

# Impact of Seasonal Variations on Foliage Penetration Experiment: A WSN-Based Device-Free Sensing Approach

Yi Zhong<sup>ID</sup>, Yang Yang, *Member, IEEE*, Xi Zhu<sup>ID</sup>, Yan Huang, Eryk Dutkiewicz, *Member, IEEE*, Zheng Zhou, *Member, IEEE*, and Ting Jiang

**Abstract**—Foliage penetration (FOPEN) has been found to be a critical mission for a variety of applications, ranging from surveillance to military. Recently, an emerging technology, namely wireless sensor network (WSN)-based device-free sensing (DFS), has been introduced to the domain of FOPEN. This technology only utilizes radio-frequency signals for target detection and classification; thus, no additional hardware is required, just a wireless transceiver. Although the feasibility of using this technology for human detection indoors has been explored to some extent, it is questionable if the same technology can be transferred to outdoors. As far as FOPEN is concerned, the impact of seasonal variations on detection accuracy can be severe. To address this concern, in this paper, an experiment is conducted in four seasons, and how to ensure reasonable detection accuracy with seasonal variations is intensively investigated. To fully evaluate the potential of using the WSN-based DFS for FOPEN, an impulse-radio ultrawideband technology-based prototype is used to collect data samples in different seasons. Unlike the conventional approach based on a combination of statistical properties of received-signal strength and a support vector machine, this approach adopts two special measures for performance enhancement. One measure is to use a higher order cumulant (HOC) algorithm for feature extraction, so that the impact on detection accuracy due to unwanted clutters can be minimized. The other one is to determine the optimal parameters of the classifier by means of a flower pollination algorithm. Consequently, the adverse effects on detection accuracy due to variations of weather conditions in four seasons can be accommodated. According to the experimental result, it is shown that the average classification accuracy of the presented approach can be improved by at least 20% under all seasons with an ensured robustness.

**Index Terms**—Device-free sensing (DFS) technology, feature extraction, flower pollination algorithm (FPA), foliage penetration (FOPEN), higher order cumulant (HOC), impulse-radio ultrawideband (IR-UWB), support vector machine (SVM).

Manuscript received July 20, 2017; revised December 9, 2017; accepted January 17, 2018. Date of publication March 2, 2018; date of current version August 27, 2018. This work was supported in part by NSFC under Grant 61671075 and Grant 61631003 and in part by ARC under Grant DE160101032. (*Corresponding author: Yan Huang*)

Y. Zhong is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: yi.zhong@student.uts.edu.au).

Y. Yang, X. Zhu, Y. Huang, and E. Dutkiewicz are with the Global Big Data Technology Center, School of Computing and Communications, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: yan.huang-3@student.uts.edu.au).

Z. Zhou and T. Jiang are with the Key Laboratory of Universal Wireless Communication, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2018.2804346

## I. INTRODUCTION

FOLIAGE penetration (FOPEN), in particular human detection, is an indispensable technique that has been widely deployed for many applications, ranging from military and surveillance to search and rescue [1]–[5]. The traditional approaches have been revealed to be deficient due to ever-increasing requests for safety, security, and environment protection at the local and national levels. Previously, most studies focused on using either synthetic aperture radar (SAR) [1], [2] or hybrid-sensor systems [3]–[5] for FOPEN-related applications. Although both approaches can effectively differentiate targets, with a low false-alarm probability, the implementation as well as the installation and maintenance costs for such systems are relatively high, which makes them unfeasible for large-scale deployment in practice. In order to effectively solve this problem, a cost-effective monitoring approach with a relatively high reliability is urgently required.

As an emerging technology, wireless sensor network (WSN)-based device-free sensing (DFS) has gained extensive attention for cost-effective sensing [6]–[8]. It only uses radio-frequency (RF) signals for sensing; thus, it requires no additional hardware but an RF transceiver. Consequently, it has a potential to dramatically reduce the overall cost. As every sensor node within a WSN inherently has an RF transceiver, it would be a natural choice to fully utilize this hardware for not only data communication but also target detection as well as classification.

Although magnificent progress has been made for indoor target detection and localization by analyzing statistical properties (SPs) of the strength of the received RF signal, deploying this approach amongst foliage or in any other environment that has time-varying clutter has not yet been fully explored. There are still some fundamental issues to be tackled [9], [10]. In particular, the classification accuracy needs to be significantly improved before it can be considered a practical approach. Unfortunately, very few articles that address this issue can be found in the literature [11].

In an outdoor environment, the classification accuracy can be dramatically degraded by many factors. First of all, most outdoor scenarios are based on nonline-of-sight, such as FOPEN. The trunks, branches, and leaves of trees can rapidly attenuate the strength of RF signals. As a result, the weak signals that may contain important features of targets are likely

to be overwhelmed. Second, foliage is an intrinsically time-variant channel environment due to changes of atmosphere and ground conditions. Strong winds could result in moving the branches and leaves of trees and generate a large amount of time-varying clutter/noise, which is quite undesirable. As far as classification is concerned, the extracted features from received RF signals must be not only sufficiently distinctive among different types of target but also robust enough against noise. Finally, different weather conditions and seasons could also affect the performance of the presented approach due to the variations of humidity and foliage density. If the classifier was only trained using the data sets obtained in a specific condition, it would be quite difficult to ensure a good accuracy. To overcome this issue, it is desirable to train the classifier with more and more data sets. Ideally, all weather conditions under four seasons should be included. The classical support vector machine (SVM) classifier has been successfully applied to target classification [12]–[15]. However, its performance is strongly dependent upon the optimized parameters of the SVM. Using an enormous amount of training data sets to obtain the best optimized parameters for SVM could significantly increase the computational cost. Thus, how to determine the optimal parameters for a classifier in a simple but effective way also needs to be investigated in detail.

In this paper, impulse-radio ultrawideband (IR-UWB) signals are utilized for sensing, and their higher order cumulants (HOCs) are extracted as “key” features. In addition, a flower pollination algorithm (FPA) is used in conjunction with an SVM to ensure a good classification accuracy achieved in an effective way. To demonstrate the performance of the presented approach, three targets that have a similar physical size, namely a human, a wooden board, and an iron cupboard, are used in a foliage environment. A variety of data samples are taken in four seasons with different weather conditions, including sunny, rainy, snowy, and foggy. These data samples are then used to extensively verify the performance of the presented approach in terms of accuracy and robustness. In particular, the impact on classification accuracy due to different classifiers, numbers of training samples, and values of signal-to-noise ratio (SNR) is carefully evaluated. According to the experimental results, an average accuracy of 94% is achieved in terms of human target classification, which is similar to other SAR and hybrid-sensor systems [1]–[5].

## II. EXPERIMENTAL SETUP AND DATA COLLECTION

To verify the feasibility of the aforementioned approach, in particular the accuracy of distinguishing between different targets in four seasons with different weather conditions, a variety of data samples has been taken in a foliage environment. The operating frequency of the used IR-UWB transceiver covers from 3.1 to 5.3 GHz with the center frequency at 4.2 GHz. The transceiver is combined with planar elliptical dipole antennas to form a prototype that is used for data collection. In addition, the prototype is connected to a PC so that the collected data can be directly acquired and recorded. The experimental setup is given in Fig. 1.

The measurements were performed in a park, which contains a mixture of hardwood trees and low-lying undergrowth.

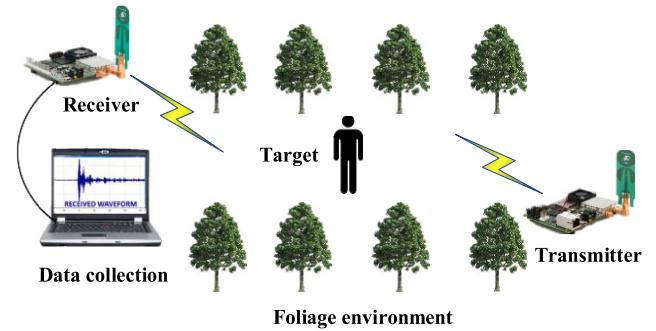


Fig. 1. Measurement setup.

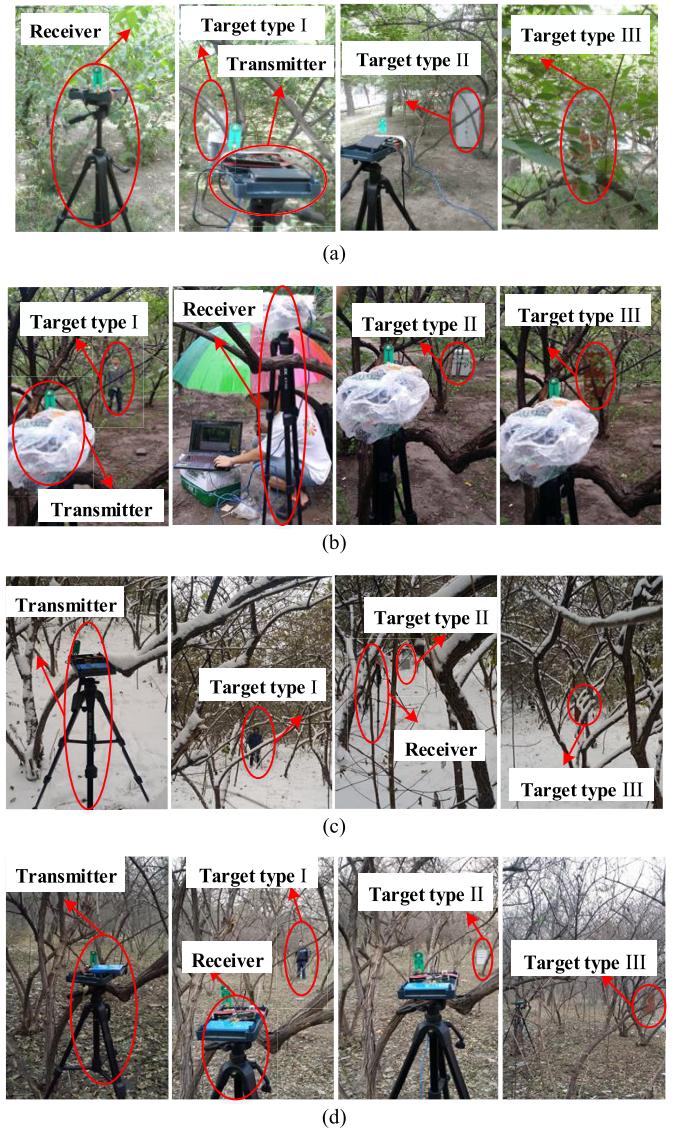


Fig. 2. Seasonal variations. (a) Sunny. (b) Rainy. (c) Snowy. (d) Foggy.

As shown in Fig. 2, measurements were performed under four typical weather conditions in different seasons, namely sunny, rainy, snowy, and foggy. Consequently, four data sets are generated and used in this paper. These data sets were collected in four seasons in 2016. For all measurements, the transceiver is placed at a fixed location at a constant height of 1.5 m

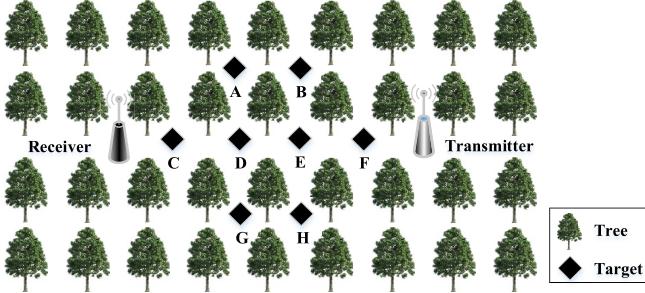


Fig. 3. Indication of the selected location for data collection in a foliage environment.

above the ground. The distance between the transmitter and the receiver is approximately 20 m in all scenarios. Three target types, including human (target type I), iron cupboard (target type II), and wooden board (target type III), are used for classification. The height and weight of the human target are 175 cm and 68 kg, respectively. The dimension of the iron cupboard is 50 cm × 40 cm × 140 cm (length × width × height). The size of the wooden board is 60 cm × 180 cm (width × height), and the thickness of the board is 2.5 cm.

To ensure that the experimental setup is appropriate and the prototype is fully functional, a penetration measurement was first taken without any placed target. Once the functionality of the prototype is confirmed, different types of target are then placed between a transmitter and a receiver for testing purposes. Since the detection as well as classification accuracy may be adversely affected by the selected locations where the target is placed, eight different locations (A, B, C, D, E, F, G, and H) are used to ensure that the system can effectively detect and classify the placed target in the entire monitoring area. These locations are summarized in Fig. 3. On the other hand, it is also worthwhile to mention that the resolution of the scanned (captured) waveform is 61 ps. Since there are 350 sampling points (limited by the used system), the overall duration of each captured waveform is  $61 \text{ ps} \times 350 = 21.35 \text{ ns}$ . However, to improve the classification accuracy, only the later part with 90 sampling points that contains strong impulse is selected for feature extraction, and signal preprocessing is required to remove unwanted waveform.

In order to collect enough data samples for the experiment, at every location, 500 data samples are taken for each type of target under each weather condition. As there are eight different locations, overall 16 000 data samples of data for each type of target are collected to form a database. These collected data are used for the purpose of target classification as will be explained in detail later on. To illustrate the challenge of target classification using a DFS-based technology, two typical scenarios are used as examples, and their corresponding waveforms of the received signals are given. In Fig. 4, the waveforms measured with different targets are compared with “no target” under a snowy condition. As can be seen, although it is relatively easy to identify if there is any target placed between the transmitter and the receiver, it may not be possible to correctly classify its type unless the features from different

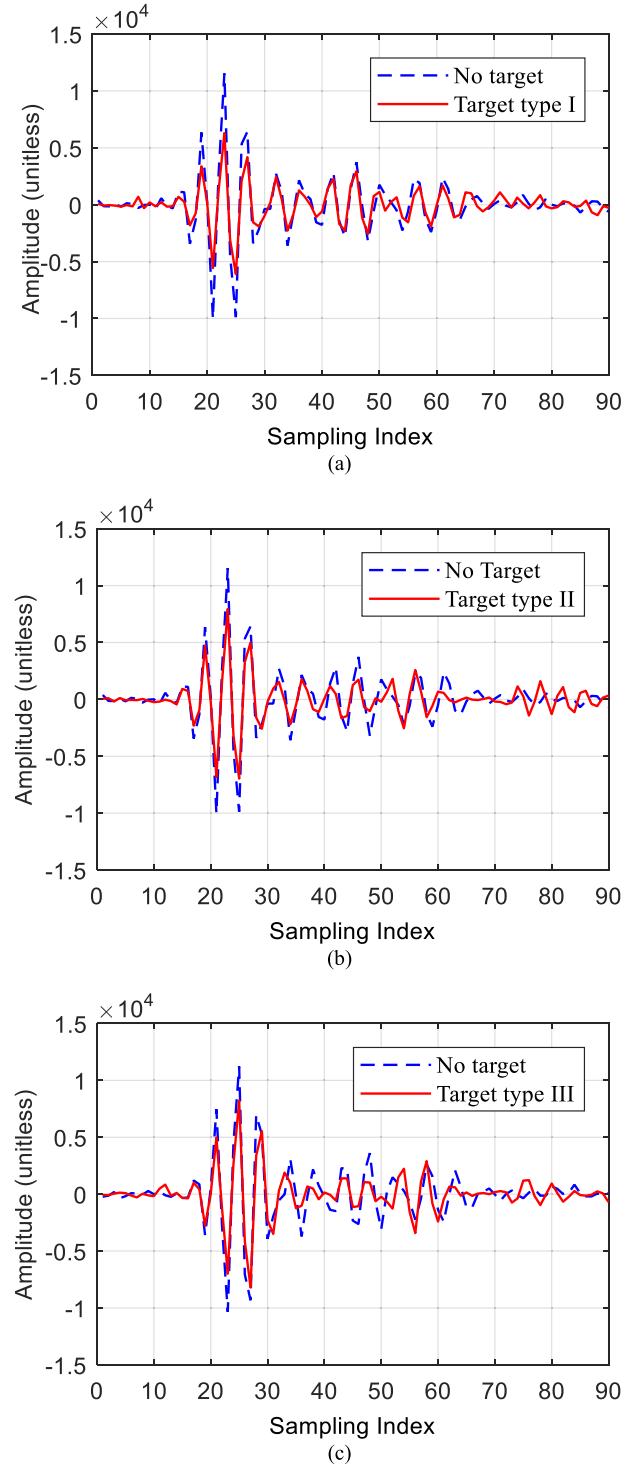


Fig. 4. Measured waveform with the different types of target under snowy conditions. (a) No target versus target type I. (b) No target versus target type II. (c) No target versus target type III.

targets can be extracted in a more distinguished way. For this reason, a HOC-based algorithm is selected for feature extraction in this paper. The detail of this method will be shown in Section III.

On the other hand, the waveforms measured for the same target (human) with different weather conditions in different seasons are shown in Fig. 5. As illustrated, the weather

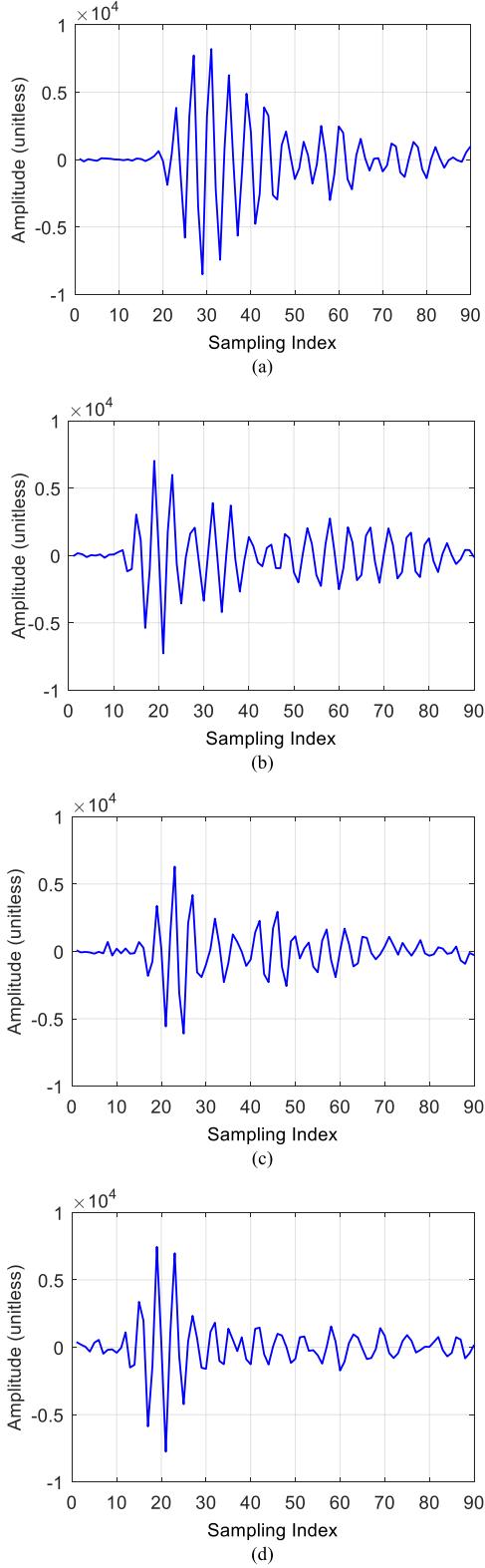


Fig. 5. Measured waveform with a human target under different weather conditions. (a) Sunny. (b) Rainy. (c) Snowy. (d) Foggy.

condition could have a significant impact on the received signals. It therefore indicates that multiple data sets must be created and then used for target classification; otherwise, the classification accuracy is very likely to be poor.

### III. HOC-BASED FEATURE EXTRACTION

Being a classical approach of signal processing, the HOC-based algorithm has been extensively used for pattern recognition [11], [16]–[20]. Since the HOC of Gaussian signals is statistically zero, this method is very well suited for the separation of additive mixtures of independent non-Gaussian signals and Gaussian noise. This feature can be exploited to not only detect but also classify non-Gaussian signals and provide a reasonable immunity in applications, where the signal source is corrupted by the Gaussian noise. As previously mentioned, one of the critical hurdles to succeed with DFS technology for human target detection in a foliage environment is to enhance the immunity of the system against noise generated by the clutter (trees and their branches). In addition, distinct features can be extracted from a higher order domain of a signal due to the fact that both amplitude and phase information are contained by the HOCs of a signal. Therefore, an HOC-based method is selected in this paper, as it has a potential to reduce false-alarm triggering [11]. The basic principle of HOC is briefly introduced here. If a collection of random variables is given as  $\{x_1, x_2, \dots, x_k\}$ , their corresponding  $k$ th-order cumulant can be defined as the coefficient of  $\{v_1, v_2, \dots, v_k\}$  in the Taylor series expansion of the cumulant-generating function

$$\psi(v) = \ln E\{\exp(jv'x)\} \quad (1)$$

where  $E[\cdot]$  is the expectation operator, representing the statistics average. The  $k$ th-order cumulant of  $x$  is defined as

$$\text{cum}(x_1, x_2, \dots, x_k) = (-j)^k [\partial/\partial v_1 \partial v_2 \cdots \partial v_k] \psi(v)|_{v=0}. \quad (2)$$

If the average of some random variables is nonzero, the expression for the third-order cumulant and above is very complicated. Thus, the zero-mean processing is required to simplify the HOCs. This is particularly true for this case, as by no means the average of received RF signal is zero. Let  $\{x(t)\}$  be a zero-mean  $k$ th-order stationary random process. The  $k$ th-order cumulant of this process is denoted as  $C_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1})$ , which is defined for the joint  $k$ th-order cumulant of the random variable  $x(t), x(t+\tau_1), \dots, x(t+\tau_{k-1})$ , i.e.,

$$C_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \text{cum}(x(t), x(t+\tau_1), \dots, x(t+\tau_{k-1})) \quad (3)$$

where  $\tau_1, \tau_2, \dots, \tau_{k-1}$  are time delays. Following from (4) and (6), the corresponding second-, third- and fourth-order cumulants of zero-mean  $x(t)$  can be expressed as:

$$C_{2x}(\tau) = E\{x(t)x(t+\tau)\} \quad (4)$$

$$C_{3x}(\tau) = E\{x(t)x(t+\tau_1)x(t+\tau_2)\} \quad (5)$$

$$\begin{aligned} C_{4x}(\tau) &= E\{x(t)x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)\} \\ &\quad - C_{2x}(\tau_1)C_{2x}(\tau_2-\tau_3) - C_{2x}(\tau_2)C_{2x}(\tau_3-\tau_1) \\ &\quad - C_{2x}(\tau_3)C_{2x}(\tau_1-\tau_2). \end{aligned} \quad (6)$$

For a zero-mean Gaussian random process, the cumulants have the following conclusions:

$$C_{1x} = 0, \quad C_{2x} = \sigma^2, \quad C_{kx} = 0 \quad (k \geq 3). \quad (7)$$

As indicated in (7), if a signal contains additive Gaussian noise, including colored or white noise, the influence of noise can be minimized by using higher order cumulants, which proves the fact that the HOC-based approach is inherently suitable for this case. However, this requires excessive number crunching. In order to extract useful information from higher order cumulants of non-Gaussian stationary processes, 1-D slices of multidimensional cumulants are often used. This is mainly due to the fact that a 1-D slice of a higher order cumulant retains all the useful characteristics of the signal. Therefore, the computational efficiency can be significantly improved by using 1-D slices. The 1-D slice of a third-order cumulant  $C_{3x}(\tau_1, \tau_2)$  can be defined as

$$C_{3x}(0, \tau) = \text{cum}(x(t)x(t)x(t + \tau)) \quad (8)$$

and

$$C_{3x}(\tau, \tau) = \text{cum}(x(t)x(t + \tau)x(t + \tau)). \quad (9)$$

If a random process is symmetrically distributed, its third-order cumulant is zero, whereas a nonsymmetrical process has an extremely small third-order cumulant but a much larger fourth-order cumulant. As a result, in this paper, the fourth-order cumulants are chosen for feature extraction due to the fact that they should be much more distinctive than their third-order counterparts. The 1-D slices of a fourth-order cumulant  $C_{4x}(\tau_1, \tau_2, \tau_3)$  can be defined as follows:

$$C_{4x}(0, 0, \tau) = \text{cum}(x(t)x(t)x(t)x(t + \tau)) \quad (10)$$

$$C_{4x}(0, \tau, \tau) = \text{cum}(x(t)x(t)x(t + \tau)x(t + \tau)) \quad (11)$$

$$C_{4x}(\tau, \tau, \tau) = \text{cum}(x(t)x(t + \tau)x(t + \tau)x(t + \tau)). \quad (12)$$

It is noted that although different types of 1-D slices of fourth-order cumulants are possible, as shown in (10)–(12), including radial, vertical, horizontal, diagonal, and offset-diagonal, the vertical slice is selected in this paper mainly, because it can provide a simple but effective way to extract important information from the received signals, which improves the accuracy of target classification without a significantly increased computational cost. Furthermore, using the platform presented in Fig. 2, only the signals that have been sampled can be extracted at the output, and it is impossible to obtain the true values of the HOCs. Consequently, the true values from a finite set of measured data samples must be estimated. It can be assumed that the data samples,  $x(1), x(2), \dots, x(n)$ , are from a true, zero-mean and stationary process. The estimated values of the fourth-order cumulants using 1-D vertical slices  $\hat{c}_{4x}(0, 0, \tau)$  can be expressed as

$$\begin{aligned} \hat{c}_{4x}(0, 0, \tau) &= \frac{1}{N} \sum_{n=1}^N [x(n) - \hat{m}_{1x}]^3 [x(n + \tau) - \hat{m}_{1x}] \\ &\quad - \frac{3}{N^2} \sum_{n=1}^N [x(n) - \hat{m}_{1x}] [x(n + \tau) - \hat{m}_{1x}] \\ &\quad \times \sum_{n=1}^N [x(n) - \hat{m}_{1x}]^2 \end{aligned} \quad (13)$$

where

$$\hat{m}_{1x} = \frac{1}{N} \sum_{n=1}^N x(n).$$

For a more detailed overview of HOC-based methods, readers can refer to [11], [21], and the references therein. To illustrate the effectiveness of using the HOC-based method for feature extraction, the two scenarios that are previously used for waveform generation in Figs. 4 and 5 are used again, in Figs. 6 and 7. As illustrated, there are 180 extracted features used in this paper.

#### IV. FPA-SVM CLASSIFIER

##### A. Classical Support Vector Machine

In order to classify the type of placed target, the extracted features need to be separated into different classes. The SVM is a discrete algorithm that can be used to find the maximum margin between different classes of data [22]. Although the SVM can be used to separate multiple classes, for ease of understanding, in this section, the discussion is limited to two-class problems [22]. If two classes are linearly separable, using the following function, an optimal hyperplane can be found to separate them:

$$f(x) = \langle w \cdot x \rangle + b = 0 \quad (14)$$

where  $w$ ,  $x$ , and  $b$  denote the weight vector, input vector, and bias term, respectively. The terms  $w$  and  $b$  are used to define the location of the optimal hyperplane, which can be found by solving the following constrained optimization problem:

$$\begin{aligned} \text{Min } & \frac{1}{2} \|w\|^2 \\ \text{s.t. } & y_i \langle w \cdot x_i + b \rangle \geq 1, \quad i = 1, 2, \dots, N \end{aligned} \quad (15)$$

where  $y_i \in \{-1, +1\}$  is the set of class labels, if  $\{x_i, y_i\}_{i=1}^N$  is the set of training data sets. However, the classes involved in most practical problems are not separable. In this case, the SVM classifier is obtained by solving the following optimization problem:

$$\begin{aligned} \text{Min } & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t. } & \begin{cases} y_i \langle w \cdot x_i + b \rangle \geq 1 - \xi_i, & i = 1, 2, \dots, N \\ \xi_i \geq 0, & i = 1, 2, \dots, N \end{cases} \end{aligned} \quad (16)$$

where  $\xi_i$  are the slack variables that allow the SVM to tolerate misclassifications and  $C$  is the penalization parameter, which is used to control the tradeoff between minimized training errors and model complexity. Using the Lagrangian multipliers algorithm, the aforementioned optimization problem can be transformed into a dual quadratic optimization problem, that is

$$\begin{aligned} \text{Max } & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ \text{s.t. } & \sum_{i=1}^N \alpha_i y_i = 0, \quad C \geq \alpha_i \geq 0, \quad i = 1, 2, \dots, N. \end{aligned} \quad (17)$$

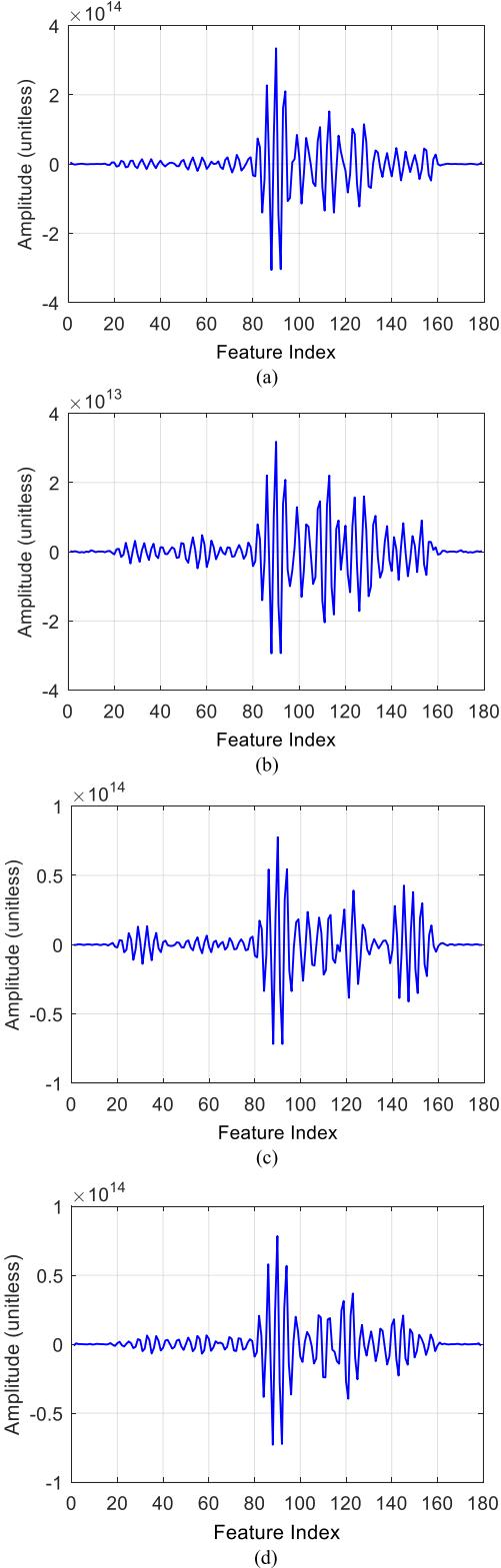


Fig. 6. Extracted features using the presented approach under snowy weather.  
(a) No target. (b) Target type I. (c) Target type II. (d) Target type III.

By solving the dual optimization problem, the linear decision function can be created as

$$f(x) = \text{sign} \left( \sum_{i=1}^N \alpha_i y_i (x_i, x) + b \right). \quad (18)$$

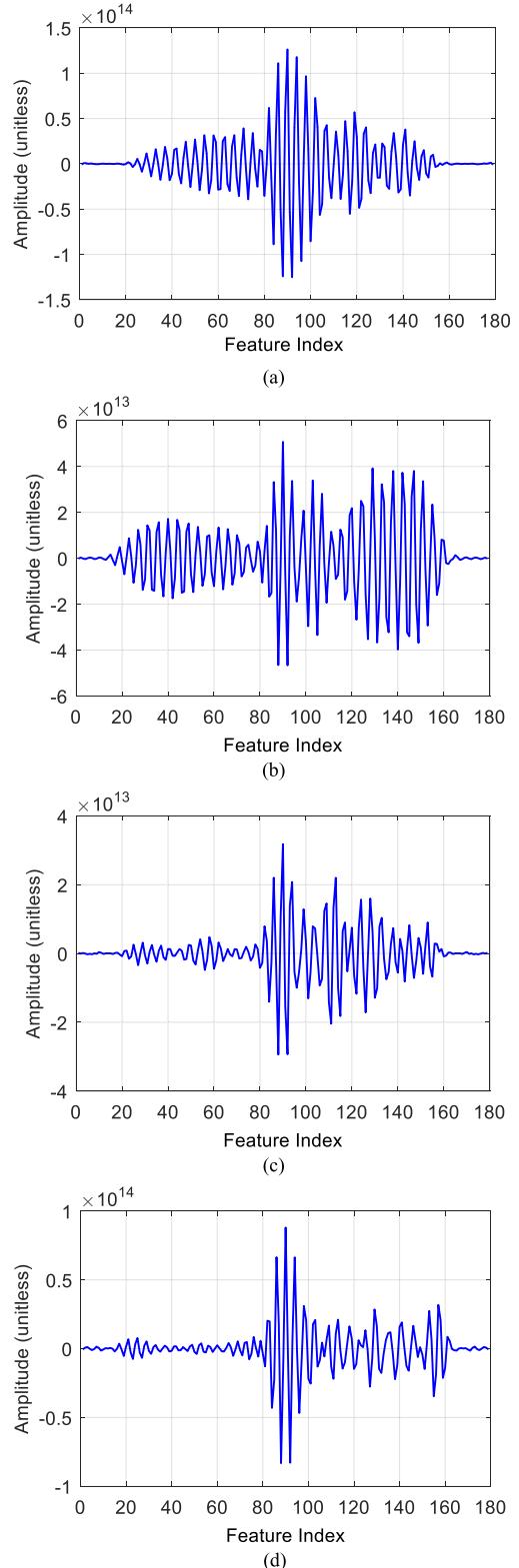


Fig. 7. Extracted features of type I target with different weather conditions.  
(a) Sunny. (b) Rainy. (c) Snowy. (d) Foggy.

In order to use an SVM in a nonlinear classification, the kernel function  $K(x_i, x)$  needs to be determined. The value of  $K(x_i, x)$  equals  $\varphi(x_i) \cdot \varphi(x)$ , where  $\varphi(\cdot)$  is the transformation function, which takes the input data into high-dimensional

feature space. Then, the nonlinear decision function is described as follows:

$$f(x) = \text{sign} \left( \sum_{i=1}^N \alpha_i y_i K(x_i, x) + b \right). \quad (19)$$

In this paper, the radial basis function (RBF) kernel function  $K(x_i, x) = \exp(-\|x_i - x\|^2/2\sigma^2)$  is selected to construct the SVM, where  $\sigma$  denotes the width of the RBF kernel function. Note that both the width of the RBF kernel function  $\sigma$  and the penalization parameter  $C$  are user-determined parameters [11]. As previously described, the SVM is an intrinsically binary classifier. However, in practice, most cases are multiclass. Several unique approaches have been presented in the literature to solve multiclass problems using a binary classifier [23]–[25].

For an  $N$ -class classification problem, multiple SVMs can be used to form a binary-tree structure (BTS), so that multiclass decisions can be made with only  $N-1$  SVMs. As a result, the training consumption is kept relatively low. For this reason, a BTS-based approach is adopted to construct the classification model in this paper.

### B. Flower Pollination Algorithm

Pollination is a natural phenomenon that plays an extremely significant role in the sexual reproduction of plants. Inspired by this phenomenon, the FPA was first presented in [26] and [27]. A distinct advantage of this algorithm is that it very well balances the defined goals between local and global optimization. The behavior of pollination can be summarized using the following four rules.

**Rule 1:** Biotic cross-pollination can be considered as a process of global pollination, and pollen-carrying pollinators move in a way that obeys Lévy flights.

**Rule 2:** Abiotic and self-pollination can be considered as a process of local pollination.

**Rule 3:** Pollinators, such as insects, can develop flower constancy, which is equivalent to a reproduction probability that is proportional to the similarity of the two flowers involved.

**Rule 4:** The interaction or switching between local and global pollination can be controlled by a switch probability  $p \in [0, 1]$ , with a slight bias toward local pollination.

To model these rules, the following equations are developed. In the step of global pollination, gametes of flowers can travel over a long distance, if they are carried biotically. Therefore, Rules 1 and 3 can be represented mathematically as

$$x_i^{t+1} = x_i^t + \gamma L(\lambda)(g_* - x_i^t) \quad (20)$$

where  $x_i^t$  is the pollen  $i$  or solution vector  $x_i$  at generation  $t$  and  $g_*$  is the current best solution found among all solutions for the current generation. Here,  $\gamma$  is a scaling factor to control step size. In addition,  $L(\lambda)$  is the parameter that corresponds to the strength of the pollination, which also essentially affects the step size. Since insects may move over a long distance with various distance steps, a Lévy flight can be used to mimic this characteristic efficiently. That is, we draw  $L > 0$  from a Lévy distribution

$$L \sim \frac{\lambda \Gamma(\lambda) \sin(\pi \lambda / 2)}{\pi} \frac{1}{s^{1+\lambda}}, \quad (s \geq s_0 > 0) \quad (21)$$

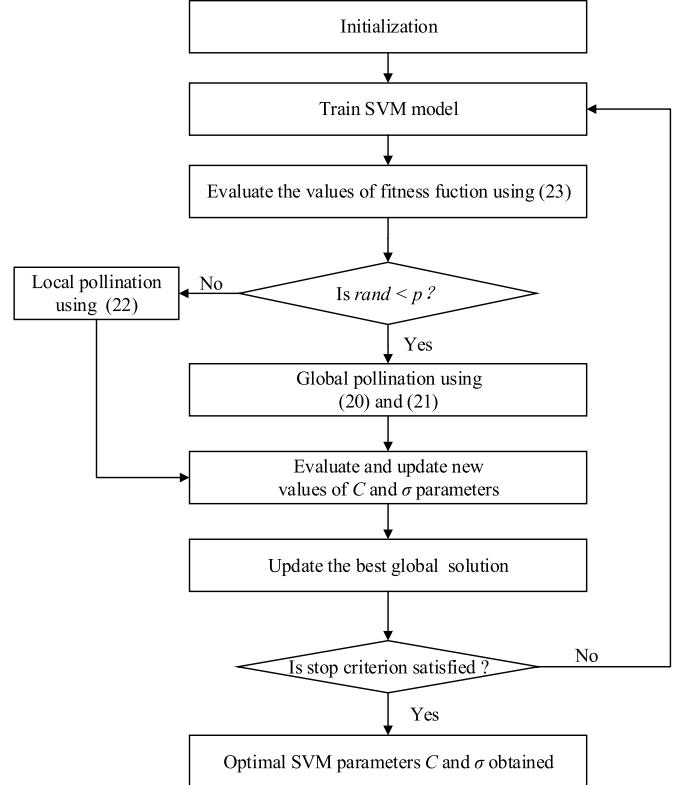


Fig. 8. Flowchart of parameters optimization using the presented FPA-SVM.

where  $\Gamma(\lambda)$  is the standard gamma function. Its distribution is valid for large steps  $s > 0$ . In order to model local pollination, both Rules 2 and 3 can be represented as

$$x_i^{t+1} = x_i^t + \epsilon (x_j^t - x_k^t) \quad (22)$$

where  $x_j^t$  and  $x_k^t$  are the pollen of the same plant species but from different flowers, which represent flower constancy with local pollination. Mathematically, if  $x_j^t$  and  $x_k^t$  come from the same species or are selected from the same population, this would become a local random pollination if  $\epsilon$  is drawn from a uniform distribution in  $[0, 1]$ . Though pollination activities can occur at all scales, both local and global, adjacent flower patches or nearby flowers are more likely to be pollinated by local flower pollen than those far away. In order to mimic this, Rule 4 or the proximity probability  $p$  can be used to switch between common global pollination and intensive local pollination.

### C. Parameters Optimization Using FPA-SVM

In this paper, the RBF kernel is used to construct the SVM classifier. Since the selection of some critical parameters could have a great influence on the performance of SVM, to find an optimized solution for these parameters in an effective way, the previously discussed FPA-based algorithm is used, because it is able to deal with an optimization issue on a large scale with relatively fast convergence [28]–[30]. The basic steps used to determine the parameters of the SVM by means of FPA are explained in the following, and the flowchart of this algorithm is shown in Fig. 8.

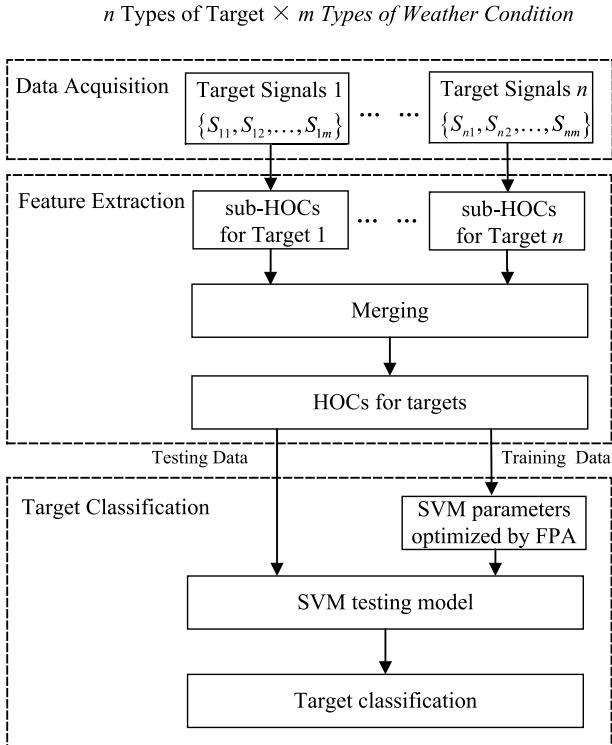


Fig. 9. Presented classification model.

**Step 1 (Initialization):** In this step, a population of pollen with random solutions ( $C$  and  $\sigma$ -parameters for SVM) is initialized first. Then, original values for other important parameters, such as the size of the initialized population, switch probability, and maximum generation number  $G_{\max}$ , are defined.

**Step 2 (Evaluation of Fitness Function):** The fitness function is used to evaluate the quality of every solution. It is based on the classification error of an SVM classifier, which is defined as follows:

$$\text{fitness} = 1 - \frac{N_t}{N_t + N_f} \quad (23)$$

where  $N_t$  and  $N_f$  note the number of true and false classifications, respectively.

**Step 3 (Global Versus Local Optimization):** Suppose  $\text{rand} \in [0, 1]$  is a random number for each solution  $x_i$  and  $p$  is the switch probability. If ( $\text{rand} < p$ ), a new solution needs to be determined using (20) and (21) for global optimization; otherwise, the local optimization needs to be conducted using (22). It is noted that each new solution generated in this step must be evaluated, and if their fitness values are better than the initial ones, the results must be updated.

**Step 4 (Iteration):** Step 3 keeps executing until the predetermined maximum number of generations  $G_{\max}$  is reached. Finally, the global optimal values of  $C$  and  $\sigma$  can be obtained for the SVM classifier.

#### D. Target Classification Using the Presented Model

The different types of targets can be treated as different classification labels. The presented model used in this paper is shown in Fig. 9.

As can be seen, this model is divided into three stages: data acquisition, feature extraction, and parameters optimization plus target classification. The design flow of this model is briefly given in this section. If  $n$  types of target need to be classified under  $m$  types of weather condition, the size of the overall scenarios needing to be considered is  $n \times m$ . In our case, 16 different scenarios need to be taken into consideration. In the data acquisition stage, identical numbers of data samples are collected for each scenario. Based on these collected samples, the distinctive features from each scenario are extracted using the HOC-based algorithm. Once sufficient features are gained for each scenario, these features are then merged according to form databases that can be used for target classification. As four types of targets need to be classified, four databases are required. It is noted that an identical number of features are used for each merged database. In the target classification stage, using the randomly selected training data from the merged databases, the parameters of the SVM classifier are optimized by means of the presented FPA. Once the optimization of these parameters is accomplished, the classifier should be able to recognize different types of target using the previously created test data. As will be shown in Section V, using this model, the type of the detected target can be classified with a good accuracy even under different weather conditions.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Confusion Matrix of the Presented Model

The data obtained in Section II are divided into two categories: training and testing data sets, in which the training data sets are used to calculate the fitness function and train the FPA-SVM classifier, whereas the testing data sets are mainly used to evaluate the performance of the classification model. Under every weather condition, from the created data set, 120 data samples are randomly selected for each target type, which indicates that a total of 1920 data samples are used, as four types of target need to be classified under four different weather conditions. It is worthwhile to mention that, once data samples are used for training, they are then removed from the overall database. Consequently, it is ensured that no similar data samples are used for both training and testing purposes, which makes the evaluation results more meaningful. To form the required testing data set, 160 data samples are randomly selected from each target type for the same weather condition. Consequently, overall 640 data samples are used to test each target type with different weather conditions. According to the preliminary parametric studies, the following values are used for the key parameters of FPA  $p = 0.8$  and  $\lambda = 1.5$  and a scaling factor  $\gamma = 0.1$ . In addition, the size of population is 20, and the number of iterations is set to be 200. All experimentally gained data samples are adopted to verify the performance of the presented approach in terms of classification accuracy. The confusion matrix of this approach is summarized in Table I.

### B. Performance Comparison With Other Approaches

In order to demonstrate the performance improvement by using the FPA-SVM classifier, other classical classifiers

TABLE I  
CONFUSION MATRIX OF THE FPA-SVM APPROACH

Target type of samples	Test results			
	No target	Target type I	Target type II	Target type III
No target	623	11	2	4
Target type I	18	583	26	13
Target type II	5	23	591	21
Target type III	4	8	24	604

TABLE II  
SUMMARY OF CLASSIFICATION ACCURACY (%) USING DIFFERENT CLASSIFIERS

Target type	Classification rate (%)			
	FPA-SVM	SVM	BPNN	KNN
No target	97.34	96.72	96.25	95.16
Target type I	91.09	86.56	81.09	80.78
Target type II	92.34	88.44	87.03	86.41
Target type III	94.37	90.31	88.91	87.34
Average	93.78	90.51	88.32	87.42

are also used as benchmarks, including the conventional SVM, the backpropagation neural networks (BPNNs), and the  $k$ -nearest neighbor (KNN) models. The comparison results are shown in Table II. For the KNN classifier, the Euclidean distance is used as a similarity measure, and the number of neighbors is selected to be 10. For the BPNN classifier, three layers are applied to the target classification. The number of input nodes is determined according to the dimensions of the features. The output layer contains four nodes, which is the number of target types in this paper. The optimal parameters of the nodes for the hidden layer, the target error, the training rate, and the momentum were determined to be 25, 0.005, 0.1, and 0.01, respectively. For the conventional SVM classifier, the selected parameters are given as follows:  $C = 16.4218$  and  $\sigma = 1.6026$ . As can be seen, the presented approach outperforms all others in target classification in a foliage environment, even with different weather conditions.

To further investigate the performance improvement from using the HOC-based algorithm for feature extraction, the conventional SP-based algorithm is used as a benchmark. For the purpose of comparison, seven SPs are used, as listed in Table III. The comparison results are shown in Table IV. As can be seen, using either the SVM or FPA-SVM classifier, it is evident that the HOC-based algorithm offers an inherently improved accuracy over the SP-based one. On the other hand, using either the HOC- or SP-based algorithm for feature extraction, any approach that adopts the FPA-SVM classifier always provides a better performance than its SVM-based counterpart. It therefore once again proves that the presented approach is superior to the conventional ones.

### C. Reliability Analysis

To further evaluate the performance of the presented approach in terms of reliability, the impact on accuracy of different sizes of training data set is investigated. The number of training samples for each target type is varied from 160 to 480 with a step of 40. The obtained results are shown in Fig. 10. As illustrated, using the FPA-SVM classifier,

TABLE III  
SEVEN TYPICAL SP-BASED PARAMETERS

Energy of $r(t)$ : $E_r = \int_{-\infty}^{+\infty}  r(t) ^2 dt$
Maximum amplitude of $r(t)$ : $A_{\max} = \max_t  r(t) $
Mean excess delay: $\tau_m = \int_{-\infty}^{+\infty} t  r(t) ^2 dt / \int_{-\infty}^{+\infty}  r(t) ^2 dt$
RMS delay spread: $\tau_{RMS} = \sqrt{\int_{-\infty}^{+\infty} (t - \tau_m)^2  r(t) ^2 dt / \int_{-\infty}^{+\infty}  r(t) ^2 dt}$
Mean of $ r(t) $ : $\mu_{ r } = \frac{1}{T} \int_T  r(t)  dt$
Standard deviation of $ r(t) $ : $\sigma_{ r }^2 = \frac{1}{T} \int_T [  r(t)  - \mu_{ r } ]^2 dt$
Kurtosis of $ r(t) $ : $k = \frac{1}{\sigma_{ r }^4 T} \int_T [  r(t)  - \mu_{ r } ]^4 dt$

Note:  $r(t)$  is the target waveform, and the resolution of the scan waveform is 61 ps.

TABLE IV  
SUMMARY OF CLASSIFICATION ACCURACY (%) USING DIFFERENT FEATURE EXTRACTION METHODS

Target type	Classification rate (%)			
	FPA-SVM		SVM	
	HOC	SP	HOC	SP
No target	97.34	91.41	96.72	90.31
Target type I	91.09	69.84	86.56	64.22
Target type II	92.34	75.16	88.44	72.50
Target type III	94.37	75.47	90.31	63.75
Average	93.78	77.97	90.51	72.69

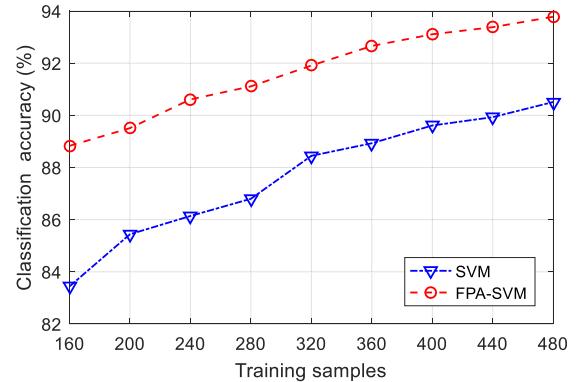


Fig. 10. Impact on average classification accuracy due to different numbers of training samples.

the classification accuracy is inherently improved compared with the classical SVM-based one, even when the size of the training data set is reduced.

As RF signals can be easily attenuated due to environmental variations, evaluating the impact of SNR variation on classification accuracy is also very critical. To do so, additional Gaussian white noise is added to the original collected data so that the SNR can be effectively varied. In the simulation, the noise level is varied from 5 to 30 dB with a step of 5 dB. The average accuracy obtained from these tests is shown in Fig. 11. As illustrated, the presented approach has not only higher classification accuracy but also

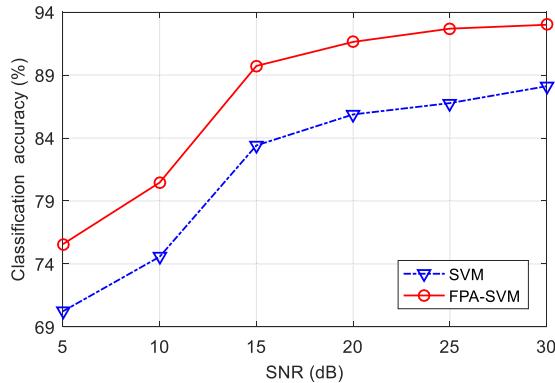


Fig. 11. Impact on average classification accuracy due to SNR variations.

excellent immunity against noise, compared with the classical SVM-based classifier.

## VI. CONCLUSION

In this paper, the performance analysis of using the presented WSN-based DFS approach for FOPEN in all seasons is investigated. To ensure a reliable performance in all seasons with different weather conditions in terms of classification accuracy, an FPA-SVM-based classifier is used in conjunction with HOC-based feature extraction. As a result, the adverse effect on accuracy due to seasonal variations, such as different weather conditions, can be mitigated. To demonstrate the feasibility of using this approach for FOPEN, a variety of data samples have been taken under 16 different scenarios, including four different types of target and four different weather conditions, in a foliage environment. Using these data samples, the robustness of the prototype is extensively verified, and the impact on classification accuracy due to the number of training samples and SNR variations is analyzed. Without using any sensing devices other than RF transceivers, the average accuracy is greater than 94%, while the conventional SP-based approach can only achieve an accuracy of 73% under the same scenario. Therefore, we can firmly conclude that the presented approach is feasible for FOPEN-related applications in all seasons, such as intelligent border control and smart search-and-rescue.

## REFERENCES

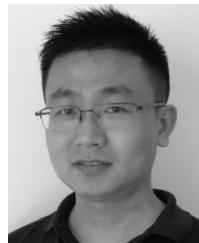
- [1] T.-K. Chan, Y. Kuga, and A. Ishimaru, "Experimental studies on circular SAR imaging in clutter using angular correlation function technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2192–2197, Sep. 1999.
- [2] P.-O. Frolind, A. Gustavsson, M. Lundberg, and L. M. H. Ulander, "Circular-aperture VHF-band synthetic aperture radar for detection of vehicles in forest concealment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1329–1339, Apr. 2012.
- [3] X. Jin, S. Sarkar, A. Ray, S. Gupta, and T. Damarla, "Target detection and classification using seismic and PIR sensors," *IEEE Sensors J.*, vol. 12, no. 6, pp. 1709–1718, Jun. 2012.
- [4] Q. Guan, X. Yin, X. Guo, and G. Wang, "A novel infrared motion sensing system for compressive classification of physical activity," *IEEE Sens. J.*, vol. 16, no. 8, pp. 2251–2259, Apr. 2016.
- [5] T. Damarla, M. Bradley, A. Mehmood, and J. M. Sabatier, "Classification of animals and people ultrasonic signatures," *IEEE Sensors J.*, vol. 13, no. 5, pp. 1464–1472, May 2013.
- [6] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 612–621, May 2011.
- [7] N. Patwari and J. Wilson, "RF sensor networks for device-free localization: Measurements, models, and algorithms," *Proc. IEEE*, vol. 98, no. 11, pp. 1961–1973, Nov. 2010.
- [8] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl, "RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals," *IEEE Trans. Mobile Comput.*, vol. 13, no. 4, pp. 907–920, Apr. 2014.
- [9] Q. Liang, "Radar sensor wireless channel modeling in foliage environment: UWB versus narrowband," *IEEE Sensors J.*, vol. 11, no. 6, pp. 1448–1457, Jun. 2011.
- [10] B. Li, C. Zhao, H. Zhang, X. Sun, and Z. Zhou, "Characterization on clustered propagations of UWB sensors in vehicle cabin: Measurement, modeling and evaluation," *IEEE Sensors J.*, vol. 13, no. 4, pp. 1288–1300, Apr. 2013.
- [11] Y. Zhong, Y. Yang, X. Zhu, E. Dutkiewicz, Z. Zhou, and T. Jiang, "Device-free sensing for personnel detection in a foliage environment," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 6, pp. 921–925, Jun. 2017.
- [12] Y. Tang, Y.-Q. Zhang, N. V. Chawla, and S. Krasser, "SVMs modeling for highly imbalanced classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 1, pp. 281–288, Feb. 2009.
- [13] J. Muñoz-Marí, F. Bovolo, L. Gomez-Chova, L. Bruzzone, and G. Camp-Valls, "Semisupervised one-class support vector machines for classification of remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3188–3197, Aug. 2010.
- [14] S. Bahrampour, A. Ray, S. Sarkar, T. Damarla, and N. M. Nasrabadi, "Performance comparison of feature extraction algorithms for target detection and classification," *Pattern Recognit. Lett.*, vol. 34, no. 16, pp. 2126–2134, Dec. 2013.
- [15] Y. Geng, J. Chen, R. Fu, G. Bao, and K. Pahlavan, "Enlighten wearable physiological monitoring systems: On-body RF characteristics based human motion classification using a support vector machine," *IEEE Trans. Mobile Comput.*, vol. 15, no. 3, pp. 656–671, Mar. 2016.
- [16] V. Chandran and S. L. Elgar, "Mean and variance of estimates of the bispectrum of a harmonic random process—An analysis including leakage effects," *IEEE Trans. Signal Process.*, vol. 39, no. 12, pp. 2640–2651, Dec. 1991.
- [17] V. Chandran, B. Carswell, B. Boashash, and S. Elgar, "Pattern recognition using invariants defined from higher order spectra: 2-D image inputs," *IEEE Trans. Image Process.*, vol. 6, no. 5, pp. 703–712, May 1997.
- [18] M. E. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Trans. Energy Convers.*, vol. 14, no. 4, pp. 1065–1074, Apr. 1999.
- [19] L. L. Khadra, A. S. Al-Fahoum, and S. Binajjaj, "A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1840–1850, Nov. 2005.
- [20] J. Muthuswamy, D. L. Sherman, and N. V. Thakor, "Higher-order spectral analysis of burst patterns in EEG," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 1, pp. 92–99, Jan. 1999.
- [21] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.
- [22] S. Abe, *Support Vector Machines for Pattern Classification*. Springer, 2010.
- [23] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
- [24] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.
- [25] G. Madzarov, D. Gjorgjevikj, and I. Chorbev, "A multi-class SVM classifier utilizing binary decision tree," *Int. J. Comput. Inform.*, vol. 33, no. 2, pp. 225–233, Jan. 2009.
- [26] X.-S. Yang, M. Karamanoglu, and X. He, "Flower pollination algorithm: A novel approach for multiobjective optimization," *Eng. Optim.*, vol. 46, no. 9, pp. 1222–1237, Apr. 2013.
- [27] X.-S. Yang, M. Karamanoglu, and X. He, "Multi-objective flower algorithm for optimization," *Proc. Comput. Sci.*, vol. 18, pp. 861–868, Jan. 2013.
- [28] E. Nabil, "A modified flower pollination algorithm for global optimization," *Expert Syst. Appl.*, vol. 57, pp. 192–203, Sep. 2016.

- [29] A. Y. Abdelaziz, E. S. Ali, and S. A. Elazim, "Flower pollination algorithm and loss sensitivity factors for optimal sizing and placement of capacitors in radial distribution systems," *Int. J. Elect. Power Energy Syst.*, vol. 78, pp. 207–214, Jun. 2016.
- [30] S. A.-F. Sayed, E. Nabil, and A. Badr, "A binary clonal flower pollination algorithm for feature selection," *Pattern Recognit. Lett.*, vol. 77, pp. 21–27, Jul. 2016.



**Yi Zhong** is currently pursuing the dual-program Ph.D. degree in communication and information system with the Beijing University of Posts and Telecommunications, Beijing, China, and the University of Technology Sydney, Sydney, NSW, Australia.

Her research interests include transistor-level radio-frequency integrated circuit design, system-level implementation for device-free sensing network, and sense-through-foliage target detection and recognition.



**Yang Yang** (S'11–M'14) was born in Inner Mongolia, China. He received the Ph.D. degree from Monash University, Melbourne, VIC, Australia, in 2013.

From 2012 to 2015, he was an Asia Pacific GSP Engineer with Rain Bird, Azusa, CA, USA. From 2015 to 2016, he served as a Senior Research Associate with the Department of Engineering, Macquarie University, Sydney, NSW, Australia. In 2016, he was a Research Fellow with the State Key Laboratory of Millimeter-Waves, City University of Hong Kong, Hong Kong. In 2016, he was involved in the National Basic Research Program of China (973 Program) and appointed as an Honorary Research Fellow with the Shenzhen Institute, City University of Hong Kong. In 2016, he joined the University of Technology Sydney, Sydney, as a Lecturer. His research interests include radio-frequency integrated circuit, microwave and millimeter-wave circuits and systems, reconfigurable antennas, wearable antennas, and wearable medical sensing devices and systems.

Dr. Yang received the Global GSP Success Award in 2014.



**Xi Zhu** received the B.E. degree (Hons.) and the Ph.D. degree from the University of Hertfordshire, Hertfordshire, U.K., in 2005 and 2008, respectively.

He is currently a Lecturer with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW, Australia. His research interests include analog baseband, radio frequency, and mm-wave circuits and systems design. He has co-authored over 60 refereed publications in the above-mentioned fields.



**Yan Huang** received the M.S. degree from the School of Computer Science and Engineering, Beihang University, Beijing, China. He is currently pursuing the Ph.D. degree with the Global Big Data Technology Center, University of Technology Sydney, Sydney, NSW, Australia.

His research interests include machine learning, deep learning, and pattern recognition.



**Eryk Dutkiewicz** (M'05) received the B.E. degree in electrical and electronic engineering and the M.Sc. degree in applied mathematics from The University of Adelaide, Adelaide, SA, Australia, in 1988 and 1992, respectively, and the Ph.D. degree in telecommunications from the University of Wollongong, Wollongong, NSW, Australia, in 1996.

His industry experiences include the management of the Wireless Research Laboratory, Motorola, Sydney, NSW, Australia, in 2000. He has held visiting professorial appointments at several institutions, including the Chinese Academy of Sciences, Beijing, China, Shanghai Jiao Tong University, Shanghai, China, and Macquarie University, Sydney. He is currently the Head of the School of Electrical and Data Engineering, University of Technology Sydney, Sydney. His research interests include 5G networks and medical body area networks.



**Zheng Zhou** (M'08) received the master's and Ph.D. degrees in electrical engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1982 and 1988, respectively.

From 1993 to 1995, he was a Visiting Research Fellow with The Chinese University of Hong Kong, Hong Kong, supported by the Hong Kong Telecom International Postdoctoral Fellowship. In 2000, he visited the Japan Kyocera DDI Research Institute, Tokyo, Japan, as an Invited Overseas Researcher supported by the Japan Key Technology Research Center. He is currently with the School of Information and Communication Engineering, BUPT. His research interests include short-range wireless technology, UWB wireless communication, and cognitive radio.

Dr. Zhou was the Committee Member of Technical Committee on Cognitive Networks IEEE ComSoc and the Steering Committee Chair of International Symposium on Communications and Information Technologies. He is a Senior Member of the China Institution of Communications (CIC) and a Radio Application and Management Technical Committee Member of CIC. He is also a Senior Member of China Computer Federation (CCF) and a Sensor Network Technical Committee Member of CCF.



**Ting Jiang** received the B.S., M.S., and Ph.D. degrees in communication and information system from Yanshan University, Qinhuangdao, China, in 1982, 1988, and 2003, respectively.

He is currently a Professor with the Beijing University of Posts and Telecommunications, Beijing, China. His research interests include the wireless broadband interconnection, the information theory, the short distance wireless communication technological theory and application, and the wireless sensor network.