Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications

Muhammad Zohaib Anwar, Zeeshan Kaleem, and Abbas Jamalipour, *Fellow, IEEE*

***Abstract*—In recent years, popularity of unmanned air vehicles enormously increased due to their autonomous moving capability and applications in various domains. This also results in some serious security threats, that needs proper investigation and timely detection of the amateur drones (ADr) to protect the security sensitive institutions. In this paper, we propose the novel machine learning (ML) framework for detection and classification of ADr sounds out of the various sounds like bird, airplanes, and thunderstorm in the noisy environment. To extract the necessary features from ADr sound, Mel frequency cepstral coefficients (MFCC), and linear predictive cepstral coefficients (LPCC) feature extraction techniques are implemented. After feature extraction, support vector machines (SVM) with various kernels are adopted to accurately classify these sounds. The experimental results verify that SVM cubic kernel with MFCC outperform LPCC method by achieving around 96.7% accuracy for ADr detection. Moreover, the results verified that the proposed ML scheme has more than 17% detection accuracy, compared with correlation-based drone sound detection scheme that ignores ML prediction.**

***Index Terms*—Amateur drone detection, acoustic-based surveillance, monitoring drone architecture, feature extraction, public safety, UAV.**

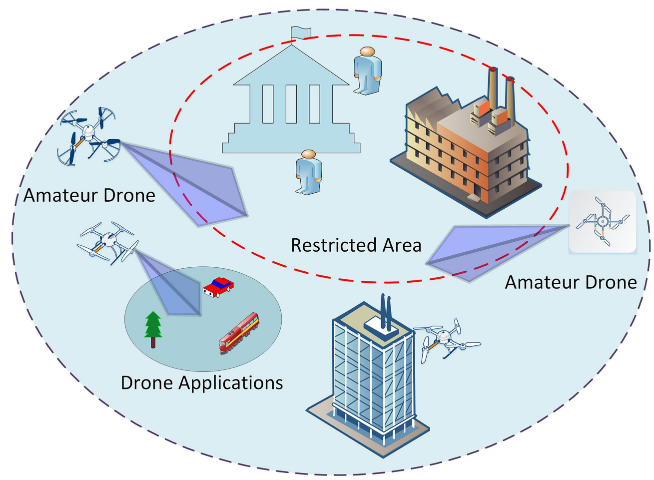


Fig. 1. Amateur drones detection around restricted area and drone applications in surveillance.

I. Introduction

UNMANNED aerial vehicle (UAV) got much attention recent years because of its countless applications [1]–[3]. Drones have applications in tracking, transportation monitoring, aerial photography, device-to-device (D2D) communications [4]–[6], disaster relief [7], [8], location-based services [9], data rate enhancement [1], and security provisioning as shown in Fig. 1. Drone can be used as an additional base stations to overcome the surge in data traffic demands during events like footaball and olympic games [10]. Moreover, drones can be deployed during disasters to provide on-demand wireless coverage by enabling multi-hop communications [11], [12].

However, drones deployment can deliberately or accidently violate the security measures of the National institutions [8]. It can act as a carrier for transfering explosive payloads and also violate the periphery of security sensitive areas. These violations might be launched from any illegal oraganization including non-certified, extremest, and terriosit or from any private operators carrying chemical and other explosive materials. To cope with these challenges, there is a rapid demand for the technologies that can timely detect and disarm amateur drones (ADr). The use of survelliance drones for detecting ADr has been addressed in [13], [14], which ultimately decreases the public safety threats. However, due to drones small size, capability to fly at low altitudes, and their low radar cross section, they can easily enters in no fly zone area as shown in Fig. 1. Considering these facts, number of solutions for ADr detection; sound, video, thermal, and radio frequency (RF) are discussed in the literature [15]. However, each schemes has its benefits and limitations. For example, video and thermal-based detection schemes can be adopted for ADr detection, but they have limitations in adverse weather conditions. ADr propeller and motor sounds can be used for ADr detection by deploying acoustic sensors. The acoutic based detection can be cost effective solution for the National security agencies as they have large ADr detection range. However, noise can be a limiting factor in sound-based ADr detection schemes [16].

Motivated by this, we propose machine learning (ML) inspired sound-based drone detection scheme that can quickly achieve higher accuracy with less complexity even in noisy environment. The proposed scheme integrates the advanced acoustic processing and ML algorithms to enhance the ADr detection performance. The feature extraction performance of

Manuscript received July 18, 2018; revised November 20, 2018; accepted January 1, 2019. Date of publication January 17, 2019; date of current version March 14, 2019. The review of this paper was coordinated by Dr. Z. Fadlullah.
(Corresponding author: Abbas Jamalipour.)

M. Z. Anwar and Z. Kaleem are with the Electrical and Computer Engineering Department, COMSATS University Islamabad, Wah 47040, Pakistan (e-mail: zohaibanwar22@yahoo.com; zeeshankaleem@gmail.com).

A. Jamalipour is with the School of Electrical and Information Engineering, University of Sydney, Sydney, NSW 2006, Australia (e-mail: a.jamalipour@ieee.org).

Digital Object Identifier 10.1109/TVT.2019.2893615

different acoustic processing algorithms like linear predictive cepstral coefficients (LPCC) and Mel frequency cepstral coefficients (MFCC) [17] are compared in terms of accuracy under the noisy environment. To accurately differentiate the ADr sound with other flying objects, support vector machine (SVM) classifier is being proved to provide better classification results even with small data set [18].

II. Related Work

ADr detection and classification can be performed through radar, sound, video, thermal, and radio frequency (RF) technologies as discussed above. The existing work for ADr monitoring using array of cameras and microphones is discussed in [19]. Here, sound and image feature extraction was performed to monitor the presence of drones, and they achieve upto 81% ADr detection accuracy. ADr detection using radars based on pseudo random binary sequence is proposed in [20]. Results show that ADr can be detected within 100 m range for 2 GHz band using radar technology. Similarly, in [21] authors demonstrates that 35 GHz frequency-modulated continuous-wave (FMCW) radar having two fixed antenna can be deployed to identify ADr presence. Results prove that drones are successfuly detected with their estimated velocity. The efficiency of this system can be further increased by using rotating antennas. Furthermore, deep learning techniques [22]–[27] with deep belief network (DBN) along with convolutional neural networks (CNN) are discussed for detecting ADr [28]. These schemes have limitations as it can only give high ADr detection accuracy for good channel conditions, and also demands huge data set for accurate detection. Similarly, in [29], authors proposed the method of rouge drones detection based on radio measurements in cellular networks using two machine learning methods: logistic regression and decision tree. The major limitation in these models are they achieved lower accuracy and generated more interference for drones flying at lower heights.

In [30], authors introduced the acoustic-based drone detection for real-time scenario by implementing deep learning algorithms such as plotted image machine leaning (PIL) and K nearest neighbors (KNN). Simulation results show that PIL detection efficiency is 22% more as compared to KNN, whereas KNN has much less complexity than PIL. These two schemes have limitations because they can only perform better with huge data set. Authors in [31] discussed the correlation-based sound detection scheme for limited database, but accuracy was quite low in real-time environment. However, none of these scheme exist that has a low-complexity, high accuracy in noisy environment, and works well with limited database. For detecting drones through video was presented in [32], where two cameras has been equipped with system to work at day and night mode to maximize the detection rate. Short wave infrared (SWIR) and high resolution visual-optical (VIS) camera has been adopted for this. But this model fails to work in strong wind situation and gives worse detection results. In [33], hidden Markov model (HMM) is used to observe the existence of drones using acoustic sensors. Due to complexity of classifier, it gives poor results in case of small training data. So, up to the author knowledge no such scheme exists that leverages both the speech processing and ML for ADr detection. The main contribution of this paper are two fold:

- The integration of the advanced acoustic processing algorithms and ML framework which can efficiently and timely detect ADr.
- We prove that with minimum data set, the accurate classification of ADr is possible under high order SVM kernels.

The rest of the paper is organized as follows. Section III discussed in detail the proposed acoustic-based detection methodology. Experimental results are presented in Section IV. Finally, conclusions are drawn in Section V.

III. Acoustic Signal-based Drone Detection Methodology

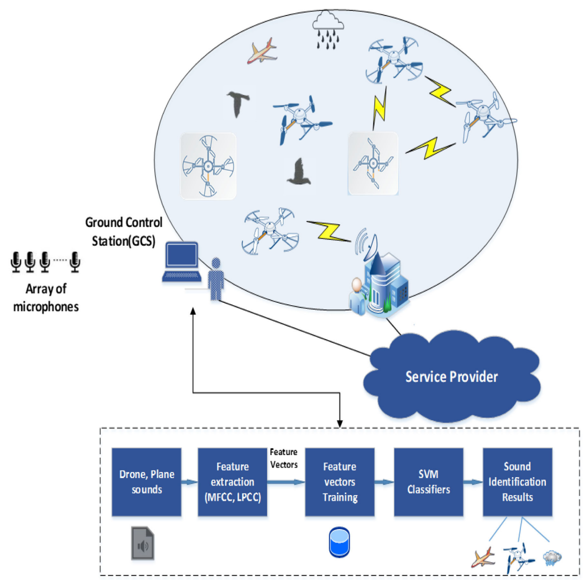
UAV has a specific acoustic signatures to differentiate them from other sounds in the surroundings. To achieve higher accuracy and efficiency in sound recognition and classification, valuable sound features plays a vital role. Hence, it is very important to select meaningful features from sound samples to get the accurate detection results. Giving a raw audio signal to sound classifier is not appropriate and will lead to inaccurate results because of too much redundant information. Since, sound contain silent regions and noise as well, feature vectors having prominent frequency components are extracted from the audio signal and given to the sound classifier. Feature vector must contain the information that can help to easily differentiate the various sounds received from different sources.

The system model adopted for sound-based drone detection is shown in Fig. 2(a). We deployed arrays of microphones to gather sounds of various objects like drones, airplanes, birds, and thunderstorm. These sounds are passed to ground control station (GCS) where each features are extracted using ML techniques. In this paper, we opted two most popular techniques such as LPCC and MFCC to extract sound prominent features. They extract high frequency features in cepstral domain which precisely extracts the peaks from spectrum envelope. After getting feature vectors, the information is given to the trained system that classifies using SVM classifier. To achieve more accuracy, the extracted features are classified using various SVM kernels. The flow chart of the proposed system is shown in Fig. 2(b).

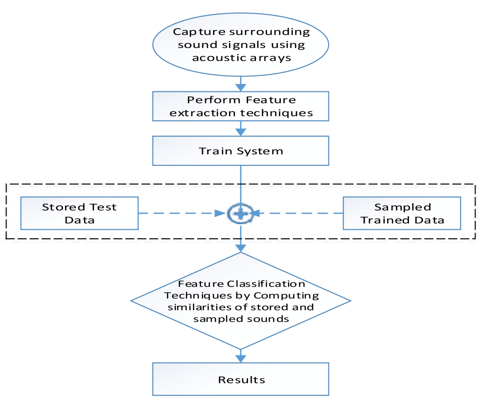
K-fold cross-validation technique is adopted here to evaluate the performance of the classifier. In K-fold cross-validation, we separate our dataset into K-folds, and out of these we use $K - 1$ folds for training and last fold as a test dataset. Then this process is randomly repeated to use another $K - 1$ fold for system training and one as a test data [34]. This process is repeated N times using different data portions to optimize the model accuracy. A 10-fold cross validation to optimize the system accuracy is adopted here.

A. Linear Predictive Cepstral Coefficients (LPCC) Computation

LPCC is one of important technique to estimate the basic parameters such as pitch of an acoustic signal. In LPCC, cepstral coefficients can be computed in two ways; one is through linear prediction (LP), and second approach is by using filter banks. In this paper, LP approach is adopted to compute the cepstral



(a) Drone detection system model



(b) Drone detection system work flow

Fig. 2. System scenarios. (a) Sound-based drone detection system model. (b) Sound-based drone detection system work flow.

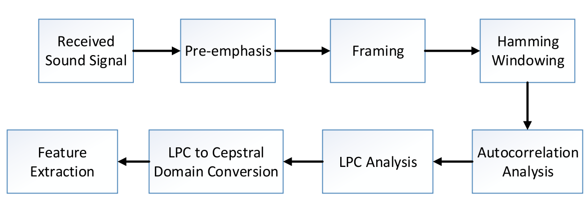


Fig. 3. LPCC coefficients computation steps.

coefficients. The main steps to calculate the cepstral coefficients are depicted in Fig. 3 and are briefly discussed here; 1) pre-emphasis; to spectrally flatten and make it less susceptible for finite precision effects, 2) frame blocking; is performed to accumulate the pre-emphasized samples into frames, 3) windowing; to window each frame in order to minimize the discontinuities at the start and end of the frames, and generally hamming window is used for this purpose. 4) Autocorrelation analysis: is performed for each frame of windowed signal to find the autocorrelation coefficients. 5) Linear prediction coding (LPC) analysis: here Durbin's method is used to convert the $p+1$ autocorrelation coefficients of each frame into LPC parameters, where p represents the order of the LPC, 6) LPC parameter to

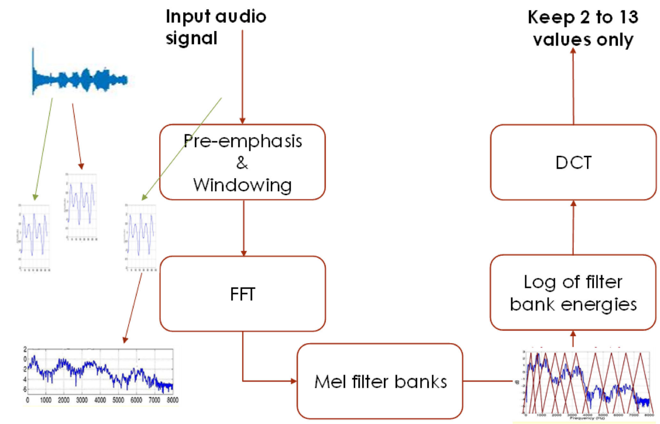


Fig. 4. MFCC computation major steps.

cepstral coefficient: finally the cepstral coefficients are directly calculated here by the following recursive relation

$$C_l = \begin{cases} a_l + \sum_{k=1}^{l-1} \left(\frac{k}{l}\right) C_k a_{l-k}, 1 \leq m \leq p \\ \sum_{k=1}^{l-1} \left(\frac{k}{l}\right) C_k a_{l-k}, l > p \end{cases} \tag{1}$$

where a_l represents the l-th LPC coefficient and C_l are the cepstral coefficients.

B. Mel Frequency Cepstrum Coefficients (MFCC) Computation

MFCC being the most widely used feature extraction scheme in frequency domain. MFCC due to its frequency domain feature is very accurate as compared to LPCC time domain features. In MFCC, Mel frequencies are wrapped using triangular filters which give better demonstration of sounds [17]. Short term analysis is used to compute MFCC, and from each frame its feature vectors are calculated. The major differentiating factor of MFCC with LPCC cepstrum is that its frequency axis is wrapped to Mel scale. Most of the steps included in MFCC computation is similar to LPCC with some additional computation steps. These are processing in frequency domain, usage of Mel filter banks, and discrete cosine transform (DCT) of filter banks output to de-correlates the overlapping coefficients as shown in Fig. 4.

The frequency in hertz can be converted into Mel frequency scale as

$$mel(f) = 2595 \log 10(f/700 + 1) \tag{2}$$

Here mel represents the frequency in Mel scale and f in hertz. By taking the log of filter banks output, we get Mel spectrum and finally DCT is performed to get Mel cepstrum coefficients as

$$c(n) = \sum_{m=0}^{M-1} (\log D(m) \cos \left(\frac{\pi n(k-0.5)}{M}\right); n = 0, 1, ...C-1 \tag{3}$$

where $c(n)$ represents the MFCC coefficient, C is the number of MFCC coefficients, $D(m)$ is the mel spectrum of the magnitude spectrum, obtained by multiplying the magnitude spectrum by triangular Mel weighting filters, and m is the m-th triangular filter coefficient. In this paper, $n = 13$ is considered because

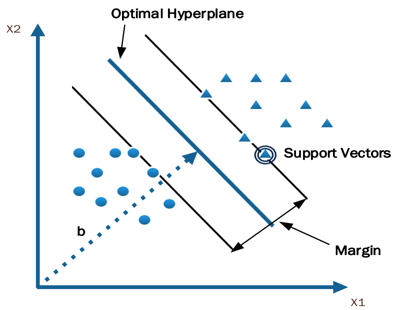


Fig. 5. SVM classifier model.

higher DCT coefficients changes rapidly and degrades the audio recognition computation, so we ignore higher coefficients.

C. Support Vector Machines Based Feature Classification

In this paper, SVM as a supervised learning model is considered for data classifications. SVM acquires an optimal hyperplane having set of positive and negative samples and work on structural risk minimization principle. That is, it minimizes the chances of misclassification between classes as shown in Fig. 5. SVM selects the optimal decision boundary depending upon maximum margin which optimally separates the data points. Classification error ratio is minimized as margin increases, and hence larger the margin which results in minimum error [35].

The training points nearer to the optimum separating hyper plane are the support vectors [36]. It can be represented as

$$w^T x + b = \pm 1 \tag{4}$$

where b is bias, w and x represents the vectors, and w is perpendicular to linear decision boundary. From training data set, w vector is computed. When input data goes into a higher dimensional space, learning takes place to improve its capability of separating data points of various classes by finding the best hyperplane. This method is usually known as kernel trick. The main advantage of SVM kernels are their capability to work in any dimensions without any additional computations and complexity [37]. SVM performs very well for noisy and high dimensional data and achieves high accuracy on small data set. This motivates us to select SVM as a classifier. For SVM classification accuracy, selecting an appropriate kernel plays a vital role. We assessed the SVM performance with three different kernels, i.e., linear, polynomial, and Gaussian kernel. The linear kernel is represented as

$$K(x_i, x_j) = x_i^T x_j \tag{5}$$

Similarly, the polynomial kernel is computed as

$$K(x_i, x_j) = \left(1 + x_i^T x_j\right)^p \tag{6}$$

where x_i and x_j are used for computing two vectors dot product, and are plotted in a space of order p. The Gaussian kernel is

$$K(x_i, x_j) = \frac{\exp\ \left(-\|x_i - x_j\|^2\right.}{2\sigma^2} \tag{7}$$

where $\|x_i - x_j\|$ tells the euclidean distance between two samples. Width of Gaussian kernel can be set by variance σ that controls the classifier performance.

IV. Experimental Results

In this section, we evaluated the proposed sound-based detection and classification schemes by gathering real-time acoustic data. The acquired data is detected and classified by implementing MFCC, LPCC and SVM using MATLAB.

A. Acoustic Data Description

We evaluated the performance of the proposed method with acoustic data set collected by performing real-time test experiments. We collected 4 different sets of sound data; birds, drones, thunderstorm, and airplanes in a real noisy environment. We trained the system with total 138 samples and for testing we used 34 sound samples. Samples varies in length but maximum time length is 55 sec.

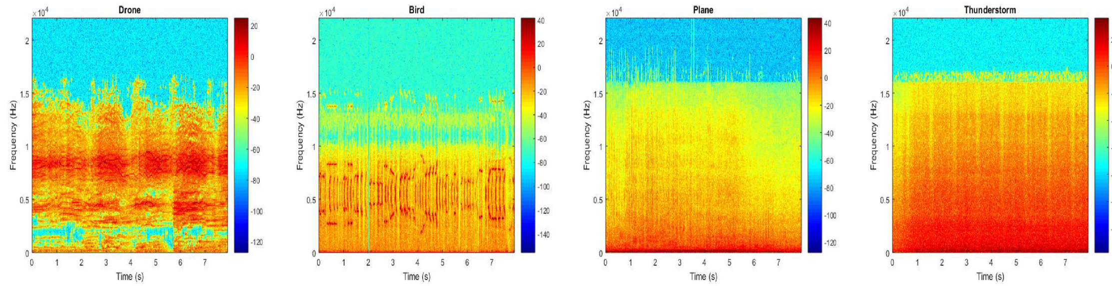
B. Drone Detection Using Spectrograms and Feature Vectors

To judge the performance of the proposed ML inspired ADr sound detection scheme, the spectrograms and Mel cepstrums are used. Fig. 6 shows the spectrograms of different classes that includes drone, bird, plane, and thunderstorm sounds. We can observe a strong red horizontal line from drone spectrum that indicates the existence of a drone sound at a particular frequency. For remaining sounds, the harmonic lines are not strong enough because they only have low frequency components. Similarly, in Fig. 7 we plotted Mel cepstrum that shows the compact view of sound spectrum, and shows the MFCC coefficients in each frame, which capture main features and aspects of spectral shape.

Fig. 8 (a)–(b) shows the feature vectors computed by LPCC and MFCC schemes. LPCC generated 10 features for each class with normalized amplitudes. For all classes, it can be observed that only first two features accompanied major sound information as compared to higher features coefficients. In MFCC, features are generated frame by frame using filter-banks for each sound files. As compared with LPCC features, MFCC features vectors are providing more detailed and accurate information due to Mel scale representation.

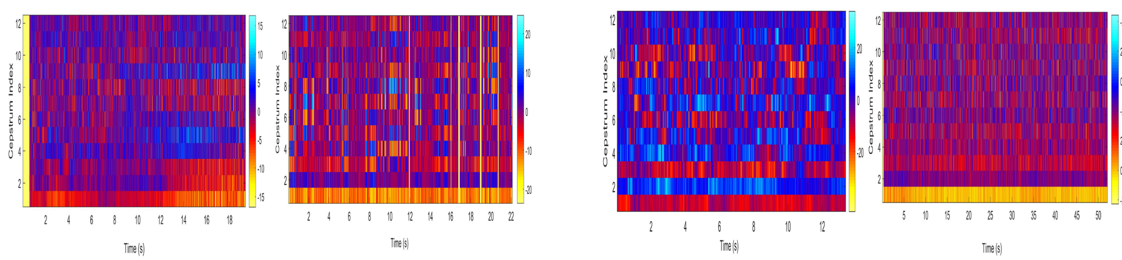
C. Detection System Evaluation Using Confusion Matrices

The performance in terms of detection accuracy of the proposed system is measured by confusion matrix. The diagonal elements in a matrix shows the number and percentage of correctly predicted class, whereas the wrong prediction assigns to wrong class. The following are the important parameters need to be defined: True positive (TP) tells us probability of correct predictions of classes relative to overall predictions; false positive (FP) tells us probability of incorrect predictions of classes relative to overall predictions. Similarly, true negative (TN) tells us probability of true predictions of classes relative to total samples belonging to that class; false negative (FN) tells us probability of wrong predictions of classes relative to total samples belonging to that class; Sensitivity also known as true positive rate (TPR) that gives us the information about the proportion of actual positives that has been correctly classified. The 100% sensitivity means false negative is zero, and all the predictions done by the



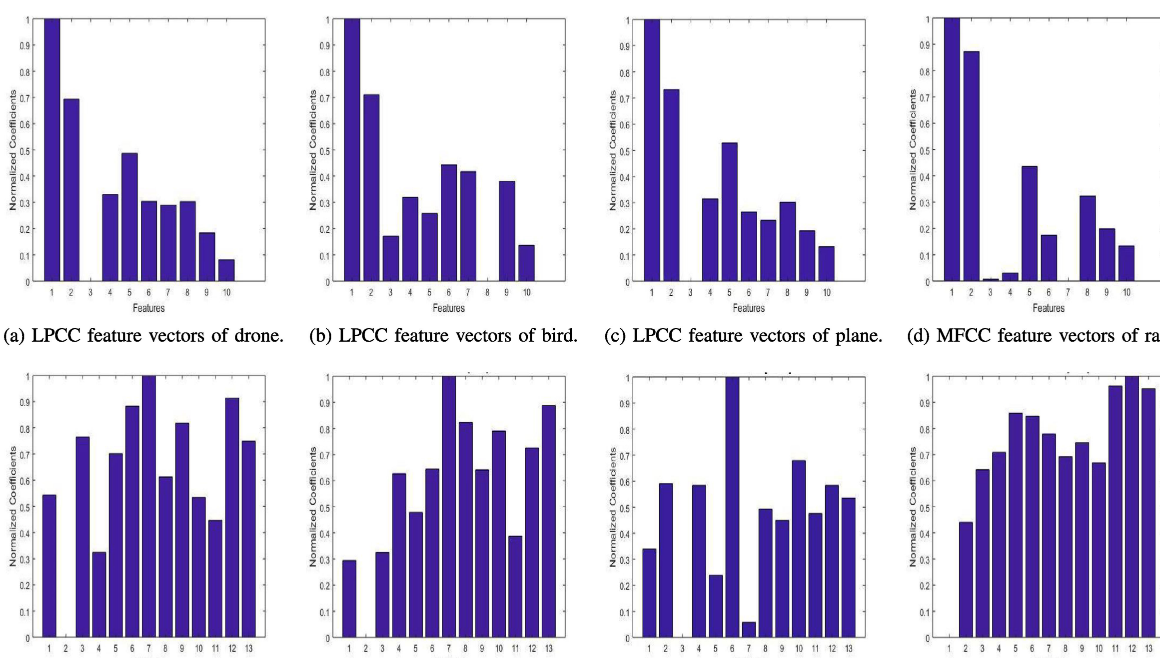
(a) drone (b) bird (c) Plane (d) Thunderstorm

Fig. 6. Spectrograms of various sounds.



(a) Mel cepstrum of plane. (b) Mel cepstrum of bird. (c) Mel cepstrum of drone. (d) Mel cepstrum of rain.

Fig. 7. Mel cepstrum of various sounds. (a) Airplane. (b) Bird. (c) Drone. (d) Rain.



(a) LPCC feature vectors of drone. (b) LPCC feature vectors of bird. (c) LPCC feature vectors of plane. (d) MFCC feature vectors of rain.

(e) MFCC feature vectors of drone. (f) MFCC feature vectors of bird. (g) MFCC feature vectors of plane. (h) MFCC feature vectors of rain.

Fig. 8. Feature vectors of LPCC and MFCC for various sounds.

classifier is correct. The sensitivity is calculated as

$$Senstivity(TPR) = \frac{TP}{TP + FN} \tag{8}$$

False negative rate (FNR) tells the probability of wrongly predicted classes as

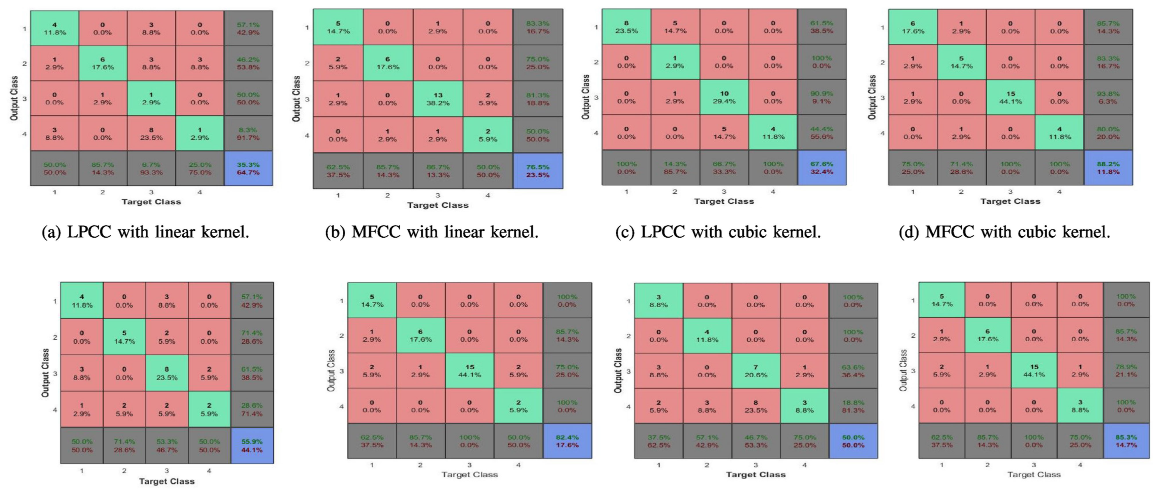
$$FNR = \frac{FN}{TP + FN} = 1 - TPR \tag{9}$$

Positive predicted value (PPV) tells the number of correct positive instances as

$$PPV = \frac{TP}{TP + FP} \tag{10}$$

False discovery rate (FDR) tells the number of false negative instances as

$$FDR = \frac{FP}{TP + FP} = 1 - PPV \tag{11}$$



(a) LPCC with linear kernel. (b) MFCC with linear kernel. (c) LPCC with cubic kernel. (d) MFCC with cubic kernel.

(e) LPCC with quadratic kernel. (f) MFCC with quadratic kernel. (g) LPCC with Gaussian kernel. (h) MFCC with Gaussian kernel.

Fig. 9. Confusion matrices for LPCC and MFCC with various kernels.

TABLE I
SVM CLASSIFICATION RESULTS WITH LPCC USING DIFFERENT KERNEL FUNCTIONS

(a) Linear Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	63.1%	57.2%	66.6%
Thunderstorm	60%	46.1%	85.7%
Drones	44.4%	50%	44%
Planes	46.1%	8.3%	78.5%

(b) Cubic Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	81.1%	61.5%	100%
Thunderstorm	78.3%	100%	78.5%
Drones	75.2%	90.9%	72.2%
Planes	80.4%	44.4%	100%

(c) Quadratic Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	73.08%	57.1%	78.9%
Thunderstorm	81.6%	71.4%	87.5%
Drones	61.2%	60.5%	61.1%
Planes	73%	28.6%	88.4%

(d) Gaussian Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	77.2%	100%	73.6%
Thunderstorm	84%	100%	81.2%
Drones	58.6%	63.6%	55.5%
Planes	54.8%	18.8%	93.3%

In (12), the specificity parameter which measures the proportion of correctly identified negatives instances as

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

Overall accuracy and error of classifier is computed as

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{13}$$

$$Error = \frac{FP + FN}{TP + FP + TN + FN} = 1 - Accuracy \tag{14}$$

TABLE II
SVM CLASSIFICATION RESULTS WITH MFCC USING DIFFERENT KERNEL FUNCTIONS

(a) Linear Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	86.6%	83.3%	87.5%
Thunderstorm	81.5%	75%	95.2%
Drones	84.8%	81.3%	86.6%
Planes	85.2%	50%	92.3%

(b) Cubic Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	90.9%	85.7%	92.3%
Thunderstorm	90.1%	83.3%	92.5%
Drones	96.7%	93.7%	100%
Planes	95.8%	80%	100%

(c) Quadratic Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	90.3%	100%	88.4%
Thunderstorm	93.3%	85.7%	95.6%
Drones	85.4%	75%	100%
Planes	92.1%	100%	92.8%

(d) Gaussian Kernel: Accuracy, Sensitivity and Specificity for each event.

Events	*Accuracy*	*Sensitivity*	*Specificity*
Birds	90.6%	100%	88.8%
Thunderstorm	93.5%	85.7%	95.8%
Drones	87.7%	78.9%	100%
Planes	96.6%	100%	96.3%

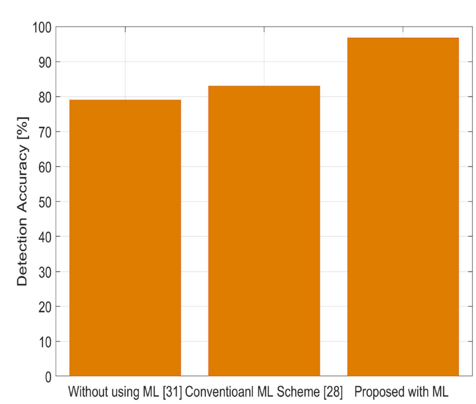
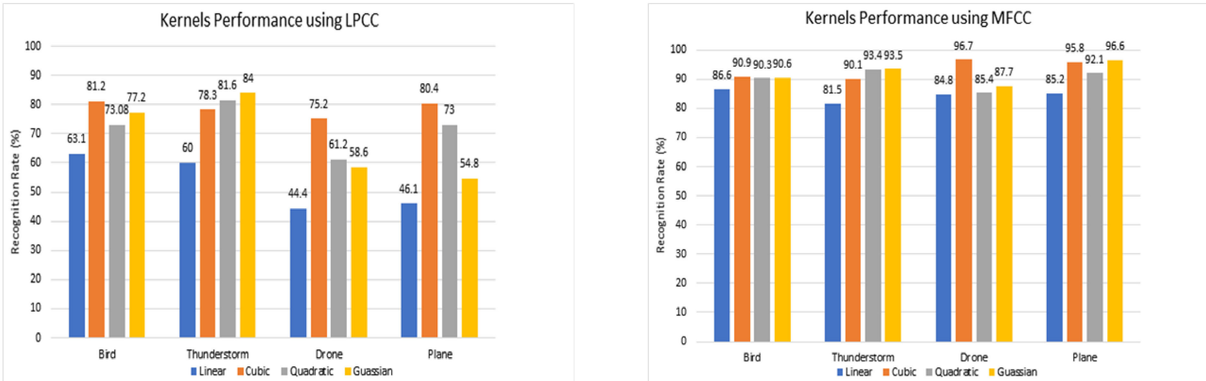


Fig. 10. Comparison among the proposed ADr detection, conventional scheme, and without using machine learning schemes.



(a) Accuracy rate using SVM kernels with LPCC.

(b) Accuracy rate using SVM kernels with MFCC.

Fig. 11. Kernels comparison. (a) SVM kernels with LPCC. (b) SVM kernels with MFCC approach.

We evaluated the detection accuracy of the proposal with different SVM kernels. The performance of the SVM classifier is analyzed using the confusion matrix. In a matrix, event 1, 2, 3 and 4 are representing birds, thunderstorm, drones and airplane classes, respectively. Fig. 9 represents the LPCC and MFCC confusion matrices with different kernels. The overall ADr detection accuracy for the cubic SVM kernel with LPCC and MFCC approaches around 67.6% and 88.2%, respectively, which is the highest among the others kernels. With cubic kernels using MFCC, our model is able to correctly predict drone instances with precision rate of 93.8% while only 6.3% FDR.

We noticed that LPCC has very low sensitivity, since for thunderstorm and drones detection it only gives 14.3% and 66.7% (TPR). For drones, five trials were wrongly predicted as bird. From this, we infer that LPCC features are not robust for ADr detection in noisy urban environment because of its high misclassification ratio.

The SVM kernels using MFCC features set outperformed LPCC, and gives better classification accuracy. The classification accuracy of linear kernel is low as compared to other classifiers because of the non-linear and non-stationary properties of the drone acoustic signals. It demonstrates that MFCC approach can be effectively applied to predict ADr despite small number of data set. Based on small data set, the proposed sound-based detection model achieved high performance for detecting drones. Hence, we conclude that the cubic SVM kernel gives better ADr detection accuracy with MFCC.

1) SVM Kernels With LPCC: Table I shows the rate of change in accuracy with various SVM kernels using LPCC. The detection performances are computed in terms of matrices like accuracy, sensitivity and specificity. It can be observed that for birds and thunderstorm events; accuracy, sensitivity and specificity has higher performance than the other two events. When comparison is done under various kernels using LPCC, the drone detection accuracy with cubic kernel is around 75% as compared to linear kernel that gives minimum accuracy of only 44.4%. Due to low sensitivity and specificity factor, overall drone detection probability is very low. The main limitation of this scheme is the noisy environment and weak feature classification set.

2) SVM Kernels With MFCC: Similarly, Table II shows the rate of change in accuracy with various SVM kernels using MFCC approach. Results demonstrates that around 80% of overall accuracy, sensitivity and specificity is achieved. As compared to other kernels, linear kernel has lowest accuracy and sensitivity rate for all instances. We also compare the performance of the proposed scheme with existing ML scheme that adopted PIL and KNN classifiers and with no ML scheme. Results in Fig. 10 demonstrate that the proposed scheme achieve around 97% detection accuracy as compared with existing ML and other conventional schemes without ML, which have around 83% and 79% detection accuracy respectively. MFCC performs well in dealing with noisy data-set because it only give importance to high frequencies using Mel scale.

3) Kernels Comparison in Terms of Accuracy: Fig. 11 (a)-(b) represents the performance of SVM kernels with LPCC and MFCC schemes. Kernels results indicates that for detecting ADr, MFCC features appears much better than LPCC features. SVM with MFCC using a cubic and Gaussian kernels shows a higher detection rates in terms of accuracy than with a linear kernel. In case of LPCC, accuracy rate relatively decreases. The experimental results show that SVM classifier with cubic kernel provides the highest detection accuracy around 96.7% with MFCC features. In comparison, cubic kernel among the other kernels achieves maximum accuracy rate of 75% using LPCC features. MFCC scheme outperforms because of its frequency domain characteristics which provide better diversity gain in terms of kernels. From these results, we inferred that the SVM is a more suitable and reliable classifier for the classification by adopting MFCC for feature extraction.

V. Conclusion

UAV deployment is raising some serious security challenges because of its capability to carry explosive materials. Hence, amateur drone detection should be timely countered with high precision rate to avoid such instances. In this paper, we compare the performance of the proposed machine learning inspired sound-based amateur drone detection scheme in terms of accuracy using MFCC and LPCC with SVM classifier. Results show that the proposed scheme can work well in noisy environment by adopting MFCC with SVM cubic kernal. Using a

SVM classifier, we classify samples from a small data set with good accuracy. Simulation results demonstrate the effectiveness and improvement in the accuracy rate with MFCC approach in noisy environment. It has the capability to correctly classify with a probability of around 85% under various kernels. Furthermore, we also compare the MFCC and LPCC methods with various SVM kernels and analyze each class performance and accuracy. Finally, we conclude that the proposed scheme achieve around 97% detection accuracy as compared with existing ML and other conventional schemes without ML, which have around 83% and 79% detection accuracy respectively. Results proved that the acoustic-based drone detection by using ML is the cost effective and accurate tool for amateur drone detection and it can be easily adopted by the National security agencies. Since, we targeted to achieve desirable accuracy with small data set, thus we implemented machine learning schemes instead of deep learning techniques, that requires huge data set for achieving high accuracy. In future, we have a plan to conduct much larger scale experiments with bigger data set to increase the detection accuracy using deep neural network.

REFERENCES

[1] Z. M. Fadlullah, D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, "A dynamic trajectory control algorithm for improving the communication throughput and delay in UAV-aided networks," *IEEE Netw.*, vol. 30, no. 1, pp. 100–105, Jan./Feb. 2016.

[2] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. Global Commun. Conf.*, 2014, pp. 2898–2904.

[3] Z. Kaleem *et al.*, "Amateur drone surveillance: Applications, architectures, enabling technologies, and public safety issues: Part 1," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 14–15, Jan. 2018.

[4] F. Tang, Z. M. Fadlullah, N. Kato, F. Ono, and R. Miura, "AC-POCA: Anticoordination game based partially overlapping channels assignment in combined UAV and D2D-based networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1672–1683, Feb. 2018.

[5] Z. Kaleem, Y. Li, and K. Chang, "Public safety users priority-based energy and time-efficient device discovery scheme with contention resolution for ProSe in third generation partnership project long-term evolution-advanced systems," *IET Commun.*, vol. 10, no. 15, pp. 1873–1883, 2016.

[6] G. Z. Khan, E.-C. Park, and R. Gonzalez, "M 3-cast: A novel multicast scheme in multi-channel and multi-rate wifi direct networks for public safety," *IEEE Access*, vol. 5, pp. 17 852–17 868, 2017.

[7] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Stochastic geometry study on device-to-device communication as a disaster relief solution," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3005–3017, May 2016.

[8] Z. Kaleem and K. Chang, "Public safety priority-based user association for load balancing and interference reduction in PS-LTE systems," *IEEE Access*, vol. 4, pp. 9775–9785, 2016.

[9] F. Tang, Z. M. Fadlullah, B. Mao, N. Kato, F. Ono, and R. Miura, "On a novel adaptive UAV-mounted cloudlet-aided recommendation system for LBSNs," *IEEE Trans. Emerg. Topics Comput.*, doi: 10.1109/TETC.2018.2792051.

[10] A.-H. Akram, S. Chandrasekharan, G. Kaandorp, W. Glenn, A. Jamalipour, and S. Kandeepan, "Coverage and rate analysis of aerial base stations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 3077–3081, Dec. 2016.

[11] D. He, S. Chan, and M. Guizani, "Drone-assisted public safety networks: The security aspect," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 218–223, Aug. 2017.

[12] Z. Kaleem, M. Z. Khaliq, A. Khan, I. Ahmad, and T. Q. Duong, "PS-CARA: Context-aware resource allocation scheme for mobile public safety networks," *Sensors*, vol. 18, no. 5, pp. 1–17, 2018.

[13] Z. Kaleem and M. H. Rehmani, "Amateur drone monitoring: State-of-the-art architectures, key enabling technologies, and future research directions," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 150–159, Apr. 2018.

[14] İ. Güvenç, O. Ozdemir, Y. Yapici, H. Mehrpouyan, and D. Matolak, "Detection, localization, and tracking of unauthorized UAS and jammers," in *Proc. 36th IEEE/AIAA Digit. Avionics Syst. Conf.*, 2017, pp. 1–10.

[15] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on the cognitive internet of things," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 29–35, Jan. 2018.

[16] I. Guvenc, F. Koohifar, S. Singh, M. L. Sichitiu, and D. Matolak, "Detection, tracking, and interdiction for amateur drones," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 75–81, Apr. 2018.

[17] A. Kumar, S. S. Rout, and V. Goel, "Speech mel frequency cepstral coefficient feature classification using multi level support vector machine," in *Proc. 4th IEEE Uttar Pradesh Sect. Int. Conf. Elect., Comput. Electron.*, 2017, pp. 134–138.

[18] L. Grama, L. Tuns, and C. Rusu, "On the optimization of SVM kernel parameters for improving audio classification accuracy," in *Proc. 14th Int. Conf. Eng. Modern Elect. Syst.*, Jun. 2017, pp. 224–227.

[19] H. Liu, Z. Wei, Y. Chen, J. Pan, L. Lin, and Y. Ren, "Drone detection based on an audio-assisted camera array," in *Proc. IEEE 3rd Int. Conf. Multimedia Big Data*, 2017, pp. 402–406.

[20] S. J. Lee, J. H. Jung, and B. Park, "Possibility verification of drone detection radar based on pseudo random binary sequence," in *Proc. IEEE Int. SoC Design Conf.*, 2016, pp. 291–292.

[21] J. Drozdowicz *et al.*, "35 GHz FMCW drone detection system," in *Proc. IEEE 17th Int. Radar Symp.*, 2016, pp. 1–4.

[22] Z. Fadlullah *et al.*, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 2432–2455, Forth Quater 2017.

[23] N. Kato *et al.*, "The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 146–153, Jun. 2017.

[24] F. Tang *et al.*, "On removing routing protocol from future wireless networks: A real-time deep learning approach for intelligent traffic control," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 154–160, Feb. 2018.

[25] Z. M. Fadlullah *et al.*, "On intelligent traffic control for large scale heterogeneous networks: A value matrix based deep learning approach," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2479–2482, Dec. 2018.

[26] F. Tang, B. Mao, Z. M. Fadlullah, and N. Kato, "On a novel deep-learning-based intelligent partially overlapping channel assignment in SDN-IoT," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 80–86, Sep. 2018.

[27] F. Tang, Z. M. Fadlullah, B. Mao, and N. Kato, "An intelligent traffic load prediction based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5141–5154, Dec. 2018.

[28] G. J. Mendis, T. Randeny, J. Wei, and A. Madanayake, "Deep learning based doppler radar for micro UAS detection and classification," in *Proc. IEEE Mil. Commun. Conf.*, 2016, pp. 924–929.

[29] H. Rydén, S. B. Redhwan, and X. Lin, "Rogue drone detection: A machine learning approach," 2018, arXiv:1805.05138.

[30] J. Kim, C. Park, J. Ahn, Y. Ko, J. Park, and J. C. Gallagher, "Real-time UAV sound detection and analysis system," in *Proc. Sensors Appl. Symp.*, 2017, pp. 1–5.

[31] J. Mezei and A. Molnár, "Drone sound detection by correlation," in *Proc. 11th IEEE Int. Symp. Appl. Comput. Intell. Inform.*, 2016, pp. 509–518.

[32] T. Müller, "Robust drone detection for day/night counter-UAV with static VIS and SWIR cameras," in *Proc. Ground/Air Multisensor Interoperability, Integration, Netw. Persistent ISR VIII*, 2017, Art. no. 1019018.

[33] L. Shi, I. Ahmad, Y. He, and K. Chang, "Hidden Markov model based drone sound recognition using MFCC technique in practical noisy environments," *J. Commun. Netw.*, vol. 20, no. 5, pp. 509–518, Oct. 2018.

[34] T. T. Wong and N. Y. Yang, "Dependency analysis of accuracy estimates in K-fold cross validation," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 11, pp. 2417–2427, Nov. 2017.

[35] P. Rai, V. Golchha, A. Srivastava, G. Vyas, and S. Mishra, "An automatic classification of bird species using audio feature extraction and support vector machines," in *Proc. IEEE Inventive Comput. Technol., Int. Conf.*, 2016, vol. 1, pp. 1–5.

[36] K. S. Parikh and T. P. Shah, "Support vector machine–a large margin classifier to diagnose skin illnesses," *Procedia Technol*, vol. 23, pp. 369–375, 2016.

[37] N. V. Mankar, A. Khobragade, and M. M. Raghuwanshi, "Classification of remote sensing image using SVM kernels," in *Proc. World Conf. Futuristic Trends Res. Innov. Social Welfare*, Feb. 2016, pp. 1–5.

Muhammad Zohaib Anwar received the M.S. degree in electrical and computer engineering from COMSATS University Islamabad, Islamabad, Pakistan. His areas of interest include 5G technologies, unmanned aerial vehicles, and machine learning application in 5G.

Zeeshan Kaleem received the Ph.D. degree in electronics engineering from Inha University, Incheon, South Korea, in 2016. He is an Assistant Professor in electrical and computer engineering, COMSATS University Islamabad, Islamabad, Pakistan. He has authored or coauthored more than 35 technical Journals and conference papers in reputable venues and also holds 18 US and Korean Patents. His current research interests includes public safety networks, 5G system testing and development, and unmanned air vehicle communications. He was the recipient of the National Research Productivity Award from Pakistan Council of Science and Technology in 2016–17 and 2017–18. He is the co-recipient of Best Research Proposal Award from SK Telecom, Korea. He is serving as an Associate Technical Editor for prestigious Journals/Magazine such as the IEEE Communications Magazine and IEEE Access, *Elsevier Computer and Electrical Engineering*, *Springer Human-Centric Computing and Information Sciences*, and *Journal of Information Processing Systems*. He was a Guest Editor for special issues of IEEE Access, *Sensors*, IEEE/KICS Journal of Communications and Networks, and *Physical Communication*, and also served as TPC for world distinguished conferences such as IEEE Vehicular Technology Conference, the IEEE International Conference on Communications, and the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.

Abbas Jamalipour (S'86–M'91–SM'00–F'07) received the Ph.D. degree in electrical engineering from Nagoya University, Nogoya, Japan. He is a Professor of ubiquitous mobile networking with the University of Sydney, Sydney, NSW, Australia. He has authored or coauthored nine technical books, 11 book chapters, more than 450 technical papers, and five patents, all in the area of wireless communications. He is a fellow of the Institute of Electrical, Information, and Communication Engineers and the Institution of Engineers Australia, an ACM Professional Member, and the IEEE Distinguished Lecturer. He is an elected member of the Board of Governors, Executive Vice-President, a Chair of Fellow Evaluation Committee, and an Editor-in-Chief of the Mobile World, IEEE Vehicular Technology Society. He was the Editor-in-Chief of the IEEE Wireless Communications, Vice President-Conferences and a member of Board of Governors of the IEEE Communications Society, and has been an Editor for several journals. He has been a General Chair or Technical Program Chair for a number of conferences, including the IEEE International Conference on Communications, Global Communications Conference, Wireless Communications and Networking Conference, and International Symposium on Personal, Indoor, and Mobile Radio Communications. He is the recipient of a number of prestigious awards such as the 2016 IEEE ComSoc Distinguished Technical Achievement Award in Communications Switching and Routing, the 2010 IEEE ComSoc Harold Sobol Award, the 2006 IEEE ComSoc Best Tutorial Paper Award, as well as 15 Best Paper awards.