Spectrum Sensing and Signal Identification With Deep Learning Based on Spectral Correlation Function

Kürşat Tekbıyık, *Graduate Student Member, IEEE*, Özkan Akbunar, *Student Member, IEEE*, Ali Rıza Ekti, *Senior Member, IEEE*, Ali Görçin, *Senior Member, IEEE*, Güneş Karabulut Kurt, *Senior Member, IEEE*, and Khalid A. Qaraqe, *Senior Member, IEEE*

***Abstract*—Spectrum sensing is one of the means of utilizing the scarce source of wireless spectrum efficiently. In this paper, a convolutional neural network (CNN) model employing spectral correlation function (SCF) which is an effective characterization of cyclostationarity property, is proposed for wireless spectrum sensing and signal identification. The proposed method classifies wireless signals without a priori information and it is implemented in two different settings entitled `CASE1` and `CASE2`. In `CASE1`, signals are jointly sensed and classified. In `CASE2`, sensing and classification are conducted in a sequential manner. In contrary to the classical spectrum sensing techniques, the proposed CNN method does not require a statistical decision process and does not need to know the distinct features of signals beforehand. Implementation of the method on the measured over-the-air real-world signals in cellular bands indicates important performance gains when compared to the signal classifying deep learning networks available in the literature and against classical sensing methods. Even though the implementation herein is over cellular signals, the proposed approach can be extended to the detection and classification of any signal that exhibits cyclostationary features. Finally, the measurement-based dataset which is utilized to validate the method is shared for the purposes of reproduction of the results and further research and development.**

***Index Terms*—Deep learning, spectrum sensing, cyclostationarity, signal classification, spectral correlation function, convolutional neural networks.**

I. Introduction

TODAY'S wireless communication systems have to bear an unprecedented increase in data transmission volume. It is essential for wireless communication networks to utilize the limited source of spectrum as efficiently and effectively as possible to meet the demand [1]. Furthermore, the efforts including the deployment of small cells, utilizing millimeter wave (mmWave) bands, effective spectrum usage algorithms, massive multiple-input multiple-output (MIMO) systems [2], and cognitive radio networks target the same goal. Cognitive radios aim to attend this purpose by sharing the spectrum dynamically among users; thus, spectrum sensing and signal identification became major techniques for cognitive radio networks.Considering joint communications, sensing, and localization demanded by 6G and beyond, efficient spectrum allocation will be more crucial for heterogeneous networks. For instance, 5G NR Release 16 introduces dynamic spectrum sharing which is novel method enabling parallel operation of 5G and Long-Term Evolution (LTE) in the same band [3]. Furthermore, it is envisioned that radar and communications system will share the same frequency band [4].

Moreover, today's vehicular communications unprecedentedly evolve and develop. Thus, the spectrum band allocated for vehicular networks gains increasing importance. Although vehicular communications is separately building on dedicated short-range communications (DSRC) and cellular communications, it is foreseen that numerous upcoming vehicular applications in dense vehicle networks will pave the way for joint utilization of DSRC and cellular networks in the near future [5], [6]. Recently, opportunistic utilization of spectrum for DSRC and cellular communications has been proposed [7]. In this regard, spectrum awareness and intelligent spectrum management would be essential in vehicular networks.

Manuscript received May 1, 2021; revised July 22, 2021 and August 19, 2021; accepted August 23, 2021. Date of publication September 2, 2021; date of current version October 15, 2021. This work was supported in part by the European Commission Horizon 2020 Research and Innovation Programme BEYOND5 under Grant Agreement 876124. This study was supported in part by NPRP12S-0225-190152 from the Qatar National Research Fund (a Member of The Qatar Foundation). The review of this article was coordinated by Prof. Khoa Le. *(Corresponding author: Kürşat Tekbıyık.)*

Kürşat Tekbıyık is with the Informatics and Information Security Research Center (BİLGEM), TÜBİTAK, Kocaeli 41470, Turkey, and also with the Department of Electronics and Communications Engineering, İstanbul Technical University, İstanbul, İstanbul 34467, Turkey (e-mail: tekbiyik@itu.edu.tr).

Özkan Akbunar is with the Informatics and Information Security Research Center (BİLGEM), TÜBİTAK, Gebze, Kocaeli 41470, Turkey, and also with the Electrical-Electronics Engineering, Hacettepe University Ankara, Ankara 06800, Turkey (e-mail: ozkan.akbunar@tubitak.gov.tr).

Ali Rıza Ekti is with the Informatics and Information Security Research Center (BİLGEM), TÜBİTAK, Kocaeli 41470, Turkey, and also with the Department of Electrical-Electronics Engineering, Balıkesir University, Balıkesir 10145, Turkey (e-mail: arekti@balikesir.edu.tr).

Ali Görçin is with the Informatics and Information Security Research Center (BİLGEM), TÜBİTAK, Kocaeli 41470, Turkey, and also with the Department of Electronics and Communications Engineering, Yıldız, İstanbul 34220, Turkey (e-mail: agorcin@yildiz.edu.tr).

Güneş Karabulut Kurt was with the Department of Electronics and Communications Engineering, İstanbul Technical University, İstanbul, Turkey. She is now with the Poly-Grames Research Center, Department of Electrical Engineering, Polytechnique Montréal, Montréal H3T 1J4, Canada (e-mail: gunes.kurt@polymtl.ca).

Khalid A. Qaraqe is with the Department of Electrical and Computer Engineering, Texas A&M University at Qatar, Doha, Qatar (e-mail: khalid.qaraqe@qatar.tamu.edu).

Digital Object Identifier 10.1109/TVT.2021.3109236

When spectrum sensing and signal identification techniques are considered, it is seen that sensing techniques of energy detection and matched filtering require a priori information such as number of second order noise statistics, cyclic frequencies and particular pulse shaping filter characteristics to operate. Moreover, after processing of the received signals, a statistical decision mechanism should be implemented to complete the sensing process [8]. Such cumbersome process can hamper the agile decision making requirements of 5G and beyond networks, thus, classical sensing paradigm can not satisfy the requirements of fast changing operation environment of contemporary and future wireless communications networks. In this context, deep learning (DL) has been proposed as a solution to the parameter adaptation issues of classical techniques. This stems from the known ability of DL techniques in extracting the intrinsic features of given inputs through a convolutional process. The use of DL based approaches also eliminates the need for a statistical decision mechanism at the end of the identification process. Along this line, the recent study shows that DL methods outperform classical approaches in signal detection in the spectrum [9]. To achieve the requirements for 5G and beyond wireless networks, an intelligent radio design for spectrum sensing and signal identification is required and such solution can be realized with the help of machine learning (ML) algorithms [10] utilizing features such cyclostationarity of signals [11].

A. Related Works

When the literature on the implementation of artificial intelligence techniques for spectrum sensing and signal identification purposes are considered, it is initially seen that CNNs are trained with high-order statistics of single carrier signals for modulation classification [12]. The CNN classifier is used for modulation and interference identification for industrial scientific medical (ISM) bands by utilizing fast Fourier transform (FFT), amplitude-phase representation (AP) and in-phase/quadrature (I/Q) features for training [13]. Moreover, a covariance matrix based CNN has been proposed for spectrum sensing in [14]. This approach utilizes sample covariance matrix as the input of CNN to further improve spectrum sensing performance. In [15], covariance matrix based transfer learning is also proposed for signal detection for ambient backscatter communications. Also, CNNs have been utilized for spectrum sensing applications under α-stable noise and for real-time hardware platforms in [16] and [17], respectively. Another study [18] focused on the protocol classification in ISM band by utilizing fully connected neural networks. As another example of the application of DL to signal classification, long short term memory (LSTM) is deployed for modulation classification and identification of digital video broadcast (DVB), Tetra, LTE, Global System for Mobile communications (GSM), wide-band FM (WFM) signals by using AP and FFT magnitude for training [19]. The performance of the proposed model is high, however, it employs synthetic data generated from MATLAB. In the real channels, there are numerous phenomenons, which further complicate the signal characteristics.

On the other hand, cyclostationarity signal analysis has been explored for modulation classification, parameter estimation and spectrum sensing for more than 20 years. In addition to being an established method for spectrum sensing in cognitive radio domain, cyclostationary features detection (CFD) is also utilized to distinguish generic modulations such as M-PSK, M-FSK, and M-QAM [11], [20]. When the radio access technology (RAT) identification [21] is considered, second order cyclostationarity is employed for classification of LTE and GSM signals [22]. Later, a tree-based classification approach is proposed to identify GSM, cdma2000, universal mobile telecommunications system (UMTS) and LTE signals [23].

B. The Contributions

1) Novelty in Terms of Numerical Studies: In terms of performance analysis, first, a comparative analysis is conducted and superiority of SCF over the features of I/Q, AP and FFT is shown for the purpose of training of DL networks. Second, comparison with the existing DL methods such as convolutional long short term memory fully connected deep neural network (CLDNN) [24], LSTM [19], DenseNet [25], ResNet [12] are given in terms of accuracy, memory consumption and computational complexity. Third, it is shown that the proposed method outperforms support vector machines (SVMs) trained with SCF, which is our previous study. Fourth, the performance of the proposed method is compared with the classical spectrum sensing technique of CFD, which requires the cyclic frequencies as a priori information. The identification results indicate important performance improvements over the aforementioned techniques.

2) Methodological Novelty: CFD depends on extracting the underlying features using likelihood-based techniques utilizing statistical decision mechanisms and for CFD to operate under the dynamically changing communication medium, an additional mechanism to adaptively adjust decision parameters such as thresholds and the number samples is required [26]. On the other hand, even though employment of DL techniques for the purposes of spectrum sensing and signal identification implies considerable advantages in terms of performance and complexity, utilization of FFT, AP and I/Q as input features to the intelligent networks do not lead stable and dependable results due to the rapidly and significantly changing wireless communications medium between the nodes. Therefore, *this study proposes application of SCF as input feature to CNNs for blind wireless signal identification*. The problems of spectrum sensing and signal identification are framed into two particular contexts which utilize a novel methodology based on CNN and SCF of wireless signals without bi-frequency mapping. Therefore, the proposed method can be employed either to decide whether the signal is present or not in the spectrum or to distinguish signals from each other. Sensing and identification performance of the method is tested and validated utilizing real-life over-the-air signal measurements of GSM, UMTS, and LTE signals.

The proposed method approaches to the problems of sensing and identification from the aspects of two cases; in `CASE1`, the

designed CNN model is fed directly with the SCFs of measurements of GSM, UMTS, LTE along with SCF of spectrum which is only comprised of noise. Sensing and classification are executed jointly for CASE1. On the other hand in CASE2, a two-step approach is adopted; first, as a spectrum sensing method to measure the spectrum occupancy is conducted and this stage is followed by a signal classification procedure.

3) Novelty in Terms of Experimental Activities: Focusing on the valuable information in the dataset is an important metric for the proposed method; thus, it is denoted that utilizing only the meaningful part of the input matrices improves the classification performance along with alleviation in training time and complexity. On the other hand, the general dataset, which has been developed from measurements taken through a comprehensive measurement campaign conducted in different locations and frequency bands, is shared publicly in [27]. Therefore, the measurement-based dataset is open to researchers as a comprehensive resource in the development and validation of their work.

4) Applicability for Future Research Problems: Even though in this work the scope of implementation is focused on cellular signals, the introduced identification system can be directly used for detection and classification of any signal that exhibit cyclostationary features. All analyses are based on the real-world measurements taken during a measurement campaign conducted at different locations with varying environmental conditions such as channel fading statistics and signal-to-noise ratio (SNR) levels. Finally, the measurement data that this work is experimented on is also shared for reproducibility of this work and to support future research and development activities in this domain.

5) Tutorial on DL-Based Cognitive Communications: Besides all the contributions above, this study provides a comprehensive tutorial for practical applications of deep learning-based cognitive communications. In this study, each step from feature extraction to points to be considered in the training process is handled by considering hardware constraints and optional approaches are discussed. For this reason, this study acts as a guide for designers as well as researchers.

C. Organization of the Paper

The rest of the paper is structured as follows. Background information on the system model, cyclostationary analysis and CNNs is presented in Section II. The problem statement is given in Section III. The proposed CNN model is described in Section IV. The details of the measurement setup and dataset utilized in this study are given in Section V. Section VI presents measurement results and details the classification performance of the proposed method. The concluding remarks are provided in Section VII.

II. Background

Assuming that received signal is down converted to baseband before further processing, first the complex baseband equivalent of the received signal, $r(t)$ should be defined. When the presence of fading environment with thermal noise, received signal can be given as

$$r(t) = \rho(t) * x(t) + \omega(t), \tag{1}$$

where $\omega(t)$ denotes the complex additive white Gaussian noise (AWGN) with $\mathcal{CN}(0, \sigma_N^2)$ in the form of $\omega(t) = \omega_I(t) + j\omega_Q(t)$ as both $\omega_I(t)$ and $\omega_Q(t)$ being $\mathcal{N}(0, \sigma_N^2/2)$ and $j = \sqrt{-1}$; the complex baseband equivalent of the transmitted signals is denoted as $x(t)$; and $\rho(t)$ stands for the impulse response for the time-invariant wireless channel because of extremely short observation time for a signal.

Depending on the idle or busy state of the mobile propagation channel in the radio frequency (RF) spectrum, the signal detection process of deep learning methods can be modelled as a binary hypothesis test

$$r(t) = \begin{cases} \rho(t)x(t) + \omega(t), & H_1 \\ \omega(t), & H_0. \end{cases} \tag{2}$$

H_0 and H_1 hypotheses stand for the presence of noise only and the unknown signal, respectively. Therefore, the problem statement can be stated as identification of the presence of the unknown signal, $x(t)$, and classification of the $x(t)$.

A. Cyclostationarity

Cyclostationary signal processing leads to extracting hidden periodicities in a received signal, $r(t)$. Since these periodicities (*e.g.*, symbol periods, spreading codes, and guard intervals) exhibit unique characteristics for different signals, they provide the necessary information for identification. Thus, the unknown signals $x(t)$ can be identified by using cyclostationary features to obtain the statistical characteristics of $r(t)$ in the presence of $\omega(t)$ and multipath fading without a priori information. A nonlinear transformation, second-order cyclostationarity of a signal can be expressed as

$$s_\tau(t) = \mathbb{E}\left\{r(t + \tau/2)r^*(t - \tau/2)\right\}, \tag{3}$$

where $\mathrm{s}_\tau(t)$ is the autocorrelation of $r(t)$. Assuming that the autocorrelation function is periodic with T_0 for second-order cyclostationary signals, a Fourier series expansion of $s_\tau(t)$ is given as

$$\mathbb{R}_r^\alpha(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} s_\tau(t)e^{-j2\pi\alpha t}dt, \tag{4}$$

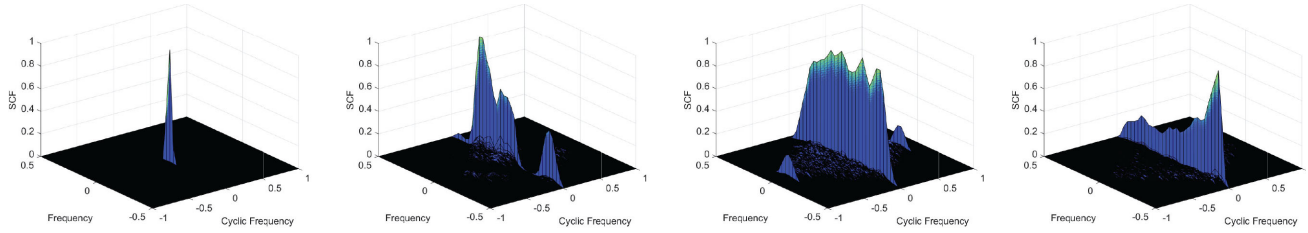
where $\mathbb{R}_r^\alpha(\tau)$ is the cyclic autocorrelation function (CAF) and α values denote the cyclic frequencies.

The Fourier transform of the CAF for a fixed α is given with the cyclic Wiener relation [11]

$$\mathbb{S}_r(f) = \int_{-T/2}^{T/2} \mathbb{R}_r^\alpha(\tau)\, e^{-j2\pi f\tau}d\tau, \tag{5}$$

where $\mathbb{S}_r(f)$ is called as SCF which is equal to the power spectral density (PSD) when α is zero.

The computational complexity of calculating SCF is relatively high. However, this complexity can be decreased by using the FAM based on time smoothing via FFT [28]. FAM estimates the



(a) (b) (c) (d)

Fig. 1. Fast Fourier transform accumulation method (FAM) based SCF estimates of cellular signals in bi-frequency plane. It is easily observed that the signals show different cyclic characteristics. The noise does not show cyclic characteristics as SCF of noise gives only peak at the center of bi-frequency plane where the cyclic frequency is zero.

SCF as

$$\mathbb{S}_{r_T} = \sum_k R_T(kL, f) R_T^*(kL, f) g_c(n-k) e^{-i2\pi kq/P}, \quad (6)$$

where $R_T(n, f)$ denotes the complex demodulates which is the N'-point FFT of $r(n)$ passed through a Hamming window and can be computed by

$$R_T(n, f) = \sum_{k=-N'/2}^{N'/2} a(k) r(n-k) e^{-i2\pi f(n-k)T_s}, \quad (7)$$

where $a(n)$ and $g_c(n)$ are both data tapering windows. The symbols N', T_s, and L denote the channelization length, sampling period, and sample size of hopping blocks, respectively. The ratio between the number of total samples and L is employed as the length of second FFT, whose length is denoted as P. The FAM has six implementation steps. These steps are respectively channelization, windowing, N'-point FFT, complex multiplication, P-point FFT and bi-frequency mapping. In the study, the unit rectangle and Hamming windows are employed as $g_c(n)$ and $a(n)$, respectively. Fig. 1 illustrates SCFs results in bi-frequency plane, which are estimated by FAM algorithm for GSM, UMTS, and LTE along with the noise. Consequently, the input matrix, $\mathbf{X}_k^{SCF}$, to be fed into classifier model is given as

$$\mathbf{X}_k^{SCF} = |\mathbb{S}_{r_T}(nL, f)|. \quad (8)$$

As seen from Fig. 1, the SCF of noise creates a peak at DC frequency because of the lack of cyclostationary characteristics in AWGN signal. Due to midambles and bursty structure of GSM frames, SCF generates unique peaks as depicted in Fig. 1(b). SCF of UMTS creates specific peaks at 3.84 MHz owing to the fact that the spreading factor of wideband code division multiple access (WCDMA) is 3.84 Mcps. Since the frame duration of LTE signals is 10 ms, the peaks at 100 Hz as given in Fig. 1(d).

B. Amplitude-Phase

The amplitude and phase values of time-domain I/Q data can be used to establish a real-valued classification feature matrix, $\mathbf{X}_k^{AP}$. This feature matrix is composed of the amplitude and phase vectors of the received signal samples. So, $\mathbf{X}_k^{AP}$ is defined as

$$\mathbf{X}_k^{AP} = \begin{bmatrix} \mathbf{x}_A^T \\ \mathbf{x}_\phi^T \end{bmatrix}, \quad (9)$$

where $\mathbf{x}_A = ({r_q}^2 + {r_i}^2)^{\frac{1}{2}}$ and $\mathbf{x}_\phi = \arctan(\frac{r_q}{r_i})$ denote the amplitude and phase vectors, respectively.

C. Fast Fourier Transform

The characteristics of signals in frequency domain can be employed as discriminating classification features. The FFT of the received signal is used to obtain a real-valued classification feature matrix $\mathbf{X}_k^{FFT}$ as

$$\mathbf{f} = \mathcal{F}(r),\ \mathbf{X}_k^{FFT} = \begin{bmatrix} \mathbf{f}_{re}^T \\ \mathbf{f}_{im}^T \end{bmatrix}, \quad (10)$$

where $\mathcal{F}(\cdot)$ stands for the FFT of the received signals; $\mathbf{f}_{re}$ and $\mathbf{f}_{im}$ are real and imaginary parts of $\mathbf{f}$, respectively.

D. The Convolutional Neural Networks

CNN is a class of deep neural networks which is mainly employed in image classification and recognition. CNN process inputs like a visual system in human. In other words, it extracts features in an input rather than fitting data [29]. In this study, we utilize the input matrices which resemble image consisting of features in a specific positions as seen in Fig. 1. Still, it has been recently extended to several application areas. CNNs have two stages: feature extraction and classification. In feature extraction, a convolutional layer is followed by a pooling layer. In the convolution layer, the feature matrix is convolved with different filters to obtain convolved feature map as follows

$$\mathrm{h}[i, j] = \sum_{p=1}^{m} \sum_{l=1}^{n} w_{p,l} \mathbf{X}_\mathbf{k}[i+p-1, j+l-1], \quad (11)$$

where $w_{p,l}$ is the element at p-th row and l-th column of the $m \times n$ filter matrix, and $\mathbf{X}_\mathbf{k}[\cdot, \cdot]$ denotes the elements of feature matrix convolved by $w_{p,l}$. The convolution layer is followed by the pooling layer to reduce computational complexity and training time, and control over-fitting due to the fact that pooling layer makes the activation less sensitive to feature locations [30]. The $u \times v$ maximum pooling operation is described as

$$\mathrm{g}[i, j] = \max\{h[i+a-1, j+b-1]\}, \quad (12)$$

where $1 \le a \le u$ and $1 \le b \le v$. The output of the pooling layer is a 3-D tensor. This output is then reshaped into a 1-D vector. This vector is fed to the dense (fully-connected) layers for the final classification decision.

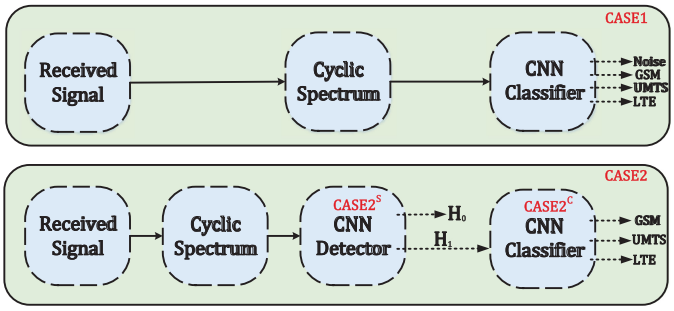


Fig. 2. Two different approaches for the sensing and classification of signals. In CASE1, signal sensing and signal classification are jointly conducted. However, CASE2 firstly sense signal in the spectrum, then classify.

III. Problem Statement

The dynamic communications environment of next generation wireless networks require fast, robust and adaptive sensing and identification of the multi-dimensional communications medium to utilize the resources quickly and efficiently [8]. In this context, spectrum sensing and signal identification becomes important means of achieving effective resource utilization. To that end, we approach the problems of sensing and identification via DL from two aspects:

CASE1: In this case, the designed CNN classifier is trained with all possible classes, in this case GSM, UMTS, LTE and empty spectrum which can be referred to as AWGN only. For each signal the cyclic spectrum is constructed based on the procedures described in Section II-A. The cyclic spectrum is then fed to the CNN classifier, which is trained with four possible inputs beforehand. Finally, the classification is made.

CASE2: In this case a two-stage approach is adopted; at the first stage a CNN detector (the same CNN model defined is employed for both detection and classification for the sake of simplicity) is utilized to decide whether a signal exists in the given band or not by training the CNN by two classes, first comprised of GSM, UMTS, and LTE signals and second part with AWGN only. Thus, in the first stage a decision is made about whether a signal exists in the spectrum or not as in the case of classical spectrum sensing. If the decision is made that there is an information bearing signal in the given band, second stage is activated utilizing a CNN classifier, which is trained in our case with three classes (*i.e.*, GSM, UMTS, and LTE) and finally a decision is made for the class of the signal occupying the spectrum.

Please note that the classification refers to identification of the signals, and at the detection part of the approach H_1 and H_0 refers to the existence and non-existence of a signal over the spectrum based on binary hypothesis testing. Both CASE1 and CASE2 are illustrated in Fig. 2. By evoking the second strategy, it is possible to differentiate spectrum occupancy and then determine whatever the signal is. By doing so, the system focuses on signal existence rather than its type at low SNR levels, especially. Moreover, the classification accuracy can be improved by utilizing the intermediary noise cancellation step between sensing and classification parts. It should be noted that this study does not employ an external noise cancellation method.

As known, some operation fields can be under impact of high noise and interference. In that case, direct signal identification cannot be accurate due to distortion in feature vectors. CASE2 is designed to disjointly sense and classify signals in the spectrum. However, CASE1 provides joint sensing and identification for signals in the spectrum.

If the signals to be detected and identified are in the low SNR region, using CASE2 increases its performance at the expense of increasing computational complexity and memory usage. On the other hand, CASE1 provides sensing and signal classification with low complexity and moderate performance.

Firstly, we can define the accuracy for CASE1, $P_{\texttt{CASE1}}$ as:

$$P_{\texttt{CASE1}} = \sum_{k=0}^{3} P(\hat{\boldsymbol{\chi}}_k|\boldsymbol{\chi}_k)P(\boldsymbol{\chi}_k), \tag{13}$$

where $\boldsymbol{\chi}_k$ denotes the label array of the transmitted signals and k represents the label of the classes AWGN, GSM, UMTS, and LTE, respectively. $\hat{\boldsymbol{\chi}}_k$ is array for the predicted classes of the received signals. In a short, $P_{\texttt{CASE1}}$ stands for the accuracy of four-classes classification problem. For CASE2, it is required to define two independent accuracy functions: the sensing accuracy, $P^{\texttt{S}}_{\texttt{CASE2}}$ and the classification accuracy, $P^{\texttt{C}}_{\texttt{CASE2}}$, which are defined as

$$P^{\texttt{S}}_{\texttt{CASE2}} = P(\hat{\boldsymbol{\chi}}_{\texttt{S}} = 1|H_1) + P(\hat{\boldsymbol{\chi}}_{\texttt{S}} = 0|H_0), \tag{14}$$

$$P^{\texttt{C}}_{\texttt{CASE2}} = \sum_{k=1}^{3} P(\hat{\boldsymbol{\chi}}_k|\boldsymbol{\chi}_k, H_1)P(\boldsymbol{\chi}_k). \tag{15}$$

$\hat{\boldsymbol{\chi}}_{\texttt{S}}$ is the prediction of $\boldsymbol{\chi}_{\texttt{S}}$ regarding to the presence of a signal in the spectrum. $\boldsymbol{\chi}_k$ stands for the predictions for the classification part of CASE2. $\boldsymbol{\chi}_{\texttt{S}}$ is defined for the transmitted signal as:

$$\boldsymbol{\chi}_{\texttt{S}} = \begin{cases} 0, & k = 0, \\ 1, k = 1, 2, 3. \end{cases} \tag{16}$$

The overall accuracy for CASE2 can be introduced in terms of $P^{\texttt{S}}_{\texttt{CASE2}}$ and $P^{\texttt{C}}_{\texttt{CASE2}}$ by

$$P_{\texttt{CASE2}} = P(\hat{\boldsymbol{\chi}}_{\texttt{S}} = 1|H_1)P^{\texttt{C}}_{\texttt{CASE2}}. \tag{17}$$

IV. The Proposed CNN Model

As indicated in Section III, the proposed method relies on a CNN model, which is designed for the problem defined in this study. Design and implementation of CNN for classification of wireless mobile communication signals is conducted via an open source machine learning library, Keras [31]. The proposed CNN model consists of three convolution and three pooling layers sequentially. The convolution layers have respectively 64, 128, and 64 filters. The network is terminated by two fully connected layers. First hidden layer includes 256 neurons. Second hidden layer consists of 4 and 3 neurons for CASE1 and CASE2, respectively. The leaky rectified linear unit (ReLU) activation function with an alpha value 0.1 is used in each convolution layer to extract discriminating features. Leaky ReLU is selected

TABLE I
THE PROPOSED CNN LAYOUT

Layer	Output Dimensions
Input	8193×16
Conv1	$8193 \times 16 \times 64$
Leaky ReLU1	$8193 \times 16 \times 64$
Max_Pool1	$4097 \times 8 \times 64$
Conv2	$4097 \times 8 \times 128$
Leaky ReLU2	$4097 \times 8 \times 128$
Max_Pool2	$2049 \times 4 \times 128$
Conv3	$2049 \times 4 \times 64$
Leaky ReLU3	$2049 \times 4 \times 64$
Max_Pool3	$1025 \times 2 \times 64$
Flatten	131200
Dense1	256
Dense2	4
Trainable Par.	$33,736,772$

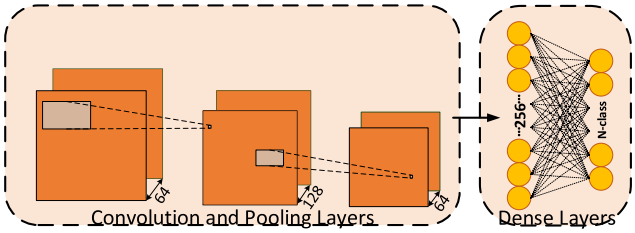


Fig. 3. The proposed CNN model consists of three convolutional layers and two dense layers with Adam optimizer with learning rate of 10^{-5}.

instead of ReLU. Unlike ReLU, leaky ReLU maps larger negative values to smaller ones by a mapping line with a small slope. In each convolution layer, 3×3 filters are used. 2×2 max pooling is used to reduce the dimension and training time. A fully connected layer is formed by 256 neurons and Leaky ReLU activation function. Following the fully connected layers, the probabilities for each class are computed by the softmax activation function. In addition, the adaptive moment estimation (ADAM) optimizer is utilized when determining the model parameters. In the training phase, early stopping is employed to prevent the model from over-fitting. The patience is chosen as 10 epochs for early stopping function and validation loss is monitored during the training. If the validation loss converges a level and remain at this level during 10 epochs, the training is terminated and the weights at the end of training are used in the test. The implementation layout for the proposed CNN model is given in Table I. The input matrices, $\mathbf{X}_k^{AP}$, $\mathbf{X}_k^{FFT}$, and $\mathbf{X}_k^{SCF}$ are used at the beginning of the proposed model by convolving with filters. The overall block diagram for the proposed CNN model is depicted in Fig. 3.

When the motivation behind designing such a CNN model is considered, it should be noted firstly that the information about changes in the local regions of the mapped output is extracted by using $3 \times 3 \times 64$ filters in the first convolution layer. In this problem, because the SCF creates local differences in frequency and cyclic frequency regions, the smaller filter size is preferred to catch peaks in the feature matrices. Thus, local differences are taken into account along the layers. After the first layer determines the cyclic characteristics of all local terms as a general process, the second layer examines the properties such as location and size related to these characteristics. Here, it is aimed to deal with cyclic features in detail by increasing the number of filters to 128. In the last layer, all properties are converted to an average of all information gathered and eventually sent to the decisive layer which is dense layer. For this reason, the number of filters in the last layer should be chosen so that sufficient information is obtained without overfitting. Therefore, the number of filters is selected as 64 in the last layer. It is customary to quantify the performance of a classifier model in terms of the precision (Π), recall (Ψ), and F_1-score performance metrics. The precision metric quantifies how much positive results are actually positive, the recall provides information on how much true positives are identified correctly as positive, and F_1-score gives an overall measure for the accuracy of a classifier model since it is the harmonic average of precision and recall. These metrics are given as

$$\Pi = \frac{\xi}{\xi + \upsilon}, \ \Psi = \frac{\xi}{\xi + \mu}, \ F_1\text{-score} = 2 \times \frac{\Pi \times \Psi}{\Pi + \Psi}, \tag{18}$$

where ξ, υ, and μ denote the numbers of true positive, false positive, and false negative, respectively.

V. MEASUREMENT METHODOLOGY AND DATASET GENERATION

The dataset to test and evaluate the proposed method is developed from the measurements taken through a measurement campaign conducted at different locations and frequency bands. In order to make the model robust against environmental changes, measurements have been conducted in different locations as illustrated in Fig. 5. The locations of transmitters and measurement points can be seen in Fig. 5. It can be seen that the signals propagate through the urban area, and then reach the receivers in sub-urban area. The measurement focuses on 800, 900, 1800, and 2100 MHz frequency bands that are allocated for cellular communications. Rohde Schwarz FSW26 spectrum analyzer and a set of Yagi-Uda antennas are employed at the receiver. The measurements are unified as follows: for each signal observed in the spectrum, 16384 I/Q samples are taken. Measurements are conducted at 15 different SNR levels. Each level consists of the same number of signals which is 4000. Therefore, 60000 signals in total are recorded and included in the dataset. Sample power spectra of these signal types, obtained with the Welch's method, are shown in Fig. 4. When the proposed method is considered, the dataset is split into test and train data with the proportion of 0.4 and 0.6, respectively.

To better understand the effects of wireless communications channels over the received signals, first, amplitude distributions of four different recordings of all three signals are given in Fig. 6. Fig. 6 indicates different power and amplitude levels. The distribution of the received power changes considerably since the measurements are taken at different locations, times and frequency bands. This result implies Rayleigh-like fading behavior stemming from the amplitude distributions of received signals. This is an expected result when the measurement area and the locations of transmitters and receivers are considered. Eventually the received power of the signal is obviously affected

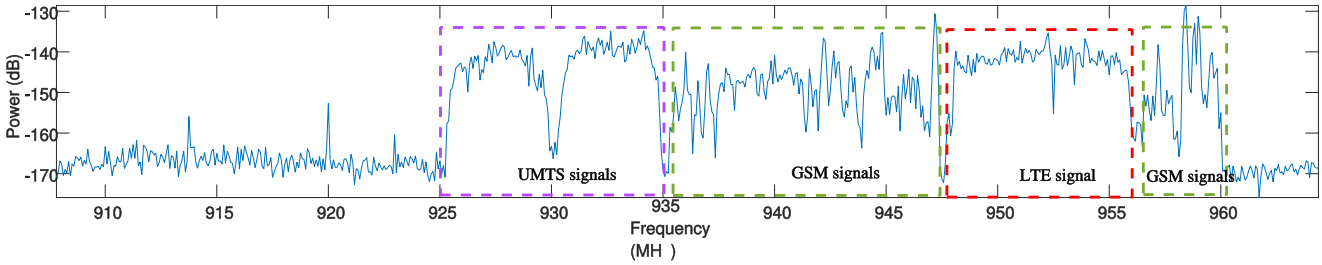


Fig. 4. This snapshot of spectrum denotes a sample from the dataset comprised of cellular signals recorded during a comprehensive measurement campaign. 900 MHz band is represented here but the measurements are not limited to that band; thus, cover all cellular bands.

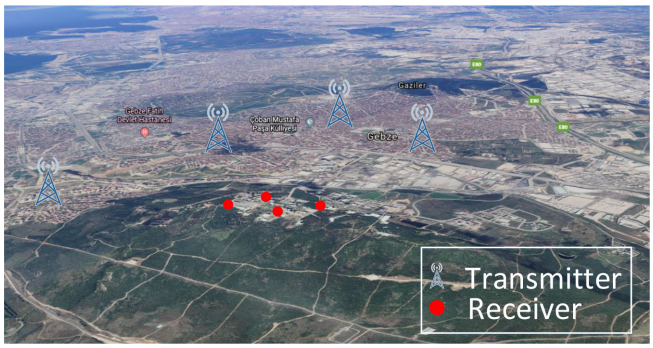


Fig. 5. An overview of the measurement area. The transmitters are located in the urban area, but the receivers are in a sub-urban area.

by the shadowing, multipath fading and path loss as depicted in Fig. 6.

For comparative analysis, we need to blindly find SNR levels of received signals. For this purpose, we used a subspace-based SNR estimation algorithm when we created the dataset. The algorithm finds the noise floor first. Then, it finds the total power of a received noisy signal. Last, the SNR regarding the received signal is estimated. As the algorithm works by finding the noise floor, it cannot show accurate performance if noise power is higher than or equal to the signal power. Therefore the dataset includes only signals whose SNR levels are higher than 0 dB. The SNR estimation method allows fair comparison without any dependency on the number of symbols.

The dataset is shared in [27] in the format of SCF. The dataset covers 60000 SCF matrices with the dimensions of 8193×16 corresponding to received I/Q samples of 16384 for each signal.

VI. Classification Performance Analysis

We evaluate the performance of the proposed classification model over the comprehensive dataset described in Section V. Therefore, the dataset is composed of GSM, WCDMA for UMTS and LTE signals which are recorded over-the-air at different locations with unique conditions in terms of the number of channel taps, and fading, again as noted in Section V. Training and test sets contain 9000 and 6000 signals for each waveform. The I/Q signal length is 16384. CNN is trained and tested on the graphics processing unit (GPU) server equipped with four NVIDIA Tesla V100 GPUs.

First, we focus on the results for CASE1. As stated before, CASE1 refers to four-classes classification problem. To investigate the problem in a basic and fundamental way, we utilized a baseline method. First, Naive Bayes classifier is employed for this task. Fig. 7 denotes comparative test accuracy. The baseline method cannot perform over 75% accuracy but, the test accuracy of CNN model exceeds 90% at 11 dB SNR. It takes a maximum accuracy value of 92% at 15 dB. The confusion matrices related to CASE1 are depicted in Fig. 8. Due to the low SNR values, the model mostly can not accurately classify the signals and identifies the signal as Noise. This case can be observed in Fig. 8(a). Therefore, dividing the problem into two parts becomes a viable alternative: first sense, then classify. In this case, we analyse both CNN detector and CNN classifier (see Fig. 2). For the sensing part of the architecture, noise signals are labeled as 0 and the rest of the set is labeled as 1. The detection results are plotted again in Fig. 7 as $P^{\mathrm{S}}_{\mathtt{CASE2}}$. The detection accuracy follows 96% at almost all SNR values.

Following the steps above, assuming that a signal is present in the spectrum at the output of CNN detector of CASE2 in Fig. 2, the performance of the CNN classifier can be investigated. This stage is labeled as $P^{\mathrm{C}}_{\mathtt{CASE2}}$ in Fig. 7 and it is observed that the classification accuracy exceeds 90% at 3 dB SNR. It gives the best performance, 98.5%, at 9 dB and it is remained stable until 15 dB. As given in (17), $P_{\mathtt{CASE2}}$ shows the cascaded sensing and classification performance under the presence of a signal in the spectrum. $P_{\mathtt{CASE2}}$ shows a better performance than $P_{\mathtt{CASE1}}$, which denotes the joint sensing and classification performance. The Fig. 9(a) depicts the confusion matrices related to CNN classifier of CASE2 and implies that even at low SNR regime, the classifier can identify GSM signals with high accuracy; however, overall precision of the classifier is low *i.e.*, in contrary to GSM signals, the classifier has difficulty in recognition of UMTS and LTE signals in low SNR regime. But the accuracy and precision of the classifier enhance as SNR increases in Fig. 9(b) and (c). This phenomena is observed due to the dominance of characteristics in feature matrices which follow Gaussian distribution. As known, GSM is associated with Gaussian minimum shift keying (GMSK); therefore, GSM signals inherently show characteristics defined by Gaussian distribution in the case of high SNR. Decreasing in SNR leverages Gaussian characteristics in the received signal because of AWGN. That is to say, UMTS and LTE signals with lower SNR values become denoting Gaussian characteristics; thus, the model is prone to learn Gaussian characteristics to decrease its loss function. When the trained

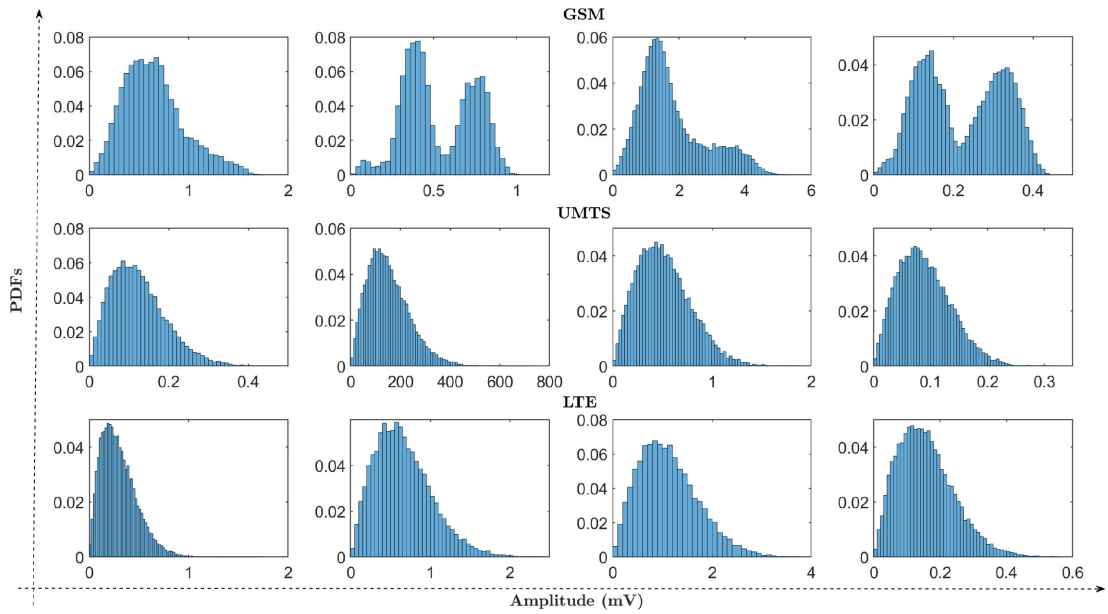


Fig. 6. Sample PDFs of the amplitude of received signals in the dataset. The example PDFs show the different channel and received power characteristics.

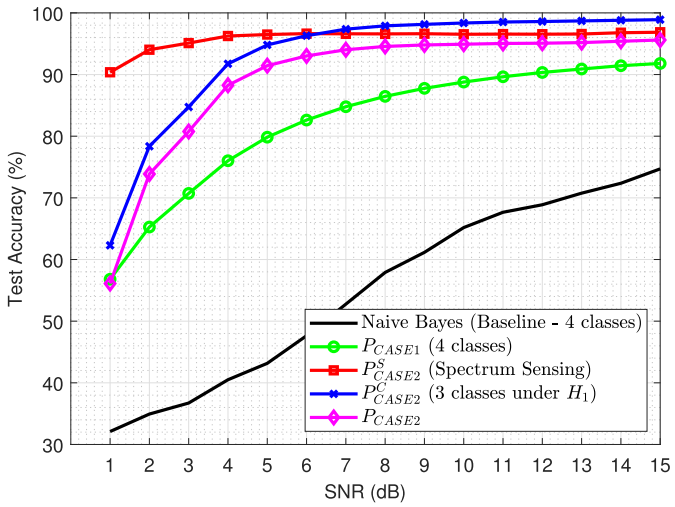


Fig. 7. Accuracy values with respect to SNR level of the received signals for both cases.

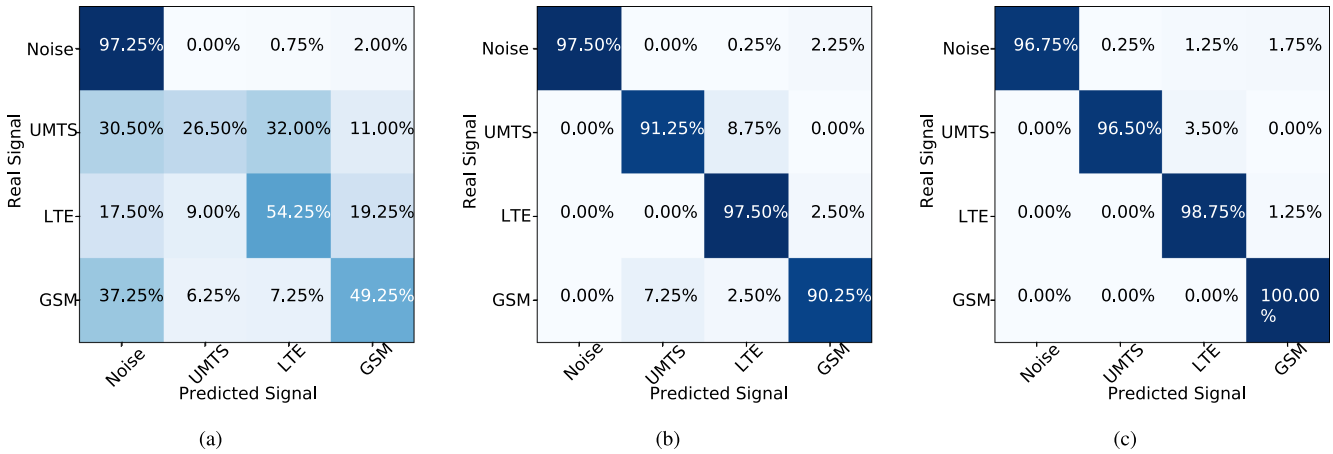
model is tested, it is expected that the model can identify the signals which have dominant Gaussian characteristics. As a result, the model can identify GSM signals, which inherently denote Gaussian attributes, in lower SNR regime where UMTS and LTE signals lose their unique features. This statement shows parallelism with the results given in CASE1 in Fig. 8(a). The model accurately identifies AWGN at lower SNR regime as given in Fig. 8(a).

The results for CASE2 are given in parts to this point. Now, we can examine the overall performance of CASE2. Obviously, there is a loss of performance due to some misdetection in the sensing phase. Both the detection rate in the sensing stage and the accuracy in the classification stage are high at 3 dB and thereafter, so overall performance does not suffer a significant loss. As shown in Fig. 7, the overall performance of CASE2 is far superior to that of CASE1. Especially at low SNR levels, the signals remaining after first detecting and separating noise from the signal set by the CNN detector can be classified with much better performance. In this way, the performance is higher in CASE2. However, it should be noted that CASE2 is more costly than CASE1 in terms of training time and the number of models. Obviously, CASE2 can be predicted to perform better than CASE1 in the presence of a jammers exhibiting Gaussian characteristics or other interfering signals.

Sensing performance can be considered that it is slightly lower than conventional spectrum sensing methods like energy detector and matched filter. However, it should be noted that this study employs real–world data rather than simulation or synthetic data. For example, energy detectors can sense a signal in a spectrum with optimal performance; however, it needs to know noise variance. But even with a slight error on estimating the noise variance, the sensing performance seriously decreases. Moreover, as the power of spread spectrum signals (e.g. WCDMA in UMTS) is spreaded in a wide band, its power is very close to the noise floor. By taking into this account, in a fading environment, it can be said that energy detector cannot perform a satisfactory detection rate for spread spectrum signals as stated in [32]. It is worth noting that our measurements follow Rayleigh distribution as seen in Fig. 6. On the other hand, matched filters are waveform-specific solution and they require the perfect knowledge for signals.

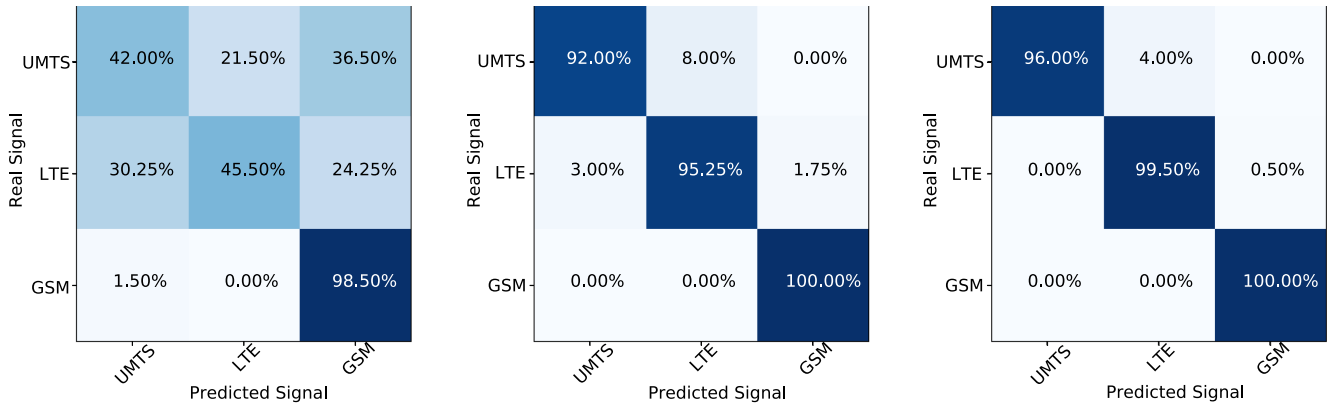
A. Investigation for the Impact of Different Features

In this section, we compare the performance of other features of I/Q, AP, and FFT which are frequently employed for sensing purposes with SCF. The features are used as detailed in Section II. The results of this test are presented in Table II. Unlike the modulation classification studies [12], [33], I/Q cannot provide a meaningful input for the model due to the severe fading effect on



(a) (b) (c)

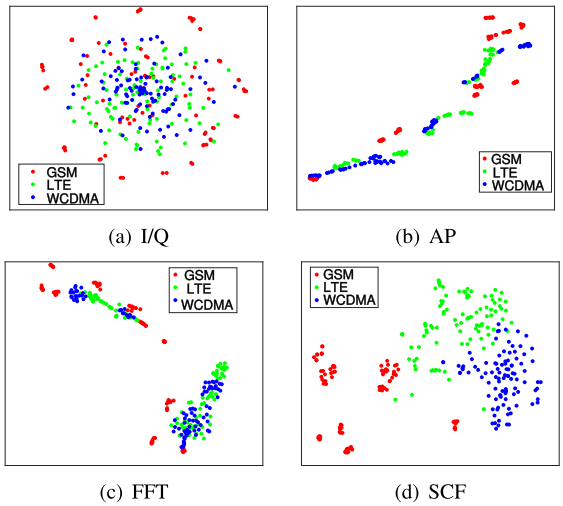
Fig. 8. Confusion matrices for CASE1 at SNR levels of (a) 1 dB, (b) 5 dB, and (c) 10 dB. It should be noticed that the model does not randomly choose only one signal at low SNR level.



(a) (b) (c)

Fig. 9. Confusion matrices for the classification part of CASE2 at SNR levels of (a) 1 dB, (b) 5 dB, and (c) 10 dB. It should be noticed that the model does not randomly choose only one signal at low SNR level.

the phase of signal. The histograms of phase imply that the signal phase is corrupted and the information on the phase is lost. That is why I/Q shows poor performance. The average performances also indicate that SCF outperforms I/Q, AP, and FFT for all SNR levels. Assuming that these two are used along with I/Q as the main features for training, these results show significant gains for real-world signals especially above 5 dB SNR level. It is observed that AP performs better than FFT. The average training time per epoch is approximately 60 s for SCF feature where both FFT and AP take 7.5 s per epoch; however, both FFT and AP cannot show an acceptable classification performance, P^{C}_{CASE2}. Although the cost of computing both features is far behind the SCF, they are far from delivering the desired performance. To visualize the vectors in input space, we employ the t-distributed stochastic neighbor embedding (t-SNE) algorithm. Although originally I/Q samples are not linearly separable, SCF clusters the vectors in the space and allows almost linear separation as depicted in Fig. 10. The analysis based on t-SNE results show that SCF better separates signal vectors in space. The results of this study are in line with the previous analysis [34].



(a) I/Q (b) AP

(c) FFT (d) SCF

Fig. 10. Two-dimensional demonstration of the features by the t-SNE algorithm. This illustration shows that in contrary to the other features, SCF can separately cluster real-world signals in space successfully.

TABLE II
CLASSIFICATION PERFORMANCE METRICS FOR THE PROPOSED CNN MODEL WITH SCF, AP, AND FFT FEATURES FOR CASE2

SNR	Feature	Signal	Precision (Π)	Recall (Ψ)	F1-Score
1dB	I/Q	UMTS	0.33	1.00	0.49
		LTE	0.00	0.00	0.00
		GSM	0.00	0.00	0.00
		Average	0.11	0.33	0.16
	AP	UMTS	0.35	0.39	0.37
		LTE	0.31	0.29	0.30
		GSM	0.30	0.29	0.30
		Average	0.32	0.32	0.32
	FFT	UMTS	0.00	0.00	0.00
		LTE	0.21	0.50	0.30
		GSM	0.00	0.00	0.00
		Average	0.07	0.17	0.10
	SCF	UMTS	0.59	0.40	0.48
		LTE	0.68	0.46	0.54
		GSM	0.62	0.98	0.76
		Average	0.63	0.61	0.59
5dB	I/Q	UMTS	0.33	1.00	0.49
		LTE	0.00	0.00	0.00
		GSM	0.00	0.00	0.00
		Average	0.11	0.33	0.16
	AP	UMTS	0.43	0.42	0.42
		LTE	0.39	0.43	0.41
		GSM	0.61	0.56	0.58
		Average	0.47	0.47	0.47
	FFT	UMTS	0.53	0.49	0.51
		LTE	0.25	0.51	0.34
		GSM	0.00	0.00	0.00
		Average	0.26	0.33	0.28
	SCF	UMTS	0.97	0.92	0.94
		LTE	0.92	0.95	0.94
		GSM	0.98	1.00	0.99
		Average	0.96	0.96	0.96
10dB	I/Q	UMTS	0.33	1.00	0.49
		LTE	0.00	0.00	0.00
		GSM	0.00	0.00	0.00
		Average	0.11	0.33	0.16
	AP	UMTS	0.50	0.52	0.51
		LTE	0.50	0.53	0.52
		GSM	0.88	0.79	0.83
		Average	0.63	0.61	0.62
	FFT	UMTS	0.49	0.37	0.42
		LTE	0.27	0.62	0.38
		GSM	0.00	0.00	0.00
		Average	0.26	0.33	0.27
	SCF	UMTS	1.00	0.96	0.98
		LTE	0.96	0.99	0.98
		GSM	1.00	1.00	1.00
		Average	0.99	0.98	0.98
15dB	I/Q	UMTS	0.33	1.00	0.49
		LTE	0.00	0.00	0.00
		GSM	0.00	0.00	0.00
		Average	0.11	0.33	0.16
	AP	UMTS	0.55	0.54	0.55
		LTE	0.55	0.57	0.56
		GSM	0.94	0.93	0.94
		Average	0.68	0.68	0.68
	FFT	UMTS	0.73	0.49	0.58
		LTE	0.35	0.82	0.49
		GSM	0.00	0.00	0.00
		Average	0.36	0.44	0.36
	SCF	UMTS	1.00	0.97	0.99
		LTE	0.97	0.99	0.98
		GSM	0.99	1.00	1.00
		Average	0.99	0.99	0.99

Another important factor in feature selection is computational complexity as well. It can generally say that the feature with high computational complexity can achieve better accuracy. In some applications, the hardware has very limited computation capacity. The designers should notice the trade-off between complexity and accuracy. It is worth saying that the feature selection criteria strictly depends on the application and its hardware. In this study, we provide a tutorial for deep learning-based spectrum sensing and signal classification systems. For example, SCF provides the highest accuracy among features; whereas, its complexity is $\mathcal{O}(N^2)$ [28]. The computational complexity of FFT can be given as $\mathcal{O}(N \log_2(N))$. Although FFT is more complex than AP, it cannot perform as high as AP. As a result, it can be said that SCF can be preferred in the aspect of complexity and accuracy trade-off.

TABLE III
PERFORMANCE COMPARISON BETWEEN THE EXISTING DL NETWORKS AND THE PROPOSED SYSTEM FOR THE CLASSIFICATION STAGE OF CASE2 AT SNR VALUE OF 15 DB

Network	Signal	**Precision**	**Recall**	F_1**-score**
CLDNN [24]	UMTS	0.33	1.00	0.50
	LTE	0.00	0.00	0.00
	GSM	0.00	0.00	0.00
	Average	0.11	0.33	0.17
LSTM [35]	UMTS	0.33	1.00	0.50
	LTE	0.00	0.00	0.00
	GSM	0.00	0.00	0.00
	Average	0.11	0.33	0.17
Proposed CNN with SCF	UMTS	0.79	1.00	0.88
	LTE	1.00	0.72	0.84
	GSM	0.99	1.00	0.99
	Average	0.93	0.91	0.91

B. Comparison With Existing Deep Learning Networks

The existing DL networks are employed to classify the cellular communication signals. We utilize CLDNN [24] and LSTM [35] models. These models are originally used in modulation classification. Without any change in the models, input matrix, and input vector as proposed in the papers are adopted in the study. CLDNN takes a 2×128 matrix which is composed of amplitude and phase values for each I/Q sample. On the other hand, LSTM model utilizes a vector reshaped version of the matrix used in CLDNN. Therefore, the length of the vector is 256. Its first half includes in-phase components while the rest of the vector is quadrature components. Other details are found in [24], [35]. The precision, recall, and F_1-score are given in Table III. It shows that CLDNN and LSTM decide that the received signal is UMTS whatever it actually is. Even though LSTM and CLDNN can be trained in a short time by using I/Q vector and matrix, employing I/Q vector and matrix give poor classification performance.

C. Comparison With SVM

In our previous work, we employed SVMs to identify real-world signals [34]. Even though utilization of SCF in SVM provides good performance, training of SVM should be conducted for each SNR level separately *i.e.*, at the end of the training, the more SNR values in the dataset, the more models

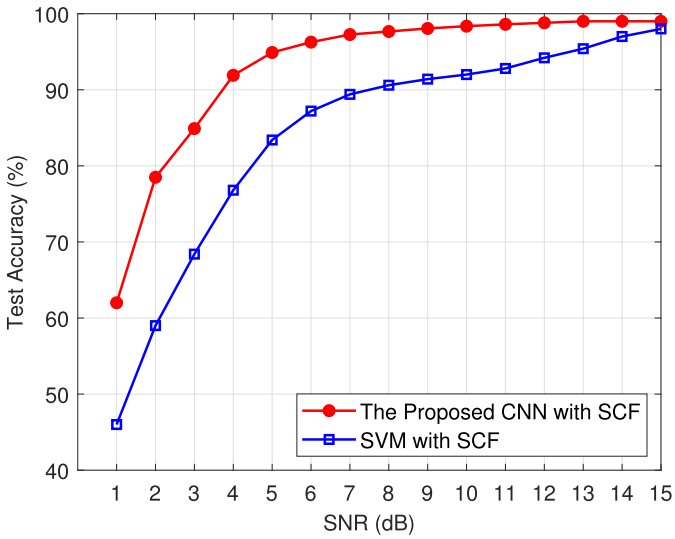


Fig. 11. The classification performance comparison between SVM in [34] and the proposed CNN structure for P^{C}_{CASE2}.

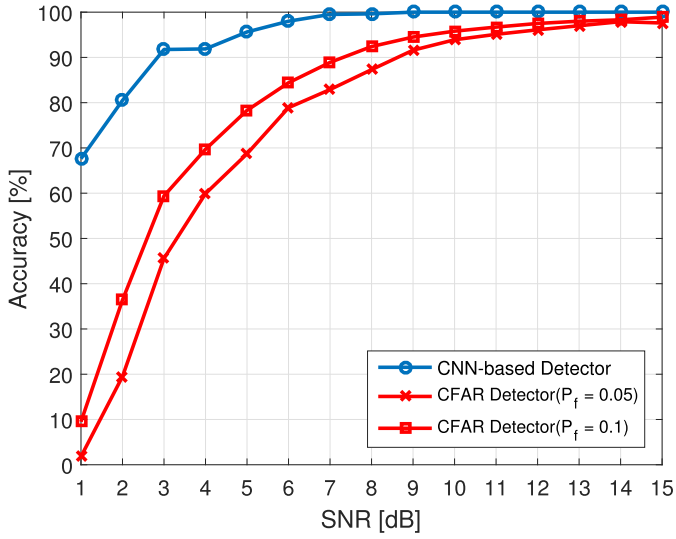


Fig. 12. Spectrum sensing performances of CFAR detectors and CNN-based detector with respect to SNR.

should be created. The real-world utilization of SVM requires an SNR estimator and loading of all pre-trained models to memory during operation; thus, reducing the applicability of the method and making improvements a necessity. As seen in Fig. 11, the CNN-based classifier shows a superior performance compared to SVM-based classifier of [34], under the conditions of the classification part of CASE2. To this end, while CNN-based classifier employs a less costly feature due to elimination of mapping of bi-frequency spectrum, it still performs with higher accuracy. Therefore, producing a model independent of the SNR is an advantage of the proposed CNN based method since the training set contains an equal number of signals from each SNR. As a result, a single model would be adequate for classification in a large SNR range at the test stage.

D. Comparison With CFD

Besides signal classification, the proposed CNN model can be used for spectrum sensing. We investigated the sensing performance of the model by training a CNN-based spectrum occupancy detector trained over 600 pure noise signals and 600 noisy WCDMA signals for each SNR value. Then, the model is tested with 400 pure noise signals and 400 noisy WCDMA signals for each SNR level and sensing results are acquired. Furthermore, for comparison purposes, we implement a constant false alarm rate (CFAR) detector utilizing classical CFD [36] to identify WCDMA signals. CFAR detector utilizes cell averaging [37] over the SCF matrix. The detector finds peaks in the SCF matrix by a peak detector with a CFAR. As nature of the CFAR detectors, the threshold is set according to a CFAR. Thus, this detector guarantees a CFAR that can be selected according to the operational requirements. Please note that UMTS signals are deliberately selected due to their known dominant SCF characteristics stemming from cyclic spreading codes. The results of this test are given in Fig. 12. In view of these results, it is clearly seen that the CNN-based detector outperforms the CFAR detector at all SNR regimes. For example, the sensing performance of the CNN-based detector is 91.75% at 3 dB while the probability of detection for the CFAR detector are 45.6% and 59.4% for the selected false alarm rates as 0.05 and 0.1, respectively.

E. Focusing on the Meaningful Region of Spectral Correlation Function

As mentioned above, SCF creates a bi-frequency feature matrix with high data size. Therefore, the training process necessitates high computation capacity. It should be noted that we employed a GPU server equipped with four NVIDIA Tesla V100 GPUs. To make training over a single GPU possible (or to accelerate the training process), we focused on the meaningful part of the feature matrices. As shown in Fig. 1, elements of the matrices have insignificantly small values but except the elements around the middle of the matrices. Moreover, this approach allows to investigate the possibility of accuracy improvement and the fair comparison with the existing DL networks. As stated in Section V, an SCF matrix has the dimension 8193×16. Therefore, it is not possible to train such a dense model in our server equipped with four NVIDIA Tesla V100 GPUs. To compare our proposed CNN architecture with a more dense model, we decrease the dimensions of the SCF matrices by using only 16×16 part in the middle of the matrices. Only in this way, we are able to train complex models such as LSTM [19] and DenseNet [25] with SCF. Moreover, the proposed CNN, CLDNN [24], and ResNet [12] are also trained with the shrunken SCF matrices. It is worth saying that, we conduct four-class classification (i.e., CASE1) in this study. The results depicted in Fig. 13 shows that the proposed CNN is favorable in terms of both low complexity (i.e., epoch time) and efficient memory allocation, as well as high test accuracy. During

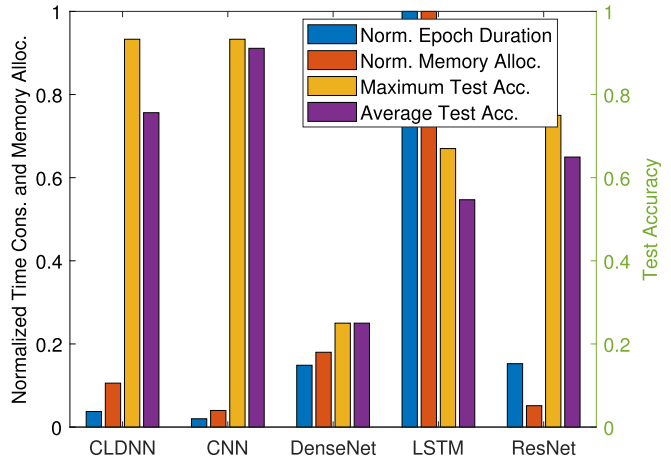


Fig. 13. Model comparison in terms of memory, complexity, and accuracy. The epoch time and memory allocation rate are normalized with their maximum values observed in these models (maximum values for both are observed in LSTM). The average accuracy is the mean accuracy in the SNR range between 1 dB and 15 dB.

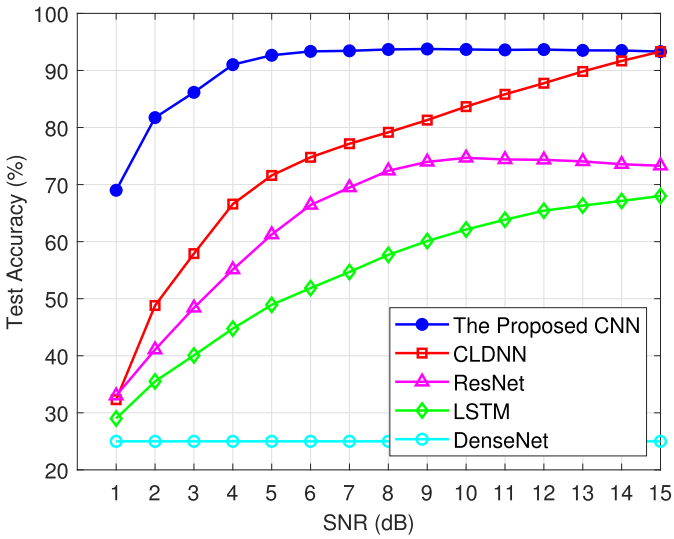


Fig. 14. Test accuracy with respect to SNR values for the proposed CNN, CLDNN, LSTM, ResNet, and DenseNet models.

this study, batch sizes are kept same for all models. The memory allocation and training time have been normalized by LSTM's memory allocation rate and training time, respectively; thus, computer-independent results are provided in Fig. 13. It should be noted that early stopping is used during training of models and the minimum number of epochs is required by the proposed CNN. Furthermore, Fig. 14 denotes the accuracy with respect to SNR levels for each model. By considering results, it can be observed that the proposed CNN is more robust and efficient than the existing models. Moreover, it is seen that CNN gives better results with this smaller matrix than the complete matrix is used. By eliminating the region except for the meaningful part of SCF, the input matrices become more distinct from each other. Fig. 1 implies that SCF matrices have similarities except for the meaningful part. The confusion matrices in Fig. 15 for 16×16 inputs denote the improvement in the precision of AWGN. This

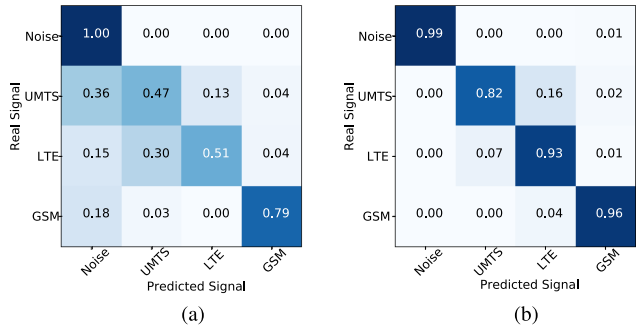


Fig. 15. Confusion matrices for CASE1 at SNR levels of (a) 1 dB and (b) 5 dB when 16×16 inputs are employed.

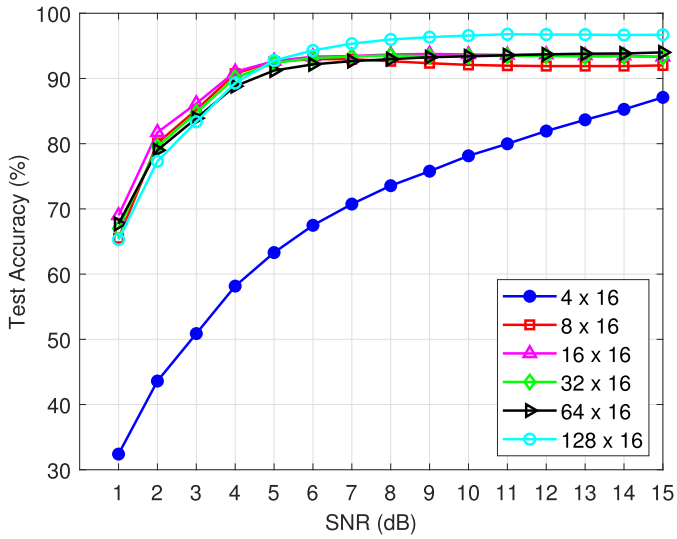


Fig. 16. The test accuracy of the proposed CNN architecture with respect to input size.

explains why the small portion of the matrix can lead to higher accuracy.

It is also explored how the dimensions of the small partition affect the performance of the CNN model. The results show that using 4 rows does not perform well enough. When using rows between 8 and 128 (as power of two), the results are satisfactory. The test accuracy with respect to input size is demonstrated in Fig. 16. It is revealed that by considering the accuracy at lower SNR regimes and the training time, 16×16 is the most suitable size for the CNN.

VII. Conclusion

In this study, a DL-based method utilizing SCF as an input to the designed CNN model to achieve spectrum sensing or signal identification interchangeably or jointly without the requirement of any a priori information is proposed. First approach investigates the joint sensing and classification of wireless signals. Second, a sequential approach is adopted. The results show that sequential approach performs better than the joint approach. Moreover, comparative analysis indicated the superiority of SCF as a distinctive feature when compare to the contemporary features utilized for currently available DL-based detector models. The results also imply that under stringent channel

conditions, the CNN model of the proposed method provides better spectrum sensing performance than other available DL models, SVMs, and classical CFD. These results indicate that applicability of DL-based techniques in the rapidly changing communications environment of contemporary wireless communications networks. It should be noted that SCF can extract the unique characteristics of the signals with low sensitivity to noise. Therefore, it is possible to train a network using a small number of signals. But the computational complexity of SCF is high. Also, since SCF produces an output in the bi-frequency plane, the data size can be large. Thus, it may require batch size reduction when training larger datasets. Moreover, it may cause *out of memory* error on some GPUs. Multiple GPUs can be used to avoid memory error. If it is required, it is possible to scale SCF to larger datasets by increasing computational capacity. Furthermore, focusing on meaningful part of SCF reduces the data size. Hence, the training complexity can be decreased.

In subsequent studies, the performance of the proposed method for sensing and identification of other wireless signals or modulation techniques with cyclic features can be explored. Also, Doppler shift is required to be investigated for vehicular applications. It is stated above that the dataset mostly includes signals received through Rayleigh fading channels. As remembered, Rayleigh fading creates signals with uniform phase distribution. Similarly, although the Doppler shift produces a phase rotation, it corresponds to the rotation of the uniformly distributed phase on a circle in polar coordinates. In addition, considering that the transmitter and receiver are different and asynchronous devices, the current system is subject to some phase shift. It is observed in the above results that the performance is high under these conditions. However, more detailed studies are needed on the effect of Doppler shift on system performance. Considering the simple implementation of `CASE1`, residual networks should be further investigated to improve sensing performance by exploiting noise features. Furthermore, the performance of the proposed method can be investigated against adversarial attacks and efforts can be made to develop various techniques to strengthen its resistance to these types of intrusions. Although this study focuses on supervised learning, it is possible to improve the performance of the proposed method by supporting unsupervised learning methods in feature extraction.

REFERENCES

[1] S. Haykin and P. Setoodeh, "Cognitive radio networks: The spectrum supply chain paradigm," *IEEE Trans. Cogn. Commun. Netw.*, vol. 1, no. 1, pp. 3–28, Mar. 2015.

[2] J. G. Andrews *et al.*, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[3] C. De Lima *et al.*, "Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges," *IEEE Access*, vol. 9, pp. 26902–26925, 2021.

[4] T. Wild, V. Braun, and H. Viswanathan, "Joint design of communication and sensing for beyond 5G and 6G systems," *IEEE Access*, vol. 9, pp. 30845–30857, 2021.

[5] K. Abboud, H. A. Omar, and W. Zhuang, "Interworking of DSRC and cellular network technologies for V2X communications: A survey," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9457–9470, Dec. 2016.

[6] K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, and Y. Zhou, "Heterogeneous vehicular networking: A survey on architecture, challenges, and solutions," *IEEE Commun. Surv. Tut.*, vol. 17, no. 4, pp. 2377–2396, Oct.–Dec. 2015.

[7] Y. Wang, J. Jose, J. Li, X. Wu, and S. Subramanian, "Opportunistic use of the DSRC spectrum," *US Pat. App. 13/921,706*, Dec. 2014.

[8] A. Gorcin and H. Arslan, "Signal identification for adaptive spectrum hyperspace access in wireless communications systems," *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 134–145, Oct. 2014.

[9] W. M. Lees, A. Wunderlich, P. Jeavons, P. D. Hale, and M. R. Souryal, "Deep learning classification of 3.5GHz band spectrograms with applications to spectrum sensing," *IEEE Trans. Cogn. Commun. Netw*., vol. 5, no. 2, pp. 224–236, Jun. 2019.

[10] Z. Qin, X. Zhou, L. Zhang, Y. Gao, Y.-C. Liang, and G. Y. Li, "20 years of evolution from cognitive to intelligent communications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 6–20, Mar. 2020.

[11] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Process. Mag.*, vol. 8, no. 2, pp. 14–36, Apr. 1991.

[12] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Signal Process.*, vol. 12, no. 1, pp. 168–179, Jan. 2018.

[13] M. Kulin, T. Kazaz, I. Moerman, and E. D. Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications," *IEEE Access*, vol. 6, pp. 18484–18501, Mar. 2018.

[14] 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.

[15] C. Liu, Z. Wei, D. W. K. Ng, J. Yuan, and Y.-C. Liang, "Deep transfer learning for signal detection in ambient backscatter communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1624–1638, Mar. 2021.

[16] A. Mehrabian, M. Sabbaghian, and H. Yanikomeroglu, "Spectrum sensing for symmetric α-stable noise model with convolutional neural networks," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5121–5135, Aug. 2021.

[17] D. Uvaydov, S. D'Oro, F. Restuccia, and T. Melodia, "Deepsense: Fast wideband spectrum sensing through real-time in-the-loop deep learning," in *Proc. IEEE Conf. Comput. Commun.*, 2021, pp. 1–10.

[18] S. Kokalj-Filipovic, R. Miller, and J. Morman, "AutoEncoders for training compact deep learning RF classifiers for wireless protocols," in *20th Int. Workshop Signal Process. Adv. Wireless Commun.*, 2019, pp. 1–5.

[19] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, May 2018.

[20] N. Han, S. Shon, J. H. Chung, and J. M. Kim, "Spectral correlation based signal detection method for spectrum sensing in IEEE 802.22 WRAN systems," in *Proc. 8th Int. Conf. Adv. Commun. Technol.*, vol. 3, 2006, pp. 1765–1770.

[21] M. Oner and F. Jondral, "Air interface identification for software radio systems," *AEU - Int. J. Electron. Commun.*, vol. 61, no. 2, pp. 104–117, 2007.

[22] E. Karami, O. A. Dobre, and N. Adnani, "Identification of GSM and LTE signals using their second-order cyclostationarity," in *Proc. IEEE Int. Instrum. Meas. Tech. Conf.*, Pisa, Italy, May 2015, pp. 1108–1112.

[23] Y. A. Eldemerdash, O. A. Dobre, O. Ureten, and T. Yensen, "Identification of cellular networks for intelligent radio measurements," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 2204–2211, Apr. 2017.

[24] S. Ramjee, S. Ju, D. Yang, X. Liu, A. E. Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," 2019, Accessed: Mar. 4, 2021. [Online]. Available: https://arxiv.org/pdf/1901.05850.pdf

[25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.

[26] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, "An overview of feature-based methods for digital modulation classification," in *Proc. Int. Commun. Signal Process., Their Appl.*, Mar. 2013, pp. 1–6.

[27] K. Tekbıyık, Ö. Akbunar, A. R. Ekti, A. Görçin, and G. K. Kurt, "COSINE: Cellular communication signal dataset," 2020, Accessed: Mar. 4, 2021. [Online]. Available: http://dx.doi.org/10.21227/safr-gh59

[28] R. Roberts, W. Brown, and H. Loomis, "Computationally efficient algorithms for cyclic spectral analysis," *IEEE Signal Process. Mag.*, vol. 8, no. 2, pp. 38–49, Apr. 1991.

[29] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11 074–11 0 82, 2017.

[30] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," in *1st Int. Conf. Learn. Representations*, 2013, pp. 1–.
[31] F. Chollet *et al.*, "Keras," 2015, Accessed: Mar. 4, 2021. [Online]. Available: https://keras.io
[32] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. 38th Asilomar Conf. Signals, Syst. Comput.*, vol. 1, 2004, pp. 772–776.
[33] K. Tekbiyik, A. R. Ekti, A. Gorcin, G. K. Kurt, and C. Kececi, "Robust and fast automatic modulation classification with CNN under multipath fading channels," in *Proc. IEEE 91st Veh. Technol. Conf.*, 2020, pp. 1–6.
[34] K. Tekbıyık, Ö. Akbunar, A. R. Ekti, A. Görçin, and G. K. Kurt, "Multi-dimensional wireless signal identification based on support vector machines," *IEEE Access*, vol. 7, pp. 138890–138903, 2019.
[35] S. Hu, Y. Pei, L. Pu, and Y.-C. Liang, "Robust modulation classification under uncertain noise condition using recurrent neural network," in *Proc. IEEE Glob. Commun. Conf.*, 2018, pp. 1–7.
[36] C. M. Spooner and A. N. Mody, "Wideband cyclostationary signal processing using sparse subsets of narrowband subchannels," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 2, pp. 162–176, Jun. 2018.
[37] M. A. Richards, *Fundamentals of Radar Signal Processing*. New York, NY, USA: McGraw-Hill Education, 2014.

Kürşat Tekbıyık (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees (with high Hons.) in 2017 and 2019, respectively, in electronics and communication engineering from Istanbul Technical University, Istanbul, Turkey, where he is currently working toward the Ph.D. degree in telecommunications engineering. He is also a Senior Researcher with HISAR Lab., TUBITAK BILGEM. His research interests include algorithm design for signal intelligence, next-generation wireless communication systems, terahertz wireless communications, and machine learning.

Özkan Akbunar (Student Member, IEEE) received the B.Sc. degree from Electrical and Electronics Engineering Department, Anadolu University, Eskişehir, Turkey. He is currently working toward the M.Sc degree with Telecommunication Engineering Department, Istanbul Technical University, Istanbul, Turkey. He is with the Scientific and Technological Research Council of Turkey, Informatics and Information Security Research Center, as a R&D Engineer since December 2017. His main research interests include wireless communication systems, digital signal processing, machine learning, and radar.

Ali Rıza Ekti (Senior Member, IEEE) is from Tarsus, Turkey. He received the B.Sc. degree in electrical and electronics engineering from Mersin University, Mersin, Turkey, (September 2002–June 2006), also studied with Universidad Politechnica de Valencia, Valencia, Spain, in 2004–2005, the M.Sc. degree in electrical engineering from the University of South Florida, Tampa, FL, USA, (August 2008–December 2009) and the Ph.D. degree in electrical engineering from the Department of Electrical Engineering and Computer Science, Texas A&M University, College Station, TX, USA, (August 2010–August 2015). He is currently an Assistant Professor with Balikesir University Electrical and Electronics Engineering Department. He also holds the Division Manager position of HISAR Lab., TUBITAK BILGEM, where he is responsible for R&D activities in the wireless communications and signal processing. His current research interests include statistical signal processing, convex optimization, machine learning, channel modeling, resource allocation, and traffic offloading in next generation wireless networks.

Ali Görçin (Senior Member, IEEE) graduated from Istanbul Technical University, Istanbul, Turkey, with the B.Sc. degree in electronics and telecommunications engineering and the master's degree on defense technologies at the same university. After working with Turkish Science Foundation (TUBITAK), on avionics projects for more than six years, he moved to the U.S., to pursue the Ph.D. degree with the University of South Florida (USF), Tampa, FL, USA, on wireless communications. He was with Anritsu Company during his tenure in USF and with Reverb Networks and Viavi Solutions after his graduation. He is a Faculty Member with Yildiz Technical University, Istanbul. He is also the President of TUBITAK BILGEM.

Güneş Karabulut Kurt (Senior Member, IEEE) received the B.S. degree (with high Hons.) in electronics and electrical engineering from Bogazici University, Istanbul, Turkey, in 2000, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Ottawa, Ottawa, ON, Canada, in 2002 and 2006, respectively. She is currently an Associate Professor of electrical engineering with Polytechnique Montréal, Montreal, QC, Canada. From 2000 to 2005, she was a Research Assistant with the University of Ottawa. Between 2005 and 2006, she was with TenXc Wireless, Canada. From 2006 to 2008, she was with Edgewater Computer Systems Inc., Canada. From 2008 to 2010, she was with Turkcell Research and Development Applied Research and Technology, Istanbul. Between 2010 and 2021, she was with Istanbul Technical University, Istanbul, Turkey. She is an Adjunct Research Professor with Carleton University, Ottawa, ON, Canada. She is a Marie Curie Fellow and was the recipient of the Turkish Academy of Sciences Outstanding Young Scientist (TÜBA-GEBIP) Award in 2019. She is also currently an Associate Technical Editor (ATE) of the *IEEE Communications Magazine* and a Member of the IEEE WCNC Steering Board.

Khalid A. Qaraqe (Senior Member, IEEE) was born in Bethlehem. He received the B.S. (Hons.) degree in electrical engineering from the University of Technology, Bagdad, Iraq, in 1986, the M.S. degree in electrical engineering from the University of Jordan, Amman, Jordan, in 1989, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 1997. From 1989 to 2004, he has held a variety positions in many companies and he has more than 12 years of experience in the telecommunication industry. He has worked on numerous projects and has experience in product development, design, deployments, testing and integration. In July 2004, he joined the Department of Electrical and Computer Engineering, Texas A&M University, Qatar, where he is currently a Professor and the Managing Director with the Center for Remote Healthcare Technology, Qatar. He has authored or coauthored more than 131 journal papers in top IEEE journals, and published and presented 250 papers at prestigious international conferences. He has authored or co-authored 20 book chapters, four books, three patents, and presented several tutorials and talks. His research interests include communication theory and its application to design and performance, analysis of cellular systems, indoor communication systems, mobile networks, broadband wireless access, cooperative networks, cognitive radio, diversity techniques, index modulation, visible light communication, FSO, telehealth, and noninvasive bio sensors. He has been awarded more than 20 research projects consisting of more than USD 13 M from local industries in Qatar and the Qatar National Research Foundation (QNRF).