Specific Emitter Identification Based on Multichannel Depth Feature Fusion

Lin Tong,1,2 Mengqing Fang,1,2 Yulu Xu,1,2 Zhengcheng Peng,1,2 Weijie Zhu,1,2 and Ke Li1,2

1School of Information & Computer, Anhui Agricultural University, Hefei, Anhui 230036, China
2Anhui Provincial Engineering Laboratory for Beidou Precision Agriculture Information, Anhui Agricultural University, Hefei, Anhui 230036, China

Correspondence should be addressed to Ke Li; like@ahau.edu.cn

Received 28 April 2022; Revised 8 June 2022; Accepted 14 June 2022; Published 30 July 2022

Academic Editor: Kuruva Lakshmanna

Copyright © 2022 Lin Tong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Specific emitter identification involves extracting the fingerprint features that represent the individual differences of the emitter. This is achieved by processing the radio-frequency signals. Feature extraction and classifier selection are key factors that affect SEI performance. This paper proposes a deep convolutional neural network model based on multisignal feature fusion to identify the emitters. As part of the implementation model, the methods of singular spectrum analysis (SSA), variational mode decomposition (VMD), and intrinsic time-scale decomposition (ITD) are used to extract various signal features of emitter signals. Finally, a multichannel deep learning model is adopted to fuse each signal feature automatically and identify different signal emitters. Experimental results show that the proposed method completely considers the complementarity and independence of varying signal features and excavates hidden deep feature information. Hence, the process is considered reliable and effective.

1. Introduction

Specific emitter identification (SEI) is a method to identify and classify different emitters according to the subtle features of their signals. The subtle differences in the hardware structure of the different emitters will affect the signal when processed within the emitters [1]. In recent years, specific emitter identification technology has been widely applied in military and civilian fields. For example, we can use it in the military area to identify the equipment of the emitter of a friend or foe and the electronic reconnaissance technology [2]. And in the civil field, it can be used to monitor the security of wireless networks [3–6].

There are many types of research for obtaining specific emitter identification features [7]. The specific transmitter identification (SEI) technology is based on transient or steady-state signals according to the type of signal. It is difficult to collect the transient signal because the existence time is fleeting and has high requirements for the receivers. In 2019, Ali et al. [8] first applied Hilbert-Huang transform to RF fingerprint identification of Bluetooth device identification and improved detection of transient signals by time-frequency analysis, helping readers evaluate the availability of RF fingerprint Bluetooth signal in physical layer security of the wireless network. In 2020, Tian et al. [9] extracted the fusion features of instantaneous amplitude, frequency, and phase of the signal and higher order spectrum to obtain a more accurate RF fingerprint. In 2020, Baldini and Gentile [1] studied the application of the general linear chirplet transform and combined with the convolutional neural network to identify wireless devices in the Internet of Things. On the contrary, the steady-state signal is easier to collect. In 2017, Song et al. in [10] used natural time scale decomposition (ITD) to obtain the time-frequency energy distribution of signals, which overcomes the limitations of empirical model decomposition (EMD) and improves the accuracy and efficiency of time-frequency energy distribution (TFED) acquisition for nonlinear or nonstationary signals. In 2019, Han et al. [11] proposed a 3D-HESMS feature extraction method based
on 3D Hilbert spectrum. In 2018, Liu et al. [12] used scale-invariant feature transform (SIFT) to obtain the radar emitter's location and scale features. In 2020, Gok et al. [13] decomposed the envelope of radar pulses and the instantaneous frequency into components with low support in the frequency domain. They then obtained a set of features that characterize these compact signals for recognition. In the nonlinear aspect of the transmitter system, some methods do not distinguish between transient signal and steady signal. In 2018, Zhu and Gan [14] converted the instantaneous amplitude, phase, and frequency of the received signal into graphs using the visibility diagram and the horizontal visibility diagram. They then calculated the normalized aroma entropy using the corresponding degree distribution as the RF fingerprint. In 2019, Zheng et al. [15] proposed a channel-independent and robust physical layer recognition system based on function modeling, with transmitted data as input and sent RF signal as output.

In the aspect of information mining for high-dimensional data, a deep learning model can effectively reduce the dimension of high-dimensional data. It can also learn the practical information under its unique structure of neural network layer [16].

Deep learning approaches have shown superior performance in recognition and classification tasks [17]. In 2018, Ding et al. [18] constructed a convolutional neural network to recognize compressed bispectral features of signals. The experimental results showed that the observed recognition rate of 5 classification tasks could reach more than 0.9 under the condition of 15 dB SNR. In 2019, Pan et al. [19] constructed the deep residual network on the Hilbert spectrum of signal grayscale identification. In 2019, Yu et al. [20] formed a multichannel convolutional neural network model to sample the original signal at multiple time scales. In 2020, Gong et al. [21] constructed a generative adversarial network (GAN) model to recognize the bispectral gray histogram of the signal and then verified the high identification accuracy of the model under various channel conditions through experiments. However, the approaches mentioned above in the literature only use the deep learning model to identify and classify signals based on a single feature. Due to the limitation of single signal feature classification, a few recently proposed identifying and classifying emitters by integrating multiple signal features.

In this paper, a multichannel deep convolutional neural network model is constructed for the fusion and classification of multiple deep signal features. The proposed method greatly improves the overall SEI accuracy.

Next, this paper will be divided into three parts: the first part is the introduction of the singular spectrum analysis (SSA) feature extraction algorithm of the original signal, the component center frequency and bandwidth feature extraction algorithm based on Visual Molecular Dynamics (VMD), the signal component skewness and kurtosis feature extraction algorithm based on intrinsic time-scale decomposition (ITD), and the introduction of the model structure of the multichannel deep convolutional neural network. The second part is the process of several experiments and the comparison of experimental results. The third part is the summary of the model algorithm.

2. Proposed Method

2.1. Singular Spectrum Analysis (SSA). Singular spectrum analysis (SSA) can extract signals' period and trend features and convert the signals into independent time series [26].

The specific steps are as follows. Step 1: build the trajectory matrix. The length of the original signal is \( N \), and the length of the sliding window is \( L_p \), \( K_p = N - L_p + 1 \). Step 2: singular value decomposition. (1) Calculate the eigenvalue and eigenvector \( U \) of the trajectory matrix. (2) Calculate the left singular vector \( U \) and the right singular vector \( V \). \( V_i = X_i U_i / \sqrt{\lambda_i}, X_i = \sqrt{\lambda_i} U_i V_i^T \). Step 3: grouping. Separate the target signal components from other signal components. Step 4: averaging of the diagonal reconstruction signal. The corresponding singular vector was reconstructed according to the grouping results:

\[
\text{rca} = \sqrt{\lambda_i} U_i V_i^T, \quad i \text{ is the selected } r \text{ singular vectors} [27].
\]

For different emitter signals, the period and trend features extracted by the singular spectrum analysis method are various to distinguish signals from different emitters.

2.2. The Component Center Frequency and Bandwidth Features of Signals Based on Variational Modal Decomposition (VMD). VMD is a nonrecursive decomposition that determines the set of all modes and their respective center frequencies. On the other hand, these modes can reconstruct the original signal [28]. VMD is one of the most predominant and efficient decomposition techniques wherein a multicomponent signal is decomposed into series of subsignals which have very specific sparsity properties. This is achieved by assessing the bandwidth of a subsignal or mono component in an iteration process using an alternate direction method of multipliers (ADMM). VMD has enhanced performance in decomposition, excellent resistance to noise and stability which has immense applicability in feature extraction and fault diagnosis. This paper uses the VMD algorithm to transform the signal decomposition process into a variational framework. The signal decomposition is realized by searching for the optimal solution of the constrained variational model. In the iterative solution process of the variational model, the frequency center and bandwidth of each intrinsic mode function (IMF) component
are constantly updated. Assuming that the original signal is decomposed into K IMF components by VMD, the corresponding expression of the constrained variational model is shown in

$$\min_{\{u_k\}, \{w_k\}} \left\{ \sum_k \left\| \partial_t \left[ \delta(t) + \frac{j}{\pi t} * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\}, \tag{1}$$

s.t. $\sum_k u_k = f$

In formula (1), $u_k$ is K IMF components decomposed by VMD method, $w_k$ is the frequency center of each IMF component $u_k$, $\partial_t$ is the partial derivative of the time of the function, $\delta(t)$ is the unit impulse function, $j$ is the imaginary unit, and $*$ means convolution. Solving the optimal solution of the constraint variational model in formula (1), the signal’s frequency band is divided according to the frequency characteristics of the signal, and multiple narrowband IMF components are obtained. Then, their frequencies are arranged from low to high.

Figure 1 can clearly show that the features of each signal are significantly different through the three-dimensional coordinate diagram of the center frequency and bandwidth of each IMF component of the five transmitted signals. It can fully prove that the center frequency and bandwidth of the signal can be identified and extracted as the features of the signal.

2.3. The Skewness and Kurtosis Features of Signal Component Based on the Intrinsic Time-Scale Decomposition (ITD). ITD is a transformation method used to analyze nonlinear or nonstationary signals [29]. The baseline and rotation components of the signals are extracted by defining the piecewise linear baseline extraction operator $\xi$. The process of decomposition calculation process of the original signal $x(t)$ by ITD method is as follows:

$$x(t) = L\xi x(t) + (1 - \xi)x(t) = L(t) + H(t) = L(t) + H(t). \tag{2}$$

In formula (2), $L(t) = \xi x(t)$ represents the baseline signal component of original signal $x(t)$, and $H(t) = (1 - \xi)x(t)$ represents the rotation signal component of original signal $x(t)$.

In the literature [29], the author puts forward that the skewness and kurtosis of each signal component after the ITD of the signal reflected the non-Gaussian feature information of the signal and could be used as the signal feature. Through experiments, we also extracted each signal component’s kurtosis and skewness features, paving the way for the subsequent training of the multichannel feature fusion model. Skewness refers to distortion that leads to deviation from the symmetrical bell curve or normal distribution for a data set. Kurtosis refers to the sharpness of the peak of the frequency distribution curve. This measured by combining the weight of a distribution tail with respect to the center of the distribution.

2.4. Multichannel Deep Convolutional Neural Network Model (MC-DCNN). This paper’s multichannel deep convolutional neural network model has multiple parallel input channels. Each input channel is composed of a convolutional layer and a pooling layer, and it is used to extract the secondary features of a signal feature of the input [30].
Figure 2 shows the multichannel deep convolutional neural network model structure. In MC-DCCN model, single dimension of multivariate time series is fed into each channel as input which learns the features individually. The MC-DCCN model aggregates the learnt features from each channel and further feeds them into a multilayer perceptron to perform the classification. The estimation of parameters is done utilizing the gradient-based method which helps to train the MC-DCNN model. The multichannels in MC-DCNN are able to learn the feature representations for each univariate time series automatically. The traditional MLP is then used to combine all these features for enhanced representation of each class. This type of feature learning and feature combining approach helps in improvement of the classification performance for multivariate time series [31, 32]. The SSA features of the original emitter signal, the component center frequency and bandwidth features of signals based on VMD, and the skewness and kurtosis features of the signal components based on the ITD are input through three different independent channels. And the secondary features obtained after the convolutional pooling operation are fused into the fully connected neural network. The backpropagation algorithm trains the whole model. Furthermore, the classification results are output by the softmax layer.

The secondary features extracted from convolutional pooling are the deeper internal signal features, and the deep convolutional neural network can mine deeper data information.

3. Experimental Results and Discussion

3.1. Signal Acquisition. In this experiment, five Kenwood models of short-wave hand-held radio stations are used as identification objects, and the communication frequency is 160 MHz. The signal acquisition equipment is Rhode Schwartz FSL3, the sampling frequency band is set to 159.75-160.25 MHz, the sampling rate is 7 kHz, and the RF signal is converted to 70 kHz IF. The sampled data are stored in hex data format on the host computer, which is a Lenovo Xiaoxin Air 14IKBR, equipped with an i7-8550U processor and 16G memory. The signal data acquisition process is shown in Figure 3.

3.2. Feature Extraction of Original Signal. First, we extract three signal features of the emitter signal: the periodic and trend features of SSA, the component center frequency and bandwidth features of signals based on VMD, and the skewness and kurtosis feature of the signal component based on the ITD. We divide the original signals collected by each emitter into 5000 signal segments according to the time length, which are used as 5000 signal samples of each emitter. Among them, 3500 samples are used as training set samples of each emitter, and 1500 samples are used as test set samples. In the process of model training, the 3500 training set samples are further divided into training set and verification set, and the ratio of the two is 8:2. Signal features of the same dimensional size are extracted for each signal sample of each station. Among them, the SSA features are in dimension $1 \times 1228$, the component center frequency and bandwidth features of signals based on VMD are in dimension $1 \times 70$, and the skewness and kurtosis features of the signal component based on the ITD are in dimension $1 \times 1228$.

3.3. Multichannel Deep Convolutional Neural Network Model Construction. In this paper, we construct a deep convolutional neural network model with three parallel convolutional modules as input channels for signal features.
convolution kernel size of the one-dimensional convolution layer of the convolution module corresponding to the periodic and trend features of SSA is used. The size of the convolution kernel of the one-dimensional convolution layer of the convolution module based on the center frequency and bandwidth features of the VMD signal is set to $1 \times 25$ and equal-width convolution is adopted. The size of the pooling kernel of the one-dimensional pooling layer is set as $1 \times 2$, and the average pooling method is used for pooling. The size of the convolution kernel of the one-dimensional convolution layer of the convolution module corresponding to the skewness and kurtosis features of the signal component of the original signal based on the ITD is set as $1 \times 3$, and the method of constant width convolution is adopted. The size of the pooling kernel of the one-dimensional pooling layer is set as $1 \times 2$, and the average pooling method is used for pooling.

The fully connected layer of the model is an ordinary fully connected feedforward neural network structure, and the activation functions are all tanh functions. The overall model uses the categorical cross-entropy function as the loss function and the Adadelta function as the optimization function.

### 3.4. The Effect of Single and Multifeature Fusion on Identification Accuracy.

To study whether emitter recognition with multifeature fusion can effectively improve the recognition accuracy, compared with the single-feature deep learning method, we construct single-channel, two-channel, and three-channel deep convolutional neural network models for experiments. Figure 4 shows the experimental results of different channel numbers of the model when the SNR = 15 dB, and the number of emitters to be identified is 5.

The recognition accuracy of a single feature based on SSA, VMD, and ITD is 91.25%, 94.12%, and 85.37%, respectively. Based on the fusion of two features, such as SSA and ITD, VMD and ITD, and SSA and VMD, their recognition accuracy is 92.34%, 96.07%, and 96.28%. Moreover, recognition accuracy based on the fusion of the three features is 99.11%. The above experimental results show that the recognition accuracy based on the fusion of the two features is slightly higher than that of the corresponding single feature. The recognition accuracy based on multifeature fusion is higher than that of multifeature fusion of two-feature fusion. Therefore, the multifeature fusion method can effectively improve the accuracy of transmitter recognition. This also justifies the fact that the increase in channel numbers in the convolutional layers leads to enhanced performance especially in accuracy. Similar results have been seen in previous studies as well wherein CNN models have been proposed with two convolutional layers and more number of channels in association with two pooling layers. It was observed that the CNN model with highest number of channels in the convolutional layers yielded better accuracy, precision, sensitivity, and F-score [33].

### 3.5. Influence of Different Convolution Layers on Identification Accuracy.

This paper conducts several experiments under different model parameters to explore the influence of different parameters in the deep learning model on identification accuracy. Among them, the input features of each experiment are the multiple signal features mentioned above, the number of emitters to be identified is 5, and the SNR is 15 dB. We set different convolution layers for the
The number of convolution kernels of each convolution layer is set as 1, and the experimental result is shown in Figure 5.

As shown in Figure 5, when the number of convolution layers is set as 2–5 layers, compared to the one layer convolution layer, the identification accuracy of the model is slightly improved. Moreover, when the number of convolution layers is 2–5, the fluctuation range of model recognition accuracy is small. However, due to the increase in the number of convolution layers, the model structure becomes jumbled, which gives rise to a longer training time for the model.

Therefore, considering the model’s identification accuracy and training time, when the number of layers of the convolution layer is set as 2, the overall model performs best.

In addition, we compare the average pooling method with the maximum pooling method in the setting of the pooling layer method. Under the condition of SNR = 15 dB, experimental results of the two pooling methods are shown in Figure 6.

Figure 6 shows that when the average pooling method is adopted in testing identification accuracy, the experimental identification rate is slightly higher for the average pooling method than the maximum pooling method. For this reason, we choose to use the average pooling method.

In the full-connection layer neural network, we set up some dropout layers. Figure 7 shows the change of experimental identification accuracy when the dropout layer is set to different parameters.

The results shown in Figure 7 show that in different dropout layers, when the dropout layer parameter is set to 0.3, the recognition accuracy of the model is the highest, reaching 89%. Therefore, we set the dropout layer ratio at 0.3.
Figure 8 shows that when the learning rate is set at 0.001 and 0.2, the identification accuracy of the model is lower. When the learning rate is set between 0.01 and 0.15, the identification accuracy of the model is higher overall, and the fluctuation range is small. Especially when the learning rate is 0.1, the recognition accuracy reaches the highest.

Table 1 records the change in model identification accuracy when SNR = 15 dB, and the different numbers of convolution kernels are used in the convolution layers.

<table>
<thead>
<tr>
<th>The number of convolution kernels</th>
<th>The number of convolution layers</th>
<th>Identification accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>89%</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>87%</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>86%</td>
</tr>
</tbody>
</table>

Figure 8 shows that when the learning rate is set at 0.001 and 0.2, the identification accuracy of the model is lower. When the learning rate is set between 0.01 and 0.15, the identification accuracy of the model is higher overall, and the fluctuation range is small. Especially when the learning rate is 0.1, the recognition accuracy reaches the highest.

Table 1 records the change in model identification accuracy when SNR = 15 dB, and the different numbers of convolution kernels are used in the convolution layers.

As shown in Table 1, when the number of convolution kernels of the convolution layers is set to 1, the model obtains the highest identification accuracy.

3.6. Identification Accuracy Obtained by Different Experimental Methods. The methods in the literatures [19–21, 23, 25] are deep residual network, multisampling convolutional neural network, long short-term memory (LSTM), recurrent neural network, and deep ensemble learning. These techniques are compared with the approach in this paper. Under different SNR conditions, the number of emitters to be identified is 5, and the average identification accuracy of different methods is shown in Figure 9.

We can see from Figure 9 that the recognition accuracy of all methods increases with the increase of SNR. ResNet method is worse than other methods when SNR is lower than 15 dB, but its recognition accuracy is improved when SNR is higher than 15 dB. Among them, the MDCNN method is generally superior to other methods.
Table 2 shows the recognition time required by different methods when the SNR is set to 15 dB. Although the recognition speed of MRNN is faster than that of MDCNN, the recognition accuracy of MDCNN is much higher than that of MRNN. The MRNN helps in computing derivatives for gradient-based learning methods automatically. Therefore, to sum up, we choose the MDCNN method.

3.7. Identification Accuracy Obtained from Different Numbers of Radiation Sources. When the number of emitters to be identified is 3, 5, and 10, respectively, the identification accuracy of the model is shown in Table 3.

From Table 3, we can see that with the increase of the number of signal emitters, the accuracy of their identification is decreasing. In this paper, we use five signal transmitters to collect signals, and the recognition accuracy can reach 91%.

### 4. Conclusion

The specific emitter identification method proposed in this paper firstly extracts multiple signal features of the radiation source signal, namely, the periodicity and trend features of SSA, the component center frequency and bandwidth features based on VMD, and the skewness and kurtosis features of the signal component based on ITD. Then, multiple signal features are fed into the multichannel deep convolutional neural network model for recognition. Multiple experiments reveal the fact that the multichannel depth feature fusion method completely considers the independence and complementarity of different signal features. The proposed method implements a convolutional neural network to extract the depth information of each signal feature and thereby improves the experimental value of each signal feature. Thus, it yields good performance in identifying specific radiation sources for field data sets. Due to constraints of limited resources, the study limits collection of the transmitted signals from five signal transmitters. The training of the deep learning models still takes long time. Thus, the future scope of study includes increase in the number of signal transmitters, collection of more signals, and implementation of new deep learning models to improve the identification accuracy of radio-frequency fingerprints of communication radiation sources. This could also reduce the time and cost of the process involved in identification.

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSA</td>
<td>Singular spectrum analysis</td>
</tr>
<tr>
<td>VMD</td>
<td>Variational modal decomposition</td>
</tr>
<tr>
<td>IMF</td>
<td>Intrinsic mode function</td>
</tr>
<tr>
<td>ITD</td>
<td>Intrinsic time-scale decomposition</td>
</tr>
<tr>
<td>SEI</td>
<td>Specific emitter identification</td>
</tr>
<tr>
<td>MC-DCNN</td>
<td>Multichannel deep convolutional neural network model [34]</td>
</tr>
</tbody>
</table>
**Data Availability**

The data in this paper are not intended to be shared. The main reasons are as follows: first, the experimental data set in this paper comes from national military equipment, which is more private and should not be shared. Secondly, due to an unavoidable delay between paper submission and paper publication, there is a risk that others will copy the data in this paper after sharing.

**Conflicts of Interest**

There is no competing interest in this paper.

**Authors’ Contributions**

Lin Tong carried out the construction of the experimental model and wrote the main article. Mengqing Fang carried out some data analysis and wrote some papers. Yulu Xu updated the drawing of the graphs in the paper, and Zhengcheng Peng updated the algorithm used in our paper and participated in the writing of the paper. Weijie Zhu analyzed the realization data of the paper and carried out some paper polishing work. Ke Li collected the experimental data and designed the overall experimental scheme.

**Acknowledgments**

This work is supported by the Open Fund of Information Materials and Intelligent Sensing Laboratory of Anhui Province (No. IMIS202009), the Anhui Agricultural University Introduction and Stabilization of Talents Research Funding (No. yj2020-74), and the Natural Science Research Key Project of Colleges and Universities in Anhui Province (No. KJ2021A0182). These three funding projects provide a good study environment and experimental equipment for this experiment.

**References**


