DRLNet: A Deep Reinforcement Learning Network for Hybrid Features Extraction and Spectrum Sensing in Cognitive Radio Networks

Usha Rani M A * and Prashanth C R

Department of Electronics & Telecommunication Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India;
Email: Prashanthcr.et@drait.edu.in (P.C.R.)
*Correspondence: Usharanima.et@drait.edu.in (U.R.M.A.)

***Abstract*—Over the past two decades, communication technologies have advanced significantly, but the growing use of various communication methods has led to a shortage of available spectrum. Cognitive Radio (CR) emerges as a solution to this challenge. A crucial aspect of CR is spectrum sensing, which detects available spectrum gaps. However, current spectrum sensing methods have limitations, including insufficient signal representation, inefficiency, and sensitivity to noise. To address these issues, this study embraces a deep learning approach and introduces an innovative spectrum detection architecture for cognitive radio networks. The method combines deep learning and reinforcement learning, leveraging deep learning for energy and energy correlation feature extraction. Additionally, a Recurrent Neural Network (RNN) module is used to capture time-shifted signal correlation. To enhance feature extraction, Short-Time Fourier Transform (STFT) feature extraction is incorporated. The combined feature vector is processed through a reinforcement learning module. Finally, these features are used to train the deep learning classifier which uses residual blocks for better representation of feature while learning. The highest prediction score is considered as the decision threshold in this work. The outcome of this work is compared with other deep learning methods in terms of P_d, P_f and sensing error for varied sample size and modulation schemes. The comparative analysis shows the robustness of proposed approach by achieving the P_f as 0.32 and 0.06 for QAM16 and QPSK modulations.**

***Keywords*—cognitive radio, spectrum sensing, machine learning, Deep Learning (DL), Reinforcement Learning (RL)**

I. INTRODUCTION

During last two decades, we have noticed a tremendous growth in the demand of wireless and mobile communication. This growth has resulted in increasing the global data traffic. This growth constantly continuing to increase. According to Ref. [1], there were 5.1 billion mobile subscribers in 2018 which is expected to grow 5.7 billion in 2023.

Similarly, Cisco presented an annual Internet report which has expected the increase in mobile devices from 8.8 billion in 2018 to 13.1 billion in 2023. By 2023, 5G connection speeds are anticipated to reach 575 Mbps, which is 13 times faster than the typical mobile connection. New wireless communications applications like the Internet of Things, wearable technology, etc. are continually creating a massive quantity of data in addition to the enormous expansion in data traffic [2]. The mobile and internet based applications urge for high speed data rate and increased Quality of Service (QoS) resulting in increase in the requirement of free/idle spectral bands. According to the current communication standards, a certain spectrum is allotted to the application.

The demand for additional spectrum resources is increasing exponentially as more wireless services are supported, which has sparked the development of new high-speed data network technologies. However, the radio spectrum is fundamentally considered as a limited resource in which the frequency bands are solely allocated to the licensed user which is known as Primary User. This prolonged allocation when the device is in ideal condition lead to increase the spectrum dearth in the certain spectrum band. Contrarily, an assessment of spectrum utilization being carried out by the Federal Communication Commission (FCC) has revealed that several areas of the radio spectrum, known as spectrum holes, are not employed for a considerable length of time, which results in underutilization of the given spectrum. Similarly, recent studies on Orthogonal Frequency Division Multiplexing (OFDM) communication standards shows that spectrum underutilization issues are faced due to fixed spectrum allocation [3].

Therefore, spectrum scarcity becomes a challenging issue which affects the communication performance of the entire network. in order to mitigate this issue, researchers have introduced cognitive radio mechanism which is considered as favourable solution to overcome the issue of spectrum scarcity. The cognitive radios examine the availability of idle or unused spectrum from Primary Users (PUs). this unused spectrum is allocated to secondary user without producing any interference for the PU. This task is done by employing Dynamic Spectrum Access (DSA) technique. This process is termed as "spectrum sensing".

Manuscript received July 18, 2023; revised August 9, 2023; accepted September 18, 2023; published December 1, 2023.

doi: 10.12720/jait.14.6.1321-1330

Cognitive radio is a technology that enables wireless communication devices to detect and intelligently adapt to their environment. The goal of cognitive radio is to improve the efficiency of the radio spectrum by allowing unlicensed devices to share unused or underused spectrum bands without interfering with licensed users. Cognitive radio devices are equipped with advanced sensing and signal processing capabilities that allow them to sense and analyze the Radio Frequency (RF) environment in real-time. Based on this analysis, cognitive radio devices can make intelligent decisions on which channels to use, how much power to transmit, and other parameters to optimize their performance while minimizing interference to other users.

One of the key features of cognitive radio is its ability to operate dynamically and adaptively to changes in the RF environment. This allows cognitive radio devices to exploit underutilized spectrum and avoid crowded or noisy channels, improving their efficiency and reducing interference. Cognitive radio technology has potential applications in a wide range of fields, including military communications, wireless networks, and public safety. The use of cognitive radio is also being explored for emerging technologies such as the Internet of Things (IoT), where the ability to adapt to dynamic RF environments can improve the reliability and performance of wireless IoT devices.

The cognitive radio networks pose several advantages in wireless cellular communication systems. Recently, several spectrum sensing approaches have been established to mitigate the spectrum scarcity issues. Generally, these methods are classified as wideband and narrowband sensing methods [4]. The wideband sensing methods focus on analysing the number of frequencies at a time whereas narrowband methods focus on analysing one frequency channel at a time. several narrowband spectrum sensing methods have been introduced such as energy detection [5, 6], matched filter detection [7], cyclostationary feature detection [8, 9], covariance based detection [10, 11], and machine learning based sensing approach [12], etc. The energy detection method relies on power of incoming signal and compares it to previously estimated threshold to estimate the presence of PUs. However, the performance of this method is affected due to low Signal-to-Noise Ratio (SNR) conditions.

Dibal *et al*. [15] reviewed several methods of spectrum sensing where waveform detection method is described which has highest reliability by correlating the reference and received signal. This is a highly efficient method but it requires precise information about the signals of PUs. moreover, Secondary Users (SUs) do not have information about PU’s signals thus it cannot be implemented for blind detection. Zhao *et al*. [16] presented eigenvalue based spectrum sensing method which has shown significant performance for low SNR conditions but its performance is affected due to high computational complexity. In contrast, the wideband methods divide the spectrum into different sub-bands which are further sensed subsequently or concurrently by employing narrowband techniques.

The sequential methods fail to achieve better performance because of time complexity and energy consumption to use of high rate analog-to-digital converters whereas the simultaneous spectrum sensing mechanism require more number of sensors and synchronization function which increases the implementation complexity [13]. Similarly, Usman *et al.* [14] presented a study where suggested that prior information of PU signals can be used to classify spectrum sensing and these methods can be classified as coherent and non-coherent schemes. Further, these approaches can be categorized based on transmitter, receiver and interference for spectrum sensing. The traditional spectrum sensing methods encounter challenges due to multipath fading and shadowing. These obstacles can be mitigated by leveraging spatial diversity via cooperative spectrum sensing. In this approach, specific CR users exchange the data with Fusion Centre (FC).

This FC integrates local information to formulate global decision. At this stage, CR users can perform soft-decision and hard decision processes to enhance the detection performance [17]. The idea of learning from the environment is an essential aspect of cognitive radios, which involve monitoring and adjusting operating characteristics to changing conditions. In order to facilitate this learning process, several researchers have explored the use of machine learning systems [11–16] for spectrum sensing. Since channel conditions can be difficult to estimate due to fading and shadowing, spectrum sensing based solely on current sensing slots may not be reliable in determining the PU status. However, machine learning and deep learning-based spectrum sensing can implicitly learn about the environment and detect PU activity without prior knowledge of the surroundings. Several machine learning based methods have been presented in literature review section. In this work, we adopt the deep learning based approach for spectrum sensing to increase the performance of cognitive radio networks. The main contributions of this work are as follows:

1. We adopt the deep learning method for feature extraction where Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) modules are incorporated.
2. In order to improve the feature extraction process, we include Short-Time Fourier Transform (STFT) feature analysis.
3. The obtained features are processed through the reinforcement learning block where residual block based deep learning scheme is employed for classification.

The traditional spectrum sensing methods rely on dynamic spectrum access techniques such as energy detection [5, 6], matched filter detection but increased demand of spectrum leads to increase the complex characteristics of spectrum usages. Therefore, these methods fail to achieve the desired performance for spectrum sensing. Moreover, traditional machine learning overcome the challenges but accuracy and reliability still remains challenging task. To overcome these issue, we adopt deep learning approach along with reinforcement learning mechanism.

The deep learning methods is used for deep feature extraction and RL approach is used to improve the training process by incorporating the self-learning characteristics during training phase. For feature extraction, we use the combination of CNN, RNN with STFT module to obtain the robust features. Finally, the obtained features are processed through a deep learning classifier to achieve the sensing outcome.

Rest of the article is organized in following subsections: Section II presents the brief literature review about machine learning and deep learning based approaches for spectrum sensing. Section III presents the proposed deep learning based solution for cooperative spectrum sensing. Section IV presents the outcome of proposed approach and its comparative analysis with existing mechanisms. Finally, Section V presents the conclusion and future scope of this research.

II. Literature Review

This section presents the brief discussion on recent researches in this domain of spectrum sensing in cognitive radio networks.

Ghazizadeh *et al*. [18] adopted cooperative Stainless Steel (SS) mechanism and introduced a novel approach by introducing an improved Support Vector Machine (SVM) classifier which is named as 2-Phase SVM. This approach uses vector containing energy levels of PU as feature vector which are further processed by employing SVM training and testing process. The labelling is done based on the availability and unavailability of channels. Similarly, the ensemble classification methods are known as the advanced machine learning methods which uses combination of two or more classifiers. These methods provide better classification accuracy. Ahmad *et al*. [19] presented machine learning approach for SS. According to this process, the feature extraction process is carried out by using cyclostationary method where Fast Fourier Transform (FFT) accumulation method is employed to obtain the final features. Finally, these features are used to train the ensemble classifier to detect the signals in low SNR conditions. This ensemble classifier is constructed with the decision tree and AdaBoost methods. Gul *et al*. [20] discussed the issue of involvement of malicious user which reports the false information to the fusion centre about PU's activity. To overcome this issue, authors introduced boosted tree algorithm to increase the detection accuracy and reliability by mitigating different attacks caused due to malicious users such as Always Yes, Always No, Always opposite and Random Opposite.

Sheng *et al.* [21] adopted narrowband spectrum sensing method and incorporated concept of machine learning. The first phase of this approach includes extraction of trace of covariance matrix and Variance of Quadratic Covariance Matrix (TCVQ) feature vectors and later, support vector machine classifier is employed to classify the feature patterns. The TCVQ approach explores the difference of Eigen values and structural characteristics of received signal which helps to increase the overall performance of the systems.

Recently, deep learning based schemes have gained huge attention in various applications and widely adopted in spectrum sensing tasks [22]. However, the traditional machine learning methods rely on the signal-noise models therefore, accuracy of these systems depends on the assumption of these models. To overcome this issue, Xie *et al*. [23] presented a combined deep learning architecture with the help of Convolutional Neural Networks (CNNs) and the Long-Short Term Memory (LSTM) networks. The CNN and LSTM model helps to extract the spatial and temporal feature of the given input data. Further, this CNN-LSTM model extracts the energy-correlation attributes by exploring the covariance matrices which are produced by sensing data. The obtained series of features is then fed to the LSTM block to learn the activity pattern of PUs. Xing *et al*. [24] reported that existing methods are fail to capture the temporal correlation attributes from spectrum data. Generally, the CNNs do not extract the temporal correlations from time series data whereas LSTMs are used to accomplish this task. However, it fails to focus on important part of the given spectrum data. Therefore, authors presented a novel deep learning approach which is developed by combining 1D CNN, BiLSTM, and self-attention. The CNN and BiLSTM extract the local features and global correlation information from time series data. Similarly, the self-attention mechanism helps to focus on the significant attributes obtained from BiLSTM.

Solanki *et al*. [25] developed a combined model by combining CNN and Recurrent Neural Network (RNN) models. Further, transfer learning mechanism is also used to improve the SS performance. Sarikhani *et al.* [26] discussed the importance of deep reinforcement learning in cooperative SS and introduced DRL based CSS approach. This approach is useful in reducing the signalling in the SUs.

Du *et al*. [27] presented a new approach which uses information geometry and deep learning for spectrum sensing. In first stage, covariance matrix is computed and later geodesic distance is computed between signals which is considered as feature vector. Finally, deep learning classification algorithm is employed.

Pati *et al*. [28] introduced a new approach which is based on the deep convolution neural network and transfer learning concept for non-cooperative spectrum sensing. This CNN architecture consists of four-layers which is helpful in reducing the computational complexity and further, the transfer learning approach is also implemented to improve the learning performance. Zheng *et al*. [29] reported that traditional spectrum sensing approaches have limitations in dealing with the complex and dynamic radio environment. To address this issue, the authors propose a deep learning-based spectrum sensing approach that uses a Convolutional Neural Network (CNN) for classification. The proposed approach involves two stages: feature extraction and classification. In the feature extraction stage, the received signal is pre-processed to obtain spectrogram images, which capture the frequency and time-domain characteristics of the signal. The CNN is then trained on the spectrogram images to learn the features that are relevant for classification. In the classification stage, the trained CNN is used to classify the received signal as either occupied or unoccupied. Nasser *et al*. [30] introduced a deep learning based hybrid approach that combines energy detection and cyclostationary feature detection techniques. The authors describe the architecture of the DotNetNuke (DNN) model, which consists of several layers of neurons that process the input signal and output the probability of spectrum occupancy. They also discuss the training process for the DNN model, which involves using a large dataset of labelled signals to adjust the weights and biases of the neurons. Several methods have been discussed but these methods suffer from several challenges. Below given Table I presents a brief overview of these mechanisms.

TABLE I. COMPARATIVE OVERVIEW OF EXISTING SCHEMES

Ref.	Contribution	Remarks
[18]	SVM ML based cooperative SS	Traditional ML do not have strong feature learning characteristics
[19]	ML based model which uses FFT based feature extraction	This method do not carry strong feature extraction thus misclassification rate increases as the data increases
[21]	It uses covariance based feature and SVM classification	Poor accuracy and convergence impacts its performance
[23]	DL method by using CNN and LSTM	It is able to extract the temporal and spatial feature but extracts only energy attributes
[24]	CNN and BiLSTM and attention based DL model	Overcomes the temporal feature extraction
[25]	CNN and RNN model with transfer learning	It uses transfer learning but requires huge dataset
[28]	CNN with Non-cooperative SS	It reduces complexity and improves the performance by using transfer learning
[29]	Extracts time and frequency domain features and classify with CNN	Converts spectrum data into image and train the network
[30]	Combined energy and cyclostationary feature detection with DL	Requires huge dataset and suffer from gradient vanish problem

III. PROPOSED MODEL

This section presents the proposed deep learning and reinforcement learning based solution to enhance the performance of spectrum sensing. The first sub-section briefly describes the various components of proposed model and next sub-section presents the proposed hybrid architecture.

A. *Main Components of Proposed Model*

Signal Model: The signal model represents the received signal at each cognitive radio device, which includes the transmitted signal from the primary user and any noise and interference in the environment. The signal model can be expressed as:

$$y[n] = s[n] + w[n] \tag{1}$$

where $y[n]$is the received signal, $s[n]$ is the transmitted signal from the primary user, and $w[n]$ is the noise and interference.

Deep Learning Model: The deep learning model is used to extract features from the received signal and classify it as either idle or busy. The deep learning model can be represented by a function f that maps the input signal x to a probability $p(x)$ of the signal being idle. The deep learning model can be trained using a dataset of labeled signals to optimize its parameters and improve its accuracy.

Reinforcement Learning: Reinforcement learning is a machine learning technique that has shown promise for improving spectrum sensing in cognitive radio systems. In reinforcement learning, an agent learns to make decisions by interacting with an environment, receiving rewards or punishments based on its actions, and adjusting its behavior accordingly. The use of reinforcement learning for spectrum sensing involves designing an agent that can learn to sense the radio spectrum efficiently and accurately. The reinforcement learning approach for spectrum sensing involves the following key components:

- State representation: The state of the environment is represented by a set of features that capture the relevant characteristics of the radio signals, such as signal strength, noise level, and interference.
- Action selection: The agent selects an action based on the current state of the environment. In spectrum sensing, the action corresponds to the frequency band or channel to sense.
- Reward function: The agent receives a reward or punishment based on the accuracy of its sensing decision. The reward function can be designed to encourage the agent to prioritize certain frequency bands or to penalize false sensing decisions.
- Learning algorithm: The agent uses a learning algorithm to adjust its behavior based on the rewards and punishments received. Reinforcement learning algorithms such as Q-learning, SARSA, or Deep Q-networks (DQNs) can be used for this purpose.

The reinforcement learning approach can be implemented in a cooperative spectrum sensing scenario where multiple cognitive radio nodes work together to sense the spectrum. The nodes can share their sensing decisions and reward signals to learn collectively and improve the overall spectrum sensing performance.

One key advantage of reinforcement learning for spectrum sensing is its ability to adapt to changing radio environments. The agent can learn from past experiences and adjust its behavior based on the current state of the environment. This makes it well-suited for dynamic and unpredictable radio environments, where traditional sensing approaches may be less effective. Overall, the use of reinforcement learning for spectrum sensing in cognitive radio systems is an active area of research with promising results. However, there are also challenges that need to be addressed, such as the design of efficient state representations and the scalability of learning algorithms for large-scale networks.

Decision Threshold: The decision threshold is used to make the final decision on whether the spectrum is idle or busy based on the output of the cooperative sensing algorithm. The decision threshold can be set based on the desired false alarm and detection probabilities.

B. Proposed Model

This section presents the proposed deep learning and Reinforcement learning architecture for spectrum sensing. The complete architecture is depicted in Fig. 1, where H_1 and H_0 are the initial datasets. These datasets are further processed through different feature extraction models such as energy, energy correlation, time shifted signal correlation [31] and STFT. Further, these features are merged together to formulate the ensemble feature vector. Later, deep reinforcement learning model with residual network is incorporated to learn these attributes and classify the patterns according to presence and absence of PUs.

1) System model

In cognitive radio networks, the spectrum sensing problem can be represented in the form of binary hypothesis problem which can be represented as:

$$H_0: Y_n = U_n$$
$$H_1: Y_n = h_n X_n + U_n \qquad (2)$$

where H_0 denotes the absence of PU and H_1 represents the presence of PU, i.e., PU is in silent state and active state, respectively. X_n denotes the transmitted signal, whereas Y_n denotes the received signal vector. The channel index value between PU and SU can be denoted as $h_n \in C_m$and U_n denotes the noise vector. By utilizing the signal vector, it is possible to create decision statistics that identify the state of the PU as H_1 in the test statistics (T) using a decision threshold (D_s). If T exceeds the value of D_s, it indicates the presence of PUs; otherwise, their absence is indicated. If $T > D_s$ condition is satisfied then it shows the presence of PU otherwise it denotes the absence of PU.

2) CNN model for feature extraction

In this research, we have adopted the CNN based model to improve the spectrum sensing performance because the traditional machine learning schemes have overfitting issue and outperformed by DL methods. This research utilizes a training set, denoted as Y, consisting of N pairs of input data $(x_N)$and labeled data (l_N) represented as $Y = \{(x_1, l_1), (x_2, l_2), (x_3, l_3), \dots, (x_N, l_N)\}$. The presence of PUs is indicated by Y. However, increasing the size of the training input leads to a rise in computational complexity. Sampling statistics used for PU sensing may include redundant data from the same distribution source. Thus, pre-processing of the input data is necessary before starting the training process. In this work, we employed energy and time shifted signal correlation, and STFT as the important attributes.

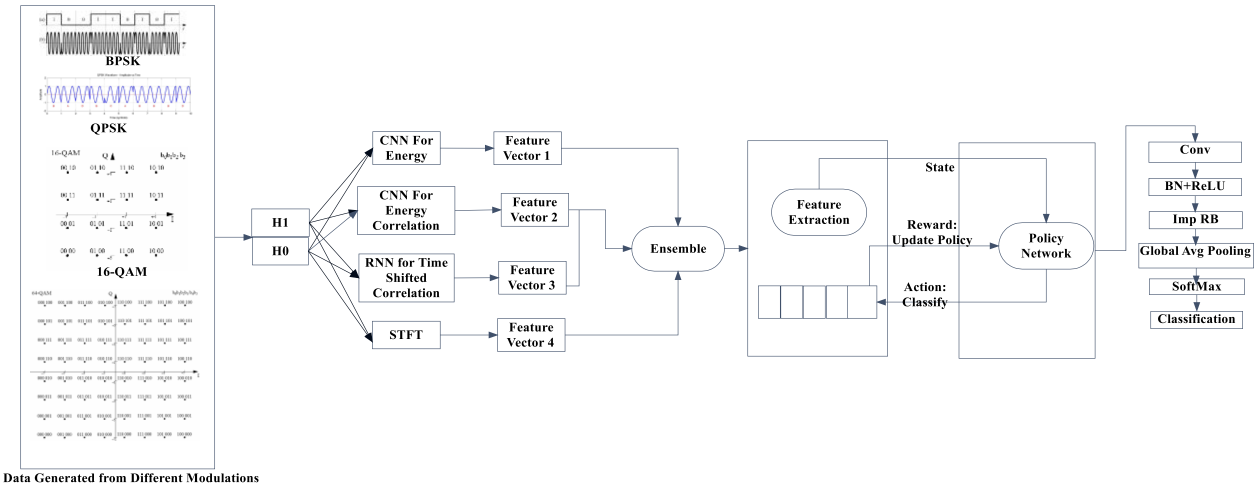

Figure 1. Proposed model.

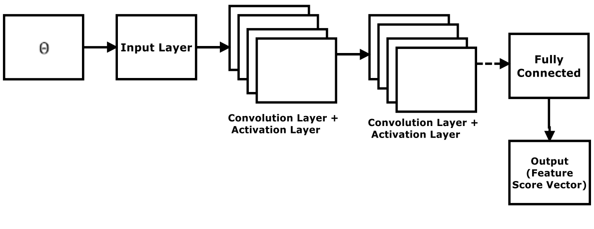

Figure 2. CNN feature extraction module.

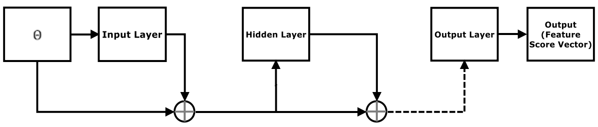

Figure 3. RNN feature extraction module.

In feature extraction phase, the covariance matrices are given as input to the CNN and RNN layers to extract the corresponding features. Similarly, these matrices are given as input to the STFT feature computation model. The convolution block consists of Leaky Rectified Linear Unit. The obtained spatial features are fed to the 2D convolution layer which contains filter of size 3×3 convolution. In order to maintain the consistency between input and output, zero padding and stride operations are also included. Finally, the fully connected layer is implemented which is used to perform the classification. Figs. 2 and 3 depicts the architecture of CNN and RNN models for feature extraction.

Similarly, we implement short time Fourier transform analysis for feature extraction which can be expressed as:

$$STFT_x(t,\omega) = \sum_{n=\infty}^{L-1} x(n)g(n-t)e^{-jm\omega} \quad (3)$$

where g(n) denotes the window function. In this work, we have adopted Hamming window function because of it achieves good performance and reduces spectrum leakage. Further, the spectrogram of the input signal can be computed as:

$$SP_x(t,\omega) = |STFT_x(t,\omega)|^2 \quad (4)$$

With the help of spectrum sensing model, the received signal can be simplified as:

$$= [x(1), x(2), \dots, x(n)] \quad (5)$$

Final, normalized feature vector is generated as:

$$X_s = \left[x_S^{(1)}, x_S^{(2)}, \dots, x_S^{(n)}\right] \quad (6)$$

The combination of these features produces the final feature vector to make the final decision about presence of PUs as $D_{H_1}(\circledcirc_n)$or absence of PU as $D_{H_0}(\circledcirc_n)$ such that $D_{H_0}(\circledcirc_n) + D_{H_1}(\circledcirc_n) = 1$.

3) State space model for RL

Further, we implement Reinforcement learning mechanisms to learn the patterns. a brief discussion or this approach is presented in aforementioned sub-section. In a wireless network, a shared bandwidth is divided into $\boldsymbol{N}$ correlated channels. At this stage, the occupied channel is (1) and vacant (0). The entire system can be modelled using a Markov model with 2^N states. MDP is a mathematical framework used for modeling decision-making problems in situations where outcomes are partly random and partly under the control of a decision maker. It consists of several components such as States (S), Actions (A), Transition Probabilities (P), Rewards (R), Policy (π) and optimal policy etc. It comprises of a finite collection of states; these states symbolize distinct scenarios or setups within a system. The decision maker engages with the environment, causing the system to shift from one state to another. In every given state, there exists a finite collection of actions accessible to the decision maker. These actions serve as the available choices or decisions that can be executed within that specific state, ultimately resulting in transitions between states. The transition probability ensures the movement from one state to another state when any particular action is performed. For each transition, there is a certain reward is assigned for the action performed. With the help of this, the state space can be represented as:

$$S = \left\{s = (s_1, \dots, s_N) \mid s_i \in \{0,1\}, i \in \{1, \dots, N\}\right\} \quad (7)$$

The state transition of each channel can be denoted as: $P = \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix}$.

4) Action space model for RL

In these communication systems, a single user requires a certain bandwidth or vacant channels to aggregate all the channels in the range of aggregation capacity. Initially, user either remain ideal or selects a certain length C from whole channel to sense. Thus, the remaining segment for selection can be denoted as $N - C + 1$. By repeating this process, the vacant channels will be aggregated for transmission. Let the action can be denoted as $A = \{0,1, \dots, N - C + 1\}$ and user selects i^{th} segment at the initialization of slot t if $a_t = i(i \in A, i \neq 0)$ or remains ideal it $a_t = 0$.

5) Reward model for RL

After taking the action a_t at time t and we consider that user can receive a binary feedback f_t as Acknowledgement (ACK) regarding successful packet delivery. For successful transmission, $f_t = 1$and failure transmission is characterized by $f_t = 0$. Then, the reward function for this action at time t, can be expressed as:

$$r_t(s_t, a_t) = \begin{cases} 0, \; if \; a_t = 0 \\ 4f_t - 2, if \; 1 \leq a_t N - C + 1 \end{cases} \quad (8)$$

where s_t denotes the state of system at a given time slot t. Here, our main aim is to find the optimal policy π which is used in drawing the observations o_t to the next action a_{t+1} at each time slot to exploit the collected reward. This policy can be expressed as:

$$V_\pi(o) = \mathbb{E}_\pi\left[\sum_{t=0}^{\infty} \gamma^t r_{t+1}\big(s_{t+1}, \pi(o_t)\big) | o_0 = o\right] (9)$$

Where $\in (0,1)$ denotes the discount factor, $\pi(o_t)$ represents the action under policy π for current observation o_t. Based on this, the optimum policy π^*can be expressed as:

$$\pi^* = arg \max_\pi V_{pi}(o)$$
$$= arg \max_\pi E_{pi}\left[\sum_{t=0}^{\infty} \gamma^t r_{t+1}\big(s_{t+1}, \pi(o_t)\big) | o_0 = o\right] \quad (10)$$

6) Classification

In order to classify the spectrum sensing patterns, we present a CNN based classification architecture for spectrum sensing. The proposed architecture uses a residual block model to improve the learning performance.

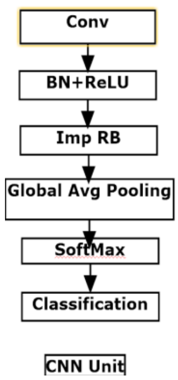

Figure 4. Proposed CNN.

Fig. 4 shows the proposed architecture and Fig. 5 shows the residual block model. The proposed residual block includes the insertion of a ReLU activation function after each convolution layer, resulting in three linear regression activation functions that are better at extracting feature information compared to a single one.

To enhance the network's extraction efficiency, Batch Normalization (BN) is introduced, which speeds up the network's convergence process and makes the entire network's training more robust. Because the spectrum sensing model can overfit with limited training data, especially when deepening the network, a dropout layer is also included in the residual block to prevent this.

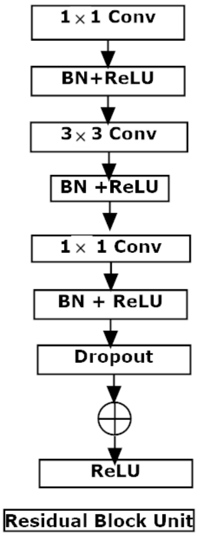

Figure 5. Residual block.

The outcome of residual block is $H(x) = f(z) + x$, and output of layer l of residual network is $H_1(x) = f(z_l) + x_l$. In order to train the network, samples of signal X_s are labelled and used for training which are represented as:

$$Y_{training} = \left\{\left(x_s^{(1)}, y^{(1)}\right), \left(x_s^{(2)}, y^{(2)}\right), \dots, \left(x_s^{(m)}, y^{(m)}\right)\right\} \quad (11)$$

Similarly, the test samples are denoted as:

$$Y_{test} = \begin{Bmatrix} \left(x_s^{(m+1)}, y^{(m+1)}\right), \left(x_s^{(m+2)}, y^{(m+2)}\right), \dots \\ , \left(x_s^{(n)}, y^{(n)}\right) \end{Bmatrix} \quad (12)$$

In this training process, we have used cross-entropy loss function which is expressed as follows:

$$L = \left[\sum_{i=1}^{N} y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)} \log\left(1 - \hat{y}^{(i)}\right)\right] \quad (13)$$

IV. Result and Discussion

This section presents the experimental analysis of proposed model where obtained performance is compared with existing scheme. First sub-section presents the dataset generation process and second section presents simulation analysis.

A. Dataset Generation

In this work we have used publically available Radio ML dataset. In this work, we have considered five different nodulation schemes such as BPSK, QPSK, 8PSK, QAM16, and QAM64. The SNR distribution variation is considered from −20 dB to +18 dB with an increment of 2 dB.

These signals are represented as positive samples whereas the complex additive Gaussian noise is used to generate the negative samples. The complete training set contains $'n'$ samples which are processed through the deep neural network in the form 2×*n* vector with in phase and quadrature phase components. Table II shows the parameter details used during dataset generation.

TABLE II. Simulation Parameters for Dataset Generation

Parameters	Considered value
Modulation	BPSK, QPSK, 8PSK, QAM 16 and QAM 64
Sample Size	64, 128, 256, 512
SNR	−20 dB to +18 dB
Training Samples	153,000
Validation samples	51,000
Testing Samples	51,000

B. Performance Measurement Parameters

The assessment criteria used in this scenario include three metrics: Probability of detection (P_d), Sensing Error (SE), and Probability of false alarm (P_f). P_d is the probability of correctly identifying the presence of the primary user in the spectrum when it is being used. P_f, on the other hand, is the probability of incorrectly identifying the presence of the primary user when the spectrum is not occupied.

These probabilities were calculated across various Signal-to-Noise Ratio (SNR) levels. To compute SE, the average of P_f and the probability of miss detection (P_m) was determined. P_m is the probability of wrongly identifying that the spectrum is not in use when the primary user is actually using it. The overall performance of this system is measured based on confusion matrix. Below given Table III shows the confusion matrix.

TABLE III. Confusion Matrix for Spectrum Sensing

		Predicted Value	
		Signal	Noise
Actual	Signal	A	B
	Noise	C	d

Based on this confusion matrix, the performance measurement parameters can be computed as presented in Table IV:

TABLE IV. Parameters for Spectrum Sensing

P_d	$\frac{a}{a+b}$
P_m	$1 - P_d$
P_f	$\frac{c}{c+d}$
SE	$average(P_f, P_m)$

C. Comparative Analysis

In this segment, we present the comparative investigation of proposed approach for varied deep learning scheme by considering different modulation techniques.

Figs. 6–9 shows the comparative analysis for different modulation schemes.

In this experiment, we measure the performance of proposed model and compared it with traditional Deep Learning model, CNN, and LSTM model for varied modulation such as BPSK, QPSK, QAM 16 and QAM 64. For varied SNR levels. The experimental analysis shows that probability detection increases for QAM modulations.

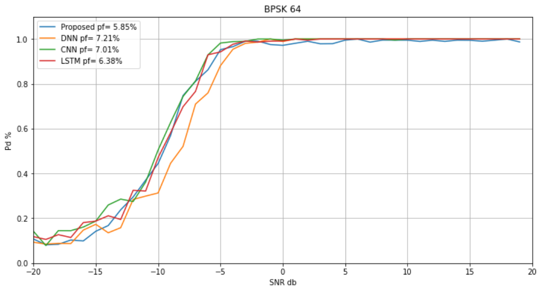

Figure 6. Performance for BSPK 64.

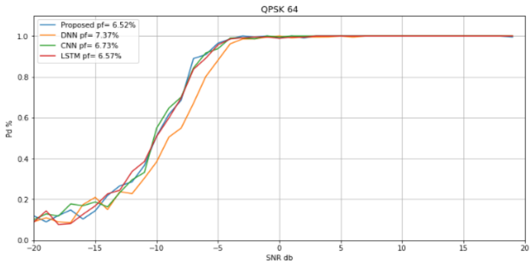

Figure 7. Performance for QSPK 64.

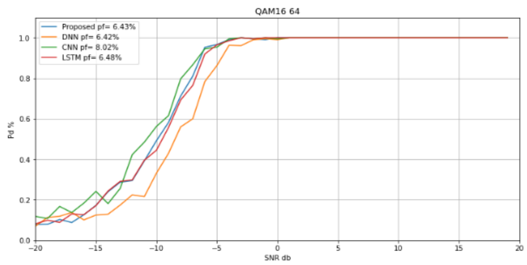

Figure 8. Performance for QAM 16.

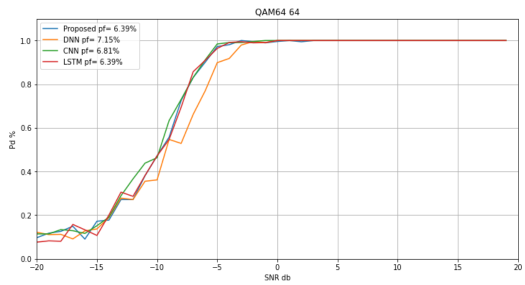

Figure 9. Performance for QAM 64.

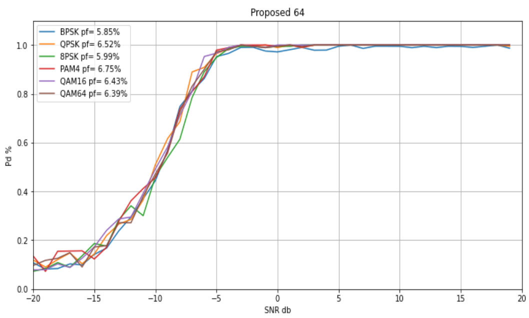

Figure 10. Comparative analysis for Proposed 64 samples considering different modulation schemes.

From this experiment, we conclude that the signals with lower SNRs carry less information in comparison with high SNR because the quality of these signals is compromised due to noise. Aforementioned figure shows the performance for 64 samples where proposed approach obtained improved performance because of its significant pattern learning nature. Further, we extended the experimental analysis for different number of samples. Fig. 10 depicts the comparative analysis for four different sample size such as 64, 128, 256, and 512 by considering different modulation schemes.

Comparative analysis for Proposed 64 samples considering different modulation schemes is shown in Fig. 10. Comparative analysis for Proposed 128 samples considering different modulation schemes is shown in Fig. 11. Similarly, we have measured the performance of proposed approach for different modulation schemes such as BPSK, QPSK, 8PSK, QAM16 and QAM 64. The obtained performance is depicted in Fig. 12. In an another experiment, we have used 512 samples and measured the performance in terms of P_d. The obtained performance is depicted in Fig. 13.

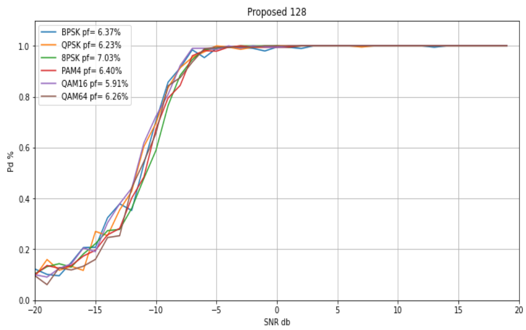

Figure 11. Comparative analysis for Proposed 128 samples considering different modulation schemes.

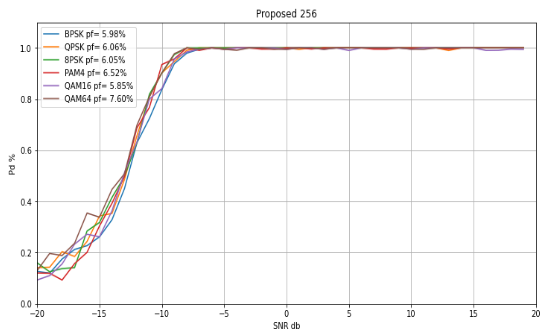

Figure 12. Comparative analysis for Proposed 256 samples considering different modulation schemes.

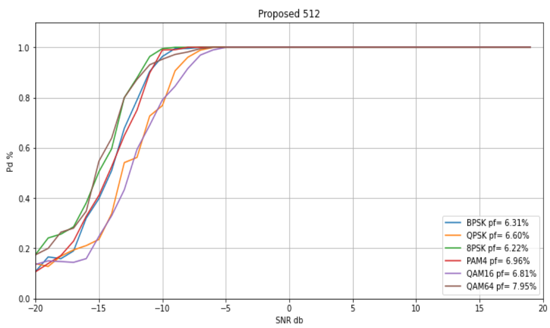

Figure 13. Comparative analysis for Proposed 512 samples considering different modulation schemes.

This experiment shows that the proposed approach achieves better performance for QAM 64 modulation scheme for all sample scenarios. In order to compare the performance of proposed model, we have extended the experimental analysis for 64, 128, 256, and 512 samples which are modulated by employing QAM16 and QPSK techniques. The obtained performance is compared with the existing deep learning mechanism. Below given Table V shows the comparative performance for QAM16 modulation for 64 and 128 sample size.

TABLE V. COMPARATIVE PERFORMANCE FOR 64 AND 128 SAMPLE LENGTH QAM16 MODULATION

	64 Sample			128 Sample		
Models	P_f	$P_d(-20\,dB)$	$SE(\%)$	P_f	P_d	$SE(\%)$
CNN	1.75	24.48	14.86	3.60	26.87	15.70
LeNet	0.89	24.96	14.51	0.84	27.15	14.63
ResNet	0.68	25.61	14.39	0.55	24.80	14.42
Inception	14.63	36.17	19.65	20.22	39.91	21.68
LSTM	0.47	24.13	14.75	3.22	28.40	15.35
CLDNN	1.25	25.15	14.87	6.93	31.20	17.30
Proposed Model	0.32	30.25	11.50	2.21	34.30	10.10

The sensing error of proposed model is less for this experiment therefore the P_f and P_d are reduced in proposed model. Thus, it improves the spectrum occupancy detection rate. Further, we measure the performance for QPSK signal for 64 and 128 samples. Below given Table VI shows the comparative performance for QPSK modulation for 64 and 128 sample size.

TABLE VI. COMPARATIVE PERFORMANCE FOR 64 AND 128 SAMPLE LENGTH QPSK MODULATION

	64 Sample			128 Sample		
Models	P_f	$P_d(-20\,dB)$	$SE(\%)$	P_f	P_d	$SE(\%)$
CNN	1.50	27.15	14.91	4.95	23.35	16.10
LeNet	0.10	23.47	14.36	0.11	26.15	14.20
ResNet	0.09	23.58	14.60	0.15	25.50	14.50
Inception	16.98	35.90	20.40	20.30	39.40	21.70
LSTM	1.29	24.70	14.90	3.86	29.40	15.60
CLDNN	1.64	26.27	15.20	6.10	31.20	16.30
Proposed Model	0.06	40.85	10.50	21.50	35.20	10.8

Similarly, we measured the performance for 256 and 512 samples for QAM16 and QPSK modulation. Below given Tables VII, VIII shows the outcome for QAM 16 and QPSK modulation schemes.

TABLE VII. COMPARATIVE PERFORMANCE FOR 256 AND 512 SAMPLE LENGTH QAM16 MODULATION

	256 Sample			512 Sample		
Models	P_f	$P_d(-20\,dB)$	$SE(\%)$	P_f	P_d	$SE(\%)$
CNN	9.54	30.70	17.80	12.90	35.31	19.05
LeNet	1.40	25.30	14.70	2.70	27.89	15.40
ResNet	0.015	23.90	14.50	0.20	25.77	14.50
Inception	21.50	38.40	22.05	17.90	39.70	18.60
LSTM	3.50	29.40	15.60	1.60	26.40	15.10
CLDNN	8.30	34.30	16.70	6.96	42.90	14.40
Proposed Model	0.05	44.50	12.90	0.03	46.95	10.50

TABLE VIII. COMPARATIVE PERFORMANCE FOR 256 AND 512 SAMPLE LENGTH QPSK MODULATION

	256 Sample			512 Sample		
Models	P_f	$P_d(-20\,dB)$	$SE(\%)$	P_f	P_d	$SE(\%)$
CNN	10.40	32.50	18.40	14.50	34.90	19.50
LeNet	0.90	26.20	14.65	2.75	26.85	15.50
ResNet	0.12	25.10	14.30	0.05	24.20	14.43
Inception	22.98	39.60	22.50	17.60	42.22	18.5
LSTM	4.51	29.20	15.81	2.60	28.50	15.32
CLDNN	9.65	35.40	17.40	8.82	43.85	15.15
Proposed Model	4.20	45.12	13.10	0.05	49.65	10.15

Therefore, based on these studies we demonstrate that the proposed model is able to achieve the desired performance with less sensing error due to its significant nature of pattern learning for different noise levels.

V. CONCLUSION

Currently, the cognitive radio networks have gained huge attention in this domain of wireless communication because the CRs have been identified as the promising solution to deal with spectrum scarcity issue. However, the traditional spectrum sensing methods suffer from various challenges.

Therefore, machine learning schemes are adopted in this domain. The proposed model uses the concept of deep learning and introduced a new architecture for spectrum sensing. This architecture uses CNN, RNN and STFT for feature extraction. The obtained feature vector is combined and processed through the reinforcement learning. Finally, a CNN based classification module is presented to classify the patterns of spectrum of noise. In future, this approach can be explored for power control, security enhancement and resource allocation to improve the overall performance of the communication system.

CONFLICT OF INTEREST

The Authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization and Methodology by Usha Rani M. A and Prashanth C. R, Formal Analysis and Investigation by Usha Rani M. A and Prashanth C. R, Software and Validation by Usha Rani M. A, Data Curation and Original Draft Preparation by Usha Rani M. A, Supervision, Project Administration, Review and Editing by Prashanth C. R, Visualization by Usha Rani M. A. All authors had approved the final version.

ACKNOWLEDGMENT

The authors would like to thank all the staff of the Department of Electronics & Telecommunication Engineering, Principal, Dr. Ambedkar Institute of Technology for the support given.

REFERENCES

[1] D. Janu, K. Singh, and S. Kumar, "Machine learning for cooperative spectrum sensing and sharing: A survey," *Transactions on*

Emerging Telecommunications Technologies, vol. 33, 2021. doi: 10.1002/ett.4352
[2] Cisco, Cisco Annual Internet Report (2018–2023), White Paper, 2020.
[3] N. Kaabouch, *Handbook of Research on Software-Defined and Cognitive Radio Technologies for Dynamic Spectrum Management*, IGI Global, 2014.
[4] P. Feng, Y. Bai, Y. Gu, J. Huang, X. Wang, and C. Liu, "A rapid coarse-grained blind wideband spectrum sensing method for cognitive radio networks," *Computer Communications*, vol. 166, pp. 234–243, 2021.
[5] X. Ling, B. Wu, H. Wen, P. H. Ho, Z. Bao, and L. Pan, "Adaptive threshold control for energy detection based spectrum sensing in cognitive radios," *IEEE Wirel. Commun. Lett.*, vol. 1, pp. 448–451, 2012.
[6] D. Capriglione, G. Cerro, L. Ferrigno, and G. Miele, "Effects of real instrument on performance of an energy detection-based spectrum sensing method," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 5, pp. 1302–1312, 2019.
[7] A. Brito, P. Sebastião, and F. J. Velez, "Hybrid matched filter detection spectrum sensing," *IEEE Access*, vol. 9, pp. 165504–165516, 2021.
[8] K. Sherbin and V. Sindhu, "Cyclostationary feature detection for spectrum sensing in cognitive radio network," in *Proc. 2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, IEEE, May 2019, pp. 1250–1254.
[9] J. Zhang, F. Ran, and D. Liu, "Maritime cognitive radio spectrum sensing based on multi-antenna cyclostationary feature detection," *International Journal of Electronics*, vol. 107, no. 7, pp. 1044–1062, 2020.
[10] C. Charan and R. Pandey, "Intelligent selection of threshold in covariance-based spectrum sensing for cognitive radio networks," *Wireless Networks*, vol. 24, pp. 3267–3279, 2018.
[11] A. Z. Chen and Z. P. Shi, "Covariance-based spectrum sensing for noncircular signal in cognitive radio networks with uncalibrated multiple antennas," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 662–665,2020.
[12] F. Obite, A. D. Usman, and E. Okafor, "An overview of deep reinforcement learning for spectrum sensing in cognitive radio networks," *Digital Signal Processing*, vol. 113, 103014, 2021.
[13] H. Sun, A. Nallanathan, C. X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 74–81,2013.
[14] M. B. Usman, R. S. Singh, and S. Rajkumar, "Stage spectrum sensing technique for cognitive radio network using energy and entropy detection," *Wireless Power Transfer*, 2022.
[15] P. Y. Dibal, E. N. Onwuka, J. Agajo, and C. O. Alenoghena, "Application of wavelet transform in spectrum sensing for cognitive radio: A survey," *Physical Communication*, vol. 28, pp. 45–57, 2018.
[16] W. Zhao, H. Li, M. Jin, Y. Liu, and S. J. Yoo, "Eigen values-based universal spectrum sensing algorithm in cognitive radio networks," *IEEE Systems Journal*, vol. 15, no. 3, pp. 3391–3402, 2020.
[17] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Communications and Mobile Computing*, 2018.
[18] E. Ghazizadeh, D. Abbasi-moghadam, and Nezamabadi-pour, "An enhanced two-phase SVM algorithm for cooperative spectrum sensing in cognitive radio networks," *International Journal of Communication Systems*, vol. 32, no. 2, e3856, 2019.
[19] H. B. Ahmad, "Ensemble classifier based spectrum sensing in cognitive radio networks," *Wireless Communications and Mobile Computing*, 2019.
[20] N. Gul, M. S. Khan, S. M. Kim, J. Kim, A. Elahi, and Z. Khalil, "Boosted trees algorithm as reliable spectrum sensing scheme in the presence of malicious users," *Electronics*, vol. 9, no. 6, 1038, 2020.
[21] J. Sheng, F. Liu, Z. Zhang, and C. Huang, "TCVQ-SVM algorithm for narrowband spectrum sensing," *Physical Communication*, vol. 50, 101502, 2022.
[22] J. Gao, X. Yi, C. Zhong, X. Chen, and Z. Zhang, "Deep learning for spectrum sensing," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1727–1730, 2019.
[23] J. Xie, J. Fang, C., Liu, and X. Li, "Deep learning-based spectrum sensing in cognitie radio: A CNN-LSTM approach," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2196–2200, 2020.
[24] H. Xing, H. Qin, S. Luo, P. Dai, L. Xu, and X. Cheng, "Spectrum sensing in cognitive radio: A deep learning based model," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 1, e4388,2022.
[25] S. U. R. E. N. D. R. A. Solanki, V. A. S. U. D. E. V. Dehalwar, J. Choudhary, M. L. Kolhe, and K. Ogura, "Spectrum sensing in cognitive radio using CNN-RNN and transfer learning," *IEEE Access*, vol. 10, pp. 113482–113492, 2022.
[26] R. Sarikhani and F. Keynia, "Cooperative spectrum sensing meets machine learning: Deep reinforcement learning approach," *IEEE Communications Letters*, vol. 24, no. 7, pp. 1459–1462, 2020.
[27] K. Du, P. Wan, Y. Wang, X. Ai, and H. Chen, "Spectrum sensing method based on information geometry and deep neural network," *Entropy*, vol. 22, no. 1, 94, 2020.
[28] B. M. Pati, M. Kaneko, and A. Taparugssanagorn, "A deep convolutional neural network based transfer learning method for non-cooperative spectrum sensing," *IEEE Access*, vol. 8, pp. 164529–164545, 2020.
[29] S. Zheng, S. Chen, P. Qi, H. Zhou, and X. Yang, "Spectrum sensing based on deep learning classification for cognitive radios," *China Communications*, vol. 17, no. 2, pp. 138–148, 2020.
[30] A. Nasser, M. Chaitou, A. Mansour, K. C. Yao, and H. Charara, "A deep neural network model for hybrid spectrum sensing in cognitive radio," *Wireless Personal Communications*, vol. 118, no. 1, pp. 281–299, 2021.
[31] S. B. Goyal, P. Bedi, J. Kumar, and V. Varadarajan, "Deep learning application for sensing available spectrum for cognitive radio: An ECRNN approach," *Peer-to-Peer Networking and Applications*, vol. 14, no. 5, pp. 3235–3249, 2021. doi: 10.1007/s12083-021-01169-4