Engineering Science and Technology, an International Journal 28 (2022) 101028

HOSTED BY

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: www.elsevier.com/locate/jestch

Full Length Article

Drone classification using RF signal based spectral features

Rabiye Kılıç, Nida Kumbasar, Emin Argun Oral, Ibrahim Yucel Ozbek *

Ataturk University, Department of Electrical Electronics Engineering, Yakutiye, Erzurum 25240, Turkey

ARTICLE INFO

Article history:
Received 12 March 2021
Revised 19 May 2021
Accepted 15 June 2021
Available online 31 July 2021

Keywords:
Drone detection
Classification
RF signal
PSD
MFCC
LFCC
SVM

ABSTRACT

Drone detection and classification, important in military and civilian applications, are performed using different sensor signals. Proposed study handles this task using Radio Frequency (RF) signals utilizing basic machine learning methods. It is composed of two main stages as feature extraction succeeded by training/testing of the model. In feature extraction stage, valuable information for classification, contained in the RF signal, is obtained. For this purpose, spectral features, frequently used in speech processing applications, are employed. Specifically, Power Spectral Density (PSD), Mel-Frequency Cepstral Coefficients (MFCC) and Linear Frequency Cepstral Coefficients (LFCC) are adopted by adjusting filter bank margins and parameters for this task. In the second stage, a Support Vector Machine (SVM) classifier is first trained based on the obtained features and finally tested to measure its performance. All experimental studies are carried out using publicly available DroneRF dataset. This dataset contains 2-Class, 4-Class and 10-Class samples for drone existence vs. background (BG), drone types and drone operation modes, respectively. The best classification results are obtained using, PSD, MFCC and LFCC based features for 2-Class, MFCC and LFCC based features for 4-Class and LFCC based features for 10-Class. Accuracy rates for 2-Class, 4-Class and 10-Class are 100%, 98.67% and 95.15%, respectively. These results show that the proposed method outperforms the results given in the literature for DroneRF dataset.

1. Introduction

Drones, also known as unmanned aerial vehicles (UAV), have a broad usage [1]. They are used in many military and civilian applications. While generally employed as an observation tool in the military field, they are also utilized in advertising, transportation, firefighting, search and rescue works, traffic monitoring and atmospheric studies in the civilian applications. The rapid spread of drones brings along security problems as well. For this reason, it is of great importance to detect its existence or its operation purpose if there is one around. For that, traditional methods such as use of sensors, acoustic signals and radars are utilized although they are not reliable enough [2]. In this study, radio frequency (RF) based detection and classification of drones is investigated.

1.1. Related studies

In the literature, there are broad range of applications and methods regarding drone detection and classification. These include use of radar sensors, electro-optical cameras, thermal cameras and acoustic sensors. In this section, first traditional then RF based drone detection studies are examined. Radar sensors are common tools for the detection of flying vehicles. Compared to other sensors, it perceives long range operation and is not affected by adverse weather conditions [3]. However, it is insufficient in detecting small and slow-moving objects such as drones [4]. A probabilistic motion model has been created in order to classify surveillance radar data as UAVs and birds [5]. Also, UAV classification using Linear Frequency Modulated Continuous Wave (LFMCW) 2D surveillance radar data is applied [6]. Another study argues that radar sensors are generally reliable for the drone detection although they are not efficient in classification [7]. Optical sensors, on the other hand, are a general tool used for image processing. For the purpose of UAV classification, images were first captured using optical cameras and they are used for classification utilizing classical methods such as Histogram of Gradients (HOG) using optical cameras [8] or deep neural network (DNN) based image processing methods such as Faster-RCNN [9], VGG-16 [10] and ZF-net [11] are used for CNN methods. All these methods require the use of high resolution cameras to be able to separate drones from background images. In addition, there may be difficulties in distinguishing UAVs from small objects such as birds [7]. Use of thermal sensors are sug-

* Corresponding author.
E-mail addresses: rabiyekilic@atauni.edu.tr (R. Kılıç), nida.kumbasar@grv.atauni.edu.tr (N. Kumbasar), eminoral@atauni.edu.tr (E.A. Oral), iozbek@atauni.edu.tr (I.Y. Ozbek).

https://doi.org/10.1016/j.jestch.2021.06.008
2215-0986/

gested as an alternative approach to solve this problem. They can capture and detect the heat emitted by objects that cannot be visualized with optical sensors. In [12], drones are detected using panoramic thermal cameras. Thermal cameras are also preferred as they perform well in night surveillance and can capture images in adverse weather conditions such as rain, snow, fog, etc. However, their sensitivity to moisture is a major disadvantage [7]. Another type of sensor used for drone detection is acoustic sensors. They are low cost sensors and may distinguish UAV sounds from other sound sources. However, they are highly affected by noise [7]. For example in [13], radar and acoustic sensors are used together to detect UAV rotor-type. In addition, high resolution cameras are used to detect UAVs [14]. Real-time drone detection and tracking is studied using acoustic sensors in K-Nearest Neighborhood (KNN) method [15]. It is reported that detection is difficult if the distance between the acoustic sensor and the UAV is more than 150 m for UAV detection [16]. As an alternative approach, drone emitted RF signals by have been recently used for drone detection [17,18]. A large data set, namely DroneRF, has been created from different drones for that purpose [17]. Using this data set, it is estimated whether there exists a drone in the environment, its type and/or its operation mode if so. In [18,19], UAV detection and classification was performed using DNN method. In other studies, that utilize the DroneRF data set, drone signal detection and classification were performed using 1-DCNN method [20,21] and 10-fold cross-validation. While these studies employ both low band and high band RF signals, [22] only use low band signals. They also used XGBoost algorithm for that purpose. The authors, in [23], proposed a Deep Complex valued Convolutional Neural Network (DC–CNN) based on RF fingerprinting to recognize different drones. They performed drone detection and classification using two different RF drone signal datasets and nine different algorithm models. Drone detection performance in terms of RF signal source distance is studied in [24] by constructing a drone detection mechanism. In [25], the loads carried by drones is estimated up to 200 m distance using five different drones. They constructed features using GMM-UBM Supervector Normalization with MFCC and classified their self-produced dataset signals with a SVM classifier.In another study, RF signal of nine different drones and Wi-Fi signals were used [26]. They performed drone classification using CNN.

1.2. Motivation and contribution

The main motivation of the proposed study is to detect the drone existence, its type and operation mode using RF signals. Since RF drone signals are spread over time and frequency domains, their properties are similar to audio signals. By examining the power spectral density of the drone signals, it has been observed that they show different properties in different frequency ranges. Therefore, it has been aimed to investigate the effectiveness of using classical sound features such as MFCC and LFCC, which can highly concentrate on different frequency bands and extract useful information, for drone classification. Moreover, the current work also investigates the applicability of the simple use of each 0.25 s long RF drone record directly instead of partitioning them into additional frames as needed in DNN based approaches [18–20]. It is also explored the use of a SVM classifier for that purpose since this method is very efficient with low computational burden that utilizes classical features. Finally, it is aimed to further improve the drone detection accuracy.

Based on these motivations, the proposed model is tested, and the corresponding experimental results are obtained. Considering the obtained results, main contributions of the proposed system for RF drone classification are as follows:

- It is shown that spectral features can be effectively used in RF signal based drone detection and classification problems
- It is defined how to adjust the spectral feature parameters such as number of filter bank channels and frequency range, number of cepstral coefficients, etc. for drone signal classification
- It has been demonstrated that direct use of 0.25 s long DroneRF dataset signals instead of partitioning them into multiple frames results in better classification performance
- It is obtained SVM classifier optimization in RF signal based drone classification problem

The rest of the paper is organized as follows: In Section 2, the DroneRF data set and feature extraction methods are provided. Also, SVM method for classification is explained. The preprocessing before the classification process is described in Section 3. Section 4 describes the experimental studies for classification. The obtained results and their comparison with the literature are given in Sections 10 and 11, respectively.

2. Methods

The proposed method performs drone classification employing RF signals. The main block-diagram of this method is shown in Fig. 1. First, the features are extracted from RF signals. For that, Power Spectral Density (PSD), Mel Frequency Cepstrum Coefficients (MFCC) and Linear Frequency Cepstral Coefficients (LFCC) are obtained for low and high frequency bands of the DroneRF signals. Then, a SVM classifier is trained for 2, 4 and 10 classes based on the obtained features. In the 2-Class problem drone existence while in the 4-Class and 10-Class problems drone types and different drone operation modes are detected, respectively. Finally, the proposed model is tested through 10-Fold cross-validation. Before proceeding with the details of the proposed method, dataset used in this study will be explained first as follows.

2.1. Dataset

Publicly available large open source DroneRF dataset [17] is used in this study. It consists of 0.25 s length RF signals, collected at a sampling rate of 40 MHz. This dataset was created by recording signals of AR, Bebop, Phantom drones as well as background at low and high frequencies, corresponds to low-band and high-band of the RF spectrum [27–29]. It consists of 454 RF drone signals of which 227 corresponds to low-band and the other 227 to high-band. All signals, consist of 10 million samples, are recorded for the different operation modes of the drones as described in [17,18] and also denoted in Table 1.

Fig. 2. shows all available sub-classes of the data set. It can be initially considered as a 2-Class dataset such as drone and no-drone. Considering the different type of drones in dataset, they can also be grouped into 4-Class, namely AR, Bebop, Phantom and BG. Including all available modes for each drone, they can be finally divided into 10-Class.Based on this information, the problem in hand can be considered as a detection problem when the DroneRF is treated as a 2-Class dataset (no drone/drone), while it is considered as a classification problem when the DroneRF is treated as a multi-class dataset.

2.2. Feature extraction

The information, contained in both low-band and high-band RF signals of the data set, is distributed over the frequency spectrum. Therefore, it is frequency dependent and the features of drone signals are better extracted in the frequency domain [30,31]. For that, PSD, MFCC and LFCC [31] are used in the proposed drone classifica-

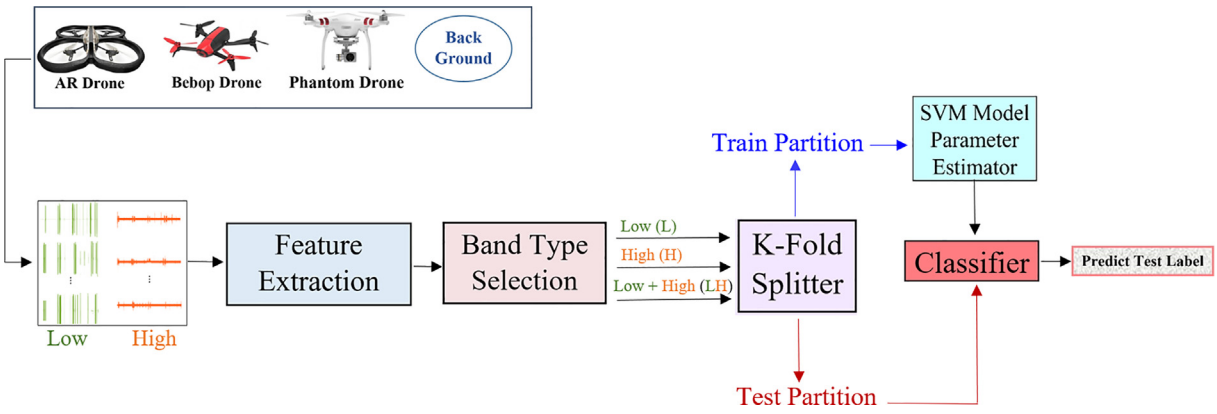

Fig. 1. The main block diagram of the proposed drone classification and detection method.

Table 1
Modes of DroneRF dataset.

M_1:	Open and connected to controller
M_2 :	Without physical intervention
M_3 :	The drone is flying without video recording
M_4:	Flying with video recording

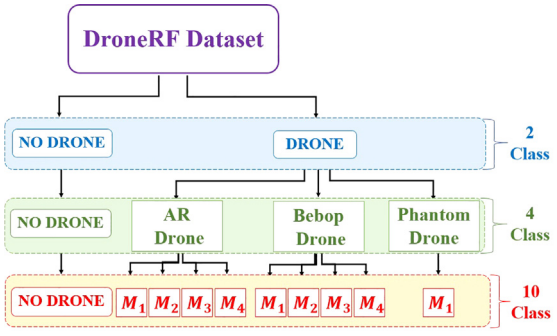

Fig. 2. DroneRF dataset sub-classes.

tion method as they widely preferred in frequency dependent signal processing studies.

2.2.1. Power Spectrum Density (PSD)

Drone RF signals have a random nature, and they have a finite average power. This power term is estimated in terms of power spectral density (PSD) of the RF signal. In this work, PSD is calculated using the Welch algorithm [32,33]. RF drone signal, represented by $x[n]$, is partitioned into frames by multiplying Hamming window function $w[n]$, to obtain windowed drone signals, denoted by $x_m[n]$, as shown in Eq. (1).

$$x_m[n] = w[n]x[n + (m-1)R] \quad n = 0, 1, \ldots, N-1, \quad m = 1, 2, \ldots, M \tag{1}$$

Here, N is the window length, R is the window skip size and M is the number of frames. The N-point DFT of windowed drone signals, denoted by $X_m[k]$, is obtained as given in . Eq. (2).

$$X_m[k] = \sum_{n=0}^{N-1} x_m[n] e^{-j2\pi k/N} \quad k = 0, 1, \ldots, N-1 \tag{2}$$

Let $S_m[k]$ be the PSD estimation of windowed drone signals based on periodogram approach as follows. (Eq. (3))

$$S_m[k] = \frac{1}{N} |X_m[k]|^2 \quad k = 0, 1, \ldots, N-1 \tag{3}$$

Finally, Welch PSD estimation, denoted by $S[k]$, is obtained by averaging the periodograms over frames as given in Eq. (4).

$$S[k] = \frac{1}{M} \sum_{m=0}^{M-1} S_m[k] \quad k = 0, 1, \ldots, N-1 \tag{4}$$

Fig. 3 shows log-scale PSDs for low-band (L) and high-band (H) components of RF signals, calculated for AR, Bebop, Phantom drone RF signals as well as background. Both graphs are obtained by averaging all DroneRF signals in the dataset.

From the evaluation of this figure, it is clear that a particular drone can be differentiated from others in particular frequency bands. As an example, in 2400 MHz region Bebop and Phantom drones are separable from others while BG and Bebop are separable from others in 2420 MHz region. Hence, it is suggested that all drones in the dataset can be classified utilizing different bands of PSD separately instead of whole PSD to obtain a better classification performance. For this purpose, in this study, it is proposed to use LFCC and MFCC features in classification task. These features successfully used in various audio processing applications since they are capable of extracting powerful features from multiple frequency bands of signals. Hence, LFCC and MFCC features are also utilized for drone detection and classification as detailed in the next sub-section.

2.2.2. Cepstral coefficients

Cepstral coefficients, obtained by frequency domain Mel and Linear scaling, are used as features in drone detection and classification. Computation block-diagram is given in Fig. 4.

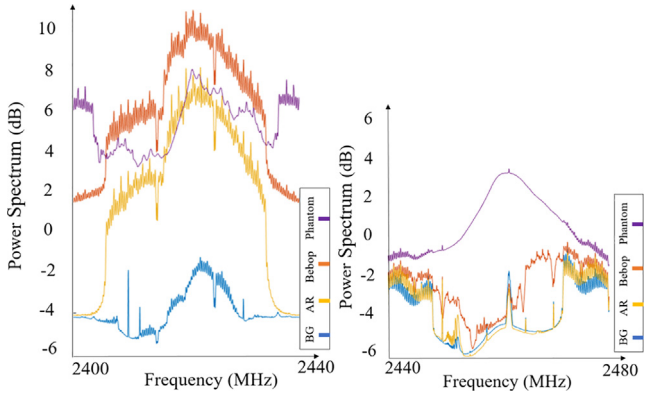

Fig. 3. Low-band and high-band PSDs of AR, Bebop, Phantom drone RF signals along with background.

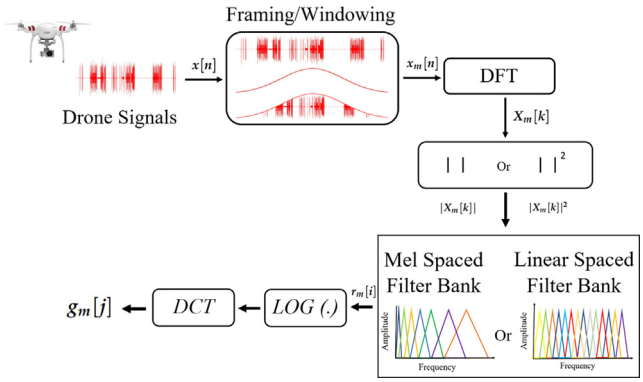

Fig. 4. Block-diagram of cepstral coefficients based feature extraction.

In cepstral feature extraction method, similar to PSD feature extraction, first the DFT of the windowed drone signal is calculated (Eq. (2)), and then its absolute DFT ($|X_m[k]|$) or absolute squared DFT ($|X_m[k]|^2$) is fed to a filter bank. Here, two different type of triangular-shaped filters are used in the filtre bank. In LFCC, fixed-width filters, defined in the frequency domain, are utilized while that of with varying-width according to mel-scale are used in MFCC process [34,35]. Three parameters define such filter banks. These are total number of filters (F), minimum frequency (f_{min}) and maximum frequency (f_{max}) in the filtering band. Corresponding filter bank examples are shown in Fig. 5a and 5b. Although, in general, f_{min} is zero while f_{max} is half of the sampling frequency, in the current study these two values are adjusted according to the drone signal properties to maximize the accuracy rate as explained in Section 4.2. The ith filter in the filter bank, represented by $H_i[k]$ in the frequency domain, are parametrized in terms of half filter width (frequency resolution) and center frequencies. These parameters are defined in [36] and summarized in Eqs. (5)–(7),

$$\Delta f = \frac{f_{max} - f_{min}}{F+1} \tag{5}$$

$$f_{c_i}^{linear} = f_{min} + i\Delta f \quad i = 1, 2, \ldots, F \tag{6}$$

$$f_{c_i}^{mel} = 700(10^{f_{c_i}^{linear}/2595} - 1) \tag{7}$$

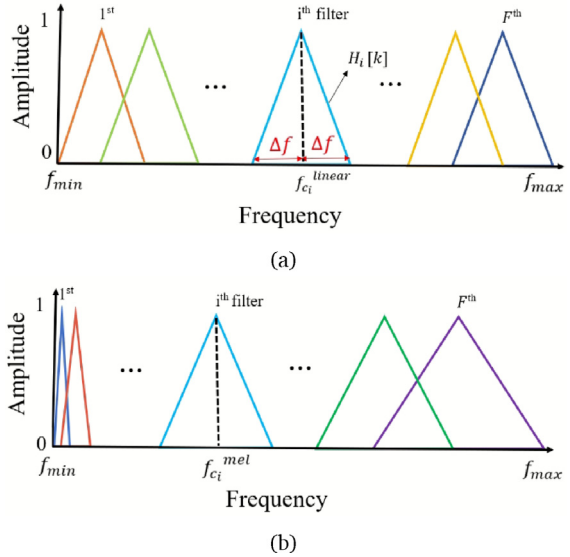

Fig. 5. Filter bank (a) Linear scale; (b) Mel scale.

where the corresponding ith linear filter bank center frequency, denoted by $f_{c_i}^{linear}$, is defined in terms of half filter bandwidth (frequency resolution) width, Δf. Also, Mel filter center frequency, denoted by $f_{c_i}^{mel}$, is obtained scaling $f_{c_i}^{linear}$, giving a bandwidth as $f_{c_{i+1}}^{mel} - f_{c_{i-1}}^{mel}$.

Let $r_m[i]$ be output of the ith filter bank. It is defined as follows.

$$r_m[i] = \sum_{k=0}^{N-1} |X_m[k]|\, H_i[k] \quad i = 1, \ldots, F \tag{8}$$

In this study we also examined the use of absolute squared DFT instead of absolute DFT given in Eq. (8). Corresponding cepstral coefficients, represented by $g_m[j]$, is obtained as the Discrete Cosine Transform (DCT) of the logarithm of the filter bank output as follows.

$$g_m[j] = \sum_{i=1}^{F} \log(r_m[i]) \cos\left(\pi j \frac{(i-0.5)}{F}\right) \quad j = 1, \ldots, F \tag{9}$$

In this study, separate filter banks are designed and used for low-band and high-band drone RF signals.

2.2.3. Feature vectors

In this study two different type of feature vectors are obtained using either spectral components, defined by Eq. (4), or cepstral coefficients, defined by Eq. (9), to form the $\mathbf{z}$ feature vector as defined follows.

$$\mathbf{z} = \begin{cases} [s[0], \ldots, s[N-1]]^T, & \text{for PSD Based Features} \\ [g_m[1], \ldots, g_m[F]]^T, & \text{for Cepstral Based Features} \end{cases} \tag{10}$$

Furthermore, the feature vector is obtained for low-band and high-band RF signals, represented by $\mathbf{z}^L$ and $\mathbf{z}^H$, respectively, as well as their concatenation, represented by $\mathbf{z}^{LH}$ as another feature set, defined in Eq. (11)

$$\mathbf{z}^{LH} \triangleq \begin{bmatrix} \mathbf{z}^L \\ \mathbf{z}^H \end{bmatrix} \tag{11}$$

Therefore, all experimental study results, presented in the following section, is named both based on the feature source and based on the sub-bands. In the former, PSD, LFCC and MFCC represent spectral based components, cepstral coefficients with linear filtering and cepstral coefficients with Mel filtering, respectively. In the latter, L, H and LH represent sub-bands, as defined in Eq. (11).

2.3. Model parameter estimation for SVM classifier

Support Vector Machine (SVM) is one of the widely used algorithms in classification problems [37]. It provides the most appropriate separating line between the samples of two classes distributed over a plane. This line is also called a decision boundary or hyper plane in higher-order dimension. In general, SVMs are used to classify samples as a 2-Class problem. Let the labeled drone data set is defined as $\{\mathbf{z_i}, y_i : i = 1, \ldots, P\}$. Here $\mathbf{z}_i \in R^d$ where i and d represent ith sample feature vector and its dimension, respectively, and $\mathbf{z}$, a generic feature vector, represents one of the above-defined $\mathbf{z^L}, \mathbf{z^H}$ and $\mathbf{z^{LH}}$ cepstral feature vector. Also, y_i and P represent the corresponding sample label and total number of samples in the data set, respectively. Support vectors, closest samples of each class, are obtained based on this data and used to define a hyperplane as given in Eq. (12). [23,38].

$$\mathbf{w}^T\mathbf{z}_i + b = 0 \tag{12}$$

where $\mathbf{w}$ and b represent the surface normal vector and distance to the origin, respectively. This is an optimization problem to maximize the margin, distance between the closest support vectors, while minimizing the classification error, and it can be solved by the Lagrangian multiplier method. This problem can be formulated as shown in Eq. (13).

$$\begin{aligned} &\underset{(\mathbf{w},b)}{\operatorname{argmin}} \left\{ \tfrac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{N}\xi_i \right\} \\ &s.t. y_i(\mathbf{w}^T\mathbf{z}_i + b) - 1 + \xi_i \geqslant 0 \;\; \xi_i \geqslant 0 \forall i \end{aligned} \tag{13}$$

where C is the trade-off parameter and ξ_i is the slack parameter representing a cost term for mis-classification.

This type of minimization problem is well suited to transform the data samples on a higher dimension when a hyperplane for linearly separating the classes is not obtainable [39]. For this purpose, an appropriate feature transformation can be obtained as shown in Eq. (14). in terms of a transformation function $\psi(.)$ [40].

$$\upsilon_i = \psi(\mathbf{z}_i) \tag{14}$$

The scalar multiplication of transformed samples, which is required in the minimization problem, can be simplified into a kernel function, $\kappa(.,.)$, as shown in Eq. (15). [41].

$$\upsilon_{\mathbf{i}}^T \upsilon_{\mathbf{j}} = \psi(\mathbf{z}_i)^T \psi(\mathbf{z}_j) = \kappa(\upsilon_{\mathbf{i}}, \upsilon_{\mathbf{j}}) \tag{15}$$

The most commonly used Kernel functions are Linear, Polynomial and Gaussian RBF (Radial Basis Function) [42]. Among these, Gaussian kernel function is defined in Eq. (16).

$$\kappa(\upsilon_{\mathbf{i}}, \upsilon_{\mathbf{j}}) = \exp(-\gamma||\vec{\upsilon_{\mathbf{i}}} - \vec{\upsilon_{\mathbf{j}}}||^2) \tag{16}$$

where γ is the kernel scale parameter. The main parameters for SVM, namely C and γ, is required to be adjusted properly in this study for a better classification performance.

3. Experimental setup

In this section performance metrics, hardware and classifier setup is introduced along with data preparation. Also, pseudo code of the proposed method is given in Table 2 for a better implementation.

3.1. Performance metrics

The classification performance of the proposed study is given in terms of accuracy, precision, recall and F1 score metrics as defined in Eqs. 17, 18, 19 and 20.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$Precision = \frac{TP}{TP + FP} \tag{18}$$

$$Recall = \frac{TP}{TP + FN} \tag{19}$$

$$F1\ score = 2\frac{PrecisionxRecall}{Precision + Recall} \tag{20}$$

In these equations, *TP* (True Positive) and *FN* (False Negative) represent data of a particular class is predicted as correct class or other classes, correspondingly. *TN* (True Negative) and *FP* (False Positive), on the other hand, represent data from other classes is predicted as other-class and the particular class, correspondingly.

Table 2
Pseudo code of the proposed method.

Start
Step-0: Load Data
Load DroneRF dataset
Create *DataMatrix* and *LabelMatrix*
Select feature extraction type (PSD or MFCC or LFCC)

Step-1: Feature Extraction
// Obtain *LowMat* feature matrix using *DataMatrix* for Low-Band
// Obtain *HighMat* feature matrix using *DataMatrix* for High-Band

if (feature extraction method = PSD) {
  Set N, M, R (given in Section 2.2.1)
  Calculate PSD based *LowMat* and *HighMat* using Eqs. 1, 2, 3 and 4,}
elseif (feature extraction method = LFCC) {
  Set $N, M = 1, f_{min}, f_{max}, F$ (given in Section 2.2.2)
  Use LINEAR filter bank in Eq. (8)
  Calculate LFCC based *LowMat* and *HighMat* using Eqs. 5, 6, 7, 8 and 9,}
elseif (feature extraction method = MFCC) {
  Set N, $M = 1$, f_{min}, f_{max}, F (given in Section 2.2.2)
  Use MEL filter bank in Eq. (8)
  Calculate MFCC based *LowMat* and *HighMat* using Eqs. 5, 6, 7, 8 and 9,}

Create *LowHighMat* feature matrix using *LowMat* and *HighMat* in Eq. (11)

Step-2: Split Data with K-Fold Cross Validation
Split *LowMat* or *HighMat* or *LowHighMat* and *LabelMatrix* into 10-folds

foreach i in [1, 2 . . ., 10] // Loop for each fold
{

Step-3: SVM Classifier Parameter Optimization
Set kernel function to Gaussian RBF
Set One-vs-One for multi-class classification
foreach C in $2^{[-15:15]}$ and γ in $2^{[-15:15]}$ {
  Optimize C and γ by using train data for ith-fold to obtain C_{opt}, γ_{opt}}

Step-4: Train SVM Classifier
Set kernel function to Gaussian RBF
Set One-vs-One for multi-class classification
Train ith-fold SVM model using C_{opt}, γ_{opt}

Step-5: Classify Test Data
Classify ith-fold test data using trained SVM model to obtain predicted test label

Step-6: Calculate Test Performance
// Calculate ith-fold test performance using predicted and true test labels
Obtain accuracy, precision, recall, F1 scores using Eqs. (17-20),
}
Calculate average performance using all 10-fold scores
End

3.2. Setup

All experiments for the proposed drone detection and classification system are performed on a computer with 10 core i9-7900X central processing unit (CPU), 128 GB of RAM, running on Ubuntu18.04.4 - LTS Linux operating system. Pre-processing of DroneRF dataset and development of proposed algorithms are performed using MATLAB 2019b software backed by two GeForce RTX 2080Ti graphics processing units (GPUs). During the implementation of the proposed study, Matlab VOICEBOX toolbox [43] was used.

3.3. Train and test data preparation

The K-Fold Cross Validation method, which is widely used in separating data sets as train and test, is used. In the current study K is chosen as 10, and hence, it is randomly divided into 10 folds. Therefore, training sessions are performed using 9 out of 10 folds during training while the remaining fold is used in testing. This procedure is repeated for 10 times to ensure that the test fold is shifted through all available 10 folds. This method plays an important role in determining the average accuracy of the proposed model.

3.4. SVM parameters determination

The SVM model is trained after the extracted features are partitioned as train and test thorough K-fold cross validation. Different kernels such as linear, polynomial and Gaussian functions are tested as SVM kernel functions. Among these the Gaussian function, yielding the highest accuracy, is determined as the kernel. Experimental studies were performed for various values of kernel scale γ and penalty parameter C in the range from 2^{-15} to 2^{15} with power increments. In the study, γ and C parameters are selected as 2^4 and 2^{14}, respectively, by trial and error. Also, one-vs-one method is used since it gives more accurate results than one-vs-all for multi-class drone classification problem.

4. Experimental results

In this section, experimental results of the proposed method are given. PSD, LFCC and MFCC are used as feature to represent RF signals. First classification results with baseline features, then experimental results with adjusted parameters are given to show how performance of the proposed drone detection method is improved. And also, final results are compared with other studies available in literature.

4.1. Baseline features

In PSD calculation as given in Eq. (4), the DFT size is defined as 4096 points. Hence, the dimension of the obtained baseline PSD feature vector is also 4096x1, both for low-band and high-band data, separately. The baseline parameters of MFCC and LFCC features are set as follows. The data is first windowed using a Hamming function, as shown in Fig. 4, and pass through a filter bank, constructed by triangular filters located on the frequency spectrum from zero to half sampling frequency point, as shown in Fig. 5a and 5b. Magnitude spectrum is used in all filtering operations, and the number of coefficients is set to 12 as baseline parameters of the study. Fig. 6. shows the accuracy results of 2-Class classification problem, drone vs. background, using PSD, MFCC and LFCC features. As shown in this figure, binary classification accuracy rate of the proposed baseline method is 97.79%, 99.55% and 98.67% with the use of high-band only PSD, MFCC and LFCC features, respectively. Also 100% accuracy rate was achieved with all low-band features as well as combining both frequency band features.

The comparison of accuracy rates for 4-Class problem, namely BG, AR, Bebop and Phantom classification, are given in Fig. 7. Using PSD, MFCC and LFCC high-band baseline features only yields an accuracy rate of 92.51, 84.14% and 92.95%, respectively, while that of using low-band baseline features results in an accuracy rate of 97.95%, 96.03% and 95.15%, respectively. Using both frequency bands, on the other hand, 97.35%, 95.15% and 94.27% accuracy rates are obtained, respectively. Fig. 8. shows the accuracy rate comparison for 10-Class problem that includes the classification task of BG, four modes of AR drone, four modes of Bebop and Phantom drone. Considering the high-band features only, 78.85% average accuracy was achieved with PSD, 62.55% with MFCC, and 77.97% with LFCC features. On the other hand, using low-band features only, 73.12% average accuracy was achieved with PSD, 72.68% with MFCC and 79.29% with LFCC features while classification using both frequency bands produced 73.12% average accuracy with PSD, 81.93% with MFCC and 87.27% with LFCC features.

Considering all results, 2-Class classification is a simpler task compared to the others such that almost 100% accuracy rates can be obtained regardless of the frequency bands or features used. Conversely, increasing number of classes introduces a more complicated problem so that a better frequency evaluation results in a better classification performance. Using MFCC and LFCC features reveals more satisfactory classification performance especially for 10-Class classification task. Considering the computational time, on the other hand, MFCC and LFCC features consumes considerably less time than PSD features due to their smaller feature vector dimensions compared to PSD as shown in Table 3. Hence, MFCC and LFCC features are prioritized and tailored to improve the drone classification problem in the rest of the proposed study.

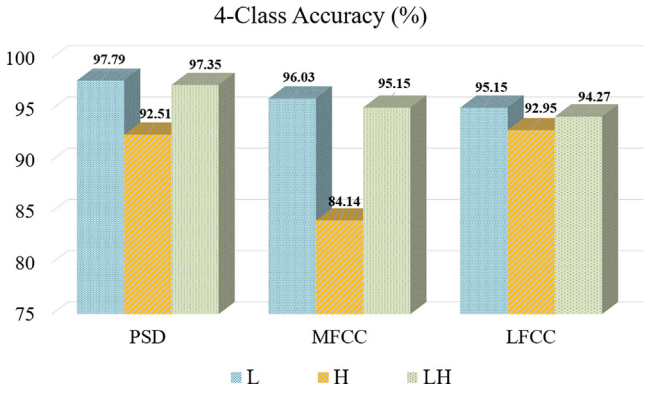

Fig. 7. Drone type classification using baseline PSD, MFCC, LFCC features.

4.2. Adjusted features

This section shows the improvement of the proposed method by adjusting the LFCC and MFCC feature parameters and use of log energy as a new feature. Feature parameters consist of number of cepstral coefficients, inclusion of zero-order coefficient, filter

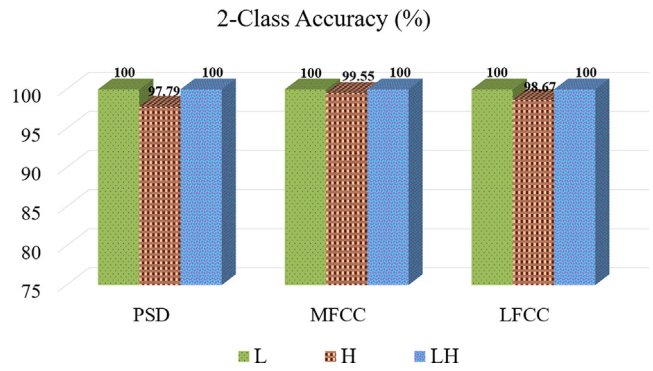

Fig. 6. Drone detection results using baseline PSD, MFCC, LFCC features.

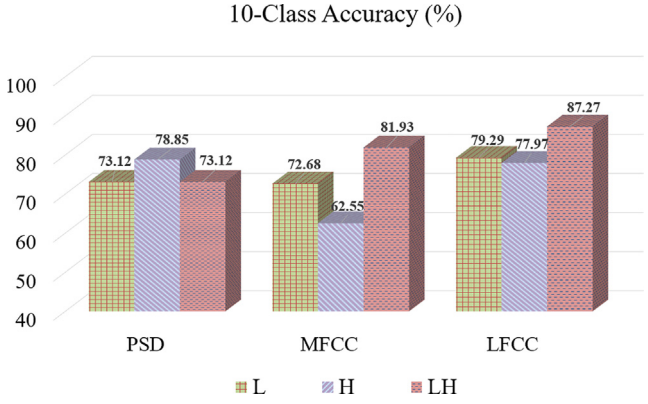

Fig. 8. Drone mode classification using baseline PSD, MFCC, LFCC features.

Table 3
Computation time for 10-Class.

	Low Band (min.)	Number of Features	High Band (min.)	Number of Features	LH Band (min.)	Number of Features
PSD	46.1	4096	45.65	4096	74.6	8192
MFCC	13.06	12	12.35	12	20	24
LFCC	12.18	12	12.5	12	19.23	24

bank frequency boundaries and use of squared DFT instead of absolute DFT before the application of filter bank. To better clarify the obtained results, abbreviation of these features are given in Table 4.

4.2.1. Use of absolute-squared/absolute DFT

The accuracy rates are compared by using absolute value of DFT, included in baseline feature set, and absolute-squared DFT (*shown as B+P*). The corresponding results are given in Table 5. It is clear that using square or absolute DFT does not change the feature dimension. Hence, all entries in this table are obtained using the same size feature vector of the baseline features. That is, accuracy rates are obtained using a 24 × 1 (low-band:12 × 1, high-band:1 2 × 1) feature vector, composed of concatenated MFCC and LFCC baseline parameters for low-band and high-band signals. This table shows that use of absolute-squared DFT instead of its absolute value improves the average accuracy rates. Therefore, absolute-squared DFT is used in the following experimental results.

4.2.2. Log energy and zero-order cepstral coefficients as additional features

The effect of inclusion log energy, *(E)*, and zero-order cepstral coefficients, *(0)*, on the success rate is also examined in the study. Expanding the baseline parameters with two new features, both for low-band and high-band signals, results in an expanded feature vector, and the corresponding results are shown in Table 6. In this table, (*B+P+0*) entry corresponds to a 26 × 1 feature vector (low-band:13 × 1, high-band:13 × 1) while (*B+P+0+E*) entry corresponds to a 28 × 1 feature vector (low-band:14 × 1, high-band:14 × 1).

Table 4
Abbreviations used in results.

B	Baseline (absolute DFT)
P	Power (absolute-squared) DFT
0	Zero-Order Coefficient
E	Log Energy (in time domain)
NC	Number of Coefficient

Table 5
Absolute-squared and absolute DFT comparison.

		MFCC Accuracy (%)	LFCC Accuracy (%)
2-Class	B	**100**	**100**
	B+P	**100**	**100**
4-Class	B	97.35	94.27
	B+P	**98.23**	**95.59**
10-Class	B	81.93	87.27
	B+P	**84.14**	**88.10**

Table 6
Log energy and the effect of zero-order cepstral coefficient.

	MFCC Accuracy (%)		LFCC Accuracy (%)	
	4-Class	10-Class	4-Class	10-Class
B+P	98.23	84.14	95.59	88.10
B+P+0	97.79	83.7	98.23	89.42
B+P+E	**98.67**	**85.46**	**98.67**	**92.51**
B+P+0+E	96.91	81.93	97.79	88.10

Based on these results, use of log energy feature improves the accuracy for both 4-Class and 10-Class problems. Hence, following sub-section results are obtained for *B +P+E*.

4.2.3. Effect of cepstral coefficients

This sub-section explores the effect of number of cepstral coefficient on accuracy rate. Fig. 9. shows the corresponding accuracy rates in terms of number of coefficients both in $MFCC_{B+P+E}$ and $LFCC_{B+P+E}$ in 4-Class and 10-Class problems. The optimum number of coefficients is determined as 12 in 4-Class classification both for MFCC and LFCC, achieving a 98.67% accuracy rate. In 10-Class problem, on the other hand, this parameter is defined as 13 for MFCC and 12 for LFCC, achieving 88.98% and 92.51% accuracy rates, respectively. Based on these results, the baseline feature set with 12 coefficients is verified as the most promising feature vector size and remains unchanged in the rest of the study.

4.2.4. Filter bank frequency range

In baseline feature extraction, the frequency range, defined by f_{min} and f_{max} (in Fig. 5a and 5b), is chosen as the whole available frequency band borders. That is, f_{min} = 2400 MHz and f_{max} = 2440 MHz for low-band signals while f_{min} = 2440 MHz and f_{max} = 2480 MHz for high-band signals. In this sub-section, it is also investigated the best performing frequency range in terms of classification accuracy both for MFCC and LFCC by scanning the available frequency range. Based on the obtained results, the best accuracy rate is 98.67% when the filter bank is placed in the full frequency range from 2400 to 2480 MHz for MFCC and LFCC in 4-Class problem. On the other hand, 2400 to 2424 MHz and 2440 to 2467 MHz bands give the best performance for low-band and high-band signals, respectively, for the 10-Class problem with LFCC, as schematically shown in Fig. 10a by shaded regions. Moreover, 2400 to 2436 MHz and 2440 to 2472 MHz bands give the best performance for low-band and high-band signals, respectively, for 10-Class problem with MFCC as schematically shown in Fig. 10b. by shaded regions. Considering these filter bank ranges, 90.33% and 95.29% accuracy rates are obtained for MFCC and LFCC, respectively. From the comparison of best accuracy rates with that of previous sub-section, it is clear that the use of adjusted frequency ranges of filter banks results in an increase of 1.35% for MFCC and 3.00% for LFCC for only 10-Class problem. Hence, one may suggest that the use of the full frequency band is more advantageous

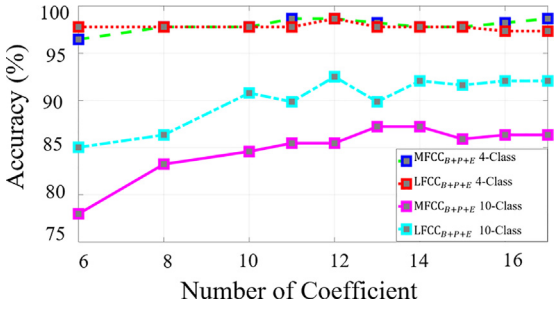

Fig. 9. Accuracy rates in terms of number of cepstral coefficients.

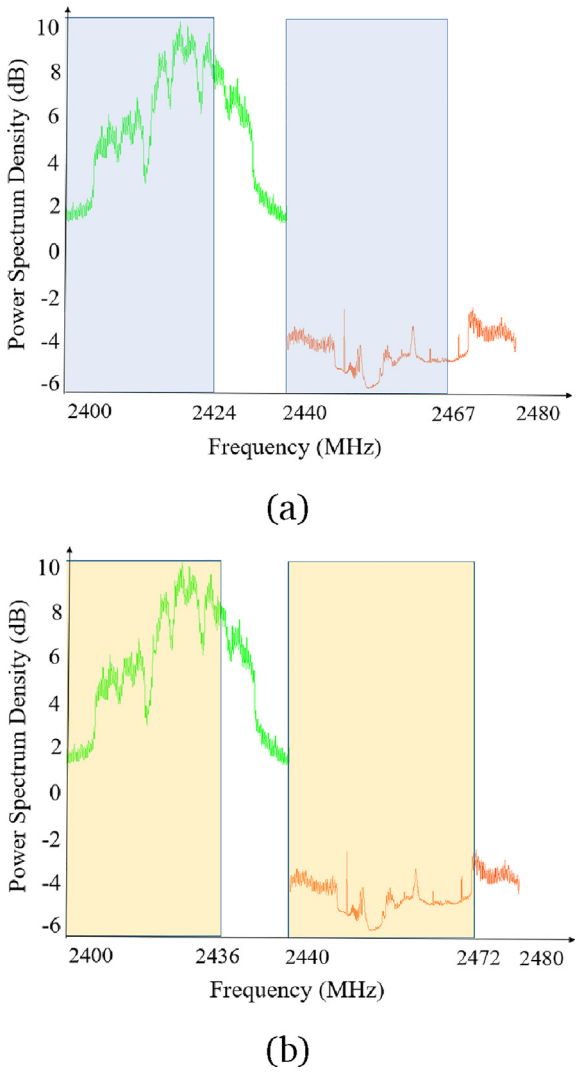

Fig. 10. Low-band and high-band signals' filter bank ranges (a) LFCC; (b) MFCC.

for 4-Class classification problem while the use of a particular frequency range is more convenient for 10-Class problem.

4.2.5. Comparison of adjusted features with baseline features

As a summary of parameter adjustment, accuracy rates with the use of baseline features are compared that of with the use of adjusted features, and corresponding results are shown in Fig. 11. for MFCC and Fig. 12. for LFCC. Two main conclusions can be driven from the evaluation of these figures. First, all accuracy results are increased with parameter adjustment. Second, LFCC offers better accuracy rates for 10-Class problem while both MFCC and LFCC give same results for 4-Class problem. In general, MFCC features perform better than in LFCC using full available frequency range in speech processing applications. On the other hand, LFCC features, captured from a particular frequency range, performs better than MFCC in drone detection and classification problem.

Two main conclusions can be driven from the evaluation of these figures. First, all accuracy results are increased with the use of parameter adjustment. Second, LFCC offers better accuracy rates for 10-Class problem while both MFCC and LFCC give same results for 4-Class problem. In general, MFCC features perform better than in LFCC using full available frequency range in speech processing applications, in this particular drone detection and classification problem it is observed the opposite for a particular frequency band.

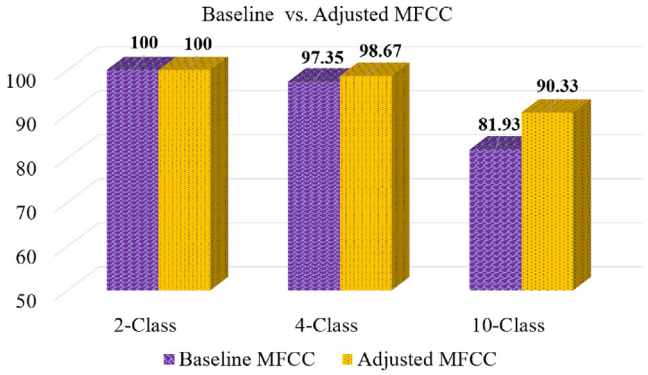

Fig. 11. Baseline and adjusted MFCC accuracy rates.

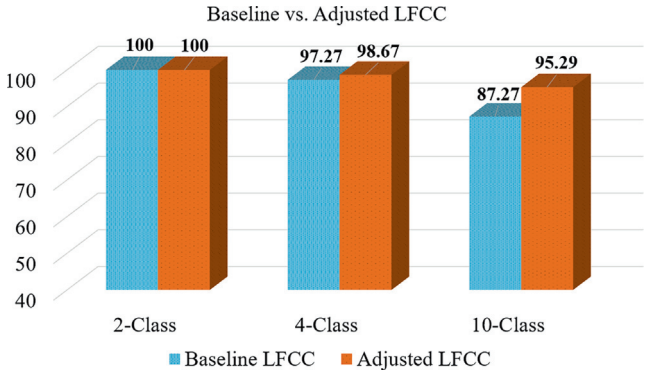

Fig. 12. Baseline and adjusted LFCC accuracy rates.

5. Discussion

This section primarily compares the obtained results in the proposed study with the literature results. In order to achieve the best comparison, average accuracy, F1 score, prediction and recall values are compared with literature that use same DroneRF dataset, as shown used in the proposed study, as shown in Table 7. Since the number of used partitions differ between the available studies, it may not be easy to compare the results directly. However, studies, performed using the DroneRF dataset with different number of partitions, compare their obtained results with each other [18,19,22], as follows. Mohammad F. Al-Sa'd et al. [18] used three DNNs through the use of 10-fold Cross-Validation. They used power spectrum as the feature set which is obtained from combined magnitude DFT of low and high-band signals. They obtained percent accuracy rates of 99.7% 84.5% and 46.8% for 2, 4 and 10-Class classification experiments, respectively. Sara Al-Emadi et al. [20] proposed a method using a CNN architecture, consists of five one-dimensional (1D) convolutional layers, utilizing 10-fold Cross-Validation. While they slightly improved accuracy results of [18] for 2 and 4-Class problems, it is increased 26.5% for 10-Class problem. They also improved F1 score about 28.1% for the same 10-Class problem. Olusiji O. Medaiyes et al. [22] proposed three machine learning models using the XGBoost algorithm and evaluated these models using 10-fold cross validation. Their major contribution to accuracy is 18.39% compared to [20] in 10-Class problem. MHD S. et al. [19] channelized the RF Drone signals' full spectrum into multiple channels and considered each one as a separate input of Multi-Channel 1-D Convolutional Neural Network. They obtained best test results of 87.4% accuracy rate and 77% F1 score for 10-Class problem. All studies of Table 7. first partition

Table 7
Comparison of obtained results with the literature.

Literature	Accuracy (%)			F1 score (%)			Recall (%)			Precision (%)			Number of Segments
	2-Class	4-Class	10-Class	2-Class	4-Class	10-Class	2-Class	4-Class	10-Class	2-Class	4-Class	10-Class	
M. F. Al-Sa'd et al. [18]	99.70	84.50	46.80	99.50	78.80	43.00	99.40	76.43	41.97	99.65	91.08	53.56	100
S. Al-Emadi et al. [20]	99.80	85.80	59.20	99.70	84.60	55.10	99.55	81.70	55.56	99.85	91.30	58.90	100
O.O. Medaiyese et al. [22]	99.96	90.73	70.09					N/A					N/A
M.S. Allahham et al. [19]	100	94.60	87.40	100	91.00	77.00	100	90.00	74.00	100	94.00	83.00	≈ 370
Proposed Study													
PSD	100	97.79	78.85	100	97.61	76.93	100	97.61	77.13	100	97.72	77.85	1
MFCC	100	98.67	90.30	100	98.58	89.68	100	98.57	89.67	100	98.75	89.89	1
LFCC	100	98.67	95.15	100	98.70	94.72	100	98.80	94.67	100	98.67	94.89	1

the RF Drone samples into varying number of segments (partitions), shown in table, as DNNs require large amount of data for training. Then, each partition is classified separately. The proposed study, on the other hand, processes each sample as a single partition considering the main goal as classifying the DroneRF dataset samples with the best performance. To achieve this, features, adjusted for drone detection and classification, are extracted from each full-length dataset sample, and they are processed by a SVM classifier. Following [18–20], the results of the proposed method is compared with the literature as follows. Proposed method improves the best 4-Class accuracy rate of 94.6%, reported in the literature, to 98.67% with an increase of 3.54%. Also, the accuracy rate of 87.4% in 10-Class classification is increased to 95.15% with an increase of 8.87%. Moreover, F1 score of the proposed method is increased from 91% to 98.70% in the 4-Class problem, yielding an increase of 8.46%. Similarly, in the 10-Class classification problem, the F1 score value, which is 77%, is increased to 94.72% with an increase of 23.01%.

6. Conclusion and future work

The spread use of drones has led to an increase in security measures. For this purpose, drone detection has become a need, and it is obtained using various techniques such as optical, radar, thermal, acoustic sensors and RF signals. This study focuses on drone detection and classification based on RF signal. Since drone RF signals have similar time and frequency-dependent characteristic of audio signals, the proposed method employs widely used spectral based audio features such as PSD, MFCC and LFCC in a SVM based machine learning algorithm for this purpose. In this manner, feature extraction parameters, namely number of cepstral coefficients, filter bank frequency range and center frequencies etc., are tailored for drone RF signals. It is also investigated the efficient use of the RF signal frequency bands, namely separate or together, for the best performance. To measure the performance of the proposed method, DroneRF dataset with 0.25 s samples are processed as single segment. This approach not only decreased the computational load but also increased the classification performance of the proposed algorithm. As the results of the experimental studies, 100% accuracy rate is obtained in drone presence problem (2-Class), and 98.67% and 95.15% accuracy rates are obtained in drone type detection (4-Class) and operational mode detection (10-Class) problems, respectively, using LFCC based features.

Based on the obtained results of the current study, handcrafted features such as PSD, MFCC and LFCC show promising results in drone detection and classification. Also, deep learning approaches give satisfactory results in classification problems especially on drone detection as given in literature. Hence, it is worth to investigate the use of such hand-crafted features with the state of art deep learning approaches as a future study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] U. Commercial, Commercial drone market analysis by product (fixed wing, rotary blade, nano, hybrid), by application (agriculture, energy, government, media & entertainment) and segment forecasts to 2022, Grand View Research: San Francisco, CA, USA..
[2] G. Ding, Q. Wu, L. Zhang, Y. Lin, T.A. Tsiftsis, Y.D. Yao, An amateur drone surveillance system based on the cognitive internet of things, IEEE Communications Magazine 56 (1) (2018) 29–35.
[3] E.F. Knott, J.F. Schaeffer, M.T. Tulley, Radar Cross Section, SciTech Publishing, 2004.
[4] P. Molchanov, K. Egiazarian, J. Astola, R. Harmanny, J. De Wit, Classification of small uavs and birds by micro-doppler signatures, in: 2013 European Radar Conference, in: 2013 European Radar Conference, IEEE, 2013, pp. 172–175.
[5] W. Chen, J. Liu, J. Li, Classification of uav and bird target in low-altitude airspace with surveillance radar data, The Aeronautical Journal 123 (1260) (2019) 191–211.
[6] M. Messina, G. Pinelli, Classification of drones with a surveillance radar signal, in: International Conference on Computer Vision Systems, Springer, 2019, pp. 723–733.
[7] S. Samaras, E. Diamantidou, D. Ataloglou, N. Sakellariou, A. Vafeiadis, V. Magoulianitis, A. Lalas, A. Dimou, D. Zarpalas, K. Votis, et al., Deep learning on multi sensor data for counter uav applications–a systematic review, Sensors 19 (22) (2019) 4837.
[8] F. Gökçe, G. Üçoluk, E. Şahin, S. Kalkan, Vision-based detection and distance estimation of micro unmanned aerial vehicles, Sensors 15 (9) (2015) 23805–23846.
[9] M. Saqib, S.D. Khan, N. Sharma, M. Blumenstein, in: A study on detecting drones using deep convolutional neural networks, in 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017, pp. 1–5.
[10] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556..
[11] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, in: European Conference on Computer Vision, Springer, 2014, pp. 818–833..
[12] A. Thomas, V. Leboucher, A. Cotinat, P. Finet, M. Gilbert, Uav localization using panoramic thermal cameras, in: International Conference on Computer Vision Systems, Springer, 2019, pp. 754–767.
[13] S. Park, S. Shin, Y. Kim, E.T. Matson, K. Lee, P.J. Kolodzy, J.C. Slater, M. Scherreik, M. Sam, J.C. Gallagher, et al., Combination of radar and audio sensors for identification of rotor-type unmanned aerial vehicles (uavs), in: 2015 IEEE SENSORS, IEEE, 2015, pp. 1–4.
[14] H. Liu, Z. Wei, Y. Chen, J. Pan, L. Lin, Y. Ren, Drone detection based on an audio-assisted camera array, in: 2017 IEEE Third International Conference on Multimedia Big Data (BigMM), IEEE, 2017, pp. 402–406.
[15] J. Kim, C. Park, J. Ahn, Y. Ko, J. Park, J.C. Gallagher, Real-time uav sound detection and analysis system, in: 2017 IEEE Sensors Applications Symposium (SAS), IEEE, 2017, pp. 1–5.
[16] S. Jeon, J.W. Shin, Y.J. Lee, W.H. Kim, Y. Kwon, H.Y. Yang, Empirical study of drone sound detection in real-life environment with deep neural networks, in: 2017 25th European Signal Processing Conference (EUSIPCO), IEEE, 2017, pp. 1858–1862.
[17] M.S. Allahham, M.F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, A. Erbad, Dronerf dataset: A dataset of drones for rf-based detection, classification and identification, Data in Brief 26 (2019) 104313.

[18] M.F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, A. Erbad, Rf-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database, Future Generation Computer Systems 100 (2019) 86–97.
[19] M.S. Allahham, T. Khattab, A. Mohamed, Deep learning for rf-based drone detection and identification: A multi-channel 1-d convolutional neural networks approach, in: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), IEEE, 2020, pp. 112–117.
[20] S. Al-Emadi, F. Al-Senaid, Drone detection approach based on radio-frequency using convolutional neural network, in: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), IEEE, 2020, pp. 29–34.
[21] R. Akter, V.S. Doan, G.B. Tunze, J.M. Lee, D.S. Kim, Rf-based uav surveillance system: A sequential convolution neural networks approach, in: 2020 International Conference on Information and Communication Technology Convergence (ICTC), IEEE, 2020, pp. 555–558.
[22] O.O. Medaiyese, A. Syed, A.P. Lauf, Machine learning framework for rf-based drone detection and identification system, arXiv preprint arXiv:2003.02656..
[23] S.R. Gunn et al., Support vector machines for classification and regression, ISIS Technical Report 14 (1) (1998) 5–16.
[24] A.H. Abunada, A.Y. Osman, A. Khandakar, M.E.H. Chowdhury, T. Khattab, F. Touati, Design and implementation of a rf based anti-drone system, in: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), IEEE, 2020, pp. 35–42.
[25] P. Nguyen, V. Kakaraparthi, N. Bui, N. Umamahesh, N. Pham, H. Truong, Y. Guddeti, D. Bharadia, R. Han, E. Frew, et al., Dronescale: drone load estimation via remote passive rf sensing, in: Proceedings of the 18th Conference on Embedded Networked Sensor Systems, 2020, pp. 326–339.
[26] S. Basak, S. Rajendran, S. Pollin, B. Scheers, Drone classification from rf fingerprints using deep residual nets, arXiv preprint arXiv:2011.13663..
[27] PhotoPoint, Online. URL: https://www.photopoint.ee/en/drones/1532414-parrot-bebop-2..
[28] Amazon, Online. URL: https://www.amazon.com/Parrot-AR-Drone-2-0-Elite%20Quadcopter/dp/B00FS7SU7K..
[29] DJI, Online. URL: https://www.dji.com/phantom-3-standard..
[30] M. Boussaa, I. Atouf, M. Atibi, A. Bennis, Ecg signals classification using mfcc coefficients and ann classifier, in: 2016 International Conference on Electrical and Information Technologies (ICEIT), IEEE, 2016, pp. 480–484.
[31] J. Rong, G. Li, Y.P.P. Chen, Acoustic feature selection for automatic emotion recognition from speech, Information Processing & Management 45 (3) (2009) 315–328.
[32] C. Gallet, C. Julien, The significance threshold for coherence when using the welch's periodogram method: effect of overlapping segments, Biomedical Signal Processing and Control 6 (4) (2011) 405–409.
[33] J.O. Smith III, Spectral Audio Signal Processing, W3K publishing, 2011.
[34] Q. Li, H. Zhu, F. Qiao, X. Liu, Q. Wei, H. Yang, Energy-efficient mfcc extraction architecture in mixed-signal domain for automatic speech recognition, in: x, in: 2018 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH), IEEE, 2018, pp. 1–3.
[35] S. Young, G. Evermann, M. Gales, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, P. Woodland, The htk book version 3.4 manual, Cambridge University Engineering Department, Cambridge, UK..
[36] S. Nagarajan, S.S.S. Nettimi, L.S. Kumar, M.K. Nath, A. Kanhe, Speech emotion recognition using cepstral features extracted with novel triangular filter banks based on bark and erb frequency scales, Digital Signal Processing 104 (2020) 102763.
[37] V. Vapnik, The Nature of Statistical Learning Theory, Springer science & business media, 2013.
[38] D. Tao, X. Tang, X. Li, X. Wu, Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (7) (2006) 1088–1099.
[39] C. Zhou, K. Yin, Y. Cao, B. Ahmed, Application of time series analysis and pso–svm model in predicting the bazimen landslide in the three gorges reservoir, China, Engineering Geology 204 (2016) 108–120.
[40] V. Vapnik, The nature of statistical learning theory. Statistics for engineering and information science, 2000.
[41] E.A. Oral, M.M. Çodur, I.Y. Ozbek, Sleep stage classification based on filter bank optimization, in: 2017 25th Signal Processing and Communications Applications Conference (SIU), IEEE, 2017, pp. 1–4.
[42] M.S. Elbisy, Sea wave parameters prediction by support vector machine using a genetic algorithm, Journal of Coastal Research 31 (4) (2015) 892–899.
[43] VOICEBOX, Online. URL: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html..