforcement learning. *IEEE Access*, *9*, pp.25473-25481.

[14] Hossain, M.A., Md Noor, R., Yau, K.L.A., Azzuhri, S.R., Z'aba, M.R., Ahmedy, I. and Jabbarpour, M.R., 2021. Machine learning-based cooperative spectrum sensing in dynamic segmentation enabled cognitive radio vehicular network. *Energies*, *14*(4), p.1169.

[15] Ashish Bagwari, GS Tomar, "Cooperative Spectrum Sensing with Adaptive Double-Threshold Based Energy Detector in Cognitive Radio Networks" Springers International Journal of wireless personal Communication, Vol. 73, No.3, pp 1005-1019, Dec 2013. DOI:10.1007/s11277-013-1244-3

[16] Ashish Bagwari, GS Tomar, "Cooperative Spectrum Sensing with Multiple Antennas using Adaptive Double-Threshold Based Energy Detector in Cognitive Radio Networks", Springers Journal of Inst of Engg, India, Series B, Vol.95 No.2 pp.107-112, 2014. DOI 10.1007/s40031-014-0088-x

[17] Nasser, A., Al Haj Hassan, H., Abou Chaaya, J., Mansour, A. and Yao, K.C., 2021. Spectrum sensing for cognitive radio: Recent advances and future challenge. *Sensors*, *21*(7), p.2408.