

Received 1 July 2023, accepted 23 July 2023, date of publication 15 August 2023, date of current version 25 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3305388

 SURVEY

Deep Neural Networks for Spectrum Sensing: A Review

SADAF NAZNEEN SYED¹, (Graduate Student Member, IEEE), PAVLOS I. LAZARIDIS¹, (Senior Member, IEEE), FAHEEM A. KHAN¹, (Member, IEEE), QASIM ZEESHAN AHMED¹, (Member, IEEE), MARYAM HAFEEZ¹, (Member, IEEE), ANTONI IVANOV^{1,2,3}, (Member, IEEE), VLADIMIR POULKOV^{1,2,3}, (Senior Member, IEEE), AND ZAHARIAS D. ZAHARIS⁴, (Senior Member, IEEE)

¹School of Computing and Engineering, University of Huddersfield, HD1 3DH Huddersfield, U.K.

²Faculty of Telecommunications, Technical University of Sofia, 1000 Sofia, Bulgaria

³Research and Development and Innovation Consortium, Sofia Tech Park, 1784 Sofia, Bulgaria

⁴School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece

Corresponding author: Pavlos I. Lazaridis (p.lazaridis@hud.ac.uk)

This work was supported by the European Union through the Horizon 2020 Marie Skłodowska-Curie Innovative Training Networks Programme “Mobility and Training for beyond 5G Ecosystems (MOTOR5G)” under grant agreement no. 861219. The authors would like to acknowledge the support of the Intelligent Communication Infrastructures Laboratory, Research and Development and Innovation Consortium, Sofia Tech Park, Bulgaria.

ABSTRACT As we advance towards 6G communication systems, the number of network devices continues to increase resulting in spectrum scarcity. With the help of Spectrum Sensing (SS), Cognitive Radio (CR) exploits the frequency spectrum dynamically by detecting and transmitting in underutilized bands. The performance of 6G networks can be enhanced by utilizing Deep Neural Networks (DNNs) to perform SS. This paper provides a detailed survey of several Deep Learning (DL) algorithms used for SS by classifying them as Multilayer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, combined CNN-LSTM architectures and Autoencoders (AEs). The works are discussed in terms of the input provided to the DL algorithm, data acquisition technique used, data pre-processing technique used, architecture of each algorithm, evaluation metrics used, results obtained, and comparison with standard SS detectors. This survey further provides an overview of traditional Machine Learning (ML) algorithms and simple Artificial Neural Networks (ANNs) while highlighting the drawbacks of conventional SS approaches for completeness. A description of some publicly available Radio Frequency (RF) datasets is included and the need for comprehensive RF datasets and Transfer Learning (TL) is discussed. Furthermore, the research challenges related to the use of DL for SS are highlighted along with potential solutions.

INDEX TERMS Autoencoders, cognitive radio, convolutional neural networks (CNNs), deep learning (DL), deep neural networks (DNNs), machine learning (ML), multilayer perceptrons (MLPs), long short-term memory (LSTM) networks, spectrum sensing, 6G.

I. INTRODUCTION

The global network traffic continues to rise exponentially due to the growing popularity and ease of accessing wireless devices [1], [2], [3], [4]. As per a forecast by Ericsson [5], the number of mobile subscriptions will rise from about 8.4 billion in 2022 to around 9.2 billion by the end of 2028. The heterogeneous demands of the emerging applications like

The associate editor coordinating the review of this manuscript and approving it for publication was Usama Mir¹.

Internet of Everything, Holographic Telepresence, Extended Reality, Industry 5.0, and Intelligent Healthcare have motivated the development of 6G technology [6]. The possibility of finding a new vacant spectrum to accommodate these service requirements has become an increasingly difficult and expensive process [7]. The problem of spectrum scarcity can be attributed to the traditional fixed spectrum allocation policies which allocated a large part of the radio spectrum to licensed Primary Users (PUs) and left a smaller portion of the spectrum for unlicensed Secondary Users (SUs) [8], [9].

These allocation techniques permitted only the PUs to utilize the spectrum despite the resources being idle [10]. Since a considerable portion of the expensive frequency resources provided by wireless systems are not always utilized in many regions, the static spectrum assignment of most current and legacy networks is quite inefficient [11]. These techniques inherently protect the PUs from interference but can deny critical radio requirements of SUs even when multiple bands of PU are unused. With the massive rise in the number of network devices, static allocation techniques will rapidly exhaust the radio resources.

Dynamic Spectrum Management (DSM) is a flexible spectrum allocation technique that allows the SUs to access the PU spectrum if it is idle, or to even share the PU spectrum if the transmission of PU is protected from interference [12]. Cognitive Radio (CR) is the principal enabler of DSM [13] and ensures that the spectrum is utilized in an efficient manner by allowing the SUs to opportunistically access the idle frequency bands of the PUs [14]. The term 'CR' was first proposed by Joseph Mitola [15] and Simon Haykin introduced the basic cognitive cycle in [16]. Wireless regional area network (IEEE 802.22) is the first international standard based on CR, and the CR technology is now being used in several standards, including Zigbee (IEEE 802.15.4) and Wi-Fi (IEEE 802.11) [17]. Spectrum Sensing (SS) is one of the most crucial techniques of CR which provides real-time occupancy information of frequency bands that are available for SUs without interfering with PUs' operations [18]. An accurate detection of the occupancy of spectrum bands is critical in CR operation, since all secondary transmission strategies are based on it [19].

The PU spectrum has been traditionally sensed with popular SS techniques like Energy Detection (ED), Matched Filter Detection (MFD) and Cyclostationary-based Detection (CBD). These conventional techniques result in inefficient use of radio resources due to missed detection of the PU and false alarms [20]. Due to the drawbacks of conventional SS techniques and increasing popularity of Artificial Intelligence (AI), many recent works have used traditional Machine Learning (ML) algorithms like Support Vector Machines (SVMs) and K-Nearest Neighbor (KNN) for SS. However, the manual feature extraction process involved in traditional ML algorithms requires expert knowledge and is a time-consuming process. Deep Learning (DL) is a data-driven approach and also a subset of ML that can automatically capture complex patterns and features from input data [21]. Deep Neural Network (DNN) techniques being quickly adaptable are robust to uncertain radio environment. The main objective of this article is to survey the latest research efforts towards the application of DL for the important task of sensing the PU spectrum.

A. INTRODUCTION TO SPECTRUM SENSING

The radio spectrum is inefficiently exploited because of the fixed allocation policies and the growth in user needs [22].

TABLE 1. List of acronyms in alphabetical order.

Acronym	Definition
ACARS	Aircraft Communications Addressing and Reporting System
AE	Autoencoder
AGM	Arithmetic to Geometric Mean
AI	Artificial Intelligence
ANN	Artificial Neural Network
APASS	Activity Pattern Aware Spectrum Sensing
ASK	Amplitude-Shift Keying
AUC	Area Under Curve
AWGN	Additive White Gaussian Noise
BPSK	Binary Phase-Shift Keying
CAV	Covariance Absolute Value
CBD	Cyclostationary-based Detection
CED	Classical Energy Detection
CFD	Cyclostationary Feature Detection
CLDNN	Convolutional Long short-term Deep Neural Network
CM	Covariance Matrix
CM-CNN	Covariance Matrix aware Convolutional Neural Network
CNN	Convolutional Neural Network
CP	Cyclic Prefix
CPFSK	Continuous Phase Frequency-Shift Keying
CR	Cognitive Radio
CRN	Cognitive Radio Network
CSCG	Circularly Symmetric Complex Gaussian
CSS	Cooperative Spectrum Sensing
DCS	Deep Cooperative Sensing
DL	Deep Learning
DNN	Deep Neural Network
DNN-LRT	Deep Neural Network-based Likelihood Ratio Test
DSM	Dynamic Spectrum Management
EBD	Eigenvalue-based Detection
EC	Estimator Correlator
ED	Energy Detection
FAR	False Alarm Rate
fc	fully connected
FLDA	Fisher Linear Discriminant Analysis
FROC	Free-response Receiver Operating Characteristic
FSK	Frequency-Shift Keying
GFSK	Gaussian Frequency-Shift Keying
GMM	Gaussian Mixture Model
HD	Hard Decision
HMM	Hidden Markov Model
IED	Improved Energy Detection
i.i.d.	independent and identically distributed
IoT	Internet of Things
I/Q	In-phase/Quadrature
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
LSTM-SS	Long Short-Term Memory-based Spectrum Sensing
MED	Maximum Eigenvalue Detector
MFD	Matched Filter Detection

In CR-based DSM, SS is a crucial step to learn the radio environment [12] and help increase spectrum utilization. SS is used to continuously sense the licensed user spectrum

TABLE 1. (Continued.) List of acronyms in alphabetical order.

ML	Machine Learning
MLD	Maximum Likelihood Detector
MLP	Multilayer Perceptron
NBC	Naïve Bayes Classifier
N-CNN	Convolutional Neural Network with unstandardized input
OFDM	Orthogonal Frequency-Division Multiplexing
OOBE	Out-Of-Band Emissions
PAM	Pulse Amplitude Modulation
PAS-SS	Primary user Activity Statistics-based Spectrum Sensing
PFA	Probability of False Alarm
PM	Probability of Miss detection
PoD	Probability of Detection
PoE	Probability of Error
PSK	Phase-Shift Keying
PU	Primary User
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying
RF	Radio Frequency
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SA	Self-Attention
SAE	Stacked Autoencoder
SAE-SS	Stacked Autoencoder-based Spectrum Sensing
SAE-TF	Stacked Autoencoder-based spectrum sensing using Time-Frequency domain signals
S-CNN	Convolutional Neural Network with Standardized input
SD	Soft Decision
SI-ED	Sweep-Integrated Energy Detection
SNR	Signal-to-Noise Ratio
SS	Spectrum Sensing
SSE	Signal Subspace Eigenvalues
SU	Secondary User
SVM	Support Vector Machine
TL	Transfer Learning
UAV	Unmanned Aerial Vehicle
UDSS	Unsupervised Deep Spectrum Sensing
USRP	Universal Software Radio Peripheral
VAE	Variational Autoencoder
WD	Waveform Detection

by the SUs to detect PU activity and spectrum holes in terms of duration, frequency, and location [23]. SS is modelled as a binary hypothesis testing problem with the two hypotheses being the presence or absence of the PU [24]. The functions of spectrum sharing, decision making, and resource allocation are implemented if the spectrum is available for transmission [17]. Fig. 1 depicts a block diagram representing sensing of the PU spectrum by an SU. The received radio signals at the SU are first passed through any signal processing, feature extraction and data pre-processing steps before being provided as input to an SS detector. The SS

detector receives and utilizes a priori information if required and generates a decision about PU presence or absence.

SS techniques can be grouped based on the number of users as multi-user or Cooperative Spectrum Sensing (CSS) and single user or local or non-CSS [25]. In CSS, a group of SUs share their sensed data (soft fusion) or decisions (hard fusion) with fusion center to improve the sensing accuracy. Local sensing results are different for SUs due to the differences in sensing capabilities [26]. The SS performance of a CR system is determined by the global decision combination rule of fusion center along with other factors like the number of SUs, the environment for SS, and the capabilities of SUs [27]. The combination of soft decisions results in an optimal detection performance but needs an infinite amount of bandwidth in theory [28]. Hard fusion on the other hand produces inferior results while saving bandwidth [29]. Performance and bandwidth efficiency can be balanced by using a combination of hard and soft decisions from the SUs [27].

Non-CSS techniques suffer from the hidden PU problem [30]. By using CSS, the performance gain of a CR system can be increased by the cooperation of multiple SUs to detect spectrum holes [27]. CSS can overcome problems like multipath fading and shadowing, ensuring that PU constraints are met for SS [31], [32].

B. RELATED SURVEYS

There are several survey papers that have reviewed the various types of SS algorithms and have discussed the research challenges and future directions associated with it. For instance, [33] emphasizes the need of exploring the code and angle dimensions along with frequency, time, and space dimensions for obtaining complete spectrum awareness. SS approaches of ED, waveform-based sensing, CBD, radio identification-based sensing and MFD are presented. The work further discusses the challenges involved in sensing the PU spectrum and the concept of CSS and its types. The survey [34] reviews SS techniques by categorizing them into three classes based on whether they need both source signal and noise power information, only noise power information (semi-blind detection) or no prior information (totally blind detection), with a particular focus on semi-blind and blind techniques. An analysis on detection threshold and test statistics distribution is provided and the challenges in developing a practical SS device are discussed. Axell et al. [35] provide the fundamentals of signal detection and conventional narrowband and wideband SS detectors. The work describes CSS in detail and discusses energy efficiency in CSS.

Sharma et al. [36] discuss the enablers of CR along with the practical imperfections in a CR system. In addition, the work provides a classification of popular SS techniques based on the signal processing techniques used, signal bandwidth, coordination between SUs and number of RF chains. Another survey [37] provides a detailed review of

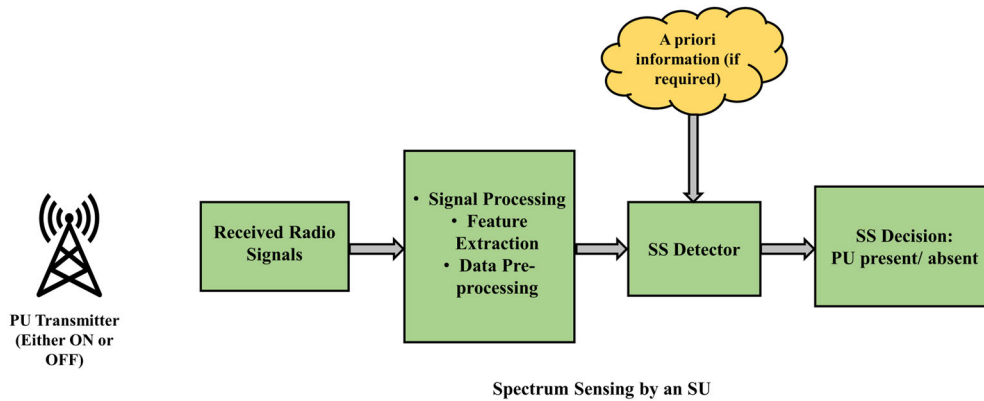


FIGURE 1. Block diagram representing sensing of the PU spectrum by an SU.

traditional SS techniques by categorizing them on the basis of bandwidth as narrowband and wideband SS approaches. Besides describing narrowband sensing techniques like ED, MFD and Eigenvalue-based Detection (EBD) and wideband sensing techniques like multiband sensing, wavelet-based sensing and compressive sensing, the paper discusses the practical implementation aspects for various sensing techniques. Work [19] first describes the different access modes for CR and then summarizes four common SS methods: ED, MFD, Covariance Absolute Value (CAV) detector and Hadamard ratio-based detector. The concept of Signal-to-Noise Ratio (SNR) wall is explained and the performance of the four SS detectors is analyzed and compared under various conditions.

The survey in [25] groups the SS techniques on the basis of bandwidth as narrowband and wideband, on the basis of number of users as non-cooperative and cooperative, on the basis of detection as transmitter and receiver detection, and finally on the basis of need of prior knowledge as blind, semi-blind and non-blind detection. In addition, a detailed comparison of popular SS techniques: ED, MFD, feature detection, Waveform Detection (WD) and EBD is provided. The work also discusses the system modelling methods for SS, the challenges associated with sensing the PU spectrum and CR standards. The work in [17] summarizes the fundamentals of CR and SS. SS techniques are described, and mathematical models are provided by classifying the schemes into conventional methods, such as ED and MFD and recent advanced sensing schemes such as wideband compressive and adaptive compressive techniques. The survey further discusses the challenges involved in sensing the PU spectrum and various applications of Cognitive Radio Networks (CRNs).

A review in [38] provides a classification of SS techniques based on bandwidth. The narrowband sensing techniques: ED, CBD, MFD, covariance-based detection and traditional ML-based SS are discussed. Works based on traditional ML techniques like K-means clustering, SVM and KNN are surveyed. Wideband sensing techniques are further grouped into Nyquist-based and compressive sensing techniques.

A study in [11] reviews probabilistic SS approaches by grouping them on the basis of what features they extract from the samples of received signals. ED, CBD, EBD, MFD and blind detection techniques and their sub-categories are described in detail along with a discussion on the implementation challenges. Based on the applications and technologies already envisioned for 6G, the role of SS is conceptualized for use in future networks.

The survey in [23] is based on the applications of ML for CSS and dynamic spectrum sharing by focusing on the feature vector extracted from the received signal, the type of ML algorithm and evaluation metrics. The ML-based sensing techniques are categorized as supervised, unsupervised and reinforcement learning techniques, and are analyzed based on the features, type of SUs, and performance metrics used. The work further classifies the spectrum sharing techniques and summarizes the use of ML algorithms for spectrum sharing. Although numerous survey papers exist in the field of SS, there is a lack of literature on the latest DL-based developments for sensing the spectrum. ML-based SS has been covered in some surveys, but the papers mainly focus on traditional ML algorithms and not on the emerging research efforts using DL. Table 2 summarizes the key focus of the surveys discussed in this sub-section.

C. CONTRIBUTION

In contrast to the existing research, this paper provides a detailed survey of the recent works that have used DNN algorithms for SS. The key contributions of this paper are as follows:

- After introducing the concept of SS, this work summarizes the contribution of several fundamental review papers in the field of SS. The survey then discusses about conventional SS algorithms and their drawbacks. This will provide a comprehensive guide for readers to develop an understanding of the SS research.
- We then provide an overview of some works that are based on traditional ML algorithms and simple Artificial Neural Networks (ANNs) before discussing DNNs for completeness. Summarizing the early works

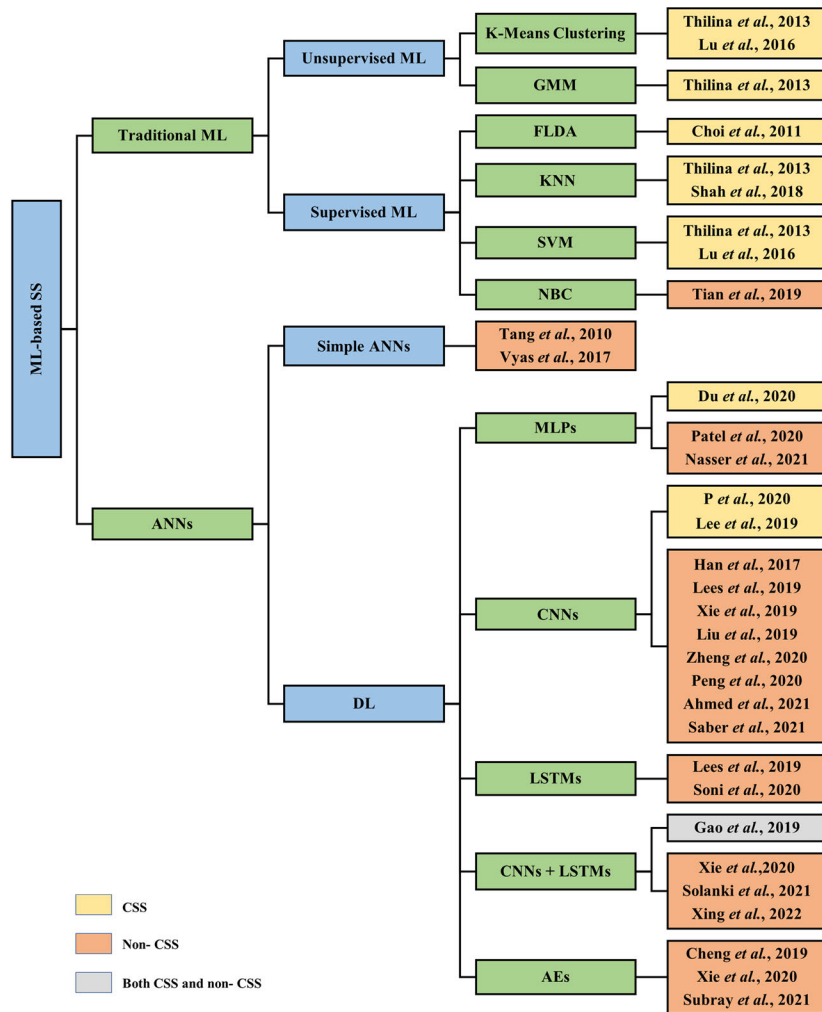


FIGURE 2. Summary of ML-based SS algorithms discussed in this paper.

that used ML for SS helps create a timeline and understanding of research efforts towards adopting AI for SS.

- The major portion of this paper is dedicated to surveying the latest works using DL algorithms for SS by categorizing them into five types of DL algorithms: Multilayer Perceptrons (MLPs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, combined CNN-LSTM architectures and Autoencoders (AEs). A detailed analysis of various DL approaches in terms of the input provided to the DL algorithm, method of data acquisition, data pre-processing technique used, architecture of each algorithm, evaluation metrics used, results obtained and comparison with standard SS detectors is presented. This discussion will create awareness about using DL techniques for SS. For instance, which pre-processing technique and DNN architecture should be selected for a high detection performance.

- We describe some publicly available Radio Frequency (RF) datasets and discuss the need for comprehensive RF datasets and the concept of Transfer Learning (TL).
- The work further highlights the research challenges related to the use of DL for SS along with potential solutions.

D. PAPER ORGANIZATION

Fig. 2 provides a summary of the ML-based SS algorithms discussed in this paper along with their years of publication. The figure also highlights if a particular work is based on CSS or non-CSS scenario. Fig. 3 summarizes the paper organization. We introduced the concept of DSM and the problem of spectrum scarcity in Section I. This section further covers the related surveys and highlights our contribution. Section II describes conventional and ML-based SS approaches. Section III reviews various DL-based SS techniques. Section IV provides information about various RF signal datasets, the concept of TL and software used. Section V includes research challenges of applying DL for

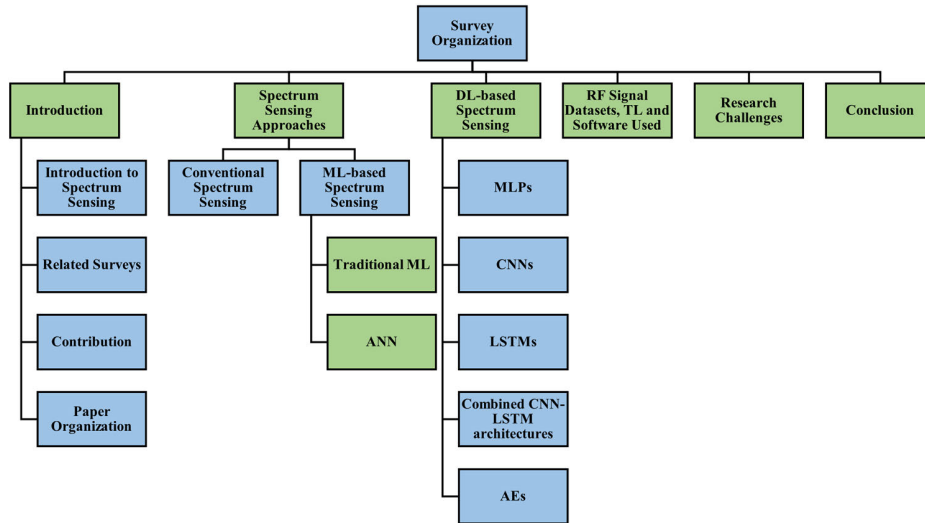


FIGURE 3. Survey paper organization.

TABLE 2. Related surveys.

Survey	Key Focus
[33], 2009	Conventional SS, complete spectrum awareness, CSS, and its types.
[34], 2010	Conventional SS (non-blind, semi-blind and blind), analysis on detection threshold and test statistics distribution, challenges in developing a practical SS device.
[35], 2012	Fundamentals of signal detection, conventional SS (narrowband and wideband), energy efficiency in CSS.
[36], 2015	Enablers and imperfections of CR systems, popular SS techniques and classification.
[37], 2017	Conventional SS (narrowband and wideband) techniques and their practical implementation aspects.
[19], 2017	Conventional SS, access modes for CR, SNR wall.
[25], 2018	Conventional SS (narrowband and wideband, non-cooperative and cooperative, transmitter and receiver detection, blind, semi-blind and non-blind detection), system modelling methods, challenges, CR standards.
[17], 2019	Fundamentals of CR and SS, conventional and advanced SS techniques and their mathematical models, applications of CR network.
[38], 2019	Narrowband SS techniques including traditional ML-based SS, wideband SS techniques.
[11], 2021	Probabilistic SS approaches, their sub-categories and implementation challenges, role of SS for future networks.
[23], 2021	Traditional ML-based SS (supervised, unsupervised and reinforcement learning).
This work	DNN-based SS, overview of traditional ML and simple ANN-based SS, need for RF datasets, concept of TL and research challenges.

SS along with potential solutions while section VI concludes this paper.

II. SPECTRUM SENSING APPROACHES

A. CONVENTIONAL SPECTRUM SENSING

The PU spectrum has traditionally been sensed using popular sensing techniques like ED, MFD, EBD, CBD, and WD. These conventional SS algorithms have been extensively reviewed in several research works, as summarized in section I. However, these techniques result in an inefficient utilization of radio resources due to the problem of missed detection of the licensed user and false alarms [20]. For instance, there is no need for prior information about the PU signal before using ED, but it has a high false alarm rate and is inefficient in environments with low SNRs [38]. The MFD technique becomes impractical to implement because it requires prior knowledge of the PU signals [25]. EBD technique has high computational complexity [37]. SS using CBD involves high power consumption, processing complexity and sensing time [17]. There is a possibility of synchronization errors with WD [25]. SS techniques must ensure a low Probability of False Alarm (PFA) and a high Probability of Detection (PoD) to lower the impact of harmful interference [38]. Due to the drawbacks of conventional SS techniques and growing popularity of the field of AI, a lot of recent works have adopted ML algorithms for sensing the spectrum accurately.

B. MACHINE LEARNING (ML)-BASED SPECTRUM SENSING

ML-based SS algorithms are based on extracting feature vectors from patterns and classifying them into either null hypothesis (absence of PU) or alternative hypothesis (presence of PU) [27]. These techniques are more adaptive than conventional sensing techniques due to their learning ability and when adopted for CSS can achieve a better detection performance due to their capacity to describe more optimized decision region on the feature space [39]. ML can help address the problem of spectrum scarcity by increasing

spectrum utilization [40]. ML techniques treat SS as a binary classification problem and use energy or probability vectors to predict the status of RF channel [38]. The ML algorithms can be broadly classified into three classes [41]:

- Supervised Learning: where model is trained using input samples and the corresponding labels.
- Unsupervised Learning: where input samples are distinguished by the model without any output labels.
- Reinforcement Learning: where an agent learns to map input to actions by communicating with an environment.

A review of the three types of ML algorithms used for CSS is provided in a recent survey [23]. Out of the three categories, supervised and unsupervised ML techniques are popular among researchers in the context of SS. In this work ML techniques for SS are analyzed by dividing them into two classes: traditional ML and ANNs. The literature on ANN-based sensing is scarce, and this survey aims to provide a detailed review of various ANN algorithms with a particular focus on DL techniques by describing their architectures, input provided to algorithms, outputs produced, evaluation metrics and comparison with standard models.

1) TRADITIONAL MACHINE LEARNING-BASED SPECTRUM SENSING

In this work, traditional ML refers to ML algorithms that are not based on neural networks. The use of traditional ML for sensing of radio spectrum has been widely adopted in several works. For example, to utilize ML algorithms for CSS, [42] explored the use of Fisher Linear Discriminant Analysis (FLDA) for fusing the sensing results from SUs. FLDA can be considered as a supervised ML technique and is used to separate two or more classes by determining a linear combination of features [43]. The PU network is modelled as a random geometric network and the SUs sense the spectrum using ED to determine spectrum availability. The sensing performance is made accurate by incorporating location information and reliability of the decision of each SU with the help of a linear fusion rule whose coefficients are determined by FLDA. The Receiver Operating Characteristic (ROC) plots of the proposed scheme are compared with equal coefficient model, AND rule, OR rule and Maximum Likelihood Detector (MLD)-based rule by considering two circular detection areas.

In equal coefficient model, the sensing results are combined in a linear manner like the proposed model but with same linear coefficients for all SUs without considering the SU network topology. The OR and AND rules are hard fusion rules. The OR rule determines the presence of PU if at least one SU reports the presence of PU whereas in AND rule the presence of PU is confirmed when all SUs detect the PU. The MLD following Neyman-Pearson criterion is the optimal detector for the problem of random PU network detection. The detection performance for all models improves with the increase in radius of detection and the proposed model outperforms the equal coefficient, AND and OR rules-based models for both circular detection areas.

Work [39] proposes unsupervised ML techniques: K-means clustering and Gaussian Mixture Model (GMM) along with supervised ML techniques: SVM and weighted KNN. The energy levels received at SUs are considered as feature vectors and fed into ML models to predict the channel availability. K-means clustering works by dividing the features into 'K' clusters and mapping the clusters to the status of PU based on the centroids of clusters. GMM is a probabilistic approach which models feature vectors as a gaussian mixture distribution so that each gaussian distribution corresponds to a cluster. For using SVM, the training energy vectors are made linearly separable by mapping them to a higher dimensional feature space by means of a non-linear mapping function. The SVM algorithm then finds a hyperplane which is at the maximum distance from data points of the two classes corresponding to availability of PU. In weighted KNN technique the nearest 'K' points are assigned weights inversely proportional to their distances and classes are predicted based on the majority voting of neighbors.

The system model consists of multiple PUs and the channel is considered available only if all PUs are inactive. Multiple SUs estimate the energy levels of received baseband complex signal samples and report it to the fusion center which generates the energy vector. The ML models are evaluated by comparing their training times, classification delays and ROC curves by modelling a system with 25 SUs and 2 PUs. KNN classifier takes the least training time while SVM needs the highest training duration. On comparing the average classification delay for various classifiers, it is observed that the classification time remains constant even on increasing the number of training samples for Fisher linear discriminant [42], K-means clustering, and GMM techniques. On comparing the ROC curves of various CSS schemes by varying the number of SUs it is found that a greater number of SUs results in better performance for the proposed classifiers. SVM with linear kernel outperforms SVM with polynomial kernel, K-means clustering, KNN with Euclidean and Cityblock distances, GMM as well as the other CSS schemes like Fisher linear discriminant, OR rule, and AND rule.

Following a similar approach, Lu et al. [44] proposed the use of two-dimensional probability vectors in place of high-dimensional energy vectors for K-means clustering and SVM-based CSS. The high dimensional feature vector is transformed into a low dimensional probability vector to achieve smaller training time and classification delay. The performance of K-means clustering, SVM with linear kernel and SVM with polynomial kernel techniques with energy vector and probability vector is compared to determine their probabilities of detection, training durations and classification delays. When considering a CRN of one PU and 2 SUs, the ML techniques report higher detection probabilities than OR and AND fusion rules while SVM with probability vector outperforms all other sensing techniques. In a CRN with one PU and 9 SUs, SVM with linear kernel and probability

vector has the highest detection accuracy. While the training duration of K-means clustering is the longest, it has the least classification delay. Due to the low dimension of the probability vector-based ML algorithms, they report lower training duration and classification delay than energy vector-based ML algorithms.

In [27], the ED of received signals is viewed as an analogy to feature vectors for training a KNN classifier [45] for CSS. The KNN algorithm considers the sensing classes from training phase and current sensing report as neighbors and the Smith-Waterman algorithm [46] calculates the distance between the neighbors. Posterior probability is used to determine the nearest neighbor and KNN determines the prior and conditional probabilities. Multiple SUs are used to provide spatial diversity and the sensing slot is divided into mini slots to add temporal diversity. At each mini slot, the energy signal is quantized into discrete zones and multiple bits corresponding to the zones are transmitted instead of transmitting soft or hard decisions. The local sensing reports from SUs are combined by the fusion center with the help of a weight-based decision combination rule in which SUs are assigned weights based on their effectiveness. The performance of the scheme with training sizes of 330 and 100 is evaluated in Additive White Gaussian Noise (AWGN) as well as fading channels and compared with the conventional OR rule.

In AWGN channel, the proposed technique with 330 training size reports the highest PoD while the scheme with 100 training size has comparable performance with the conventional OR method for low SNR values. When comparing the Probability of Error (PoE), the proposed model with higher training size reports low PoE for low SNR values. For high SNR ranges, both proposed schemes have similar performance, confirming that the model gives more reliable performance than conventional SS even with a smaller training phase. It was concluded that for low SNR values, a higher number of training samples is needed to accurately predict the status of the PU. When considering a fading environment, the proposed method with larger training size reports a better detection performance and a lower PoE than the OR scheme.

In a recent work, Tian et al. [47] formulated the SS problem as an SNR-based multi-class classification problem to adapt to SNR variations. A Naïve Bayes Classifier (NBC) is trained for the SS of Orthogonal Frequency Division Multiplexing (OFDM) signals. NBC is a supervised ML algorithm which performs classification based on Bayesian decision theorem [48]. The work also proposed a class-reduction assisted prediction method to reduce the time needed for SS. On comparing the PoD versus SNR curves and the ROC curves, the NBC-based detection outperforms ED, Cyclic Prefix (CP)-based detector [49], asymptotic simple hypothesis test-based detector [50] and a neural network with two hidden layers. Furthermore, the performance bounds of the SS error rates are calculated, and the performance of NBC-based detector is found within the bounds.

Table 3 summarizes the traditional ML-based research works discussed in this subsection by providing a description of the input to the traditional ML algorithm, any pre-processing technique used, the ML algorithm used, key evaluation metrics used and performance of the SS model. Several studies have used both supervised and unsupervised traditional ML techniques to sense the PU spectrum. Unsupervised learning is practically easier to implement than supervised learning as it does not require information about the PU availability, but supervised learning techniques demonstrate better sensing performance as they have additional information about the status of PU [39]. Despite their popularity, the manual feature extraction in traditional ML techniques requires the knowledge of field experts and is a time-consuming process. Additionally, the accuracy of the selection of input features influences the detection results [23]. Moreover, most of the works based on traditional ML are limited to CSS scenario and rely on features extracted from conventional SS techniques to provide input to the sensing algorithms. These disadvantages of traditional ML algorithms have accelerated the research towards adopting neural networks for SS.

2) ARTIFICIAL NEURAL NETWORK (ANN)-BASED SPECTRUM SENSING

ANNs are inspired by the structure and functioning of a human brain and are widely used to model complex real-world problems in many disciplines [51]. These networks primarily consist of an input layer, hidden layers, and an output layer. Neural networks are efficient in learning non-linear functions and adapting the non-linear features of PU signals [52]. In this paper the ANN-based SS is grouped into simple ANN techniques which include neural networks with a single hidden layer and DL which is covered in detail in the next section.

In an early attempt to utilize simple ANN for SS, [53] proposed a joint detection method which combined ED and Cyclostationary Feature Detection (CFD) with ANN. The PU signal is assumed to be an amplitude modulated signal in AWGN. Four features are extracted to provide input to the ANN, one of which is energy and the other three are cyclostationary feature values. These features form the input of a neural network having one hidden layer and determines the status of the PU.

Vyas et al. [52] proposed a hybrid SS scheme by utilizing energy from traditional ED and Zhang test statistics from likelihood ratio test statistic [54] as features for training a simple ANN with a single hidden layer. The model is evaluated with the help of real-world PU signals obtained with an experimental test setup inspired from [55]. The hardware part of the platform consists of a Universal Software Radio Peripheral (USRP-N210), WBX daughter board, RF-Explorer and D3000N Super Discone antenna. GNU Radio and MATLAB form the software part of the measurement setup.

The combined features and labels from four radio technologies are utilized for training four different ANN

TABLE 3. Summary of traditional ML-based approaches for spectrum sensing.

Research Work	Description of Input to Traditional ML algorithm	Key Data Pre-processing Used	Traditional ML Algorithm Used	Key Evaluation Metrics	Key Performance Notes
[42], 2011	Sensing results from SUs, topology information of SUs	-	FLDA	ROC curves	PoD = 0.9, PFA = 0.37 for radius of detection area = 150 meters and ideal case with no hidden PU problem.
[39], 2013	Vector of energy levels received at SUs	-	Unsupervised ML techniques: K-means clustering, GMM Supervised ML techniques: SVM, weighted KNN	Training times, classification delays, ROC curves	PoD > 0.7 for PFA > 0.1 for SVM with linear kernel when considering 2 PUs and 25 SUs.
[44], 2016	Two-dimensional probability vector	Down-conversion of high dimensional feature vector to two-dimensional probability vector, scaling of probability vector	Unsupervised ML technique: K-means clustering Supervised ML technique: SVM	PoD, training duration, classification delay	PoD > 0.9 for PFA = 0.6 for SVM with linear kernel and probability vector and considering 1 PU and 9 SUs.
[27], 2018	Sensing report	-	KNN	PoD, PoE	PoD about 20% higher than OR rule at SNR = -16 dB for fading channel and training size of 330.
[47], 2019	Average received signal power and CP-induced correlation features of OFDM signals	-	NBC	PoD versus SNR and ROC curves	PoD > 0.8 for SNR = -20 dB for NBC methods.

architectures and the best ANN model is identified with the help of a cross-validation set. The model is evaluated for the four radio technologies by considering different sets of features: only energy of current sample, only Zhang statistics of current sample, energy and Zhang statistics of current sample, and energy and Zhang statistics of current and previous samples. For all radio technologies, using only the current sensing event gives the worst sensing performance while using energy and Zhang statistics of current and previous sample achieves the best accuracy. The proposed technique achieves a better detection performance than the Improved Energy Detection (IED) [56] and Classical Energy Detection (CED) [55] approaches for all radio technologies.

Table 4 summarizes the simple ANN-based research works discussed in this subsection. Even though simple ANN methods improve sensing performance, their detection results depend directly on the accuracy of the input features obtained from the received signals [57]. The extraction of specific features from the original received signal is limited to obtaining partial information, which inevitably leaves out implicitly hidden but helpful features. As simple ANNs do not have multiple layers, they cannot perfectly capture complex data features and non-linearity. To overcome these shortcomings of simple ANN algorithms, DL algorithms have been widely adopted for the task of SS.

III. DEEP LEARNING (DL)-BASED SPECTRUM SENSING

DL is a subset of ML that can automatically capture complex patterns and features from input data [21]. As these algorithms are quickly adaptable, they are robust to uncertain

radio environments. DL enhances model performance by utilizing the non-linear relationship in training data optimally [58]. The multiple hidden layers enable the DNN to learn patterns from datasets layer by layer. Low-level data features are transformed into high-level abstract features as the output from a lower layer serves as input for a higher layer [59].

Fig. 4 depicts a generalized representation of DNN-based SS in a non-CSS scenario. With DL, the SS problem is treated as a binary classification or hypothesis testing problem with the two classes representing absence of PU or null hypothesis (H_0) and presence of PU or alternative hypothesis (H_1). Firstly, either real-world data is acquired, or samples are synthetically generated to represent spectrum data. This data can be collected in the format of In-phase/Quadrature (I/Q) samples, spectrograms, Covariance Matrices (CMs), etc., or various features like energy and cyclostationary features can be derived from this data. Secondly, the acquired spectrum data can be pre-processed by techniques like data standardization, data normalization, filtering, matrix manipulation, etc., to be in an appropriate form and increase the detection performance of the DNN. The pre-processed data is split into training, validation, and test sets. The model is trained by the training set and then tuned by the validation data in an offline process. Next, the well-trained model with its hyperparameters optimized is used to classify the test data into either H_0 or H_1 in an online detection process.

Fig. 5 represents the generalized use of a DNN for decision fusion in a CSS scenario involving multiple SUs. For CSS, the DNN is used to combine the sensing results obtained from

TABLE 4. Summary of simple ANN-based approaches for spectrum sensing.

Research Work	Description of Input to Simple ANN algorithm	Key Data Pre-processing Used	Key Features of Simple ANN Architecture	Key Evaluation Metrics	Key Performance Notes
[53], 2010	One energy and three cyclostationary feature values of amplitude modulated signal in AWGN	-	Single hidden layer	PoD, PFA	PoD can reach 1 when testing SNR \geq -10 dB and training SNR is between -30 dB and 10 dB.
[52], 2017	Four features obtained from the energy and Zhang statistics of current and previous samples from four radio technologies	Filtering of signals and removal of transient peaks	Single hidden layer	PoD, PFA	PoD close to 0.4 at SNR = -20 dB for FM Broadcasting signals with PFA = 0.022 and sample size of 500.

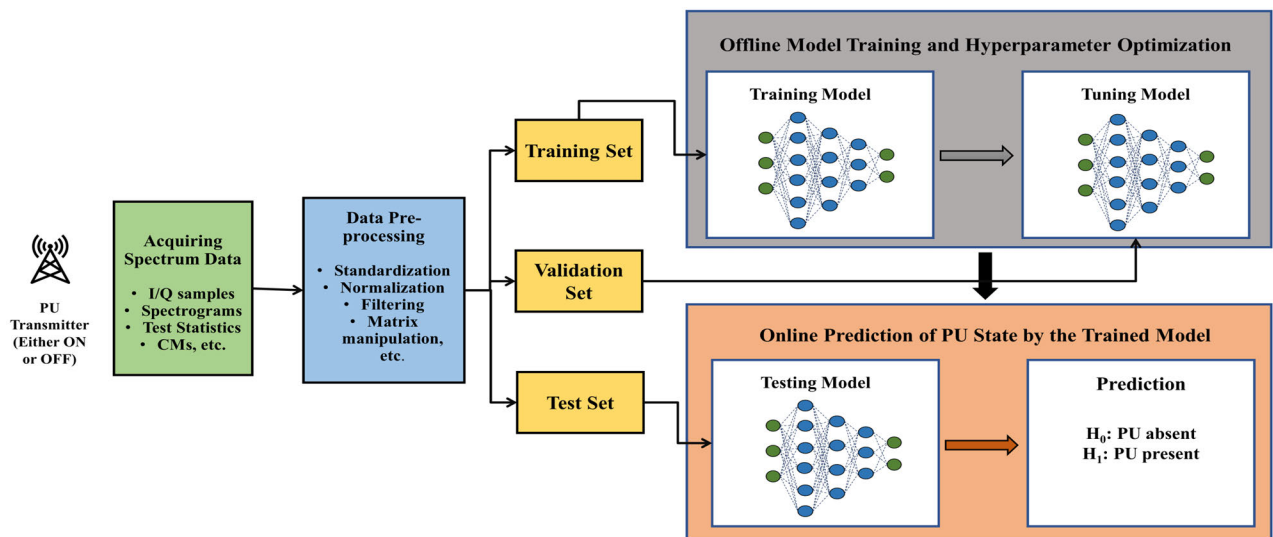


FIGURE 4. Generalized representation of DNN-based spectrum sensing in a non-CSS scenario.

individual SUs. The dataset formed of the sensing results can be pre-processed and is then divided into training, validation, and test sets. The DL model acts as a fusion center and generates prediction about the status of the PU by combining sensing information from multiple SUs.

DL algorithms are data-driven and can efficiently capture the non-linearities of input data. With data-driven methods, the model has the advantage of being trained by extracting inherent patterns from data and not relying on signal and noise assumptions, thus ensuring reliable performance when used in practical with real signals [60]. In this paper, the DNN architectures used for SS have been classified as MLPs, CNNs, LSTM networks, combined CNN-LSTM architectures, and AEs, as shown in Fig. 6. As a result of their simple architecture, MLPs are easy to implement and are preferred for the task of CSS for fusion of the sensing data from multiple SUs. CNNs are also preferred for fusion of sensing results in CSS as the sensing outcomes are correlated. To efficiently capture the spatial and temporal details of sensing data, CNNs and LSTMs are popularly used with diverse and complex radio data in non-CSS scenarios. AEs being an unsupervised ML approach are suitable when only a limited amount of labeled data is available.

A. MULTILAYER PERCEPTRONS (MLPs)

An MLP is a feedforward ANN having one or more hidden layers. When the number of hidden layers is higher than one, an MLP is considered as a DNN. The number of neurons in the input layer is determined by the dimensions of dataset. The neurons in output layer are equal to the number of output labels or classes. The number of hidden layers and the count of neurons in each hidden layer are determined with the aim of optimizing the MLP accuracy. The architecture of a simple MLP having multiple hidden layers is depicted in Fig. 7.

Du et al. [61] proposed an MLP with 3 hidden layers for the centralized CSS of the PU spectrum by combining information geometry with DL. Input to this method called ‘IG-DNN’ is a dataset consisting of geodesic distances derived from covariance matrices of sensing signals and noise. The energy values of noise and signal mixed with noise in a range of SNR values are sensed by multiple SUs to form the input for the MLP. The experiments were performed using multiple signals and by fixing the PFA it was concluded that a greater number of SUs and a higher SNR result in better sensing performance. IG-DNN outperformed IG-FCM [62] and MME-K-means [63] algorithms when their performances were compared under various simulation settings.

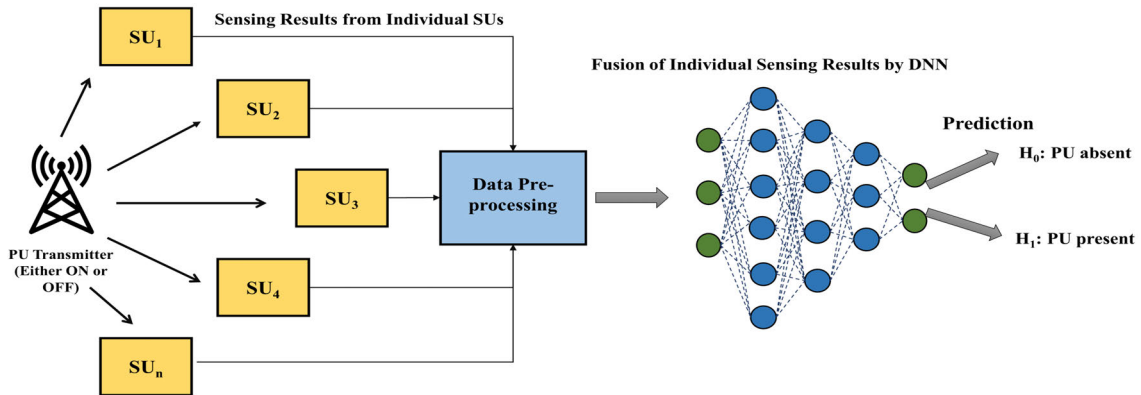


FIGURE 5. Representation of decision fusion using a DNN in CSS scenario.

TABLE 5. Summary of MLP-based approaches for spectrum sensing.

Research Work	Description of Input to MLP algorithm	Key Data Pre-processing Used	Key Features of MLP Architecture	Key Evaluation Metrics	Key Performance Notes
[61], 2020	Geodesic distance between signals simulated using MATLAB	-	MLP with 3 hidden layers	ROC curves	PoD > 0.6 for PFA = 0.2, SNR = -20 dB ~ SNR = -14 dB and 2 SUs.
[64], 2020	Four radio technologies captured using USRP added with AWGN and AWGN	Filtering of captured signals and removal of transient peaks	MLP with 2 hidden layers	PoD versus SNR curves	PoD > 0.8 for SNR = -20 dB, number of training samples per class = 100.
[65], 2021	Test statistics from six detectors and SNRs	-	MLP with 2 hidden layers	PoD, FAR	Average PoD of 0.93, almost zero FAR with more than three detectors.

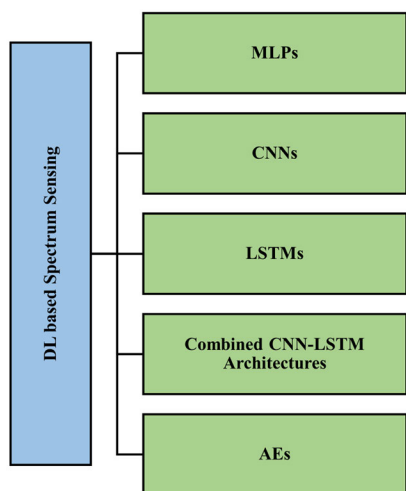


FIGURE 6. Deep Learning-based architectures used for SS.

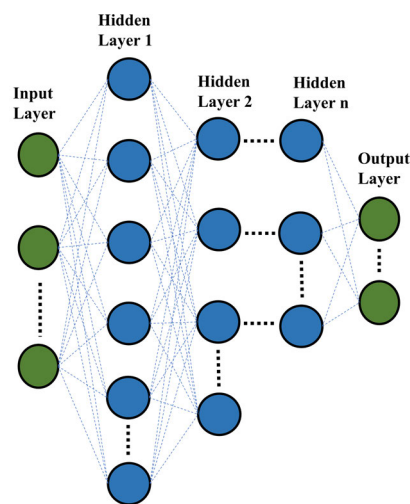


FIGURE 7. An MLP having multiple hidden layers.

In another work [64], an MLP with two hidden layers is designed after optimizing the number of hidden layers, the neurons in each hidden layer, optimization algorithm, activation function and learning rate for SS. The PU data is formed by four different radio technologies captured using an empirical setup similar to that in work [52]. The signals were filtered, the transient peaks were removed and AWGN was added to obtain desired SNR levels. The

input to the MLP are four features: the energy values of the current and previous sensing event and the Zhang statistics of current and previous sensing event while the output is the status of PU channel. A total of four ANN architectures are utilized, each assigned to a particular radio technology. The proposed approach is compared with a neural network without hyperparameter tuning, CED, IED and with

NBC-based sensing [47]. On comparing the PoD versus SNR curves, it is found that the proposed scheme has similar performance to NBC at low SNRs. CED and IED techniques are computationally simple, but the MLP reports a better PoD. On averaging the detection performance of the model over four radio technologies, it reports a 63% performance improvement over CED and IED.

Nasser et al. [65] proposed a hybrid SS scheme wherein the test statistics of six detectors were combined to train an MLP having 2 hidden layers. By considering 16- Quadrature Amplitude Modulation (16QAM) signals and AWGN, the data pertaining to the test statistics of ED, autocorrelation detector [66], maximum eigenvalue detector [67], cumulative power spectral density detector [68], maximum–minimum eigenvalue detector [67], and goodness-of-fit detector [69] and the SNR values were used as features to train the MLP. The performance of the DNN is evaluated with metrics PoD and False Alarm Rate (FAR). It is observed that increasing the number of detectors used for training the MLP increases PoD, while decreasing the FAR. More than three detectors result in an average PoD of 0.93 and an almost zero FAR.

Table 5 summarizes the research works discussed in this subsection by providing details about the dataset used, key pre-processing technique used, key features of the MLP architecture, key evaluation metrics used and performance notes on the SS model.

B. CONVOLUTIONAL NEURAL NETWORKS (CNNs)

CNNs are DNNs popularly deployed in the areas of computer vision and natural language processing. Besides input and output layers, a simple CNN network consists of convolution, pooling and fully connected (fc) layers [70]. Features are extracted automatically from input samples with the help of kernels or filters in convolutional layers and are called feature maps. Pooling layers perform down-sampling to decrease the complexity for further layers in the network and to avoid overfitting [71]. Fully connected layers are further used to classify the input data with the help of features extracted from previous layers. Fig. 8 shows the architecture of a basic CNN model.

There are some standard CNN architectures most of which have been trained on a large image database called ImageNet [72]. By utilizing these models, a different task can be recognized without the need to train from scratch, making these models particularly useful in cases where limited training data is available [73]. This is the process of TL, and the pre-trained weights of these models can be accessed with the help of various DL libraries like Keras and PyTorch. For instance, the GoogLeNet/Inception-V1 architecture proposed in [74] consists of 22 layers having a total of 9 inception blocks. The work defines inception modules as blocks that facilitate better computation and deeper networks by reducing the dimensionality through stacked 1×1 convolutions. By using these modules, the dimensions are reduced thus lowering the computational costs and addressing the issue

of overfitting. MobileNetV2 [75] was proposed for mobile devices and has 19 residual bottleneck layers following the initial fully convolutional layer of 32 filters. Its architecture is inverted residual with bottleneck layers having residual connections between them.

VGG-16 and VGG-19 [76] are classical CNN architectures designed by exploring the impact of architecture depth on model performance. VGG-16 has 16 layers out of which 13 layers are convolutional layers and 3 layers are fc layers. VGG-19 has 19 layers among which 16 are convolutional layers and 3 layers are dense layers. ResNet-18 and ResNet-50 [77] were proposed to facilitate the training of deep networks with the help of a deep residual learning framework. In deeper networks, as the network depth increases, accuracy becomes saturated and then rapidly degrades due to a higher training error. In ResNet-18 there are 17 convolutional layers while ResNet-50 has 49 convolutional layers [78]. As the network becomes deeper, it suffers from the problem of vanishing gradients. To address this, in DenseNet-121 [79] all layers are interconnected and the feature maps of all preceding layers form inputs for each layer. DenseNet-121 has a total of 120 convolutional layers.

The similarity between images and signal covariance matrices makes CNNs widely suitable for SS problems [21]. Another reason for their popularity is that the operation of a CNN filter or kernel is similar to the filtering operations at communications receivers [80]. In an early attempt to use CNN architecture for SS, [81] proposed a CNN-based SS algorithm having a single convolutional layer to sense the presence of PU signal in environments having low SNRs. The energy and cyclostationary features of signals are extracted, standardized, and provided as input to the CNN model. Based on [82], it is considered that the PU signals are cyclostationary signals while noise is stationary and hence, they can be distinguished on the basis of cyclostationary features. The proposed model is evaluated by using Binary Phase-Shift Keying (BPSK) signals in an AWGN channel. The CNN with standardized input (S-CNN) is compared with CNN model trained without standardizing the input (N-CNN) and CFD models by varying the SNR between -20 dB and -5 dB. While S-CNN has a better detection performance than N-CNN, both S-CNN and N-CNN outperform the conventional CFD algorithm at all SNRs.

In [60], a CNN-based algorithm called Activity Pattern Aware SS (APASS) is proposed which learns the PU activity pattern to perform SS. The input to the algorithm includes CM of current frame and a matrix formed by stacking CMs from past frames to enable the CNN model to exploit the PU activity pattern and improve detection accuracy. The model architecture is inspired from the standard CNN architecture called LeNet [83] and consists of a total of seven layers, including 2 convolutional layers and 2 dense layers. Both correlated and uncorrelated signal models are adopted, and PU signal vector is considered Gaussian with zero mean. The work analyzes the convergence behavior of loss function at different SNR levels, and it is observed that at high SNR

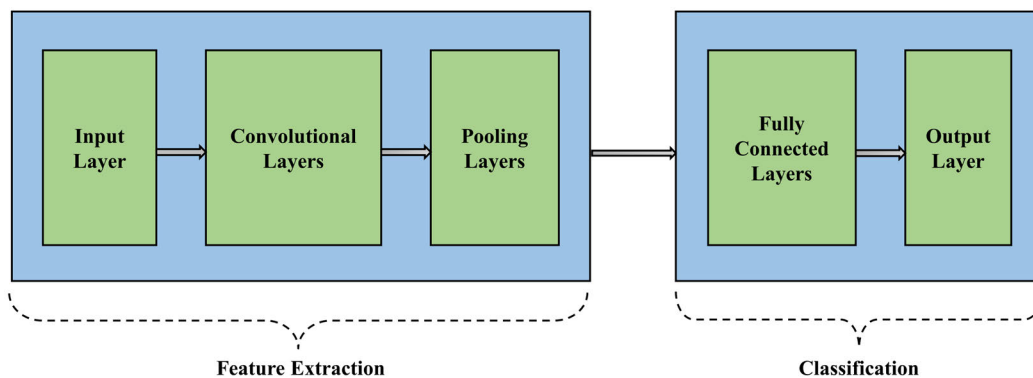


FIGURE 8. Architecture of a basic CNN model.

levels the loss function quickly converges to zero due to a large difference in CMs of the two hypotheses. APASS detector is compared with optimal Estimator-Correlator (EC) detector and Hidden Markov Model (HMM) with the help of PoD versus SNR curves. When considering the uncorrelated signal model, the APASS detector with both deterministic and random PU activity model outperforms the EC and HMM models. The detection performance of all algorithms improves with the increase in the number of receiving antennae. On adopting a correlated signal model, similar trends are observed with an improvement in performance of all three detectors.

Liu et al. [84] designed DNN-based Likelihood Ratio Test (DNN-LRT) and proposed the CM-aware CNN (CM-CNN) model for SS. Sample CMs form the input of the DNN which is based on LeNet-5 [83] architecture and consists of 2 convolutional layers. The PU signals are generated by simulation using independent and identically distributed (i.i.d.) and exponential correlation models. Noise samples either follow a Gaussian distribution or are real sea clutter samples [85]. The performance of CM-CNN is compared with EC, ED [86], Maximum Eigenvalue Detector (MED) [87], blindly combined energy detector [88] and CAV-based detector [89] with the help of ROC and PoD versus SNR curves. Under Gaussian noise, CM-CNN obtains a comparable performance with the optimal EC detector for both i.i.d. and exponential correlation models. When considering sea clutter and exponential correlation model, CM-CNN has a satisfactory detection performance while EC detector cannot be implemented as sea clutter lacks statistical model. CM-CNN is robust in conditions of low SNR and outperforms EC detector when noise uncertainty is 1 dB.

Lees et al. [78] compared the narrowband and wideband detection performance of 13 different models for detection of SPN-43 air traffic control radar. With the help of 3.5 GHz band low-resolution spectrograms [90], [91], the performance of conventional algorithms: ED and Sweep-Integrated Energy Detection (SI-ED) [92], traditional ML algorithms: SVM, KNN and GMM, standard CNN architectures: ResNet-18, ResNet-50, Inception-V1, VGG-16, VGG-19 and

DenseNet-121, along with CNN and LSTM models proposed in this work is evaluated. Two test sets, i.e., Test A and Test B, were defined. Test A had an equal proportion of spectrograms in categories: SPN-43, Radar 3 Out-Of-Band Emissions (OOBE), both SPN-43 and Radar 3 OOBE and neither. Test B was a subset of Test A without Radar 3 OOBE.

For the proposed LSTM architecture, the 10 MHz channel is divided into sequential slices along the time axis. The outputs of LSTM cells are passed to dropout cells with a 50% probability. An fc layer with 50 neurons receives the output of the last cell after all the time slices have been provided to the LSTM. Next a layer with a single neuron and sigmoid activation generates a prediction between 0 and 1. The proposed CNN architecture is called CNN-3 and consists of a convolutional layer, a dense layer with 150 neurons and a layer with a single neuron which generates output between 0 and 1. After the convolutional layer, the work uses a novel averaging step for activation maps which is equivalent to single filter 1×1 convolutional layer.

For narrowband detection, all 13 models are evaluated and compared based on their ROC curves. Out of all the models, CNN-3 has the best performance for Test A set while it closely follows the performance of the Inception-V1 model for Test B set. For both sets A and B, the best performing model amongst standard CNN models is Inception-V1, traditional ML models is SVM with linear kernel and full input, conventional models is SI-ED and proposed models is CNN-3. The best performing models from the single channel evaluation for each model category are compared for wideband detection of SPN-43 across multiple channels observed simultaneously with a single receiver by using Free-response ROC (FROC) curves. For set A, CNN-3 outperforms the other 3 models while for set B Inception-V1 reports the best area under the curve value followed by SVM and then CNN-3. It is further observed that the CNN-3 has the fastest detection time among ML models. CNN-3 is further used for the classification of the complete set of spectrograms following which a spectrum occupancy estimate for SPN-43 is provided, and the power of non-SPN-43 emissions is characterized.

By presenting SS as a binary classification problem, Zheng et al. [24] proposed a CNN architecture which in addition to 2 basic convolution layers consists of 6 residual blocks in cascade. The model was made robust to noise power uncertainty by normalizing the received signal power. Signal data for eight modulation techniques is simulated, including 32QAM, 16QAM, 8-Pulse Amplitude Modulation (8PAM), 4PAM, Quadrature Phase-Shift Keying (QPSK), 4-Frequency-Shift Keying (4FSK), 2FSK and BPSK. An equal number of AWGN or colored noise samples are simulated and used for training the DL model. An accuracy of 90.55% was achieved on the test data. The model outperformed two conventional SS models, frequency domain entropy-based method and maximum-minimum eigenvalue ratio-based method by reporting a higher PoD.

To test the performance of the model with untrained data, the work also simulated test samples having modulation types 8- Phase-Shift Keying (8PSK), 8FSK and 64QAM. After fixing the PFA at 0.01 it was observed that the signals were sensed with high probability. To observe the model's performance with real-world signals it was tested against Aircraft Communications Addressing and Reporting System (ACARS) signals. The ACARS samples when used to fine tune the model by adopting a TL-based approach outperforms the two conventional techniques. The model is robust to pink noise in contrast to the conventional approaches whose performance degrades with pink noise proving that DL can automatically extract noise characteristics from data.

The 'Deep Sensing' CNN introduced in work [80] comprises of 2 convolutional and 2 dense layers. The received radio signals are filtered to limit noise using a rectangular band limited filter and sampled in MATLAB to produce a discrete time sequence. The work detects narrowband Gaussian-distributed signal in AWGN to compare the performance of Deep Sensing with an optimal sensing algorithm whose analytical expression is available in accordance with log-likelihood ratio. Deep Sensing outperforms ED and has a performance close to the optimal sensing algorithm when compared by using PoD and PFA. The authors further examined the robustness of Deep Sensing using narrowband Gaussian signals with zero mean in AWGN and QPSK signals. Since Deep Sensing was not effective on communication scenarios that differed from the training set, experiments with TL without labels and fine tuning were conducted. When a limited amount of labelled data is available, the model has proven to be robust across various domains.

Ahmed et al. [45] proposed a CNN-based approach called 'Deep-CRNet' which consists of 85 layers in total and has 5 convolution blocks coupled with 2 intermediate residual- inception blocks. In a communication network comprising of Internet of Things (IoT) and Unmanned Aerial Vehicles (UAVs), the model performs opportunistic spectrum access-based SS for SUs. Complex waveforms of PU and noise signals are generated artificially. The authors

generate complex signal frames using eight different modulation techniques: 64-QAM, 16-QAM, Continuous Phase Frequency Shift-Keying (CPFSK), Gaussian Frequency-Shift Keying (GFSK), BPSK, 8-PSK, QPSK and PAM4 and each of these frames is separately impacted by an independent Rayleigh multipath fading channel, clock offset, and AWGN. The SNR range is kept between -20 dB to $+25$ dB, in increments of 5 dB. Equal number of AWGN samples are generated by changing the noise power in a range of -100 dBm to -5 dBm in steps of 5 dB.

The model achieves an accuracy of 99.74% in differentiating between the signal and noise frames. The performance of Deep-CRNet for over the air signals is assessed by using signal frames from RadioML [93], [94], [95], [96], [97] dataset having the modulation techniques of 64-QAM, 16-QAM, 8-PSK, BPSK and QPSK. The model outperforms other state-of-the-art pre-trained DNN architectures of GoogLeNet [74] and MobileNetV2 [75]. Deep-CRNet demonstrates superior detection performance when compared with other benchmark traditional and DL-based SS schemes.

In [22], 2000 FSK and Amplitude-Shift Keying (ASK) modulated signals are synthetically generated by using an Arduino Uno microcontroller board and a 433 MHz transmitter. The signals are received by an RTL-SDR receiver which is connected to MATLAB. The received signals are transformed into time-frequency representations and are classified as PU signals or noise by a CNN classifier having two convolutional layers. If an SU is positioned closer to PU it can sense the spectrum more reliably therefore the authors generate their data by varying the distances between the sender and receiver. CNN surpasses the performance of ED, ANN, and SVM models.

CNNs are also suitable for CSS as just like the adjacent pixels in an image are correlated, the sensing outcomes from nearby SUs and adjacent bands have spectral and spatial correlations [98]. Lee et al. [98] used a CNN structure having three convolutional and two dense layers, 'Deep Cooperative Sensing' (DCS) in a CSS scenario to combine the sensing decisions from multiple mobile SUs. It was assumed that the PU can simultaneously occupy multiple bands and that SUs do not transmit when sensing is performed. Each SU senses the spectrum using ED and DCS combines the decisions regardless of if they are Hard Decisions (HDs) or Soft Decisions (SDs). DCS is evaluated with a parameter 'sensing error' which is determined by averaging the probabilities of missed detection and false alarm. The performance of DCS is compared with the conventional sensing methods: K-out-of-N and SVM with linear kernel and it is found that DCS with SD shows the lowest sensing error followed by DCS with HD. DCS is robust in the conditions of high noise power densities and lower count of training samples or SUs but takes the highest computation time.

In another example of CNN-based CSS, [32] used a CNN with one convolutional layer for performing data fusion with five SUs and a mobile PU. BPSK modulated random bits with

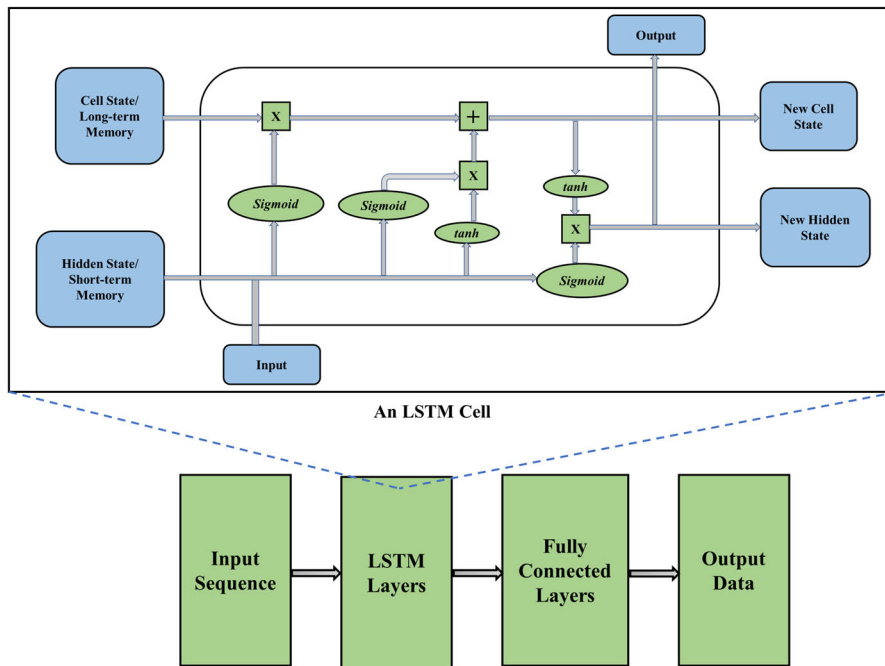


FIGURE 9. Architecture of a simple LSTM model with an LSTM cell [102].

Rayleigh and Nakagami- m fading are transmitted and raw I/Q samples for SUs are generated and split into training and test sets. The performance of the proposed model is analyzed with the help of classification accuracy and is compared with CSS with ED-based HD fusion rules: AND majority rules and OR. CNN-based CSS scheme outperforms ED-based AND and majority rules. It is observed that for all models, performance under Rayleigh fading is superior to Nakagami- m fading.

Table 6 provides a comprehensive summary of the CNN-based SS algorithms surveyed in this work.

C. LONG SHORT-TERM MEMORY (LSTM) NETWORKS

Recurrent Neural Networks (RNNs) are a class of ANNs used with time series data in which output of a layer is provided as feedback to the input to determine the output of that layer. LSTMs are a type of RNNs which can capture long term time dependencies and can exploit correlation from the time-series spectrum data. Fig. 9 depicts the architecture of a simple LSTM model and an LSTM cell. The LSTM algorithm proposed in [78] was summarized in the previous subsection on CNNs.

In another work, Soni et al. [99] analyzed the temporal correlation within the spectrum data captured through an empirical setup with the help of an LSTM network. To make the LSTM unbiased, data at very low SNR values is included. An LSTM with a single hidden unit has the best validation accuracy. LSTM-based SS (LSTM-SS) which can capture the temporal correlation from the input data and PU Activity Statistics-based SS (PAS-SS) models are introduced in this work. PU activity statistics and occupancy patterns can be estimated using the sequence of sensing decisions

and in a CR network, this statistical information will be helpful in predicting spectrum occupancy trends, planning SS, selecting the right spectrum band and channel for CR system, maximizing system performance, and improving spectral efficiency [29]. PAS-SS consists of an LSTM model with 3 hidden layers for prediction and an ANN with a single hidden layer for classification. For acquiring data for experiments with LSTM-SS and PAS-SS, two empirical bed setups are used with a USRP and a digital spectrum analyzer respectively. When LSTM-SS is compared with CNN and ANN with the help of PoD versus SNR curves, LSTM-SS achieves the best detection performance. On comparing the classification accuracy versus SNR curves, training and execution times of LSTM-SS with the ML techniques ANN, Gaussian Naïve Bayes, and Random Forest it is observed that LSTM-SS achieves the best classification accuracy but reports longer training and execution times.

A summary of research works utilizing LSTM networks for SS has been provided in Table 7.

D. COMBINED CNN-LSTM ARCHITECTURES

A CNN can extract spatial features from the input data while an LSTM network captures temporal variations. The CNN and LSTM techniques can be combined to extract complex features from data. A simple CNN-LSTM model is shown in Fig. 10. In the presence of noise uncertainty, when the SNR value falls below a threshold called the SNR-wall, an SS detector will fail to perform. The problem of SNR-wall can be solved by utilizing the structure of the PU signal, adding diversity, and reducing the noise uncertainty [100].

As an alternative to conventional energy detectors that suffer from the SNR-wall problem, work [101] uses DL

TABLE 6. Summary of CNN-based approaches for spectrum sensing.

Research Work	Description of Input to CNN algorithm	Key Data Pre-processing Used	Key Features of CNN Architecture	Key Evaluation Metrics	Key Performance Notes
[81], 2017	Cyclostationary and energy features	Standardization of cyclostationary and energy feature sets	CNN with 1 convolution layer	PoD versus SNR curves	PoD about 0.5 higher than CFD at SNR = -20 dB.
[98], 2019	Sensing data generated from multiple SUs by simulation	-	CNN with 3 convolution layers called 'Deep Cooperative Sensing' (DCS)	Sensing error, ROC curve, computation time	Area Under Curve (AUC) for DCS with SD = 0.952 and DCS with HD = 0.95.
[60], 2019	CM of current frame and a matrix formed by stacking CMs from past frames	Appropriate stacking of CMs from past frames	CNN with 2 convolution layers inspired from LeNet architecture	PoD versus SNR curves	PoD > 0.725 with deterministic PU activity and uncorrelated model, SNR = -20 dB, antennae = 8.
[84], 2019	Sample CMs PU signals: i.i.d. and exponential correlation models Noise: Gaussian and sea clutter	-	CNN with 2 convolutional layers inspired from LeNet-5 architecture	ROC curves, PoD versus SNR curves	PoD = 96.7% at PFA = 1.9% and SNR = -18 dB.
[78], 2019	3.5 GHz band spectrograms	-	CNN with 1 convolution layer called CNN-3	ROC and FROC curves	CNN-3 has FROC- AUC of [0.994, 0.998] for Set A and 95% confidence interval.
[32], 2020	BPSK modulated randoms bits incorporating Rayleigh and Nakagami-m fading, AWGN	Data normalization	CNN with 1 convolution layer	Classification accuracy	Classification accuracy of 77.99% for Rayleigh fading and 77.33% for Nakagami fading.
[24], 2020	Signals of 8 modulation types and noise generated by simulation	Power normalization	CNN with 2 convolution layers, 6 residual blocks in cascade	PoD versus SNR curves	Classification accuracy: 90.55% PoD > 0.6 for SNR = 10 dB and pink noise.
[80], 2020	Filtered and sampled radio signals generated in MATLAB	-	CNN with 2 convolution layers called 'Deep Sensing'	ROC Curves	PoD > 0.8 for PFA > 0.05, SNR = -4 dB.
[45], 2021	Synthetically generated complex waveforms of PU signal and noise	-	CNN with 85 layers with convolution and residual-inception blocks, called 'Deep-CRNet'	ROC and precision recall curves	Classification accuracy = 94.47% at SNR = 0 dB.
[22], 2021	Artificially generated ASK and FSK modulated signals with an Arduino Uno card and a 433 MHz wireless transmitter	Transformation into time- frequency representations	CNN with 2 convolution layers	PoD, PFA, classifier accuracy.	PoD = 0.994, PFA = 0.0023 and classifier accuracy = 100%.

TABLE 7. Summary of LSTM-based approaches for spectrum sensing.

Research Work	Description of Input to LSTM algorithm	Key Data Pre-processing Used	Key Features of LSTM Architecture	Key Evaluation Metrics	Key Performance Notes
[78], 2019	3.5 GHz band spectrograms	-	LSTM cells with 50% dropout between them followed by an fc layer and output layer	ROC curves	ROC- AUC of [0.987, 0.998] for Set B and 95% confidence interval.
[99], 2020	Data acquired using USRP added with AWGN and AWGN, digital spectrum analyzer	-	LSTM cells followed by sigmoid activation called 'LSTM-SS', LSTM with 3 hidden layers + ANN with one hidden layer called 'PAS-SS'	PoD versus SNR curves, ROC curves, classification accuracy vs SNR curves	Classification accuracy > 90% for SNR >= -10 dB.

to extract the hidden structures of the PU signals and propose 'DetectNet' which consists of convolutional, LSTM, and fc layers. The signal samples are generated using the

RadioML2016.10a dataset [93], [94], [95], [96], [97] in eight modulation types: 8PSK, BPSK, QPSK, GFSK, CPFSK, QAM64, QAM16 and PAM4. The negative samples are

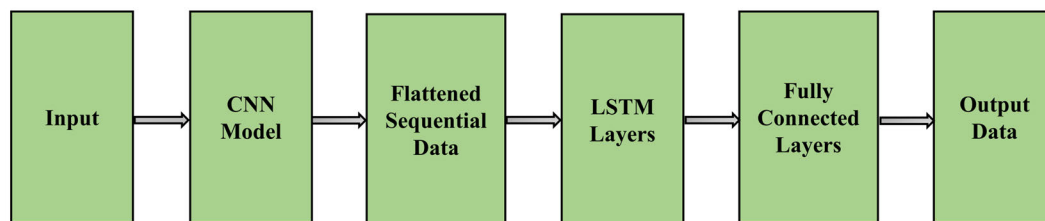


FIGURE 10. Architecture of a simple CNN-LSTM model.

additive noises having zero mean Circularly Symmetric Complex Gaussian (CSCG) distribution. Energy normalization is performed on the training samples as the simulation results are minimally impacted by energy. Furthermore, the signal modulation structure can be better exploited, and detector model can have better generalization ability. The paper uses a customized two-stage training strategy based on constant false alarm rate detector to control the performance of the DL detector. The first stage of the model involves early stopping at 6 epochs while in the second step the model observes metrics trade-off characteristic where the validation loss and accuracy are both kept stable while PFA and PoD at different SNRs vary with each epoch.

The authors further propose and optimize other DNN models which are an MLP with four fc layers, a CNN with two convolutional and one fc layer and an LSTM with two LSTM layers. The detection performance of these models is compared with that of DetectNet by considering QAM16 modulated signals. It is found that DetectNet and CNN have superior performance than MLP and LSTM and DetectNet outperforms CNN by achieving a low value of PFA for similar PoD. For a fixed sample length, DetectNet gives the best performance for FSK modulated signals. While evaluating the generalization ability of DetectNet it is found that when the training and testing data have similar modulation schemes, the model has good generalization power, and the performance deteriorates when the training and test sets have different modulation types. For a higher sample length of data, the performance of DetectNet improves. The work further proposes ‘SoftCombinationNet’ for a CSS scenario which provides a global decision by combining the soft information from sensing nodes. DetectNet is deployed at each sensing node and the probability vectors from these nodes are processed by a neural network having three fc layers to make the final decision about the status of the PU.

Xie et al. [103] proposed a CNN-LSTM-based model which first extracts energy correlation features from covariance matrices generated by sensing data with the help of CNN layers and then inputs the series of energy correlation features corresponding to multiple sensing periods into an LSTM network to learn PU activity patterns. This DNN architecture consists of two convolutional layers followed by an LSTM layer and a dense layer. The work mentions that DL-based SS architectures are not susceptible to the signal-noise model assumptions since they learn directly from the

sensing data. The PU signals are QPSK modulated signals having unit energy while the noise signals are simulated by following Gaussian and Laplace distributions. For the PU activity pattern, the work considers lognormal state sojourn time model where state transitions are represented by a semi-Markov process [104], [105] and the real state sojourn time model for which real sensing data is collected using the USRP-2922. The CNN-LSTM detector outperforms the detectors MED [87], Signal Subspace Eigenvalues (SSE) detector [106], Arithmetic to Geometric Mean (AGM) detector [106] and the DL-based APASS detector [60] in scenarios with and without noise uncertainty.

In [20] ‘DLsenseNet’ which is based on combined CNN-LSTM architecture is used to capture both spatial and temporal details of data and for sensing the spectrum. DLsenseNet is made of a modified inception block, LSTM layers and fc layers. Samples of eight different types of digital modulation from the RadioML2016.10b dataset [93], [94], [95], [96], [97] are used to represent PU signals and the absence of PU is represented by CSCG noise vector with zero mean. I/Q components of these signals form input to the model which then predicts the status of the PU. The signals are energy normalized to make DLsenseNet independent of energy and have greater generalization capacity in environments where background noise changes. The work analyzed the effect of various modulation schemes and length of data samples and compared the performance of DLsenseNet with other DNN models: CNN, residual network (ResNet), LeNet, inception module, LSTM, Convolutional Long short-term Deep Neural Network (CLDNN) and previously reported SS techniques of DetectNet [101] and CNN-LSTM [103]. DLsenseNet outperforms other models because it integrates the advantages of CNN, LSTM, and inception models. The performance of DLsenseNet is compared by using eight different modulation schemes and there is very little difference in detection performance between various modulated signals, suggesting that DLsenseNet is insensitive to the order of modulation.

Xing et al. [21] proposed a BiLSTM-based DNN by combining convolutional, concatenated, BiLSTM, Self-Attention (SA) layers and fc layers. The model utilizes the hidden states by simultaneously scanning data in opposite directions. The PU signals in eight modulation schemes: BPSK, QPSK, 8PSK, CPFSK, GFSK, QAM16, QAM64 and PAM4 are generated using GNU radio while complex

AWGN signals are used as noise signals. In an ablation study, the impact of BiLSTM, SA, and concatenated layers is observed on sensing error by comparing five models: CNN, CNN-LSTM, CNN-BiLSTM, CNN-BiLSTM-SA, and CNN-BiLSTM-SA-CONCAT by using QAM 16 signals. The CNN model comprises of two convolutional blocks each having a one-dimensional convolutional layer which helps extract sufficient local features, especially in environments with low SNR [80]. The features lost during transmission by the one-dimensional CNN are compensated by the concatenated layer which is denoted by CONCAT and combines the input and output from the CNN. BiLSTM layers help capture the long- and short-term time dependencies from the input data in opposite directions and with the help of the SA layer, the model highlights the most important features obtained by the BiLSTM layers. It is found that for SNR values of less than -5 dB, there is a significant difference in the performance of the five models. CNN-BiLSTM-SA-CONCAT achieves the best sensing performance followed by CNN-BiLSTM-SA, CNN-BiLSTM, CNN-LSTM and CNN. The models achieve comparable sensing results in an SNR range above -5 dB.

The CNN-BiLSTM-SA-CONCAT model is compared against the CNN architecture used in [80], the ResNet model in [24], the LSTM network in [99] and the CLDNN proposed in [101] with their default hyperparameter settings using QAM 16 signals. It is observed that the models using LSTM take more time for both training and detection as unlike CNNs, LSTM operations cannot be parallelized by a GPU. The CNN-BiLSTM-SA-CONCAT model takes the longest training and detection times but reports the least sensing error under all SNR conditions. The authors further observed the results of the model under eight modulation schemes and different SNR values by keeping a fixed sample length of 128. The model has comparable performance for all modulation schemes, but the GFSK modulated signals have the lowest sensing error in the SNR range of -12 dB to -5 dB. The robustness of CNN-BiLSTM-SA-CONCAT is studied by training it with QAM16 modulated signals and testing it individually with signals having a different modulation scheme (BPSK, QAM64, PAM4 and GFSK). The classifier achieves good detection results with minimal deterioration in performance. When the model is trained individually with QAM16 signals having sample lengths of 64, 128, 256 and 512, it is established that the sample length of 512 reports the least sensing error as longer signal sample consists of more temporal information.

Table 8 provides a summary of CNN-LSTM-based approaches discussed for SS by describing the input to the CNN-LSTM algorithm, the data pre-processing technique used, key features of the CNN-LSTM architecture, main evaluation metrics and performance of the SS model.

E. AUTOENCODERS (AEs)

An AE neural network is an unsupervised learning algorithm that reduces the dimensionality of input data and reconstructs

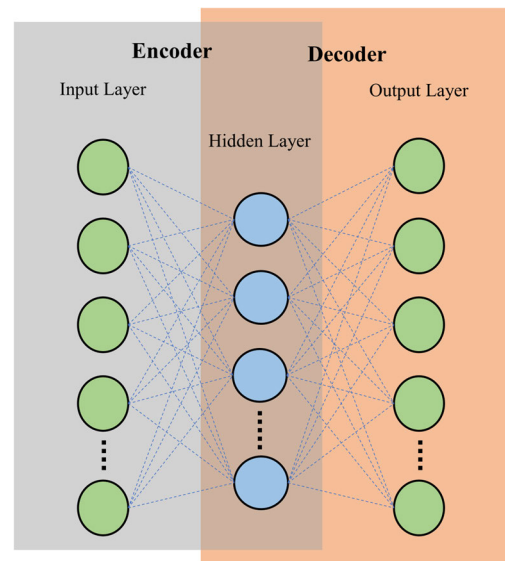


FIGURE 11. Structure of a conventional autoencoder.

the original data [107]. An AE has three layers: input, hidden, and output or reconstruction layers. Each AE undergoes an encoding-decoding process during training. The encoding process maps the input data into a hidden representation, and the decoding process reconstructs input data from the hidden representation [108]. Fig. 11 shows the structure of a conventional AE. AE architectures having two or more hidden layers in both encoder and decoder have been utilized for SS.

Cheng et al. [57] proposed two novel Stacked Autoencoder (SAE) frameworks having two hidden layers for sensing of OFDM signals. To create the trained SAE model, the input and hidden layers of all the trained AEs are stacked together layer by layer. The first framework is termed Stacked Autoencoder-based SS (SAE-SS). SAE-SS is first pre-trained wherein the features of the PU are extracted in an unsupervised manner following which a Logistic Regression classifier is used to fine tune the model. As the sensing accuracy of SAE-SS degrades for low SNR values, another model, Stacked Autoencoder-based SS using Time-Frequency domain signals (SAE-TF) is proposed. With SAE-TF, both time and frequency domain signals are used as input which allows the model to extract more hidden features and perform better in conditions of low SNR. The received signals in time domain are transformed to frequency domain with the help of Fast Fourier Transform and then manipulated to provide input to the SAE-TF model along with the time domain input.

SAE-TF has a better sensing performance than SAE-SS but has twice as many input units as SAE-SS resulting in a higher training complexity. The OFDM system is generated with BPSK modulation. For training of SAE-SS and SAE-TF models, the training data is divided according to the SNR values for training different SAE architectures for a higher detection performance. The performance of SAE-SS

TABLE 8. Summary of CNN-LSTM-based approaches for spectrum sensing.

Research Work	Description of Input to CNN-LSTM algorithm	Key Data Pre-processing Used	Key Features of CNN-LSTM Architecture	Key Evaluation Metrics	Key Performance Notes
[101], 2019	Positive samples: data generated in 8 modulation techniques based on 'RadioML2016.10a' dataset, negative samples: additive noises based on zero mean circularly symmetric complex Gaussian distribution	Energy normalization	CNN-LSTM with 2 convolution and 2 LSTM layers called 'DetectNet' 'SoftCombinationNet' for CSS: 3 dense layers	PoD versus SNR curves	SNR-wall of DetectNet = -8.5 dB for PFA = 5.92% and PoD = 90%.
[103], 2020	Sample covariance matrix from the self-generated samples	Use of sample covariance matrix to reduce computational complexity	CNN-LSTM with 2 convolution layers and 1 LSTM layer	ROC Curves, PoD versus SNR curves	PoD close to 0.4 at SNR = -20 dB, PFA = 0.1, for real state sojourn time model without noise uncertainty.
[20], 2021	'RadioML2016.10b', CSCG noise vector with zero mean	Energy normalization	CNN-LSTM having an Inception block followed by an LSTM block, called 'DLSenseNet'	PoD, PFA, sensing error	For QAM-16 signals with sample length of 512, PFA = 0.04%, sensing error = 12.87% and PoD = 46.98% at SNR = -20 dB.
[21], 2022	Data generated using GNU Radio	Data normalization	CNN-LSTM having convolution, BiLSTM and SA layers	Sensing error versus SNR curves, ROC curves, detection, and training time	Sensing error < 0.35 at SNR = -20 dB.

and SAE-TF models are compared with conventional OFDM signal sensing techniques: ED, CP [109] and CM [110]-based SS and with neural network-based techniques: ANN model in [52] and CNN model in [81] under various conditions. When comparing the Probability of Miss detection (PM) while varying the SNR values, it is observed that the proposed SAE models have the least PM values, with SAE-TF outperforming SAE-SS.

In work [111], a Variational Autoencoder (VAE) and unsupervised DL-based detector termed Unsupervised Deep SS (UDSS) is designed to limit the amount of labelled data required for training the SS model. The VAE proposed in [112] has a probabilistic model parameter layer after the hidden layer as opposed to a conventional AE [113]. A VAE-GMM-based approach having three hidden layers each in encoder and decoder is used to separate data into two clusters and a small amount of labeled noise data is then used to identify the clusters representing PU signals and noise. PU signals are represented by unit energy QPSK modulated symbols while noise samples are either Gaussian or Laplacian. CMs are calculated and vectorized before being input to the neural network. The sensing performance of UDSS is compared to MED, SSE detector, AGM detector, CNN [84] and Kernel K-Means [114] with the help of ROC curves. Under all simulation settings, the performance of UDSS is close to that of CNN even with limited labelled data and it outperforms all other detectors.

Subray et al. [107] classified LTE and Wi-Fi signals for SS with three types of AE neural networks: Deep, Variational and LSTM AEs. Both encoders and decoders of the deep AE and VAE consist of two hidden layers. In LSTM AE, the input data is encoded and decoded using LSTM cells. The

LSTM AE consists of three hidden layers both in encoder and decoder. USRP B210 and GNU Radio were used to capture the LTE signals while MATLAB was used to generate Wi-Fi signals. The signal strength of LTE signals was matched to that of Wi-Fi signals by addition of 20 dB. Four features were provided as inputs for all AEs: I/Q samples and the phase and amplitude values derived from the I/Q samples. The AEs were evaluated by taking different combinations of signals from 802.11ax and 802.11ac Wi-Fi protocols. With the main evaluation metrics being precision and recall, the Deep AE using exponential linear unit activation was found to be most efficient for the classification task.

The AE-based SS algorithms discussed in this subsection are summarized in Table 9.

IV. RADIO FREQUENCY (RF) SIGNAL DATASETS, TRANSFER LEARNING (TL) AND SOFTWARE USED

The application of ML algorithms, especially DL algorithms is becoming increasingly popular in the field of SS. The training, validation, and testing of these models requires huge amounts of data. There are a few publicly available RF datasets having a collection of various modulations with 'RadioML' released by DeepSig Inc. [93], [94], [95], [96], [97] being the most popular. RadioML datasets have been widely used by researchers for addressing modulation classification and SS problems [20], [45], [101] with the help of ML. The latest version is 'RadioML 2018.01A' which comprises of synthetic simulated channel effects and over-the-air recordings of 24 digital and analog modulation techniques. This data is available in hdf5 format as complex floating-point values, with 2 million examples, each having a sample length

TABLE 9. Summary of AE-based approaches for spectrum sensing.

Research Work	Description of Input to AE algorithm	Key Data Pre-processing Used	Key Features of AE Architecture	Key Evaluation Metrics	Key Performance Notes
[57], 2019	OFDM system generated with BPSK modulation	FFT of time domain signals for SAE-TF	SAE-SS and SAE-TF both have 2 hidden layers followed by Logistic Regression classifier	PM versus SNR, PFA versus PM, and PM versus ratio of incorrect labels curves, training time	PM of SAE-TF increases from 0.0098 to 0.0114 with an increase in noise uncertainty from 0.5 dB to 1 dB at SNR = -10 dB.
[111], 2020	Flattened CMs of Unit energy QPSK modulated signals and Gaussian or Laplacian noise	Calculation and vectorization of CMs	VAE with 3 hidden layers in both encoder and decoder	ROC curves, PoD versus SNR curves	PoD > 0.3 with Laplacian noise and uncorrelated model, SNR = -14 dB, PFA = 0.001, antennae = 16.
[107], 2021	I/Q samples, phase and amplitude values derived from the I/Q samples of LTE signals acquired using a USRP and Wi-Fi signals generated in MATLAB	Addition of 20 dB to LTE signals	Deep AE: 2 hidden layers VAE: 2 hidden layers LSTM AE: 3 hidden layers	Precision and recall	Precision= 99.98%, recall= 87.76% for Deep AE using IEEE 802.11ax samples with 64-QAM modulation.

TABLE 10. Description of some publicly available RF datasets.

RF Dataset	Total Modulations Included	SNR Range	Sample Length
RadioML 2018.01A	24	-20 dB to +30 dB at intervals of 2 dB	1024
RadioML 2016.10A	11	-20 dB to +18 dB at intervals of 2 dB	128
Panoradio SDR	18	-10 dB to +25 dB at intervals of 5 dB	2048
MIGOU-MOD	11	22 dB, 37 dB	128

of 1024. A previous version of this dataset, ‘RadioML 2016.10A’ consists of 8 digital modulated and 3 analog modulated signals in varying SNRs and is available as a pickle file.

Apart from RadioML datasets, there are some other RF datasets that can be explored by researchers for testing the robustness of their SS models. For example, an RF dataset by Panoradio SDR [115] was generated synthetically by the application of Gaussian noise, Watterson fading and with random frequency and phase offsets to speech, music, and text signals. It is designed for signal and modulation classification tasks using novel ML algorithms and has signals from 18 different transmission modes. It consists of 172,800 signal vectors with each vector having 2048 I/Q samples. Another dataset, MIGOU-MOD [116], [117] is a dataset acquired from a low-power IoT platform called ‘MIGOU’ and has over-the-air measurements of real radio signals modulated with 11 modulation types. The signals were generated with the help of a USRP and GNU Radio software and recorded using the MIGOU platform in an office environment. The main properties of these datasets are summarized in Table 10.

In addition to these published RF datasets, many works have generated data specific to their experiments using MATLAB and Python. Some works have captured real-world data for their experiments with the help of a USRP and GNU Radio Software. However, the wireless communications and RF signals domain lacks robust and comprehensive datasets comparable to those in domains like speech, handwriting, and object recognition [40]. The use of TL can be an effective strategy when a limited amount of training data is available. ImageNet [72] is a large database of images that has been used to train many standard CNN architectures such as VGG-16 [76] and ResNet-50 [77]. With the help of TL, the knowledge gained by these standard models to classify the ImageNet dataset can be transferred to RF datasets for SS. Works [45] and [78] have explored some standard CNN models and have evaluated their performances for SS. Table 11 provides a description of the architectures of the standard CNN models used in these works.

V. RESEARCH CHALLENGES RELATED TO THE USE OF DL FOR SS

A. REQUIREMENT OF LARGE AMOUNTS OF DATA AND LACK OF COMPREHENSIVE RF DATASETS

Although DL-based approaches are becoming increasingly popular in SS, the lack of comprehensive RF datasets poses a major challenge to the deployment of DL algorithms [40]. DL techniques are data-hungry and require huge amounts of data for training, validating, and testing their models. The concept of TL can be explored to utilize the model weights of pre-trained DL architectures to enhance the sensing performance. In addition, data augmentation could be used to increase the amount and quality of data and add diversity to the dataset [118]. By augmenting the dataset with relevant data, the trained model can be made more robust, improving its overall performance significantly [119]. Furthermore, the use of unsupervised DNN architectures which require only

TABLE 11. Description of the architectures of standard CNN models.

Standard CNN Architectures and Research Works that have used them	Description of Architecture
GoogLeNet/ Inception-V1 [45], [78]	22 layers having a total of 9 inception blocks
MobileNetV2 [45]	19 residual bottleneck layers following the initial fully convolutional layer of 32 filters.
VGG-16 [78]	16 layers out of which 13 layers are convolutional layers and 3 layers are fc layers.
VGG-19 [78]	19 layers among which 16 are convolutional layers and 3 layers are dense layers.
ResNet-18 [78]	Deep residual learning framework, 17 convolutional layers.
ResNet-50 [78]	Deep residual learning framework, 49 convolutional layers.
DenseNet-121 [78]	All layers are interconnected, 120 convolutional layers.

a limited amount of labelled data like in work [111] is a potential solution for this research problem.

B. LACK OF VALIDATION OF RESULTS ON REAL-WORLD DATA

Just as in the case of modulation classification, most of the literature on ML-based SS is based on simulations or theoretical results, so it is unclear how well a model performs in real-life settings [120]. To add to it, the algorithms are often trained for specific modulation and noise considerations and their performance cannot be generalized. To ensure the robustness of DL algorithms in practical environments, real-world datasets should be used to validate their performance [23].

C. HIGHER OFFLINE TRAINING TIME

Even though DNN techniques result in enhanced detection performance, these models have high computational complexity. DNN-based models require huge amounts of data for training which results in higher offline training times. Consequently, more research should be conducted on enhancing detection performance through limited training data to decrease training time. This issue can also be mitigated by the use of specialized hardware with high processing capabilities like GPUs which can significantly cut down the training time [57]. Moreover, the training is usually required only occasionally when the sensing conditions change, and the well-trained model can then generate fast predictions.

D. POOR PERFORMANCE AT LOW SNR

DNN-based SS algorithms generally perform well in conditions of high SNR, but their performance drastically falls at low SNRs. The SNR for SS is usually low, and the received signal energy and noise levels at the SU fluctuate over time [111]. Work [57] addressed this issue by training

different SAE architectures with data in different SNR conditions. Along with it, the SAE-TF model used both time and frequency domain samples as input to ensure robust performance in regions of low SNRs. The focus of future DL-based SS research should be to ensure robust performance at low SNRs as a model that performs better at low SNRs also performs well in conditions of high SNR [20]. It is also important for researchers to explore the interpretability of radio signal data for achieving a higher detection performance.

VI. CONCLUSION

As a principal enabler of DSM, CR technology helps SUs access the unutilized bands of the PUs and improves spectrum utilization. The SS capability of CR determines the availability of radio resources of PUs in order for the SUs to utilize the vacant frequency bands. The use of DL algorithms for SS can enhance the performance of 6G networks. This paper surveyed various DNNs used for the task of SS by classifying them as MLPs, CNNs, LSTM networks, combined CNN-LSTM architectures, and AEs. The DL algorithms are compared based on the dataset used, data acquisition technique, data pre-processing method, algorithm architecture, evaluation metrics, obtained results and comparison of results with standard SS detectors. DNNs are increasingly being used to sense the PU spectrum due to their automated feature extraction capabilities and their ability to adapt to changing radio environments. This work also highlighted the shortcomings of conventional SS and traditional ML approaches while presenting an overview of traditional ML algorithms and simple ANNs. Traditional ML techniques involve manual feature extraction, while conventional sensing techniques suffer from missed detection of PUs and generate false alarms. In addition, this paper summarized some publicly available RF signal datasets, the concept of TL and the need to have diverse RF signal datasets. Finally, the research challenges associated with the use of DL techniques to identify vacant frequency bands were discussed along with potential solutions.

REFERENCES

- [1] S. N. Syed, P. I. Lazaridis, F. A. Khan, Q. Z. Ahmed, M. Hafeez, V. Holmes, I. P. Chochliouros, and Z. D. Zaharis, "Deep learning approaches for spectrum sensing in cognitive radio networks," in *Proc. 25th Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Oct. 2022, pp. 480–485.
- [2] G. K. Papageorgiou, K. Voulgaris, K. Ntougias, D. K. Ntaikos, M. M. Butt, C. Galitto, N. Marchetti, V. Frascolla, H. Annouar, A. Gomes, A. J. Morgado, M. Pesavento, T. Ratnarajah, K. Gopala, F. Kaltenberger, D. T. M. Slock, F. A. Khan, and C. B. Papadias, "Advanced dynamic spectrum 5G mobile networks employing licensed shared access," *IEEE Commun. Mag.*, vol. 58, no. 7, pp. 21–27, Jul. 2020.
- [3] A. Ahmed, Z. Elsaraf, F. A. Khan, and Q. Z. Ahmed, "Cooperative non-orthogonal multiple access for beyond 5G networks," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 990–999, 2021.
- [4] F. C. Okogbaa, Q. Z. Ahmed, F. A. Khan, W. B. Abbas, F. Che, S. A. R. Zaidi, and T. Alade, "Design and application of intelligent reflecting surface (IRS) for beyond 5G wireless networks: A review," *Sensors*, vol. 22, no. 7, p. 2436, Mar. 2022.

- [5] Ericsson. *Mobile Subscriptions Outlook*. Accessed: Feb. 28, 2023. [Online]. Available: <https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-subscriptions-outlook>
- [6] C. D. Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 836–886, 2021.
- [7] Y.-C. Liang, J. Tan, H. Jia, J. Zhang, and L. Zhao, "Realizing intelligent spectrum management for integrated satellite and terrestrial networks," *J. Commun. Inf. Netw.*, vol. 6, no. 1, pp. 32–43, Mar. 2021.
- [8] K. A. Yau, J. Qadir, C. Wu, M. A. Imran, and M. H. Ling, "Cognition-inspired 5G cellular networks: A review and the road ahead," *IEEE Access*, vol. 6, pp. 35072–35090, 2018.
- [9] S. T. Muntaha, P. I. Lazaridis, M. Hafeez, Q. Z. Ahmed, F. A. Khan, and Z. D. Zaharis, "Blockchain for dynamic spectrum access and network slicing: A review," *IEEE Access*, vol. 11, pp. 17922–17944, 2023.
- [10] Q. Z. Ahmed, M. Hafeez, F. A. Khan, and P. Lazaridis, "Towards beyond 5G future wireless networks with focus towards indoor localization," in *Proc. IEEE 8th Int. Conf. Commun. Netw. (ComNet)*, Hammamet, Tunisia, Oct. 2020, pp. 1–5.
- [11] A. Ivanov, K. Tonchev, V. Poulkov, and A. Manolova, "Probabilistic spectrum sensing based on feature detection for 6G cognitive radio: A survey," *IEEE Access*, vol. 9, pp. 116994–117026, 2021.
- [12] Y.-C. Liang, *Dynamic Spectrum Management*. Singapore: Springer, 2019, doi: [10.1007/978-981-15-0776-2](https://doi.org/10.1007/978-981-15-0776-2).
- [13] L. Li, F. Khan, M. Pesavento, T. Ratnarajah, and S. Prakriya, "Sequential search based power allocation and beamforming design in overlay cognitive radio networks," *Signal Process.*, vol. 97, pp. 221–231, Apr. 2014.
- [14] G. Ding, J. Wang, Q. Wu, Y.-d. Yao, R. Li, H. Zhang, and Y. Zou, "On the limits of predictability in real-world radio spectrum state dynamics: From entropy theory to 5G spectrum sharing," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 178–183, Jul. 2015.
- [15] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [16] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [17] M. S. Gupta and K. Kumar, "Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues," *J. Netw. Comput. Appl.*, vol. 143, pp. 47–76, Oct. 2019.
- [18] Z. Zhang, W. Zhang, S. Zeadally, Y. Wang, and Y. Liu, "Cognitive radio spectrum sensing framework based on multi-agent arc hitecure for 5G networks," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 34–39, Dec. 2015.
- [19] L. Claudino and T. Abrão, "Spectrum sensing methods for cognitive radio networks: A review," in *Wireless Pers. Commun.*, vol. 95, no. 4, pp. 5003–5037, Apr. 2017.
- [20] S. Solanki, V. Dehalwar, and J. Choudhary, "Deep learning for spectrum sensing in cognitive radio," *Symmetry*, vol. 13, no. 1, p. 147, Jan. 2021.
- [21] H. Xing, H. Qin, S. Luo, P. Dai, L. Xu, and X. Cheng, "Spectrum sensing in cognitive radio: A deep learning based model," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 1, p. e4388, Jan. 2022.
- [22] M. Saber, A. Chehri, A. E. Rharas, R. Saadane, and M. Wahbi, "A cognitive radio spectrum sensing implementation based on deep learning and real signals," in *Innovations in Smart Cities Applications*, vol. 4. Cham, Switzerland: Springer, 2021, pp. 930–941, doi: [10.1007/978-3-030-66840-2_70](https://doi.org/10.1007/978-3-030-66840-2_70).
- [23] D. Janu, K. Singh, and S. Kumar, "Machine learning for cooperative spectrum sensing and sharing: A survey," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 1, p. e4352, Sep. 2021.
- [24] S. Zheng, S. Chen, P. Qi, H. Zhou, and X. Yang, "Spectrum sensing based on deep learning classification for cognitive radios," *China Commun.*, vol. 17, no. 2, pp. 138–148, Feb. 2020.
- [25] V. Sharma and S. Joshi, "A literature review on spectrum sensing in cognitive radio applications," in *Proc. 2nd Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Madurai, India, Jun. 2018, pp. 883–893.
- [26] M. Usman and K. Insoo, "Secure cooperative spectrum sensing via a novel user-classification scheme in cognitive radios for future communication technologies," *Symmetry*, vol. 7, no. 2, pp. 675–688, May 2015.
- [27] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–17, Sep. 2018.
- [28] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.
- [29] S. Kyperountas, N. S. Correal, and Q. Shi, "A comparison of fusion rules for cooperative spectrum sensing in fading channels," EMS Res., Motorola, Tech. Rep., 2010. [Online]. Available: <https://api.semanticscholar.org/CorpusID:16716527>
- [30] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, no. 1, pp. 40–62, Mar. 2011.
- [31] R. Fan and H. Jiang, "Optimal multi-channel cooperative sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1128–1138, Mar. 2010.
- [32] S. P. K. R. Sudhindra, and S. M. N., "Deep learning for cooperative spectrum sensing," in *Proc. 2nd PhD Colloq. Ethically Driven Innov. Technol. Soc. (PhD EDITS)*, Nov. 2020, pp. 1–2.
- [33] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.
- [34] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 1–16, Dec. 2010.
- [35] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.
- [36] S. K. Sharma, T. E. Bogale, S. Chatzinotas, B. Ottersten, L. B. Le, and X. Wang, "Cognitive radio techniques under practical imperfections: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1858–1884, 4th Quart., 2015.
- [37] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1277–1304, 2nd Quart., 2017.
- [38] Y. Arjoune and N. Kaabouch, "A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions," *Sensors*, vol. 19, no. 1, p. 126, Jan. 2019.
- [39] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2209–2221, Nov. 2013.
- [40] T. A. Hall, R. Caromi, M. Souryal, and A. Wunderlich, "Reference datasets for training and evaluating RF signal detection and classification models," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–5.
- [41] J. Kaur, M. A. Khan, M. Ifikhar, M. Imran, and Q. E. U. Haq, "Machine learning techniques for 5G and beyond," *IEEE Access*, vol. 9, pp. 23472–23488, 2021.
- [42] K. W. Choi, E. Hossain, and D. I. Kim, "Cooperative spectrum sensing under a random geometric primary user network model," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1932–1944, Jun. 2011.
- [43] T. Li, S. Zhu, and M. Ogihara, "Using discriminant analysis for multi-class classification: An experimental investigation," *Knowl. Inf. Syst.*, vol. 10, no. 4, pp. 453–472, Mar. 2006.
- [44] Y. Lu, P. Zhu, D. Wang, and M. Fattouche, "Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Doha, Qatar, Apr. 2016, pp. 1–6.
- [45] R. Ahmed, Y. Chen, and B. Hassan, "Deep learning-driven opportunistic spectrum access (OSA) framework for cognitive 5G and beyond 5G (B5G) networks," *Ad Hoc Netw.*, vol. 123, Dec. 2021, Art. no. 102632.
- [46] T. F. Smith and M. S. Waterman, "Comparison of biosequences," *Adv. Appl. Math.*, vol. 2, no. 4, pp. 482–489, Dec. 1981.
- [47] J. Tian, P. Cheng, Z. Chen, M. Li, H. Hu, Y. Li, and B. Vucetic, "A machine learning-enabled spectrum sensing method for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11374–11378, Nov. 2019.
- [48] T. Kim and J.-S. Lee, "Exponential loss minimization for learning weighted naive Bayes classifiers," *IEEE Access*, vol. 10, pp. 22724–22736, 2022.

- [49] Z. Lei and F. P. S. Chin, "Sensing OFDM systems under frequency-selective fading channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1960–1968, May 2010.
- [50] M. Jin, Q. Guo, J. Xi, Y. Li, and Y. Li, "On spectrum sensing of OFDM signals at low SNR: New detectors and asymptotic performance," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3218–3233, Jun. 2017.
- [51] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000.
- [52] M. R. Vyas, D. K. Patel, and M. Lopez-Benitez, "Artificial neural network based hybrid spectrum sensing scheme for cognitive radio," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–7.
- [53] Y.-J. Tang, Q.-Y. Zhang, and W. Lin, "Artificial neural network based spectrum sensing method for cognitive radio," in *Proc. 6th Int. Conf. Wireless Commun. Netw. Mobile Comput. (WiCOM)*, Chengdu, China, Sep. 2010, pp. 1–4.
- [54] D. K. Patel and Y. N. Trivedi, "LRS-G2 based non-parametric spectrum sensing for cognitive radio," in *Proc. CROWNCOM*, 2016, pp. 330–341.
- [55] M. López-Benitez, F. Casadevall, and C. Martella, "Performance of spectrum sensing for cognitive radio based on field measurements of various radio technologies," in *Proc. Eur. Wireless Conf. (EW)*, Lucca, Italy, 2010, pp. 969–977.
- [56] M. López-Benitez and F. Casadevall, "Improved energy detection spectrum sensing for cognitive radio," *IET Commun.*, vol. 6, no. 8, p. 785, 2012.
- [57] Q. Cheng, Z. Shi, D. N. Nguyen, and E. Dutkiewicz, "Sensing OFDM signal: A deep learning approach," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7785–7798, Nov. 2019.
- [58] Z. Elsaraf, F. A. Khan, and Q. Z. Ahmed, "Deep learning based power allocation schemes in NOMA systems: A review," in *Proc. 26th Int. Conf. Autom. Comput. (ICAC)*, Portsmouth, U.K., Sep. 2021, pp. 1–6.
- [59] X. Du, Y. Cai, S. Wang, and L. Zhang, "Overview of deep learning," in *Proc. 31st Youth Academic Annu. Conf. Chin. Assoc. Autom. (YAC)*, Wuhan, China, Nov. 2016, pp. 159–164.
- [60] J. Xie, C. Liu, Y.-C. Liang, and J. Fang, "Activity pattern aware spectrum sensing: A CNN-based deep learning approach," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1025–1028, Jun. 2019.
- [61] 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, p. 94, Jan. 2020.
- [62] S. Zhang, Y. Wang, J. Li, P. Wan, Y. Zhang, and N. Li, "A cooperative spectrum sensing method based on information geometry and fuzzy c-means clustering algorithm," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–12, Jan. 2019.
- [63] Y. Zhang, P. Wan, S. Zhang, Y. Wang, and N. Li, "A spectrum sensing method based on signal feature and clustering algorithm in cognitive wireless multimedia sensor networks," *Adv. Multimedia*, vol. 2017, pp. 1–10, Oct. 2017.
- [64] D. K. Patel, M. López-Benitez, B. Soni, and Á. F. García-Fernández, "Artificial neural network design for improved spectrum sensing in cognitive radio," *Wireless Netw.*, vol. 26, no. 8, pp. 6155–6174, Jul. 2020.
- [65] 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 Pers. Commun.*, vol. 118, no. 1, pp. 281–299, Jan. 2021.
- [66] M. Naraghi-Pour and T. Ikuma, "Autocorrelation-based spectrum sensing for cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 718–733, Feb. 2010.
- [67] Y. Zeng and Y.-c. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.
- [68] A. Nasser, A. Mansour, K. C. Yao, H. Abdallah, and H. Charara, "Spectrum sensing based on cumulative power spectral density," *EURASIP J. Adv. Signal Process.*, vol. 2017, no. 1, pp. 1–19, May 2017.
- [69] D. Teguig, V. Le Nir, and B. Scheers, "Spectrum sensing method based on likelihood ratio goodness-of-fit test," *Electron. Lett.*, vol. 51, no. 3, pp. 253–255, Feb. 2015.
- [70] MathWorks. *What is a Convolutional Neural Network? 3 Things You Need to Know*. Accessed: Mar. 15, 2022. [Online]. Available: <https://uk.mathworks.com/discovery/convolutional-neural-network-MATLAB.html>
- [71] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Antalya, Turkey, Aug. 2017, pp. 1–6.
- [72] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255.
- [73] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, Mar. 2021, Art. no. 53.
- [74] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.
- [75] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520.
- [76] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [77] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
- [78] W. M. Lees, A. Wunderlich, P. J. Jeavons, P. D. Hale, and M. R. Souryal, "Deep learning classification of 3.5-GHz band spectrograms with applications to spectrum sensing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 2, pp. 224–236, Jun. 2019.
- [79] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
- [80] Q. Peng, A. Gilman, N. Vasconcelos, P. C. Cosman, and L. B. Milstein, "Robust deep sensing through transfer learning in cognitive radio," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 38–41, Jan. 2020.
- [81] D. Han, G. C. Sobabe, C. Zhang, X. Bai, Z. Wang, S. Liu, and B. Guo, "Spectrum sensing for cognitive radio based on convolution neural network," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Shanghai, China, Oct. 2017, pp. 1–6.
- [82] V. Prithiviraj, B. Sarankumar, A. Kalaiyaran, P. P. Chandru, and N. N. Singh, "Cyclostationary analysis method of spectrum sensing for cognitive radio," in *Proc. 2nd Int. Conf. Wireless Commun., Veh. Technol., Inf. Theory Aerosp. Electron. Syst. Technol. (Wireless VITAE)*, Chennai, India, Feb. 2011, pp. 1–5.
- [83] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [84] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, "Deep CM-CNN for spectrum sensing in cognitive radio," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 10, pp. 2306–2321, Oct. 2019.
- [85] S. Haykin, C. Krasnor, T. J. Nohara, B. W. Currie, and D. Hamburger, "A coherent dual-polarized radar for studying the ocean environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 1, pp. 189–191, Jan. 1991.
- [86] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.
- [87] Y. Zeng, C. L. Koh, and Y.-C. Liang, "Maximum eigenvalue detection: Theory and application," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, 2008, pp. 4160–4164.
- [88] Y. Zeng, Y. Chang Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Process. Lett.*, vol. 15, pp. 649–652, 2008.
- [89] Y. Zeng and Y.-C. Liang, "Spectrum-sensing algorithms for cognitive radio based on statistical covariances," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1804–1815, May 2009.
- [90] P. Hale, J. Jargon, P. Jeavons, M. Lofquist, M. Souryal, and A. Wunderlich, "3.5 GHz radar waveform capture at Point Loma," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NIST TN 1954, 2017. [Online]. Available: <https://doi.org/10.6028/NIST.TN.1954>
- [91] P. Hale, J. Jargon, P. Jeavons, M. Lofquist, M. Souryal, and A. Wunderlich, "3.5 GHz radar waveform capture at fort story," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Note 1967, 2017. [Online]. Available: <https://doi.org/10.6028/NIST.TN.1967>

- [92] M. A. Richards, *Fundamentals of Radar Signal Processing*. Chicago, IL, USA: McGraw-Hill, 2005.
- [93] X. Liu, D. Yang, and A. E. Gamal, "Deep neural network architectures for modulation classification," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2017, pp. 915–919.
- [94] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.* Cham, Switzerland: Springer, 2016, pp. 213–226.
- [95] DeepSig Inc. *RF Datasets For Machine Learning*. Accessed: Jun. 8, 2022. [Online]. Available: <https://www.deepsig.ai/datasets>
- [96] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [97] T. J. Shea and N. West, "Radio machine learning dataset generation with GNU radio," in *Proc. GNU Radio Conf.*, vol. 1, no. 1, Sep. 2016. Accessed: Jul. 2, 2022. [Online]. Available: <https://pubs.gnuradio.org/index.php/grcon/article/view/11>
- [98] W. Lee, M. Kim, and D.-H. Cho, "Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3005–3009, Mar. 2019.
- [99] B. Soni, D. K. Patel, and M. López-Benítez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97437–97451, 2020.
- [100] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.
- [101] J. Gao, X. Yi, C. Zhong, X. Chen, and Z. Zhang, "Deep learning for spectrum sensing," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1727–1730, Dec. 2019.
- [102] G. Loye. *Long Short-Term Memory: From Zero to Hero With PyTorch*. Accessed: Mar. 22, 2022. [Online]. Available: <https://blog.floydhub.com/long-short-term-memory-from-zero-to-hero-with-pytorch/>
- [103] J. Xie, J. Fang, C. Liu, and X. Li, "Deep learning-based spectrum sensing in cognitive radio: A CNN-LSTM approach," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2196–2200, Oct. 2020.
- [104] Y. Saleem and M. H. Rehmani, "Primary radio user activity models for cognitive radio networks: A survey," *J. Netw. Comput. Appl.*, vol. 43, pp. 1–16, Aug. 2014.
- [105] T. Nguyen, B. L. Mark, and Y. Ephraim, "Spectrum sensing using a hidden bivariate Markov model," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4582–4591, Sep. 2013.
- [106] R. Zhang, T. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.
- [107] S. Subray, S. Tschimben, and K. Gifford, "Towards enhancing spectrum sensing: Signal classification using autoencoders," *IEEE Access*, vol. 9, pp. 82288–82299, 2021.
- [108] G. Liu, H. Bao, and B. Han, "A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis," *Math. Problems Eng.*, vol. 2018, pp. 1–10, Jul. 2018.
- [109] E. Axelsson and E. G. Larsson, "Optimal and sub-optimal spectrum sensing of OFDM signals in known and unknown noise variance," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 290–304, Feb. 2011.
- [110] W. Xu, W. Xiang, M. El-kashlan, and H. Melpouyan, "Spectrum sensing of OFDM signals in the presence of carrier frequency offset," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6798–6803, Aug. 2016.
- [111] J. Xie, J. Fang, C. Liu, and L. Yang, "Unsupervised deep spectrum sensing: A variational auto-encoder based approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5307–5319, May 2020.
- [112] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.
- [113] Q. Xu, Z. Wu, Y. Yang, and L. Zhang, "The difference learning of hidden layer between autoencoder and variational autoencoder," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, Chongqing, China, May 2017, pp. 4801–4804.
- [114] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel k-means: Spectral clustering and normalized cuts," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2004, pp. 551–556.
- [115] *Machine Learning Dataset for Radio Signal Classification*. Accessed: Jun. 12, 2022. [Online]. Available: <https://panoradio-sdr.de/radio-signal-classification-dataset/>
- [116] R. Utrilla. *MIGOU-MOD: A Dataset of Modulated Radio Signals Acquired With MIGOU, a Low-Power IoT Experimental Platform*. Accessed: Jun. 8, 2022. [Online]. Available: <https://data.mendeley.com/datasets/fkwr8mzndr/1>
- [117] R. Utrilla, E. Fonseca, A. Araujo, and L. A. Dasilva, "Gated recurrent unit neural networks for automatic modulation classification with resource-constrained end-devices," *IEEE Access*, vol. 8, pp. 112783–112794, 2020.
- [118] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, Dec. 2022, Art. no. 100258.
- [119] L. Li, Z. Zhang, and L. Yang, "Influence of autoencoder-based data augmentation on deep learning-based wireless communication," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2090–2093, Sep. 2021.
- [120] C. d. Vrieze, L. Simic, and P. Mahonen, "The importance of being earnest: Performance of modulation classification for real RF signals," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Seoul, South Korea, Oct. 2018, pp. 1–5.



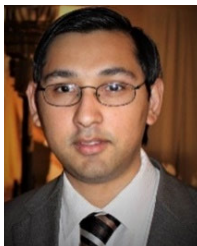
SADAF NAZNEEN SYED (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering and the M.Tech. degree in digital communication from the Maulana Azad National Institute of Technology (MANIT), India, in June 2018 and June 2021, respectively. She is currently pursuing the Ph.D. degree with the University of Huddersfield. Her M.Tech. thesis was based on breast cancer classification from histopathology images using transfer learning and deep neural networks. She was a System on Chip Verification Intern with Intel Corporation, Bengaluru, from August 2020 to July 2021, where she performed data analysis using machine learning techniques and got familiar with formal verification. She is an MSCA Early Stage Researcher for the EU H2020 MOTOR5G Consortium. Her research interests include the application of deep learning techniques in wireless communication and healthcare.



PAVLOS I. LAZARIDIS (Senior Member, IEEE) received the Diploma degree in electrical engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1990, the M.Sc. degree in electronics from Université Pierre and Marie Curie (Paris 6), Paris, France, in 1992, and the Ph.D. degree from École Nationale Supérieure des Télécommunications (ENST) Paris and Université Paris 6, in 1996. From 1991 to 1996, he was involved in research with France Télécom and teaching with ENST Paris. In 1997, he became the Head of the Antennas and Propagation Laboratory, Télédiffusion de France/the France Télécom Research Center (TDF-C2R Metz). From 1998 to 2002, he was a Senior Examiner with the European Patent Office (EPO), The Hague, The Netherlands. From 2002 to 2014, he was involved in teaching and research with the ATEI of Thessaloniki, Thessaloniki, and Brunel University, London, U.K. He is currently a Professor of electronics and telecommunications with the University of Huddersfield, U.K. He has been involved as a Principal Investigator in several international research projects, such as EU Horizon 2020 MOTOR5G and RECOMBINE and NATO-SIP ORCA. He has published over 200 research articles and several national and European patents. He is a member of IET (MIET), a Senior Member of URSI, and a fellow of the Higher Education Academy (FHEA). He is serving as an Associate Editor for IEEE Access.



FAHEEM A. KHAN (Member, IEEE) received the bachelor's degree in electronics engineering from Aligarh Muslim University, India, the master's degree in communication and radar engineering from the Indian Institute of Technology Delhi, India, and the Ph.D. degree in electrical and electronic engineering from Queen's University Belfast, in 2012. He is currently a Senior Lecturer in electronic engineering with the School of Computing and Engineering, University of Huddersfield. He has more than 15 years of teaching and research experience at academic and research institutions in the U.K., Sultanate of Oman, United Arab Emirates, and India. Before joining as a Lecturer with the University of Huddersfield, in 2016, he was a Research Associate in wireless communications and signal processing with the Institute for Digital Communications, The University of Edinburgh, where he contributed to research and project management in several EU/EPSCRC research projects. His research is funded by the following research grants: Proof of Commercial Concept (PoCC) research grant on "non-orthogonal multiple access techniques in beyond 5G networks," EU H2020 ETN Research Project MOTOR5G, and EU H2020 RISE Research Project. His current research interests include spectrum sharing/spectrum management for beyond 5G wireless networks, licensed shared access and cognitive radio, non-orthogonal multiple access, machine learning and deep learning in wireless communications, full-duplex communications, MIMO and millimeter wave communications, and the Internet of Things (IoT) networks. He is a fellow of HEA.



QASIM ZEESHAN AHMED (Member, IEEE) received the Ph.D. degree from the University of Southampton, Southampton, U.K., in 2009. He was an Assistant Professor with the National University of Computer and Emerging Sciences (NUCES-FAST), Islamabad, Pakistan, from November 2009 to June 2011. He was a Postdoctoral Fellow with the Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, from June 2011 to June 2014. He joined the University of Kent, U.K., as a Lecturer, from January 2015 to January 2017. He was a Lecturer, then a Senior Lecturer, and currently a Reader of electronic engineering with the School of Computing and Engineering, University of Huddersfield, U.K., in 2017, 2018, and 2020, respectively. He is also a Principal Investigator of the U.K. for Erasmus + DigiHealth-Asia Project and the MSCA Staff Exchanges EVOLVE Project. He is a Co-Investigator of the EU H2020 ETN Research MOTOR5G Project and the EU H2020 RISE Research RECOMBINE Project. His research interests include ultrawide bandwidth systems, millimeter waves, device-to-device, digital health, and cooperative communications. He is a fellow of HEA.



MARYAM HAFEEZ (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Leeds, U.K., in 2015. From 2015 to 2018, she was a Research Fellow with the Institute of Robotics, Autonomous Systems and Sensing (IRASS), University of Leeds. Since 2018, she has been a Senior Lecturer with the Department of Engineering and Technology, University of Huddersfield. Her current research is funded by the EU Horizon 2020 Program. Her research interests include the design and analysis of protocols for next-generation green intelligent wireless networks by employing tools from game theory and stochastic geometry along with the Internet of Things (IoT) and Industry 4.0 related research. She is also serving as a member of the Editorial Board for *Frontiers in Communications and Networks*. She has worked in the area of dynamic spectrum access had received the Best Paper Award from the IEEE International Conference on Communications (ICC), in 2013.



ANTONI IVANOV (Member, IEEE) received the master's degree in innovative communication technologies and entrepreneurship from the Technical University of Sofia (TUS), Bulgaria, and Aalborg University, Denmark, in 2016, and the Ph.D. degree in communication networks and systems from TUS, in 2020. He is currently a Postdoctoral Researcher with the Teleinfrastructure Laboratory, Faculty of Telecommunications, TUS. His research interests include cognitive radio networks, adaptive algorithms for dynamic spectrum access, deep learning-based solutions for cognitive radio applications, radio frequency data analytics, volumetric spectrum occupancy assessment, and graph signal processing for resource allocation in current and future wireless networks.



VLADIMIR POULKOV (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees from the Technical University of Sofia (TUS), Sofia, Bulgaria. He has been the Dean of the Faculty of Telecommunications, TUS, and the Vice Chairman of the General Assembly of the European Telecommunications Standards Institute (ETSI). He is currently a TUS Professor with more than 30 years of teaching, research, and industrial experience, managing numerous educational, research and development, and engineering projects in the field of telecommunications. He is also the Head of the Teleinfrastructure Research and Development Laboratory, TUS; the Head of the Intelligent Communication Infrastructures Research and Development Laboratory, Sofia Tech Park; and the Chairperson of the Cluster for Digital Transformation and Innovation, Bulgaria. He is the author of many scientific publications, tutoring B.Sc., M.Sc., and Ph.D. courses in the field of information transmission theory and wireless access networks. He is a fellow of the European Alliance for Innovation (EAI).



ZAHARIAS D. ZAHARIS (Senior Member, IEEE) received the B.Sc. degree in physics, the M.Sc. degree in electronics, the Ph.D. degree in antennas and propagation modeling for mobile communications, and the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1987, 1994, 2000, and 2011, respectively. From 2002 to 2013, he was with the Administration of the Telecommunications Network, Aristotle University of Thessaloniki. Since 2013, he has been with the School of Electrical and Computer Engineering, Aristotle University of Thessaloniki. He has been involved in several international research projects, such as EU Horizon 2020 MOTOR5G and RECOMBINE. He is the author of 78 scientific journal articles, 62 international conference papers, one national patent, five book chapters, and one textbook. His current research interests include the design and optimization of antennas and microwave circuits, signal processing on smart antennas, development of evolutionary optimization algorithms, and neural networks. He is a member of the Technical Chamber of Greece. Recently, he was elected as the Chair of the Electron Devices/Microwave Theory and Techniques/Antennas and Propagation Joint Chapter of the IEEE Greece Section. He is also serving as an Associate Editor for IEEE ACCESS.

...