e-ISSN: 2149-4584

Erzincan University Journal of Science and Technology

EJSAT 2025, 18 (3) 892-916

Research Article

Multi-frame Fusion Methods Based on Cepstral Coefficients for Drone Classification

Nida Kumbasar[1,3], Rabiye Kılıç[2,3*], Emin Argun Oral[3], Ibrahim Yucel Ozbek[3,4]

[1] Informatics and Information Security Research Center (BİLGEM), TUBİTAK, 41470, Kocaeli, Turkey
[2] Department of Computer Engineering, Ataturk University, 25240, Erzurum, Turkey
[3] Department of Electrical Electronics Engineering, Ataturk University, 25240, Erzurum, Turkey
[4] Department of Artificial Intelligence and Data Engineering, 25240, Erzurum, Turkey

Received: 20/11/2024, **Revised:** 07/02/2025, **Accepted:** 18/02/2025, **Published:** 31/12/2025

Abstract

The increasing popularity of drones in recent years has resulted in privacy and security vulnerabilities. Today, drones can be easily purchased and used, leading to concerns about intrusion into private areas. Detecting the presence of drones and identifying their operation mode is of great importance. Various detection techniques, including video, sound, thermal imaging, and Radio Frequency (RF) signals, are employed for drone detection and classification. In this study, RF signals are utilized for classifying drones using DroneRF dataset, which is a publicly available open-source dataset. It consists of four main classes for drone types, namely AR, Bebop, Phantom drones and back ground for no drone existence. The dataset is further divide into ten sub-classes representing different operating modes of each drone. Operating mode classification is crucial for security reasons since they represent drones' specific activity. To achieve high performance in drone classification, we propose the multi-frame majority voting method using cepstral coefficients. Drone signals are divided into multiple frames (2, 4, and 8), and each frame is analyzed using Mel Frequency Cepstral Coefficients (MFCC) and Linear Frequency Cepstral Coefficients (LFCC) attributes. Each frame is classified by Support Vector Machine (SVM), and majority voting is applied to the predictions from the frames. Results show 100% accuracy for drone classification (4-Class) and 99.11% accuracy for defining operating modes (10-Class). The proposed method outperforms existing methods in drone classification using the DroneRF dataset.

Keywords: Drone classification, RF signal, MFCC, LFCC, SVM

Kepstral Katsayılara Dayalı Çoklu Çerçeve Füzyon Yaklaşımıyla Drone Sınıflandırması

Öz

Son yıllarda dronların artan popüleritesi gizlilik ve güvenlik açıklarına yol açmıştır. Günümüzde dronlar kolayca satın alınıp kullanılabilmekte ve bu da özel alanlara izinsiz girilmesi konusunda endişelere yol açmaktadır. Dronların varlığını tespit etmek ve çalışma modlarını belirlemek büyük önem taşımaktadır. Dron tespiti ve sınıflandırması için video, ses, termal görüntüleme ve Radyo Frekansı (RF) sinyalleri dahil olmak üzere çeşitli tespit teknikleri kullanılmaktadır. Bu çalışmada, DroneRF veri kümesi kullanılarak dronların sınıflandırılması için RF sinyalleri kullanılmıştır. Bu veri kümesi, kamuya açık açık kaynaklı bir veri kümesidir. AR, Bebop, Phantom dronları ve dron olmayan arka plan olmak üzere dron tipleri için dört ana sınıftan oluşmaktadır. Veri kümesi ayrıca her bir dronun farklı çalışma modlarını temsil eden on alt sınıfa ayrılmıştır. Çalışma modu sınıflandırması, dronların belirli aktivitesini temsil ettikleri için güvenlik açısından önemlidir. Dron sınıflandırmasında yüksek performans elde etmek için, kepstral katsayıları kullanarak çok çerçeveli çoğunluk oylama yöntemini öneriyoruz. İHA sinyalleri birden fazla çerçeveye (2, 4 ve 8) bölünür ve her çerçeve Mel Frekans Kepstral Katsayıları (MFCC) ve Doğrusal Frekans Kepstral Katsayıları (LFCC) nitelikleri kullanılarak analiz edilir. Her çerçeve Destek Vektör Makinesi (SVM) tarafından sınıflandırılır ve çerçevelerden gelen tahminlere çoğunluk oyu uygulanır. Sonuçlar, İHA sınıflandırması için %100 doğruluk (4-Sınıf) ve çalışma modlarını tanımlamak için %99,11 doğruluk (10-Sınıf) gösterir. Önerilen yöntem, DroneRF veri setini kullanarak İHA sınıflandırmasında mevcut yöntemlerden daha iyi performans gösterir.

Anahtar Kelimeler: Drone sınıflandırması, RF sinyali, MFCC, LFCC, SVM

Corresponding Author: rabiyekilic@atauni.edu.tr

Cite this Article: Erzincan University Journal of Science and Technology 2025, 18(3)892-916. https://doi.org/10.18185/erzifbed.1587701

1. Introduction

Significant technological advances in the drone industry have led to an increasing number of such devices used in different applications [1, 2]. Drones are increasingly being equipped with the latest technologies and sensors, such as Global Positioning System (GPS), Laser Imaging Detection and Ranging (LIDAR), radar, and vision sensors. Today, with the development of drone technology, drones are used for various purposes such as cinematography, agriculture, photography, and entertainment [3]. Drones, despite their advantageous functions, also serve illegal purposes and pose a threat to public safety [4-5]. The increasing use of Unmanned Aerial Vehicle (UAVs) creates both security and privacy vulnerabilities. It is known that UAVs are used for malicious purposes such as cybercrime, terrorism, and drug trafficking [6]. For this reason, detecting foreign drones is extremely important today. Automatically detecting and identifying a drone is a difficult process. Radar sensing [7], vision sensing [8], acoustic sensing [9], and Radio Frequency, (RF) fingerprint-based sensing [10] are some of the most widely used technologies to detect and classify drones. The methods used in the literature for drone detection and classification are briefly summarized as follows. Radar detection uses an active sensor in detection systems to determine the range, angle, or speed of the drone. A radar system consists of a transmitter, a receiver, and a processor. Traditional radar systems will fail to detect a mini drone due to the small radar cross-sectional area. To overcome this problem, researchers use a multi-static radar or a Frequency Modulated Continuous Wave (FMCW) radar [10, 11]. To distinguish drones from birds, rotor blade rotation and clear micro-Doppler signatures are used in radar rotation analysis [12, 13]. Video/image detection includes both visual and thermal sensing. In the studies [15, 16, 17], researchers perform drone detection by analyzing color, shape, and edge information. The detection method is reliable but requires a line of sight between the drone and the camera, and performance is highly dependent on daylight conditions and weather conditions such as dust, rain, fog, and cloud. The similarity of a bird to a drone complicates the task for a video detector. In the study in [18], a 2-Dimensional (2D) scale was classified as a drone or bird with a neural network using the invariant Generic Fourier Descriptor (GFD) features with rotation and translation. In addition, drone model classification was done by proposing a Convolutional Neural Network (CNN) that describes drone models in real-life videos [19]. In addition to these studies, recently Ahmad et al. [44] proposed an architecture consisting of two main modules on RGB drone images. The first one is a Visual Geometry Group based convolutional neural network (VGG-CNN) framework for detection, and the second one is a mask region based double convolutional neural network (MR-DCNN) framework for identification of drone type, payload and flight characteristics. Some studies in the literature [20, 21, 22] have utilized acoustic detection to detect drone presence. In this approach, the sounds produced by flying drones with the help of microphones are used. Acoustic sensing generally works well in a quiet or less noisy environment, but performance suffers when the environment is noisy, such as in urban or industrial areas. In another study, Mel-Frequency Cepstral Coefficients (MFCCs) derived from sound signatures are then fed to CNN, which predicts the presence of a drone [23]. In another study, a new frequency feature extraction method based on the Prony's algorithm was utilized for drone detection and classification with acoustic signals [47]. This study by Najafi et al. the performance of the

method was compared with traditional audio features such as MFCCs, Gamma Tone Cepstral Coefficients (GTCCs) and Fast Fourier Transform (FFT). As an alternate approach, drone-emitted RF signals have been recently used for drone detection. A large dataset, namely DroneRF, has been created from different drones for that purpose [24]. Al-Sa'd et al. [10] used Deep Neural Networks (DNN) to detect the presence, type, and flight mode of a drone. The performance of each DNN is verified by 10-fold cross-validation and evaluated using a variety of metrics. Al-Emadi et al. [25] proposed a new RF-based drone detection solution using the CNN method. Medaiyese et al. [26], a machine learning model was developed using the XGBoost algorithm to detect and identify the presence of a drone and its type. Evaluated using 10-fold cross-validation the XGBoost model. Allahham et al. [27] propose a deep learning-based approach for drone detection, type identification, and state identification using a multi-channel 1-Dimensional (1D) CNN. Swinney et al. [28] plot the spectrogram, histogram, raw IQ, and Power Spectral Density (PSD) as well as graphical representations of RF signals. Features are extracted using a pre-trained ResNet50 CNN in the ImageNet data. Finally, drone RF signals are classified with the machine learning model Logistic Regression (LR) as a classifier. Ensemble learning was performed with the dataset K-Nearest Neighbor (KNN) and XGBoost preprocessed with FFT drone modes, and classification was done with the hierarchical learning method by Nemer et al. [29]. Kılıç et al. [30] extracted features from PSD, MFCC, and Linear Frequency Cepstral Coefficients (LFCC) improved filter banks and parameters, and classification was with the Support Vector Machine (SVM). Hybrid Model with Feature Fusion Network (HMFFNet) model structure is proposed by Kumbasar et al. [31]. In that study, the RF signals were converted into images using spectrogram, persistence spectrum, and percentile spectrum; the next features were obtained by using VGG19. Obtained features were classified by SVM. Huynh et al. [32] created a CNN structure called RF-UAVNet. In other study RF signals are adapted according to the signal properties; sampling structure based on the compressively sensed theory and data preprocessing method is performed by Mo et al. [33]. Next, the neural network structure is designed to detect and classify RF fingerprints. In another study [34], the loads carried by drones are estimated up to 200 m 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. AlKhonaini et al. [43] trained the reinforcement learning algorithm with an entropy regularization term hierarchically (with 2,4,10 classes respectively) with multiple policies. They also comprehensively evaluated the performance during training and testing using several metrics, including loss, average return per episode, time, and energy consumption, in addition to accuracy. Zahid et al. [45] combined automatic and manual feature extraction techniques for drone signal classification with RF signals, integrated wavelet-based noise removal into the framework, and proposed a compound ensemble learning (CEL)-based neural network. Haque et al. [46] proposed an explainable AI-based approach for drone recognition and identification with RF signals using SHapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) approaches. Table 1 presents a comparative analysis of drone detection methods using RF signals.

In this study, a particular focus is placed on the detection of drones using RF signals, employing advanced methods like majority voting, utilized to improve accuracy performance by combining the results of multiple tests, and cepstral coefficients such as MFCC and LFCC. These methods offer promising results by enabling more accurate identification and classification of drones in various environmental conditions.

The main motivation of the proposed work is the classification of drones using drone RF signals. For this purpose, RF signals are split into multiple frames (2, 4, and 8). RF signals are divided into multi-frames, and experimental studies are carried out for each frame. MFCC and LFCC features are obtained. Estimation of each frame using SVM, drone types, and drone modes are classified with majority voting.

- There are studies in the literature as a single frame. In this study, a single frame was defined as the baseline. Performance was measured from frames divided into multi frames.
- Unlike the multi-segment numbers (100, 1000, etc.) used in the literature, the RF signals were divided into a small number of segments (2, 4, and 8) in this study.
- The drone RF signal was estimated by voting the prediction obtained from each frame by majority voting.
- The results obtained with MFCC and LFCC were compared in terms of drone classification performances.

Table 1. Comparative analysis of drone detection methods using RF signals

Study	Dataset Used	Techniques Applied	Evaluation Method	Number of Class	Accuracy (%)
Al-Sa'd et al. [10]	DroneRF	Deep Neural Networks (DNN)	10-fold cross-validation	2-4-10	99.7-84.5-46.8
Al-Emadi et al. [25]	DroneRF	Convolutional Neural Network (CNN)	10-fold ross-validation	2-4-10	99.8-85.8-59.20
Medaiyese et al. [26]	DroneRF	XGBoost	10-fold cross-validation	2-4-10	99.96-90.73-70.09
Allahham et al. [27]	DroneRF	Multi-channel 1D CNN	Not specified	2-4-10	100-94.6-87.4
Swinney et al. [28]	DroneRF	ResNet50, Logistic Regression (LR)	5-fold cross-validation	10	91
Nemer et al. [29]	Custom RF Dataset	Ensemble Learning (KNN, XGBoost, FFT)	Hierarchical Learning Method	2-10	99.5-99.2

Study	Dataset Used	Techniques Applied	Evaluation Method	Number of Class	Accuracy (%)
Kılıç et al. [30]	DroneRF	SVM with PSD, MFCC, LFCC features	10-fold ross-validation	2-4-10	100-98.67-95.15
Kumbasar et al. [31]	DroneRF	VGG19-SVM Hybrid model	10-fold ross-validation	2-4-10	100-99.55-97.75
AlKhonaini et al. [43]	DroneRF	Reinforcement Learning with Hierarchical Training	Cross-entropy loss, Multiple Metrics	2-4-4	99.7
Zahid et al. [45]	RF Drone Signals	Compound Ensemble Learning (CEL)	Machine Learning Evaluation	15	99.0
Najafi et al. [47]	Acoustic Signals	Prony's Algorithm for Frequency Feature Extraction	Comparison with MFCCs, GTCCs, FFTs	4	93.6

The remainder of the paper is organized as follows: Section 2 presents the materials and methods used in this study. Section 3 details the experimental setup, while Section 4 provides the experimental results and numerical analysis. Section 5 compares the proposed method with existing approaches from the literature, and Section 6 concludes the study.

2. Material and Methods

This section describes the materials and methods used for drone detection and classification using RF signals. We focus on the multi-frame classification method proposed in this study, which differs from the single-frame method used in prior research. The dataset used in the proposed method, DroneRF, and the multi-frame classification approach are explained in detail. The process begins with the division of RF signals into multiple frames. For each frame, feature extraction techniques are applied to obtain low-band and high-band features. These extracted features are then classified using Support Vector Machine (SVM). To enhance the accuracy of the classification, majority voting is employed, where predictions from each frame are combined to determine the final classification result.

DroneRF Database

In this study, the open-source DroneRF dataset is used [10, 24]. The dataset consists of 227 recorded segments collected from 3 different drones (Bebop, AR, and Phantom drones) and

background. The dataset is in comma separated values (csv) files. It includes RF signals when there is no drone and RF signals when there is a drone in different modes. It consists of 4 different modes, such as off, on and connected, hovering, flying, and video recording, as listed below.

- Mode 1: On and connected to the controller
- Mode 2: No physical intervention
- Mode 3: Drone flying without video recording
- Mode 4: Flying with video recording

Baseline Method (Single Frame)

In the literature, there are studies [30, 31] carried out without dividing drone RF signals into multiple frames, or when the RF signals are split into multiple frames, each frame is estimated. The success rate is presented on a frame-by-frame basis, not for the entire signal.

In this study, unlike the literature, the RF signals divided into multi-frames are estimated for each frame, and the final class of the entire signal is determined by majority voting. Fig. 1 shows the comparison chart of the single frame used in the literature and the proposed method (multi frame). The multi-frame method used in the study is explained in detail in the following sections.

Fig. 1 illustrates a comparative analysis between the baseline single-frame method and the proposed multi-frame approach for drone detection and classification. The baseline method processes low-band and high-band RF signals as a single frame, extracting features (F_L, F_H) and feeding them into a single SVM model (M) to produce a prediction (P). In contrast, the proposed multi-frame method incorporates multiple frames and extracts features (F_L^1, F_L^2,..., F_L^m and F_H^1, F_H^2,..., F_H^m) from each frame. These extracted features are then independently processed using multiple SVM models (M_1, M_2,..., M_m), generating individual predictions (P_1, P_2,..., P_m). A final classification result (P_F) is obtained through a voting mechanism, which aggregates the predictions of such multiple frames. This methodology enhances classification accuracy by leveraging temporal information and mitigating potential misclassifications associated with single-frame analysis.

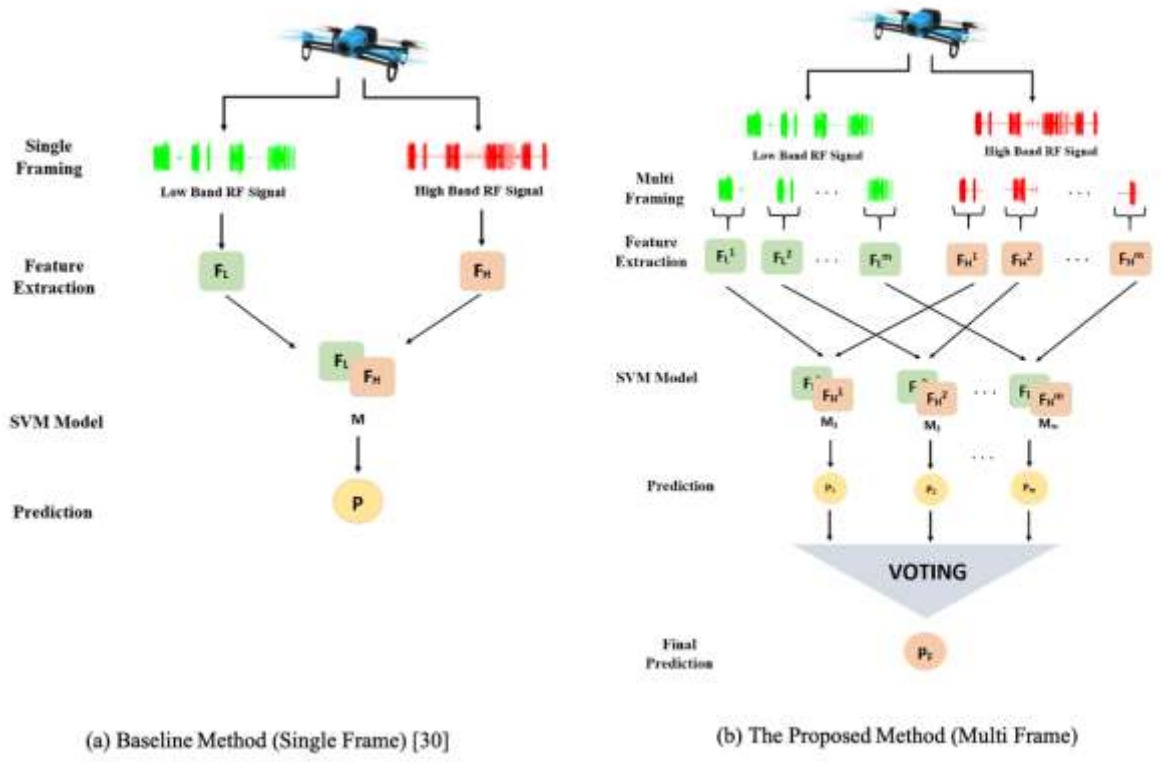


Fig. 1. Comparison of drone detection methods. (a) The baseline method uses single-frame feature extraction and SVM Classification [30]. (b) The proposed multi-frame method enhances accuracy through feature extraction across multiple frames and majority voting.

The Proposed Method (Multi-Frame)

In this study, multi-frame is suggested as the proposed method. RF signals are received and processed by receivers when there is no drone and when there is a drone. Bebop, AR, and Phantom drones used in receiving RF signals operate in different bandwidth ranges. It is assumed that they use WiFi operating at 2.4 GHz while creating this dataset. Two NI USRP-2943R RF receivers were used. The first receiver captures the lower half (low) of the frequency band, and the second receiver records the upper half (high). Thus, there are 454 RF signal files in total. Recordings are taken as 10.25 seconds in total. In 5.25 seconds of this received signal, the RF drone is in communication. Bebop, AR, and Phantom drones were used while receiving RF signals. The drone RF dataset has more than 40 GB of data.

Drone RF signals are divided into 2, 4, and 8 frames to be used in experimental studies. The same process is applied to both the low and high bands of a drone signal. Fig. 2 depicts the decomposition of the drone's RF signal into low-band and high-band components, which are further divided into 2-frame, 4-frame, and 8-frame segments. This multi-frame approach enhances temporal feature extraction, improving the accuracy and robustness of drone detection compared to single-frame methods.

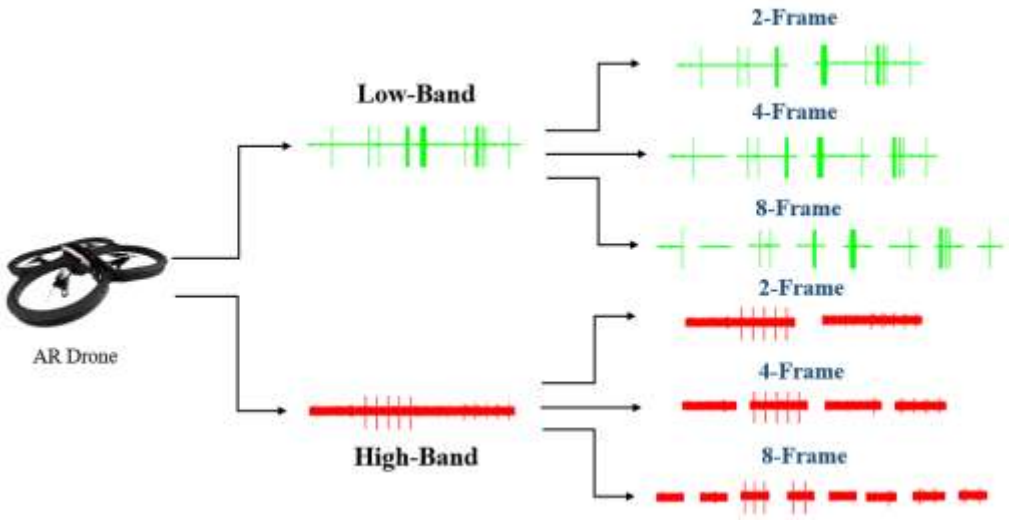


Fig. 2. Drone RF signal separated into its frames.

Feature Extraction with Cepstral Coefficients

Cepstral coefficients are commonly used to extract features from time-varying signals. By switching to the frequency space, signal processing is provided [35, 36].

Step 1 Framing: It can be seen that time-varying signals do not change much in short time intervals. For this reason, signals are processed by framing them at certain lengths.

Step 2 Windowing: The framing process causes discontinuity in the signal. In order to prevent these discontinuities, windowing is done. Hamming windowing, which is widely used, was used in this study. Hamming windowing is given in Eq. 1.

$$w(n) = 0.54 - 0.46\cos(2\pi n/N - 1) \quad (1)$$
$$N - 1 \geq n \geq 0$$

where $w(n)$ represents the value of the Hamming window for n, frame time-index, N is the total length of the window, typically the length of the segment of the signal that is being analyzed.

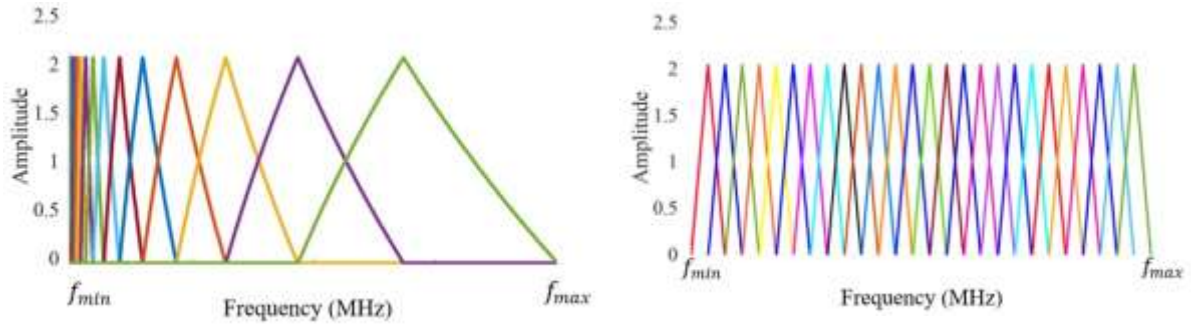
Step 3 FFT: The transition from the time domain to the frequency domain of the windows consisting of (N) samples is realized with the FFT. The (x_n) signals in the time domain are converted to the frequency space expressed by (X_k) in Eq.2 with the following formula.

$$X_k = \sum_{n=0}^{N-1}(x_n)e^{-j2\pi kn/N}, \quad \text{k}= 0, 1, \dots, N-1 \quad (2)$$

Step 4 Filter Banks: Cepstral coefficient scales obtained with Mel and linear filters are used as features in drone detection and classification. LFCC, fixed-width filters, defined in the frequency domain, are utilized, while those with varying width according to Mel-scale are used

in the MFCC process. Three parameters define such filter banks. These are the total number of filters F, minimum frequency (f_{min}), and maximum frequency (f_{max}) in the filtering band. These parameters are defined in [37].

The widths of Mel frequency filters are determined according to the sensitivity in the frequency space. Corresponding filter bank examples are shown in Fig. 3. In a mel-scale filter bank, narrow triangular filters are used at low frequencies and wider filters are used at high frequencies. Thus, the signal in the frequency space is converted into a mel-scale power spectrum. In linear-scale filter bank filters of equal width are used at each frequency value.



(a) (b)

Fig. 3. Filter banks used in cepstral coefficients; (a) Mel-scale filter bank, (b)Linear-scale filter bank.

Step 5 Cepstrum Coefficients: The last step is to obtain the cepstrum coefficients. It reveals the differences and similarities with the cepstral coefficients. There are cepstrum coefficients with Discrete Cos Transform (DCT). First of all, $H_m(k)$ Mel frequency filter bank, is multiplied by $X_n(k)$, the input signal, as shown in Eq. 3.

$$S_n^m = \sum_{k=0}^{N-1} X_n(k).H_m(k) \tag{3}$$

$$log(S_n^m) = log\big(X_n(k).H_m(k)\big) = log\big(X_n(k)\big) + log\big(H_m(k)\big) \tag{4}$$

$$\text{cepstrum(k)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} log(S_n^m).e^{jwk}dw \tag{5}$$

Eq.4 applies a logarithmic transformation to the power spectrum of the signal. This is usually utilized to make the dynamic range of the signal more manageable and to better analyze the signal at different frequencies. Eq. 5 uses the inverse Fourier transform to extract the cepstrum coefficients from the logarithmic spectrum. This transforms the frequency components of the signal into a representation in the time domain.

Support vector machine (SVM) Classifier

SVM (Support Vector Machine) classification is one of the widely used methods [38]. It provides drawing the most appropriate line to separate the elements of each class shown on the

plane. These lines are also called decision boundaries or hyperplanes. Support vector machines are basically divided into linear and non-linear. Linear support vector machines are often used to distinguish between the two classes. There are support vectors representing each class. Usually these class labels are used as $(-1, +1)$. A straight line is drawn to center the boundary planes. This line is called hyper plane [39,40]. It is decided which class the incoming data belongs to, depending on whether (y) is (-1) or $(+1)$.

Nonlinear SVMs do not have a hyperplane. Data that is not separated linearly is based on the principle of classification by mapping in high [41]. It is the tradeoff parameter used to increase the distance of the support vectors used in the equation from the hyperplane and to reduce the classification error. The optimization problem can be solved by the Lagrangian multiplier method. Another space conversion process is performed using kernel tricks known as the kernel plane [42]. The most commonly used kernel methods are polynomial kernel and Radial Basis Function (RBF) kernel. Thus, each kernel function corresponds to a new feature space. In this study, classification process was carried out by experiments with SVM parameters linear, polynomial, and RBF kernel parameters.

Majority Voting Based Multi-Frames

Within the scope of this study, drone detection and classification, majority voting is done based on multi-frames. First, the low- and high-band signals received from the drone are divided into multiple frames (m frames), and attributes are obtained for each frame. Then, the features obtained from the low-high band were combined and used to create the model with SVM. Each frame of the RF signal is estimated. And finally, the RF signal's class is estimated by voting by majority voting for each frame.

3. Results and Discussion

3.1. Results

In this section, performance metrics of experimental studies, experimental setup, preparation of train and test data, and SVM parameters are explained.

Performance Metrics

The classification performance of the proposed study is evaluated in terms of Accuracy, Precision, Recall and F1 score metrics as defined in Eqs. 6-9.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1\ score = 2\frac{Precision.Recall}{Precision+Recall} \quad (9)$$

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

Experimental Setup

All experiments for the proposed drone detection and classification system are performed on a 10-core i9-7900X central processing unit (CPU) with 128GB of RAM and the Ubuntu 18.04.4 LTS Linux operating system platform. Preprocessing of the DroneRF dataset and proposed algorithms are developed using MATLABR2019b software backed by three GeForce RTX 2080Ti graphics processing units (GPUs).

Train and Test Data Preparation

It is one of the commonly used methods for separating the dataset into train and test. The dataset is randomly divided into k parts. When k is taken as 10, the dataset is divided into 10 parts. With the dataset, it is possible to train 10 times with different parts and test 10 times on different parts. This method ensures the accuracy of the prepared model.

SVM Parameters

The SVM model is created after the extracted features are grouped as train and test with cross-validation. Different kernel functions, such as linear, polynomial, and Gaussian functions, are tested as kernel functions, and the Gauss function that performed the highest accuracy is determined as the kernel function. Experimental studies are performed for various values of kernel scale (v) and penalty parameter (C) in the range from 2^{-15} to 2^{15} with the power increments. In this study, the one-vs-one method is preferred for SVM because it offered better results than one-vs-all.

Performance Analysis

In this section, the experimental results of the proposed multi-frame late fusion methods are described. LFCC and MFCC are represented as features of RF signals. First, the methods used to develop MFCC and LFCC features and their results are presented. Experimental studies are repeated for 2, 4, and 8 frames, and the results are shown. Finally, the proposed method is compared with the literature.

Adjusted Features for MFCC and LFCC

This section shows the improvement of the proposed method by adjusting the LFCC and MFCC feature parameters and adding new features. For that, to better clarify the obtained results, abbreviations used in the reported results are given in Table 2.

Table 2. Abbreviations used in the tables.

D	Default
O	Zero-Order Coefficient
E	Log Energy (in time domain)
NC	Number of cepstral Coefficient

Single-Frame Classification Results

Drone signals are classified as a single frame by Kılıç et al. [30]. The features of each signal were extracted using MFCC and LFCC, and the success rates obtained as a result of 2, 4, and 10 class classification with SVM are shown in Table 3.

The bottom row of the Tables (3-4-5-6) presents the parameter combinations for which the relevant result was obtained. The plus sign (+) indicates to be used together.

Table 3. Single-frame based accuracy rates with MFCC and LFCC features [30]

	Accuracy (%)		
	2-Class	**4-Class**	**10-Class**
MFCC	100	98.67	90.33
LFCC	100	98.67	95.29
Parameters:	D, NC=12	D+O+E, NC=12	D+E, NC=13

When Table 3 is examined, without dividing drone signal frames; 100% success rate has been achieved in the classification of drone with or without (2-Class). A success rate of 98.67% is achieved in drone types (4-Class) and 95.29% in classification between drone modes (10-Class) by using LFCC.

The Proposed Multi-Frame Classification Results

In this section, since 100% accuracy was achieved 2-class (drone detection) single frame [30], experimental studies are carried out for 4 and 10 classes. Each drone signal is divided into 2, 4, and 8 multi-frame, and the features obtained from each frame are classified by SVM, and the success rates obtained by using majority voting are shown in Table 4.

Table 4. Multi-frame based accuracy rates with MFCC and LFCC features

	Accuracy (%)			
Number of Frames	**MFCC**		**LFCC**	
	4-Class	10-Class	4-Class	10-Class

2	98.23	89.43	98.23	94.86
4	**99.55**	93.90	**100**	**97.09**
8	**99.55**	**94.79**	99.89	95.27
Parameters:	D+O+E, NC=12	D+E, NC=13	D+O+E, NC=12	D+E, NC=13

From the comparison of MFCC and LFCC results presented in Table 4, the accuracy rate of the multi-frame based classification depends on the number of frames. The best results of 100% and 97.09% are obtained for 4 frames in 4-Class problem and for 4 frames in 10-Class problem in LFCC, respectively.

Figs. 4 and 5 present the classification percent accuracy comparisons between the single-frame method and the proposed multi-frame approach for 4-class and 10-class classification tasks, respectively. In Figure 4, the multi-frame approach significantly improves accuracy compared to the single-frame method, with the 4-frame configuration achieving 100% accuracy with LFCC features and 99.55% with MFCC features. The 8-frame configuration also maintains high accuracy (99.89% for LFCC and 99.55% for MFCC), demonstrating the effectiveness of incorporating temporal information. Similarly, Figure 5 shows that the multi-frame approach enhances accuracy in the 10-class classification task, with the 4-frame configuration achieving the highest accuracy (97.09%) using MFCC features, followed by the 8-frame method (95.27% for MFCC and 94.79% for LFCC). However, the 2-frame configuration exhibits a slight decrease in accuracy compared to the single-frame method in LFCC-based classification, indicating that a higher number of frames is more beneficial. These results confirm that multi-frame processing significantly improves classification accuracy, particularly when an optimal number of frames is utilized, making it more effective than the traditional single-frame approach.

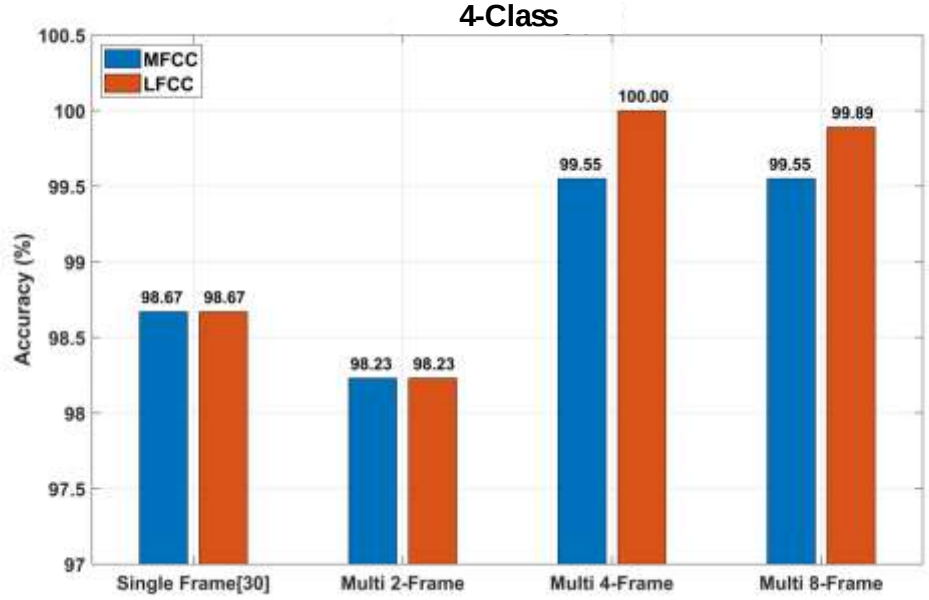


Fig. 4. Comparison of single frame [30] and multi-frame results for 4-Class

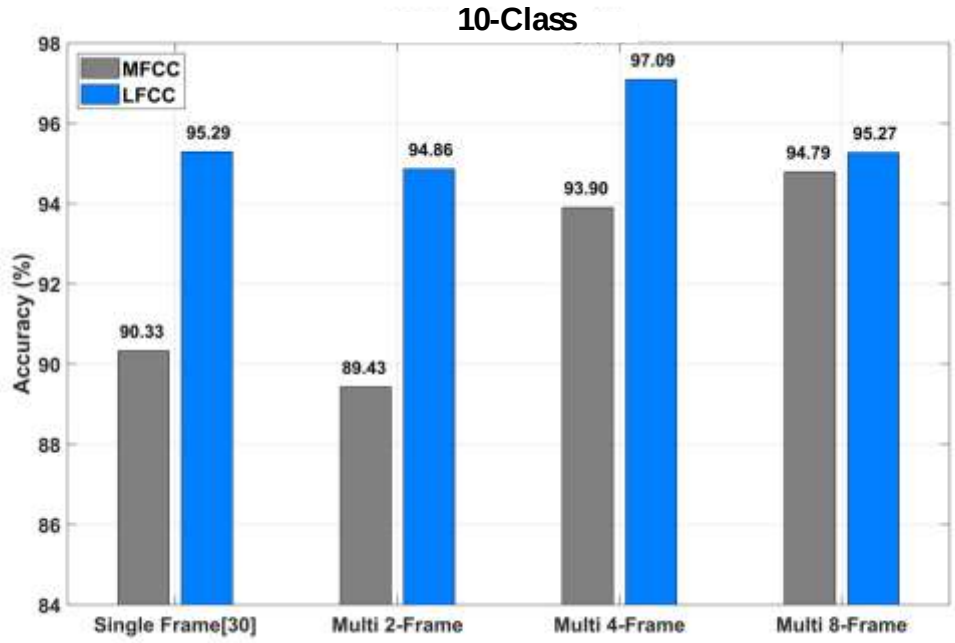


Fig. 5. Comparison of single frame [30] and multi-frame results for 10-Class

The Effect of Overlap-Rate on Success Rate

In this section, the effect of overlap on the success rate is analyzed. Classification is done with no overlap (0%), 50%, and 75% overlap. In Table 5, the success rate obtained with MFCC is presented by overlapping.

Table 5. Overlapped multi-frame accuracy rates with MFCC

	Accuracy (%)					
	4-Class			**10-Class**		
	Overlap rate			**Overlap rate**		
Number of frames	**0%**	**50%**	**75%**	**0%**	**50%**	**75%**
2	98.23	100	99.55	89.43	95.15	96.91
4	99.55	100	99.55	93.90	97.35	98.67
8	99.55	100	100	94.79	98.80	99.11
Parameters:	D+O+E, NC=12			D+E, NC=13		

From the evaluation of Table 5, the 4-Class problem yields a 100% accuracy rate with the use of a 50% overlap ratio, while that of the 10-Class problem increases to 99.11% with a 75% overlap ratio, all with MFCC features.

It is clear that the accuracy rate of MFCC and LFCC feature-based classification is improved to 100% in the 4-Class problem with multi-frame.

On the other hand, the accuracy rate dependency on overlap ratio is only investigated for the 10-Class problem using LFCC features since full accuracy is obtainable without overlapping frames in the 4-Class problem. It is determined that the accuracy rate increases from 97.09% to 97.35% with an overlapping ratio of 75% as shown in Table 6.

Table 6. Overlapped multi-frame accuracy rates with LFCC for 10-Class

	Accuracy (%)		
Number of	**Overlap rate**		
Frames	**0%**	**50%**	**75%**
2	94.86	93.39	93.39
4	97.09	96.18	97.35
8	95.27	96.47	97.35
Parameters:	D+E, NF=13		

In the 10-Class problem, the accuracy rate changes from 94.79% to 99.11% with frame overlapping for MFCC-based features, yielding a 4.55% increase. That of LFCC, on the other hand, shows a 0.31% increase for the 97.35% best accuracy rate. Fig. 6 presents a comparative analysis of classification accuracy for the 10-class task using single-frame, multi-frame, and multi-frame with overlap approaches. The results indicate that the multi-frame approach outperforms the single-frame method, with accuracy increasing from 90.33% (MFCC) and 95.29% (LFCC) to 94.79% (LFCC) and 97.09% (MFCC), respectively. Furthermore,

introducing frame overlap approach improves the accuracy with 99.11% (MFCC) and 97.35% (LFCC). These results demonstrate that multi-frame processing significantly improves classification performance, while incorporating overlap further refines feature extraction, leading to higher accuracy in drone detection and classification tasks.

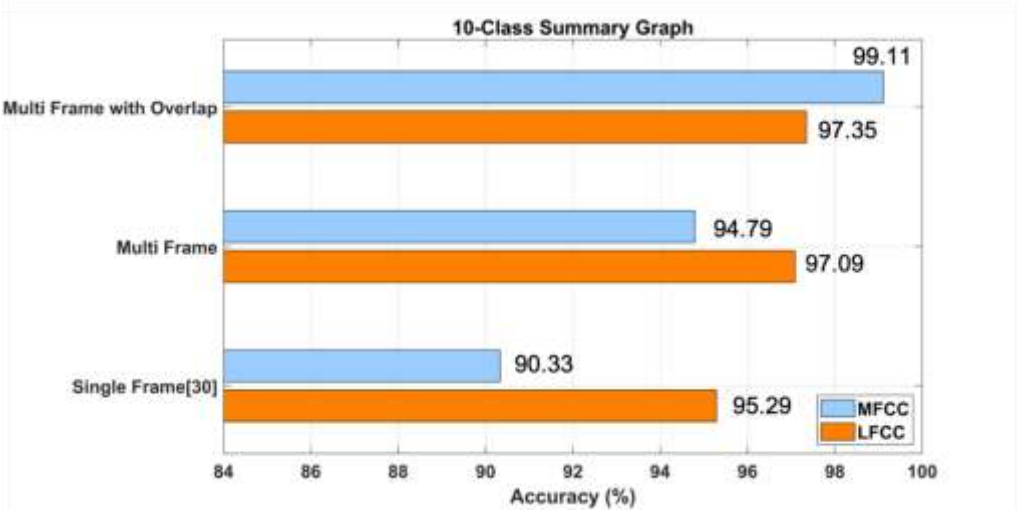


Fig. 6. The effect of multi-frame on accuracy rate for 10-Class.

By using the DroneRF dataset, a 100% success rate is achieved in the classification of 2 and 4 class drone RF signals and 99.11% in the classification with 10-Class. Only the drone operation mode (10-Class) confusion matrix is shown in Fig. 7, since 100% success was finally achieved in drone detection (2-Class) and drone classification (4-Class) problems.

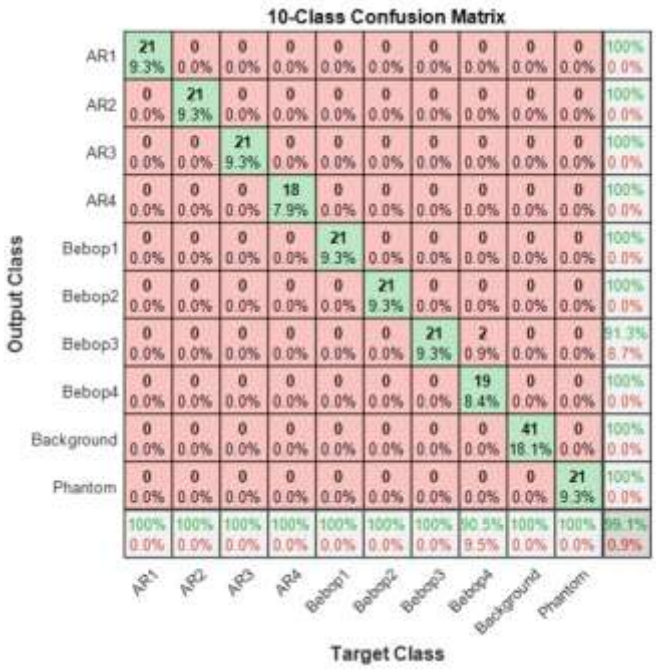


Fig. 7. Operation mode recognition (10-Class) confusion matrix.

The Accuracy, F1 Score, Recall and Precision values of the proposed method are obtained as in Table 7.

Table 7. Accuracy, F1 score, Recall and Precision values of the proposed method.

		MFCC	**LFCC**
Accuracy (%)	2 Class	100	100
	4 Class	100	100
	10 Class	99.11	97.35
F1 score (%)	2 Class	100	100
	4 Class	100	100
	10 Class	99.06	97.11
Recall (%)	2 Class	100	100
	4 Class	100	100
	10 Class	99.07	97.06
Precision (%)	2 Class	100	100
	4 Class	100	100
	10 Class	99.09	97.34

3.2. Discussion

In this section, the proposed method is compared with the studies in the literature using the DroneRF dataset and is shown in Table 8.

Al-Sa'd et al. [10] used three DNNs through signals divided into 100 segments, using 10-fold cross-validation. They performed the power spectrum as the feature set, which is obtained from the combined magnitude DFT of low- and high-band signals. Al-Emadi et al. [25] proposed a method using a CNN architecture that consists of five 1D convolutional layers, utilizing 10-fold cross-validation. While they slightly improved the accuracy results of Al-Sa'd et al. [10] for 2, 4, and 10-Class problems. Medaiyese et al. [26] proposed three machine learning models utilizing the XGBoost algorithm and evaluated these models using 10-fold cross validation. Their major contribution to accuracy is compared to Al-Emadi et al. [25] in the 10-Class problem. Allahham et al. [27] channelized the RF Drone signals' full spectrum into multiple channels and considered each one as a separate input of multi-channel 1D CNN. It is used with RF signals divided into approximately 370 segments. Swinney et al. [28] RF signals as spectrogram, histogram, raw IQ constellation, and PSD graphical representations. The features are extracted using ResNet50 and classify by LR utilizing 5-fold cross-validation as10-Class drone signals. Kılıç et al. [30] extract features with PSD, MFCC, and LFCC and classify them with SVM. In that study, RF signals are handled as 1 segment, and 10-fold cross-validation is

used in classification. By Kumbasar et al. [31], RF signals are represented by images using spectrogram, persistence spectrum, and percentile spectrum with HMFFNet model structure classified by SVM and the features obtained using VGG19. In that study, DroneRF data was discussed into 1, 10, and 100 segments, and 10-fold cross-validation was carried out in their studies. Huynh et al. [32] with the CNN architecture they call RF-UAVNet 10-fold cross validation performed in the study, which is done by dividing the approximate signals into 1000 segments.

With the proposed MFCC and LFCC feature multi-frame majority voting method, 100% accuracy has been achieved in the 4-Class classification. An accuracy rate of 99.11% is achieved in 10-Class classification. Experimental studies are carried out according to 2, 4, and 8 segments. 10-fold cross validation is used to calculate the accuracy rate.

Table 8. Comparison of studies using the DroneRF dataset

	Accuracy (%)			**F1 score (%)**		
Literature	**2-Class**	**4-Class**	**10-Class**	**2-Class**	**4-Class**	**10-Class**
Al-Sa'd et al.[10]	99.70	84.50	46.80	99.50	78.80	43.00
Al-Emadi et al.[25]	99.80	85.80	59.20	99.70	84.68	55.10
Medaiyese et al.[26]	99.96	90.73	70.09		N/A	
Allahham et al.[27]	100	94.60	87.40	100	91	77.00
Swinney et al.[28]	N/A		91.00	N/A		91.20
Kılıç et al.[30]	100	98.67	95.15	100	98.70	94.72
Kumbasar et al.[31]	100	99.55	97.75	100	99.69	97.57
Huynh et al.[32]	99.85	98.55	95.33	99.75	98.48	95.06
The Proposed Study (Multi Frame Majority Voting)						
MFCC	N/A	**100**	**99.11**	N/A	**100**	**99.06**
LFCC	N/A	**100**	97.35	N/A	**100**	97.11

Analysis of time dependent signal requires the use of an optimum frame-length to better extract features as there exists a trade of between the temporal and frequency resolution, known as Heisenberg Uncertainty Principle [48], and this frame-length, in general, depends on the specific data. Hence, in the proposed study, all experimental studies were performed for different frame-lengths to gather best results to suit DroneRF data. For that purpose, different frame lengths were examined by splitting the data into 2, 4 and 8 frames, and the corresponding results yielded that the best performance was acquired for the length obtained with 4-frame splitting. Another parameter in the feature extraction process is the filtering characteristic to select different frequency content of the signal. For that, MFCC and LFCC approaches were

utilized, and LFCC based feature set performed better compared to MFCC based feature set. This is probably due to the fact that Mel-frequency filtering is optimized for human hearing system as varying band-widths are utilized in MFCC [49]. On the other hand, it is observed that the linear band-width approach of LFCC, which equally partitions the frequency range, yielded better result compared to MFCC, and hence, the use of LFCC approach is better suited for drone rf signals. A final discussion should be addressed regarding the effect of framing approach on the computational complexity of the proposed method. This is important as the overlapping is also utilized in the study, and this operation results in the increased amount of data to process yielding slightly increased processing time at a gain of increased accuracy performance. The results demonstrate that the proposed multi-frame approach improves drone detection accuracy compared to single-frame methods. However, the study has certain limitations. The computational complexity increases due to the multi-frame processing and majority voting mechanism, which extends the inference time compared to single-frame classification. Additionally, the approach relies on the quality of RF signals, making it sensitive to environmental variations and potential external interference. Moreover, the study primarily focuses on specific drone models and RF signal characteristics, and further validation with a broader dataset is necessary to improve generalizability. Addressing these limitations in future work will enhance the efficiency and robustness of the proposed approach.

4. Conclusion

In this study, the problem of detection and classification of different types of drones is investigated. In the proposed approach, the multi-frame majority voting method is used. In the study using drone RF signals, 2, 4, and 8 frames are separated, and each frame min is extracted with MFCC and LFCC and classified by SVM. Then the estimation of each signal frame is voted, and the RF signal is estimated. In the study, a success rate of 100% and 99.11% was achieved in the classification of drone types and drone modes, respectively. The results obtained were compared with studies using the same dataset in the literature. It is seen that the proposed method gives better results than classifications with drone types in the current literature.

The proposed approach holds great potential for real-world applications, including security monitoring, airport surveillance, and restricted area monitoring, where real-time detection and classification of drones are crucial. The ability to detect drones using RF signals makes it particularly useful in environments with limited visibility or challenging conditions such as fog or darkness, where optical detection methods may fail.

In future studies, it is planned to develop the dataset by taking RF signals from different drone models and to work on the problem of drone detection and classification.

Ethics in Publishing

There are no ethical issues regarding the publication of this study.

Author Contributions

Kumbasar N.: Conceptualization, Methodology, Software, Validation, Editing original draft.
Kılıç R.: Conceptualization, Literature search, Data collection, Writing original draft.
Oral E.A.: Analysis, Supervision.
Ozbek I.Y.: Interpretation, Project administration All authors have read and agreed to the published version of the manuscript.

References

[1] Zeng, Y., Zhang, R. & Lim, T. J. (2016). Wireless communications with unmanned aerial vehicles: Opportunities and challenges. *IEEE Communications Magazine*, 54(5), 36-42.

[2] Akter, R., Doan, V.-S., Lee, J.-M. & Kim, D.-S. (2021). CNN-SSDI: Convolution neural network inspired surveillance system for UAVs detection and identification. *Computer Networks*, 201, 108519.

[3] Hassanalian, M. & Abdelkefi, A. (2017). Classifications, applications, and design challenges of drones: A review. *Progress in Aerospace Sciences*, 91, 99-131.

[4] Taha, B. & Shoufan, A. (2019). Machine learning-based drone detection and classification: State-of-the-art in research. *IEEE Access*, 7, 138669-138682.

[5] Ritchie, M., Fioranelli, F. & Borrion, H. (2017). Micro UAV crime prevention: Can we help Princess Leia? In Crime prevention in the 21st century (pp. 359-376). Springer.

[6] Harkins, G. (2020). Illicit drone flights surge along us-mexico border as smugglers hunt for soft spots. accessed.

[7] Mehta, V., Dadboud, F., Bolic, M. & Mantegh, I. (2023). A Deep Learning Approach for Drone Detection and Classification Using Radar and Camera Sensor Fusion. *IEEE Sensors Applications Symposium (SAS)*, 01-06.

[8] Rozantsev, A., Lepetit, V. & Fua, P. (2016). Detecting flying objects using a single moving camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(5), 879-892.

[9] Busset, J., Perrodin, F., Wellig, P., Ott, B., Heutschi, K., Rühl, T. & Nussbaumer, T. (2015). Detection and tracking of drones using advanced acoustic cameras. *Unmanned/Unattended Sensors and Sensor Networks XI; and Advanced Free-Space Optical Communication Techniques and Applications*, 9647, 53-60.

[10] Al-Sa'd, M. F., Al-Ali, A., Mohamed, A., Khattab, T. & Erbad, A. (2019). RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database. *Future Generation Computer Systems*, 100, 86-97.

[11] Fioranelli, F., Ritchie, M., Griffiths, H. & Borrion, H. (2015). Classification of loaded/unloaded micro-drones using multistatic radar. *Electronics Letters*, 51(22), 1813-1815.

[12] Drozdowicz, J., Wielgo, M., Samczynski, P., Kulpa, K., Krzonkalla, J., Mordzonek, M., Bryl, M. & Jakielaszek, Z. (2016). 35 GHz FMCW drone detection system. *2016 17th International Radar Symposium (IRS)*, 1-4.

[13] Fuhrmann, L., Biallawons, O., Klare, J., Panhuber, R., Klenke, R. & Ender, J. (2017). Micro-Doppler analysis and classification of UAVs at Ka band. *2017 18th International Radar Symposium (IRS)*, 1-9.

[14] Ritchie, M., Fioranelli, F., Griffiths, H. & Torvik., B. (2016). Monostatic and bistatic radar measurements of birds and micro-drone. *2016 IEEE Radar Conference (RadarConf)*, 1-5.

[15] Zhang, Z., Cao, Y., Ding, M., Zhuang, L. & Yao, W. (2016). An intruder detection algorithm for vision based sense and avoid system. *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, 550-556.

[16] Ganti, S. R. & Kim, Y. (2016). Implementation of detection and tracking mechanism for small UAS. *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1254-1260.

[17] Stolkin, R., Rees, D., Talha, M. & Florescu, I. (2012). Bayesian fusion of thermal and visible spectra camera data for region based tracking with rapid background adaptation. *2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 192-199.

[18] Unlu, E., Zenou, E. & Riviere, N. (2018). Using shape descriptors for UAV detection. *Electronic Imaging*, 2018(9), 121-128.

[19] Wisniewski, M., Rana, Z. A. & Petrunin, I. (2022). Drone Model Classification Using Convolutional Neural Network Trained on Synthetic Data. *Journal of Imaging*, 8(8), 218.

[20] Kim, J., Park, C., Ahn, J., Ko, Y., Park, J. & Gallagher, J. C. (2017). Real-time UAV sound detection and analysis system. *2017 IEEE Sensors Applications Symposium (SAS)*, 1-5.

[21] Siriphun, N., Kashihara, S., Fall, D. & Khurat, A. (2018). Distinguishing drone types based on acoustic wave by IoT device. *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, 1-4.

[22] Anwar, M. Z., Kaleem, Z. & Jamalipour, A. (2019). Machine learning inspired sound-based amateur drone detection for public safety applications. *IEEE Transactions on Vehicular Technology*, 68(3), 2526-2534.

[23] Alaparthy, V., Mandal, S. & Cummings, M. L. (2021). A comparison of machine learning and human performance in the real-time acoustic detection of drones. *IEEE Aerospace*.

[24] Allahham, M. H. D. S., Al-Sa'd, M. F., Al-Ali, A., Mohamed, A., Khattab, T. & Erbad, A. (2019). DroneRF dataset: A dataset of drones for RF-based detection, classification and identification. *Data in Brief*, 26, 104313.

[25] Al-Emadi, S. & Al-Senaid, F. (2020). Drone detection approach based on radio-frequency using convolutional neural network. *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 29-34.

[26] Medaiyese, O. O., Syed, A. & Lauf, A. P. (2021). Machine learning framework for RF-based drone detection and identification system. *2021 2nd International Conference On Smart Cities, Automation and Intelligent Computing Systems (ICON-SONICS)*, 58-64.

[27] Allahham, M. S., Khattab, T. & Mohamed, A. (2020). Deep learning for rf-based drone detection and identification: A multi-channel 1-d convolutional neural networks approach. *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 112-117.

[28] Swinney, C. J. & Woods, J. C. (2021). Unmanned aerial vehicle operating mode classification using deep residual learning feature extraction. *Aerospace*, 8(3), 79.

[29] Nemer, I., Sheltami, T., Ahmad, I., Yasar, A. U.-H. & Abdeen, M. A. R. (2021). RF-based UAV detection and identification using hierarchical learning approach. *Sensors*, 21(6), 1947.

[30] Kılıç, R., Kumbasar, N., Oral, E. A. & Ozbek, I. Y. (2022). Drone classification using RF signal based spectral features. *Engineering Science and Technology, an International Journal*, 28, 101028.

[31] Kumbasar, N., Kılıç, R., Oral, E. A. & Ozbek, I. Y. (2022). Comparison of spectrogram, persistence spectrum and percentile spectrum based image representation performances in drone detection and classification using novel HMFFNet: Hybrid Model with Feature Fusion Network. *Expert Systems with Applications*, 206, 117654.

[32] Huynh-The, T., Pham, Q.-V., Nguyen, T.-V., Da Costa, D. B. & Kim, D.-S. (2022). RF-UAVNet: High-Performance Convolutional Network for RF-Based Drone Surveillance Systems. *IEEE Access*, 10, 49696-49707.

[33] Mo, Y., Huang, J. & Qian, G. (2022). Deep Learning Approach to UAV Detection and Classification by Using Compressively Sensed RF Signal. *Sensors*, 22(8), 3072.

[34] Nguyen, P., Kakaraparthi, V., Bui, N., Umamahesh, N., Pham, N., Truong, H., Guddeti, Y., Bharadia, D., Han, R., Frew, E., Massey, D. & Vu, T. (2020). DroneScale: drone load estimation via remote passive RF sensing. *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 326-339.

[35] Paul S, B. S., Glittas, A. X. & Gopalakrishnan, L. (2021). A low latency modular-level deeply integrated MFCC feature extraction architecture for speech recognition. *Integration*, 76, 69-75.

[36] Kumaran, U., Radha Rammohan, S., Nagarajan, S. M. & Prathik, A. (2021). Fusion of mel and gammatone frequency cepstral coefficients for speech emotion recognition using deep C-RNN. *International Journal of Speech Technology*, 24(2), 303-314.

[37] Nagarajan, S., Nettimi, S. S. S., Kumar, L. S., Nath, M. K. & Kanhe, A. (2020). Speech emotion recognition using cepstral features extracted with novel triangular filter banks based on bark and ERB frequency scales. *Digital Signal Processing*, 104, 102763.

[38] Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), 121-167.

[39] Oral, E. A., Çodur, M. M. & Ozbek, I. Y. (2017). Sleep stage classification based on filter bank optimization. *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 1-4.

[40] Alsarhan, A., Alauthman, M., Alshdaifat, E., Al-Ghuwairi, A.-R. & Al-Dubai, A. (2021). Machine Learning-driven optimization for SVM-based intrusion detection system in vehicular ad hoc networks. *Journal of Ambient Intelligence and Humanized Computing*, 1-10.

[41] Gani, E. & Manzie, C. (2004). Intelligent computing methods for indicated torque reconstruction. *Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference*, 259–264.

[42] Achirul Nanda, M., Boro Seminar, K., Nandika, D. & Maddu, A. (2018). A comparison study of kernel functions in the support vector machine and its application for termite detection. *Information*, 9(1), 5.

[43] AlKhonaini, A., Sheltami, T., Mahmoud, A. & Imam, M. (2024). UAV Detection Using Reinforcement Learning. *Sensors*, 24(6), 1870.

[44] Ahmad, I., Narmeen, R., Alawad, M. A., Alkhrijah, Y. & Ho, P. H. (2024). Integrating Visual Geometry and Mask Region CNN for Enhanced UAV Detection and Identification. In 2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall) (pp. 1-6). *IEEE*.

[45] Zahid, M. U., Nisar, M. D., Fazil, A., Ryu, J. & Shah, M. H. (2024). Composite Ensemble Learning Framework for Passive Drone Radio Frequency Fingerprinting in Sixth-Generation Networks. *Sensors*, 24(17), 5618.

[46] Haque, E., Hasan, K., Ahmed, I., Alam, M. S. & Islam, T. (2024). Towards an Interpretable AI Framework for Advanced Classification of Unmanned Aerial Vehicles (UAVs). In 2024

IEEE 21st Consumer Communications & Networking Conference (CCNC) (pp. 644-645). *IEEE*.

[47] Najafi, J., Mirzakuchaki, S. & Shamaghdari, S. (2025). Autonomous Drone Detection and Classification Using Computer Vision and Prony Algorithm-Based Frequency Feature Extraction. *Journal of Intelligent and Robotic Systems*, 111(1), 1-21.

[48] Dias, N. C., de Gosson, M. & Prata, J. N. (2024). A metaplectic perspective of uncertainty principles in the Linear Canonical Transform domain. *Journal of Functional Analysis*, 287(4), 110494.

[49] Kinoshita, Y., Osanai, T. & Clermont, F. (2022). Sub-band cepstral distance as an alternative to formants: Quantitative evidence from a forensic comparison experiment. *Journal of Phonetics*, 94, 101177.