Deep factor analysis for weather varied sense-through-foliage target detection

Wenling Xue[1,2], Ting Jiang[1]* , Xuebin Sun[1], Xiaokun Zheng[1,2] and Xue Ding[1]

*Correspondence:
tjiang@bupt.edu.cn
[1] Key Laboratory of Universal
Wireless Communication,
BUPT, Xitucheng Road 10,
Beijing 100876, China
Full list of author information
is available at the end of the
article

Abstract

In this paper, the influence of seasonal variation on target detection accuracy and the effectiveness of deep factor analysis (DFA) in signal denoising are studied. To extensively verify the universality of the DFA_based approach, a variety of target objects, including no target, human, wood board and iron cabinet targets, are measured in foliage environment under four different weather conditions. Then, after removing background noise from the collected data, deep factor analysis is carried out to reduce the impact of noise. The experimental results show that the influence of weather variation on target detection can be effectively eliminated by DFA_based algorithm, which can improve the average classification accuracy in all seasons. Finally, by means of cross validation, the effectiveness of DFA_based algorithm on signal denoising and the influence on target detection accuracy are further studied. The method is stable and universal in any weather conditions, even in hazy and snowy days, which can be stable at about 93%.

Keywords: Impulse-radio ultra-wideband (IR-UWB), Deep factor analysis, Target detection

1 Introduction

In recent years, the detection of hidden objects, such as human beings, animals and vehicles in foliage environment has attracted enormous attention. However, due to the special conditions of foliage environment, the traditional detection technology has revealed the shortcomings. For example, optical or image-based detection technology is difficult to meet the accuracy requirements of detection and classification because of the influence of light and occlusion in the foliage environment. The cost of the synthetic-aperture radar (SRA)-based system is relatively high which greatly reduces the feasibility of application. At present, a new type of wireless sensor technology, device free sensing(DFS), has been proved to be feasible in target detection in many cases [1, 2]. This method combines communication with target detection, which greatly reduces the cost of sensor network. The commonly used DFS technologies for target recognition are based on UWB signal and WiFi signal. However, the most widely used scene of WiFi signal is indoor, such as indoor positioning [3, 4], human activity through the wall detection [5], occupancy estimation [6] and so on. Ultra-wideband (UWB) signals are universally worked for intrusion detection in sense-through-foliage environment due to their

Xue *et al. J Wireless Com Network*   (2021) 2021:41

Page 2 of 19

advantages such as high transmission rate, strong anti-interference ability and low cost as well as good target detection in foliage environment [7–10]. These application include not only static object recognition [11–13] but also moving target detection like human activity recognition [14]. Static target recognition includes targets with similar dielectric constants [2, 15, 16] and targets with large difference in in dielectric constant.

However, the actual foliage environment is complex and variable. Many practical problems need to be studied and solved. For example, when the target detection system is installed outdoors, it is necessary to ensure that any environmental change, such as rain, snow, lightning and other weather changes will not affect the detection accuracy of the system. Environmental noise and multipath fading will also have significant interference with the signal. Therefore, Xue [17, 18] and Zhong [19, 20] analyzed the problem of low target recognition rate caused by weather changes and proposed solutions to the problem. The solutions adopted in these literatures is feature extraction and classification recognition, mixing all signals collected in all weather conditions, and the collected data is divided into a training set and a test set, and the training set is trained to obtain an optimized model to achieve accurate target recognition. Although the recognition rate is higher, these methods are not universal. It is not applicable for the data set collected under any weather conditions, which has a large amount of calculations and strict requirements on the training set.

The focus of this study is how to automatically recognize targets without retraining the model to minimize the impact of weather changes on target recognition. A new approach to solve the problem of repeated training in practical application is adopted. The influence of different weather is used as clutter noise of the target object, and the depth factor analysis (DFA) method is used to remove the clutter noise, so as to eliminate the influence of different weather in the data preprocessing stage without any impact on the classifier in the system. According to the method in reference [21], we distinguish the signal with target from the signal without target. We take the received signal without target as the reference signal. Signals collected on the targeted scene are cross correlated with the reference signals to remove most of the clutter noise. After that, the DFA_based algorithm is used to remove the residual noise. Therefore, after the system is installed outdoors, only the data collected under various weather conditions need to preprocessed, and the trained model can be used for target detection and recognition without repeated training. The accuracy and stability are verified by experiments and DFA_based algorithm in this paper has universality and practicability.

The next section describes data acquisition and preprocessing. Section 3 provides the theory of DFA. Section 4 analysis the process of target detection and discusses the experimental results. Finally, the paper conclusions and future works are shown in Sect. 5.

2 Data acquisition and preprocessing

2.1 Data acquisition

In this study, an IR-UWB transceiver is used to acquire target signals in different weather conditions and data acquisition is carried out in the foliage cluster environment. PulsON 400 (P400) of Time Domain Co., Ltd is adopted for measurement. It's operating band is from 3.1 to 5.3 GHz with center frequency 4.2 GHz. We affixed the UWB transmitter and receiver to a bracket, approximately 1.5m from the ground and the transmitter

Xue *et al. J Wireless Com Network*   (2021) 2021:41

Page 3 of 19

and receiver are nearly 20m apart from each other. After receiving the data acquisition instruction sent by personal computer (PC), P400 begins to collect the data, and then sends the collected data to PC for recording. The schematic diagram of the measurement system is illustrated in Fig. 1. In order to acquire signals at any position between TX and RX, the same target object is placed in eight different positions marked as A, B, C, D, E, F, G and H. For example, people stand at 8 different positions as shown in Fig. 1, and we collect signals from these 8 different locations. 70 samples were collected from each location, and then the samples from 8 different locations were mixed together to form 560 samples as a data set for Human being.

The three types of targets employed in this study are (a) human being(a standing man with 65 kg weight and 175 cm height), (b) wood board (approximately 60 cm × 142 cm × 2.5 cm(width × height × thickness)) and (c) iron cabinet [approximately 50 cm × 40 cm × 130 cm (length × width × height)] [22]. The data sets used in the study are mainly collected under four weather conditions, as shown in Fig. 2, which are sunny, snowy, rainy and haze days. Figure 2a is a snapshot of sunny day, whereas Fig. 2b–d are that of snowy day, rainy day and hazy day. The measurement process at the other seven locations is a repetition of the above process.

For a target placed in a single position, the target signal is transmitted to a computer for digital recording, and 70 samples for each location are collected. Measurements are performed at 8 different locations to obtain 560 samples of a certain type of target. The same process is repeated for other target objects, and the same process is carried out for four types of weather. So, a total of 8,960 data samples under various weather conditions are obtained. The samples collected in this section will be used in the following target classification, which will be explained in detail later.

The main advantage of IR-UWB technology is its high spatial-temporal resolution. When objects with different materials (different conductivity and dielectric constant), sizes and shapes are placed in the foliage environment, the transmission path between the transmitter and the receiver varies depending on the absorption and reflection of
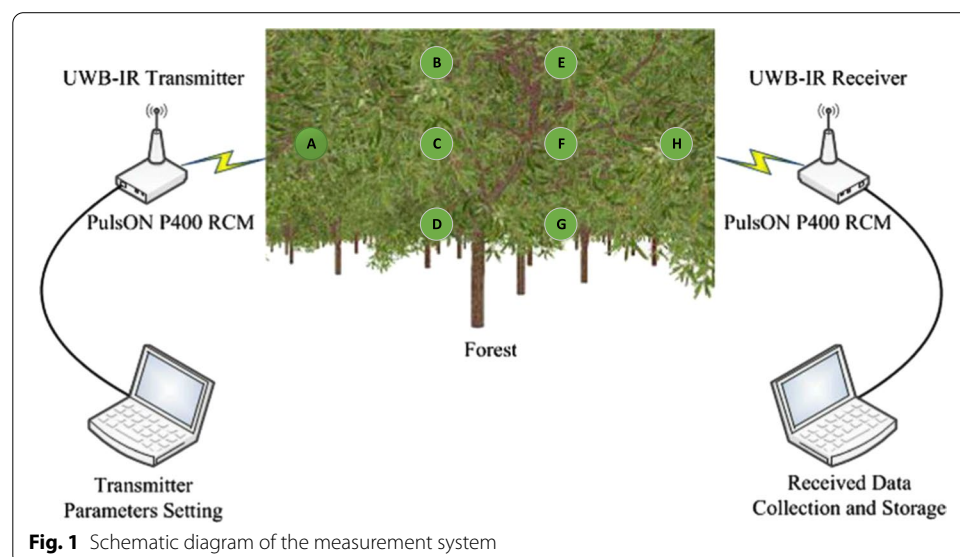

Fig. 1 Schematic diagram of the measurement system

Fig. 2 Actual test scenarios

objects, as well as the size of particle in the air and air humidity. And, the received UWB signal is a multi-path signal generated by different reflection and scattering in the foliage environment. The common characteristic of UWB channel in foliage environment is that the transmission attenuation of received UWB signal is time-varying. Therefore, even for the same target, the waveform of the received signal is different under different weather conditions. In this paper, three types of target object are human being, wood board and iron cabinet and the dielectric constants of them varies greatly due to their materials. Thus, the target object absorbs and reflects UWB signals differently in the foliage environment. The received signal $r(t)$ can be written as

Xue *et al. J Wireless Com Network*   (2021) 2021:41

Page 5 of 19

$$r(t) = \sum_{i=1}^{n} A_i s(t - t_i) + \epsilon(t) \tag{1}$$

where A_i and t_i are the multipath amplitude and the delay of the received signal passing over the i_{th} path respectively, A_i and t_i are random varying, $\epsilon(t)$ is stochastic noise, n is the number of scattering paths and $s(t)$ is the transmitted signal. The model can include multipath reflection from the target as well as direct reflection from the target. Through the above experimental device, the received UWB signal is shown in Fig. 3. According to the P400 documentation,the first part of the received UWB signal waveform is background noise, which is almost static. Useful information containing target object is in the second part of the signal waveform. We collected data under four weather conditions: sunny, rainy, snowy and haze days. Two kinds of scene signals are collected in each weather, which are no target signal and target signal. The signal with target is divided into three kinds: human, wood board and iron cabinet. The first portion of the received UWB signal is removed and its amplitude is reduced to 1/10,000 of the original magnitude.The processed received signal is illustrated in Figs. 4 and 5.

From Fig. 4, it is observed that when different target objects are placed between TX and RX, the received UWB signals vary in amplitude, which means that the received signals contain different target information, even in the same weather. The collected UWB signals of same target object (an example of Human target) in different weather conditions are illustrated in Fig. 5. (a) is in sunny and rainy weather conditions and (b) is in snowy and hazy weather conditions.

2.2 Data preprocessing

The purpose of data preprocessing is to lessen the influence of background noise. There are four scenarios for data acquisition in each weather condition. The first scenario is that no target is placed, and the other three are that three different objects are placed separately. According to the method in reference [21], we distinguish the signal with target from the signal without target. Here, we take the signal
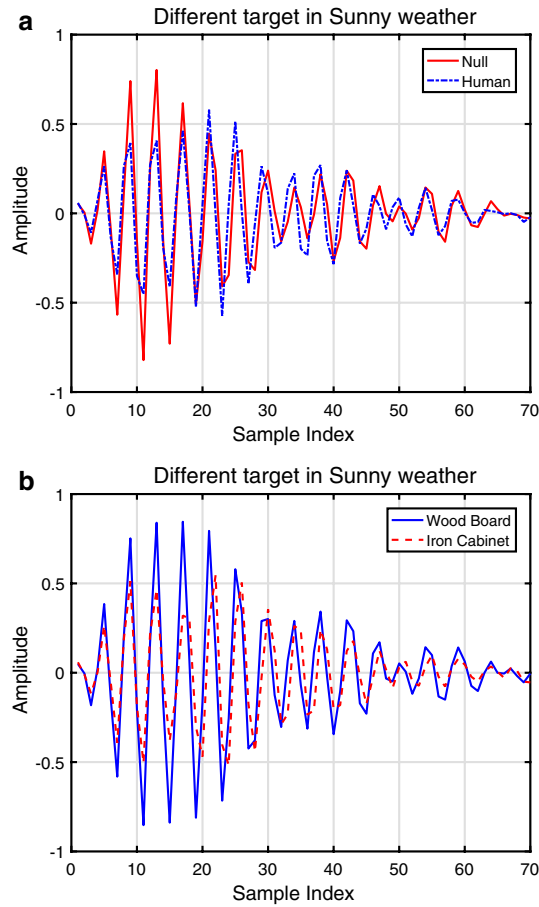

Fig. 3 Received UWB signal

Xue *et al. J Wireless Com Network*   (2021) 2021:41

Page 6 of 19
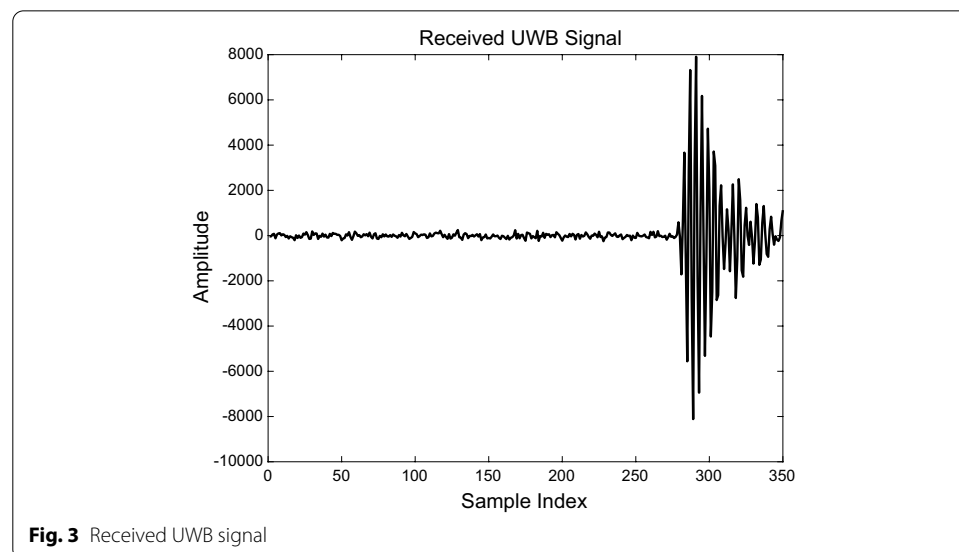
Fig. 4 Received UWB signals of different target in sunny weather

collected in the first scene as background noise $r_1(t) = \sum_{k=1}^{n} A_k s(t - t_k) + \epsilon_1(t)$, and the signals collected in the other three scenes with the target object as the received signal $r_2(t) = \sum_{p=1}^{n} A_p s(t - t_p) + \epsilon_2(t)$. According to formula (2), the cross-correlation is carried out on two signals: the signal with target $r_2(t)$ and background noise $r_1(t)$, and the alignment point τ_0 with the largest correlation value of the two signals is found. After locating the maximum correlation point τ_0, the received target signal $r_2(t)$ and background noise signal $r_1(t)$ are subtracted to obtain the pre-processing signal $y(t)$, as shown in formula (3). Therefore, as illustrated in Fig. 6, the number of sampling points of the obtained preprocessed signal is less than 70.

$$R(\tau) = \int_{-\infty}^{+\infty} r_1(t) r_2(t + \tau)\, dt, \qquad R(\tau_0) = R_{max}(\tau) \tag{2}$$

$$y(t) = r_2(t - \tau_0) - r_1(t - \tau_0) \tag{3}$$

Figure 6 shows the received UWB signals of human in different weather conditions after clutter noise removal.
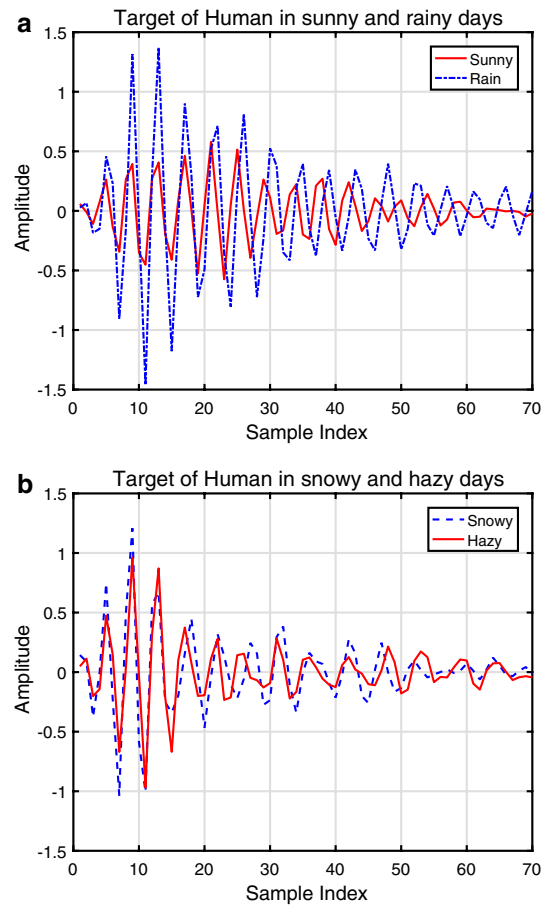
Fig. 5 Received UWB signals of human in different weather before preprocessing

Comparing Figs. 5 and 6, it can be seen that the difference of the same target signal collected under different weather conditions increases after the clutter noise is removed by cross-correlation. The signal after clutter noise removal still contains noise. How to eliminate the influence of noise as much as possible before feature extraction is the main problem of this paper.

In another article [17], the collected UWB signal has been proven to obey Gaussian distribution. Therefore, the residual noise signal in UWB signal can also be considered to obey Gaussian distribution. So, the method of factor analysis is proposed to to remove residual noise in the paper. The next section will discuss the principle of factor analysis and introduce the concept of deep factor analysis.

3 Deep factor analysis

Factor analysis (FA) [23] is a probabilistic model, which separates different target from background noise features to obtain robust target-invariant features. In maximum likelihood factor analysis, a k-dimensional vector of real-valued factor z is used to model a
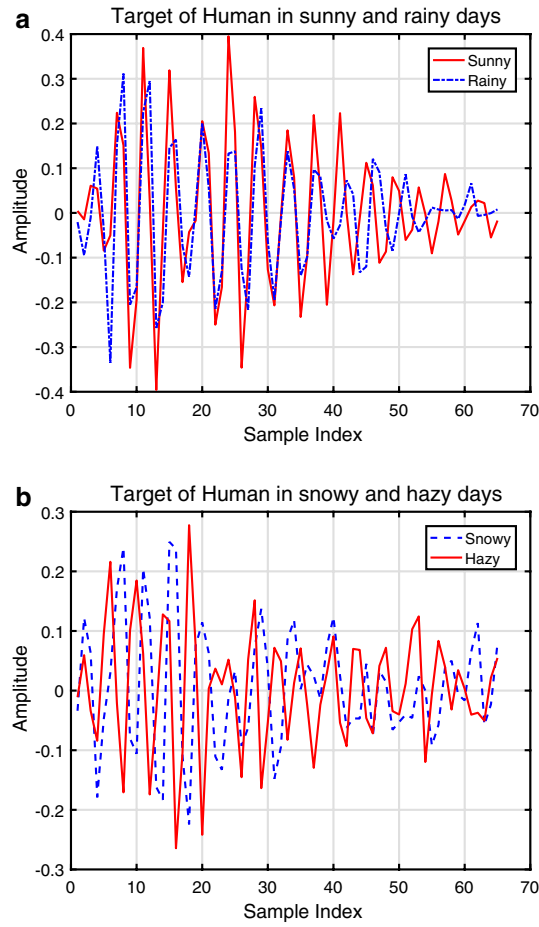
Fig. 6 Received UWB signals of human in different weather conditions after preprocessing

p-dimensional real-valued data vector y, where k is typically much smaller than p. The data vector y can be expressed as (4) [24]:

$$y = \mu + \Lambda z + \epsilon \tag{4}$$

where Λ is the factor loading matrix, ϵ is error signal. The distribution of random variable ϵ is $N(0, \Psi)$, where Ψ is a diagonal matrix. The factor z satisfies the Gaussian distribution with zero-mean independent normals and unit variance, which is expressed as $N(0, I)$ [24]. One of the key assumptions of factor analysis is that Ψ is a diagonal matrix. Find the mathematical expectation and variance of Eq. (3):

$$E(y) = E(\mu + \Lambda z + \epsilon) = \mu + \Lambda E(z) + E(\epsilon) = \mu \tag{5}$$

$$
\begin{aligned}
D(y) &= E[(y - E(y))(y - E(y))^{-1}] \\
&= E[(\mu + \Lambda z + \epsilon - \mu)(\mu + \Lambda z + \epsilon - \mu)^{-1}] \\
&= \Lambda\Lambda^{-1} + \Psi
\end{aligned}
\tag{6}
$$

The above analysis demonstrates that y obeys Gaussian distribution with μ mean and its covariance is $\Lambda\Lambda' + \Psi$. The purpose of factor analysis is to find the Λ and Ψ which can

best simulate the covariance structure of y. The variables z construct a model for the correlation between the elements of y, while the ϵ variables account for independent noise among the elements of y [25]. Given Λ and Ψ, the expected values of each factor can be calculated by linear projection [24, 26]:

$$E(z/y) = \Lambda'(\Psi + \Lambda\Lambda')^{-1}s = \beta s \tag{7}$$

where $\beta = \Lambda'(\Psi + \Lambda\Lambda')^{-1}$. The joint normality of data and factors can be obtained by (8):

$$P\left(\begin{bmatrix} y \\ z \end{bmatrix}\right) = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Lambda\Lambda' + \Psi & \Lambda \\ \Lambda' & I \end{bmatrix}\right) \tag{8}$$

since Ψ is a diagonal matrix, the $p \times p$ matrix $(\Psi + \Lambda\Lambda')$ can be efficiently inverted using the matrix inversion lemma:

$$(\Psi + \Lambda\Lambda')^{-1} = \Psi^{-1} - \Psi^{-1}\Lambda(I + \Lambda'\Psi^{-1}\Lambda)^{-1}\Lambda'\Psi^{-1} \tag{9}$$

where I is the $k \times k$ identity matrix. In addition, it is possible (actually necessary for expectation maximization (EM)) to calculate the second moment of the factors,

$$E(zz'/y) = Var(z/y) + E(z/y)E(z/y)' = I - \beta\Lambda + \beta yy'\beta' \tag{10}$$

Maximum likelihood factor analysis is carried out for (7) and (10) based on EM algorithm to obtain the parameters Λ and Ψ. EM algorithm is an iterative optimization strategy. Each iteration is divided into two steps, one is the expectation step (E-step), the other is the maximization step (M-step):

Expectation Step (E-step)

Given Λ, μ, Ψ, calculate $E(z/y_i)$ and $E(zz'/y_i)$ for each data point y_i.

Maximization Step (M-step)

$$\Lambda^{new} = \left(\sum_{i=1}^{n} y_i E(z/y_i)'\right)\left(\sum_{l=1}^{n} E(zz'/y_l)\right)^{-1} \tag{11}$$

$$\Psi^{new} = \frac{1}{n}diag\left\{\sum_{i=1}^{n} y_i y_i' - \Lambda^{new}E[z/y_i]y_i'\right\} \tag{12}$$

The *diag* operator in formula (12) sets all non-diagonal elements of the matrix to zero.

3.1 Deep factor analysis model

The probabilistic factor analysis model can be expressed as [27]. This article performs deep factor analysis on the signals after the previous preprocessing.

$$y = \mu + \Lambda y_1 + \varepsilon \tag{13}$$

where y is the preprocessed signal, μ is the mean of feature vector, y_1 is the target factor , Λ is the projection matrix whose columns span as the subspace of cross-target variation,

ε is the additive noise. Therefore, the essence of FA is to decompose any target into linear combinations of factors, which can be used to interpret the observed data.

FA is also bidirectional [28]. By constructing the multi-layer factor analysis model of the signal, the purpose of noise removal is achieved. Then, the signal y is reconstructed in reverse direction. The removal of surface noise ε is small, and there is still a lot of noise in the first signal factor y_1. In addition, since the factor analysis model has no special restriction on input data, the factor analysis method can be used to construct the upper level model for the first signal factor y_1 which still contains noise. So, a multi-level factor analysis model is obtained, which we call deep factor analysis.

$$y_0 = \mu_0 + \Lambda_0 y_1 + \varepsilon_0 \tag{14}$$

$$y_1 = \mu_1 + \Lambda_1 y_2 + \varepsilon_1 \tag{15}$$

And so forth, the factor analysis model of layer n signal is obtained.

$$y_n = \mu_n + \Lambda_n y_{n+1} + \varepsilon_n \tag{16}$$

For the factor y_n analysis model constructed by n-layer factor, μ_n is the mean value of the nth layer factor, Λ_n is the load matrix of the nth layer factor , and ε_n is the nth layer noise.

Therefore, Eq. (14) can be rewritten as:

$$
\begin{aligned}
y_0 &= \mu_0 + \Lambda_0 y_1 + \varepsilon_0 = \mu_0 + \Lambda_0(\mu_1 + \Lambda_1 y_2 + \varepsilon_1) + \varepsilon_0 \\
&= \mu_0 + \Lambda_0(\mu_1 + \Lambda_1(\mu_2 + \Lambda_2 y_3 + \varepsilon_2) + \varepsilon_1) + \varepsilon_0 \\
&= \mu_0 + \Lambda_0(\mu_1 + \Lambda_1(\mu_2 + \cdots + \Lambda_{n-1}(\mu_n + \Lambda_n y_{n+1} + \varepsilon_n) \\
&\quad + \varepsilon_{n-1}) + \cdots + \varepsilon_1) + \varepsilon_0 \\
&= (\mu_0 + \Lambda_0 \mu_1 + \Lambda_0 \Lambda_1 \mu_2 + \cdots + \Lambda_0 \Lambda_1 \cdots \Lambda_{n-1} \mu_n) \\
&\quad + \Lambda_0 \Lambda_1 \cdots \Lambda_{n-1} \Lambda_n y_{n+1} \\
&\quad + \boxed{(\Lambda_0 \cdots \Lambda_{n-1} \varepsilon_n + \Lambda_0 \cdots \Lambda_{n-2} \varepsilon_{n-1} + \cdots + \Lambda_0 \varepsilon_1 + \varepsilon_0)}
\end{aligned}
\tag{17}
$$

As can be seen from (17), the boxed part is the noise in the signal after deep factor decomposition. From this part, we can know that noise is no longer pure Gauss noise, but a weighted mixed noise. So after the deep factor decomposition, this part of noise is filtered out, and the filtered signal is reconstructed to obtain a pure target signal. In the forward process of constructing factor analysis model, the target signal is separated from noise signal by factor analysis model to obtain the hidden factors in the model. The hidden factors still contains noise, and the factor analysis model is constructed for the hidden factor. Therefore, the method of constructing factor analysis model for hidden factors is adopted step by step to gain the noise of each layer. After removing the noise of each layer, the pure target signal is obtained.

3.2 Signal reconstructing

In the process of factor analysis model construction, the signal $\hat{y}_n$ is reconstructed according to the the mean μ_n, load matrix Λ_n after removing the noise ε_n. As obtained in (17).

The hidden factor of the top layer y_{n+1} was used to reconstruct the hidden factor of the lower layer $\hat{y}_n$.

$$\hat{y}_n = \mu_n + \Lambda_n y_{n+1} \tag{18}$$

Then, using the same reconstruction rules, the reconstructed signal $\hat{y}_{n-1}$ of the upper layer is obtained from $\hat{y}_n$.

$$\hat{y}_{n-1} = \mu_{n-1} + \Lambda_{n-1}\hat{y}_n \tag{19}$$

And so on, the reconstruction signals of the upper layer are obtained in turn.

$$\hat{y}_1 = \mu_1 + \Lambda_1 \hat{y}_2 \tag{20}$$

$$\hat{y}_0 = \mu_0 + \Lambda_0 \hat{y}_1 \tag{21}$$

Then signal $\hat{y}_0$ is the reconstructed signal after the deep factor analysis model. The signal filters out most of the Gaussian and weighted Gaussian mixture noise. Figure 7 shows the architecture of a simple DFA, which consists of two procedure, named as deep factor analysis and signal reconstruction.

4 Target detection and results discussion

Figure 8 shows the procedure of target detection used in this study. As mentioned in Sect. 2, after data preprocessing, a total of 6720 data samples of three target objects are used for the experiment (2240 data samples are collected for each target in four weather conditions). They are roughly divided into training set and test set according to the 7:3

Fig. 7 Deep factor analysis denoising

Fig. 8 The procedure of target detection
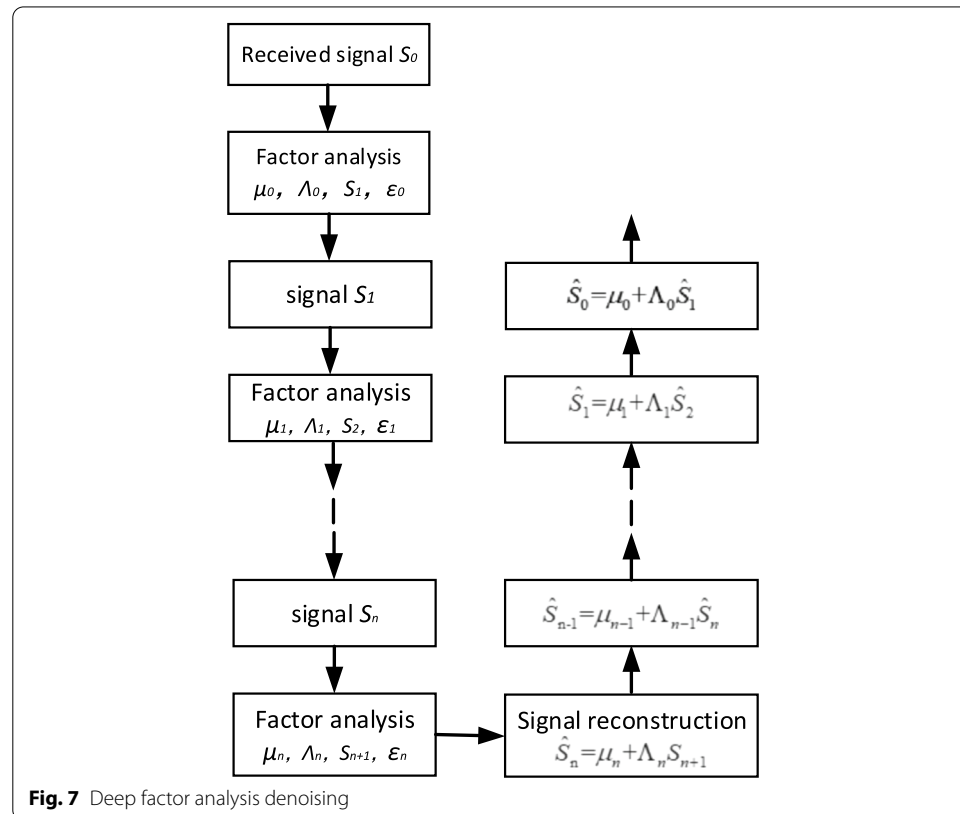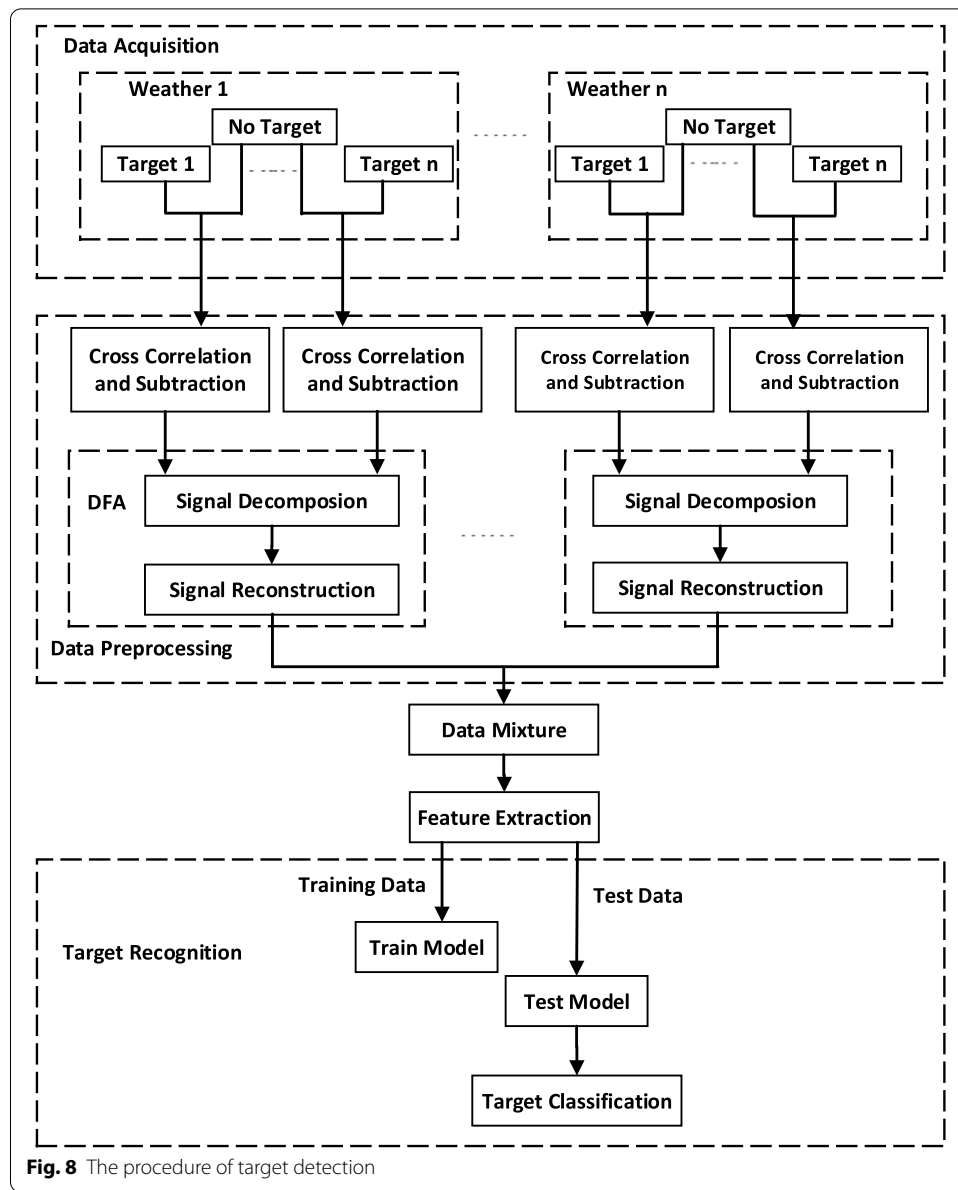
principle, which contain 1600 samples and 640 samples respectively. The training set is used to calculate the fitness function, train the classifier effectively, and the test set is used to verify the performance of the classifier.

4.1 Construction of deep factor analysis model

In the deep factor analysis model, it is very important to determine the number of factors of each layer and factor analysis layers. According to the dimensionality reduction principle of factor analysis, if there are too many layers, the useful signal will be lost while the noise is removed. It is necessary to select the appropriate layer number of deep factor analysis. Similarly, the same problem also exists when setting the number of factors in each layer. In this section, the target person is taken as an example, and the

Xue *et al. J Wireless Com Network* (2021) 2021:41

Page 13 of 19

recognition rate is taken as an evaluation index. Through a large number of experiments, the layers of factor analysis model and the factors of each layer are determined. In the first scenario, we measured the recognition rate of different target using support vector machine (SVM) classifier. The measured layers are 2 layers, 4 layers, 6 layers, 7 layers,8 layers and 10 layers. In general, Table 1 shows the classification accuracy increases with the increase of the number of layers due to the variation of noise. When the number of factor analysis layers increases to 7, the target object recognition rate increases gradually with the increase of the number of layers. However, when the number of layers increases to 8, the recognition rate decreases. The reason is that the useful information is removed while removing noise. Therefore, a 7-layer deep factor analysis model is applied to eliminate noise in all of the following experiments.

According to the previous factor analysis theory, the dimension of the remaining factor z is smaller than that of the original signal y. The dimension of y in this paper is 65, so the dimension of z is smaller than 65. After a large number of experiments, the factors z of layer 1–7 are 60, 50, 40, 30, 20, 10, 5 respectively. Taking human being as the target object as an example, according to the number of decomposition layers and the decomposition coefficient of each layer, the received UWB signal is decomposed by seven-layer factor, as shown in Fig. 9. Waveform (a) in Fig. 9 is the preprocessing UWB signal. Because the seven-layer factor analysis is set, (b)–(h) shows the seven-layer factor morphology. It can be clearly seen that the sampling points of each layer are different and decrease layer by layer.

After the decomposition of UWB signal containing noise is completed, the signal can be reconstructed according to the bidirectional nature of the factor analysis model. After removing the noise signal ε from each layer, the signal is reconstructed, as shown in Fig. 10. From the experimental results of Fig. 10, it can be seen that the noise elimination is obvious after signal decomposition and reconstruction.

4.2 The performance of DFA

After the clutter noise is removed by preprocessing and deep factor analysis, the features of UWB signals are extracted. The paper mainly studies the denoising performance of deep factor analysis algorithm, classic time-domain statistical features are selected, as shown in Table 2 [19]. Then the validity of deep factor analysis is verified by different classifiers, such as SVM, k-NearestNeighbor (KNN) and Back Propagation Neural Network (BPNN). It can be seen from the confusion matrix in Tables 3, 4 and 5 that after deep factor analysis, the influence of weather changes on target recognition is well

Table 1 Classification accuracy (%) of different number of layers

	Human being (%)	Wood board (%)	Iron cabinet (%)
10 layers	73.39	85.76	71.48
8 layers	89.15	92.60	74.45
7 layers	97.23	99.69	96.83
6 layers	95.20	99.63	97.78
4 layers	91.81	98.49	88.41
2 layers	86.99	92.60	74.45

Fig. 9 Signal decomposition

Fig. 10 Signal reconstruction

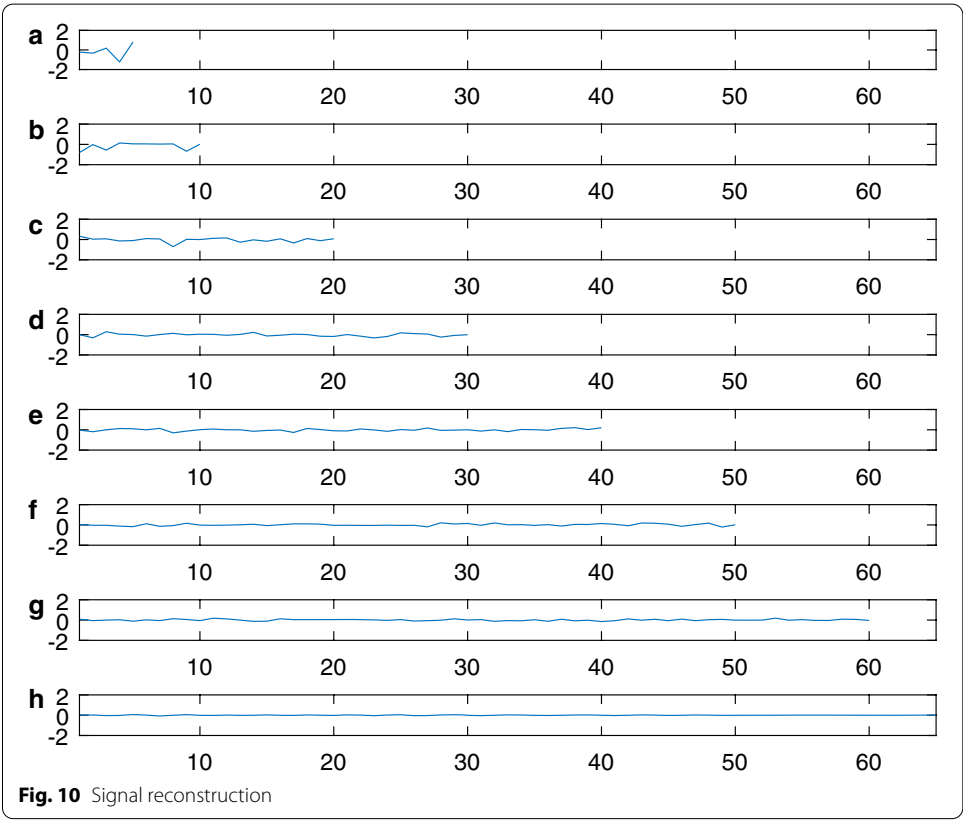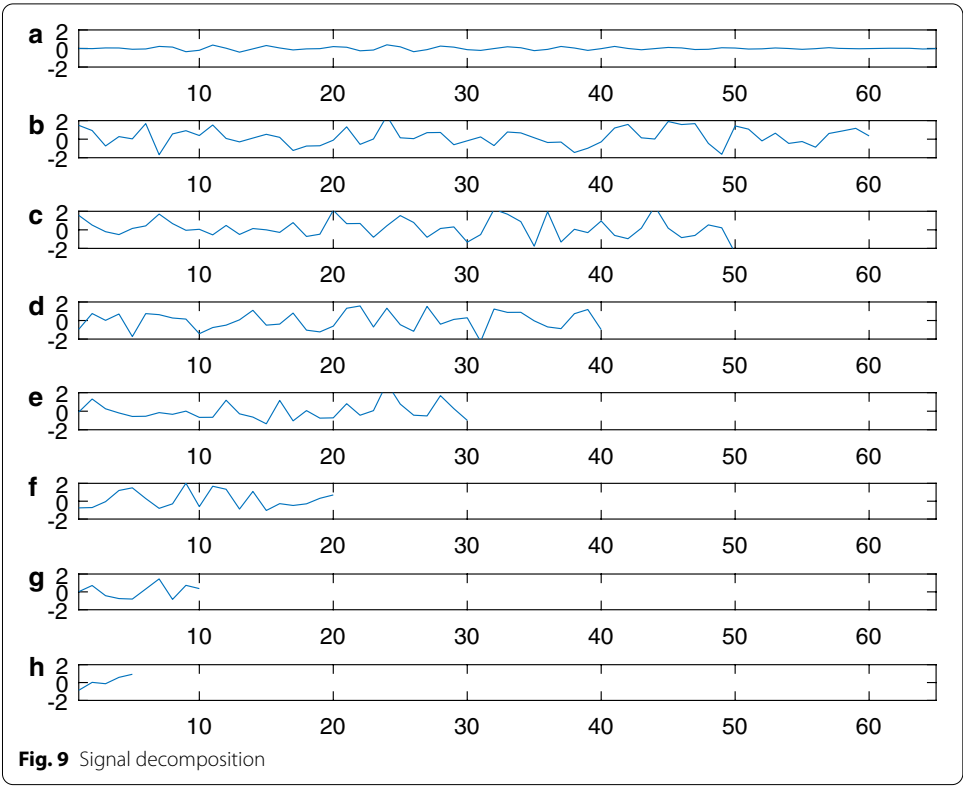**Table 2 Seven typical SP-based parameters [19]**

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 $|r(t)|$: $\sigma_{|r(t)|}^2 = \frac{1}{T} \int_T |r(t) - \mu_{|r|}|^2 \, dt$

Kurtosis of $|r(t)|$: $k = \frac{1}{\sigma_{|r|}^4} \int_T (|r(t)| - \mu_{|r|})^4 \, dt$

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

Table 3 Confusion matrix of the DFA-SVM approach

	Human being	Wood board	Iron cabinet
Human being	622.29	1.72	15.99
Wood board	0.97	638.00	1.03
Iron cabinet	15.38	4.89	619.73

Table 4 Confusion matrix of the DFA-KNN approach

	Human being	Wood board	Iron cabinet
Human being	602.87	4.23	32.90
Wood board	0.03	632.33	7.64
Iron cabinet	9.95	14.64	615.41

Table 5 Confusion matrix of the DFA-BPNN approach

|	Human being	Wood board	Iron cabinet
Human being	579.66	4.66	55.68
Wood board	0.06	625.04	14.90
Iron cabinet	16.50	16.49	607.01

eliminated. The data in the four weathers have been mixed and the recognition rate has reached more than 90%.

In order to further verify the effectiveness of DFA_based denoising algorithm, samples are denoised by DFA_based and non-denoising respectively, and then the recognition rate is compared by the same classifier. The results are presented in Table 6. The recognition accuracies for all three types of target are equal to or higher than 90% which is adequate for many other high-level applications. The main reason for this result is that from the time-domain waveform shown in Figs. 4 and 5, the difference between target signals after DFA becomes larger. The experimental results further verify the effectiveness of DFA_based algorithm.

As can be seen from the results in Table 6, that the average recognition rate of the system is increased by nearly 25% when the DFA_based algorithm is applied for data preprocessing. Similarly, the complexity of the system is higher than that without

Table 6 Summary of classification accuracy (%) using different classifiers

Algorithm description	Human being (%)	Wood board (%)	Iron cabinet (%)
DFA+SVM	97.23	99.69	96.83
SVM	74.22	67.66	60.00
DFA+KNN	97.11	99.19	91.08
KNN	74.49	68.75	58.77
DFA+BPNN	90.57	97.66	95.02
BPNN	62.37	58.00	44.75
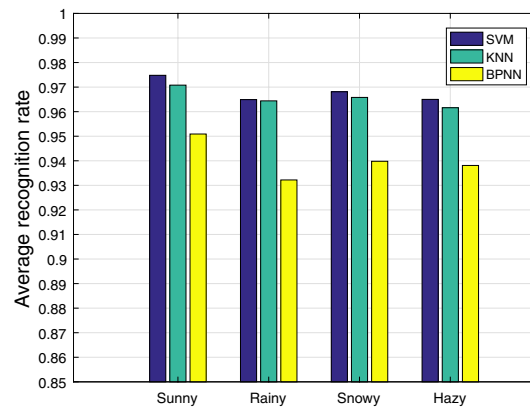
Fig. 11 Stability of DFA algorithm by cross-validation (mixed data set as training set)

DFA_based algorithm. The DFA_based algorithm in this study had a lower complexity compared with others. For N sampling points, the complexity of the factor analysis noise reduction algorithm was O(N) [28]. In addition, from the point view of the system, the running time of the system increased by about 17 s after the addition of DFA algorithm.

4.3 Cross-validation through different weather conditions

4.3.1 Mixed dataset as training set

In the previous measurement, we mixed the target measurement data under four kinds of weather conditions, taking one part for training and the other part for testing. In order to verify the validity of DFA_based algorithm in any weather, the mixed data is trained, and then the measurement data in any weather is randomly selected as test set for cross-validation. Therefore, we set up four actual scenarios. For each scenario, we make multi-point measurements of three kinds of target data and investigate the average recognition rate. The four single scenarios include (1) sunny days, (2) rainy days, (3) snowy days, and (4) haze days. In the validation process, we randomly select 1000 samples from mixed data as training samples. Then, the data collected in sunny, rainy, snowy and haze days are respectively taken as the input of the trained model for testing. The classification is performed through the use of three different algorithm, SVM, KNN and BPNN. Each test is executed 100 times, and the average recognition rate of each classification algorithm for three kinds of targets in each weather is obtained as shown in the Fig. 11.

It can be found from Fig. 11 that the DFA algorithm can eliminate the influence of clutter noise and achieve a high recognition rate for the data collected in any weather. No matter which classifier is used, the recognition rate tends to be stable in four different weather conditions. It further demonstrates that the proposed DFA denoising algorithm can effectively eliminate the influence of weather changes on target recognition.

4.3.2 Single dataset as training set

The above experiments demonstrate the effectiveness of DFA_based denoising algorithm. However, the models are trained by mixing data collected from various weather conditions. In this section, the collected data in a single weather is used as the input of the training model, and then the data collected in any other weather will be used as the test set to verify the effectiveness of the DFA_based denoising algorithm.

Figure 12 is the recognition result obtained by taking the data collected in sunny days as training set. It can be seen from Fig. 12 that when the data collected in any weather conditions such as rainy, snowy and haze days are used as test set, no matter which classifier is used, a result with a recognition rate of more than 93% can be obtained. Thus, it can be seen that the DFA_based denoising algorithm adopted in this paper can effectively eliminate the influence of weather changes on target recognition in the data preprocessing stage so as to simplify the subsequent processing process. This method has universality in practical application. Comparing Figs. 11 and 12, we can see that the correct recognition rate of the target is very stable no matter we use the mixed data or single weather data as training set. It can be concluded that after the weather effect is removed as background noise in the data preprocessing stage, even the most common statistical features and the simplest classifier are used, good recognition results can be obtained. Furthermore, the practicability and universality of the method are illustrated, which can be popularized in practical application.

5 Conclusion and future work

In this work, the performance analysis of signal preprocessing and DFA_based approach for noise reduction has been investigated under different weather conditions. Data are collected using UWB transceiver and preprocessed by correlation for removing most
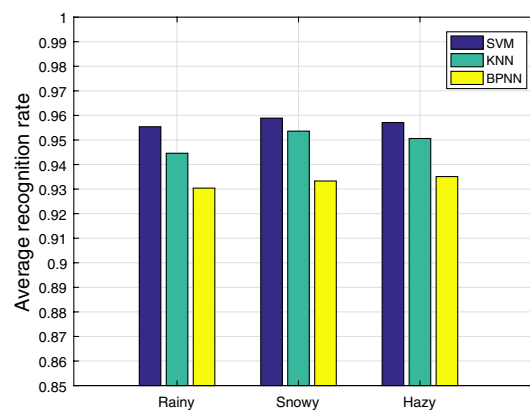

Fig. 12 Stability of DFA algorithm by cross-validation (Sunny data set as training set)

Xue *et al. J Wireless Com Network* (2021) 2021:41

Page 18 of 19

background noise. To extensively verify the universality of the proposed DFA_based approach, a variety of target objects, including human being, wood board and iron cabinet targets, are measured in foliage environment under four different weather conditions. The impact of the number of layers in DFA_based model for target classification is analyzed. In addition, cross-validation is performed to verify the effectiveness of the DFA_based algorithm in any weather. Through cross validation, it can be seen that the DFA_based algorithm can eliminate the influence of weather change on target recognition, and the target recognition rate tends to be constant under any weather, reaching more than 93%. Therefore, we can firmly conclude that the DFA_based noise reduction method is a feasible solution, and the method in this paper can be used to improve the classification and recognition rate of different types of targets under different weather conditions in a foliage environment.

For future work, the proposed algorithm will be extended to the moving object. More elaborate tests on new type of target and large datasets will be performed.

Abbreviations
DFA: Deep factor analysis; IR-UWB: Impulse-radio ultra-wideband; SVM: Support vector machine; HMMs: Hidden Markov models; KNN: K-NearestNeighbor; BPNN: Back Propagation Neural Network.

Author details
[1] Key Laboratory of Universal Wireless Communication, BUPT, Xitucheng Road 10, Beijing 100876, China. [2] The College of Electronic Information Engineering, Hebei University, Qiyi dong Road 2666, Baoding 071002, Hebei Province, China.

References
1. Y. Zhong, Y. Yang, X. Zhu, E. Dutkiewicz, Z. Zhou, T. Jiang, Device-free sensing for personnel detection in a foliage environment. IEEE Geosci. Remote Sens. Lett. **14**(6), 921–925 (2017)
2. Y. Zhong, E. Dutkiewicz, Y. Yang, X. Zhu, Z. Zhou, T. Jiang, Internet of mission-critical things: human and animal classification-a device-free sensing approach. IEEE Internet Things J. **5**(5), 3369–3377 (2018). https://doi.org/10.1109/JIOT.2017.2760322
3. X. Yang, Z. Liu, W. Nie, W. He, Q. Pu, AP optimization for Wi-Fi indoor positioning based on RSS feature fuzzy mapping and clustering. IEEE Access **8**, 153599–153609 (2020)
4. J. Wang, Z. Tian, X. Yang, M. Zhou, Twpalo: through-the-wall passive localization of moving human with Wi-Fi, in *2019 IEEE Global Communications Conference (GLOBECOM)* (2020)
5. X. Yang, R. Cao, M. Zhou, L. Xie, Temporal-frequency attention based human activity recognition using commercial WiFi devices. IEEE Access **8**, 137758–137769 (2020)
6. S. Depatla, A. Muralidharan, Y. Mostofi, Occupancy estimation using only WiFi power measurements. IEEE J. Sel. Areas Commun. **33**(7), 1381–1393 (2015)
7. G. Zhao, Q. Liang, T.S. Durrani, An EMD based sense-through-foliage target detection UWB radar sensor networks. IEEE Access **6**, 29254–29261 (2018). https://doi.org/10.1109/ACCESS.2018.2841900

8. S. Zhai, T. Jiang, Target detection and classification by measuring and processing bistatic UWB radar signal. Measurement **47**(1), 547–557 (2014)
9. S. Zhai, T. Jiang, A novel particle swarm optimization trained support vector machine for automatic sense-through-foliage target recognition system. Knowl. Based Syst. **65**(4), 50–59 (2014)
10. M. You, T. Jiang, New method for target identification in a foliage environment using selected bispectra and chaos particle swarm optimisation-based support vector machine. IET Signal Proc. **8**(1), 76–84 (2014)
11. J. Liang, Q. Liang, Sense-through-foliage target detection using UWB radar sensor networks. Pattern Recognit. Lett. **31**(11), 1412–1421 (2010)
12. D.W. Yang, S.J. Lee, T.H. Kang, J.H. Yoon, J.H. Ko, An improved classification method of concealed obstacles using UWB radar and stereo cameras, in *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)* (2011), pp. 1–4
13. S. Zhai, T. Jiang, Sparse representation-based feature extraction combined with support vector machine for sense-through- foliage target detection and recognition. IET Signal Proc. **8**(5), 458–466 (2013)
14. J.D. Bryan, J. Kwon, N. Lee, Y. Kim, Application of ultra-wide band radar for classification of human activities. IET Radar Sonar Navigation **6**(3), 172–179 (2012)
15. W. Xue, J. Ting, J. Shi, Animal intrusion detection based on convolutional neural network, in *The 17th International Symposium on Communications and Information Technologies, ISCIT 2017* (2017), pp. 1–5. https://doi.org/10.1109/ISCIT.2017.8261234
16. Y. Zhong, Z. Zhou, T. Jiang, M. Heimlich, E. Dutkiewicz, G. Fang, Classification of animals and people based on radio-sensor network, in *International Symposium on Communications and Information Technologies* (2016), pp. 113–116. https://doi.org/10.1109/ISCIT.2016.7751603
17. W. Xue, T. Jiang, An adaptive algorithm for target recognition using Gaussian mixture models. Measurement **124**, 233–240 (2018)
18. W. Xue, T. Jiang, Sense-through-foliage target detection based on sparse representation and Gaussian mixture models, in *IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia* (2018), pp. 6376–6379
19. Y. Zhong, Y. Yang, X. Zhu, Y. Huang, E. Dutkiewicz, Z. Zheng, T. Jiang, Impact of seasonal variations on foliage penetration experiment: a WSN-based device-free sensing approach. IEEE Trans. Geosci. Remote Sens. **56**(9), 5035–5045 (2018)
20. Y. Huang, Y. Zhong, Q. Wu, E. Dutkiewicz, J. Ting, Cost-effective foliage penetration human detection under severe weather conditions based on auto-encoder/decoder neural network. IEEE Internet Things J. **6**, 6190–6200 (2018). https://doi.org/10.1109/JIOT.2018.2878880
21. J. Liang, Q. Liang, S.W. Samn, A differential based approach for sense-through-foliage target detection using UWB radar sensor networks, in *ICC 2008-IEEE International Conference on Communications, Beijing* (2008), pp. 1952–1956. https://doi.org/10.1109/ICC.2008.374
22. S. Zhai, T. Jiang, A novel sense-through-foliage target recognition system based on sparse representation and improved particle swarm optimization-based support vector machine. Measurement **46**(10), 3994–4004 (2013)
23. D. Gong, Z. Li, D. Lin, J. Liu, X. Tang, Hidden factor analysis for age invariant face recognition, in *2013 IEEE International Conference on Computer Vision* (2013), pp. 2872–2879. https://doi.org/10.1109/ICCV.2013.357
24. Z. Ghahramani, E.G. Hinton, The EM algorithm for mixtures of factor analyzers. Technical report, CRG-TR-96C1, University of Toronto (1997)
25. M. Yang, N. Ahuja, D. Kriegman, Face detection using mixtures of linear subspaces, in *IEEE International Conference on Automatic Face & Gesture Recognition* (2000), pp. 70–76
26. H. Kaya, A.A. Salah, Adaptive mixtures of factor analyzers. arXiv:abs/1507.02801 (2015)
27. H. Li, H. Zou, H. Hu, Modified hidden factor analysis for cross-age face recognition. IEEE Signal Process. Lett. **24**(4), 465–469 (2017)
28. G. Wang, L. Yang, M. Liu, X. Yuan, P. Xiong, F. Lin, X. Liu, ECG signal denoising based on deep factor analysis. Biomed. Signal Process. Control **1–9**, 101824 (2019)

Publisher's Note