Abstract—Specific emitter identification (SEI) is a highly potential technology for physical-layer authentication that is one of the most critical supplements for the upper-layer authentication. SEI is based on radio frequency (RF) features from circuit difference, rather than cryptography. These features are inherent characteristics of hardware circuits, which are difficult to counterfeit. Recently, various deep learning (DL)-based conventional SEI methods have been proposed, and achieved advanced performances. However, these methods are proposed for close-set scenarios with massive RF signal samples for training, and they generally have poor performance under the condition of limited training samples. Thus, we focus on few-shot SEI (FS-SEI) for aircraft identification via automatic dependent surveillance-broadcast (ADS-B) signals, and a novel FS-SEI method is proposed, based on deep metric ensemble learning (DMEL). Specifically, the proposed method consists of feature embedding and classification. The former is based on metric learning with a complex-valued convolutional neural network (CVCNN) for extracting discriminative features with compact intracategory distance and separable intercategory distance, while the latter is realized by an ensemble classifier. Simulation results show that if the number of samples per category is more than 5, the average accuracy of our proposed method is higher than 98%. Moreover, feature visualization demonstrates the advantages of our proposed method in both discriminability and generalization. The code and the dataset can be downloaded from https://github.com/BeechburgPieStar/FS-SEI.

Index Terms—Automatic dependent surveillance-broadcast (ADS-B), ensemble learning, few-shot specific emitter identification (SEI), metric learning, SEI.

I. INTRODUCTION

Wireless communication is rapidly developed for realizing the great vision of the Internet of everything. Meanwhile, there are various unprecedented challenges, especially the security, because of the open nature of wireless channel [1]–[3]. Authentication is one the most critical techniques for the secure wireless communications, and the existing authentication methods rely on the cryptography-based authentication mechanisms above the physical layer [4], [5]. However, since the inherent defects of these cryptography-based authentication mechanisms, such as vulnerable to spoofing, forgery, replay, eavesdropping, and reflection attacks [6], [7], the requirement of additional communication overhead and high complexity [8], they may not be applicable for emerging wireless communication systems.

Therefore, physical-layer authentication, as an effective supplement for the upper-layer authentication, is widely concerned [9]. In this article, we mainly focus on a passive physical-layer authentication technology, named as specific emitter identification (SEI), which is a greatly potential technology in both military and civilian scenarios. SEI relies on radio frequency (RF) features, usually parasitic in transmitted signals and originate from hardware differences from circuits [10], for the identification of different transmitters. Due to that the hardware difference is an inherent attribute in the process of circuit design or circuit manufacturing, RF features are tamper-resistant and difficult to counterfeit, which is also the basis of SEI.

The typical RF feature-based SEI methods generally consist of preprocessing, feature extraction, and identification [11]. The preprocessing usually contains various operations, including filtering, power normalization, synchronization, target signal interception, and so on [12], while the feature extraction and the identification are the key steps of SEI. Before the development of deep learning (DL), classical SEI methods mainly design features through statistical methods, and then machine learning-based classifiers are applied for the identification. However, there are many disadvantages of the RF features designed by artificial knowledge and experience, such as the weak performances, the poor environmental adaptability, and the lack of capabilities to deal with increasingly complex and increasing wireless transmitters.

In recent years, DL has been proved to have the powerful and effective data analysis capability [13], it has been widely applied into wireless communication technologies [14]–[16], [18], including SEI [11], [19]–[30].
These DL-based SEI methods relied on massive historical RF signal samples and deep neural networks for extracting more robust and effective RF features, which have been demonstrated to have better performances than artificial feature-based methods.

However, most of the DL-based SEI methods focused on the close-set SEI problem [11], [19]–[30], which will be specifically introduced in Section II. In other words, the transmitter categories are the same in the stages of offline training and online deployment, but it is invalid in actual applications. Because it is scarcely possible to collect signals samples from all categories of transmitters for training. Although this problem can be solved by finetuning, it is constraint by few training signal samples of new categories and limited retraining time in the process of online deployment. Hence, we focus on the few-shot SEI (FS-SEI) for aircraft identification task by automatic dependent surveillance-broadcast (ADS-B) signals, which are applied for aircraft monitoring [31]. There are massive ADS-B signal samples from multiple aircrafts as auxiliary data set for the offline training, and few ADS-B signal samples from new aircrafts as few-shot training samples in the stage of the online deployment. In this article, a deep metric ensemble learning (DMEL)-based SEI method is proposed, consisting of feature embedding based on metric learning and complex-valued convolutional neural network (CVCNN) and ensemble learning-based classifier.

Specifically, the former aims to extract discriminative features from ADS-B signal samples by combining Softmax loss and two metric-based contrastive losses (including triplet loss and center loss). Different from the separable features extracted by the conventional DL methods with Softmax loss, discriminative features are more generalized, which is not only applicable to the ADS-B signal samples from seen categories, but also suitable to the ADS-B signal samples from unseen categories or categories that are rarely seen. In addition, based on the discriminative features, the identification can be easy, but we also introduce ensemble learning to combine multiple machine learning classifiers for further enhancing the identification performance. The main contributions of this article are listed as follows.

1) We introduce the paradigm of “constructing feature embedding offline and establishing classifier online” for FS-SEI. In the offline training phase, feature embedding is constructed by CVCNN and massive historical ADS-B signal samples, because the feature embedding usually contains large numbers of hyper-parameters, which requires to be trained by massive samples. In the online deployment phase, a simple classifier with few hyper-parameters or even no hyper-parameters can be established by limited ADS-B samples from new categories of aircrafts.

2) We propose a metric learning-based feature embedding scheme for extracting generalized and discriminative features. Here, we combine two contrastive losses and Softmax loss to ensuring compact intracategory distance and separable intercategory distance in the feature space.

3) We use a simple machine learning classifier followed by the above good feature embedding to realize identification, and ensemble learning is introduced for further improving identification performance.

II. RELATED WORKS

In this section, conventional artificial features and DL-based SEI methods and four few-shot learning (FSL) algorithms with their applications are introduced.

A. Artificial RF Feature-Based SEI Methods

Artificial RF feature-based SEI methods are usually based on artificial RF features and machine learning-based classifiers, and there are various RF features for SEI, which can be mainly summarized as instantaneous features, modulation domain features, and transform domain features, which are introduced as follows.

The first kind of RF features is the instantaneous statistics of RF signal, and it usually contains the mean, variance, skewness, and kurtosis of amplitude, phase, and frequency about signals or signals reconstructed by variational mode decomposition (VMD) [32]. Next, modulation domain features are nonlinear characteristics introduced in the modulation process, including in-phase and quadrature (IQ) imbalance [33], modulation shape [34], constellation error [35], and so on. Finally, transform domain features are the features based on various transforms, including wavelet transform [36], Hilbert–Huang transform [37], short-time Fourier transform [38], and so on.

B. DL-Based SEI Methods

DL-based SEI methods apply neural networks, such as convolutional neural network (CNN), recurrent neural network (RNN), generative adversarial network (GAN), for joint feature extraction and classification. Here, the DL-based SEI methods can be mainly summarized into three categories by sample format, i.e., time domain signal, spectrogram, and constellation.

1) Time Domain Signal-Based SEI Methods: Time domain signals refer to raw IQ signals or signal components decomposed by empirical mode decomposition (EMD), VMD, and so on. Merchant et al. [19] first introduced a multilayer CNN for SEI based on seven ZigBee devices and their raw IQ signal samples, which demonstrated that CNN can effectively distinguish different transmitters, even for different devices with the same type and the same model. Similarly, Yu et al. [20] also proposed a multisampling CNN (MSCNN) for SEI based on raw IQ signal samples, using multiple downsampling transformations for joint multiscale feature extraction and classification. It is worth noting that they applied 54 ZigBee devices for simulations, and their proposed MSCNN has achieved 97% accuracy at high SNR, which illustrated the effectiveness of DL in SEI.

Different from the above CNN-based SEI methods, Wang et al. [21] proposed an LSTM and raw IQ sample-based SEI method, but they focused on the problem that the identification performance of SEI declines with time, and they introduced transfer learning for solving this problem. He and Wang [11] also proposed an SEI method based on
Specifically, the former used the bispectrum as the sample supervised and unsupervised SEI methods, respectively. Xie et al. [21] also revealed, and the estimated CFO is introduced to mod-ify the identification result, so as to avoid the performance degradation. The difference is that the former applied the real-valued CNN and the optimization method is alternating direction method of multipliers (ADMMs), while the latter focused on the CVCNN and it is optimized by proximal gradient descent. Simulation results demonstrated that there is the redundance in model parameters of the previous proposed DL-based SEI methods.

2) Spectrogram-Based SEI Methods: In these spectrogram-based SEI methods, bispectrum analysis, and Fourier transform are the common solutions to transform signals into spectrograms. Ding et al. [24] first applied compressed bispectrums by the supervised dimensionality reduction method as samples, and they used a simple four-layer CNN as identification method, which has achieved nearly 100% accuracy at high SNR. Shen et al. [25] proposed a short-time Fourier transform-based spectrogram and CNN-based SEI method, which has been demonstrated to have the better identification performance than IQ-based methods and fast Fourier transform-based methods. In addition, the influence of carrier frequency offset (CFO) on the identification performance is also revealed, and the estimated CFO is introduced to modify the identification result, so as to avoid the performance degradation.

Different from the previous supervised methods, Xie et al. [26] and Gong et al. [27] proposed the semi-supervised and unsupervised SEI methods, respectively. Specifically, the former used the bispectrum as the sample and virtual adversarial training-based CNN for realizing semi-supervised SEI; The latter is based on information maximized GAN (InfoGAN) with the priori information of the wireless channels, and the input for this method is the gray histogram constructed by the bispectrum, which is to enhance the discriminability between different transmitters.

3) Constellation-Based SEI Methods: Peng et al. [28] designed a novel differential constellation trace figure (DCTF) as the input of CNN, and their simulation results demonstrated that these signal representation methods are useful for SEI. Based on [28], Yin et al. [29] proposed a multichannel CNN (MCCNN)-based SEI method for wireless terminal authentication, which used multipart DCTFs as training and test samples. In detail, a set of multipart DCTFs is composed of three DCTFs from the transient-on, modulation, and transient-off parts, respectively. In the training and test phase, these three DCTFs are fed into three independent CNNs, and then the features extracted from these CNNs are concatenated for classification. This simulation is based on six long-term evolution (LTE) mobile phones and their simulation results indicated that multipart DCTF is better than single-part DCTF.

Except DCTF, Peng et al. [30] introduced a novel heat constellation trace figure (HCTF) for SEI. Specifically, the principle of HCTF is that the distribution density of different regions is calculated on the basis of the common constellation trace figure, and the regions with different densities are given different colors. The simulation results demonstrated that HCTF with Inception v3 has the advanced and robust performance under various channel conditions.

The reason why the above DL-based SEI schemes can achieve advanced performance is not only their excellent DL algorithms, but also the large number of RF signal samples. Gong et al. [27] revealed the relationship between the number of training samples and the identification performance in their article, i.e., the DL-based SEI method need sufficient samples, otherwise its performance is far worse than the traditional feature-based methods. Specifically, if there are 500 samples per device for training, the LSTM and ITD-based SEI method can have 96.70% accuracy, but the accuracy will decline to 65.50%, if there are only ten samples per device for training. The detailed information is given in Table I. Clearly, the number of the training samples per category are usually

| Conventional methods | Model | Transmitter type | Data format | Number of samples |
|----------------------|-------|-----------------|-------------|-------------------|
| K. Merchant et al. [19] | CNN | 7 Zigbee devices | Raw IQ | 119,000 (90% for training and validation) |
| J. Yu et al. [20] | MS CNN | 54 Zigbee devices | Raw IQ | 62,860 (80% for training and validation) |
| X. Wang et al. [21] | LSTM | 4 RF devices | Raw IQ | 12,000 (training)/250 (transferring) |
| B. He and F. Wang [11] | LSTM | 5 Emitters | Signal components | 500 samples per device for training |
| T. Jian et al. [22] | ResNet | 500 WIFI devices and 50 ADS-B transmitters | Raw IQ | 218 transmissions per device |
| X. Wang et al. [23] | CVCNN | 7 Power amplitudes | Raw IQ | 262,000 samples (70% for training) |
| L. Ding et al. [24] | CNN | 3 USRPs | Bispectrum | 3,000 samples (50% for training) |
| L. Ding et al. [25] | CNN | 20 LoRa devices | Bispectrum | 1,000 packets per device for training |
| C. Xie et al. [26] | CNN | 6 USRPs | Bispectrum | 10,000 unlabeled samples per category and 200−1200 labeled samples per category |
| J. Gong et al. [27] | InfoGAN | 5 routers | Gray histogram | 50,000 samples |
| L. Peng et al. [28] | CNN | 54 Zigbee devices | DCTF | 8,262 samples |
| F. Yin et al. [29] | MCCNN | 6 LTE mobile phones | Multi-part DCTF | 2,340 samples (80% for training) |
| Y. Peng et al. [30] | CNN | 7 Power amplitudes | HCTF | 310−410 samples per device |
hundreds, thousands, or even tens of thousands. If there are not enough samples in the above methods, they may hardly achieve these excellent results.

C. Typical FSL Methods and Their Applications

There are four typical FSL methods, including data augmentation, generative model, metric model, and meta model [39], which are specifically introduced below. In addition, the applications are only related to the signal identification.

1) Data Augmentation: Data augmentation aims to expand the few-shot training data set manually by making limited samples produce more equivalent samples, and it is an effective method to overcome the lack of training samples, but no matter what kind of expansion schemes, it will bring the problem of sample noise. Zhou et al. [40] proposed a data union augmentation method for modulation signal classification. In detail, authors applied GAN to generate samples, and then selected the generated samples with high similarity with the original samples for reducing the impact of sample noise. In addition, the samples are further expanded by spatiotemporal flip. Simulation results demonstrate that this method can improve classification performance under the condition of limited samples.

It is noted that most of data augmentation methods are only based on few-shot training data sets from their tasks to expand samples, but other three schemes will introduce auxiliary data sets from various auxiliary tasks. Due to the additional information brought by the auxiliary data sets, the performance of the other three schemes is usually better than that of data augmentation, but data augmentation can be a trick to be added into other FSL methods.

2) Generative Model: Generative methods apply autoencoder or GAN to extract robust features, which can represent and recover the corresponding samples. For instance, Wang et al. [41] used a feature Wasserstein GAN for RF-based human activity recognition under the few shot condition, which can recognize new activities with high accuracy using few samples.

3) Metric Model: Metric methods are based on the ideal of “learn to compare” or “learn to measure,” which aims to compare or measure the similarity between different samples. One of the most famous metric methods is the Siamese network (SiameseNet) [42]. SiameseNet is a coupling framework based on two weight-sharing subneural networks. Paired samples are, respectively, fed into subneural networks for their corresponding features, and then their similarity can be compared by calculating the distance between features, such as the Euclidean distance. The detailed structure of SiameseNet is given in Fig. 1.

SiameseNet can effectively distinguish the sample from one category from other categories. Wu et al. [43] and Sun [44] introduced it into SEI for the tasks of open-set SEI and FS-SEI, respectively. In addition, Zhang et al. [45] and Man et al. [46] applied the relation network, which is an improved SiameseNet and uses a fully connected neural network to judge the degree of similarity, for signal modulation recognition and zero-shot SEI.

4) Meta Model: The basic ideal of meta methods is “learn to learn.” The essence of meta learning is an optimization method to increase the generalization of neural networks in the multitasks, and its purpose is to adapt new tasks on the basis of acquiring existing knowledge. Yang et al. [47] proposed a model-agnostic meta learning (MAML)-based SEI method, and their simulation on the identification of different types of transmitters demonstrated its ultrahigh performance under the condition of limited samples. Dong et al. [48] also applied MAML with complex-valued attention for modulation recognition, which also achieved amazing recognition performance.

Compared with metric methods, meta methods, especially MAML, has the disadvantages of high training complexity and unstable gradient [49], because of their complex training process. Due to the training processing of the metric methods is simple, their implementation is not much different from DL, and these methods just need to modify the loss function. Thus, we focus on the metric methods in this article.

III. SYSTEM MODEL, DATA SET GENERATION, AND PROBLEM DESCRIPTION

A. System Model

In this article, the SEI technology is applied to identify aircrafts based on ADS-B signals. The SEI system for different aircraft identity authentication is shown in Fig. 2, and there are three steps: 1) data collection; 2) model training; and 3) model deployment. In general, data collection is conducted in the previous airspaces, and SEI models are trained on the historical ADS-B signal samples from these airspaces, but model deployment is usually conducted in a new airspace.

Considering that there are new aircrafts in the new airspace that have never appeared in the previous airspaces, it is difficult to directly deploy the SEI model based on historical ADS-B signal samples from the previous airspace in the new airspace. In detail, there are at least two problems for the quick deployment: 1) It is scarcely possible to collect enough supervised ADS-B signal samples from new aircrafts for the SEI model updating in a short time and 2) even if ADS-B signal samples are sufficient, it is difficult to quickly update the SEI model for identifying new aircrafts. Therefore, under the condition of limited ADS signal samples, FSL is introduced for rapid update and real-time deployment of SEI.
B. ADS-B Signal Data Set

The received ADS-B signal model can be written as follows:

\[ x(t) = h(t) \ast s(t) + n(t) \]  

where \( x(t) \) is the received signal; \( h(t) \) and \( n(t) \) represent wireless channel and noise, respectively; \( s(t) \) is the transmitted ADS-B signal, which consists of the preamble and the pulse position modulation-based data block. There are four pulses with fixed positions and 0.5 ± 0.05 \( \mu \) s duration per pulse in the preamble, and the duration of the data block is 112 \( \mu \) s.

The ADS-B signal collector consists of Signal Hound SM200B and 1090-MHz omnidirectional antenna. In addition, high-performance computing terminal is applied to detect ADS-B signal from original signal data, and construct the ADS-B signal data set. Specifically, the data block contains aircraft identification code, position, altitude, speed, and so on. So, sample category (i.e., aircraft identification code) can be obtained by the demodulation of the data block. Moreover, the preprocessed ADS-B signals are as IQ samples. The more detailed information of the ADS-B signal data set has been open-source, which can refer to [50].

C. Problem Description

Here, \( x \) represents the input sample, which is a ADS-B signal sample from one aircraft with IQ component format; \( y \) represents the real category of the corresponding aircraft; \( D = \{x_n, y_n\}_{n=1}^N \) denotes the data set containing ADS-B signal samples and their corresponding categories. In addition, \( x_n \) comes from a specific domain \( D = \{X, P_X\} \), where \( X \) \((x_n \in X)\) is the sample space and \( P_X \) is the marginal probability distribution of the sample space. Moreover, we use \( Y \) \((y_n \in Y)\) to represent the category space, and \( P_X \times Y \) to represent the joint probability distribution of the sample space and the category space.

1) SEI Problem: The SEI problem can be defined as a maximum-a-posteriori (MAP) criterion-based multiclass pattern recognition problem, which can be written as follows:

\[ \hat{y} = \arg \max_{y \in Y} f_{SEI}(y|x; W) \]  

where \( \hat{y} = f_{SEI}(x; W) \) is the predicted category, and \( f_{SEI} \) and \( W \) are the mapping function and the set of hyper-parameters for SEI, respectively. When the mapping function is fixed, the goal of SEI is to find a set of suitable hyper-parameters \( W \) such that it can realize the mapping from the sample space to the category space, i.e., \( f_{SEI}(W) : X \to Y \). \( W \) belongs to the hyper-parameter space \( W \), and it can minimize the expected error \( \varepsilon \), i.e.,

\[ \min_{W \in W} \varepsilon_{ex} = \min_{W \in W} \mathbb{E}(x,y) \cdot p_{X \times Y}(\hat{y}, y) \]  

where \( L(\cdot) \) is to measure the difference between the predicted category and the ground-truth category. However, \( P_{X \times Y} \) is usually unknown, and thus DL algorithms generally minimize the empirical error \( \varepsilon_{em} \) as the replacement of \( \varepsilon_{ex} \), which can be written as follows:

\[ \min_{W \in W} \varepsilon_{em} = \min_{W \in W} \mathbb{E}(x,y) \cdot d(\hat{y}, y). \]  

Moreover, the generalization error between \( \varepsilon_{ex} \) and \( \varepsilon_{em} \) is \( \varepsilon_{ge} = |\varepsilon_{em} - \varepsilon_{ex}| \). If the hyper-parameter space \( W \) to be searched is too huge, \( \varepsilon_{ge} \) will be large, and there is a high probability of overfitting. In order to minimize \( \varepsilon_{ge} \), many constraints will be imposed on \( W \) to narrow the hyper-parameter space \( W \), and thus the above problem can be written as follows:

\[ \min_{W \in W} \varepsilon_{em} \quad \text{s.t.} \quad f_{SEI}(x_n; W) = y_n, \forall (x_n, y_n) \in D. \]  

2) FS-SEI Problem: The FS-SEI problem is how to train an excellent SEI method with few ADS-B signal training samples. This problem can be described by a set of few-shot ADS-B signal data sets, consisting of training data set and test data set, i.e., \( D_{tr} = \{D_{tr1}, D_{tr2}\} \). \( D_{tr} = \{(x_i, y_i)\}_{i=1}^{N_{tr}} \) contains \( C \) categories with \( K \) samples per category, where \( N_{tr} = C \times K \) and \( K \) is usually small (for instance, \( K = 1 \) or \( 5 \)), while there are sufficient samples of the same category as \( D_{tr} \) in \( D_{te} = \{(x_i, y_i)\}_{i=1}^{N_{te}} \), where \( N_{te} \gg N_{tr} \). In addition, it is noted that \( x_t, x_j \in X_{fs} \subset X \) and \( y_i \in Y_{fs} \subset Y \), where \( X_{fs} \) and \( Y_{fs} \) are the few-shot sample space and the few-shot category space.

However, according to (5), it is almost impossible to construct an effective \( f_{SEI}(W) \) only based on \( D_{te} \) with extremely limited constraints. Thus, an auxiliary data set \( D_{au} \) is generally introduced for narrowing the hyper-parameter space \( W \) and achieving generalized model. In general, \( D_{au} = \{(x_a, y_a)\}_{a=1}^{N_{au}} \) is a supervised data set containing sufficient historical ADS-B signal samples with their corresponding categories, i.e., \( N_{au} \gg N_{tr} \) and \( |Y_{au}| \gg |Y_{fs}| \). Here, \( x_a \in X_{au} \subset X \) and \( y_a \in Y_{au} \subset Y \), where \( X_{au} \) and \( Y_{au} \) are the auxiliary sample space and the auxiliary category space.

Moreover, it is noted that the categories in \( D_{tr} \) do not appear in \( D_{au} \), i.e., \( Y_{fs} \cap Y_{au} = \emptyset \), but the few-shot ADS-B signal...
sample \( x \) and auxiliary ADS-B signal sample \( x_s \) come from the same domain, that is, \( D_f = D_{au} \), \( X_f = X_{au} \) and \( P_X_f = P_X_{au} \), where \( D_f = \{X_f, P_X_f\} \) and \( D_{au} = \{X_{au}, P_X_{au}\} \) are the few-shot domain and the auxiliary domain, respectively, and \( P_X_f \) and \( P_X_{au} \) are the marginal probability distributions of the few-shot sample space and the auxiliary sample space, respectively.

In general, the idea of FS-SEI is to construct \( f_{\text{SEI}}(W) \) with the limited supervised knowledge in \( D_f \) and the auxiliary knowledge in the \( D_{au} \)-irrelevant data set \( D_{au} \). The assumption for FS-SEI or other FS problems is reasonable and feasible, since \( D_{au} \) is easily available.

### IV. PROPOSED FS-SEI METHOD BASED ON DMEL

#### A. Framework of FS-SEI

A typical SEI can be decomposed into two parts: 1) feature embedding and 2) classification. The former is to build the feature space, that is, embedding, and 2) classification. The latter is to map the features into categories, i.e., \( f_{\text{cl}}(W_{\text{cl}}) : F \rightarrow Y \), where \( F \) is the feature space, \( f_{\text{SEI}}(W) = f_{\text{cl}}(f_{\text{FE}}(W_{\text{FE}}) : W_{\text{cl}}) \) and \( W = [W_{\text{FE}}, W_{\text{cl}}] \).

FS-SEI also has the similar implementation. However, it is difficult to construct the feature embedding based on the DL model with huge hyper-parameters in a scenario where samples are scarce, while the realization of the classification has almost no difficulties, because there are lots of classifiers with few parameters or even no parameters. (For instance, some ML classifier and distance-based classifiers.) Thus, the key of FS-SEI is constructing a mapping for a good feature embedding, i.e., \( f_{\text{FE}}(W_{\text{FE}}) : X_f \rightarrow F \). In addition, considering that the direct construction is rarely possible and \( X_f = X_{au} \), the general alternative is to construct \( f_{\text{FE}}(W_{\text{FE}}) : X_{au} \rightarrow F \).

In this article, the feature embedding is realized by CVCNN and metric learning, while the classification is based on ensemble learning. The overview of the proposed DMEL-based FS-SEI method is shown in Fig. 3.

#### B. Feature Embedding via CVCNN and Metric Learning

1) Structure of CVCNN for SEI: In this article, CVCNN works as the feature embedding \( f_{\text{FE}}(W_{\text{FE}}) \) that is used to map ADS-B signal samples to features, because it has much more powerful and effective feature representation and classification capability than real-valued CNN in signal identification [23, 48]. The specific structure and parameters of the CVCNN are shown in Table II. Here, CVCNN consists of nine convolution layers and one fully connected layer. There are three essential components of CVCNN: 1) complex-valued convolution layer; 2) complex-valued batch normalization (BN) layer; and 3) complex-valued activation function, which will be introduced as follows.

**Complex-Valued Convolution Layer:** The input of the CVCNN is a complex-valued matrix with dimensionality \( N_{in} \times C_{in} \), and it can be denoted as \( I_{in} \). The complex-valued convolution operation is given as follows:

\[
I_{out} = \sum_{n=1}^{N_{in}} W_n \ast I_{in}(\cdot, n)
\]

| Structure                                      | The number of layers |
|------------------------------------------------|----------------------|
| CVConv1D (\( N_{ne}, 3 \) + CVReLU + CVBN + Maxpooling1D (2)) | \( \times 9 \)     |
| Flatten                                       | \( \times 1 \)       |
| Dense (1024)+ ReLU + Dropout (0.5)            | \( \times 1 \)       |

Note: “\( \times 9 \)” represents that there are nine same layers, while the meaning of “\( \times 1 \)” is similar.

\[
= \sum_{n=1}^{N_{in}} [\Re(W_n) + j \cdot \Im(W_n)] \\
\ast [\Re(I_{in}(\cdot, n)) + j \cdot \Im(I_{in}(\cdot, n))] \\
= \sum_{n=1}^{N_{in}} [\Re(W_n) \ast \Re(I_{in}(\cdot, n)) + j \cdot \Re(W_n) \ast \Im(I_{in}(\cdot, n)) + j \cdot \Re(W_n) \ast \Im(I_{in}(\cdot, n)) - \Im(W_n) \ast \Im(I_{in}(\cdot, n))] \\
\]

where \( W_n \) are the complex-valued convolution kernels in the \( n \)th channel; “\( \ast \)” represents convolution operation.

The convolution operation for SEI in this article is 1-D convolution, and the dimensionality of convolution kernel is \( N_{ke} \). Assuming that there are \( N_{ne} \) neurons and the convolution stride is 1, the time complexity of one complex-valued convolution layer can be written as follows:

\[
\text{Time} \sim O(4N_{ke} \cdot N_{out} \cdot N_{in} \cdot N_{ne})
\]

where \( N_{out} \) is equal to \( (N_{in} - N_{ke} + 1) \) (the padding mode is “valid”) or \( N_{in} \) (the padding mode is “same”), respectively.

2) Complex-Valued Batch Normalization: BN is generally used to accelerate the training process and avoid overfitting, and there is the corresponding complex-valued BN (CVBN). Assuming the \( i \)th element of the complex-valued input in a batch is \( I_{b,i} \). Then, the real part and imaginary part of \( I_{b,i} \) are...
split to form a matrix
\[ I_{b,i} = \begin{bmatrix} \Re(I_{b,i}) \\ \Im(I_{b,i}) \end{bmatrix} \_{2 \times 1} \]
Similar with the real-valued BN, CVBN can be expressed as follows:
\[
f_{CVBN}(I_{b,i}) = \gamma \left( \sigma^2 + \epsilon \right)^{-\frac{1}{2}} I_{b,i} - \mu + \beta \tag{8}\]
where
\[ \sigma^2 = \begin{bmatrix} \Var_{\Re} & \Cov_{\Re, \Im} \\ \Cov_{\Re, \Im} & \Var_{\Im} \end{bmatrix} \]
\[ \mu = \begin{bmatrix} \Re \sum_{b=1}^{B} \Re(I_{b,i}) \\ \Im \sum_{b=1}^{B} \Im(I_{b,i}) \end{bmatrix} \]
\[ \gamma = \begin{bmatrix} \Re \gamma_{\Re} \\ \Im \gamma_{\Im} \end{bmatrix} \]
\[ \beta = \begin{bmatrix} \beta_{\Re} \\ \beta_{\Im} \end{bmatrix} \]
and \((-1)^{-1/2}\) is an inverse operation with taking square root; \(\Cov_{\Re, \Im}\) (or \(\Cov_{\Im, \Re}\)) is the covariance between \(\Re(I_{b,i})\) and \(\Im(I_{b,i})\); \(\Var_{\Re}\) and \(\Var_{\Im}\) are the variance of \(\Re(I_{b,i})\) and \(\Im(I_{b,i})\), respectively; \(\gamma\) and \(\beta\) are the learnable parameters; and \(B\) is the number of batch size.

**Complex-Valued Activation Function:** Complex-valued rectified linear unit (CVReLU) can be expressed as follows:

\[
f_{CVReLU}(I_{in}) = \max[\Re(I_{in}), 0] + j \cdot \max[\Im(I_{in}), 0]. \tag{9}\]

**2) Hybrid Metric for Discriminative Feature Embedding:**
Most of conventional SEI methods are close-set identification method, i.e., the categories of samples for testing are within training data set. In this case, the training DL model by Softmax loss can generally obtain good performance. Therefore, the feature embedding based on Softmax loss can only obtain features that are prone to separate different seen categories. However, the close-set SEI is nearly impossible in realistic deployment, since that it is almost impractical to precollect ADS-B signals samples from all possible aircrafts for training. It means that the separable feature is not effective enough, and the features for SEI need to be discriminative and generalized enough for characterizing the ADS-B signal samples from unseen categories or categories seen a few times. The difference between separable feature and discriminative feature is given in Fig. 4.

The core characteristics of discriminative features are the intraclass consistency and the intercategory separability. Thus, we proposed an effective discriminative feature embedding method based on hybrid metric, which consist of Softmax loss, and two typical contrastive losses, and this method can effectively narrow the distances between different ADS-B signal samples from the same aircrafts in the feature space, meanwhile enlarge the distances between ADS-B signal samples from different aircrafts.

**Softmax Loss:** The feature, extracted from ADS-B signal sample \(x_i\) in \(D_{au}\), can be expressed as \(f_{FE}(W_{FE}; x_i)\). Then, this feature is fed into a Dense layer for mapping from feature space into category space corresponding to \(D_{au}\), and the output can be expressed as \(z_a = f_{De}(W_{De}; f_{FE}(x_i))\), where \(W_{De}\) is the weight of the Dense layer. Thus, Softmax loss can be written as follows:

\[
\mathcal{L}_{Softmax} = -\mathbb{E} \left[ \log \frac{\exp(y_a)(z_a)}{\sum_n \exp(y_a)(z_n)} \right] \tag{10}\]
where \(z_a(\cdot)\) is the \(\cdot \)th element, which represents the nonnormalized probability that the sample belongs to the \(a\)th category. In addition, it is noted that the above function is a simplified form.

**Contrastive Loss:** Softmax loss is separable enough, but its generalization and discrimination are insufficient. Therefore, two contrastive losses are introduced as auxiliary losses to supplement these deficiencies. The one of the contrastive losses is triplet loss, which can narrow the intraclass distance and enlarge the intercategory distance in the feature space. Triplet loss, applied in this article, is originate from the triplet network [51], which is shown in Fig. 5.

Triplet network divides samples in the data set \(D_{au}\) into three parts: 1) anchor samples; 2) positive samples; and 3) negative samples, where the positive samples are the samples with the same categories of the anchor samples, and the negative samples are the samples that have different categories with the anchor samples. Three kinds of samples are fed into the CVCNN-based feature embedding in pairs, and triplet network
encourages the CVCNN-base feature embedding to shrink the distance between $x^{as}$ and $x^+$, meanwhile expand the distance between $x^{as}$ and $x^-$ in the feature space. Here, the Euclidean distance is generally applied to measure the distance in the feature space. So, triplet loss can be expressed as follows:

$$L_{Triplet} = E \left[ \max\left( ||f_{FE}(W_{FE}; x^+) - f_{FE}(W_{FE}; x^{as})||_2 - ||f_{FE}(W_{FE}; x^-) - f_{FE}(W_{FE}; x^{as})||_2 + \xi, 0 \right) \right]$$  \hspace{1cm} (11)

where $|| \cdot ||_2$ is the $\ell_2$ norm and $\xi$ is the tunable margin.

The other one of the contrastive losses is center loss [52], which works for obtaining more compact intracategory distances in the feature space. Its formula can be written as follows:

$$L_{Center} = \frac{1}{2} E \left[ ||f_{FE}(W_{FE}; x_a) - c_{ya}||_2^2 \right]$$  \hspace{1cm} (12)

where $c_{ya}$ is the learnable center feature of the $y_a$th category. Moreover, the sketch maps of triplet loss and center loss are given in Fig. 6.

Finally, we give the hybrid metric-based loss function for discriminative feature embedding in SEI, which is shown as follows:

$$L_{HM} = L_{Softmax} + \lambda \left( L_{Contrastive} + L_{Center} \right)$$  \hspace{1cm} (13)

where $\lambda$ is a factor to balance Softmax loss and contrastive loss.

3) Optimization Methods: There are two parts that require to be optimized: 1) the weight of feature embedding $W_{FE}$ and 2) the center features $c_{ya}$. The specific optimization methods are shown as follows.

**Optimization Method for Feature Embedding:** The weight of feature embedding $W_{FE}$ is updated by the stochastic gradient descent algorithm, which can be written as follows:

$$W_{FE}^{t+1} = W_{FE}^t - \eta \cdot \frac{\partial L_{HM}}{\partial W_{FE}}$$  \hspace{1cm} (14)

where $\eta$ is the learning rate in SGD.

**The Optimization Method for Center Feature:** In the ideal case, if the features are changed, $c_{ya}$ should be updated. However, it is impractical to take the entire auxiliary data set into account, and calculate the average feature of each category. Thus, the center features are updated based on the mini-batch method, and it means that the feature centers are calculated by averaging the features of the categories containing in the mini-batch. The update formula of $c_{ya}$ can be written as follows:

$$c_{ya}^{t+1} = c_{ya}^{t} - \alpha \cdot \Delta c_{ya}$$  \hspace{1cm} (15)

$$\Delta c_{ya} = \frac{1}{1 + \sum_{b=1}^{B} \delta(y_b == y_a) \cdot ||f_{FE}(W_{FE}; x_b) - c_{ya}||^2}$$  \hspace{1cm} (16)

where $\alpha$ is the learning rate and $\delta(\cdot)$ is a conditional selection function. If the condition is established, $\delta(\cdot)$ will be equal to 1, otherwise it will be 0. In other world, if there are no samples corresponding to the $y_a$th category in a mini-batch data set $\{x_b, y_b\}_{b=1}^{B}$, the feature center of the $y_a$th category remains unchanged.

C. Ensemble Classifier Based on Probability Average Method

As mentioned above, FS-SEI can be described as the cascaded feature embedding and a simple classifier, i.e., $f_{SEI}(W) = f_{cl}(f_{FE}(W_{FE}); W_{cl})$. Thus, after the implement of the hybrid metric-based feature embedding $f_{FE}(W_{FE})$, a simple classifier $f_{cl}(W_{cl})$ is applied for FS-SEI. Here, taking logistic regression (LR) as an example, the training and testing processing of classifier is given as follows.

The training process is based on the few-shot ADS-B signal training data set $D_{tr}$, which can be described as follows:

$$W_{cl}^t = \arg \min_{W_{cl}} \left\{ -\frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} \log P_i(y_i) \right\}$$  \hspace{1cm} (17)

where $P_i = f_{SEI}(x_i; W)$ is the predicted probability distribution about $x_i$, where $x_i \in D_{tr}$, and $P_i(y_i)$ is the $y_i$th element, which represents the predicted probability of the $y_i$th categories. Next, the test process on $D_{te}$ can be written as follows:

$$\hat{y}_j = \arg \max_{1 \leq y_j \leq C} P_j$$  \hspace{1cm} (18)
Algorithm 2 Identification Method via Ensemble Learning

Input: Predicted results \( \hat{y}_{fc}^{en} \) on test dataset \( D_{te} \).

Output: Predicted results \( \hat{y}_j^{en} \) on test dataset \( D_{te} \).

1. Set the number of base classifiers \( M \); 
2. Obtain \( W_{m}^{FE} \), \( m = 1, 2, \ldots, M \) by multiple independent model initialization and training: 
   \[ \text{Train on } D_{tr} = \{x_i, y_i\}_{i=1}^{N_{tr}}; \]
3. for \( m = 1 \) to \( M \) do: 
   4. Add LR classifier \( f_{cl} \); 
   5. Obtain \( f_{m}^{en} = f_{FE}(x_i; W_{m}^{FE}), i \in \{1, 2, \ldots, N_{tr}\}; \)
   6. Obtain \( W_{m}^{CL} \) based on \( f_{cl} \) and \( y_i \) by minimizing (17); 
7. end for 
8. Test on \( D_{te} = \{x_i\}_{i=1}^{N_{te}}; \)
9. Cascade \( f_{FE} \) and \( f_{cl} \), i.e., \( f_{SEI}(W_{m}^{en}) = f_{cl}(f_{FE}(W_{m}^{en}); W_{m}^{cl}); \)
10. Calculate the predicted probability distribution \( P_j^{en} = f_{SEI}(x_i; W_{m}^{en}); \)
11. Calculate the the predicted category \( \hat{y}_j^{en} \) by (19); 
12. return \( \hat{y}_j^{en} \)

where \( \hat{y}_j \) is the predicted category, and \( P_j = f_{cl}(f_{FE}(x_i; W_{m}^{FE}); W_{m}^{cl}); x_j \in D_{te} \). This process means that the category with the highest probability is as the predicted category.

To further improve the identification performance and robustness, ensemble learning is introduced, which consists of two steps: 1) construction of multiple base FS-SEI models with diversity and 2) joint decision. In the former step, base SEI models are built by multiple independent training for diversity. These model weights can be written as \( \{W_m^{en}\}_{m=1}^{M} \), where \( M \) is the number of base FS-SEI models. The latter step is based on the probability average method, which can be expressed as follows:

\[
\hat{y}_j^{en} = \arg \max_{\{1 \leq y_j \leq C\}} \sum_{m=1}^{M} P_j^{en} \quad (19)
\]

where \( \hat{y}_j^{en} \) is the predicted category by ensemble learning and \( P_j^{en} = f_{SEI}(x_i; W_{m}^{en}); x_j \in D_{te}. \)

The training of classifier is completed online. It is noted that the online training and deployment can be realized quickly, because the classifiers for FS-SEI can be few-parameter and even parameterless, and the number of training samples is limited.

D. Comparison Methods

In this article, the proposed FS-SEI based method on Softmax loss, triplet loss, and center loss is named as “STC CVCNN,” and other SEI methods for comparison are introduced as follows.

1) Instantaneous Feature: Instantaneous features are extracted from the amplitude, phase, and frequency components of the received signal, and the features include the mean, variance, skewness, and kurtosis [32], [38].

2) Softmax CVCNN and Direct CVCNN: The CVCNN with only Softmax loss, based on pe-training on \( D_{au} \) and fine-tuning on \( D_{tr} \), is named as “Softmax CVCNN,” and it, based on direct training on \( D_{tr} \), is named as “Direct CVCNN.”

3) Two Contrastive Loss-Based Methods: Here, we adopt two contrastive loss-based signal recognition methods for comparison. The one of the contrastive losses is originate from SiameseNet [43]–[46], which has been introduced in Part II. The other is triplet loss, which has been specifically introduced above. It is noted that the network structures in SiameseNet and TCNN are replaced by the structure in Table II for a fair comparison. So, they are named as “Siamese CVCNN” and “Triplet CVCNN,” and their training schemes are similar with Softmax CVCNN.

4) SR2CNN: SR2CNN is an advanced zero-shot signal recognition based on combined metric and generative methods, and it has a dual branch structure, respectively, responsible for classification and reconstruction. Mean square error is applied as loss function for the reconstruction part, while Softmax loss with center loss is used for classification. It is also a hybrid metric method for signal recognition. Similarly, for a fair comparison, the CNN structure in SR2CNN is also replaced with the structure in Table II.

5) Ablation Studies: In addition to the above-compared methods, there are two ablation studies, i.e., “ST CVCNN” and “SC CVCNN.” The former is the CVCNN based on Softmax loss and triplet loss, while the latter is the CVCNN based on Softmax loss and center loss.

V. EXPERIMENTAL RESULTS

A. Simulation Environment, Parameters, and Performance Metrics

These DL models are based on Tensorflow, and these machine learning classifiers are based on Scikit-learn. The other simulation parameters are listed in Table III. In addition, considering that the samples used in this article are noisy samples in the real scenarios, and different few-shot training samples have a certain impact on the performance, we randomly selected 1000 groups of different few-shot training samples as \( D_{tr} \) for Monte Carlo simulations, and each simulation is tested on the same test data set, i.e., \( D_{te}. \) The average accuracy of these tests is applied as an indicator of the identification performance of these algorithms.

However, the average accuracy can not directly show the discriminative degrees of different feature embedding methods. So, we also apply the silhouette coefficient (SC) to
directly measure the discriminative degree \[55\]. The SC can measure both cohesion and separability, i.e., the former can indicate the intracategory distance, while the latter can show the intercategory distance. The formula of the SC is given as follows:

\[
C_{\text{silhouette}} = \frac{1}{N_{te}} \sum_{j=1}^{N_{te}} \frac{D_j^{\text{inter}} - D_j^{\text{intra}}}{\max(D_j^{\text{inter}}, D_j^{\text{intra}})} \tag{20}
\]

where the range of \(C_{\text{silhouette}}\) is between \(-1\) and \(1\), \(D_j^{\text{intra}}\) is the average distance between the feature vector of the \(j\)th sample and the feature vectors of other samples, which have the same category with the \(j\)th sample, i.e.,

\[
D_j^{\text{intra}} = \frac{\sum_k \delta(y_j = y_k) \cdot \|f_{FE}(x_j) - f_{FE}(x_k)\|_2}{\sum_k \delta(y_j = y_k)} \tag{21}
\]

where \(j \neq k, (x_j, y_j) \in D_{te}\) and \((x_k, y_k) \in D_{te}\). In addition, \(D_j^{\text{inter}}\) is the minimum distance between the feature vector of the \(j\)th sample and the feature vectors of other samples, which have the different categories with the \(j\)th sample, i.e.,

\[
D_j^{\text{inter}} = \min[k \delta(y_j \neq y_k) \cdot \|f_{FE}(x_j) - f_{FE}(x_k)\|_2]. \tag{22}
\]

It is obvious that the closer the SC approaches to one, the better the cohesion and separability. In addition, the average accuracy curves are given in Figs. 7–9, while the SCs are shown in Table IV.

B. Softmax Loss Versus Contrastive Loss

Here, we mainly compare the performance of the CVCNN based on Softmax loss with that based on two contrastive losses. The detailed results are shown in Fig. 7 and Table IV. First, when comparing with two contrastive loss-based methods, we can find that the performance of Triplet CVCNN is obviously better than that of Siamese CVCNN in the average accuracy and the SCs. It can be demonstrated that the
TABLE IV
SCS OF DIFFERENT FEATURE EMBEDDING METHODS

| Silhouette Way | Softmax CVCNN | Siamese CVCNN | Triplet CVCNN | SR2CNN | ST CVCNN | SC CVCNN | STC CVCNN (proposed) |
|----------------|--------------|---------------|--------------|--------|----------|---------|----------------------|
| 10             | 0.0793       | 0.1300        | 0.2557       | 0.2558 | 0.1683   | 0.2846  | 0.4629               |
| 20             | 0.0713       | 0.1300        | 0.2385       | 0.2246 | 0.1663   | 0.2771  | 0.4558               |
| 30             | 0.0588       | 0.0800        | 0.1860       | 0.2092 | 0.1318   | 0.2306  | 0.3722               |

Finally, it is noted that the performance of the center loss-based CVCNN is not given, because the feature embedding method, only based on center loss, has little separability, and the identification performance of the center loss-based method is close to that of a random guess. It can be demonstrated by feature visualization in Fig. 10(d).

C. Comparison Between Different Hybrid Losses

The performance between different combinations of Softmax loss and contrastive losses is shown in Fig. 8 and Table IV.

1) STC CVCNN Versus SR2CNN Versus Single Loss-Based Methods: Compared with single Softmax loss or contrastive loss, the combination of both has better performance in few-shot scenarios. In most few-shot scenarios, STC CVCNN and SR2CNN outperform single Softmax loss or the contrastive loss-based method in the average accuracy. More importantly, our proposed STC CVCNN far exceeds SR2CNN and other methods in the average accuracy and the SCs. Moreover, with the increasing of the number of aircrafts to be identified, the performance gap between STC CVCNN and other methods also increases. It demonstrated that STC CVCNN has stabler performance than other methods under the condition of different ways.
2) Ablation Studies of STC CVCNN: It is obvious that STC CVCNN outperforms ST CVCNN and SC CVCNN in the average accuracy and the SCs. Specifically, compared with ST CVCNN and SC CVCNN, STC CVCNN has 0.9%–9.7% and 1.3%–7.6% identification performance improvement in “10 ways” case, respectively. Furthermore, with the increasing of the number of aircrafts to be identified, their identification performance gaps become larger and larger. Similarly, the SCs of STC CVCNN are higher than that of ST CVCNN and SC CVCNN, which demonstrates that STC CVCNN has better separability and cohesion than other methods.

D. Single Classifier Versus Ensemble Classifier

The above simulation results are based on a single machine learning classifier, i.e., LR, and the following ensemble classifier is also applied LR as the base classifier. The detailed results between a single classifier and ensemble classifier are shown in Fig. 9. Obviously, regardless of the number of base classifiers, ensemble classifiers perform better than a single classifier. In detail, under the one-shot condition, the performance improvement can exceed 3%–7%, while that is only 0.5%–1%. With the increasing of base classifiers, the performance of the ensemble classifier can be slightly improved only in case of one shot, and the maximum improvements are 1.58%, 0.54%, and 0.95%, but the performances in other cases are barely modified. The above analysis illustrates the effectiveness of ensemble learning. Meanwhile, it also shows that the increasing of base classifiers will not bring much improvement, but only increase the cost of training and deployment, except in the one-shot scenario.

E. Feature Visualization

The dimensionality of the extracted features is reduced to two dimensions by t-distributed stochastic neighbor embedding (t-SNE) for visualization, which is shown in Fig. 10. Here, we only give the scenario of “10 ways,” because if there are too many categories, the figure will be too complex and difficult to analyze.

These single loss-based methods are shown in Fig. 10(a)–(c). It is obvious that the feature boundary of each category can be easily determined in the feature space from Softmax CVCNN, but these features of each category are loosely distributed, and do not have compact intradistance. Instead, Siamese CVCNN and Triplet CVCNN can significantly reduce the intradistance, especially the latter, but their extracted features have poor separability, for instance, “category 7” and “category 8” almost overlap in Siamese CVCNN, and they have no obvious separation boundary in Triplet CVCNN.

Thus, we combine Softmax loss and contrastive loss to extract the discriminative features. The visualization of these multiloss-based methods is shown in Fig. 10(e)–(h). Compared with Triplet CVCNN, ST CVCNN only adds extra Softmax loss, and it not only maintains compact intracategory distance, but also slightly enlarge intercategory distance to realize the separability of features. SR2CNN and SC CVCNN also have the similar visualization results.

More than that, STC CVCNN adds extra center loss on the basis of ST CVCNN for further narrowing intracategory distance. It can be observed in Fig. 10(h) that the intracategory distance of features extracted by STC CVCNN is more compact than that extracted by other methods, and STC CVCNN realizes the excellent discriminability of features, which is the key to its performance far surpassing other methods.

F. Existing Problem and Analysis

Here, the detailed analysis of 1000 Monte Carlo simulations is adopted by box plot, which is shown in Figs. 11 and 12. The red “+” represents the outlier, and the value of the red horizontal line is median; two black horizontal lines represent the maximum and minimum values in the case of removing outliers, respectively; the upper horizontal line of the blue box is the upper quartile, while the lower one represents the lower quartile.

Fig. 11 revealed the existing problem that different few-shot training samples have different identification performances. A group of excellent few-shot training samples can realize more than 90% accuracy, while the identification performance based on a set of poor-quality samples just has 60%–70%. This phenomenon is the most serious in the one-shot scenario, and the more training samples, the slighter the performance fluctuation. For example, the identification performance ranges from 69.2% to 98.8% under the “1 shot, 10 ways” condition, while the fluctuation range of accuracy is 98.3%–99.5% under the “20 shot, 10 ways” condition. The possible reason for the above phenomena is that the training data set is noisy data set.
collected from the real open world, and the sample quality is inconsistent. The sample quality may be related to wireless channel, noise, interference, even wrong category caused by wrong demodulation.

The above analysis is based on a single LR classifier, and we also give the detail results about ensemble classifier in Fig. 12. Except improving the performance, another original intention of using the ensemble classifier is to reduce the effect of sample quality on the stability of identification performance. Excluding a few outliers, ensemble classifier can obviously reduce the performance fluctuation in the scenario of 10 ways, and more-based classifiers, more stable identification performance. However, the ensemble classifier has little effect on improving the stability in other scenarios.

VI. CONCLUSION

In this article, we proposed an effective FS-SEI method for aircraft identification based on metric learning and ensemble learning. Specifically, the proposed FS-SEI method consists of feature embedding and ensemble classifier. The former is to map from ADS-B signal samples into features by CVCNN and hybrid metric loss combined with Softmax loss and two contrastive losses, while the latter is to construct the mapping from features into categories by a simple ensemble classifier. Simulation results demonstrated the effectiveness of our proposed FS-SEI via metric learning and CVCNN-based feature embedding and ensemble classifier. Feature visualization also showed the compact intracategory distance and separable intercategory distance in the features extracted by our proposed method. Finally, we also revealed the impact of noisy samples on the stability of the algorithm, and we expect to use some schemes, such as the attention mechanism [56], to reduce the impact of sample quality on identification performance in the future works.

REFERENCES

[1] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, “A survey on space-air-ground-sea integrated network security in 6G,” IEEE Commun. Surveys Tuts., vol. 24, no. 1, pp. 53–87, 1st Quart., 2022.

[2] X. Sun, D. W. K. Ng, Z. Ding, Y. Xu, and Z. Zhong, “Physical layer security in UAV systems: Challenges and opportunities,” IEEE Wireless Commun., vol. 26, no. 5, pp. 40–47, Oct. 2019.

[3] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi, “6G: Opening new horizons for integration of comfort, security, and intelligence,” IEEE Wireless Commun., vol. 27, no. 5, pp. 126–132, Oct. 2020.

[4] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, “Physical layer authentication in wireless communication networks: A survey,” J. Commun. Inf. Netw., vol. 5, no. 3, pp. 237–264, Sep. 2020.

[5] N. Wang, W. Li, P. Wang, A. Alipour-Fanid, L. Jiao, and K. Zeng, “Physical layer authentication for 5G communications: Opportunities and road ahead,” IEEE Netw., vol. 34, no. 6, pp. 198–204, Nov./Dec. 2020.

[6] C. Mitchell and C. He, “Security analysis and improvements for IEEE 802.11,” in Proc. 12th Annu. Netw. Distrib. Syst. Security Symp., 2005, pp. 90–110.

[7] X. Liu, L. Zheng, S. Helal, W. Zhang, C. Jia, and J. Zhou, “A broad learning-based comprehensive defence against SSDP reflection attacks in IoTs,” Digit. Commun. Netw., to be published. [Online]. Available: doi.org/10.1016/j.dcan.2022.02.008

[8] M. Muhammad and G. A. Safdar, “Survey on existing authentication issues for cellular-assisted V2X communication,” Veh. Commun., vol. 12, pp. 50–65, Apr. 2018.

[9] B. Li, Z. Fei, C. Zhou, and Y. Zhang, “Physical-layer security in space information networks: A survey,” IEEE Internet Things J., vol. 7, no. 1, pp. 33–52, Jan. 2020.

[10] B. He, F. Wang, Y. Liu, and S. Wang, “Specific emitter identification via multiple distorted receivers,” in Proc. IEEE Int. Conf. Commun. Workshops, 2019, pp. 1–6.

[11] B. He and F. Wang, “Cooperative specific emitter identification via multiple distorted receivers,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 3791–3806, 2020.

[12] Q. Li, X. Sun, P. Lv, and C. Sun, “The identification of power IoT devices on differential constellation trajectory map,” in Proc. Int. Conf. Meas. Technol. Mechatronics Autom., 2021, pp. 140–144.

[13] Y. Sun, S. Abeywickrama, L. Jayasinghe, C. Yuen, J. Chen, and M. Zhou, “Micro Doppler signature based detection, classification, and localization of small UAV with long short term memory neural network,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 8, pp. 6285–6300, Aug. 2021.

[14] H. Ye, G. Y. Li, and B.-H. Jiang, “Power of deep learning for channel estimation and signal detection in OFDM systems,” IEEE Wireless Commun. Lett., vol. 7, no. 1, pp. 114–117, Feb. 2018.

[15] H. He, C.-K. Wen, S. Jin, and G. Y. Li, “Deep learning-based channel estimation for beamspace mmWave massive MIMO systems,” IEEE Wireless Commun. Lett., vol. 7, no. 5, pp. 852–855, Oct. 2018.

[16] W. Xia, G. Zheng, Y. Zhu, J. Zhang, J. Wang, and A. P. Petropulu, “A deep learning framework for optimization of MISO downlink beamforming,” IEEE Trans. Commun., vol. 68, no. 3, pp. 1866–1880, Mar. 2020.

[17] C. Wen, W. Shih, and S. Jin, “Deep learning for massive MIMO CSI feedback,” IEEE Wireless Commun. Lett., vol. 7, no. 5, pp. 748–751, Oct. 2018.

[18] L. Liang, H. Ye, G. Y. Li, and G. Y. Li, “Deep-learning-based wireless resource allocation with application to vehicular networks,” Proc. IEEE, vol. 108, no. 2, pp. 341–356, Feb. 2020.

[19] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, “Deep learning for RF device fingerprinting in cognitive communication networks,” IEEE J. Sel. Topics Signal Process., vol. 12, no. 1, pp. 160–167, Feb. 2018.

[20] J. Yu, A. Hu, G. Li, and L. Peng, “A robust RF fingerprinting approach using multisampling convolutional neural network,” IEEE Internet Things J., vol. 6, no. 4, pp. 6786–6799, Aug. 2019.
X. Wang, Y. Zhang, H. Zhang, Y. Li, and X. Wei, “Radio frequency signal identification using transfer learning based on LSTM,” Circuits Syst. Signal Process., vol. 39, no. 11, pp. 5514–5528, Apr. 2020.

T. Jian et al., “Radio frequency fingerprinting on the edge,” IEEE Trans. Mobile Comput., to be published, doi: 10.1109/TMC.2021.3064466.

Y. Wang, G. Gui, H. Gacanin, T. Ohtsuki, O. A. Dobre, and H. V. Poor, “A multi-view specific emitter identification method based on complex-valued neural networks and network compression,” IEEE J. Sel. Areas Commun., vol. 39, no. 8, pp. 2305–2317, Aug. 2021.

L. Ding, S. Wang, F. Wang, and W. Zhang, “Specific emitter identification via convolutional neural networks,” IEEE Commun. Lett., vol. 22, no. 12, pp. 2591–2594, Dec. 2018.

G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, “Radio frequency fingerprint identification for LoRa using spectrogram and CNN,” in Proc. IEEE Conf. Comput. Commun., 2021, pp. 1–10.

C. Xie, L. Zhang, and Z. Zhong, “Virtual adversarial training-based semisupervised specific emitter identification,” Wireless Commun. Mobile Comput., vol. 2022, pp. 1–14, Jan. 2022.

J. Gong, X. Xu, and Y. Lei, “Unsupervised specific emitter identification method using radio-frequency fingerprint embedded InfoGAN,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 2898–2913, 2020.

L. Peng, J. Zhang, M. Liu, and A. Hu, “Deep learning based RF fingerprint identification using differential constellation trace figure,” IEEE Trans. Veh. Technol., vol. 69, no. 1, pp. 1091–1095, Jan. 2020.

P. Yin, L. Peng, J. Zhang, M. Liu, H. Fu, and A. Hu, “LTE device identification based on RF fingerprint with multi-channel convolutional neural network,” in Proc. IEEE Global Commun. Conf., 2021, pp. 1–6.

Y. Peng, P. Liu, Y. Wang, G. Gui, B. Adebisi, and H. Gacanin, “Radio frequency fingerprint identification based on slice integration cooperation and heat constellation trace figure,” IEEE Wireless Commun. Lett., vol. 11, no. 3, pp. 543–547, Mar. 2022.

H. Fu, S. Abeywickrama, L. Zhang, and C. Yuen, “Low-complexity portable passive drone surveillance via SDR-based signal processing,” IEEE Commun. Mag., vol. 56, no. 4, pp. 112–118, Apr. 2018.

A. Aghnaiya, Y. Dalveren, and A. Kara, “On the performance of variational mode decomposition-based radio frequency fingerprinting of Bluetooth devices,” Sensors, vol. 20, no. 6, pp. 1704–1713, Mar. 2020.

Y. Huang, “Radio frequency fingerprint extraction of radio emitter based on I/Q imbalance,” Procedia Comput. Sci., vol. 107, pp. 472–477, Apr. 2017.

B. Danev, T. S. Heydt-Benjamin, and S. Çapkun, “Physical-layer identification of RFID devices,” in Proc. USENIX Security Symp., 2009, pp. 199–214.

Y. Huang and H. Zheng, “Radio frequency fingerprinting based on the constellation errors,” in Proc. IEEE Asia-Pacific Conf. Commun. (APCC), 2012, pp. 900–905.

J. Li, H. Ji, and L. Wang, “Specific radar emitter recognition based on wavelet packet transform and probabilistic SVM,” in Proc. IEEE Int. Conf. Inf. Autom., 2009, pp. 1308–1313.

Y. Yuan, Z. Huang, H. Wu, and X. Wang “Specific emitter identification based on Hilbert–Huang transform-based time–frequency–energy distribution features,” IET Commun., vol. 8, no. 13, pp. 2404–2412, Sep. 2014.

Y. Tu, Z. Zhang, Y. Li, C. Wang, and Y. Xiao, “Research on the Internet of Things device recognition based on RF-fingerprinting,” IEEE Access, vol. 7, pp. 37426–37431, 2019.

J. Lu, P. Gong, J. Ye, and C. Zhang, “Learning from very few samples: A survey,” 2020, arXiv:2009.02653.

H. Zhou et al., “Few-shot electromagnetic signal classification: A data union augmentation method,” Chin. J. Aeronaut., to be published. [Online]. Available: doi.org/10.1016/j.cja.2021.07.014.

Y. Wang, L. Yao, Y. Wang, and Y. Zhang, “Robust CSI-based human activity recognition with augment few shot learning,” IEEE Sensors J., vol. 21, no. 21, pp. 24297–24308, Nov. 2021.

L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, and P. H. S. Torr, “Fully-convolutional Siamese networks for object tracking,” in Proc. Eur. Conf. Comput. Vis. 2016, pp. 880–895.

Y. Wu, Z. Sun, and G. Yue, “Siamese network-based open set identification of communications emitters with comprehensive features,” in Proc. IEEE Int. Conf. Comput. Image Signal Process. (CCISP), 2021, pp. 408–412.

G. Sun, “RF transmitter identification using combined Siamese networks,” IEEE Trans. Instrum. Meas., vol. 71, pp. 1–13, Mar. 2022.

Z. Yu, X. Wu, Y. Li, and M. Gao, “Few-shot learning of signal modulation recognition based on attention relation network,” in Proc. Eur. Signal Process. Conf., 2021, pp. 1372–1376.

P. Man, C. Ding, W. Ren, and G. Xu, “A specific emitter identification algorithm under zero sample condition based on metric learning,” Remote Sens., vol. 13, no. 23, pp. 4919–4928, Dec. 2021.

N. Yang et al., “Specific emitter identification with limited samples: A model-agnostic meta-learning approach,” IEEE Commun. Lett., vol. 26, no. 2, pp. 345–349, Feb. 2022.

Y. Dong, Y. Peng, M. Yang, S. Lu, and Q. Shi, “Signal transformer: Complex-valued attention and meta-learning for signal recognition,” 2021, arXiv:2106.04392.

A. Antoniou, H. Edwards, and A. Sterkow, “How to train your MAML,” 2018, arXiv:1810.09502.

Y. Tu et al., “Large-scale real-world radio signal recognition with deep learning,” Chin. J. Aeronaut., to be published, doi: 10.1016/j.cja.2021.08.016.

E. Hoffer and N. Ailon, “Deep metric learning using triplet network,” in Proc. Int. Workshop Similarity Based Pattern Recognit., 2015, pp. 84–92.

Y. Wen, K. Zhang, Z. Li, and Y. Qiao, “A discriminative feature learning approach for deep face recognition,” in Proc. Eur. Conf. Comput. Vis., 2016, pp. 499–515.

J. Liu et al., “Multi-scale triplet CNN for person re-identification,” in Proc. ACM Int. Conf. Multimedia, 2016, pp. 192–196.

X. He, Y. Zhou, Z. Zhou, S. Bai, and X. Bai, “Triplet-center loss for multi-view 3D object retrieval,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 1945–1954.

S. Aranganayagi and K. Thangavel, “Clustering categorical data using silhouette coefficient as a relocating measure,” in Proc. Int. Conf. Comput. Intell. Multimedia Appl., 2007, pp. 13–17.

T. Gao, X. Han, Z. Liu, and M. Sun, “Hybrid attention-based prototypical networks for noisy few-shot relation classification,” in Proc. AAAI Conf. Artif. Intell., 2019, pp. 6407–6414.

Han Gui (Senior Member, IEEE) received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. From 2009 to 2014, he joined Tohoku University, Sendai, Japan, as a Research Assistant as well as a Postdoctoral Research Fellow, respectively. From 2014 to 2015, he was an Assistant Professor with Akita Prefectural University, Akita, Japan. Since 2015, he has been a Professor with Nanjing University of Posts and Telecommunications, Nanjing, China. He has published more than 200 international peer-reviewed journal/conference papers and won eight best paper awards, e.g., ICC 2017, ICC 2014, and VTC 2014-Spring. His recent research interests include intelligent sensing and signal recognition.

Prof. Gui received the IEEE Communications Society Heinrich Hertz Award in 2021, the Clarivate Analytics Highly Cited Researcher in 2021, and the Highly Cited Chinese Researchers by Elsevier in 2020 and 2021. He is serving or served on the editorial boards of several journals, such as IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. In addition, he served as the IEEE VTS Ad Hoc Committee Member of AI Wireless, an Executive Chair of VTC 2021-Fall, the Vice Chair of WCNC 2021, and the TPC member of many IEEE international conferences. Since 2022, he has been a Distinguished Lecturer with the IEEE Vehicular Technology Society.
Yun Lin (Member, IEEE) received the B.S. degree in electrical engineering from Dalian Maritime University, Dalian, China, in 2003, the M.S. degree in communication and information system from Harbin Institute of Technology, Harbin, China, in 2005, and the Ph.D. degree in communication and information system from Harbin Engineering University, Harbin, in 2010.

From 2014 to 2015, he was a Research Scholar with Wright State University, Dayton, OH, USA. He is currently a Full Professor with the College of Information and Communication Engineering, Harbin Engineering University. He has authored or coauthored more than 200 international peer-reviewed journal/conference papers, such as IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, TR, INFOCOM, GLOBECOM, ICC, VTC, and ICNC. His current research interests include machine learning and data analytics over wireless networks, signal processing and analysis, cognitive radio and software-defined radio, artificial intelligence, and pattern recognition.

Hsiao-Chun Wu (Fellow, IEEE) received the B.S.E.E. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1990, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 1993 and 1999, respectively.

From March 1999 to January 2001, he was a Senior Electrical Engineer with Motorola Personal Communications Sector Research Labs. Since January 2001, he has been with the Faculty of the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA, where he is currently a Distinguished Professor. From July 2007 to August 2007, he was a Visiting Assistant Professor with the Television and Networks Transmission Group, Communications Research Centre, Ottawa, ON, Canada. From August 2008 to December 2008, he was a Visiting Associate Professor with the Department of Electrical Engineering, Stanford University, Stanford, CA, USA. He is a Visiting Professor with the International College of Semiconductor Technology, National Chiao Tung University, Hsinchu, Taiwan. He has authored or co-authored more than 280 peer-reviewed technical journals and conference articles in electrical and computer engineering. His research interests include wireless communications and signal processing.

Chau Yuen (Fellow, IEEE) received the B.Eng. and Ph.D. degrees from Nanyang Technological University, Singapore, in 2000 and 2004, respectively.

He was a Postdoctoral Fellow with Lucent Technologies Bell Labs, Murray Hill, NJ, USA, in 2005, and a Visiting Assistant Professor with The Hong Kong Polytechnic University, Hong Kong, in 2008. From 2006 to 2010, he was with the Institute for Infocomm Research, Singapore. Since 2010, he has been with Singapore University of Technology and Design, Singapore.

Dr. Yuen was a recipient of the Lee Kuan Yew Gold Medal, the Institution of Electrical Engineers Book Prize, the Institute of Engineering of Singapore Gold Medal, the Merck Sharp and Dohme Gold Medal, and twice a recipient of the Hewlett Packard Prize. He received the IEEE Asia–Pacific Outstanding Young Researcher Award in 2012 and the IEEE VTS Singapore Chapter Outstanding Service Award in 2019. He serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, where he was awarded as the top Associate Editor from 2009 to 2015. He served as the Guest Editor for several special issues, including IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE Communications Magazine, and IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He is a Distinguished Lecturer of IEEE Vehicular Technology Society.

Fumiyuki Adachi (Life Fellow, IEEE) received the B.S. and Dr.Eng. degrees in electrical engineering from Tohoku University, Sendai, Japan, in 1973 and 1984, respectively.

In April 1973, he joined NTT Laboratories, Kanagawa, Japan, and conducted various researches on digital cellular mobile communications. From July 1992 to December 1999, he was with NTT DoCoMo, Tokyo, Japan, where he led a research group on Wideband CDMA for 3G systems. Since January 2000, he has been with Tohoku University.

He is currently a Professor Emeritus with Tohoku University, and is leading a resilient wireless communication research group aiming at beyond 5G systems with the International Research Institute of Disaster Science. His research interests are in the area of wireless signal processing and networking, including multiaccess, equalization, antenna diversity, cooperative transmission, channel coding, and radio resource management.