APPLIED RESEARCH

RF Drone Detection System Based on a Distributed Sensor Grid With Remote Hardware-Accelerated Signal Processing

PRZEMYSŁAW FLAK[ID]1 **AND ROMAN CZYBA**[ID]2
1Department of Automatic Control and Robotics, Faculty of Automatic Control, Electronics and Computer Science, Ph.D. School, Silesian University of Technology, 44-100 Gliwice, Poland
2Department of Automatic Control and Robotics, Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 44-100 Gliwice, Poland

Corresponding author: Przemysław Flak (przemyslaw.flak@polsl.pl)

ABSTRACT Unmanned Aerial Vehicles (UAVs), sometimes known as drones, evolved from military to civilian applications, opening up novel perspectives in a variety of everyday services. The rapidly growing consumer interest in amateur drones equipped with high-end cameras compromises the everyday safety and privacy of people. In the literature, a variety of sensing techniques based on different physical phenomena have been proposed for drone detection. Among acoustic, optical, or radar detection systems, passive radiofrequency sensing is the only one that can identify a drone even before it takes off and additionally indicate the operator's location. A spectrogram-based method is developed and optimised in terms of computing location, resulting in the possibility of sensor grid deployment over a standard Ethernet network. The detection phase involves hardware-accelerated energy sensing to extract the data frames from the background noise. Drone presence is then identified using machine learning based solely on preamble pattern recognition, which reduces the computational effort. The presented procedure is evaluated in an isolated setting employing an open-source dataset and tuned across multiple neural network architectures. Next, the complete sensor processing chain is examined in a real-life scenario. The analytical energy detector stage reaches a margin of roughly -8.7 dB in the signal-to-noise (SNR) ratio. With 1.1 M parameters, the proposed neural network achieves 99.93% simulation accuracy in up to -9.5 dB SNR range. Even after quantization for embedded platform implementation, the device can be used as a stand-alone early intrusion detector or as part of a distributed sensor grid.

INDEX TERMS Convolutional neural network, drones, field programmable gate array, software defined radio, spectrogram, surveillance, unmanned aerial vehicles.

I. INTRODUCTION

The history of Unmanned Aerial Vehicles (UAVs), commonly referred to as drones, reveals that while the initial interest in the technology was purely military, a variety of civilian applications have emerged. Since the Federal Aviation Administration (FAA) of the United States granted the first commercial licence in 2006, global public awareness of drone technology has progressively expanded. Despite

The associate editor coordinating the review of this manuscript and approving it for publication was Chengpeng Hao[ID].

the significant economic implications of COVID-19, the commercial drone sector has grown steadily over the past two years and is expected to reach USD 54.81 billion by 2030 [1]. Furthermore, according to the most recent FAA report [2], the number of recreational drones in the United States is predicted to peak at roughly 1.82 million units by 2027, achieving a total annual growth rate of 1.6%.

The unique attributes of UAV operations provide new potential in various domains, including agriculture, inspection, package delivery, and mining. However, the increasing demand for aerial services and the growing number of

amateur drones in the airspace directly endanger people's safety and privacy. Because the majority of common consumer models include high-quality cameras with real-time video transmission capabilities, they may be used for harassment or stalking. In addition to everyday minor incidents, extremely hazardous events like drone intrusions into nuclear power plants, airport operations holdups, drug trafficking, and weapon smuggling are becoming increasingly frequent. Over the past two years, reports of UAV sightings from pilots, residents, and law enforcement have expanded rapidly, and the FAA currently receives more than 100 notifications each month [3]. The preceding examples highlight the significance of developing drone surveillance systems to operate in both highly specialised and widely applicable general scenarios.

Numerous sensing strategies based on various physical phenomena have been presented in the literature for drone detection. The researchers utilise electromagnetic [4], acoustic [5], video [6], and radar [7] data for algorithm development. Each of these techniques has drawbacks, and as a result, several systems with heterogeneous sensors and data fusion methods are demonstrated to satisfy practical requirements [8].

Based on the drone propeller sound distinction, the acoustic method is affected by noisy environments in metropolitan areas, which significantly reduces the sensing range. Furthermore, wind direction and temperature might affect the sensor's useful area even further. The acoustic approach is purely passive, thus it does not require any additional licences and basic detection tasks can be performed with easily accessible low-cost equipment even at embedded devices [9]. While sensor range can be extended with sophisticated microphone arrays [10], this method provides relatively limited coverage in comparison to other techniques and is ineffective when it comes to gliding drones. Also, acoustic sensors can recognise a swarm, but they could have trouble identifying specific drones inside it.

To accurately identify drones using video in low-light conditions, high-resolution cameras and high-quality optics are needed, which have a significant financial impact. To achieve a detection range of around one kilometre with a high optical zoom, it is necessary to move the camera continuously to monitor the surroundings [11]. However, when an object is spotted, a video feed enables accurate tracking and location of the intrusion. Advanced Deep Learning (DL) techniques have demonstrated strong performance in object detection lately, however, the majority of current approaches struggle to achieve a good balance between detection performance and model size. Despite a lot of effort being made in the area of optimisation in [12], the network's inference time still falls short of the standards for real-time performance. In summary, this technique is also passive, sensitive to weather conditions such as fog or rain, and faces the standard challenge of occlusion handling in object detection. Moreover, one camera can only follow one drone at a time, making swarm incursions hard to manage.

The radar technique, based on the detection of electromagnetic backscattering, is the most sophisticated and expensive option, yet to cover great distances, an active beam is needed. It requires additional frequency licences, and system deployment near an airport may be constrained. Apart from that, radar can detect and track multiple objects in the protected area, regardless of weather conditions. It is also possible to use the micro-Doppler shifts in radar echoes generated by rotating blades to distinguish between the drone object and the bird [13]. Although the passive radar concept is feasible [14], ever smaller drone sizes, low altitude operation, low flying speed, and the urban environment highly reduce this detector's accuracy.

Passive Radio Frequency (RF) sensing, based on the assumption that the drone is communicating wirelessly with the controller, is, in the authors' opinion, one of the most effective approaches to civilian surveillance. This technique requires no licences and can detect multiple objects regardless of weather conditions and daytime. With a simple antenna, the omnidirectional detection range could reach over one kilometre, while with a more sophisticated antenna set, it can be extended even over ten kilometres. In comparison to acoustic and visual, the equipment required to accomplish real-time wideband frequency monitoring is more complicated and requires more computational power. The unique feature of the RF approach is the possibility of identifying a drone startup procedure, even before it takes off and indicating the operator's location. This technique is only ineffective against autonomous drones that do not transmit telemetry or video data and operate on a pre-programmed route. This scenario is more common in military operations than in regular life.

The current article introduces a novel three-stage approach to passive RF drone detection based on a distributed sensor grid concept. The first stage refers to signal acquisition and hardware-accelerated time-frequency-domain transform calculation in a Software Defined Radio (SDR) device. Following that, we propose drone presence identification onboard an embedded computer as part of a remote sensor. Finally, drone model recognition can be performed in the data fusion centre. The first stage is characterised by the author's previous works that deal with expanding the real-time bandwidth of a low-cost SDR device [15] and increasing the broadband scan speed [16]. The current study focuses on the implementation of the drone presence identification method and serves as an opportunity to present the whole system idea. The algorithm is tuned in a laboratory environment with an augmented open-source dataset across multiple machine learning models that are evaluated in terms of performance and computational complexity. Additionally, a real-world dataset that the authors explicitly prepared for this study is used to assess the effectiveness of the entire remote processing chain. The provided system architecture allows for the development of a scalable distributed detection network that can be customised to various scenarios, which, together

with fusion centre drone model recognition, will be the subject of future study.

The structure of this paper is organised as follows: the system model, including the problem statement, is introduced in Section II, and the related work is discussed in Section III. The detection algorithm and implementation details are presented in Section IV. Finally, an experimental setup with test results is provided in Section V, whereas conclusions and ideas for future development are outlined in Section VI.

II. SYSTEM MODEL

The majority of commercially available drones transmit RF signals using unlicensed Industrial, Scientific, and Medical (ISM) frequency ranges. The 2.4 GHz and 5.8 GHz bands are used for flight control communication with the Ground Control Station (GCS), as well as to broadcast live video and transfer telemetry data. Even autonomously operating models require an active radio connection to send telemetry and real-time video streams. Monitoring over-the-air transmissions, in conjunction with advanced processing, can be applied to develop a drone surveillance system. Although the effectiveness of this technique can degrade due to signal coexistence with WiFi, Bluetooth, or ZigBee broadcasts, it has the indisputable benefit of being licence-free and independent of ambient environmental conditions.

Certain simple drone variants rely on WiFi control channels and may be operated entirely with a smartphone application. We typically deal with smaller objects, whose operational distance is limited by the power of a mobile phone radio module in this situation. Using more advanced drone models with larger payload capacities may result in more hazardous scenarios. Those high-end versions employ proprietary communication protocols, and a specialised external device is required for control in order to provide a greater effective flight range. Although the specifics of these transmissions are not part of any standard, heuristic observations suggest that Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS) and Orthogonal Frequency Division Multiplexing (OFDM) techniques are often applied. For spectrum sensing in situations where little or no prior information exists, a blind signal detection scheme has to be implemented. The high-frequency oscilloscope [17], spectrum analyzer [18], or SDR [19] are possible base devices for the RF capture instrument. Despite the fact that both the oscilloscope and the spectrum analyzer provide excellent precision, they are better suited for work in laboratory environments due to their high cost, size, and power consumption. Considering the aforementioned, the SDR platform is the foundation for the proposed system.

A fundamental challenge for the RF detection process is the large concentration of physical obstacles affecting radio propagation mechanisms and the high density of wireless operating devices in urban or industrial areas. Establishing electromagnetic situation awareness by employing a distributed spectrum monitoring system can offer greater area coverage than a single sensor unit in this case

FIGURE 1. Concept of a distributed detection system based on an example allocation of nine sensors for the protection of an industrial facility. The arrangement of the sensors can be selected so that the detection areas overlap. Intrusion detection allows coarse localisation to take place immediately.

[20]. A powerful data fusion centre that can analyse both ambiguous and fragmented input data might be part of the infrastructure, in addition to a potentially unlimited number of simpler sensing elements [21]. Furthermore, the possibility of a more flexible zone coverage configuration will be advantageous if external high-power transmitters affecting the sensor's dynamic range are present within the detection range. The cost of each individual node must be kept as low as possible for such a notion to be practical, hence, the purpose of this research is to develop a method based on low-cost SDR and commercially available components. This concept is illustrated in Fig. 1.

III. RELATED WORK

This section is focused on a brief overview of RF-based drone detection approaches, whereas a more detailed survey of all existing techniques can be found in [22]. The methods of interest may be divided into two categories, depending on how the information is acquired. Firstly, WiFi signal interception is driven by Media Access Control (MAC) layer statistical metrics analysis. When a manufacturer's MAC address range is available, packet sniffing can be used to collect more precise data for further drone type classification [23]. Received signal strength (RSS) statistics from a commercially available WiFi device can also be used to detect intrusions in domestic scenarios [24]. Despite the reliability and variety of commercial receivers, the strategy is still beyond the scope of this study due to the modest market share of drones that feature WiFi control and the restricted operational range compared to proprietary protocol-driven variants. Therefore, this work is focused on the principle of spectrum sensing, which relies on raw radio signal analysis and the development of sophisticated detection algorithms.

The physical-layer traffic analysis, focused directly on corresponding In-Phase and Quadrature components (IQ) signal samples, is presented in [25]. To extract radio packets, Power Spectral Entropy (PSE) and a hysteresis thresholding approach are combined in the paper. Following that, a Physical-Layer Protocol Statistical Fingerprint (PLPSF)

is calculated for classification based on temporal features such as packet length and inter-packet duration. Although the method provides scale permanence, frequency invariance, and a high detection rate even at a low signal-to-noise ratio (SNR), the 100 ms signal segments have to be captured for analysis. The technique is highly effective at distinguishing between drone models from various vendors, though some misclassifications occur because of WiFi traffic's significant unpredictability and user-dependent nature.

As features for classification, [17] employs a variety of Machine Learning (ML) methods in conjunction with time-domain statistical characteristics such as slope, kurtosis, skewness, shape factor, and variance. However, due to the heavy exploitation of unlicensed spectrum, time-domain approaches have limited performance. Signal transformation into the frequency-domain reduces out-of-band interference and can increase classification accuracy. frequency-domain features are extracted and used to assess six different ML techniques in [26]. Both domain strategies are analysed in [27]. Moreover, a procedure that employs a merged training set of data at various noise levels has been proven to be successful for drone classification even at SNRs not explicitly included in the model training phase. Despite the fact that only a 250 μs time segment is required for processing, the proposed algorithm's multistage architecture makes it challenging to characterise in terms of computational cost and final detection latency.

To enhance the detection and classification performance of drone monitoring algorithms, multiple DL models have recently been studied and compared. The predefined networks analysed in [28] reveal that MobileNet V2 is able to outperform SqueezeNet and ResNet in terms of accuracy at the expense of computational complexity. Three custom networks from [29], [30], and [31] are evaluated in [32] on the same dataset against a novel model that is presented and extensively optimised. Although the proposed approach is based on time-domain 250 μs signal segments with minor input signal processing, it has the potential to achieve high accuracy and be used in portable real-time systems while having a relatively low level of complexity. In contrast, to increase accuracy in low SNR conditions, special preprocessing techniques are applied to the time-frequency signal representation in [33] prior to utilising DL. The method's findings imply that distinguishing using the spectrogram is more accurate, however, the classification of multiple drone controller signals remains challenging.

A two-stage strategy for noise-resilient detection based on spectral entropy and Convolutional Neural Network (CNN) classification is presented in [34]. At first, the coarse Fast Fourier Transform (FFT) is performed to monitor the radio spectrum. When the possible UAV presence is detected by entropy drop, a higher-resolution time-frequency transform is recalculated on the same data segment and passed to classification. A verification made on a publicly available dataset demonstrates improved classification accuracy at an SNR of -10 dB when time is prioritised over frequency

resolution during the execution of a transform. Furthermore, in [19], a two-stage technique with a combined detection and classification framework is compared in a variety of scenarios. The two-stage approach employs Goodness of Fit (GoF) sensing for detection and the Deep Residual Neural Network (DRNN) for classification. Despite the fact that the single-step strategy based on the presented You Only Look Once (YOLO-lite) framework is faster, both strategies yield satisfactory detection and classification results. According to the authors, it is difficult to extract signal features essential for further jamming implementation when applying the dual procedure. Understanding the relevance of this issue, we shall also investigate this conclusion further in this study.

Implementing an early warning system for UAV intrusion based on commercially available components is an interesting and practical solution presented in [35]. The results show that a classifier could be effectively implemented on the Raspberry Pi and BladeRF-SDR, hardware that does not require extensive financial outlays. The processing times of the captured 80 ms data segment range from 15 to 28 seconds. While the overall system accuracy is satisfactory, the authors concluded that the low-cost computer would be more suitable to operate as a repeater of the raw SDR data for solutions requiring a fast response. The concept of a central control unit with enhanced processing capabilities is also discussed in this study, and the challenge of transferring unprocessed SDR data across a Local Area Network (LAN) is described in greater detail in the following section.

IV. PROPOSED SOLUTION

This section describes the architecture of the proposed two-step procedure for a drone RF monitoring solution in terms of where the computations are physically performed. Aside from the algorithm itself, one of the key concepts of this research is to preprocess data locally inside remote sensors to reduce the amount of information transmitted to the central classification unit. The first phase of the proposed framework, which is based on frequency-domain operations, is to capture the signal and create the real-time spectrogram inside the SDR platform with hardware acceleration. The USB input is then analysed by a portable computer, which performs signal detection to distinguish the relevant data from the noise. The transfer of the truncated spectrum fragment to the central unit is initiated only after a successful detection in the previous phase, thus reducing the LAN traffic. Fig. 2 reveals the spectrogram of drone control and video signals captured in an isolated laboratory environment where no other transmission occurs. In this example, it can be observed that the signal of interest occupies under 8% of the analysed spectrum. Assuming 24-bit IQ standard SDR precision and 83 MHz sampling mandatory for ISM 2.4 GHz imaging, raw time-domain data transmission requires around a 1992 Mbps transfer rate. Since the majority of embedded computers feature 1000 Mbps LAN ports, transfer minimisation is a key challenge to the deployment of a distributed sensor grid. In addition, while the presence of a drone in the protected
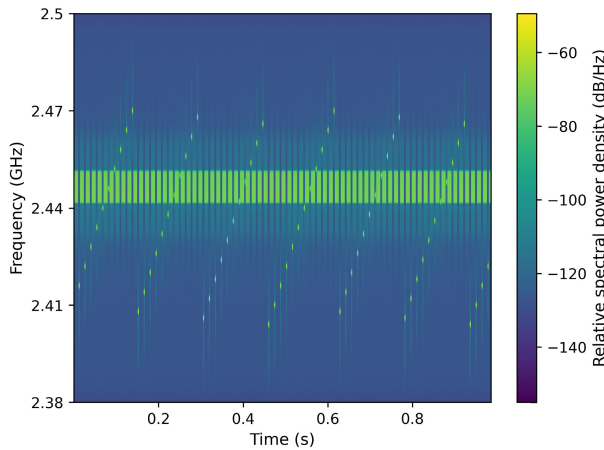
FIGURE 2. Spectrogram of drone signals captured in an isolated laboratory environment in 2.4 GHz ISM band. The fixed-frequency wideband signal belongs to the video transmission and the variable-frequency narrowband one is the control channel. The relevant information is contained in 8% of the analysed spectrum.
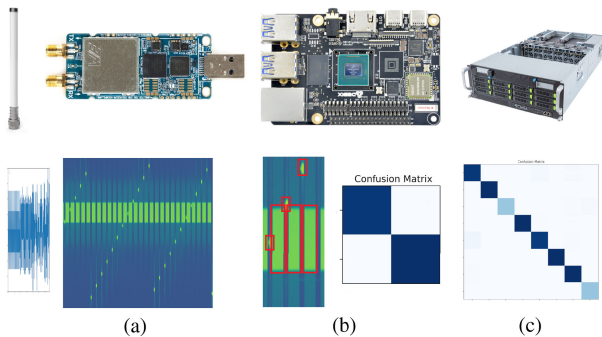

FIGURE 3. Elements of the detection and identification chain. The main components, such as the antenna, the SDR, the embedded computer, and the fusion centre, are shown with data the product of each section's processing: (a) spectrogram, (b) signal detection and drone presence identification, (c) further drone model classification. While the drone is not observed the LAN transfer rate from the multisensor grid is negligible.

zone is an abnormal situation, maintaining a constant network load at a high level is unacceptable. Fig. 3 shows a flowchart of the process in more detail.

A. SDR SPECTROGRAM

The standard off-the-shelf SDR delivers the raw time-domain IQ [36] samples without performing any sophisticated signal processing. While most high-end SDRs use 10 Gigabit Ethernet [37], mid-range ones often employ direct USB 3 connection [38]. As mentioned before, capturing a continuous frequency range from 2.400 GHz to 2.483 GHz in the form of dual 12-bit digital samples generates around a 2 Gbps stream. This is a substantial amount of data for the embedded computer to analyse in real-time. Furthermore, executing the frequency-domain transform is merely the initial step of a more comprehensive signal processing procedure. Thus, we first implemented a hardware spectrogram accelerator in [15] as a firmware enhancement for a sensor with full 83 MHz bandwidth coverage based on a low-cost SDR. The spectral estimator is obtained by separating a constant input data stream into segments with 50% overlap and computing the Fourier Transform on every segment.

To provide a theoretical basis, suppose the received signal is represented in discrete-time notation as:

$$x[n] = s[n] + z[n], \tag{1}$$

where $s[n]$ is signal of interest and $z[n]$ stands for noise. The discrete-time form of the short-time Fourier Transform (STFT) can be expressed as:

$$\text{STFT} \equiv X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n - m]e^{-j\omega n}, \tag{2}$$

where $x[n]$ represents the sampled received signal, $w[n]$ states for the window function, n indicates the sample number, and m is the position of the analysis window. The spectrogram representation of the function's power spectral density (PSD) is given by the squared magnitude of the STFT as:

$$\text{spectrogram} \equiv |X(m, \omega)|^2. \tag{3}$$

The characteristic of the window shape is selected to precisely measure the processed signal's duration in order to identify transients such as the start or end of a frame. According to [39], many window functions attenuate valuable information around the sample frame's boundaries. When applying a window type other than a uniform one, the overall signal power is reduced as a result of the time-domain tapering. The magnitudes in the frequency-domain after using the Fourier transform are no longer equal to their true values. This is explained by Parseval's Theorem [40] about energy preservation across time and frequency domains:

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2, \tag{4}$$

where $x[n]$ is signal and $X[k]$ represents its Discrete Fourier Transform (DFT). The reduction in the magnitude of the spectral components is defined as Coherent Power Gain (CPG) [41] and for a window $w[n]$ of length L is given by:

$$\text{CPG (dB)} = 20 \log_{10} \frac{1}{L} \sum_{n=0}^{L-1} w[n]. \tag{5}$$

For a uniform window expressed as:

$$w[n] = \begin{cases} 1 & \text{for } 0 \leq n \leq L - 1, \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

the CPG parameter is equal to unity or 0 dB, while for other windows is actually a loss.

Apart from its low complexity, ideal for hardware acceleration, a uniform window is applied in SDR firmware to enhance the accuracy of time event detection and energy conservation in exchange for some spectrogram distortions. As a result of the pipeline processing within the Field Programmable Gate Array (FPGA) onboard SDR, the spectral estimator is delivered for further analysis in real-time with a latency equal to twice the window length.
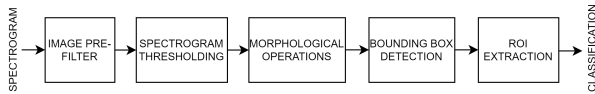
FIGURE 4. Components of the proposed signal processing chain for wideband spectrum sensing and ROI extraction.

B. PORTABLE COMPUTER SPECTRUM SENSING

In this study, spectrum sensing is directly related to differentiating the signal of interest from noise. The signal processing chain implemented to extract the relevant data is presented in Fig. 4. A number of approaches for completing this task have been proposed in the literature. The fundamental benefit of the Energy Detection (ED) approach is that, unlike Entropy Detection, Goodness of Fit Test, or Eigenvalue-Based Detection, it does not require complex signal processing. Another notable advantage is that, unlike Cyclostationary Feature Detection, Waveform-based Detection, and Matched Filter Detection, it does not require prior knowledge of the characteristics of the signal under analysis [42]. There are also several frameworks that explicitly use machine learning for signal detection [43]. These methodologies are unfeasible for the sensing component proposed in the current study due to the restricted computational power of portable devices and the emphasis on real-time processing. In addition, to improve response time in the future, the detection algorithms will be transferred to an FPGA chip, onboard a selected SDR, where there is no room for neural operations.

The ED technique simply entails measuring the received signal energy at the receiver side and comparing it to a threshold. Following spectrogram computation, this process is executed in the frequency-domain in this study. Assuming that H_0 relates to the energy of the noise-only region and H_1 is the signal plus noise region of the spectrum, the sensing decision θ is given by:

$$\theta = \begin{cases} H_1 & \text{for } E \geq \lambda, \\ H_0 & \text{for } E < \lambda, \end{cases} \tag{7}$$

where, if the energy of the frequency bin E is higher than the threshold λ, the measured sample is accepted as signal and noise, otherwise classified as noise only.

To address the issue of noise reduction before the thresholding process and minimise the false alarm detection rate, some smoothing filter architectures were instigated in the author's previous study [15]. The best solution for suppressing high frequencies while simultaneously minimising spatial spread was a Gaussian filter, which preserved the signal edge without compromising crucial details like the occupied bandwidth and the precise length of the signal. The 2D Gaussian convolution can be written as:

$$G(x, y) = S(x, y) \otimes k(x, y), \tag{8}$$

where $\otimes$ denotes convolution, $k(x, y)$ is a Gaussian kernel, $S(x, y)$ is the input signal and $G(x, y)$ is the convolved output. In the digital domain, the discrete kernel forms a matrix of size $n \times n$, with radially symmetric coefficients. Different
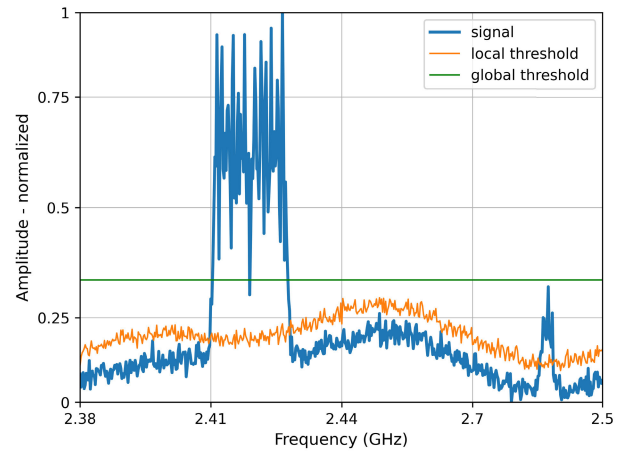

FIGURE 5. Comparison of the effectiveness of global and local thresholding in the case of a non-flat wideband noise floor. Some signals may not be detected with a global threshold applied.

kernel sizes and their effects on the signal detection process are further examined in the experimental part. Exemplary kernel $k(x, y)$ can be expressed as:

$$k(x, y) = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{9}$$

The common signal thresholding variations examined in the literature can be considered fixed or adaptive. Besides being straightforward, the fixed technique, in which the threshold is not dynamically recalculated, suffers from fluctuating background noise, especially in the crowded ISM band. As the proposed sensor should provide signal detection in various environments, the adaptive technique is chosen for implementation. Moreover, when dealing with non-flat responses caused by impedance mismatches in wideband measurements, filter roll-off, or group delay of analogue elements, local thresholding for every single frequency bin is preferred. Fig. 5 shows an example of a received signal spectrum where the Noise Floor (NF) is not flat. In reality, global threshold detection is insufficient over a wide bandwidth due to those fluctuations, especially when some signals have low SNR. Multiple approaches for local thresholding are examined using real-life drone recordings since the subject of ED itself has been thoroughly explored in numerous studies. Based on the fixed nature of sensor usage and the accidental style of drone incursions, we can optimise the threshold calculation with regard to long-term analysis.

The approach proposed in [44] is capable of handling the slow time-varying NF characteristic, frequency dependency, presence of a signal, and computing cost minimisation all at once. In comparison to the current state-of-the-art two-dimensional Forward Consecutive Mean Excision (FCME) [45] algorithm, the proposed method delivers comparable NF estimation performance at a reduced runtime, according to the authors. The primary concept behind the work shown above is to propose a metric that tracks NF changes and is easier to compute than generating a new threshold for every

subsequent signal frame. Numerical experiments indicate that the suggested strategy identifies signals properly while taking into account the low fluctuation dynamics of background noise and has a 10 times faster runtime. This feature is especially important for embedded device implementation because, considering at least 83 MHz sampling by SDR and 1024 FFT length with 50% overlap, the threshold recalculation has to be performed in around 6.2 μs.

In this paper, we use the Non-Parametric Amplitude Quantization Method (NPAQM) [46] to distinguish between H_0 and H_1 states in power spectrum data. We utilised this technique in our prior research [15] because, unlike FCME, it is not iterative. To provide the starting point for NPAQM threshold selection, the sample set obtained by the FFT frame of length L is sorted in ascending order as $y_{(1)} < y_{(2)} < \ldots < y_{(L)}$ and expressed as:

$$Y_l = \{y_1, y_2, \ldots, y_L\}. \qquad (10)$$

Next, the quantization of the amplitude range is suggested to choose the most suitable threshold candidate while reducing the computation work. With ceiling function $\lceil \cdot \rceil$, the quantization level M, is determined as:

$$M = \lceil 1 + \log_2 L + \log_2 (1 + \frac{|g|}{\sigma_g}) \rceil, \qquad (11)$$

where the modulus function $|\cdot|$ of distribution skewness g is obtained with:

$$g = \frac{\frac{1}{L} \sum_{i=1}^{L} (Y_i - \overline{Y_l})^3}{(\frac{1}{L} \sum_{i=1}^{L} (Y_i - \overline{Y_l})^2)^{\frac{3}{2}}}, \qquad (12)$$

and $\overline{Y_l}$ is the mean of the input sample set Y_l, where:

$$\sigma_g = \sqrt{\frac{6(L-2)}{(L+1)(L+3)}}. \qquad (13)$$

The calculation step q, is then written as:

$$q = \lceil \frac{y_{(L)} - y_{(1)}}{M} \rceil, \qquad (14)$$

where $y_{(L)}$ and $y_{(1)}$ relates to $\max(Y_l)$ and $\min(Y_l)$ respectively, since the incoming data was already sorted. The list of potential threshold candidates can now be prepared as follows:

$$Y(m) = \{y_{(1)}, y_{(1)} + q, \ldots, y_{(1)} + (M-1)q\}, \qquad (15)$$

and simplified to:

$$\gamma_i = \{\gamma_1, \gamma_2, \ldots, \gamma_M\}. \qquad (16)$$

At this point, the data is divided into two subsets: noise subset $\omega(\gamma_i)$ and the signal plus noise subset $S(\gamma_i)$. It is presumed that samples with values below the threshold are noise-only, whereas samples with higher values are signal-plus-noise. The sets of $\omega(\gamma_i)$ and $S(\gamma_i)$ are calculated as follows:

$$\omega(\gamma_i) = \{Y_{(l)} < \gamma_i\} = \{y_{(1)}, y_{(2)}, \ldots, y_{(k)}\}, \qquad (17)$$

$$S(\gamma_i) = \{Y_{(l)} \geq \gamma_i\} = \{y_{(k+1)}, y_{(k+2)}, \ldots, y_{(L)}\}, \qquad (18)$$

for all elements in (16), where k is the number of elements smaller than γ_i. Calculation of the between-class variance of both subsets for each threshold element in (16) is determined as:

$$\sigma^2(\gamma_i) = P_s(\gamma_i) P_\omega(\gamma_i) [\overline{S}(\gamma_i) - \overline{\omega}(\gamma_i)]^2, \qquad (19)$$

where $P_s(\gamma_i)$ is the probability of signal elements and $\overline{S}(\gamma_i)$ is the mean of the subset (18), $P_\omega(\gamma_i)$ is the probability of noise elements and $\overline{\omega}(\gamma_i)$ is the mean of the subset (17). Now, the probabilities can be declared as:

$$P_s(\gamma_i) = \frac{k}{L}, \quad P_\omega(\gamma_i) = 1 - \frac{k}{L}, \qquad (20)$$

where k is the number of elements in each subset. Finally, the first-order difference function is carried out with (19) in range $m = 1, 2, \ldots, M-1$ written as:

$$\lambda_m(\gamma_i) = |\sigma_{m+1}^2(\gamma_i) - \sigma_m^2(\gamma_i)|. \qquad (21)$$

The optimal threshold is determined by the value that minimises (21), which can be presented as:

$$\lambda(\gamma^{eff}) = \min \lambda(\gamma_i), \text{ for } \gamma_i = \{\gamma_1, \gamma_2 \ldots, \gamma_M\}. \qquad (22)$$

The authors in the same article [46] also presented a novel heuristic technique as a supplement to the NPAQM, leveraging the degree of proximity between the calculated threshold $\lambda(\gamma^{eff})$ and the mean of the total sample set Y_l, to enhance performance in noise-only regimes. The heuristic implies that the sample collection contains components that are nothing but noise when:

$$\lceil \gamma^{eff} \rceil \leq \lceil \mu_Y \rceil, \qquad (23)$$

and:

$$\left| \mu_Y - \gamma^{eff} \right| \leq 10 \times \left[y_{(F)} - y_{(1)} \right], \qquad (24)$$

where mean μ_Y is defined as:

$$\mu_Y = \frac{1}{L} \sum_{l=1}^{L} Y_l \qquad (25)$$

Although determining an appropriate threshold is a crucial step in the detection chain described here, additional morphological image processing is also carried out to further facilitate the entire procedure [47]. Following the basic nonlinear convolutions of morphological signal processing [48], erosion can eliminate noise, while dilatation can then restore any incorrectly eliminated signal areas. Considering spectrogram as A and B as structuring element (kernel) in Z^2, the dilatation of A by B can be expressed as:

$$A \oplus B = \left\{ z | \left[(\hat{B})_z \cap A \right] \subseteq A \right\}, \qquad (26)$$

while erosion is defined as:

$$A \ominus B = \{z | (B)_z \subseteq A\}. \qquad (27)$$

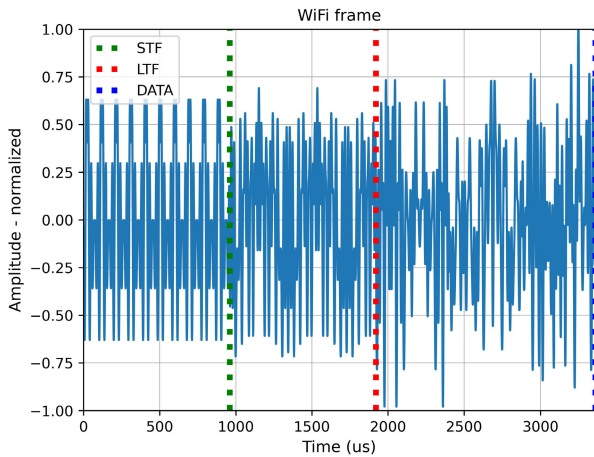A non-linear time-frequency filtering occurs when the two methods are combined. Both erosion and dilation are

FIGURE 6. Time-domain representation of a wireless signal in the 802.11n/ac protocol with region boundaries. STF is Short Training Field, LTF is Long Training Field, and DATA is the actual signal. STF and LTF together form a repeatable preamble.

processes that may be used repeatedly, but neither procedure is often reversible. The purpose of the morphological opening is to dilate the eroded image in an effort to restore as much of the original image as possible:

$$A \circ B = (A \ominus B) \oplus B \qquad (28)$$

The closing is the dual operator of the morphological opening:

$$A \bullet B = (A \oplus B) \ominus B \qquad (29)$$

It also fuses tiny cracks and long, thin gulfs, fills in gaps, and eliminates small holes in the spectrum image. The selection of the order of these operations is investigated in the experimental section.

C. SIGNAL CLASSIFICATION

Consider the basic wireless digital communication system, where the transmitter is responsible for signal generation while the receiver has the more difficult function of determining what message is being sent. Estimating and correcting for frequency or phase offset between local and remote clock generators in this configuration requires specialised procedures. Most widely known wireless systems operating in the ISM band, such as Bluetooth and WiFi, use algorithms based on the premise that some portion of the data is known. When a sequence of these distinctive symbols, generally with high correlation properties, is placed before the message itself, it is called the preamble. Fig. 6 illustrates an example of a WiFi frame in the time-domain.

While the preamble sequence is consistent and data-independent, it may be used to identify protocols in the ISM band as well as distinguish drone signals. The primary challenge in this strategy is to give a comprehensive picture of the spectrum while using a fast sample rate and optimum time-frequency transform length to achieve the necessary resolution. To avoid blurring the spectrum representation,
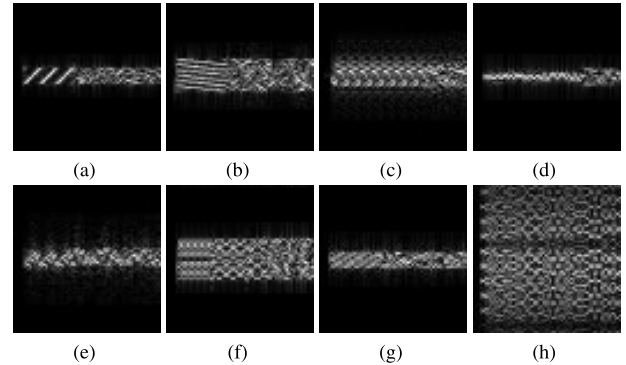

FIGURE 7. Various preambles displayed as ROI spectrograms cutout with 64 × 64 pixel areas: (a) Dji Mavic 2, (b) Dji Matrice 300, (c) Yuneec Typhoon, (d) Bluetooth, (e) Dji Phantom 4, (f) Dji Mavic Mini, (g) Dji Mavic Pro, (g) Wifi.

no temporal averaging should be applied to reveal an individual frame pattern. Additionally, the USB data stream from the SDR sensor cannot be interrupted due to the possibility of transient miss detection and preamble omission. In previous studies, a portion of the spectrum image was typically scaled down on the input layer of a neural network. In contrast, the proposed approach relies on Region-of-Interest (ROI) cropping. Due to the significantly smaller size of the input data, this presents an opportunity to optimise the machine learning model for embedded computing. Fig. 7 provides a comparison of various preamble patterns for drone control signals. Following successful drone signal categorisation, the entire radio frame is packaged and transferred to the central computer via the LAN interface. To ensure that no vital information about the hopping pattern is lost, additional detail about the centre frequency is appended to the ROI metadata.

V. EXPERIMENTAL RESULTS AND DISCUSSION

We executed a number of experiments and tune procedures to demonstrate the potential of our two-step approach for distributed drone intrusion identification. Initially, an open-source dataset [49] was used to alter the detection performance and preamble extraction mechanism. It provides high-quality recordings of ten drone signals, both video and telemetry, acquired in an isolated laboratory environment. The constant background noise and high SNR of the attached data enabled straightforward ground-truth annotation of the signal frames using global thresholding and bounding box recognition from the OpenCV library [50]. To conduct experiments with the classification framework, the dataset was augmented with WiFi and Bluetooth signals generated in a Matlab environment. Finally, the proposed approach was validated using real-world signals captured with a cost-effective enhanced SDR sensor.

A. DETECTOR PERFORMANCE

To evaluate the detector's performance, various levels of Additive White Gaussian Noise (AWGN) were injected into the raw signal to model the drone and the distant position of the receiving antenna. In most cases, signal and noise

power from the time-domain are used to calculate SNR. Alternatively, the power of any type of noise or wideband signal can be evaluated in the frequency-domain by summing the FFT bins on a linear scale, according to [51]. The advantage of this approach is that the selection of the bins that contribute to the signal, and hence the spectrum components, is flexible and can be used for a very wide range of measurements. The signal and noise power can be estimated independently as follows:

$$P_{singal}(dB) = 10 \log_{10}(\sum_{s=1}^{S} |X(s)|^2), \qquad (30)$$

$$P_{noise}(dB) = 10 \log_{10}(\sum_{n=1}^{N} |X(n)|^2), \qquad (31)$$

where $X(s)$ and $X(n)$ are the frequency bins for signal and noise, respectively. The final SNR may be determined as:

$$\text{SNR (dB)} = P_{signal}(dB) - P_{noise}(dB). \qquad (32)$$

To assess the quality of each signal detection result, the Intersection-over-Union (IoU) was utilised. The metric is defined as the proportion of overlap between a frame's true location and a predicted position in the spectrogram. The computation of the IoU ratio can therefore be determined as follows:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}. \qquad (33)$$

Various IoU levels for drone control signals are compared in Fig. 8. It can be observed that with SNR degradation, the predicted area is increasing. The signal can be considered detected if the IoU between the exact and estimated spectrum positions is greater than a threshold. Dual threshold values are provided since future classification is expected to function at 64×64 and 32×32 pixel ROIs. A smaller analysis region obtained with better detection performance will aid in further neural model size reduction.

In statistics, accuracy refers to the extent to which a binary classification test precisely detects or eliminates a condition. In other words, accuracy is the ratio of valid predictions, both true positives and true negatives, to the total number of instances studied. Binary accuracy and other metrics for performance evaluation may be expressed using the following formulas:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \qquad (34)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \qquad (35)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \qquad (36)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \qquad (37)$$

where TP are the true positives, TN are the true negatives, FP are the false positives, and FN are the false negatives. Finally, the detection accuracy in terms of the Gaussian
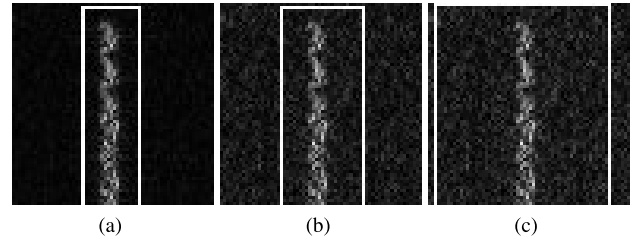

FIGURE 8. Various IoU values and bounding box detection displayed on ROI spectrograms cutout with 64 × 64 pixel areas: (a) IoU 1.0, (b) IoU 0.8, (c) IoU 0.6.

filter kernel is presented for 0.8 and 0.6 IoU thresholds in Fig. 9. It is shown that even at positive SNR, accuracy is not perfect in the absence of any smoothing filter, and the detection performance improves in negative SNR as the kernel size grows. Additionally, as the detection curve gradually saturates, increasing the mask size above 7×7 does not provide much benefit. The proper sequence of applying morphological operators led to further improvement, as seen in Fig. 10. With open and close functions following threshold application, a kernel of seven was finally selected to facilitate hardware implementation in future FPGA acceleration. The proposed classification approach, which is based on the spectrogram analysis of a single radio frame, offers more than 150 detection opportunities per second for a typical drone model. Thus, even if some frames remain undetected, the 20% accuracy attained at -8.2 dB SNR still gives a high potential for intrusion detection.

B. CLASSIFICATION SIMULATION

This section employs both laboratory and simulation techniques to assess the performance of the proposed binary classification model. Laboratory recordings of ten drone models, the same as in the previous section, were supplemented with WiFi and Bluetooth signals created in the MATLAB environment. The same sample rate and initial SNR were maintained during this process. Additional waveforms contained 100 frames of each protocol: 802.11ax, 802.11n/ac, Bluetooth 2.0, and Bluetooth Low Energy, with 500k, 1M, and 2M data rates. Again, all the signals were subjected to the AWGN channel up to the detection limit of -9.5 dB SNR, with a 0.5 dB step. The entire dataset was split randomly into train, validation, and test sets in a 6:2:2 ratio. The procedure of network fitting was then performed ten times to evaluate the whole dataset, and the total performance was calculated by averaging the results of all trials. The network was evaluated on a platform using an AMD Ryzen 5 3.4 GHz six-core CPU without GPU support.

The major purpose of similar studies is to categorise a particular drone model, even though the presence of an intrusion is frequently the first step in a hierarchical framework [52]. The authors of [30] accomplish both tasks with the same network, although when dealing with binary-only problems, a more efficient architecture can be provided, like in [31]. The approaches for classifying drone signals
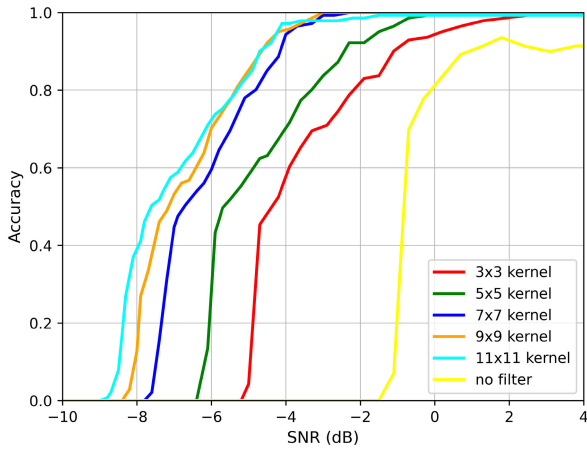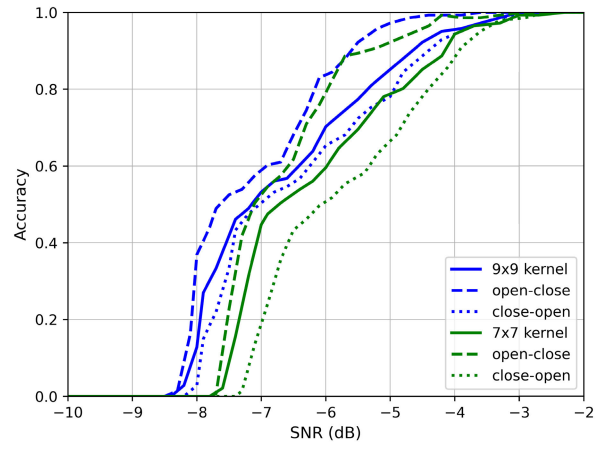
(a) Detection accuracy when IoU = 0.8
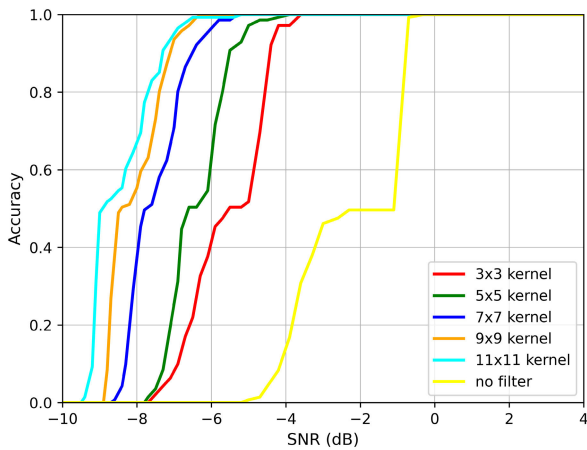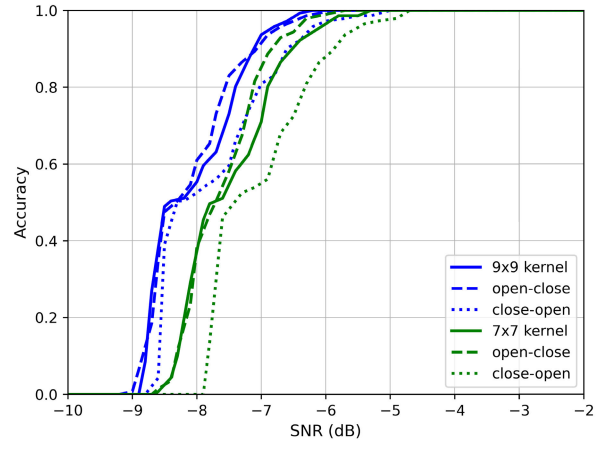

(b) Detection accuracy when IoU = 0.6

FIGURE 9. Detection accuracy in terms of SNR for various Gaussian filter kernel sizes and reference comparison without a filter.

(a) Detection accuracy when IoU = 0.8

(b) Detection accuracy when IoU = 0.6

FIGURE 10. Detection accuracy in terms of SNR showing the impact of additional morphological operations sequence.

TABLE 1. Binary classification results and model parameters for time-domain signals input or periodogram analysis. Speed measurements according to [32]. The results of the proposed method are included for reference.

Model	Accuracy	F1-score	Params	FLOPs	Speed
DNN [29]	99.71%	99.57%	5.10 M	5.2 M	1.24 ms
1D-CNN [31]	99.81%	99.61%	0.06 M	23.0 M	1.98 ms
MC-CNN [30]	99.95%	99.46%	0.05 M	47.0 M	2.64 ms
RF-UAVnet [32]	99.85%	99.75%	0.01 M	4.4 M	1.31 ms
Proposed DNN	99.93%	99.92%	1.11 M	2.2 M	0.31 ms

can be performed in the time-domain [29] or frequency-domain [33]. Although the proposed sensor design, where the spectrogram is produced by SDR hardware, does not suit the time-domain technique, certain accuracy and computational complexity metrics are included for comparison. The essential characteristics of binary classification models, explored in interesting research [32] that additionally demonstrate customised network evaluation, are summarised in Table 1. The attached findings of the time-domain and periodogram-based methods show high accuracy and fast processing times with 1D CNN layers, however, neither has been tested under various SNR conditions. Therefore, only general parameters of the proposed method are provided for comparison. The previous techniques use bandwidths that are significantly broader than the signal being identified, which can reduce performance in complex electromagnetic scenarios. On the other hand, spectrogram-based frameworks that handle negative SNR have more complicated two-dimensional input tensors and, consequently, 2D CNN layers.

In this study, Keras was used to train neural networks, with Tensorflow serving as the backend in the Python

environment. Spectrogram classification networks from [27] and [33] were revised to accommodate the reduced ROI pixel input size and retrained on the current dataset. Although the general structure of the original solutions remained unchanged, the activation function in the last layer was modified to be sigmoid. The aforementioned networks served as a reference in the subsequent evaluation, while the particular proposed design is shown in Fig. 11. The hyperparameters of the training process included batch size of 128, Adam optimiser with default learning rate, and binary cross-entropy loss function.

Table 2 illustrates the characteristics of various network models adapted to the 64 × 64 ROI classification technique.
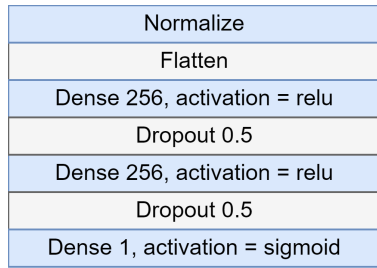
| Normalize |
| Flatten |
| Dense 256, activation = relu |
| Dropout 0.5 |
| Dense 256, activation = relu |
| Dropout 0.5 |
| Dense 1, activation = sigmoid |

FIGURE 11. Proposed neural network structure for ROI detection.

TABLE 2. Binary classification results and model parameters based on spectrogram analysis for reduced input tensor size to 64 × 64 area. Speed measurements according to own desktop evaluation platform.

Model	Accuracy	F1-score	Params	FLOPs	Speed
DNN [29]	99.92%	99.91%	0.55 M	1.2 M	0.87 ms
CNN [33]	99.55%	99.53%	0.39 M	213.4 M	2.28 ms
CNN [27]	99.60%	99.59%	1.27 M	56.8 M	0.68 ms
Proposed DNN	99.93%	99.92%	1.11 M	2.2 M	0.31 ms

The observed differences in trainable parameters and Million Floating Point Operations (MFLOPs) are essential for binary drone identification complexity analysis due to the small performance variations between the evaluated methods. For example, the original full spectrogram approach, based on CNN, offers 0.485 M trainable parameters with 507.9 MFLOPs while utilising a 356 × 452 input tensor size, according to [33]. The same network structure with a reduced ROI input of 64 × 64 allows for parameter reduction to 0.39 M and 213.4 MFLOPs. Using convolution filters in conjunction with pooling is a standard approach applied to image classification or object detection due to its inherent translation invariance property. Utilising a pre-processed spectrogram region does not benefit from CNN's ability to disregard positional changes or translations, so convolution layers were eliminated from the final solution.

The reference DNN [29], constructed upon three dense layers with 128 filters each, has fewer parameters and lower estimated MFLOPs than the proposed network. Nevertheless, the proposed, only two-layer layout with 256 filters, performs better in terms of speed without accuracy loss. A calculation period of 0.31 ms is particularly important because it is still less than the 0.625 ms minimum recorded Bluetooth transmission time slot.

To further tune the performance, the 32 × 32 ROI variation was also evaluated with the same dataset. Following this method, the parameter count was reduced to 0.32 M and the accuracy was dropped to 99.78% with a computation time of 0.14 ms. The reliability of the proposed network for both ROI versions in the entire detection range up to −9.5 dB is compared in Fig. 12. In reference to 64 × 64 ROI, the smaller variation offers computation speeds that are 0.17 ms faster with an overall accuracy drop of 0.15%. More detailed information regarding particular accuracy in terms of SNR is outlined in Fig. 13. It can be demonstrated that the greatest contribution to accuracy loss emerges in the SNR region of −3 dB to −9 dB in all cases.
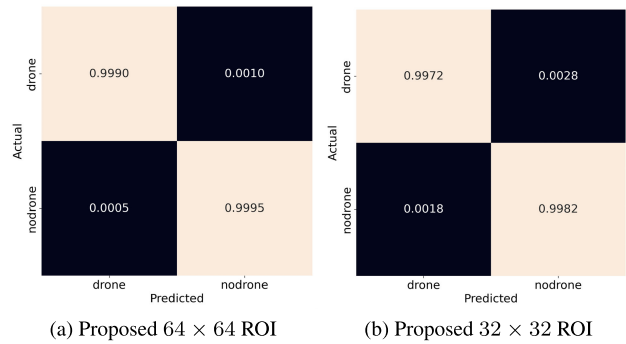

(a) Proposed 64 × 64 ROI   (b) Proposed 32 × 32 ROI

FIGURE 12. Confusion matrix for average drone detection accuracy in entire 20 dB to −9.5 dB SNR signal range obtained with the proposed model.

The proposed methodology compared to prior publications that also address the issue of determining the presence of a drone in the various SNR values, is shown in Fig. 14. The approach provided in [53] operates on the raw RF time-domain signals fed directly into the network model and does not require any feature extraction techniques. The model trained with the data from 0 dB to 30 dB SNR achieves an average accuracy of 97.53%. Another approach [54] uses the Wavelet Transform (WT) that utilises a variable window function, in contrast to the STFT which slides a fixed window through the signal. The authors employed the SqueezeNet machine learning model, which has a 98.9% accuracy at 10 dB SNR. A multistage detector is implemented in another system [55], where a set of 15 statistical factors is employed to characterise the drone's RF frame in both the time and frequency-domains. At 10 dB SNR, the detection accuracy is 99.8%, with a false alarm rate of 2.8%. In comparison, the proposed framework preserves over 99% drone presence identification accuracy in the entire presented SNR range.

C. REAL-LIFE RESULTS

In order to prepare the provided technique for real-life application, an additional process had to be conducted. The neural network deployment on the embedded computer was possible after the transformation to the TensorFlow Lite model. The first step of this procedure was the post-training model conversion to lite format, which initially kept float precision unchanged. This operation decreased the 64 × 64 stored model size from 12.9 MB to 9 MB without compromising accuracy.

Quantisation is the process of transforming a neural network operating on 32-bit floats to a smaller representation, such as 8-bit integers, resulting in decreased model size and memory utilisation. This strategy reduces network latency and improves power efficiency since operations may be conducted using integer data types, requiring fewer computations on processor cores. This improves network speed and reduces memory access, thus improving total power efficiency. Quantisation has advantages, but it can reduce overall precision. Additional 8-bit optimisation of the network structure, tailored to embedded computer
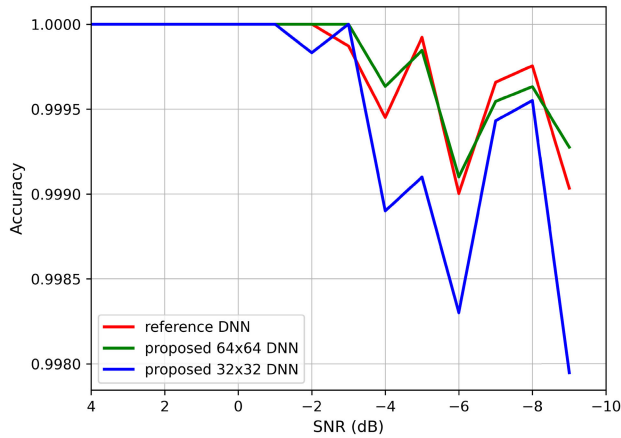
FIGURE 13. Binary detection accuracy for different network models in terms of SNR. In every situation, the SNR range of −3 dB to −9.5 dB contributes the most to accuracy loss.
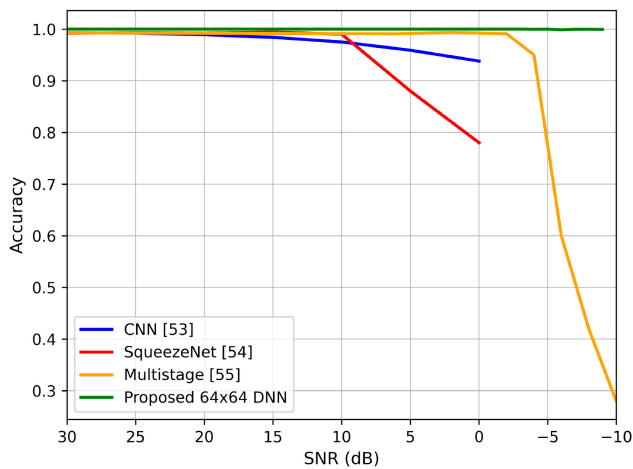

FIGURE 14. Binary detection accuracy for different methods in terms of SNR. The proposed framework preserves over 99% drone presence identification accuracy even at low SNR.

(a) up to 150 m range   (b) up to 350 m range

FIGURE 15. Drone detection accuracy for complete detection chain in a range of 150 m (a) and 350 m (b). The 0.17% and 1.51% of false-positive observations will not have a significant impact on LAN load.

FIGURE 16. Binary detection accuracy for the quantised network model in terms of object distance from a sensor.

deployment, provided further model size reduction to 2.3 MB in the cost of 0.02% accuracy loss in relation to the original.

Comprehensive laboratory experiments facilitated the development of the final method for embedded platform implementation and complete processing chain verification. Frequency-domain signals from six drone models, together with WiFi and Bluetooth devices, were captured with an enhanced SDR sensor. Data were collected in an open space at the local airport area in Gliwice, Poland, where the final system may be installed. To construct a real-life dataset, all transmitters were moved in a range of 10 m to 350 m from the receiving device. The single board embedded computer with an i.MX 8M Plus processor [56] was selected as a capture and calculation platform to achieve real-time performance. The integrated onboard Neural Processing Unit (NPU) complies only with the integer quantised model.

The sampling frequency of the proposed detection chain was 96 MHz, which is quite different from the simulated environment's sampling of 120 MHz. As a result, annotation
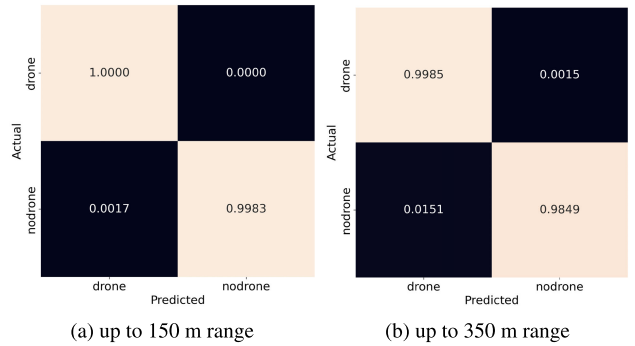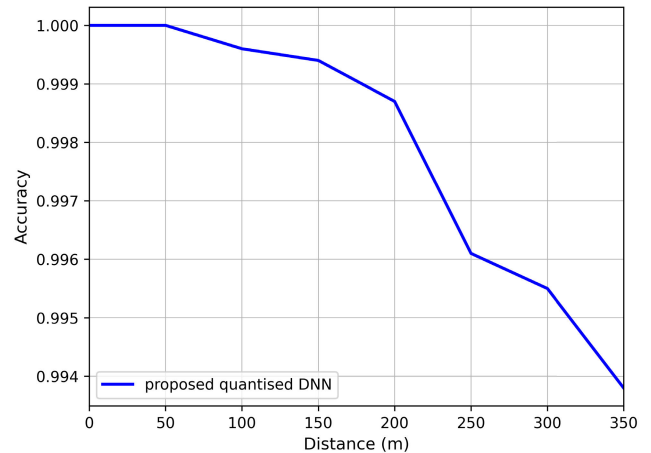
and training are needed as the prior network was unable to be utilised in its previous form. Regardless of ROI size, the processing time was shorter than the period of a single radio frame, hence the approach with a 64×64 area was chosen for improved overall reliability. From a sensor detection range of up to 150 m, a final accuracy of 99.94% is noted, and more specific characteristics are shown in Fig. 15 and Fig. 16. It can be concluded that all drone signals were perfectly detected and classified, with only a 0.17% false alarm rate for 150 m distance. Whereas examining the extended range of up to 350 m, a final accuracy of 99.38% and a 1.51% false alarm rate is noted. Considering the device's intended use case, such an insignificant number of false positive observations will not have a vast influence on the transmission rate to the fusion data centre.

VI. CONCLUSION AND FUTURE WORK

In this paper, a drone surveillance framework based on an RF sensor grid with local intrusion detection and a fusion centre concept for further classification was presented. While transferring raw time-domain data from the entire 2.4 GHz ISM frequency band requires about 2 Gbps in a single sensor scenario, setting up a multisensor grid will result in a massive LAN load even without drone presence.

Local signal preprocessing and transferring only the relevant information, which is contained in 8% of the complete spectrum, make the distributed sensor grid idea practical.

Instead of raw IQ data, the cost-effective USB SDR receiver was utilised to deliver a spectrum estimator to the embedded computer. Therefore, the entire software processing power was utilised for signal sensing and neural computations. A local binary classification strategy built on customised neural network architecture resulted in substantially reduced LAN traffic when no drone presence was detected due to a 0.17% false alarm rate in the range up to 150 m and 1.51% in the range of 350 m. For higher negative SNR immunity and operation in complex electromagnetic environments, the proposed method utilised two-dimensional spectrogram representation rather than time-domain or periodogram approaches. Furthermore, the separation of sensing tasks and trimming of data to preamble regions of interest simplified the overall network architecture. The results were verified in both a simulated environment with noise augmentation and an outdoor scenario. The final signal detection performance indicated an SNR margin of -8.7 dB and drone identification with 99.93% accuracy in simulation. Real-life experiments showed 99.94% accuracy in the sensor range of up to 150 m and 99.38% in the range of 350 m with classification tasks computed on an embedded device.

Future work will focus on further optimising the framework, including the implementation of the signal detection procedure in FPGA hardware. This will result in even lower latency and improved embedded computer offloading for drone countermeasure algorithm development. In addition, the simulation environment will be expanded to include more complex radio channel impairment models and multi-object presence scenarios. Apart from drone intrusion detection, multi-class classification and sensor grid fusion problems will be investigated with more advanced neural network architectures. Finally, after the detection challenges are accomplished, neutralisation by radiofrequency jamming techniques will be established to form a complete counter-drone system.

REFERENCES

[1] Fortune Business Insights. *Commercial Drone Market Size*. Accessed: Jul. 30, 2023. [Online]. Available: https://www.fortunebusiness insights.com/commercial-drone-market-102171

[2] Federal Aviation Administration. *FAA Aerospace Forecast FY 2023-2043*. Accessed Jul. 30, 2023. [Online]. Available: https://www.faa.gov/dataresearch/aviation/aerospaceforecasts/faa-aerospace-forecast-fy-2023-2043

[3] Federal Aviation Administration. *UAS Sightings Report*. Accessed Jul. 30, 2023. [Online]. Available: https://www.faa.gov/uas/resources/public-records/uas-sightings-report

[4] F.-L. Chiper, A. Martian, C. Vladeanu, I. Marghescu, R. Craciunescu, and O. Fratu, "Drone detection and defense systems: Survey and a software-defined radio-based solution," *Sensors*, vol. 22, no. 4, p. 1453, Feb. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/4/1453

[5] C. Dumitrescu, M. Minea, I. M. Costea, I. C. Chiva, and A. Semenescu, "Development of an acoustic system for UAV detection," *Sensors*, vol. 20, no. 17, p. 4870, Aug. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/17/4870

[6] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, "Real-time and accurate drone detection in a video with a static background," *Sensors*, vol. 20, no. 14, p. 3856, Jul. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/14/3856

[7] V. Semkin, M. Yin, Y. Hu, M. Mezzavilla, and S. Rangan, "Drone detection and classification based on radar cross section signatures," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, Jan. 2021, pp. 223–224.

[8] J. Dudczyk, R. Czyba, and K. Skrzypczyk, "Multi-sensory data fusion in terms of UAV detection in 3D space," *Sensors*, vol. 22, no. 12, p. 4323, Jun. 2022.

[9] A. Brighente, M. Conti, G. Peruzzi, and A. Pozzebon, "ADASS: Anti-drone audio surveillance sentinel via embedded machine learning," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Jul. 2023, pp. 1–6.

[10] P. M. Baggenstoss, M. Springer, M. Oispuu, and F. Kurth, "Efficient phase-based acoustic tracking of drones using a microphone array," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

[11] A. Rozantsev, V. Lepetit, and P. Fua, "Detecting flying objects using a single moving camera," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 5, pp. 879–892, May 2017.

[12] H. Sun, J. Yang, J. Shen, D. Liang, L. Ning-Zhong, and H. Zhou, "TIB-Net: Drone detection network with tiny iterative backbone," *IEEE Access*, vol. 8, pp. 130697–130707, 2020.

[13] J. Yan, H. Hu, J. Gong, D. Kong, and D. Li, "Exploring radar micro-Doppler signatures for recognition of drone types," *Drones*, vol. 7, no. 4, p. 280, Apr. 2023. [Online]. Available: https://www.mdpi.com/2504-446X/7/4/280

[14] M. Ummenhofer, L. C. Lavau, D. Cristallini, and D. O'Hagan, "UAV micro-Doppler signature analysis using DVB-S based passive radar," in *Proc. IEEE Int. Radar Conf. (RADAR)*, Apr. 2020, pp. 1007–1012.

[15] P. Flak, "Drone detection sensor with continuous 2.4 GHz ISM band coverage based on cost-effective SDR platform," *IEEE Access*, vol. 9, pp. 114574–114586, 2021.

[16] P. Flak, "Hardware-accelerated real-time spectrum analyzer with a broadband fast sweep feature based on the cost-effective SDR platform," *IEEE Access*, vol. 10, pp. 110934–110946, 2022.

[17] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, "Micro-UAV detection and classification from RF fingerprints using machine learning techniques," in *Proc. IEEE Aerosp. Conf.*, Mar. 2019, pp. 1–13.

[18] B. Kaplan, I. Kahraman, A. Görçin, H. A. Çirpan, and A. R. Ekti, "Measurement based FHSS-type drone controller detection at 2.4 GHz: An STFT approach," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, May 2020, pp. 1–6.

[19] S. Basak, S. Rajendran, S. Pollin, and B. Scheers, "Combined RF-based drone detection and classification," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 1, pp. 111–120, Mar. 2022.

[20] J. Schuette, B. Fell, J. Chapin, S. Jones, J. Stutler, M. Birchler, and D. Roberson, "Performance of RF mapping using opportunistic distributed devices," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2015, pp. 1624–1629.

[21] P. Skokowski, K. Malon, and J. Łopatka, "Building the electromagnetic situation awareness in MANET cognitive radio networks for urban areas," *Sensors*, vol. 22, no. 3, p. 716, Jan. 2022.

[22] C. Koulouris, P. Dimitrios, I. Al-Darraji, G. Tsaramirsis, and H. Tamimi, "A comparative study of unauthorized drone detection techniques," in *Proc. 9th Int. Conf. Inf. Technol. Trends (ITT)*, May 2023, pp. 32–37.

[23] 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.

[24] S. Birnbach, R. Baker, S. Eberz, and I. Martinovic, "PrettyFlyForAWiFi: Real-world detection of privacy invasion attacks by drones," *ACM Trans. Privacy Secur.*, vol. 24, no. 4, pp. 1–34, Nov. 2021.

[25] L. Morge-Rollet, D. Le Jeune, F. Le Roy, C. Canaff, and R. Gautier, "Drone detection and classification using physical-layer protocol statistical fingerprint," *Sensors*, vol. 22, no. 17, p. 6701, Sep. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/17/6701

[26] Y. Zhang, "RF-based drone detection using machine learning," in *Proc. 2nd Int. Conf. Comput. Data Sci. (CDS)*, Jan. 2021, pp. 425–428.

[27] E. Ozturk, F. Erden, and I. Guvenc, "RF-based low-SNR classification of UAVs using convolutional neural networks," *ITU J. Future Evolving Technol.*, vol. 2, no. 5, pp. 39–52, 2021.

[28] R. Akter, M. Golam, A. Zainudin, V.-S. Doan, and D.-S. Kim, "RF signal-based multipurpose UAV surveillance system using deep neural network," in *Proc. 13th Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2022, pp. 555–559.

[29] 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, pp. 86–97, Nov. 2019.

[30] M. S. Allahham, T. Khattab, and A. Mohamed, "Deep learning for RF-based drone detection and identification: A multi-channel 1-D convolutional neural networks approach," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 112–117.

[31] S. Al-Emadi and F. Al-Senaid, "Drone detection approach based on radio-frequency using convolutional neural network," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 29–34.

[32] T. Huynh-The, Q.-V. Pham, T.-V. Nguyen, D. B. D. Costa, and D.-S. Kim, "RF-UAVNet: High-performance convolutional network for RF-based drone surveillance systems," *IEEE Access*, vol. 10, pp. 49696–49707, 2022.

[33] D.-I. Noh, S.-G. Jeong, H.-T. Hoang, Q.-V. Pham, T. Huynh-The, M. Hasegawa, H. Sekiya, S.-Y. Kwon, S.-H. Chung, and W.-J. Hwang, "Signal preprocessing technique with noise-tolerant for RF-based UAV signal classification," *IEEE Access*, vol. 10, pp. 134785–134798, 2022.

[34] F. B. Mehouachi, Q. Yang, J. Galvis, S. Morales, M. Meriac, F. Vega, and C. Kasmi, "Detection of UAVs based on spectrum monitoring and deep learning in negative SNR conditions," *URSI Radio Sci. Lett.*, vol. 3, p. 43, Jan. 2021.

[35] C. Swinney and J. Woods, "Low-cost raspberry-Pi-based UAS detection and classification system using machine learning," *Aerospace*, vol. 9, no. 12, p. 738, Nov. 2022. [Online]. Available: https://www.mdpi.com/2226-4310/9/12/738

[36] A. Wolke. *What's Your IQ—About Quadrature Signals*. Accessed May 30, 2022. [Online]. Available: https://www.tek.com/en/blog/quadrature-iq-signals-explained

[37] Lime Microsystems. *The LimeSDR*. Accessed: Jul. 30, 2023. [Online]. Available: https://limemicro.com/products/boards/limesdr

[38] Ettus Research. *The USRP x310*. Accessed: Jul. 30, 2023. [Online]. Available: https://www.ettus.com/all-products/x310-kit

[39] Tektronix. *Understanding FFT Overlap Processing Fundamentals*. Accessed: Jan. 10, 2023. [Online]. Available: https://www.tek.com/en/documents/primer/understanding-fft-overlap-processing-fundamentals-0

[40] P. des Chênes., "Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaire du second ordre, à coefficients constants," *Sciences, mathématiques et physiques. (Savants étrangers.)*, vol. 1, pp. 638–648, Jan. 1806.

[41] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.

[42] J. Lorincz, I. Ramljak, and D. Begušić, "A survey on the energy detection of OFDM signals with dynamic threshold adaptation: Open issues and future challenges," *Sensors*, vol. 21, no. 9, p. 3080, Apr. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/3080

[43] W. Li, K. Wang, and L. You, "A deep convolutional network for multitype signal detection and classification in spectrogram," *Math. Problems Eng.*, vol. 2020, pp. 1–16, Sep. 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:224886322

[44] H. Iwata, K. Umebayashi, A. Al-Tahmeesschi, and J. Lehtomäki, "High-efficiency FCME-based noise power estimation for long-term and wide-band spectrum measurements," *IEEE Access*, vol. 9, pp. 149883–149893, 2021.

[45] H. Iwata, K. Umebayashi, A. Al-Tahmeesschi, M. López-Benítez, and J. J. Lehtomäki, "Time and frequency varying noise floor estimation for spectrum usage measurement," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshop (WCNCW)*, Apr. 2019, pp. 1–6.

[46] A. J. Onumanyi, A. M. Abu-Mahfouz, and G. P. Hancke, "Amplitude quantization method for autonomous threshold estimation in self-reconfigurable cognitive radio systems," *Phys. Commun.*, vol. 44, Feb. 2021, Art. no. 101256. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1874490720303335

[47] N. W. Evans, J. S. Mason, and M. J. Roach, "Noise compensation using spectrogram morphological filtering," in *Proc. 4th IASTED Int. Conf. Signal Image Process*, 2002, pp. 157–161.

[48] J. Serra and J. Serra, *Image Analysis and Mathematical Morphology*. New York, NY, USA: Academic Press, 1982.

[49] M. Vuorenmaa, J. Marin, M. Heino, M. Turunen, and T. Riihonen, Nov. 2020, "Radio-frequency control video signal recordings drones," *Zenodo*, doi: 10.5281/zenodo.4264467.

[50] G. Bradski, "The OpenCV library," *Dr. Dobb's J. Softw. Tools*, vol. 120, pp. 122–125, Nov. 2000.

[51] S. Scholl. (Oct. 2016). *Exact Signal Measurements Using FFT Analysis*. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:hbz:386-kluedo-42930

[52] I. Nemer, T. Sheltami, I. Ahmad, A. U.-H. Yasar, and M. A. R. Abdeen, "RF-based UAV detection and identification using hierarchical learning approach," *Sensors*, vol. 21, no. 6, p. 1947, Mar. 2021.

[53] S. S. Alam, A. Chakma, M. H. Rahman, R. Bin Mofidul, M. M. Alam, I. B. K. Y. Utama, and Y. M. Jang, "RF-enabled deep-learning-assisted drone detection and identification: An end-to-end approach," *Sensors*, vol. 23, no. 9, p. 4202, Apr. 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/9/4202

[54] O. O. Medaiyese, M. Ezuma, A. P. Lauf, and I. Guvenc, "Wavelet transform analytics for RF-based UAV detection and identification system using machine learning," *Pervas. Mobile Comput.*, vol. 82, Jun. 2022, Art. no. 101569. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574119222000219

[55] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, "Detection and classification of UAVs using RF fingerprints in the presence of Wi-Fi and Bluetooth interference," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 60–76, 2020.

[56] NXP. *I.MX 8 Series Products*. Accessed: Jul. 30, 2023. [Online]. Available: https://www.nxp.com/products/processors-and-microcontrollers/arm-processors/i-mx-applications-processors/i-mx-8-applications-processors:IMX8-SERIES

PRZEMYSŁAW FLAK was born in Katowice, Poland. He received the B.S. and M.S. degrees in electronic engineering from the Silesian University of Technology, Gliwice, Poland, in 2010, where he is currently pursuing the Ph.D. degree. Since 2009, he has been with Flytronic SA, WB Group, Poland, where he is working on research and development topics related to drones, programmable systems, and radio telecommunication. His current research interests include unmanned aerial systems (UAS), counter-UAS systems (C-UAS), software-defined radio (SDR), and field programmable gate arrays (FPGA). He was a recipient of the 5th Annual Digilent Design Contest Award in 2009, organized within the Technical University of Cluj-Napoca, Romania, sponsored by Digilent and Xilinx.

ROMAN CZYBA was born in Zabrze, Poland. He received the B.S. and M.S. degrees in automatic control, the Ph.D. degree in robotics, and the D.Sc. degree in automatic control, electronics, and electrotechnics discipline from the Silesian University of Technology (SUT), Gliwice, Poland, in 1998, 2004, and 2019.

From 2004 to 2005, he was an Adjunct Professor with WSB University, Dąbrowa Górnicza, Poland. From 2004 to 2019, he was an Assistant Professor with the SUT, where he has been an Associate Professor, since 2019. He is currently the Head of the Laboratory of Prototyping Control Systems, Department of Automatic Control and Robotics, SUT. He is the author of 96 articles, one patent, and six inventions. His research interests include control theory and its application to complex dynamical systems, robust controllers, multivariable control systems, mathematical modeling of dynamical systems, unmanned aerial vehicles, autonomous control of unmanned aerial systems, and design and prototyping of embedded control systems.

Prof. Czyba was a recipient of the Minister of Science and Higher Education of Poland Award, in 2014, the 14th International Invention and Innovation Show INTARG®2021 Grand Prix Award and Platinum Award, in 2021, the 15th International Invention and Innovation Show INTARG®2022 Gold Medal Award, in 2022, and the International PRIX EIFFEL Award of the French Federation of Inventors, in 2022.

● ● ●