REVIEW OF DEEP LEARNING MODEL FOR SPECTRUM SENSING IN UHF BANDS FOR 5G-TVWS NETWORK

Pallam, S.W., Thuku, I.T., Luka, M.K. and Ibrahim, V.M.*

Department of Electrical and Electronics Engineering, Faculty of Engineering,
Modibbo Adama University Yola, Adamawa State, Nigeria

*Corresponding email: visaibrahim@gmail.com

ABSTRACT

Conventional methods of spectrum sensing, especially energy detection, though simple to design, have certain limitations including high sensitivity to noise uncertainty, which affect the efficacy of spectrum sensing in practical applications. Renowned deep learning network models, including Convolutional Neural Network (CNN), Support Vector Machine (SVM), Long Short-Term Memory (LSTM), Short-Time Fourier Transform (STFT), and Artificial Neural Networks (ANN) have been utilized for extracting different characteristics of signals for spectrum-sensing tasks. Although these models have demonstrated superior performances compared to conventional spectrum-sensing methods, many of the deep learning approaches lack complete awareness of the underlying source signal structure. Most of the conventional and deep learning-based spectrum sensing algorithms seem to have high perceptual performance under high-SNR environments, but their performances often fall short of expectations under low-SNR environments. Due to difficulties in obtaining large training datasets from the field, several deep learning methods depend on computer-generated datasets to train the models, which cannot reflect the practical application scenarios. This paper reviews a deep learning spectrum sensing model for spectrum sensing in UHF bands for 5G-TVWS network using a hybrid CNN-STFT-LSTM network to obtain local information for a single node spectrum sensing. While the CNN effectively extracts the spatial features from the frequency domain representations produced by STFT, the LSTM captures temporal dependencies (patterns and anomalies) in the time series data. The STFT plays a crucial role of transforming time-domain signals into time-frequency representation (spectrogram), which serves as an input dataset to CNN and LSTM, enabling the extraction of spectral features that are essential for detecting and classifying signals. Hybridizing CNN, STFT and LSTM networks can enhance the model's ability to learn complex, non-linear relationships in data, which is particularly useful in the challenging environments of signal detection. This multi-feature combination can lead to improved accuracy in identifying the presence of signals in noisy environment and ultimately enhancing overall performance in spectrum sensing tasks. The proposed model can be integrated into 5G-TVWS network for delivery of broadband internet services to rural and underserved areas. The 5G-TVWS network is suitable for rural deployment of broadband internet services due to its long-distance transmission capabilities, obstacle penetration, large coverage area, and relatively little infrastructure requirement.

Keywords: Deep learning, Television white space, Spectrum sensing, Convolutional Neural Network, Long Short-Term Memory, Short Time Fourier Transform, 5G-TVWS network

1 INTRODUCTION

As demands for wireless communication continue to grow, efficient spectrum management becomes increasingly significant. This growth, coupled with the Federal Communications Commission (FCC) report that most of the usable Radio Frequencies (RF) have been licensed to commercial operators and government entities [1], may lead to scarcity of available spectrum resource. Since the early 20th century, RF regulators have presumed that new radio transmitters will interfere with other uses of radio spectrum, forcing every wireless service provider to request for exclusive license from the regulatory agencies, thereby leading to the "doctrine of spectrum scarcity" [2]. Consequently, the traditionally fixed spectrum allocation policies are getting outdated, and giving way to the concept of Opportunistic Spectrum Access (OSA) that allows the licensed spectrum to be accessed by unlicensed users when temporarily not being used. In reality, studies have shown that most of the licensed spectrum is grossly underutilized, with over 95% of white space spectrum availability in rural areas most of the time, while the unlicensed spectrum is saturated [3]. Moreover, the recent transition from Analog to Digital Television (DTV) has freed up a huge portion of the spectrum band and is available for opportunistic access. The freed-up spectrum along with other unused spectrum in the TV broadcast bands form the Television White Spaces (TVWS) [4]. The Nigerian Communications Commission guidelines outline that the unutilized spectrum in TV broadcast bands in the range of 470 to 694MHz can be utilized for TVWS networks to provide internet access to underserved and unserved regions of the country, mostly in rural service areas at broadband speed [5]. The Ultra High Frequency (UHF) bands, traditionally used for TV broadcasting, offer significant potential for opportunistic use by Fifth Generation (5G) networks through TVWS. The UHF operating in low-frequency band has excellent features that make it particularly attractive for delivery of wireless communication services to long distant areas, dealing with variations in terrain and environmental barriers [6]. Although 5G networks have some attractive features, including high data rate, low latency, high connectivity density and reliability, providing broadband internet access in rural areas have been a great challenge due to its high operational cost and limited coverage area [7]. While the 5G networks require extensive bandwidth resource to support high data rates, low latency, and improved connectivity across cities and regions, TVWS network exploits the unutilized spectrum in the licensed UHF TV broadcast bands to provide bandwidth broadband internet at relatively cheaper cost and extended coverage area. Indeed, covering large cities and extending 5G broadband services to rural areas require massive deployment of network infrastructure, which can be very costly. This challenge can be overcome by 5G networks having opportunistic access to the unused spectrum in the UHF TV broadcast bands, while ensuring zero harmful interference to the primary users [8]. However, TVWS varies with location and time, making it difficult to predict or obtain accurate information on frequencies occupied by primary licensed users and those that are vacant for secondary users [9]. Ensuring zero interference to licensed primary users poses a fundamental issue for efficient and safe utilization of TVWS spectrum. A reliable spectrum sensing technique is therefore essential for optimized spectrum utilization and cost-effective deployment of 5G-TVWS network in rural areas. Spectrum sensing, which primary goal is to detect the unused licensed spectrum, is one of the critical elements in a Cognitive Radio (CR) system. Effective spectrum sensing can significantly improve spectrum utilization and internet connectivity in rural and underserved areas. Several spectrum sensing techniques such as Energy Detection (ED), Matched Filtering (MF), Cyclo-Stationary Detection (CSD), and Likelihood Ratio Test (LRT) are used to detect the unused spectrum in the primary networks and reallocate them to secondary wireless communication networks with minimized interferences [10-11]. However, many of these models face certain challenges including issues of noise uncertainty [12], non-linearity [13], vanishing gradient [14], ineffective performance under low Signal-Noise Ratio (SNR) [15-16], non-adaptability to dynamic environment and limited generalization capabilities [17]. Moreover, many of the available deep learning-based sensing methods rely on synthetic datasets for their model training, which do not accurately represent real-world applications. A more robust spectrum sensing method, such as deep learning model based on combination of Convolutional Neural Network (CNN), Short Time Fourier Transform (STFT), and Long Short-Term Memory (LSTM) networks, is therefore very essential for

identifying and reallocating the unused TV spectrum to secondary users, while ensuring that primary networks are not interrupted.

2 RADIO FREQUENCY CHANNELING

Radio frequency which ranges from 3KHz to 300GHz is broadly divided into fixed number of different frequency channels in which the segments of the band are allocated for wireless communication services such as broadcasting signals for TV, radio, mobile, Wi-Fi, radar systems, GPS, and satellite transponders. The International Telecommunication Union (ITU), a United Nations specialized Agency for Information and Communications Technologies (ICTs), is responsible for allocating global radio spectrum and satellite orbits, setting technical standards, and promoting global ICT access in 194 member states. For the purpose of global radio spectrum management, ITU has divided the world into three regions as presented in table 1.

Table 1. International Telecommunication Union Regions [18]

	ITU-Region 1	ITU- Region 2	ITU- Region 3
Member Nations	Europe, Africa, Middle East & Russia	North America and South America	Asia, Pacific, Australia, and Oceania

Similarly, for efficient utilization of TV frequencies in Nigeria, the country is subdivided into ten TV broadcast zones, namely: Abuja, Benin, Enugu, Ibadan, Jos, Kaduna, Lagos, Maiduguri, Sokoto, and Uyo. A TV broadcast frequency allocated to a broadcasting station in one zone can be re-allocated to another station in another zone. In 2019, the National Frequency Allocation Table outlined the range of frequencies allocated for broadcasting in ITU-R1, extending from 47-960MHz [19]. These frequencies are divided into 5 bands as presented in Table 2.

Table 2. ITU-R1 broadcast band division

Band No.	Freq. Range (MHz)	Channel Bandwidth	Spectrum Band	Spectrum Service
1	87.5 – 108	108KHz	VHF	FM radio broadcasting
2	47 – 68	7 MHz	VHF	Cordless land mobile broadcast
3	174–230	7 MHz	VHF	TV broadcasting on VHF
4	470 – 582	8 MHz	UHF	TV broadcasting on UHF
5	582-960	8 MHz	UHF	TV broadcasting on UHF

The frequency channeling in the VHF and UHF TV broadcasting bands are presented in Tables 3 and 4 respectively.

Table 3. VHF band frequency channeling for TV broadcast in Nigeria

Band No.	Frequency Range (MHz)	Chan. No.	Channel Bandwidth	Centre Frequency	Spectrum Band	Type of Service
	174–181	5	7 MHz	177.5 MHz	VHF	TV broadcast
	181-188	6	7 MHz	184.5 MHz	VHF	TV broadcast
	188-195	7	7 MHz	191.5 MHz	VHF	TV broadcast
3	195-202	8	7 MHz	198.5 MHz	VHF	TV broadcast
	202-209	9	7 MHz	205.5 MHz	VHF	TV broadcast
	209-216	10	7 MHz	212.5 MHz	VHF	TV broadcast
	216-223	11	7 MHz	219.5 MHz	VHF	TV broadcast
	223-230	12	7 MHz	226.5 MHz	VHF	TV broadcast

Table 4. UHF band frequency channeling for TV broadcast in Nigeria

Band No.	Frequency Range (MHz)	Chan. No.	Channel Bandwidth	Centre Frequency	Spectrum Band	Type of Service
4	470 – 478	21	8 MHz	474 MHz	UHF	TV broadcast
	478 – 486	22	8 MHz	482 MHz	UHF	TV broadcast
	486 – 494	23	8 MHz	490 MHz	UHF	TV broadcast
	:	:	:	:	:	:
	:	:	:	:	:	:
	566 – 574	33	8 MHz	570 MHz	UHF	TV broadcast
	574 – 582	34	8 MHz	578 MHz	UHF	TV broadcast
5	582 – 590	35	8 MHz	586 MHz	UHF	TV broadcast
	590 – 598	36	8 MHz	594 MHz	UHF	TV broadcast
	598 – 606	37	8 MHz	602 MHz	UHF	TV broadcast
	606 – 614	38	8 MHz	610 MHz	UHF	Not allocated to TV
	614 – 622	39	8 MHz	618 MHz	UHF	TV broadcast
	:	:	:	:	:	:
	:	:	:	:	:	:
	678 – 686	47	8 MHz	682 MHz	UHF	TV broadcast
	686 – 694	48	8 MHz	690 MHz	UHF	TV broadcast
	694 – 702	49	8 MHz	698 MHz	UHF	TV broadcast
	:	:	:	:	:	:
	:	:	:	:	:	:
	838 – 846	67	8 MHz	842 MHz	UHF	TV broadcast
	846 – 854	68	8 MHz	850 MHz	UHF	TV broadcast
	854 – 862	69	8 MHz	858 MHz	UHF	TV broadcast

2.1 Television White Spaces

In wireless communication, white spaces are portions of radio spectrum that are vacant or unused by licensed primary users at specific times and locations. These vacant or interleaved frequencies of inactive channels can be utilized to provide high-speed broadband internet to rural areas at low cost. In broadcast industries, buffer channels are intentionally placed between two active TV broadcast channels to protect against interferences. The white spaces between two TV channels are usually 7MHz for VHF and 8MHz for UHF. In Nigerian TV broadcast system, VHF channel 5 begins transmission at 174MHz while channel 6 begins at 181MHz. Likewise, in UHF bands, channel 21 begins transmission at 470MHz while channel 22 begins transmission at 478MHz. These buffer frequencies together with the vacated frequencies arising from migration from analog to digital broadcasting form the Television white space (TVWS). Figure 1 shows an illustration of the availability of white spaces in broadcast TV bands.

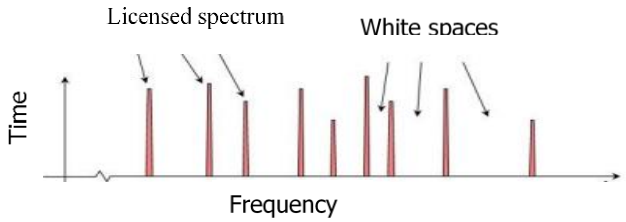

Figure 1. Illustration of white spaces in a licensed TV spectrum

2.2 Opportunistic Access by 5G in TVWS Network

The primary target of IEEE802.22 standard is to identify vacant spectrum bands in TV broadcast channels and utilize them for broadband network in rural areas. Cognitive radio (CR) system has been used to identify unused spectrum and maximize the use of limited radio bandwidth in wireless communication networks [20]. The need for Opportunistic Spectrum Access (OSA) was first proposed by FCC in 2003 as a result of limited spectrum resource for future wireless networks. The OSA model could accommodate CR equipped devices to safely coexist with licensed primary users without causing harmful interference. Cognitive radio can be implemented in TVWS network to safely provide broadband internet to rural areas without causing interference to the primary network users. This process involves spectrum sensing, decision making, resource sharing, and handoff. However, CR faces some technical challenges of sensing TV signals and providing reliable service in a dynamic environment when deployed over the TVWS [21]. With the explosive increase in mobile subscribers and the recent deployment of 5G network across some major cities in Nigeria, a better solution for extending coverage to rural areas can be provided. While the 5G technology has excellent features such as high data rate, ultra-low latency, and ultra-dense network capability, it is also characterized by limited coverage area, low penetration capability, and high deployment cost. On the other hand, TV band white spaces have excellent features including long-distance coverage, high penetration capability into environmental barriers, no-line-of-site (NLOS) transmission, low power consumption, high resilience to mobility, and low deployment cost. Combining the 5G and TVWS technologies will birth a better and cost-effective broadband services solution suitable for rural and unserved areas. The architecture of 5G-TVWS network is shown in Figure 2.

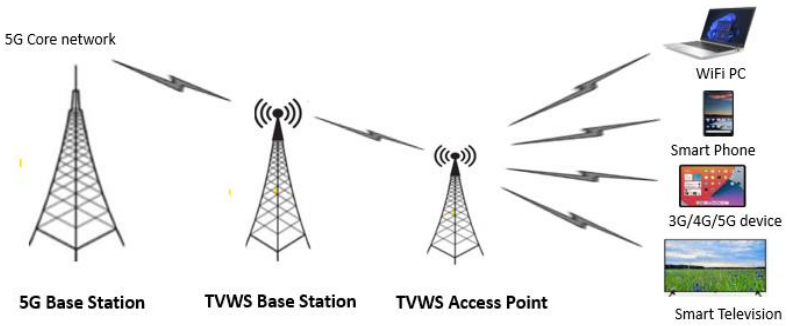

Figure 2. Architecture of 5G-TVWS Network

2.3 Deep Learning-Based Spectrum Sensing Techniques

To overcome the challenges of conventional spectrum sensing methods, several techniques such as Artificial Neural Networks (ANN) have exploited machine learning (ML) capabilities for detecting spectrum holes in the TV broadcast bands, using pertinent parameters such as channel SNR, channel capacity, bandwidth, and distance for model training. The performances of most of these networks depend on Radio Frequency (RF) parameters such as SNR threshold, causing the probability of detection to degrade due to noise uncertainty. In 2018, [22] proposed a machine learning-based spectrum sensing scheme using K-Nearest Neighbor algorithm, where each cognitive radio user either transmits or remains silent. The local

decisions are combined at the fusion center by a novel decision combination scheme that takes into account the reliability of the CR users. Based on the quantitative variables, the probability of each sensing class is calculated and classified into either presence or absence of PU signal. The probability of detection of the proposed scheme outperformed the conventional spectrum sensing schemes, with 80% of the times at -10 dB SNR in both fading and non-fading environments. However, the highest mean probability of detection that can be achieved using the proposed scheme is 0.7 using five CR users. A Covariance Matrix-aware CNN (CM-CNN) deep learning model, utilizing the sample Covariance Matrix (CM) as input to CNN structure was proposed in 2019 to mitigate the challenges faced by Energy Detection (ED) technique [23]. The experimental results show that the proposed CM-CNN model substantially outperforms the conventional method, such as Energy Detection in terms of probability of detection. However, the model faces certain challenges such as multi-dimensional input (both time-domain and frequency-domain data) handling, large clean training dataset requirement, and adaptability to dynamic environment issue. Improving on the feature extraction method can ensure continuous adaptation to changing network conditions, and optimize the effectiveness of Deep CM-CNN model. To optimize the probability of detection at low SNRs, [24] proposed a blind spectrum sensing method based on deep learning using a combination of three kinds of neural networks: CNN, LSTM, and Fully Connected Neural Networks (FCNN), in which prior knowledge of the signal characteristics is not required. Though the experimental results show that the method outperforms energy detector technique in terms of detection probability at low SNRs, the network does not have the capability of identifying which spectrum channel is vacant or occupied. To improve spectrum sensing performance and minimize the total error probability, a model for optimal threshold selection based on online learning algorithm was proposed by [25]. The model which was tested over different fading and non-fading channels under low SNRs shows little improvement in sensing performance. Moreover, a deep learning spectrum sensing model utilizing STFT-CNN network was proposed to obtain rich information from the time-frequency matrix for improving detection accuracy under low SNRs [26]. The simulation results indicate that the model achieved detection probability of 90.2% and false alarm probability of 10% at SNR of -15dB. However, STFT implements a universal time-frequency domain transformation, which does not vary with time and frequency, making it unadaptable to highly dynamic electromagnetic environment. Thus, this scheme experiences sharp decline in detection performance at lower SNRs. Within the last few years, deep learning has become the most widely used computational standard among the machine learning community for automation of feature learning and achieving substantive results in complex cognitive tasks [27]. In 2021, a new Two-Dimensional (2D) dataset of received signals was used to train three Cooperative Spectrum Sensing (CSS) schemes using classical CNN such as LeNet, AlexNet, and VGG-16 [28]. Furthermore, [29] proposed a CNN-based spectrum sensing and identification using Spectral Correlation Function (SCF), and utilizing the SCF as input to the CNN model. In this scheme, two separate scenarios were experimented: the first scenario involves training the CNN for signal classification by extracting cyclic spectra of Global System for Mobile communication (GSM), Universal Mobile Telecommunication Service (UMTS), Long Term Evolution (LTE), and the spatial spectrum, while the second involves inputting the cyclic spectra of relevant signals to the CNN for detection of PU. Once the PU is detected, the CNN classifier is activated to identify the type of PU signal, thereby enabling both spectrum sensing and signal identification in the model. The experimental results show better performance in challenging channels than those in Support vector machines and Cyclostationary feature detection. The ML algorithms such as Support Vector Machine (SVM), k-Nearest Neighbor (kNN), and Decision Tree (DT) were evaluated by [30] for designing spectrum sensing models to detect the existence of PU signal over the TV band. The authors incorporated the Principal Component Analysis (PCA) to speed up the learning process of the classifiers and reduce the duration of training while maintaining the good performance. To improve the feature extraction capabilities and performance of spectrum sensing techniques, a spectrum-sensing method based on Residual Networks (ResNets) using two-branch convolution was proposed [31]. The enhanced feature extraction capabilities led to higher detection probabilities and lower Bit Error Rates (BERs) than the traditional methods,

especially at low SNR ratios. However, achieving cognitive radio classification tasks using traditional machine learning methods require numerous successive steps, such as pre-processing, feature extraction, feature selection, learning, and classification. The performance of the machine learning models can be greatly impacted upon, leading to incorrect judgement between classes, if the feature selection is biased. To improve on the efficiency of spectrum sensing, [32] proposed an efficient CNN-based spectrum sensing model that directly uses the spectrogram of received signal as input to the CNN for feature extraction, and PU signals at different SNRs used for training of the network. The experiment was divided into two models. The experimental results of the first model with a layer of 32 convolution kernels and 40 epochs show the accuracy of the model dropping from 65% to 52% as SNR decreases from -5 dB to -20 dB respectively. As the number of layers gradually increases from the first model, the experimental results of the second model gradually achieved 98% accuracy and about 1% model loss, inferring that as the SNR of the PU signal decreases, the accuracy of the model is not affected, and the model can also achieve 98% accuracy at SNR = -20 dB, which is a large performance improvement compared with the first model. In 2022, [33] proposed a Spectrum-Sensing Convolutional Neural Network (SS-CNN) by using STFT generated spectrogram as input to the CNN, introducing Deep Convolutional Generative Adversarial Network (DCGAN) to generate augmenting data, and finally using total dataset for training the network. Since the proposed model is a data-driven model, an efficient threshold-based detection mechanism is designed for online testing, while the model without augmented data is used for offline testing. The experimental results show that the DCGAN-based SS-CNN has 10% accuracy improvement over the CNN model without data augmentation, and that the detection of the modulated signals using Bi-Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), and 8-Phase Shift Keying (8PSK) shows no difference in the performance, which suggests that the model is insensitive to the modulation order of the modulated signal. Moreover, [34] proposed a multi-band spectrum-sensing method based on a CNN with a covariance matrix, using the extracted feature of the covariance matrix of the multiband to effectively learn the correlation information between each sub-band. Since spectrum sensing for each sub-band individually leads to under-utilization of the information between two consecutive frequency bands, the technique seeks to achieve joint multi-band spectrum sensing by connecting all covariance matrices into the frequency band covariance matrix array. The experimental results show that the proposed model significantly outperforms the CM-CNN method under the SNR range -20 dB to -16 dB, and the previous traditional spectrum-sensing methods. [35] proposed a spectrum sensing method based on CNN and ResNet to predict the spatio-temporal spectrum usage in a large region using two scenarios: single transmitter and multiple transmitters. The experimental results show that the model achieved high PU prediction while maintaining lower error rate. Although CNN and ResNet architectures proved effective for spatio-temporal spectrum load prediction, they are often affected by challenges such as large quality dataset dependency, varying spatial and temporal resolution, environmental variability, and computational complexity which limit their robustness and applicability in real-world scenarios. To predict non-stationary signals at low SNRs across different frequency bands without the priori information about the signals, [36] proposed a WT-ResNet spectrum sensing model based on Wavelet Transform (WT) and Residual Network (ResNet). The simulation results indicate that the model achieved 85% detection probability and 10% false alarm probability at SNR of -16dB, outperforming other previous models. However, challenges such as wavelet function selection, noise sensitivity, generalization to different environments, overfitting on small datasets, latency and energy efficiency issues still prevail. To improve the robustness of signal detection in dynamic topology, a CNN-LSTM model utilizing long-short-term memory (LSTM) network and the CNN structure for cognitive vehicular network was proposed by [37]. While the CNN extracts spatial features, the LSTM processes temporal or sequential data. The model which is particularly effective for tasks like video classification, time series prediction, and natural language processing eliminates the impact of predefined threshold value setting prevalent in the conventional sensing methods, achieving satisfactory detection accuracy, robustness in low SNR environments and low sensing time compared to traditional methods. However, some unique features such as high speed and dynamic topology pose a serious challenge to this model, leading to a

mismatch between the spatial features extracted by CNN layers and the temporal patterns that LSTM layers are designed to process. When the spatial dimensions are over decreased in CNN, LSTM will be left with insufficient information to model the temporal dynamics, resulting to degraded performance of the network. Other challenges include difficulty in handling long time dependencies in sequential data, and sensitivity to sequence length. Table 5 shows a classification of the research papers based on the specified criteria (ML Technique, Feature Extraction, Spectrum Sensing Approach, Performance Metrics, Dataset, and Computational Efficiency).

Table 5. Classification of the research papers based on various criteria

Study	ML Technique	Feature Extraction	Spectrum Sensing Approach	Performance Metrics	Dataset	Computational Efficiency
[22]	K-Nearest Neighbour (kNN)	Quantitative variables (SNR, reliability)	Cooperative sensing with fusion center	Detection probability (0.7 at -10 dB)	Not specified	Moderate (fusion center decision combination)
[23]	CM-CNN (Deep Learning)	Sample covariance matrix	Covariance matrix-based detection	Outperforms Energy Detection (ED)	Not specified	Lower (multi-dimensional input handling challenges)
[24]	Hybrid CNN-LSTM-FCNN	Raw signals (blind sensing)	Blind sensing (no prior signal knowledge)	Better detection probability than ED at low SNRs	Not specified	High (complex hybrid architecture)
[25]	Online learning algorithm	SNR, error probability	Threshold-based optimization	Minimal improvement in sensing performance	Tested over fading/non-fading channels	Efficient (online learning framework)
[26]	STFT-CNN (Deep Learning)	Time-frequency matrix (STFT)	Time-frequency domain analysis	90.2% detection probability, 10% false alarm at -15 dB SNR	Not specified	Moderate (fixed STFT resolution limits adaptability)
[28]	Classical CNN (LeNet, AlexNet)	2D received signals	Cooperative spectrum sensing (CSS)	98.78% detection probability, 1% false alarm at -15 dB SNR	2D dataset of received signals	Varies (depends on CNN architecture, e.g., VGG-16 is computationally heavy)
[29]	CNN	Spectral Correlation Function (SCF)	Cyclostation-ary feature detection	Outperforms SVM and cyclostationary detectors in challenging channels	GSM, UMTS, LTE signals	Moderate (SCF extraction and classification)
[30]	SVM, kNN, DT (Classical ML)	PCA-reduced features	Feature selection-based sensing	Maintained accuracy with reduced training time	TV band signals	Improved (PCA reduces feature dimensionality)
[31]	ResNet (two-branch CNN)	Enhanced feature extraction	Feature learning-based detection	Higher detection probability, lower BER than traditional methods at low SNRs	Not specified	Efficient (improved feature extraction)
[32]	CNN	Spectrogram of received signals	Direct spectrogram input for classification	98% accuracy at -20 dB SNR	Spectrograms of PU signals	High (deeper layers improve efficiency)
[33]	SS-CNN with DCGAN	STFT spectrogram + augmented data	Data-augmented sensing	10% accuracy improvement over non-augmented CNN	DCGAN-augmen-ted dataset	Moderate (offline/online detection mechanisms)
[34]	CNN with covariance matrix	Multi-band covariance matrix	Joint multi-band sensing	Outperforms CM-CNN and traditional methods (-20 dB to -16 dB SNR)	Multi-band signals	Moderate (joint covariance matrix processing)
[35]	CNN + ResNet	Spatio-temporal spectrum data	Spatio-temporal prediction	High prediction accuracy, low error rate	Spatio-temporal data	Lower (computational complexity limits robustness)
[36]	WT-ResNet	Wavelet Transform (WT)	Non-stationary signal detection	85% detection probability, 10% false alarm at -16 dB SNR	Not specified	Lower (sensitive to wavelet function selection)
[37]	CNN-LSTM	Spatial-temporal features	Cognitive vehicular network sensing	Robust detection at low SNR, low sensing time	Vehicular network data	Moderate (challenges in dynamic topology and sequence-length dependency)

3 THE CNN ARCHITECTURE

Convolutional Neural Network (CNN) is a type of Deep Learning (DL) neural network that has become one of the most popular neural network architectures for image recognition and is predominantly used to extract the feature from the grid-like matrix dataset like images or videos where data patterns play

an extensive role. It consists of a convolution layer that extracts the features of the image and a fully connected layer that determines which class the input image belongs to. CNN has the capabilities of handling spatial characteristics of radio frequency signals, turning them into detectable signals. Thus, it is a promising DL solution that can improve effectiveness and precision in spectrum sensing. It mainly operates on the principle of convolutional layers, where a set of learnable filters convolve around the input data to produce feature maps [38]. These feature maps undergo pooling operations, which reduce their spatial dimensions while retaining crucial information. The stacked convolutional and pooling layers enable the network to recognize progressively complex patterns, making it capable of identifying complicated characteristics of signals in spectrum sensing [39]. The radio frequency signals can then be visualized as a single-dimensional sequence, where spatial attributes can be extracted effectively. This facilitates the recognition of complicated frequency patterns, thereby enhancing the precise detection of primary users in various spectral conditions. In a regular CNN model, there are three types of layers: Input Layer which takes in a time-series or image representation of the signal. The number of neurons in this layer is equal to the total number of features in the data (e.g. number of pixels in an image). Hidden Layer that receives input from the Input layer. There can be many hidden layers depending on the model and data size. Each hidden layer can have different numbers of neurons which are generally greater than the number of features. The output from each layer is computed by matrix multiplication of output of the previous layer with learnable weights of that layer and then by the addition of learnable biases followed by activation function which makes the network nonlinear. The hidden layers in a CNN model are convolutional Layers, Pooling Layers, Rectified linear unit (ReLu), and Fully Connected Layers. The feature detection layers of the model include convolution, pooling, and Rectified linear unit (ReLu). These three main operations are repeated on hundreds of hidden layers with each layer learning to detect a different feature. The Output Layer produces the final prediction of channel availability. The output of the hidden layer is fed into a logistic function such as Sigmoid, SoftMax or classification which converts the output of each class into the probability score of each class. While multiple convolution layers may more effectively learn images with complex features, the number and performance of the convolutional layers are not always proportional. Since the correlation between the number of convolutional layers and its performance depends on the characteristics of the input images [40], increasing the depth of a CNN can improve its ability to sense and classify signals in the spectrum by learning more complex features. However, it also introduces potential challenges such as overfitting and increased computational demands. Balancing these factors is key to optimizing spectrum sensing performance. Figures 3 shows the convolutional neural network architecture.

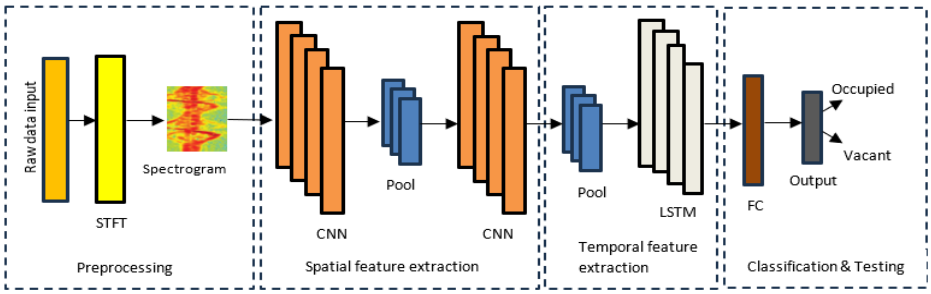

Figure 3. CNN architecture

3.1 Short Time Fourier Transform

The STFT is a linear transform algorithm that transforms time-dependent representations into two-dimensional (2D) time-frequency representation known as spectrogram, transforming the spectrum-sensing problem into an image recognition problem. The signals with different modulations have different spectrogram distributions with unique properties, letting the CNN pull out the most important features of the signal for the classification task [41]. The formula for the STFT is shown in equation (1).

$$STFT = \int_{-\infty}^{\infty}[z(u)g(u-t)]e^{-j2\pi fu}du \quad (1)$$

where z(u) is the source signal and g(u-t) the window function.

3.2 Long Short-Term Memory

The LSTM layer extracts temporal features (attributes of data related to time). The layer tries to "remember" all the past knowledge that the network has seen so far and to "forget" irrelevant data. This is done by introducing different activation function layers called "gates" for different purposes. LSTM Network consists of four different gates for different purposes as follows: Forget Gate (f): At forget gate the input is combined with the previous output to generate a fraction between 0 and 1, that determines how much of the previous state need to be preserved or how much of the state should be forgotten. This output is then multiplied with the previous state. Note: An activation output of 1.0 means "remember everything" and activation output of 0.0 means "forget everything." Input Gate (i): Input gate operates on the same signals as the forget gate, but here the objective is to decide which new information is going to enter the state of LSTM. The output of the input gate (again a fraction between 0 and 1) is multiplied by the output of tanh block that produces the new values that must be added to previous state. This gated vector is then added to previous state to generate current state. Input Modulation Gate (g): It is used to modulate the information that the Input gate will write onto the Internal State Cell by adding non-linearity to the information and making the information Zero-mean. Output Gate (o): At output gate, the input and previous state are gated as before to generate another scaling fraction that is combined with the output of tanh block that brings the current state.

3.3 Spectrum Data Streaming

While conducting the RF sweep test, spectrum data is streamed to separate text files according to the number of UHF TV broadcast channels in the study areas. The received signal strength of these broadcasting stations is measured using a setup of Software Defined Radio (SDR) device, an omnidirectional antenna, Global Positioning System (GPS), and an internet enabled Laptop computer running Windows 11 Operating system, Python programming language, Air Sleuth Spectrum Analyzer, MATLAB and Simulink computing software. Depending on the antenna, the SDR captures electromagnetic signal transmitted within the operating range of the tuner, down-converts the RF signal to baseband signal, digitizes and outputs the baseband signal samples at the USB terminal. Figure 4 shows the setup of materials for the measurement and processing of the Received Signal Strength (RSS).

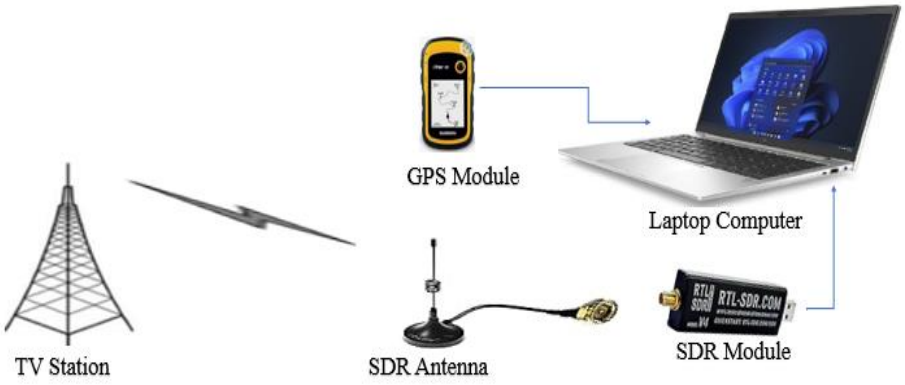

Figure 4. RF measurement materials

The parameters captured during the measurement are date, time, receiver coordinates, and Received Signal Strength (RSS). Table 6 presents a streamed raw dataset of Startimes TV station channel 23 during the RF sweep campaign in Yola North local government area of Adamawa state. The RF sweep is conducted using three different modalities: stationary, walk-across and drive-through tests.

Table 6. Startimes channel 23 TVWS data streaming record during RF sweep test

Date	Time	Latitude	Longitude	Altitude	Speed (m/s)	Power (W)	Power (dBm)	RSS (dBm)
06-Jan-25	16:16:48	9.30187	12.48432	199.8	0.04836	0.00219	3.40407	-26.29593
06-Jan-25	16:17:36	9.30186	12.48431	199.2	0.06482	0.00169	2.28694	-27.41306
06-Jan-25	16:18:25	9.30185	12.48429	194.8	0.03395	0.00159	2.00667	-27.69333
06-Jan-25	16:19:22	9.30185	12.48428	198.5	0.01235	0.00113	0.51257	-29.18743
06-Jan-25	16:20:10	9.30187	12.48430	197.0	0.03498	0.00239	3.7891	-25.91090
06-Jan-25	16:20:59	9.30187	12.48429	193.7	0.07717	0.00186	2.69500	-27.00500
06-Jan-25	16:21:48	9.30185	12.48427	200.7	0.08848	0.00233	3.67996	-26.02004
06-Jan-25	16:22:36	9.30183	12.48426	210.2	0.02315	0.00108	0.32165	-29.37835
06-Jan-25	16:23:25	9.30184	12.48426	209.5	0.05042	0.00130	1.13960	-28.56040
06-Jan-25	16:24:13	9.30186	12.48426	200.7	0.06842	0.00144	1.58438	-28.11562
06-Jan-25	16:25:02	9.30188	12.48425	189.6	0.06122	0.00119	0.77307	-28.92693
06-Jan-25	16:25:50	9.30187	12.48427	190.9	0.0782	0.00121	0.83030	-28.86970
06-Jan-25	16:26:39	9.30185	12.48427	195.9	0.04218	0.00131	1.18229	-28.51771
06-Jan-25	16:27:46	9.30184	12.48426	203.7	0.05247	0.00105	0.20236	-29.49764
06-Jan-25	16:28:34	9.30185	12.48427	206.6	0.03755	0.00118	0.70147	-28.99853
06-Jan-25	16:29:23	9.30184	12.48428	206.1	0.13221	0.00136	1.33348	-28.36652
06-Jan-25	16:30:11	9.30182	12.48424	206.8	0.08900	0.00131	1.16622	-28.53378
06-Jan-25	16:31:00	9.30181	12.48425	205.1	0.03035	0.00089	-0.51648	-30.21648
06-Jan-25	16:31:49	9.30180	12.48426	202.8	0.15073	0.00113	0.54218	-29.15782
06-Jan-25	23:54:02	9.23908	12.43830	229.0	0.02778	0	-32.69102	-62.39102
06-Jan-25	23:54:51	9.23911	12.43834	235.0	0.31433	0	-32.88559	-62.58559

3.4 Feature Extraction Process

Feature extraction aims to identify the most relevant features from the preprocessed datasets. This can be achieved by incorporating the attention mechanism layers (such as self-attention and channel attention) in the CNN architecture. These layers allow the model to focus on specific parts of the input data by assigning different weights to them, prioritizing the extraction of relevant features within the preprocessed data while suppressing the irrelevant ones. The feature extraction process in the proposed CNN-STFT-LSTM spectrum sensing model involves the following key steps. Application of STFT to transform the time-domain RF signal into the frequency domain, allowing for the extraction of relevant spectral features. This representation captures both frequency and time information, resulting in a spectrogram. Feeding the spectrogram into the CNN, which employs 2D convolutions to detect patterns, features, and relevant structures in the spectral data. CNN utilizes convolutional filters to capture patterns such as edges, textures, and shapes for feature representations. Adding an activation function to the input of the preceding layer introduces non-linearity into the network. This helps the model to learn complex patterns by transforming the input to the next layer while maintaining the output size. The ReLU layer applies an element-wise activation function to the output of CNN layer and outputs zero for any negative input and the input value itself for positive values. This allows for efficient training and prevents the vanishing gradient problem commonly seen in other activation functions. Incoporating pooling layer into the model helps in dimensionality reduction by decreasing the spatial dimensions of the feature maps while retaining important features, reducing computational load and making the model effective in extracting sophisticated characteristics from the input data. Pooling can be implemented by either Max_pooling or Average_pooling, making the computation faster, reducing memory size, and preventing overfitting. Integrating the LSTM layers to learn temporal dependencies in the data can effectively capture the sequential nature of the features extracted from the CNN, allowing the model to make predictions based on time-series information. Fully Connected (FC) Layer combines the features extracted by previous layers

to perform high level reasoning and classification task. Output Layer from the fully connected layers is fed into a logistic function for classification tasks like sigmoid or Softmax which converts the output.

4 EVALUATION METRICS FOR DEEP LEARNING MODEL

To evaluate the proposed deep learning model, any of the following performance metrics such as accuracy, precision, recall, and F1-score can be used. Accuracy is defined as the measure of all the correctly predicted samples. It is mostly used when all the classes are equally important and it is calculated as:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{2}$$

where TP (True Positive) is the outcome where the model correctly predicts samples that belong to the designated class, e.g., class 2, TN (True Negative) is an outcome where the model correctly predicts samples that do not belong to that particular class, e.g., class 2, FP (False Positive) is an outcome where the model incorrectly predicts samples of other classes as samples of class 2, while FN (False Negative) is an outcome where the model incorrectly predicts samples of the designated class, e.g., class 2 as they belong to other classes. Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives. It shows the ability of a classification model to return only actual samples belong to the class.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3}$$

Recall is defined as the percentage of actual positives that are correctly identified. It is also called the true positive rate (TPR), the sensitivity, or the probability of detection and is calculated as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{4}$$

F1-score is the harmonic mean of precision and recall and gives a measure of how many times a model makes correct predictions across the entire dataset. It is calculated as:

$$\text{F1} - \text{score} = 2 * \frac{\text{Precision}*\text{Recall}}{\text{Precision+Recall}} = \frac{2\text{TP}}{2\text{TP+FP+FN}} \tag{5}$$

5 CONCLUSIONS

In this paper, we have studied the problems of the existing deep learning spectrum sensing models and developed a more robust deep learning model based on CNN-STFT-LSTM network. The proposed model is capable of effectively learning complex features in a multi-channel environment and can identify occupied and vacant channels in TVWS network for opportunistic access by 5G networks. This demonstrates a substantial improvement in spectrum sensing tasks over the existing models, especially at low SNRs. Implementing the proposed model in a 5G-TVWS network can significantly enhance the broadband internet penetration in rural areas.

REFERENCES

1. FCC Fact Sheet (2020). Unlicensed White Space Device Operations in the Television Bands. *Federal Communications Commission*, *FCC-CIRC2010-03* (ET Docket No. 20-36), pp. 1-51.
2. Staple, G. and Werbach, K. (2004). The End of Spectrum Scarcity. *IEEE Spectrum*, vol. 41, no. 3, pp. 48-52.
3. Alonso, R.M., Guerra, A.C., Pupo, E.F., Plets, D., Nieto, G.G., Martens, L. and Joseph, W. (2020). Assessment of White Spaces Quality in Rural Areas: A large-scale spectrum survey. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2020, pp. 1-5.
4. Zhang, W., Yang, J., Zhang, G., Yang, L. and Kiat Yeo, C. (2018). TV white space and its applications in future wireless networks and communications: A survey. *IET Communications*, vol. 12, no. 20, pp.

2521-2532.

5. Danbatta, U.G. (2019). Guidelines on the use of Television White Spaces. *Nigerian Communications Commission*, pp. 1-20.
6. Suarez, M. (2020). In the Matter of Unlicensed White Space Device Operations in the Television Bands: Reply Comments of Dynamic Spectrum Alliance. *Dynamic Spectrum Alliance*, pp. 1-21.
7. Ferreira, A., Mendes, L., Dias, W., Marins, T., Gaspar, D., Matos, A., Silva, C. and Sokal, B. (2019). 5G-RANGE Project Field Trial. *2019 European Conference on Networks and Communications (EuCNC), Valencia, Spain*, vol. 20, pp. 490-494.
8. Kim, S. (2021). A New Bargaining Game-Based Unlicensed Spectrum Sharing Scheme for TVWS Platform. *IEEE Access*, vol. 9, pp. 95528-95537.
9. Abu, K., Ikpotokin, C., Igberaese, E.S. and Ehigie, E. (2023). Design and Implementation of Geo-location Database for TV White Spaces in Edo State, Nigeria. *Internal Journal of Latest Technology in Engineering, Management and Applied Science*, vol. 9, issue 5, pp. 10-23.
10. Brito, A., Sebastiao, P. and Fernando, J.V. (2021). Hybrid Matched Filter Detection Spectrum Sensing. *IEEE Access*, vol. 9, pp. 165504 -165516.
11. Luo, J., Zhang, G. and Yan, C. (2022). An Energy Detection-Based Spectrum-Sensing Method for Cognitive Radio. *Wireless Communications and Mobile Computing*, pp. 1-10.
12. Singh, B.K. and Mukhopadhyay, M. (2021). Energy Detector Based Spectrum Sensing Performance Analysis Over Fading Environment. *JSIR*, vol. 8, no. 3. http://nopr.niscpr.res.in/handle/123456789/56474.
13. Gai, J., Zhang, L. and Wei, Z. (2022). Spectrum Sensing Based on STFT-ImpResNet for Cognitive Radio. *Electronics*, vol. 11, issue 15, pp. 1-14.
14. Chandra, S.S., Upadhye, A., Saravanan, P., Gurugopinath, S. and Muralishankar, R. (2022). Deep Neural Network Architectures for Spectrum Sensing Using Signal Processing Features. *IEEE Conference Publication, IEEE Xplore*, pp. 129-134.
15. Mahajan, K. and Garg, (2022). An Enhancement to the Existing Cyclostationary Feature Detection in CRN. *AIoT Support and Applications with Multimedia*, vol. 81, pp. 37087-37099.
16. Wang, Q. and Guo, B. (2022). CNN-SVM spectrum sensing in cognitive radio based on signal covariance matrix. *Journal of Physics, Conference Series*, vol. 2395, no. 1, pp. 1-7.
17. Chae, K. and Kim, Y. (2023). DS2MA: A Deep Learning-Based Spectrum Sensing Scheme for a Multi-Antenna Receiver. *IEEE Journals & Magazine, IEEE Xplore*, vol. 12, issue 6, pp. 952-956.
18. Ohishi, M. (2024). The ITU Structure and the ITU Study Groups. *National Astronomical Observatory of Japan*, pp. 104-113.
19. National Frequency Management Council of the Federal Republic of Nigeria (2019). National Frequency Allocation Table, 192p.
20. Venkatesh, G.M., Latha, K.S., Reddy, Y.N., Vasaiah, T. and Varna Kumar Reddy, P.G. (2021). Simulation and Analysis of Spectrum Sensing in Cognitive Radio. *International Journal of Electronics, Communication & Instrumentation Engineering Research and Development*, vol. 11, no. 2, pp. 1-12.
21. Stevenson, C.R., Chouinard, G., Lei, Z., Hu, W., Shellhammer, S.J. and Caldwell, W. (2009). IEEE 802.22: The first cognitive radio wireless regional area network standard. *IEEE Communications Magazine*, vol. 47, no. 1, pp. 130-138.

22. Shah, H.A. and Koo, I. (2018). Reliable Machine Learning Based Spectrum Sensing in Cognitive Radio Networks. *Wireless Communications and Mobile Computing*, pp. 1-17.

23. Liu, C., Wang, J., Liu, X. and Ying-Chang, L. (2019). Deep CM-CNN for Spectrum Sensing in Cognitive Radio. *IEEE Journal on Selected Areas in Communications*, vol. 37, issue 10, pp. 2306-2321.

24. Yang, K., Huang, Z., Wang, X. and Li, X. (2019). A Blind Spectrum Sensing Method Based on Deep Learning. *Sensors*, vol. 19, issue 10, pp. 1-17.

25. kockaya, K. and Develi, I. (2020). Spectrum sensing in cognitive radio networks: Threshold optimization and analysis. *EURASIP Journal on Wireless Communications and Networking*, pp. 1-19.

26. Chen, Z., Yi-Qun, X., Wang, H. and Guo, D. (2020). Deep STFT-CNN for Spectrum Sensing in Cognitive Radio, vol. 25, issue 3, pp. 864-868.

27. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, vol. 8, no. 1, pp. 1-74.

28. Tan, Y. and Jing, X. (2021). Cooperative Spectrum Sensing Based on Convolutional Neural Networks. *Applied Sciences*, vol. 11, issue 10, pp. 1-13.

29. Tekbıyık, K., Akbunar, O., Ekti, A.R., Gorcin, A., Kurt, G.K. and Khalid, A.Q. (2021). Spectrum Sensing and Signal Identification with Deep Learning Based on Spectral Correlation Function. in *IEEE Transactions on Vehicular Technology*, vol. 70, issue 10, pp. 10514-10527.

30. Mohammad, A., Awin, F. and Abdel-Raheem, E. (2022). Case study of TV spectrum sensing model based on machine learning techniques. *Ain Shams Engineering Journal*, vol. 13, issue 2, pp. 1-9.

31. Gai, J., Xue, X., Li, Z. and Zhang, L. (2022). Spectrum Sensing Method Based on Residual Cellular Network. *IEEE Access*, vol. 10, pp. 61354-61365.

32. Elshafie, H., Fisal, N., Abbas, M., Hassan, W., Mohamad, H., Ramli, N., Jayavalan, S. and Zubair, S. (2015). A survey of cognitive radio and TV white spaces in Malaysia. *Transactions on Emerging Telecommunications Technologies*, vol. 26, issue 6, pp. 975-991.

33. Cai, L., Cao, K., Wu, Y. and Zhou, Y. (2022). Spectrum Sensing Based on Spectrogram-Aware CNN for Cognitive Radio Network. *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2135-2139.

34. Zhang, J., He, Z.Q., Rui, H. and Xu, X. (2022). Multiband Joint Spectrum Sensing via Covariance Matrix-Aware Convolutional Neural Network. *IEEE Communications Letters*, vol. 26, no. 7, pp. 1578-1582.

35. Ren, X., Mosavat-Jahromi, H., Cai, L. and Kidston, D. (2022). Spatio-Temporal Spectrum Load Prediction Using Convolutional Neural Network and ResNet. *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 502-513.

36. Zhen, P., Zhang, B., Chen, Z., Guo, D. and Ma, W. (2022). Spectrum Sensing Method Based on Wavelet Transform and Residual Network. *IEEE Wireless Communications Letters*, vol. 11, issue 12, pp. 2517-2521.

37. Lu, L., Li, X. and Ni, W. (2023). Multiband Cooperative Spectrum Sensing Meets Vehicular Network: Relying on CNN-LSTM Approach. *Wireless Communications and Mobile Computing*, pp. 1-8.

38. Shen, F., Wang, Z., Ding, G., Li, K. and Wu, Q. (2022). 3D Compressed Spectrum Mapping with Sampling Locations Optimization in Spectrum-Heterogeneous Environment. *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 326-338.

39. Kumar, A. and Himanshu Sharma, H. (2022). Intelligent Cognitive Radio Spectrum Sensing Based on

Energy Detection for Advanced Waveforms. *Radioelectronics and Communications Systems,* vol. 65, pp. 149-154.

40. Kim, J., Kim, J., Kim, H., Shim, M. and Choi, E. (2020). CNN-Based Network Intrusion Detection Against Denial-of-Service Attacks. *Electronics,* vol. 9, no. 6, pp. 1-21.

41. Zhang, Y. and Luo, Z. (2023). A Review of Research on Spectrum Sensing Based on Deep Learning. *Electronics,* vol. 12, no. 21, pp. 1-42.