Combined RF-Based Drone Detection and Classification

Sanjoy Basak, Sreeraj Rajendran, Sofie Pollin, *Senior Member, IEEE*, and Bart Scheers

***Abstract*—Despite several beneficial applications, unfortunately, drones are also being used for illicit activities such as drug trafficking, firearm smuggling or to impose threats to security-sensitive places like airports and nuclear power plants. The existing drone localization and neutralization technologies work on the assumption that the drone has already been detected and classified. Although we have observed a tremendous advancement in the sensor industry in this decade, there is no robust drone detection and classification method proposed in the literature yet. This paper focuses on radio frequency (RF) based drone detection and classification using the frequency signature of the transmitted signal. We have created a novel drone RF dataset using commercial drones and presented a detailed comparison between a two-stage and combined detection and classification framework. The detection and classification performance of both frameworks are presented for a single-signal and simultaneous multi-signal scenario. With detailed analysis, we show that You Only Look Once (YOLO) framework provides better detection performance compared to the Goodness-of-Fit (GoF) spectrum sensing for a simultaneous multi-signal scenario and good classification performance comparable to Deep Residual Neural Network (DRNN) framework.**

***Index Terms*—Signal detection and classification, sensor systems and applications, UAV.**

I. Introduction

THERE has been a tremendous technological improvement in the drone industry. Drones are now being equipped with state-of-the-art (SoA) technologies and sensors such as GPS, LIDAR, radar and visual sensors. These technologies facilitate drones to support numerous applications like cinematography, farming, surveillance and recreational activities. Drones equipped with advanced technologies have great potential for damaged infrastructure inspections, urgent aid supply, search and rescue operations to remote and unreachable places. Apart from these beneficial applications, drones are also being used for illegal activities which impose risks to public safety. The illegal activities include but are not limited to violation of public privacy, drug trafficking, firearm smuggling, bombing, and invading security-sensitive places like airports and nuclear power plants.

Several Counter Unmanned Aircraft Systems (C-UAS) have been proposed to disable the attack from a drone, which are mainly divided into two categories: hard and soft interception (kinetic or non-kinetic solution). The kinetic solutions include intercepting a drone using (i) a trained bird of prey (ii) a net gun [1] (iii) a laser beam, and (iv) a firearm. The non-kinetic solutions include: (i) GPS spoofing [1] to deceive a drone's localization system and (ii) RF jamming. Irrespective of the chosen solution for any environment, the presence of a drone should be detected and classified beforehand.

Detecting and classifying a drone automatically is a challenging task. Some popular technological approaches to detect and classify a drone include: (i) Radar detection, (ii) Video detection, (iii) Acoustic detection, and (iv) RF-based detection. A comprehensive literature review on the current SoA Machine Learning-based drone detection and classification using these technologies is presented in [2]. Researchers also proposed to integrate multiple technologies [3] for the detection and classification of UAVs.

The radar detection exploits the back-scattered RF signal to detect and classify a drone. The conventional radar systems will fail to detect a mini-drone due to its small radar cross section (RCS). To overcome this problem, researchers utilized the micro-Doppler signature of a Quadcopter or a Multi-rotor UAV to detect and classify it using a multi-static radar [4] or a Frequency Modulated Continuous Wave (FMCW) radar [5], [6]. A complete review of the detection and classification strengths of the current SoA FMCW radars is presented in [6].

The video/image detection includes both visual and thermal detection, and in [7]–[9] researchers proposed several drone detection methods using this technology. With this technique, drone detection is performed by analyzing its color, shape and edge information [7]. The detection method is reliable, however, it requires a line of sight (LOS) between the drone and camera and the performance is highly dependent on daylight conditions and weather conditions like dust, rain, fog and cloud. Furthermore, the resemblance of a bird to a drone makes it more challenging for a video detector. In [8], the authors utilized the motion and trajectory information of a drone to differentiate it from a bird. A brief overview of the frameworks capable of differentiating a drone from a bird is presented in [10].

The acoustic detection system utilizes the sound generated by flying drones to detect its presence using microphones.

Manuscript received February 1, 2021; revised May 5, 2021 and June 28, 2021; accepted July 13, 2021. Date of publication July 26, 2021; date of current version March 8, 2022. This work was supported by the Belgian Ministry of Defence. The associate editor coordinating the review of this article and approving it for publication was H. T. Dinh. *(Corresponding author: Sanjoy Basak.)*

Sanjoy Basak is with the Department CISS, Royal Military Academy, 1000 Bruxelles, Belgium, and also with the Department ESAT, KU Leuven, 3001 Leuven, Belgium (e-mail: sanjoy.basak@rma.ac.be).

Bart Scheers is with the Department CISS, Royal Military Academy, 1000 Bruxelles, Belgium (e-mail: bart.scheers@rma.ac.be).

Sreeraj Rajendran and Sofie Pollin are with the Department ESAT, KU Leuven, 3001 Leuven, Belgium (e-mail: sreeraj.rajendran@esat.kuleuven.be; sofie.pollin@esat.kuleuven.be).

Digital Object Identifier 10.1109/TCCN.2021.3099114

In [11], the authors proposed a framework using Hidden Markov Models (HMM) to perform phoneme analysis and identify a flying drone from its emitted sounds. Furthermore, the detection and tracking of a drone using an array of antennas are also proposed in the literature. A small tetrahedral array [12] or a microphone array consisting of 120 elements [13] are used for drone detection and tracking. The acoustic detection generally works well in a quiet or less noisy environment, however, the performance deteriorates if the environment is noisy such as urban or industrial areas or near seashores.

One of the most promising approaches to detect the presence of a drone is through RF sensing. The commercial drones perform RF communication with their ground control station (GCS) for flight control and navigation, live video transmission and transfer telemetry information. The autonomous drones also perform active RF communication to transfer live videos and telemetry messages. An RF drone detection system can detect a drone by monitoring the communication frequency spectrum. There are a few RF-based drone detection techniques proposed in the literature [14]–[19]. In [14], the presence of a drone is detected by monitoring how frequently the data packet is being transmitted at 2.4 GHz. Since most of the drones use different non-standardized protocols for their communication with their controller, the data packet transmission rate varies from WiFi and other Access Points (AP) [14]. In [15], the detection is performed by measuring the data packet length of a drone's communication link. These detection methods are inefficient since the detector can be easily spoofed by an application communicating with an AP with the same packet transfer rate or having the same packet length as a drone. In [16], a WiFi based drone surveillance method is proposed, where the identification is performed by a WiFi statistical fingerprinting technique. In [20], we observed that several commercial drone's GCS uses Frequency Hopping Spread Spectrum (FHSS) transmission as the radio control (RC) signal, which should also be accounted for in the identification method.

The detection and identification of a drone using the frequency signatures is presented in [17], [18], using Deep Neural Network (DNN) based classifiers. In [17], the author developed a dataset using three commercial drones and used a simple feedforward DNN to detect and identify them. In [18], the author presented the detection, identification and classification on the same dataset using a Convolutional Neural Network (CNN). These studies were performed on a limited dataset and the impact of noise on the detection performance was not studied. Moreover, the detection performance in presence of multiple signals or interference was not investigated.

We presented an RF-based drone detection using GoF spectrum sensing and DoA estimation using MUSIC algorithm in [21]. Drone signal detection using wideband CFAR-based energy detection and the feature extraction performance is presented in [20]. In [22], we presented drone signal classification using a DRNN framework. The classification was performed assuming a signal is already detected by a spectrum sensing algorithm. A complete solution for drone detection and classification based on RF fingerprints was not presented in our previous works, which we address in this paper. We propose two complete solutions for drone signal detection and classification, and provide an in-depth performance comparison. The detection and classification is performed in two different ways: (i) Two-stage detection and classification process: where the signal detection is initially performed using an efficient spectrum sensing method. This will detect all of the signals present in the spectrum. The detected signals are passed to a SoA classifier to provide a robust classification. (ii) Combined detection and classification: where the signals are detected and classified simultaneously. For both proposed methods, we perform detection and classification with the received signal from a single receiver and by using frequency domain fingerprints. The advantages are (i) The possibility of using the received signal from a single receiver eliminates the requirement of calibration of multiple receivers. This makes both methods easily deployable with a low cost SDR and a computational unit. (ii) The frequency domain detection and classification provide the necessary information for a possible RF jammer. Both methods can perform fast detection and classification, even in the presence of overlapping signals both in the time and frequency domain. They are more generalized and robust as they understand the position and type of signal.

The main contributions of this paper are the following.

1) A novel and realistic multi-signal dataset is created using nine commercial drones and non-drone signals (i.e., WiFi communication signals). The dataset will be made public for future research.
2) The YOLO-lite architecture is recreated from scratch and modified to perform the combined drone signal detection and classification. The two-stage detection and classification is performed using GoF spectrum sensing and DRNN classifier.
3) The simultaneous multi-signal detection, spectrum localization and classification in the ISM band is presented in this paper. We are the first to propose a framework for simultaneous multi-signal drone detection and classification.
4) The detection and classification performance of both frameworks are evaluated on our dataset. Through detailed comparisons, we show that the YOLO-lite framework provides better detection performance compared to the GoF sensing and a good classification performance comparable to the DRNN classifier.

The rest of the paper is organized as follows: A mathematical model of the received signal in an ISM band is presented in Section II. Section III provides an overview of the SoA techniques for two-stage and combined signals detection and classification. The technical details to perform two stage, and combined detection and classification is presented in Section IV. The dataset development and the experiment strategies are explained in Section V. The performance analysis is presented in Section VI and the concluding remarks are provided in Section VII.

II. Problem Statement

The ISM bands are generally populated by several homogeneous and heterogeneous RF transmissions. The transmitters generally use spread spectrum technology to perform the

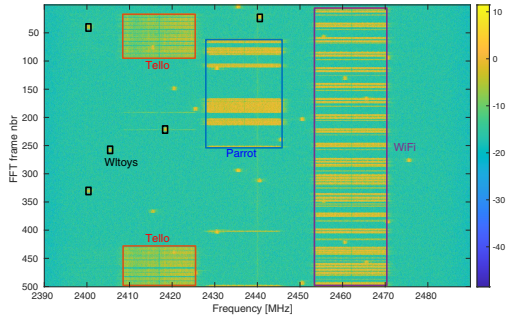


Fig. 1. Presence of multiple signals at 2.4 GHz ISM band.

communications. The FHSS transmissions are generally blind whereas the Direct Sequence Spread Spectrum (DSSS) transmissions are often cognitive in nature. Most commercial drones use DSSS signal for video signal transmission. Unlike FHSS transmissions, they perform sensing to find a free (or relatively free) channel before starting to transfer video signal. One example of such heterogeneous transmission at 2.4 GHz is shown in Fig. 1. As it can be seen from the figure, four transmissions are occurring at the same time where three transmitters use the DSSS technology and one transmitter uses the FHSS technology.

The received signal from a drone can be expressed as:

$$r(t) = \sum_{k=1}^{k=K} y_k(t) * h_k(t) + n(t), \tag{1}$$

where, $y_k(t)$ is complex baseband transmitted signal, $h_k(t)$ is the time varying impulse response, k denotes the index of transmitted signal, K is the total number of available transmitters in the ISM band, $n(t)$ is the Additive White Gaussian Noise (AWGN). With the help of discrete Fourier transform (DFT), the complex time domain receive signal can be converted into M_t consecutive segments, each with length N_f (= FFT size). The magnitude of the DFT matrix gives us the spectrogram matrix. This study aims to detect and classify drone and WiFi communication signals from the spectrogram matrix. The spectrogram representation provides more information compared to a Power Spectral Density (PSD) or IQ representation of the signal. It enables the classifier to determine useful RF signal features like frequency, bandwidth, dwell time and hop rate. Since commercial drones use pseudo-random number generators to generate the communication signals, the hopping pattern or the signal position will vary. The objective of our work is not to learn the hopping pattern, but rather to learn how the data/signal is distributed in the spectrum to detect and classify them. Deep Learning (DL) algorithms learn the signal distribution which is dependent upon factors like: (i) frequency, (ii) bandwidth, (iii) modulation, (iv) filtering parameters (v) device nonlinearities etc. We aim to utilize the DL algorithm to learn these factors from the spectrogram matrix.

III. Background

A. Signal Detection

The conventional spectrum sensing methods can be classified into parametric and non-parametric methods. The parametric methods require prior knowledge of the transmitted signal for the detection, whereas the non-parametric methods also known as blind sensing do not require any knowledge of the signal. Most popular spectrum sensing methods are energy detection, eigenvalue detection, matched filter and cyclostationary feature detection [23]. The energy detection and eigenvalue detection are non-parametric methods. The energy-based detection have been widely used for decades mainly due to its simplicity. The signal is detected if the measured energy is greater than the threshold corresponding to the specified false alarm rate. The eigenvalue detection method estimates the ratio of maximum and minimum eigenvalue and compares it with the threshold to determine if any signal is present. The cyclostationary and matched filtering are parametric methods, they require perfect knowledge of the transmitted signal and can work better at lower signal-to-noise ratio (SNR) compared to the energy-based detection [24]. The cyclostationary spectrum sensing method exploits the periodicity introduced in the transmitted signal. The matched filtering method detects a signal by correlating a known template (extracted from the transmit signal) with the received signal. Both methods are targeted towards specific (or known) signals and require a high computational cost.

For the drone signal detection within two-stage detection and classification, we choose wideband GoF based blind spectrum sensing algorithm [25]. The GoF sensing can provide better performance compared to the conventional energy detection using fewer samples of the received signal, at low SNR conditions and in presence of non-Gaussian noise [26]. The wideband GoF sensing uses DFT to divide the frequency band into small frequency bins and perform narrowband GoF sensing on each bin. In this paper, we have used Anderson-darling test statistic for the GoF sensing [25].

B. Signal Classification

DL methods have shown SoA performances in the classification of wireless signals and outperformed the conventional classification methods. Some remarkable works have been published [27], [28] in the past few years regarding the classification of modulated signals and device fingerprinting using merely the raw received signals. In the recent years, CNN frameworks have been widely investigated for wireless signal recognition and classification problems [27], [29]. Among different variants of CNNs, the residual network-based CNN [30] have shown great performance and outperformed other classifiers with equivalent network depth. In this paper, we adapt the DRNN proposed in [27] for the drone and WiFi signal classification.

C. Combined Detection and Classification

The DNN based visual object detection and classification techniques provide great tools such as YOLO [31] for the combined RF signal detection, frequency localization and classification. Signal detection/classification from a spectrogram image is analogous to the visual object detection and classification. A spectrogram image provides the time and frequency information of a spectrum instance, which can be utilized

to perform the detection and classification. Since we are interested in wideband signals, the localization and bounding box will also enable to determine important features from the detected signal, such as center frequency, bandwidth, dwell time and hop interval. Such information can be used within a cognitive radio to perform dynamic spectrum access functionality and avoid collisions in a spectrum sharing environment. YOLO was first used in [32] to perform signal detection and frequency localization. In [33], WiFi and LTE signal detection, feature extraction and classification was performed using a pre-trained YOLO framework. In this paper, we develop a YOLO framework from scratch to perform the combined drone RF signal detection and classification on the spectrogram image.

IV. Technical Approach

A. Two Stage Detection and Classification

1) Signal Detection: The signal detection is performed using the Anderson-Darling (AD) GoF test [25]. The complex time-domain receive signal is converted into the frequency domain using an N-point DFT operation on K consecutive segments. This results in a sequence of X_k of length K for every frequency bin. We perform a hypothesis test using an AD tester for each frequency bin to decide whether only noise or a signal is present in the bin. We assume, there is only noise present in the frequency bin if the normalized power spectral coefficient ($\frac{2|X_k|^2}{N\sigma^2}$) follows a χ^2 distribution. The length of DFT is N and σ^2 is the noise power. We estimated the noise power by exploiting the histogram of the power spectrum [34].

The AD test statistic (A_n^2) is calculated for each frequency bin as:

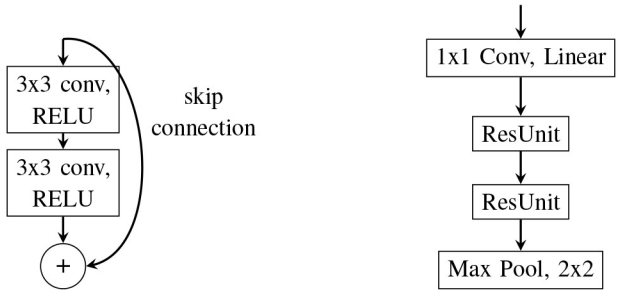
$$A_n^2 = -n - \frac{\sum_{i=1}^{n}(2i-1)\Big(\ln z_i + \ln\Big(1 - z_{(n+1-i)}\Big)\Big)}{n} \quad (2)$$

with $z_i = F_o(x_i)$.

Here, F_o represents the Cumulative Distribution Function (CDF) of a chi-square distribution with 2 degrees of freedom, $x_1 \leq x_2 \leq \cdots \leq x_n$ are the samples under test and n represents the total number of samples.

If $A_n^2 > \lambda$, we assume the signal is present in the frequency bin, otherwise, we assume that the frequency bin only contains noise. Here, λ corresponds to the detection threshold. A detailed explanation of the AD GoF sensing, the derivation of A_n^2 and the procedure to perform the hypothesis test on complex received signals from a Software Defined Radio (SDR) are provided in [25], [26]. The value of λ is determined considering 5% False Alarm Rate (FAR). The suitable value for λ was calculated through the AD GoF tests on the drone dataset. The value of $\lambda = 3.89$ provided us an approximate 5% FAR.

2) Signal Classification: The classification is performed using an adaptation of the DRNN framework proposed in [27]. The architecture is depicted in Table I and the building blocks are shown in Fig. 2. The DRNN framework consists of N residual stack units, two fully connected (FC) layers and a softmax layer. For all convolution operations, we have used 32 filters with a kernel size of 3x3, apart from the first layer of the residual stack where the kernel size is 1x1. For the max pooling, a kernel size of 2x2 is used with a stride factor of



(a) ResUnit (b) Residual stack

Fig. 2. Building blocks of the deep residual network architecture.

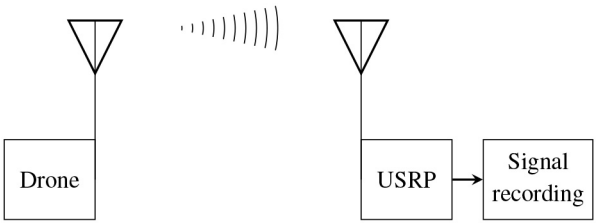


Fig. 3. Signal recording schematic.

TABLE I
DRNN Architecture

Layer no.	Layer name	Output dimensions
1.	Input	256 x 256
2.	Residual stack	128 x 128 x 32
3.	Residual stack	64 x 64 x 32
4.	Residual stack	32 x 32 x 32
5.	Residual stack	16 x 16 x 32
6.	Residual stack	8 x 8 x 32
7.	FCL/SELU	128
8.	FCL/SELU	128
9.	FCL/Softmax	10

1. For each FC layer, we have employed a scaled exponential linear unit (SeLU) activation and mean response scaled initialization [35]. To prevent overfitting, we have performed 50% dropout after each FC layer. A softmax activation is used at the final layer to give the prediction probability. We haven't performed any batch normalization operation, since, we did not observe any additional improvement with it in the classification performance.

If any signal is detected from the spectrum sensing algorithm, the complete spectrum (i.e., the time domain RX signal) is passed to the classification stage. The time-domain signal is converted to a spectrogram signal of size 256 x 256 and the classification is performed on the signal. Since we are interested in comparing the classification performance with the YOLO framework, we kept the input size the same for both algorithms.

B. Combined Detection and Classification With YOLO

We implement one of the variants of the YOLO architecture to perform the simultaneous detection and classification of drone and WiFi communication signals. One of the biggest strengths of the YOLO framework is that it can detect the signal, determine spectral features like frequency, bandwidth, dwell time and predict the class of the detected signal simultaneously. The raw spectral power values of the RF signal

TABLE II
YOLO NETWORK ARCHITECTURE

Layer	Filters	Size/Stride
Conv2D (C1) + Max Pool (MP)	16	3x3/1 + 2x2/2
C2 + MP	32	3x3/1 + 2x2/2
C3 + MP	64	3x3/1 + 2x2/2
C4 + MP	128	3x3/1 + 2x2/2
C5 + MP	128	3x3/1 + 2x2/2
C6	256	3x3/1
C7	125	1x1/1
FCL/Sigmoid activation		

in time and frequency domain are used as the input. Since recognizing such time-frequency domain spectral events are relatively simpler than the visual object recognition [32], a smaller network may be sufficient for the YOLO detection and classification task. In our experiments, we have adapted the YOLO-lite [36] architecture, which is a smaller and faster network and can be deployed on a non-GPU computer. Our adaptation of the YOLO-Lite architecture is shown in Table II. We have used leaky-relu activation after each convolution operation (C1 - C6) and linear activation on the C7 layer. The max-pooling operation is performed after the convolutions (C1 - C5). Finally, a fully connected layer is employed, and sigmoid activation is performed.

A spectrogram dataset with a dimension of 256x256 is used as the input. The network produces an output grid containing the detection probability, bounding box coordinates and class probabilities as shown following:

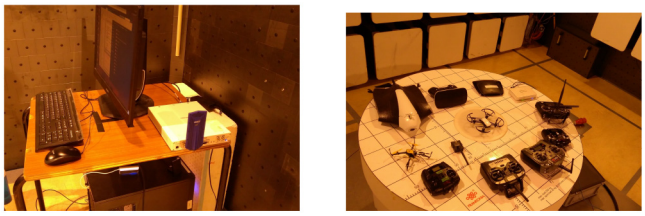
$$\mathrm{O/Pshape} = \mathrm{SxSx(Bx[C, x, y, w, h] + P)} \tag{3}$$

Here, S denotes the size of the grid. Each grid cell contains B bounding boxes, the confidence score of the detection: c, 2D coordinates: x, y, width: w, height: h of the object and the class probability P. We have used grid size 16, 2 bounding boxes and 10 different classes (Table III) for our tests. In the original YOLO-lite architecture output grid size of 8 was used, however, we found during our tests that in order to annotate the DSSS spectrograms (e.g., Tello, Parrot), a higher number of grid size is required. With the above specified parameters, the output dimension becomes 16 x 16 x 20. We have used adam optimizer [37] to optimize the training loss. The training loss involves the minimization of the sum of mean squared error loss between the ground truth and network prediction. The complete training loss function provided in [31] is used for the training optimization.

V. EXPERIMENTS

A. Experimental Setup

The drone and WiFi signals were recorded in an anechoic chamber. For this experiment, we only considered the transmitters operating at 2.4 GHz. A universal software radio peripheral (USRP) X310 was used with an omnidirectional antenna. The receive sampling rate of 100 MSps was used to receive instantaneously from the complete 2.4 GHz ISM band. Nine commercial radio controllers with drones and two WiFi routers (Table III) were used to develop the dataset. The devices were placed seven meters apart from the receiver.



(a) Signal recording wih USRP (b) Drones, GCS and WiFi routers

Fig. 4. Test setup in an anechoic chamber.

TABLE III
DRONES, RADIO CONTROLLERS AND WIFI SOURCES USED FOR THE DATASET DEVELOPMENT

Nr.	Name	Signal Type	Freq(GHz)
1	Parrot Disco	RC+Video	2.4
2	Q205	Video	2.4
3	Tello	RC+Video	2.4
4	MultiTx	RC	2.4
5	Nineeagles	RC	2.4
6	Spektrum DX4e	RC	2.4
7	Spektrum DX6i	RC	2.4
8	Wltoys	RC	2.4
9	Taranis Q X7 (S500)	RC	2.4
10 (i)	Linksys router	IEEE802.11b/g	2.4
10 (ii)	Netgear router	IEEE802.11n	2.4

Since a UAV controller generally uses a pseudorandom generator to generate the FHSS sequences as the RC signal, we included all possible hop sequences from each controller in our database. The controllers were turned off and on several times during data collection to inspect if the hop position changes and include that as well in our database.

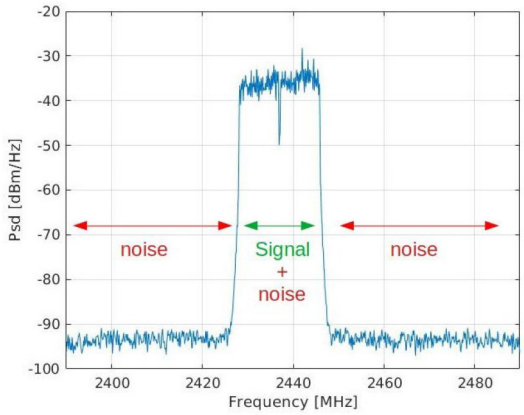
B. Dataset Development

To test the classification performance at lower SNRs, we introduced AWGN to the signal in the simulation environment. Generally, the SNR is calculated in the time domain by measuring the transmission power of the signal. Since the signal bandwidth of different drones is different, it becomes difficult to calculate the SNR in the time domain. Therefore, we calculated the signal SNR in the frequency domain as shown in Fig. 5(a). Since the bandwidth and transmission power are different for different drones, the calculated SNR was also slightly different for different noise values as presented in Fig. 5(b). To keep the performance analysis of different frameworks on the dataset simple, we considered the average SNR for the introduced AWGN values as shown in Fig. 5(c).

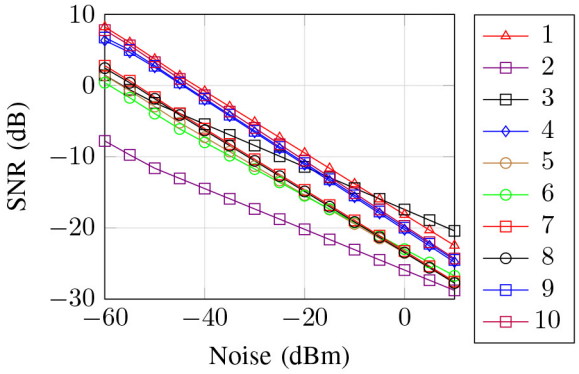
In [22], we evaluated the classification performance in Rician and Rayleigh fading simulation environment where the classifier was trained with AWGN faded dataset. We did not observe any significant deviation in the classification performance due to the channel variation. Therefore, in this paper, we only evaluated the classification performance under AWGN conditions.

C. Implementation Details

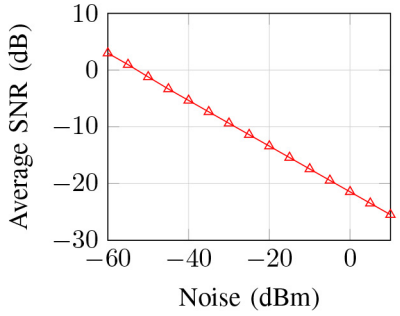
The GoF sensing was implemented in MATLAB. The DRNN framework was implemented using Tensorflow-Keras and the YOLO-lite framework was implemented using TFlearn



(a) SNR calculation in frequency domain



(b) Relation between noise vs SNR. The legends correspond to the device no. given in Table III



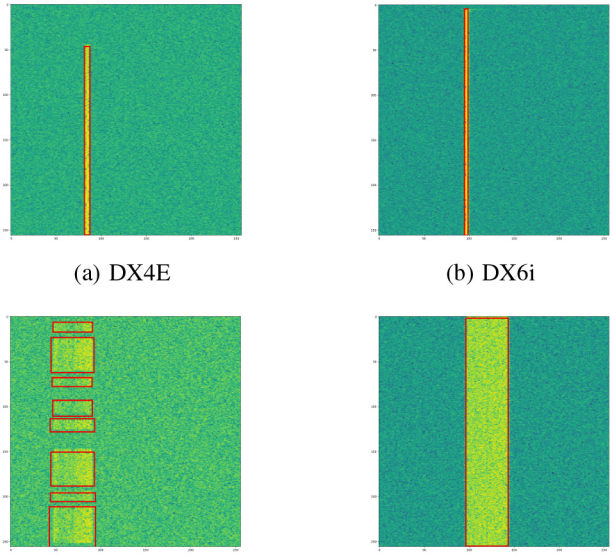
(c) Noise vs average SNR

Fig. 5. SNR variation using AWGN in the simulation environment.

in Python, both running on top of Tensorflow [38]. The simulations and the neural network training and testing were performed on an Intel Core i7 computer equipped with Nvidia RTX 2080 GPU. The adam optimizer with a learning rate of 0.001 was used for the training optimization for both networks. The training was performed with a batch size of 32.

VI. Performance Analysis

The performance analysis was performed for two scenarios: (i) detection and classification in presence of one signal at a time under AWGN conditions, (ii) detection and classification in presence of multiple signals simultaneously with frequency overlapped and non-overlapped cases.



(a) DX4E (b) DX6i

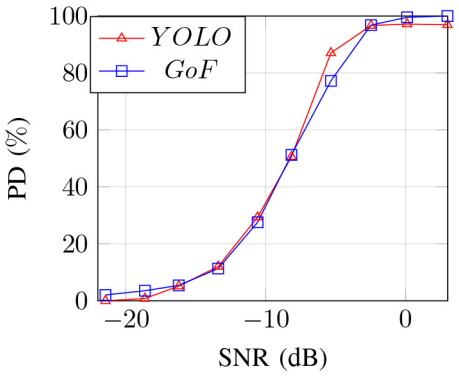
(c) Tello (d) Parrot

Fig. 6. Signal detection with YOLO-Lite.

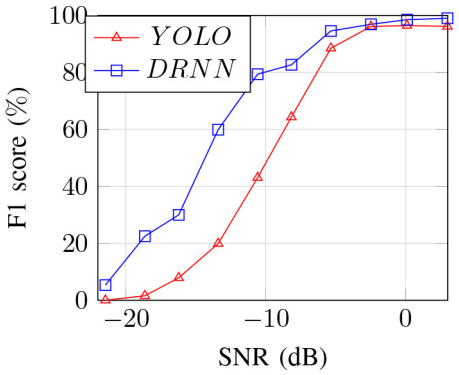
A. Single Signal Detection and Classification

Signal detection and spectrum localization with the YOLO-lite framework are presented in Fig. 6. As Fig. 6 shows, the signals were detected, localized in the spectral domain and classified by the framework accurately. The average probability of detection (PD) of YOLO under different SNR conditions is presented in Fig. 7(a). A detection from the YOLO-Lite prediction is considered to be true if it satisfies the following two conditions: (i) the confidence score of any bounding box is greater than the specified threshold (i.e., $C > 0.4$) and (ii) the Intersection over Union (IoU) is greater or equal to 0.50 (i.e., IoU ≥ 0.50). The PD for the GoF test is calculated by comparing the true frequency bins with the predicted frequency bins from the AD test. As Fig. 7(a) displays, the PD from YOLO is comparable with the GoF test. The detection probability increases for both frameworks as the SNR increases. The YOLO PD saturates around 96%, where it reaches at around -3 dB SNR. This saturation happens because YOLO often could not detect all closely spaced hops of signals from Tello, Parrot and WiFi. One example of such a spectrum is shown in Fig. 6(c). The GoF test also provides around 96% PD around SNR -3 dB, however, it increases further and reaches 99.9% at SNR 3 dB.

To evaluate the classification performances, we used the F1-score parameter, which is a harmonic mean of precision and recall. Since the F1-score takes both precision and recall into account, it allows us to compare the performance of different classifiers using just one metric. The performance of classification with the YOLO and DRNN framework are plotted in Fig. 7(b). In order to compare the classification performance of YOLO with the DRNN, we performed the classification with the DRNN framework independent of the GoF detection. At lower SNR, DRNN provides a better F1-score compared to YOLO. This is expected since YOLO-lite is much shallower framework compared to the DRNN framework. The



(a) Detection performance



(b) Classification performance

Fig. 7. Detection probability and classification performance under AWGN conditions with YOLO and DRNN framework. Train samples: 4.9k, Test samples: 2.1k.

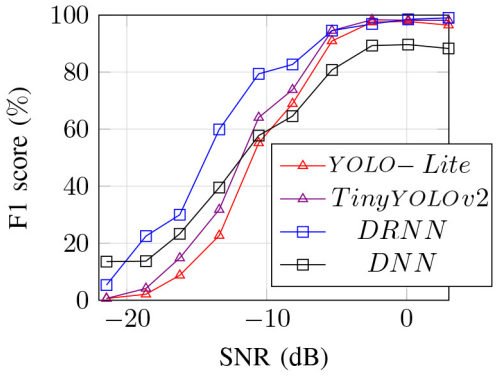


Fig. 8. Classification comparison with other frameworks under AWGN conditions. Train samples: 4.9k, Test samples: 2.1k.

F1-scores increases with the increase in SNR for both frameworks. The F1-score reaches approx. 97% at SNR −3 dB for both frameworks. The F1-score from the YOLO-lite saturates around 97% at higher SNRs. The F1-score from the DRNN framework increases to 99% at SNR 3 dB.

The classification performance of our YOLO-Lite and DRNN framework is compared with other existing frameworks namely the DNN framework proposed in [17] and the Tiny-YOLOv2-VOC framework [36]. The F1-scores of the classification performance under AWGN conditions are plotted in Fig. 8. The DRNN framework provided the best classification

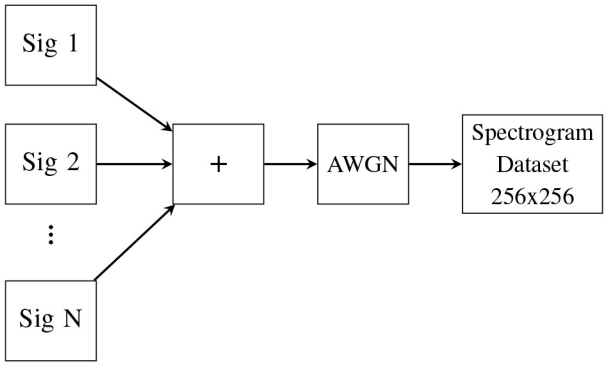
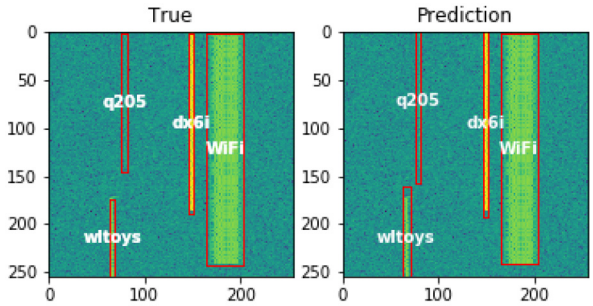
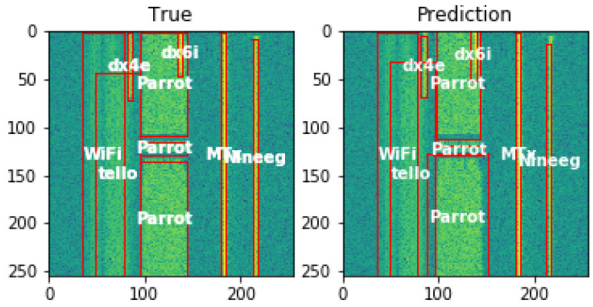


Fig. 9. Simulation schematic for multi-signal detection and classification.



(a) Detection and classification in presence of 4 signals



(b) Detection and classification in presence of 7 signals

Fig. 10. Detection and classification with YOLO in presence of multiple signals.

performance compared to the other frameworks. The Tiny-YOLOv2 provided slightly better classification performance compared to the YOLO-Lite framework. This is expected since it is a slightly deeper framework compared to the YOLO-lite framework. We also recreated Tiny-YOLOv2 from scratch for this classification test. The DNN model provided lower F1-score compared to other frameworks from SNR −10 dB and onwards. The F1-score of this framework was saturated at around 89%, whereas YOLO-lite, Tiny-YOLOv2 and DRNN provided around 97%, 98% and 99% F1-score respectively at higher SNRs.

B. Simultaneous Multi-Signal Detection and Classification

In order to test the detection and classification performance in a simultaneous multi-signal scenario, the signals were added in the simulation environment as shown in Fig. 9. To ensure

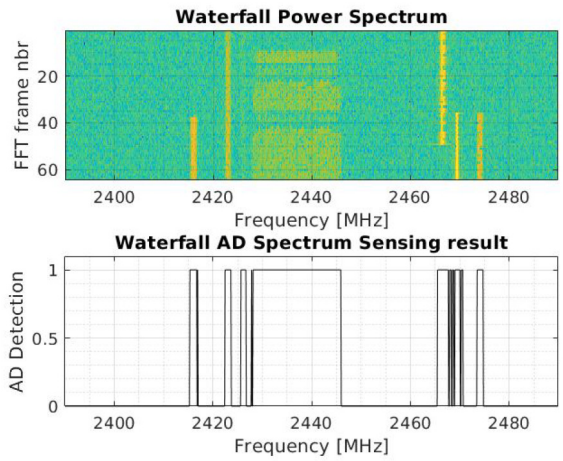


Fig. 11. Signal detection using GoF spectrum sensing in presence of multiple signals.

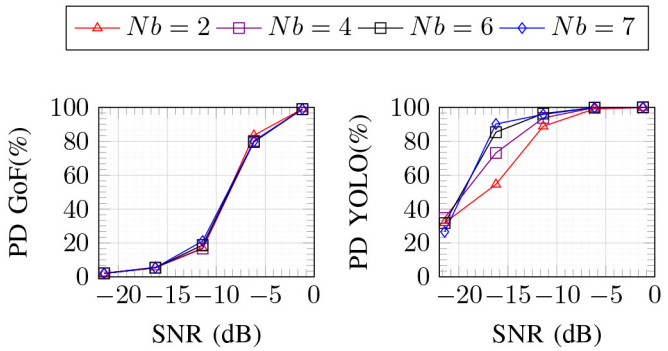


Fig. 12. Detection probability under AWGN conditions in presence of multiple-signals simultaneously. Train samples: 30.7k, Test samples: 13.2k.

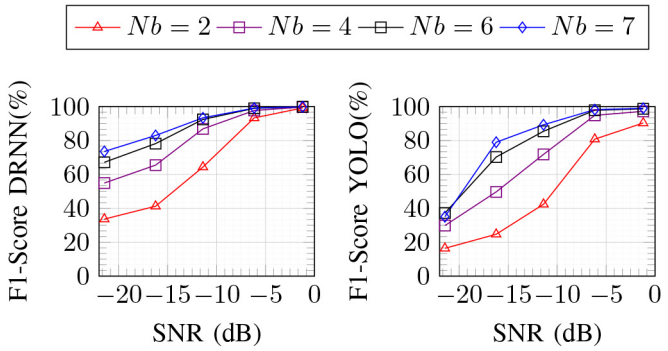


Fig. 13. Classification performance under AWGN conditions in presence of multiple-signals simultaneously. Train samples: 30.7k, Test samples: 13.2k.

the number of specified signals present in the spectrogram, we performed spectrum sensing before adding the RX signals. Each signal burst was a vector of 256*256 = 65536 complex samples in the time domain. Later AWGN was introduced after adding the signals and it was converted to the frequency domain. The detection and classification from YOLO-lite is shown in Fig. 10. In order to observe the prediction accuracy, the ground truth and prediction are plotted side by side and the true and predicted classes are also annotated in white color. As it can be seen from Fig. 10, the signals were detected, localized in spectral domain and classified accurately. YOLO-lite could also detect and classify signals which are overlapped in the frequency domain. It can be observed from Fig. 10(b), that WiFi and tello signals are overlapped in frequency domain. Similarly, DX6i and Parrot signals are also overlapped in the frequency domain. All of the signals are localized and classified accurately.

The multi-signal detection from the GoF sensing is shown in Fig. 11. The spectrogram image is plotted in Fig. 11 (top) and the AD GoF result is plotted in Fig. 11 (bottom). As it can be seen from the figure, that all seven signals are detected correctly.

The detection threshold for YOLO-lite was chosen to be 0.4. This threshold was chosen such that the maximum FA rate for the lowest SNR remains below 5%. The threshold was kept the same for single and simultaneous multi-signal detection.

The PD of the GoF sensing and YOLO-lite over different SNRs are plotted in Fig. 12. As it can be seen from the figure, YOLO-lite showed better detection performance compared to the GoF sensing. For GoF sensing, the detection performance over different SNRs did not vary depending on the number of sources. On the contrary, an increase in the detection performance was observed with YOLO-lite. If we look at Fig. 5(b), the SNRs for different signals are different. With wideband GoF sensing, if we have multiple signals with different SNR present in the spectrum, it impacts the noise floor estimation. This may result in an overestimation of the threshold. This issue was not observed with the YOLO-lite spectrum sensing.

The classification performance of DRNN and YOLO over different SNRs are plotted in Fig. 13. Similar to the PD, the F1-score of the YOLO classification increases with the increase in the number of sources. We can also observe the same phenomena with the F1-score of the DRNN classification. After detailed investigation, we found that the classification accuracy of any individual signal remains the same for single signal or simultaneous multi-signal scenario. At lower SNRs, the classification accuracy for different signals are generally different. There are two reasons behind this: (i) the actual SNR of the signals are not the same (Fig. 5(b)) (ii) classifiers can classify some signals better than other signals at lower SNRs. Fig. 5(b) shows that the actual SNR of the signals is different for different classes. As it can be observed, one of the signals has a very low SNR compared to the other signals. This is because we added a constant AWGN noise and considered the average SNR for ease of analysis. When we calculate the average F1-score for the single-signal scenario, it averages equally having independent F1 scores for each class. However, for the multi-signal scenario, as the number of signals increases, the relatively lower F1 score of a particular class cannot make the average performance as worse as the single-signal scenario. Therefore, as the number of signals in the multi-signal scenario increases the average score also goes higher. Again, similar to the single drone scenario, DRNN showed better classification performance compared to YOLO.

C. Comparison Summary

The comparisons are summarized here.

- *Signal Detection:* We have obtained better signal detection performance with YOLO compared to the GoF spectrum sensing on our dataset. At lower SNR regions, the GoF sensing showed high false alarm rate compared to the YOLO detection. Since YOLO is a supervised detection framework, the performance may deviate while detecting unknown signals. This issue can be resolved with transfer learning using a small labeled dataset of the new signal. On the contrary, the GoF sensing is a blind spectrum sensing method, it is able to detect any signal present in the spectrum.
- *Signal Classification:* The DRNN framework provided better classification performance compared to the YOLO framework. It is expected from a deep residual network since it utilizes the skip connection feature in the architecture and the network is deeper than the YOLO framework.
- *Signal Localization and Feature Extraction:* The signal localization and feature extraction is the best feature of the YOLO framework. The localization feature of YOLO enables us to detect multiple signals simultaneously and extract the useful features from the received signal. It can provide the center frequency, bandwidth, hop rate and dwell time of the detected signal. This information is required for an RF jammer to perform soft neutralization of a drone. With the two-stage detection and classification framework, it is not possible to extract all of these features. The DRNN framework does not provide the spectral position of the signal under classification. The GoF sensing provides the frequency and bandwidth of the signal, however, in presence of multiple signals, it is going to be difficult to associate the signal features with the classification label.
- *Complexity Analysis:* A complexity analysis was performed to give an overview of the computational complexity and the inference time required for each framework. The total network parameter (that is, trainable + non trainable parameter), the mean inference time on total test samples and the mean prediction time per sample are presented in Table IV. The inference time calculation was performed with the dataset used in the single signal detection scenario, where the test sample size was 2.1k samples. We used a batch size of 100 and performed the test on Nvidia GeForce RTX 2080 Ti GPU. As it can be seen from Table IV, YOLO-lite is approximately 3.4 times faster than the DRNN framework. The GoF spectrum sensing is a computationally simpler algorithm compared to the DRNN and YOLO-lite frameworks. It does not require any high-end computational unit. During the SafeShore [39] project, we implemented the GoF sensing using C++ on an Odroid-XU4 platform and performed real-time tests.
- *Limitation:* One of the limitations of both proposed methods is the classification of completely unknown signals. Since the classification is performed in a supervised manner, the classifier may not be able to classify or provide a label to the transmitted signals from newer drones. There are two possible outcomes in such a case: (i) classifier will label it as an existing drone signal if the TX signal has a similar frequency fingerprint, (ii) classifier will be confused and provide a very low classification score for all classes. Similarly, some specific models of UAV controllers may use a completely different hopping sequence compared to another controller under the same model. The YOLO detection and classification performance for such cases are not yet tested. We are going to investigate and address these issues in our future work.

TABLE IV
COMPLEXITY COMPARISON

Parameter	DRNN	YOLO Lite
Total parameters	469,322	41,537,149
Mean inference time	3.98 s	1.16 s
Mean prediction time per sample	1.89 ms	0.552 ms

VII. CONCLUSION

In this paper, we performed drone signal detection, spectrum localization and classification using two stages and combined detection and classification methods. Under the two-stage technique, we used the GoF sensing for the detection and the DRNN framework for the classification. The YOLO-lite framework was recreated from scratch to perform the combined drone RF signal detection, spectrum localization and classification. A detailed performance comparison between both of the techniques is presented using a novel drone dataset that was prepared for this study. We obtained good detection and classification performances with both techniques. Since the classification is performed in a supervised manner, the performance may deviate in presence of unknown or newer drone signals which we mentioned in detail in the limitation discussion. In the future work, we are going to investigate the unsupervised scenarios, since we are interested in developing a robust framework that can detect and classify all drone signals irrespective of the dataset it is trained with.

REFERENCES

[1] D. Sathyamoorthy, "A review of security threats of unmanned aerial vehicles and mitigation steps," *J. Defence Security*, vol. 6, no. 1, pp. 81–97, 2015.
[2] B. Taha and A. Shoufan, "Machine learning-based drone detection and classification: State-of-the-art in research," *IEEE Access*, vol. 7, pp. 138669–138682, 2019.
[3] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on the cognitive Internet of Things," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 29–35, Jan. 2018.
[4] F. Fioranelli, M. Ritchie, H. Griffiths, and H. Borrion, "Classification of loaded/unloaded micro-drones using multistatic radar," *Electron. Lett.*, vol. 51, no. 22, pp. 1813–1815, 2015. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/el.2015.3038
[5] J. Drozdowicz *et al.*, "35 GHz FMCW drone detection system," in *Proc. 17th Int. Radar Symp. (IRS)*, 2016, pp. 1–4.
[6] A. Coluccia, G. Parisi, and A. Fascista, "Detection and classification of multirotor drones in radar sensor networks: A review," *Sensors*, vol. 20, no. 15, p. 4172, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/15/4172
[7] Z. Zhang, Y. Cao, M. Ding, L. Zhuang, and W. Yao, "An intruder detection algorithm for vision based sense and avoid system," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, 2016, pp. 550–556.
[8] S. R. Ganti and Y. Kim, "Implementation of detection and tracking mechanism for small uas," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, 2016, pp. 1254–1260.

[9] R. Stolkin, D. Rees, M. Talha, and I. Florescu, "Bayesian fusion of thermal and visible spectra camera data for mean shift tracking with rapid background adaptation," in *Proc. IEEE SENSORS*, 2012, pp. 1–4.
[10] A. Coluccia *et al.*, "Drone-vs-bird detection challenge at IEEE AVSS2019," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveillance (AVSS)*, 2019, pp. 1–7.
[11] M. Nijim and N. Mantrawadi, "Drone classification and identification system by phenome analysis using data mining techniques," in *Proc. IEEE Symp. Technol. Homeland Security (HST)*, 2016, pp. 1–5.
[12] M. Benyamin and G. H. Goldman, *Acoustic Detection and Tracking of a Class I UAS with a Small Tetrahedral Microphone Array*, Army Res. Lab., Adelphi, MD, USA, Sep. 2014.
[13] J. Busset *et al.*, "Detection and tracking of drones using advanced acoustic cameras," in *Proc. Unmanned/Unattended Sens. Sens. Netw. XI Adv. Free-Space Opt. Commun. Techn. Appl.*, vol. 9647, Oct. 2015, Art. no. 96470F.
[14] P. Nguyen, M. Ravindranatha, A. Nguyen, R. Han, and T. Vu, "Investigating cost-effective RF-based detection of drones," in *Proc. 2nd Workshop Micro Aerial Veh. Netw. Syst. Appl. Civilian Use*, 2016, pp. 17–22. [Online]. Available: https://doi.org/10.1145/2935620.2935632
[15] P. Kosolyudhthasarn, V. Visoottiviseth, D. Fall, and S. Kashihara, "Drone detection and identification by using packet length signature," in *Proc. 15th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, 2018, pp. 1–6.
[16] I. Bisio, C. Garibotto, F. Lavagetto, A. Sciarrone, and S. Zappatore, "Blind detection: Advanced techniques for WiFi-based drone surveillance," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 938–946, Jan. 2019.
[17] M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database," *Future Gener. Comput. Syst.*, vol. 100, Nov. 2019, pp. 86–97.
[18] S. Al-Emadi and F. Al-Senaid, "Drone detection approach based on radio-frequency using convolutional neural network," in *Proc. IEEE Int. Conf. Inform. IoT Enabling Technol. (ICIoT)*, 2020, pp. 29–34.
[19] M. M. Azari, H. Sallouha, A. Chiumento, S. Rajendran, E. Vinogradov, and S. Pollin, "Key technologies and system trade-offs for detection and localization of amateur drones," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 51–57, Jan. 2018.
[20] P. Stoica, S. Basak, C. Molder, and B. Scheers, "Review of counter-uav solutions based on the detection of remote control communication," in *Proc. 13th Int. Conf. Commun. (COMM)*, 2020, pp. 233–238.
[21] S. Basak and B. Scheers, "Passive radio system for real-time drone detection and DoA estimation," in *Proc. Int. Conf. Military Commun. Inf. Syst. (ICMCIS)*, May 2018, pp. 1–6.
[22] S. Basak, S. Rajendran, S. Pollin, and B. Scheers, "Drone classification from RF fingerprints using deep residual nets," in *Proc. Int. Conf. Commun. Syst. Netw. (COMSNETS)*, 2021, pp. 548–555.
[23] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.
[24] I. Kakalou, D. Papadopoulou, T. Xifilidis, K. E. Psannis, K. Siakavara, and Y. Ishibashi, "A survey on spectrum sensing algorithms for cognitive radio networks," in *Proc. 7th Int. Conf. Modern Circuits Syst. Technol. (MOCAST)*, 2018, pp. 1–4.
[25] B. Scheers, D. Teguig, and V. Le Nir, "Wideband spectrum sensing technique based on goodness-of-fit testing," in *Proc. Int. Conf. Military Commun. Inf. Syst. (ICMCIS)*, 2015, pp. 1–6.
[26] D. Teguig, B. Scheers, V. Le Nir, and F. Horlin, "Spectrum sensing method based on the likelihood ratio goodness of fit test under noise uncertainty," *Int. J. Eng. Res. Technol.*, vol. 3, no. 9, pp. 488–494, 2014.
[27] 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.
[28] 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, Sep. 2018.
[29] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 146–152, Sep. 2018.
[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.
[31] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788.
[32] T. O'Shea, T. Roy, and T. C. Clancy, "Learning robust general radio signal detection using computer vision methods," in *Proc. 51st Asilomar Conf. Signals Syst. Comput.*, 2017, pp. 829–832.
[33] E. Fonseca, J. F. Santos, F. Paisana, and L. A. DaSilva, "Radio access technology characterisation through object detection," *Comput. Commun.*, vol. 168, pp. 12–19, Feb. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366420320272
[34] S. Couturier and D. Rauschen, "Energy detection based on long-term estimation of Gaussian noise distribution," in *Proc. 8th Karlsruhe Workshop Softw. Radios*, 2014, pp. 89–95.
[35] S. Ren, K. He, X. Zhang, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. ICCV*, 2015, pp. 1026–1034.
[36] R. Huang, J. Pedoeem, and C. Chen, "YOLO-LITE: A real-time object detection algorithm optimized for non-GPU computers," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, 2018, pp. 2503–2510.
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2015. [Online]. Available: arXiv:1412.6980.
[38] M. Abadi *et al.* (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [Online]. Available: https://www.tensorflow.org/
[39] European Commission. *Horizon2020. The SafeShore Project*. [Online]. Available: http://safeshore.eu

Sanjoy Basak received the M.Sc. degree in electrical engineering and information technology from the Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2016. He is currently pursuing the joint Doctoral degree with the Royal Military Academy and the Department of Electrical Engineering, KU Leuven. He joined the Royal Military Academy, Belgium, as a Researcher in 2016. His research interests include deep learning algorithms for wireless signal detection and classification.

Sreeraj Rajendran received the master's degree in communication and signal processing from the Indian Institute of Technology Bombay, India, in 2013, and the Ph.D. degree from KU Leuven, Belgium, in 2019, where he is a Postdoctoral Researcher with the Networked Systems Group. In 2013, he was a Senior Design Engineer with the Baseband Team, Cadence (Tensilica). He was also an ASIC Verification Engineer with Wipro Technologies from 2007 to 2010. His main research interests include machine learning algorithms for wireless spectrum awareness and low power wireless sensor networks.

Sofie Pollin (Senior Member, IEEE) received the Ph.D. degree (Hons.) from KU Leuven, Leuven, Belgium, in 2006. From 2006 to 2008, she continued her research on wireless communication, energy-efficient networks, cross-layer design, coexistence, and cognitive radio with UC Berkeley. In November 2008, she returned to IMEC, Leuven, where she become a Principal Scientist with the Green Radio Team. Since 2012, she has been a tenure track Assistant Professor with the Department of Electrical Engineering, KU Leuven. Her research centers around networked systems that require networks that are ever more dense, heterogeneous, battery-powered, and spectrum constrained. She is a fellow of the BAEF and Marie Curie.

Bart Scheers was born in Rumst, Belgium, in November 1966. He received the M.S. degree in engineering, with a specialization in communication from the Royal Military Academy, Brussels, Belgium, in 1991, and the joint Ph.D. degree from the Université Catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium, and Royal Military Academy, in 2001, where he presented his Ph.D. dissertation on the use of ground penetrating radars in the field of humanitarian demining. He was an Officer with the Territorial Signal Unit, Belgian Army. In 1994, he was an Assistant of Signal Processing with Royal Military Academy, where he has been a Military Professor with the Communications Information Systems and Sensors Department, since 2003, and is also the Director of the Research Unit on Radio Networks. His current domains of interest are mobile ad hoc networks (layers two and three), cognitive radio, and Internet of Things.