BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 71(4), 2023, Article number: e145766
DOI: 10.24425/bpasts.2023.145766

Research on communication emitter identification based on semi-supervised dimensionality reduction in complex electromagnetic environment

Wei GE[1,2], **Lin QI**[1,2], **Lin TONG**[1,2], **Jun ZHU**[1,2], **Jing ZHANG**[1,2], **Dongyang ZHAO**[3], and **Ke LI**[1,2]*

[1] School of Information & Computer Science, Anhui Agricultural University, Hefei, Anhui, 230036, China
[2] Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, Hefei, Anhui, 230601, China
[3] Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, ShenZhen, GuangDong, 518000, China

Abstract. The individual identification of communication emitters is a process of identifying different emitters based on the radio frequency fingerprint features extracted from the received signals. Due to the inherent non-linearity of the emitter power amplifier, the fingerprints provide distinguishing features for emitter identification. In this study, approximate entropy is introduced into variational mode decomposition, whose features performed in each mode which is decomposed from the reconstructed signal are extracted while the local minimum removal method is used to filter out the noise mode to improve SNR. We proposed a semi-supervised dimensionality reduction method named exponential semi-supervised discriminant analysis in order to reduce the high-dimensional feature vectors of the signals, and LightGBM is applied to build a classifier for communication emitter identification. The experimental results show that the method performs better than the state-of-the-art individual communication emitter identification technology for the steady signal data set of radio stations with the same plant, batch and model.

Key words: communication emitter identification; feature extraction; dimensionality reduction; VMD; ESDA.

1. INTRODUCTION

The individual identification of communication emitters refers to specific individual recognition of communication emitters from the same factory, batch, and model based on the characteristics of radio frequency (RF) fingerprint [1]. The fingerprint is produced by the internal physical characteristics of the device and the interaction between the circuit components [2]. A variety of electromagnetic signals are dense and interfered with when communicating in the real world, and the signal received in noncooperative communication has fewer labels which makes it difficult to identify them. Hence, the individual identification of communication emitters in a complex electromagnetic environment is of great research value and significance in both military and civilian fields.

In the military field, the challenge of modern communication defense technology is how to identify the RF fingerprints of enemy communication emitters timely and accurately in case of non-cooperative communication, so as to ensure the targeted surveillance, electronic interference, or firepower attack on important enemy electronic communications equipment at critical moments [3, 4]. In the civil field, as far as wireless network information security is concerned, after the security information is embedded in the physical layer of the communication emitter device, the RF fingerprint on the received signal is identified. With the security information authentication technology, the security of the wireless communication system can be greatly improved [5]. In terms of cognitive radio spectrum management, it monitors whether legitimate radio stations comply with frequency usage regulations, intercepts the interference from illegal radio stations, and identifies signal sources. Through the radio frequency fingerprint identification technology of the communication emitter, the hardware characteristics of the radio station are extracted, and the legitimate radio station and illegal radio station are distinguished, so as to achieve the purpose of spectrum resource management [6, 7]. With respect to equipment fault diagnosis, the equipment fault can be diagnosed quickly and effectively by detecting the radio frequency fingerprint of the radiation signal of the equipment [8].

According to the way the signal changes at different times, the signal can be divided into transient and steady state. Because of the short duration of transient signals, it is difficult to detect and capture transient signal under non-cooperative conditions, the practical application of transient signals [6]. On the other hand, compared with transient signals, steady-state signals are easier to capture, and it can provide a statistically more stable RF fingerprint than transient signals [9]. The methods of RF fingerprint extraction of steady-state signals can be roughly divided into methods based on time-frequency analysis and methods based on the nonlinearity of transmitter system.

In the aspect of time-frequency analysis, Liu [10] extracted the location and scale features of radar emitter by using Scale Invariant Feature Transform (SIFT). The disadvantage is that the recognition rate of this method is low when the signal-to-noise ratio is lower than 2.5 dB. Han [11] proposed a method of multi-scale segmentation (3D-HESMS) radio frequency fin-

*e-mail: like@ahau.edu.cn

Manuscript submitted 2022-10-20, revised 2023-03-29, initially accepted for publication 2023-04-23, published in August 2023.

gerprint extraction based on a three-dimensional Hilbert energy spectrum but it did not solve the problem of high computational complexity. Zhang *et al.* [12] applied the feedback classification algorithm composed of dynamic curve fitting and back propagation neural network to transient signal classification, and the recognition rate reached more than 95% in the unsupervised state.

In terms of nonlinearity of transmitter system, Huang [13] extracted the normalized permutation entropy (NPE) of the output signal of the transmitter system as the radio frequency fingerprint of the communication emitter, and the individual recognition rate of different radio stations reached 95%. However, this method depends on the selection of embedding dimension. Wu [14] used the fractal synthesis method based on the box dimension and the variance dimension to extract the radio frequency fingerprint, but it was necessary to increase the number of signal segments to improve the recognition performance, resulting in increased computational complexity. Jeong [15] developed probability moments and approximate entropy of radar signals as effective features, and the recognition accuracy of 100 radar signals was 99%. This method effectively reduces the dimension of the input signal and does not distinguish between transient signals and steady-state signals. Lu *et al.* [16] used permutation entropy and spectral features to mine the nonlinear features of the data, and this method achieved good recognition performance.

A method based on the nonlinearity of the transmitter system is studied in the this paper to solve the problem of individual identification of communication emitters in a complex electromagnetic environment. The signal is denoised based on Approximate Entropy local minimum removal of different modes after being decomposed by Variational Mode Decomposition (VMD) and VMD is also used to extract the Information Entropy of the emitters based on the preprocessed signal. Then we proposed the Exponential Semi-supervised Discriminant Analysis (ESDA) for feature dimensionality reduction, and finally, used the LightGBM method to classify different individual communication emitters.

The rest of the paper is organized as follows. The second section introduces the method we proposed. In the third part, the experimental results of our method are presented. The comparison with the state-of-art communication emitter identification methods is provided in Section 4, followed by a conclusion in Section 5.

2. PROPOSED METHOD

2.1. Variational Mode Decomposition

VMD is a nonlinear signal analysis method which adaptively determines the relevant wavebands of the steady-state signal of emitters by setting the number of decomposition modes [8].

Let the signal $u(t)$ be

$$u(t) = E(t)\cos(\theta(t)), \tag{1}$$

where $\theta(t)$ is a non-decreasing phase function, $E(t)$ is the envelope, and $\omega(t) = \theta'(t)$ is the instantaneous frequency. VMD can be represented as a constrained variational problem as follows [17]:

$$\begin{aligned} &\min_{u_k,\omega_k}\left\{\sum_k \left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|^2\right\} \\ &s.t.\ \sum_k u_k = f, \end{aligned} \tag{2}$$

where u_k and ω_k are the k-th mode component and center frequency respectively after the corresponding decomposition, K is the number of modes to be decomposed, ∂_t represents gradient for t, $\delta(t)$ is the Dirac constant and $*$ is the convolution operator. The Intrinsic Mode Functions of IMF (Intrinsic Mode Functions) decomposed by VMD method have independent center frequency and show sparse characteristics in frequency domain, possessing the characteristics of sparse research. In the process of solving the IMF, the VMD adopts the mirror extension to avoid the end effect similar to that in EMD decomposition. It can effectively avoid mode aliasing (if the k value is selected appropriately).

[17] provides more details about VMD, and the Algorithm 1 shows how to extract the Information Entropy of the communication emitters based on VMD.

Algorithm 1. Extraction of Information Entropy based on VMD

Input signal sample f(t)
Output Information Entropy feature E_k

1: Initialization: $\hat{u}_k$, ω_k, $\hat{\lambda}$, $n \leftarrow 0$
2: **repeat**
3: $n \leftarrow n+1$
4: **for** $k = 1 : K$ **do**
5: Update mod signals:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{j\neq k}\hat{u}_j(\omega) + \hat{\lambda}^n(\omega)/2}{1+2\alpha(\omega-\omega_k^n)^2}$$

6: Update mod frequencies:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega\,|\hat{u}_k(\omega)|^2\,\mathrm{d}\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2\,\mathrm{d}\omega}$$

7: **end for**
8: Update

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \zeta\left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)\right)$$

9: until convergence $\sum_k \left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2 \Big/ \left\|\hat{u}_k^n\right\|_2^2 < \sigma$
10: obtain mode component set $\{U_k | u_1, u_2, \cdots, u_k\}$;
11: Calculate the Hilbert spectrum $H_l(\omega,t)$ of the variational mode u_l

$$H_l(\omega,t)\begin{cases} a_l(t), & \omega = \omega_l(t), \\ 0, & \text{others.} \end{cases}$$

12: Calculate the energy ε_{ij} of each time-frequency frame of the Hilbert spectrum $H_l(\omega,t)$

$$\varepsilon_{ij} = \int_{(i-1)\delta_t}^{i\delta_t}\int_{(j-1)\delta_\omega}^{j\delta_\omega} \mathrm{d}\omega\,\mathrm{d}t$$

13: Calculate the Information Entropy $H(X_l)$ of the Hilbert spectrum $H_l(\omega,t)$

$$H(X_l) = -\sum_{i=1}^{n} p(\varepsilon_{ij})\log_2 p(\varepsilon_{ij})$$

Figure 1 shows the mode components obtained from the simulation signal decomposition of different SNRs. The structure of the mode components is the same, which indicates that the VMD has the advantages of strong decomposition ability and good denoising performance. It is robust to noise. The phase, envelope, and modulation information in each mode component of the emitter signal can help extract individual characteristics and improve the analyzability of the signal.

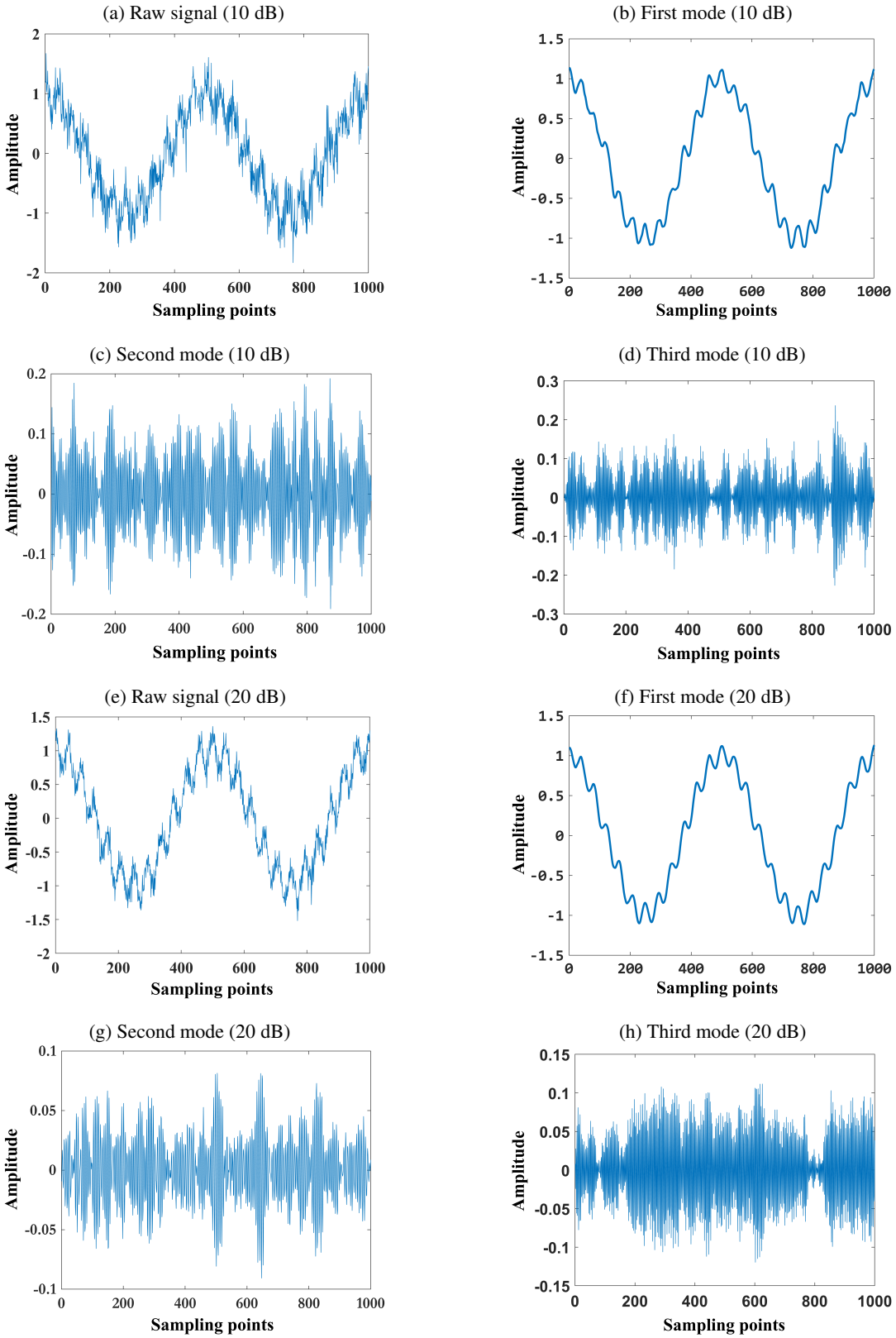


Fig. 1. The waveform of different modes after Variational Mode Decomposition

2.2. Approximate Entropy (ApEn) and Information Entropy

ApEn [18] is an algorithm used to quantify the unpredictability of variability and the number of regular patterns, which has good analysis results when the data is similar or irregular. We capture the change in the value of Approximate Entropy to find the dividing point between noise and radio signal.

Given a set of raw data: $x(1),x(2),\ldots,x(n)$, a total of n data points (m and r are preset parameters, m is the embedding dimension, r is the similarity tolerance, also known as filtering level). Form an m vector in order [18]: $X(i) = [x(i),x(i+1),\ldots,x(x+m-1)], i = 1,2,\ldots,n-m+1$.

The one with the largest difference between the corresponding elements of $X(i)$ and $X(j)$ is defined as the distance between them, denoted as [18] $\mathrm{d}[X(i),X(j)]$.

$$\mathrm{d}[X(i),X(j)] = \max_{K=0\sim m-1}[|x(i+k)-x(j+k)|]. \tag{3}$$

Given the similarity tolerance r, calculate the number g that $\mathrm{d}[X(i),X(j)] \leq r$ and $g/(N-M)$ is the similarity probability corresponding to the value of i.

Calculate the average of the logarithm of the probability of all the values of i and j, and we get the entropy value $\Phi^m(r)$ [18].

$$ApEn = \Phi^m(r) - \Phi^{m+1}(r). \tag{4}$$

The Lagrange multiplication operator λ is introduced to transform the constrained variational problem into an unconstrained variational problem and update the mode in the frequency domain and the center frequency. The k-th mode components of the signal $u_1,u_2,\ldots,u_K$ are arranged sequentially from high frequency to low frequency, and the dominance of noise on each component is gradually reduced. Hence, the dominance of the signal for each component is continuously enhanced. The method of determining the mode component demarcation u_i between the signal and noise is to calculate the approximate entropy value corresponding to each mode component firstly. Secondly, Finding the mode component u_i of the first signal sample corresponding to the approximate entropy value when taking the local minimum value. Finally, the previous mode component $u_1,\ldots,u_{i-1}$ is filtered out as noise, and the remaining mode components $u_i,u_{i+1},\ldots,u_K$ are reconstructed.

Information entropy is used to measure the amount of information contained in the data, and the amount of information varies with the probability of its occurrence [19]. In order to improve the representation ability of features, we use information entropy to calculate the information of each mode. The calculation method of information entropy is as follows [19]:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i), \tag{5}$$

where $p(x_i)$ is the probability of each vector.

2.3. Exponential Semi-supervised Discriminant Analysis

The emitter signal sample set $\mathbf{X}$ can be embedded into the optimal low-dimensional space through $\mathbf{X} \to \mathbf{F} = \mathbf{V}^T\mathbf{X}$ ($\mathbf{V}$ is the projection vector set $\mathbf{V} = [\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_c]$). The disadvantage of this method is that the linear method is used to map the information entropy features. For the nonlinear emitter signal, it is difficult to excavate its inherent geometry, which will affect the recognition accuracy of individual features. The proposed exponential semi-supervised discriminant analysis method is a semi-supervised learning method based on graph embedding. This method uses a matrix index to nonlinearly amplify the distance between different label sample sets, so as to improve the separability between classes, which can effectively make up for the deficiency of linear methods. It performs semi-supervised learning on the label information of a few labeled signal samples and the internal structure information of a large number of unlabeled signal samples. The high-dimensional Variational Mode Decomposition and Information Entropy(VMD-InEn) features are mapped to the low-dimensional subspace, thereby reducing the amount of calculation for identification and improving the accuracy of identification [20, 21].

a. Exponential Semi-supervised Discriminant Analysis (ESDA)

This paper proposes the matrix index for nonlinear processing. The matrix index is a special matrix function whose input and output object of the function are all square matrices [22]. For example, in $n \times n$a square matrix $\mathbf{A}$, the index of $\mathbf{A}$ is $\exp(\mathbf{A}) = \sum_{m=0}^{\infty}\frac{\mathbf{A}}{m!} = \mathbf{I}+\mathbf{A}+\frac{\mathbf{A}^2}{2!}+\frac{\mathbf{A}^3}{3!}+\cdots$, where $\mathbf{I}$ is the unit square matrix of $n \times n$. The matrix index has four important properties:

1. The matrix index is a finite full-rank non-singular matrix and always converges.
2. The exponent of any square matrix $\mathbf{A}$ is reversible, that is $\exp(\mathbf{A})^{-1} = \exp(-\mathbf{A})$.
3. If $\mathbf{B}$ is a non-singular matrix, then
$$\exp(\mathbf{B}^{-1}\mathbf{A}\mathbf{B}) = \mathbf{B}^{-1}\exp(\mathbf{A})\mathbf{B}.$$
4. If the eigenvectors of A are $t_1, t_2, \ldots, t_n$, and the eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the eigenvectors of $\exp(\mathbf{A})$ are also $t_1, t_2, \ldots, t_n$, and the eigenvalues of $\exp(\mathbf{A})$ are $\exp(\lambda_1)$, $\exp(\lambda_2), \ldots, \exp(\lambda_n)$.

Based on the above properties, matrix index is applied to semi-supervised discriminant analysis(SDA) in the paper. The matrix exponent is processed by the scattering matrix of the information entropy characteristics of the emitter signal, and the SDA optimization function of the exponent is as follows [22].

$$\mathbf{v}_{\mathrm{opt}} = \arg\max_{\mathbf{v}} \frac{\mathbf{v}^T\exp(\beta S_b)\mathbf{v}}{\mathbf{v}^T(\exp(\gamma S_t)+\alpha\mathbf{X}\mathbf{L}\mathbf{X}^T)\mathbf{v}}, \tag{6}$$

where β and γ are two balance parameters, the eigenvalue equation corresponding to the optimization function is given as [22]:

$$\exp(\beta\mathbf{S}_b)\mathbf{v} = \lambda\left[\exp(\gamma\mathbf{S}_t)+\alpha\mathbf{X}\mathbf{L}\mathbf{X}^T\right]\mathbf{v}. \tag{7}$$

The purpose of matrix exponential processing of scattering matrix is to embed the inter-class scatter matrix $\mathbf{S}_b$ and the overall scatter matrix $\mathbf{S}_t$ of the Information Entropy feature extracted from the emitter signal into the new space by using the

nonlinear mapping function Ψ, i.e. [22]:

$$\begin{cases} \mathbf{S}_b \to \Psi(\mathbf{S}_b) = \exp(\mathbf{S}_b), \\ \mathbf{S}_t \to \Psi(\mathbf{S}_t) = \exp(\mathbf{S}_t), \end{cases} \tag{8}$$

$\mathbf{S}_b$ and $\mathbf{S}_t$ can be written as $\mathbf{S}_b = \sigma_b^T \Lambda_b \sigma_b$, $\mathbf{S}_t = \sigma_t^T \Lambda_t \sigma_t$ when solving the eigenvalue equation, where σ_b, and σ_t are orthogonal matrices, Λ_b and Λ_t are diagonal matrices, which contain the eigenvalues of $\mathbf{S}_b$ and $\mathbf{S}_t$, respectively, and the diagonal elements are $\lambda_{b1}, \lambda_{b2}, \cdots, \lambda_{bn}$, and $\lambda_{t1}, \lambda_{t2}, \cdots, \lambda_{tn}$, respectively.

ESDA can find the best projection by maximizing the distance d_b between the information entropy features of different label samples and minimizing the distance d_t between the information entropy features of the same label emitter sample in the local neighborhood. In the input space [23], d_b and d_t are given as

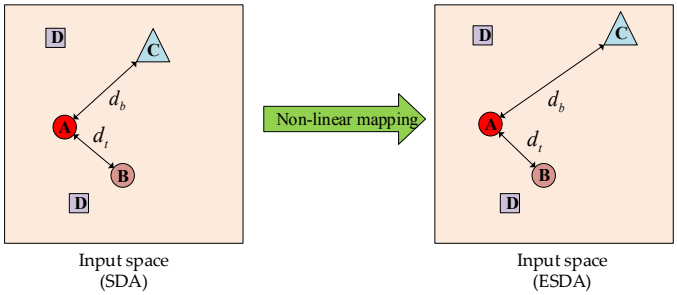
$$\begin{cases} d_b = \mathrm{tr}(\mathbf{S}_b) = \lambda_{b1} + \lambda_{b2} + \cdots + \lambda_{bm}, \\ d_t = \mathrm{tr}(\mathbf{S}_t) = \lambda_{t1} + \lambda_{t2} + \cdots + \lambda_{tm}. \end{cases} \tag{9}$$

After being embedded in a new space through equation (8), according to property 4 of matrix index, d_b and d_t can be defined as:

$$\begin{cases} d_b = \mathrm{tr}(\mathbf{S}_b) = \exp(\lambda_{b1}) + \exp(\lambda_{b2}) + \cdots + \exp(\lambda_{bm}), \\ d_t = \mathrm{tr}(\mathbf{S}_t) = \exp(\lambda_{t1}) + \exp(\lambda_{t2}) + \cdots + \exp(\lambda_{tm}). \end{cases} \tag{10}$$

Since $d_t < d_b$, based on the characteristic that the exponential function increment is exponentially increasing $\Delta d_t < \Delta d_b$, hence ESDA can significantly affect the amplification performance in determining the distance between different label data samples.

Figure 2 shows the geometric schematic diagram of ESDA. After the nonlinear mapping of the matrix index, the distance between samples A and C in the input space is significantly expanded.



A,B: samples with label 1; C: samples with label 2; D: samples without any label

Fig. 2. Geometric sketch of non-linear mapping

2.4. Light Gradient Boosting Machine (LightGBM)

LightGBM is a classification algorithm based on the decision tree. It can be used in sorting, classification, regression, and many other machine learning tasks. Traditional XGBoost requires pre-sorted, which consumes a lot of memory space. However, LightGBM is based on the histogram algorithm instead of the data structure built by pre-sorted, and there are many useful tricks after utilizing histogram. e.g. histogram does the difference and improves the cache hit rate. We use sampling (based on sample weights) to improve the training speed when facing large data volumes. Or we can assign sample weights to a certain class of samples during training. By introducing Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). LightGBM solves the shortcomings of traditional gradient boosting decision trees, which are time-consuming and inefficient [24].

In the GOSS method, the training examples are sorted in descending order according to the absolute value of the gradient. Select the first instance with a larger gradient a×100% to form the instance subset A. The remaining instances with smaller gradients are randomly sampled to obtain a subset B of instances with the size of $b \times |A^C|$. Finally, split the instances [24].

$$\overline{V}_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i \in A_l} g_i + \frac{1-a}{b}\sum_{x_i \in B_l} g_i\right)^2}{n_l^j(d)} + \frac{\left(\sum_{x_i \in A_r} g_i + \frac{1-a}{b}\sum_{x_i \in B_r} g_i\right)^2}{n_r^j(d)} \right), \tag{11}$$

where $A_l = \{x_i \in A : x_{ij} \le d\}$, $A_r = \{x_i \in A : x_{ij} > d\}$, $B_l = \{x_i \in B : x_{ij} \le d\}$, $B_r = \{x_i \in B : x_{ij} > d\}$.

In the EFB method, one of the features is selected as a vertex, and other features are selected in turn. Meanwhile, adding edges for every two features if they are not mutually exclusive, the greedy algorithm is used, which can produce reasonably good results.

3. EXPERIMENTAL RESULTS

In this experiment, 10 ultrashort-wave hand-held radio stations (SHRS) with BF 5518 model were used as identification objects. The signal acquisition equipment is the Egret RM200 receiver, and the signal was collected 100 m away from the SHRS in an open environment to reach a complex electromagnetic environment like non-cooperative communication and low SNR conditions, etc. The zero-IF I/Q orthogonal signal of SHRS was collected. The center frequency of the signal is 410 MHz and 160 MHz, and the two signals come from the same speaker, with a signal bandwidth of 25 kHz and a receiving channel bandwidth of 100 kHz. After the receiver receives the signal, it performs zero-IF processing on the signal and performs sampling at a sampling frequency of 204.8 kHz. The number of sampling points for each sample is 614, and 5000 signal samples of each SHRS are collected. All experiments were carried out on a Windows laptop computer and the CPU of Intel Core I 5-8250U 1.8 GHz.

Based on the traditional communication emitter identification process, a corresponding verification process is established

in this paper, as shown in Fig. 3. This method consists of three parts: signal denoising, semi-supervised feature extraction and emitter identification. In the part of signal denoising, the signal is decomposed into multiple modes based on VMD, and then the approximate entropy value corresponding to each mode is calculated, and the local minimum signal denoising method is used to remove noise. In the part of semi-supervised feature extraction, the RFF of communication emitters is obtained by VMD-InEn feature extraction method, and then the high-dimensional information entropy features are mapped to low-dimensional subspace by ESDA to form individual features, so as to make full use of signal samples with or without label information. In the part of emitter identification, a classifier for individual identification of communication emitters is constructed based on LightGBM.

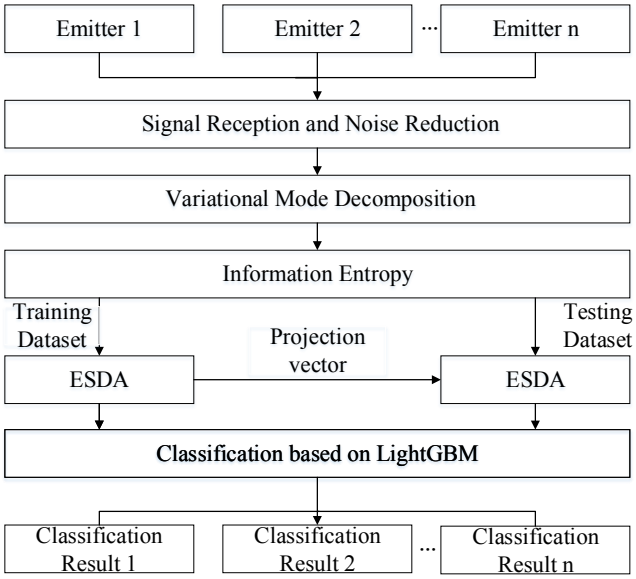


Fig. 3. The verification process of individual communication emitters

3.1. Performance Evaluation

Based on the communication emitter verification model established above, two sets of verification experiments are designed in this section to verify the performance and feasibility of the method we proposed. The first set of experiments verifies the performance of the Information Entropy extracted by VMD in characterizing the individual characteristics of communication emitters. The second group of experiments verifies the influence of the number of VMD modes on the recognition accuracy, which proves the necessity of using ESDA.

In the signal preprocessing step, the approximate entropy of the sampled signal is calculated based on VMD, and the signal is denoised by local minimum removal method [25]. In order to verify that the information entropy feature extracted based on the VMD can represent the information carried by the signal, we randomly selected 4 SHRS signal samples from the 10 SHRS with the center frequency of 410 MHz for information entropy feature extraction.

Figure 4 shows the difference in the distribution of Information Entropy features extracted from four randomly selected

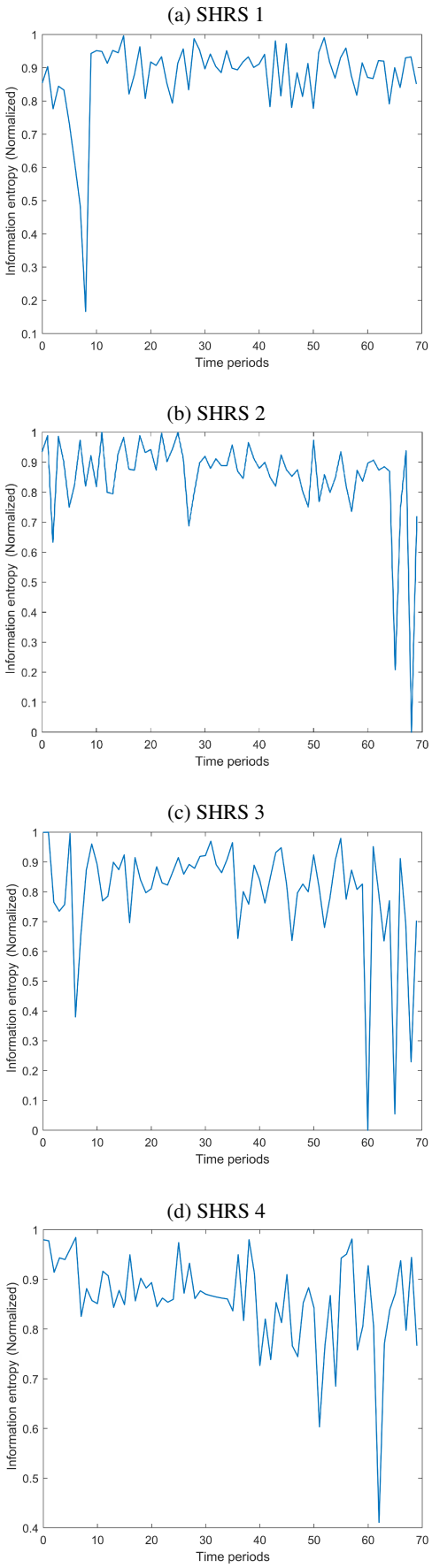


Fig. 4. Information Entropy feature of 4 SHRS

SHRS under the same mode decomposition number. According to the calculation, the correlation coefficients between the information entropy features of the four SHRS are all less than 0.5, and the correlation coefficients between the information entropy features of the SHRS 1 in different periods are all close to 0.9, which indicates that the correlation coefficients of the information entropy features of different SHRS are weak, but the same SHRS is strong.

We extract VMD-InEn features from the steady signals of 10 SHRS. The feature distribution of 10 SHRS is visualized to show the difference in fingerprint features between communication emitters. As can be seen from Fig. 5, the characteristics of the 10 SHRS are significantly different. Therefore, the information entropy extracted based on VMD can be used to distinguish SHRS.

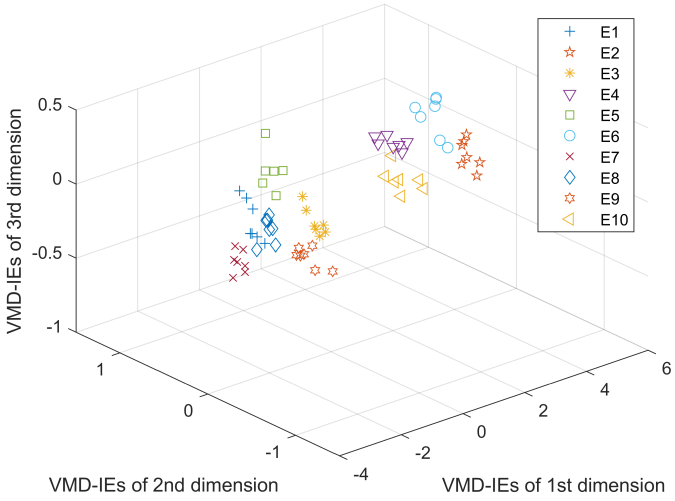


Fig. 5. VMD-InEn values for 10 SHRS

3.2. Select the Decomposition Number

When using VMD to extract signal features, if the number of modes is too small, it may not be able to fully excavate the individual information differences carried by the signal. If there are too many modes, it may cause a lot of information redundancy, and at the same time increase the scale of the system and the calculation time. Therefore, the purpose of this experiment is to select a suitable number to balance the performance of the model.

We randomly selected 2500 and 2500 samples in each SHRS as the training set and the test set, respectively, and set the number of modes to 3, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, and then extracted the signal features, and then classified them by using LightGBM.

As can be seen from Fig. 6, when the number is less than 15, the extracted features have insufficient representation ability and the recognition rate is lower than 60%. When the number is higher than 70, the information redundancy is too large, and the recognition rate decreases. The optimal number is 60–70, so 70 is chosen in this paper. ESDA is used to reduce the dimensionality of signal features, which not only keeps the original high-dimensional feature representation ability, but also reduce the calculation time.

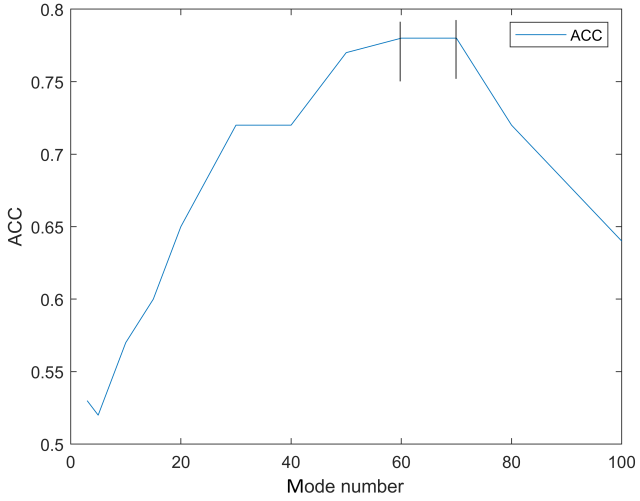


Fig. 6. Identification of different decomposition numbers

4. DISCUSSION

To illustrate the effectiveness of the communication emitter identification method in this paper, we compare our method with the state-of-art methods under two kinds of datasets in Table 1 and Table 2, and the ratio of the training set and test set is 5:5. The recognition performance on the basis of 10-fold cross-validation, and the average result were calculated to reduce the influence of random errors in the experiment.

Table 1
Dataset I

Number of samples	Center frequency	Number of SHRS	Speaker
50000	410 MHz	10	1

Table 2
Dataset II

Number of samples	Center frequency	Number of SHRS	Speaker
50000	160 MHz	10	1

a) The recognition rate under different SNRs is one of the methods to measure the identification of individual communication emitters. Figure 7 shows the performance of different identification schemes using dataset I, including RFFE-InfoGAN [2], SST [26], SSA [27], from which we can see that the method proposed in this paper has obvious advantages under higher SNR, and the recognition rate can reach nearly 100% when the SNR is higher than 20 dB.

b) We compared our method with other dimensionality reduction methods: SDA [28], SOPDA [29], SS-KLFDA [30]. These methods use LightGBM classifier and use data set I for identification. When the dimensionality of SHRS 1 is reduced, the number of training sets and test sets is 4986, respectively. In the training set, 2493 samples are from SHRS 1, and the remaining 2493 samples are from the other 9 SHRs.

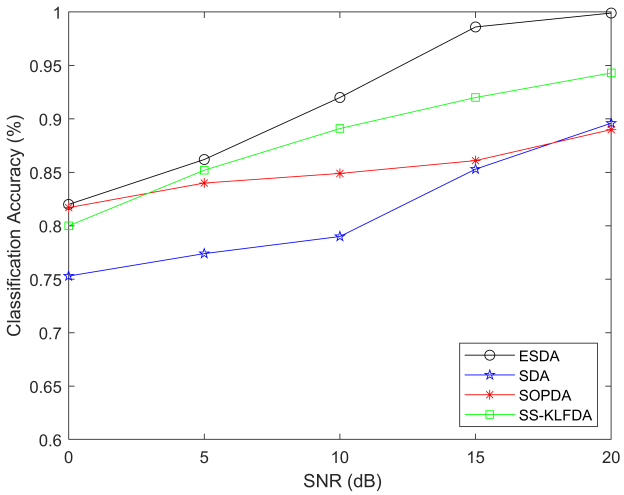


Fig. 7. Classification performance comparison for different schemes under different SNRs

The selection rules for the sample number of each SHRS in the test set are the same as those in the training set. The ratio of labeled data to unlabeled data of each SHRS is 1:100. The dimensionality reduction methods of the remaining 9 SHRS can be deduced by analogy. The recognition rates of the four methods under the SNR of 0–20 dB are shown in Fig. 8. It can be seen that the recognition rate increases as the increase of SNR. However, when the SNR is lower than 10 dB, the recognition rate of SDA, SOPDA, and SS-KLFDA are all lower than 90%, while the recognition rate of ESDA can remain above 98%. Therefore, the dimensionality reduction method proposed in this paper can be proved to be superior to the state-of-art dimensionality reduction methods.

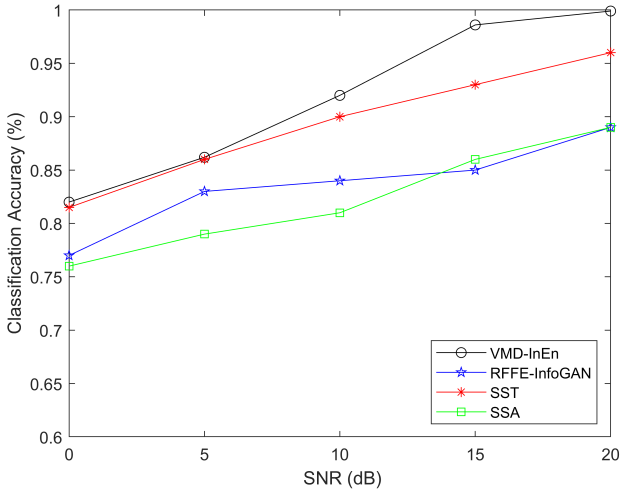


Fig. 8. Comparison of identification accuracy of four dimensionality reduction methods

On the training set, we performed 10-fold cross-validation on the VMD-InEns features before and after dimensionality reduction at a signal-to-noise ratio of 15 dB. The recognition rate, recognition time, and standard deviation are shown in Table 3. After dimensionality reduction, the redundancy of the feature matrix can be effectively reduced and the feature carried by the original matrix is preserved. The standard deviation of the classification accuracy shows that our model has a strong stability. Therefore, the ESDA method proposed in this paper can significantly improve the accuracy of identification and reduce recognition time.

Table 3
Comparison of identification accuracy and running time of LightGBM and XGBoost

Method	Recognition rate (%)	Identification time (s)	Standard deviation
LightGBM	98.5	6.22	0.0217
XGBoost	97.3	20.73	0.0234

In order to verify the robustness and effectiveness of our dimensionality reduction method, we compare the recognition results of two kinds of ratios of labeled data to unlabeled data (LUR) and four kinds of sample number. Figure 9 shows the result that the LUR is 1/100. We randomly selected 1000 and 9720 samples from data set I and data set II. The recognition rate of performing tenfold cross-validation can be maintained at more than 82.4%. Meanwhile, the identification time is shorter than 3.31 s. As can be seen from Fig. 10, we randomly selected

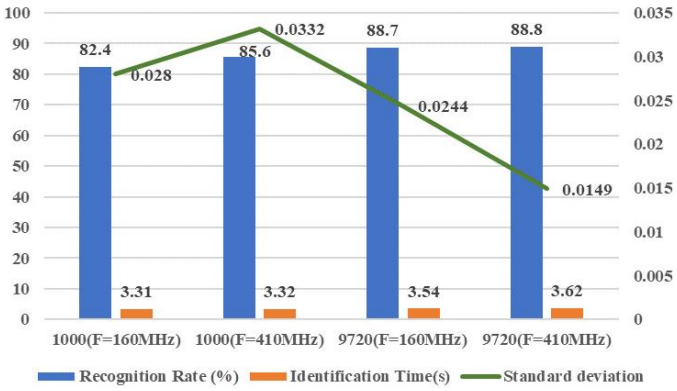


Fig. 9. Comparison of identification accuracy and running time under 1/100 LUR and different samples

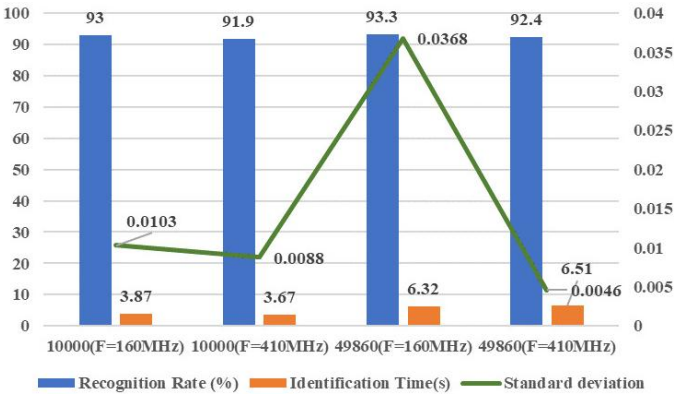


Fig. 10. Comparison of identification accuracy and running time under 1/1000 LUR and different samples

10000 and 49860 samples from data set I and data set II with 1/1000 LUR. The recognition performance is much improved on the basis of 10-fold cross-validation, and the recognition rate can be maintained at more than 91.9% and the identification time is still not long. Therefore, our method is suitable for data sets with different scales and different LURs.

c) We set up a classifier for individual identification of communication emitters based on LightGBM, which uses grid search. The optional parameters are: learning_rate = 0.1, lambda_l1 = 0.2, lambda_l2 = 0.2, max_depth = 7, num_leaves = 20. XGBoost [31], MDA [32], and KNN-P [33] are introduced to compare with the method we proposed. The ratio of a test set to a training set is 5:5. As can be seen from Fig. 11 and Table 3, compared with MDA and KNN-P, the classifier used in this paper performs well. The recognition rate of LightGBM is as good as XGBoost in data set I, but the recognition time of LightGBM is only 70% of XGBoost under the SNR of 15 dB. Therefore, the LightGBM has a better recognition rate and shorter recognition time under different signal-to-noise ratios.

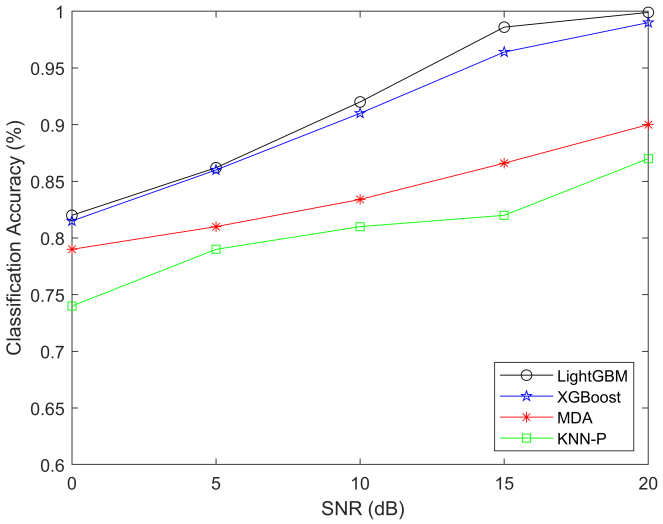


Fig. 11. Classification performance comparison for different classifiers under different SNRs

d) The number of communication emitters will affect the recognition results of the model.

The result shows that VMD-InEns can classify 10 communication emitters with a classification accuracy of more than 98%, but it can achieve better classification performance (nearly 99.9%) when identifying fewer emitters. Our method has been proved to maintain good classification performance in the case of more SHRS. In order to improve the performance of an enormous number of emitters, the training method needs to be more refined, which leads to higher computational costs. In future work, an effective algorithm needs to be considered to overcome this difficulty.

e) In addition, considering that when identifying multiple SHRS, the individual recognition results may be misjudged as other SHRS, the experiment uses the recognition confusion matrix shown in Fig. 13. The n-th row of the matrix represents the label distribution of the n-th SHRS individual feature data set

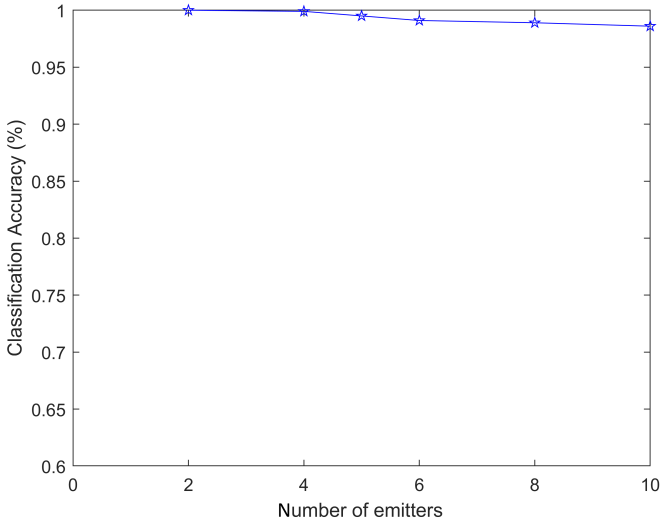


Fig. 12. Classification accuracy under different numbers of emitters

identified by the method we proposed, while the m-th column of the matrix represents the distribution of individual characteristic data sets corresponding to the m-th SHRS labels.

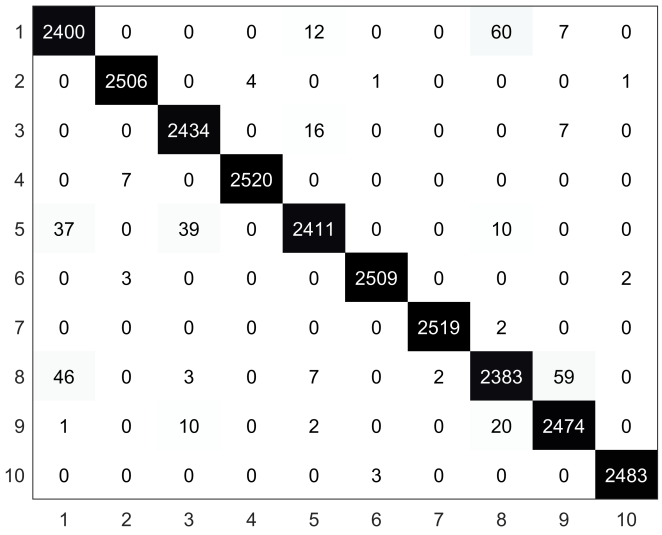


Fig. 13. Identification confusion matrix of 10 SHRS with 410 MHz center frequency

It can be seen that Fig. 13 shows the confusion matrix of communication emitter with the center frequency of 410 MHz in data set I, and the recognition performance is misjudged as SHRS 5 and 8 when identifying SHRS 1, and the recognition performance is misjudged as SHRS 1 and 3 when identifying SHRS 5 etc., while Fig. 14 shows the confusion matrix of communication radiator with the center frequency of 160 MHz in data set II. When identifying the same SHRS, there are a few differences in misjudged SHRS, which may be due to the different signal frequencies. However, the identification methods of different individual communication emitters can maintain excellent classification accuracy, which shows the methods we proposed to have better robustness performance.

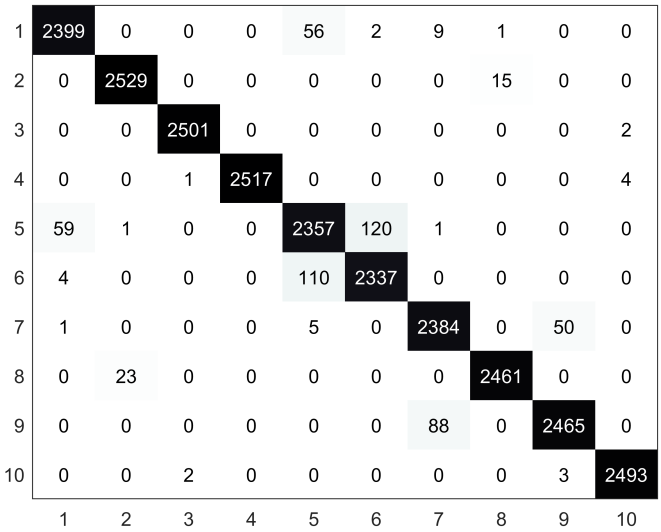


Fig. 14. Identification confusion matrix of 10 SHRS with 160 MHz center frequency

5. CONCLUSION

An individual identification method of communication emitters in complex electromagnetic environment based on system nonlinearity is proposed. In this method, the information entropy features based on VMD are extracted from the noise reduction signal, and the high-dimensional features are reduced by ESDA. We can clearly distinguish the features of each SHRS from the three-dimensional feature distribution map. Then, based on LightGBM, a classifier for individual identification of communication emitters is constructed. The experimental results show that the method achieves the identification of the same plant, same batch, and same model of radio under low SNR, whose performance is better than the state-of-art method. The recognition rate on the dataset is 98.5% based on the ten-fold cross validation. In the future, we intend to extend our method to other communication emitters and make corresponding improvements in order to improve the recognition rate and efficiency. It is also a good choice to combine our method with transfer learning or online learning, etc., which can make our identification method more real-time and transferable.

ACKNOWLEDGEMENTS

This work was supported by the Open-end Fund of Information Materials and Intelligent Sensing Laboratory of Anhui Province (IMIS202009), and by Anhui Agricultural University Introduction and Stabilization of Talents Research Funding Project (No. yj2020-74) and by the Natural Science Research Project of Colleges and Universities of Anhui Province (KJ2021A0182).

REFERENCES

[1] J. Dudczyk and A. Kawalec, "Specific emitter identification based on graphical representation of the distribution of radar signal parameters," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 63, no. 2, pp. 391–396, 2015.

[2] J. Gong, X. Xu, and Y. Lei, "Unsupervised specific emitter identification method using radio-frequency fingerprint embedded InfoGAN," *IEEE Trans. Inf. Forensic Secur.*, vol. 15, pp. 2898–2913, 2020.

[3] T.J. Bihl, K.W. Bauer, M.A. Temple, "Feature selection for RF fingerprinting with multiple discriminant analysis and using ZigBee device emissions," *IEEE Trans. Inf. Forensic Secur.*, vol. 11, no. 8, pp. 1862–1874, Aug. 2016.

[4] Y. Chen, Z. Wu, and Y. Lei, "Individual Identification of Radar Emitters Based on a One-Dimensional LeNet Neural Network," *Symmetry*, vol. 13, no. 7, p. 1215, 2021.

[5] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 1, pp. 94–104, 2016.

[6] Z. Xiao and Z. Yan, "Radar Emitter Identification Based on Novel Time-Frequency Spectrum and Convolutional Neural Network," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2634–2638, 2021.

[7] Q. Wu, C. Feres, D. Kuzmenko, D. Zhi, and Z. Yu, "Deep learning based RF fingerprinting for device identification and wireless security," *Electron. Lett.*, vol. 54, no. 24, pp. 1405–1407, Nov. 2018.

[8] N. Zhou, L. Luo, G. Sheng, and X. Jiang, "High accuracy insulation fault diagnosis method of power equipment based on power maximum likelihood estimation," *IEEE Trans. Power Deliv.*, vol. 34, no. 4, pp. 1291–1299, Aug. 2019.

[9] S. Udit, T. Nikita, B. Gagarin, and R. Barathram, "Specific Emitter Identification Based on Variational Mode Decomposition and Spectral Features in Single Hop and Relaying Scenarios," *IEEE Trans. Inf. Forensic Secur.*, vol. 14, no. 3, pp. 581–591, 2019.

[10] S. Liu, X. Yan, P. Li, X. Hao, and K. Wang, "Radar Emitter Recognition Based on SIFT Position and Scale Features," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 65, no. 12, pp. 2062–2066, Dec. 2018.

[11] J. Han, T. Zhang, Z. Qiu, and X. Zheng, "Communication emitter individual identification via 3D-Hilbert energy spectrum-based multiscale segmentation features," *Int. J. Commun. Syst.*, vol. 32, no. 1, p. e3833, Jan. 2019.

[12] Q. Zhang, Y. Guo, Z.Y. Song, "Dynamic Curve Fitting and BP Neural Network With Feature Extraction for Mobile Specific Emitter Identification," *IEEE Access*, vol. 9, pp. 33897–33910, 2021.

[13] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific Emitter Identification Based on Nonlinear Dynamical Characteristics," *Can. J. Electr. Comp. Eng.*, vol. 39, no. 1, pp. 34–41, 2016.

[14] L. Wu, Y. Zhao, Z. Wang, F.Y.O. Abdalla, and G. Ren, "Specific emitter identification using fractal features based on the box-counting dimension and variance dimension," in *2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2017, pp. 226–231.

[15] C.M. Jeong, Y.G. Jung, S.J. Lee, "Neural network-based radar signal classification system using probability moment and ApEn," *Soft Comput.*, vol. 22, no. 13, pp. 4205–4219, 2018.

[16] J. Lu and X. Xu, "Multiple-Antenna Emitters Identification Based on a Memoryless Power Amplifier Model," *Sensors*, vol. 19, no. 23, p. 5233, Dec. 2019.

[17] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[18] Pincus, Steve, "Approximate entropy (ApEn) as a complexity measure," *Chaos*, vol. 5, pp. 110–117, 1995.

[19] C.E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.

[20] P. Zhang, X. Tan, Y.U. Xuchu, and X. Wei, "Hyperspectral Imagery Feature Extraction Based on Kernel Semi-Supervised Discriminant Analysis," *J. Geomat. Sci. Technol.*, vol. 33, no. 3, pp. 258–262, 2016.

[21] L. Ma, M.M. Crawford, X. Yang, and Y. Guo, "Local-Manifold-Learning-Based Graph Construction for Semisupervised Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 53, no. 5, pp. 2832–2844, 2015.

[22] F. Dornaika, Y. El Traboulsi, "Matrix exponential based semi-supervised discriminant embedding for image classification," *Pattern Recognit.,* vol. 61, pp. 92–103, Jan. 2017.

[23] Z. Zhao, W. Qi, J. Han, Y. Zhang, and L.F. Bai, "Semi-supervised classification via discriminative sparse manifold regularization," *Signal Process.-Image Commun.*, vol. 47, pp. 207–217, Sep. 2016.

[24] Q. Meng, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," 2017.

[25] X. Zhang, X. Lu, R. Jia, and S. Kan, "Micro-seismic signal denoising method based on variational mode decomposition and energy entropy," *J. China Coal Soc.*, vol. 43, no. 2, pp. 356–363, 2018.

[26] M. Zhu, Z. Feng, X. Zhou, R. Xiao, and Y. Qi, "Specific Emitter Identification Based on Synchrosqueezing Transform for Civil Radar," *Electronics*, vol. 9, no. 4, p. 658, Apr. 2020.

[27] D. Sun, Y. Li, Y. Xu, and J. Hu, "A Novel Method for Specific Emitter Identification Based on Singular Spectrum Analysis," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.

[28] C. Deng, X. He, and J. Han, "Semi-supervised Discriminant Analysis," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on. IEEE*, 2007.

[29] H.S. Hu, D.Z. Feng, and Q.Y. Chen, "A novel dimensionality reduction method: Similarity order-preserving discriminant analysis," *Signal Process.*, vol. 182, p. 107933, May. 2021.

[30] X. Tao, C. Ren, Q. Li, W. Guo, R. Liu, and Q. He, "Bearing defect diagnosis based on semi-supervised kernel Local Fisher Discriminant Analysis using pseudo labels," *ISA Trans.*, vol. 110, pp. 394–412, Apr. 2021.

[31] T. Chen, and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *22nd ACM SIGKDD International Conference. ACM*, 2016.

[32] S. Guo and H. Tracey, "Discriminant Analysis for Radar Signal Classification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3134–3148, Aug. 2020.

[33] A. Jh, P.B. Hong, C. Jw, W. Y.D., "kNN-P: A kNN classifier optimized by P systems," *Theor. Comput. Sci. e*, vol. 817, pp. 55–65, 2020.