

# Deep adversarial neural network for specific emitter identification under varying frequency

Keju HUANG\*, Junan YANG, Hui LIU, and Pengjiang HU

College of Electronic Engineering, National University of Defense Technology, Hefei, Anhui 230037, China

**Abstract.** Specific emitter identification (SEI) is the process of identifying individual emitters by analyzing the radio frequency emissions, based on the fact that each device contains unique hardware imperfections. While the majority of previous research focuses on obtaining features that are discriminative, the reliability of the features is rarely considered. For example, since device characteristics of the same emitter vary when it is operating at different carrier frequencies, the performance of SEI approaches may degrade when the training data and the test data are collected from the same emitters with different frequencies. To improve performance of SEI under varying frequency, we propose an approach based on continuous wavelet transform (CWT) and domain adversarial neural network (DANN). The proposed approach exploits unlabeled test data in addition to labeled training data, in order to learn representations that are discriminative for individual emitters and invariant for varying frequencies. Experiments are conducted on received signals of five emitters under three carrier frequencies. The results demonstrate the superior performance of the proposed approach when the carrier frequencies of the training data and the test data differ.

**Key words:** specific emitter identification; unsupervised domain adaptation; transfer learning; deep learning.

## 1. Introduction

Specific emitter identification refers to the process of identifying individual emitters by analyzing radio frequency (RF) emissions, which contain specific characteristics of electronic circuits and radio frequency components [1]. As the hardware imperfections are determined by the production and manufacturing processes, the characteristics of individual emitters are unique and difficult to counterfeit [2]. Therefore, SEI is widely applied in both battlefield spectrum management [3] and civilian wireless network security [4, 5].

A crucial task of SEI is to obtain features discriminating different individual emitters, which are also known as radio frequency fingerprints (RFFs). According to ways of obtaining RFFs, SEI approaches can be classified into two main categories: hand-crafted features based and deep learning based. Approaches based on hand-crafted features utilize expert knowledge to extract RFFs, and train classifiers based on the RFFs. RFFs can be obtained by comparing received signals to the ideal ones in the modulation domain [6–8]. Time-frequency analysis methods, such as short-time Fourier transform (STFT) [9], wavelet transform [10], Wigner and Choi-Williams distribution [11, 12] and Hilbert-Huang transform (HHT) [13–15] are also applied to extract RFFs. Besides, some research extracts RFFs using nonlinear dynamics and complexity theory [16–18]. There are also methods extracting features from graphical representations [19] or using geometric features [20, 21]. Different from approaches based on hand-crafted features, deep learn-

ing based approaches train deep neural networks to learn RFFs and identify individual emitters end-to-end, with raw data or images transformed from raw data as input. Wu [22] used a recurrent neural network (RNN) based on long short-term memory for RF fingerprinting and achieved high detection accuracy even in the presence of strong noise. Ding [23] adopted a convolutional neural network (CNN) trained on compressed bispectrum, presenting higher accuracy than conventional methods. Merchant [24] proposed a CNN method with baseband error signal as input for ZigBee devices. Pan [25] constructed a deep residual network operating on Hilbert spectrum image, and the simulation results demonstrated superior performance of the proposed method under various channel conditions. Matuszewski [26] implemented an electromagnetic source recognition system based on deep neural networks. Wong [8, 27] proposed a method combining CNN and clustering algorithms and a method based on IQ imbalance estimation using CNN. Baldini [28] compared various approaches to turn signals into images as CNN input, including recurrence plots, continuous wavelet transform and short-time Fourier transform, with experiments showing that the wavelet-based approach outperformed other approaches.

While the majority of previous research focuses on the discriminativeness of RFFs, the reliability of RFFs is rarely considered. Most research assumes that training data and test data are independently identically distributed. In practice, however, the conditions under which the RF observables are collected for identification or testing may be different from the ones the SEI system is trained on. Changes in conditions such as channel, temperature, bandwidth, carrier frequency and modulation type lead to variations of signals, which causes current RFFs to fail [29, 30]. Unfortunately, condition changing is ubiquitous for wireless communications, especially with applications of cog-

\*e-mail: huangkeju@nudt.edu.cn

Manuscript submitted 2020-09-24, revised 2020-11-02, initially accepted for publication 2020-11-19, published in April 2021

nitive radios, which are enabled to adjust their transmitting parameters for spectrum efficiency and protection of communications [31]. Therefore, reliability of SEI approaches plays a vital role for real-world applications.

Transfer learning, which aims to extract knowledge from one or more source tasks and applies the knowledge to a target task [32], is considered as a solution for reliable SEI [30, 33]. A transfer learning method for SEI is proposed in [33]. The method transfers knowledge from past time frames to current time frame, assuming RFFs gradually change over time. Hence, the method is not suitable for abrupt variations of features, e.g. changes of carrier frequencies. In [30] a method to identify emitters with changing bandwidth is proposed. However, the method requires labeled data of a set of fixed reference emitters at each bandwidth, which is not always available in practice. Unsupervised domain adaptation (UDA), a subcategory of transfer learning, transfers knowledge between labeled source domain data and unlabeled target domain data to improve accuracy on target domain data. As UDA requires no labels for target domain data, it is more practical in reality. Recently, methods based on deep learning have produced state-of-the-art results on UDA tasks [34], with DANN [35, 36] as one of the fundamental works.

In this paper, we propose a more reliable SEI approach, namely CWT-DANN, to identify emitters under varying frequencies. The proposed approach first turn signals to time-frequency distributions by CWT for more stable convergence. Then DANN is utilized to learn representations that are both discriminative for individual emitters and invariant for varying frequencies by aligning distributions of training data and test data. Based on the learned representations, the proposed approach performs preferably against approach without transfer learning. Furthermore, due to the unsupervised setting of the DANN, only labeled data of source domain and unlabeled data of target domain are required, which is more practical than other transfer learning approaches because the labels of the test data are commonly unknown.

The contributions of this paper are as follows:

- To the best of our knowledge, this is the first attempt to utilize unsupervised domain adaptation for SEI.
- We aim to improve performance of SEI under the condition that the carrier frequencies of test data differ from training data, which has not been adequately studied.
- We propose an approach, namely CWT-DANN, which performs unsupervised domain adaptation through DANN with CWT of signals as input, for SEI under varying frequencies.
- Experiments are conducted on real data of 5 emitters and 3 carrier frequencies. The results demonstrate the superior performance of the proposed approach. When the training data and test data are of different carrier frequencies, CWT-DANN shows superior performance against baseline methods.

The rest of the paper is organized as follows: In Section 2, we introduce the background and the main ideas of unsupervised domain adaptation methods. In Section 3, the problem to be solved is formally defined. Then, in Section 4, we explicate the details of CWT-DANN. In Section 5, experiments on data of

5 emitters with different carrier frequencies are conducted to evaluate the performance of the proposed approach. Finally, we conclude the paper and discuss the future work in Section 6.

## 2. Unsupervised domain adaptation

In supervised learning, it is typically assumed that the training data comes from the same distribution as the test data. However, in practice, the distributions of training data and test data are not guaranteed to be identical, which may produce shifts between the training domain and the test domain. If not properly considered, the domain shifts can degrade the performance seriously. Domain adaptation, a subcategory of transfer learning, aims to deal with domain shifts and improve performance in the target domain utilizing data in one or more relevant source domains, with the assumption that the domains are different but similar. Particularly, unsupervised domain adaptation methods, which only requires labeled source domain data and unlabeled target domain data, are more practical when dealing with domain mismatch between training and test data.

The past few years have witnessed significant progress in UDA methods, with a growing emphasis on deep learning based methods [37]. A plethora of recent work aligns source and target domains by creating a domain-invariant feature representation. A feature representation is domain-invariant if the features follow the same distribution regardless of whether the input data is from the source or target domain. With domain-invariant features, a classifier trained to perform accurately on source domain data may also generalize well on target domain data (Fig. 1).

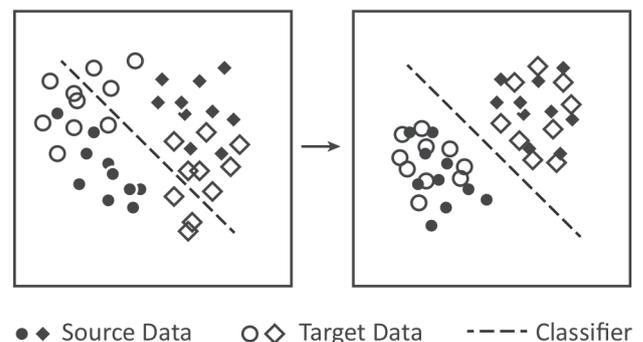


Fig. 1. Domain-invariant feature learning methods

DANN is a typical deep learning based UDA method, which learns domain-invariant features through adversarial training of neural networks [35, 36]. DANN consists of three sub-networks, namely feature extractor, label classifier and domain discriminator. The output of feature extractor is fed into label classifier and domain discriminator. Label classifier is a trivial network which predicts the class label of the input. Domain discriminator is trained to discriminate the domain of the input feature. Feature extractor is trained to assist the label classifier and fool the domain discriminator at the same time. On one hand, by assisting the label classifier, the feature extractor learns features

discriminative for class labels. On the other hand, by fooling the domain discriminator, the feature extractor learns features invariant across domains. Feature extractor and Domain discriminator are trained adversarially analogous to generator and discriminator in generative adversarial network (GAN) [38]. When domain discriminator can not be trained to determine the domain of the input feature anymore, the feature from feature extractor is deemed to be domain-invariant.

### 3. Problem description

Most previous research of SEI assumes that training data and test data are collected under the same condition. However, this assumption does not always hold in practice. For example, for wireless communications, the emitters may change carrier frequencies, which leads to variations of received signals from the same emitter. On one hand, the characteristics of hardware imperfections of one emitter are not constant at different carrier frequencies, because of phase noises of the local oscillator and frequency response of the amplifier. On the other hand, influences of wireless channels on signals vary with frequencies. These two factors lead to shifts of RFFs when the carrier frequencies of training data and test data differ and the performance of SEI may deteriorate seriously.

We denote the data space as  $\mathbf{X}$  and the label space as  $\mathbf{Y}$ . The training data, which contains received signals of emitters with certain carrier frequency, is denoted as source domain data set  $X^s = \{\mathbf{x}_1^s, \dots, \mathbf{x}_{n_s}^s\}$  of  $\mathbf{X}$ , where  $\mathbf{x}_i^s$  indicates the  $i$ th record of source domain data, with corresponding label set  $Y^s = \{y_1^s, \dots, y_{n_s}^s\}$  of  $\mathbf{Y}$ , where  $y_i^s$  indicates the emitter identity of  $\mathbf{x}_i^s$ . The test data, which contains the received signals of emitters with carrier frequency different from training data, is denoted as the target domain data set  $X^t = \{\mathbf{x}_1^t, \dots, \mathbf{x}_{n_t}^t\}$  of  $\mathbf{X}$ , with corresponding labels unknown. The task is to find an objective predictive function  $f(\cdot)$ , which can also be viewed as a conditional distribution  $P^t(y|\mathbf{x})$ , by utilizing  $X^s$ ,  $Y^s$  and  $X^t$ .

The conventional assumption that the training data and test data are collected under the same condition indicates that the

training data and the test data follow the same distribution, i.e.  $P^s(y|\mathbf{x}) = P^t(y|\mathbf{x})$ . In this case, a classifier built to learn  $P^s(y|\mathbf{x})$  using only  $X^s$  and  $Y^s$  also approximates  $P^t(y|\mathbf{x})$  and hence performs well on the test data. However, when the training data and the test data are collected under different conditions, the domain shifts lead to  $P^t(y|\mathbf{x}) \neq P^s(y|\mathbf{x})$ . Therefore, the classifier trained using only  $X^s$  and  $Y^s$  to learn  $P^s(y|\mathbf{x})$  does not approximate  $P^t(y|\mathbf{x})$  and generalize poorly on the test data. To learn  $P^t(y|\mathbf{x})$  and generalize on the test data, additional information of test data has to be introduced. UDA aims to alleviate the domain shifts and to transfer knowledge between training data and test data by exploiting  $X^t$  besides  $X^s$  and  $Y^s$ .

### 4. CWT-DANN

In this section, we introduce the details of CWT-DANN for SEI under varying carrier frequencies. The proposed approach is summarized in Fig. 2. We first turn the raw I/Q signals into time-frequency energy distributions by continuous wavelet transform (CWT) as suggested by [28]. Then the time-frequency distributions are used to train a network framework analogous to DANN, which consists of a feature extractor, a label classifier and a domain discriminator. After training, the feature extractor and the label classifier are deployed to predict the labels of test data.

**4.1. Signal preprocessing.** DANN employs CNN to extract features from the input samples. As CNNs are more suitable to extract features from images, the samples of signals are transformed into image-like representations [25]. Among several 2D representations of signals, the time-frequency energy distribution has shown better performance than other forms such as recurrence plots and bispectrums [28]. Short-time Fourier transform (STFT) is a simple way to get time-frequency distribution but has a fixed resolution. Hilbert-Huang transform (HHT) is another time-frequency distribution method, the performance of which is satisfying on simulated data [15, 25], but the obtained time-frequency distribution is sparse and hence not suit-

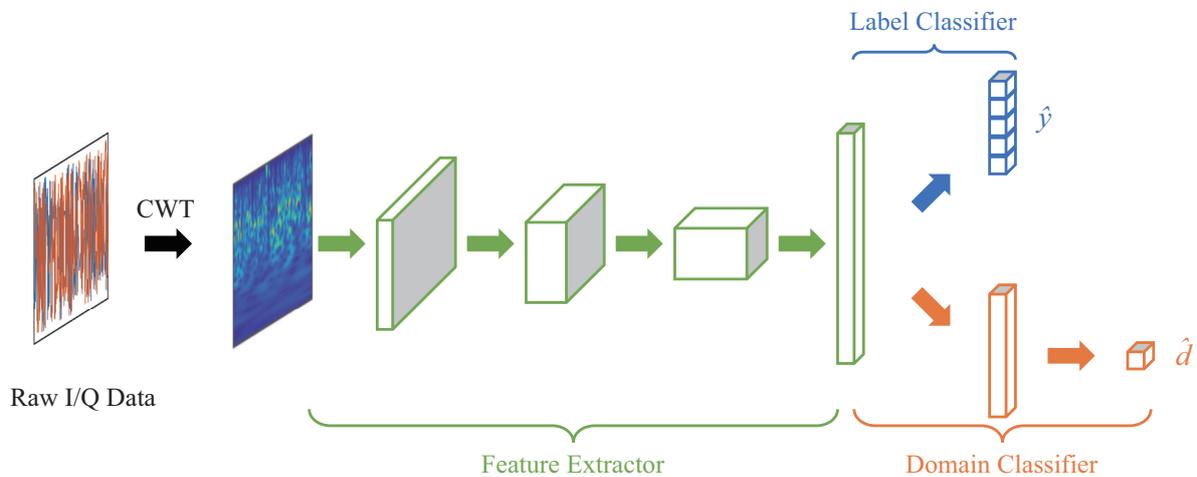


Fig. 2. DANN based SEI

able for CNN training. CWT is a classical method generating time-frequency distributions with adaptive resolution. Therefore, in this paper we adopt CWT to generate time-frequency distributions.

The samples of raw I/Q signals are preprocessed in the following procedure: First, each sample is normalized to mitigate the influence of transmitting power. Next, CWT of each normalized signal produces the scaleogram. The magnitude of the scaleogram representing time-frequency energy distribution is used for CNN training. Before being fed into a CNN, the magnitude of scaleogram is normalized for more stable convergence. Each sample contains 8192 sampling points and the number of scales in CWT is set as 101, therefore each input of DANN is of size  $8192 \times 101$ .

**4.2. Loss functions.** The basic idea of DANN is that predictions must be made based on domain-invariant features to achieve effective domain adaptation. In DANN, features are deemed to be domain-invariant if we cannot train a classifier to discriminate the domains of the features. Therefore, DANN introduces a domain discriminator network in addition to a feature extractor network and a label classifier network. The output of the feature extractor is guaranteed to be discriminative for emitter identities by co-operative training with the label classifier using labeled source domain data. At the same time, the features are supposed to be domain-invariant across carrier frequencies by adversarial training against the domain discriminator using both source domain data and target domain data without class labels. As the RFFs that the feature extractor outputs are both discriminative and domain-invariant after training, the label classifier that performs well on source domain data also generalizes effectively on target domain data.

More formally, the feature extractor can be regarded as a function mapping the input  $\mathbf{x}$  to a feature vector  $\mathbf{f}$ , i.e.  $\mathbf{f} = G_f(\mathbf{x}; \theta_f)$ , where  $\theta_f$  indicates the parameters of the feature extractor. The label classifier with parameters  $\theta_y$  takes the feature vector  $\mathbf{f}$  as input and outputs a vector  $\hat{\mathbf{y}}$  of dimension  $K$ , where  $K$  indicates the number of emitter identities and  $\hat{y}_i$  represents the estimated probability that the sample belongs to class  $i$ , i.e.  $\hat{\mathbf{y}} = G_y(\mathbf{f}; \theta_y)$ . Similarly, the feature vector  $\mathbf{f}$  is mapped by the domain discriminator with parameters  $\theta_d$  to a number  $\hat{d}$ , which represents the predicted probability of the input belonging to the source domain, i.e.  $\hat{d} = G_d(\mathbf{f}; \theta_d)$ .

The objective of the label classifier is to predict emitter identities of the inputs, therefore the loss is defined as the cross entropy between the predicted probabilities and the true emitter identity labels:

$$L_{cls}(X^s, Y^s, \theta_f, \theta_y) = -\mathbb{E}_{(\mathbf{x}, \mathbf{y}) \sim (X^s, Y^s)} \sum_{k=1}^K \mathbb{1}_{[k=y]} \ln \hat{y}_k, \quad (1)$$

$$\hat{\mathbf{y}} = G_y(G_f(\mathbf{x}; \theta_f); \theta_y),$$

In is natural logarithm function. Since the domain discriminator tries to discriminate the domains of inputs, the loss of the domain discriminator is defined as the binary cross entropy be-

tween the predicted probability and the true domain labels:

$$L_{dsc}(X^s, X^t, \theta_f, \theta_d) = -\frac{1}{2} \left( \mathbb{E}_{\mathbf{x} \sim X^s} [\ln \hat{d}] + \mathbb{E}_{\mathbf{x} \sim X^t} [\ln(1 - \hat{d})] \right), \quad (2)$$

$$\hat{d} = G_d(G_f(\mathbf{x}; \theta_f); \theta_d).$$

The loss of the feature extractor consists two parts. One required property of the outputs of the feature extractor is to be discriminative for emitter identities. This property can be guaranteed by updating the feature extractor to minimize  $L_{cls}$ . The other demanded property of the features is domain-invariance, i.e. the distributions  $S(\mathbf{f}) = \{\mathbf{f} = G_f(\mathbf{x}^s; \theta_f) | \mathbf{x}^s \sim X^s\}$  and  $T(\mathbf{f}) = \{\mathbf{f} = G_f(\mathbf{x}^t; \theta_f) | \mathbf{x}^t \sim X^t\}$  should be similar. In DANN this is achieved by training the feature extractor adversarially with the domain discriminator. If the domain discriminator can not be trained to distinguish which domain  $\mathbf{x}$  comes from based on  $\mathbf{f}$ , then it is assumed that  $S(\mathbf{f})$  and  $T(\mathbf{f})$  are similar. In [35], the adversarial loss of the feature extractor is defined to maximize the loss of the domain discriminator, i.e.  $L_{adv} = -L_{dsc}$ . However, this objective may cause the problem of gradient vanishing because the domain discriminator converges quickly early on during training. In this paper, we define  $L_{adv}$  as the cross entropy between the predicted probability and the uniform distribution as suggested by [40]:

$$L_{adv}(X^s, X^t, \theta_f, \theta_d) = - \sum_{c \in \{s, t\}} \mathbb{E}_{\mathbf{x} \sim X^c} \left[ \frac{1}{2} \ln \hat{d} + \frac{1}{2} \ln(1 - \hat{d}) \right]. \quad (3)$$

The total loss of the feature extractor is  $L_{cls} + \lambda L_{adv}$ , with  $\lambda$  as a hyperparameter weighting the relative importance of  $L_{cls}$  and  $L_{adv}$ .

Overall, the loss functions of CWT-DANN are summarized as Eqs. (4) and (5). In each training iteration, we first sample a mini-batch of labeled source domain data and a mini-batch of unlabeled target domain data. Then,  $\theta_f$  and  $\theta_y$  are updated through backpropagation of the loss function defined as Eq. (4), with parameters of the domain discriminator fixed as  $\hat{\theta}_d$ . At last, we update  $\theta_d$  by the loss in Eq. (5), with parameters of the feature extractor fixed as  $\hat{\theta}_f$ .

$$\theta_f^*, \theta_y^* = \arg \min_{\theta_f, \theta_y} L_{cls}(X^s, Y^s, \theta_f, \theta_y) + \lambda L_{adv}(X^s, X^t, \theta_f, \hat{\theta}_d), \quad (4)$$

$$\theta_d^* = \arg \min_{\theta_d} L_{dsc}(X^s, X^t, \hat{\theta}_f, \theta_d). \quad (5)$$

**4.3. Network architecture.** As the feature extractor is supposed to be deep enough to learn features that are both discriminative and domain-invariant, we adopt a simplified version of ResNet-18 [41] as the feature extractor, which contains ten convolution layers (Table 1). For each sample, the feature extractor takes the time-frequency distribution of size  $1 \times 8192 \times 101$  as input and outputs a feature vector of 512

dimensions. For the label classifier, we simply use one fully connected layer ( $512 \rightarrow K \rightarrow \text{Softmax}$ ). The domain discriminator is specified as a network of two fully connected layers ( $512 \rightarrow \text{Batchnorm} \rightarrow \text{ReLU} \rightarrow 256 \rightarrow 1 \rightarrow \text{Sigmoid}$ ), which is more complex than the label classifier, for the reason that it is theoretically proved that the hypothesis class generated by domain discriminator should include the hypothesis class generated by label classifier for effective domain adaptation [36].

Table 1  
Architecture of the feature extractor

Layer Name	Filter Size	Output Size
conv1	$7 \times 7, 32, \text{stride } 2$	$32 \times 4096 \times 51$
average pool	stride 2	$32 \times 2048 \times 26$
conv2_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix}, \text{stride } 1$	$64 \times 2048 \times 26$
conv3_x	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix}, \text{stride } 2$	$128 \times 1024 \times 13$
conv4_x	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix}, \text{stride } 2$	$256 \times 512 \times 7$
conv5_x	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix}, \text{stride } 2$	$512 \times 256 \times 4$
global average pool		512

## 5. Experimental evaluation

**5.1. Data collection.** Signals are collected from 5 emitters with carrier frequencies of 380 MHz, 450 MHz and 512 MHz respectively. The emitters transmit signals with 8 phase shift keying modulation. The sampling frequency of the receiver is 4 MHz. A signal and spectrum analyzer connected to an antenna receives the signals, converting the RF signals to the baseband. The baseband signals include in-phase signals and quadrature signals.

The signals are segmented by 8192 sampling points for each sample and then turned into matrices of size  $8192 \times 101$  by CWT. After preprocessing, we obtain 600 samples for each emitter at each carrier frequency. For the evaluation of training data and test data from the same carrier frequency, 500 samples of each emitter are randomly selected as training data and the rest 100 samples as test data. For evaluation of training data and test data from different carrier frequencies, 500 samples of each emitter from one frequency are randomly selected as training data and 500 samples of each emitter from another frequency as test data.

To investigate the influence of carrier frequency on emissions, we compare power spectrums of signals of the same emitter at different carrier frequencies. We denote the mean power spectrum of 300 samples of emitter 1 at 380 MHz as  $\hat{S}_1(f)$ , the mean power spectrum of the other 300 samples as

$\hat{S}'_1(f)$ , and the mean power spectrum of 300 samples of emitter 1 at 450 MHz as  $\hat{S}_2(f)$ . We calculate the relative difference of mean power spectrum at the same carrier frequency and different carrier frequencies as  $(\hat{S}_1(f) - \hat{S}'_1(f))/\hat{S}_1(f)$  and  $(\hat{S}_1(f) - \hat{S}_2(f))/\hat{S}_1(f)$  respectively. Figure 3 demonstrates that the spectrum difference at the same carrier frequency is roughly 0, while the difference at different carrier frequencies is not centered around 0 and shows patterns relative to frequency. The same phenomena also appears for emitter 2–5 and other frequency pairs. As only carrier frequency is changed in Fig. 3(b), this phenomena indicates that hardware characteristics vary with carrier frequency, which may cause shifts of RFFs and degrade the performance of SEI approaches.

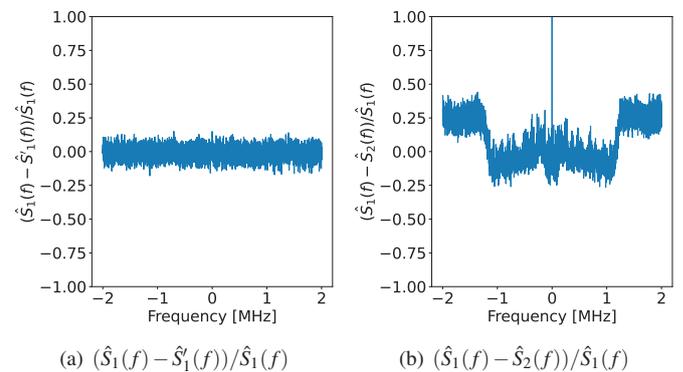


Fig. 3. Relative difference of mean power spectrum at (a) the same carrier frequency and (b) different carrier frequencies

**5.2. Baselines.** We utilize four methods to extract hand-crafted features for comparison. VMD-EM<sup>2</sup> and VMD-SF utilize variational mode decomposition (VMD) to decompose the received signal, then extract entropy, first and second order moments (EM<sup>2</sup>) of Hilbert transforms and spectral features (SF) of each mode [39]. We also replace VMD by empirical mode decomposition (EMD), obtaining feature extraction methods EMD-EM<sup>2</sup> [15], and EMD-SF. For each feature extraction method, we train a multi-class support vector machine (SVM) classifier for evaluation.

As for deep learning based methods for comparison, we simply use the source-only model of our approach, noted as CWT-CNN [28], meaning that only the feature extractor and the label classifier are trained using the training data, without domain discriminator in the network. The trained feature extractor and label classifier are then evaluated on test data without domain adaptation.

**5.3. Pre-experiment.** To verify the effectiveness of the baseline methods, we first evaluate their performance under the condition that the training data and test data are of the same carrier frequency. The CWT-CNN model is trained for 100 epochs, with batch size of 16, by RMSprop optimization method at a learning rate of  $10^{-4}$ . The decomposition order of VMD and EMD is set as 8. Each experiment is repeated 5 times for evaluation.

Table 2 shows the test accuracy of different methods under constant frequency. Data were reported as mean  $\pm$  standard error of replicate mean values. CWT-CNN is superior to methods based on hand-crafted features, justifying the effectiveness of training CNN to learn RFFs based on time-frequency distributions of signals transformed by CWT. For methods based on hand-crafted features, spectral features are better than entropy and moments, which agrees with the experimental results of [39]. The results also show that the accuracy relates to carrier frequency. For all methods based on hand-crafted features, the accuracy under carrier frequency of 450 MHz is obviously higher than that of 512 MHz, implying that the data distributions vary with carrier frequency.

Table 2  
Test accuracy under constant frequency

Method	380 MHz	450 MHz	512 MHz
VMD-EM <sup>2</sup>	0.8116 $\pm$ 0.0238	0.8596 $\pm$ 0.0168	0.8004 $\pm$ 0.0214
VMD-SF	0.9338 $\pm$ 0.0106	0.9550 $\pm$ 0.0077	0.9128 $\pm$ 0.0141
EMD-EM <sup>2</sup>	0.5862 $\pm$ 0.0255	0.6650 $\pm$ 0.0173	0.5752 $\pm$ 0.0185
EMD-SF	0.9454 $\pm$ 0.0131	0.9424 $\pm$ 0.0136	0.9046 $\pm$ 0.0132
CWT-CNN	<b>0.9924 <math>\pm</math> 0.0032</b>	<b>0.9988 <math>\pm</math> 0.0008</b>	<b>0.9996 <math>\pm</math> 0.0002</b>

**5.4. Results.** In this section, we compare the results of CWT-DANN and baseline methods under varying frequency. The training data are 500 random samples for each emitter of one carrier frequency and the test data are 500 random samples for each emitter of another carrier frequency. The training procedure of CNN-SEI and hand-crafted feature based methods is the same as the way under constant frequency. For CWT-DANN, the network is trained for 300 epochs by RMSprop. The learning rate is initially  $10^{-4}$ , and is reduced by half for every 100 epochs. Each experiment is repeated 5 times for evaluation.

**Convergence evaluation.** We first plot the learning curves of CWT-DANN to demonstrate its convergence. The learning curves of CWT-DANN at 380 MHz  $\rightarrow$  450 MHz (data of 380 MHz for training and data of 450 MHz for testing) are shown in Fig. 4 with 5 replicates of the same experiment. Although the test accuracy of each replicate oscillates seriously at early stage, as training proceeds the accuracy converges to around 0.95. The  $L_{cls}$  of each replicate converges to nearly zero, which is trivial since both the feature extractor and the label classifier are optimized to minimize  $L_{cls}$ . Based on the defini-

tion of  $L_{dsc}$  (Eq. (2)) and  $L_{adv}$  (Eq. (3)), we can deduce that the maximum of  $L_{adv}$  is  $\ln 2$ , which is represented as dashed lines in Fig. 4(c) and Fig. 4(d). During the first several epochs,  $L_{dsc}$  drops heavily and  $L_{adv}$  increases sharply, indicating that the domains can be easily discriminated by the domain discriminator at early stage. As training proceeds,  $L_{dsc}$  increases gradually and  $L_{adv}$  decreases gradually towards  $\ln 2$ , suggesting that the domain discriminator can not be trained to discriminate the domains and the feature extractor learns domain-invariant features.

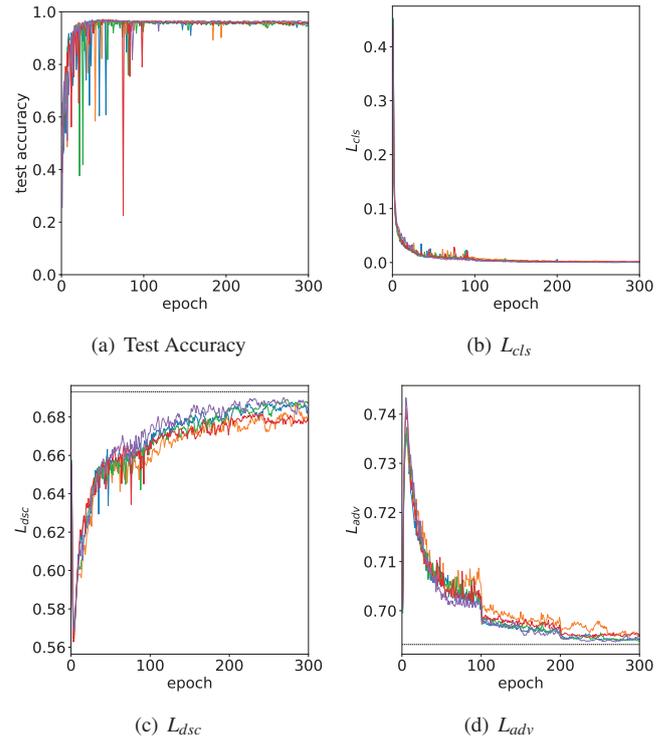


Fig. 4. Learning curves of CWT-DANN at 380 MHz  $\rightarrow$  450 MHz. Each curve correspond to one replicate (shown in a separate color), with dashed lines indicating the value of  $\ln 2$ .

**Accuracy comparison.** We successively select data of one carrier frequency for training and data of another frequency for testing, constructing 6 transfer tasks: 380 MHz  $\rightarrow$  450 MHz, 450 MHz  $\rightarrow$  380 MHz, 450 MHz  $\rightarrow$  512 MHz, 512 MHz  $\rightarrow$  450 MHz, 380 MHz  $\rightarrow$  512 MHz and 512 MHz  $\rightarrow$  380 MHz. The test accuracy of different methods are shown in Table 3.

Table 3  
Test accuracy under varying frequencies

Method	380 MHz $\rightarrow$ 450 MHz	450 MHz $\rightarrow$ 380 MHz	450 MHz $\rightarrow$ 512 MHz	512 MHz $\rightarrow$ 450 MHz	380 MHz $\rightarrow$ 512 MHz	512 MHz $\rightarrow$ 380 MHz
VMD-EM <sup>2</sup>	0.6456 $\pm$ 0.0109	0.5492 $\pm$ 0.0088	0.3130 $\pm$ 0.0054	0.4066 $\pm$ 0.0108	0.2656 $\pm$ 0.0088	0.2710 $\pm$ 0.0103
VMD-SF	0.7795 $\pm$ 0.0057	0.7830 $\pm$ 0.0077	0.5852 $\pm$ 0.0051	0.5325 $\pm$ 0.0131	0.5196 $\pm$ 0.0070	0.4477 $\pm$ 0.0062
EMD-EM <sup>2</sup>	0.4765 $\pm$ 0.0076	0.4104 $\pm$ 0.0104	0.2163 $\pm$ 0.0035	0.2436 $\pm$ 0.0111	0.1976 $\pm$ 0.0039	0.2615 $\pm$ 0.0069
EMD-SF	0.7258 $\pm$ 0.0082	0.7368 $\pm$ 0.0044	0.5696 $\pm$ 0.0077	0.5665 $\pm$ 0.0110	0.5393 $\pm$ 0.0084	0.5077 $\pm$ 0.0086
CWT-CNN	0.6827 $\pm$ 0.0185	0.7036 $\pm$ 0.0072	0.6826 $\pm$ 0.0068	0.7198 $\pm$ 0.0063	0.5069 $\pm$ 0.0146	0.6558 $\pm$ 0.0164
CWT-DANN	<b>0.9568 <math>\pm</math> 0.0056</b>	<b>0.9380 <math>\pm</math> 0.0091</b>	<b>0.9996 <math>\pm</math> 0.0004</b>	<b>0.9764 <math>\pm</math> 0.0081</b>	<b>0.6612 <math>\pm</math> 0.0098</b>	<b>0.7620 <math>\pm</math> 0.1152</b>

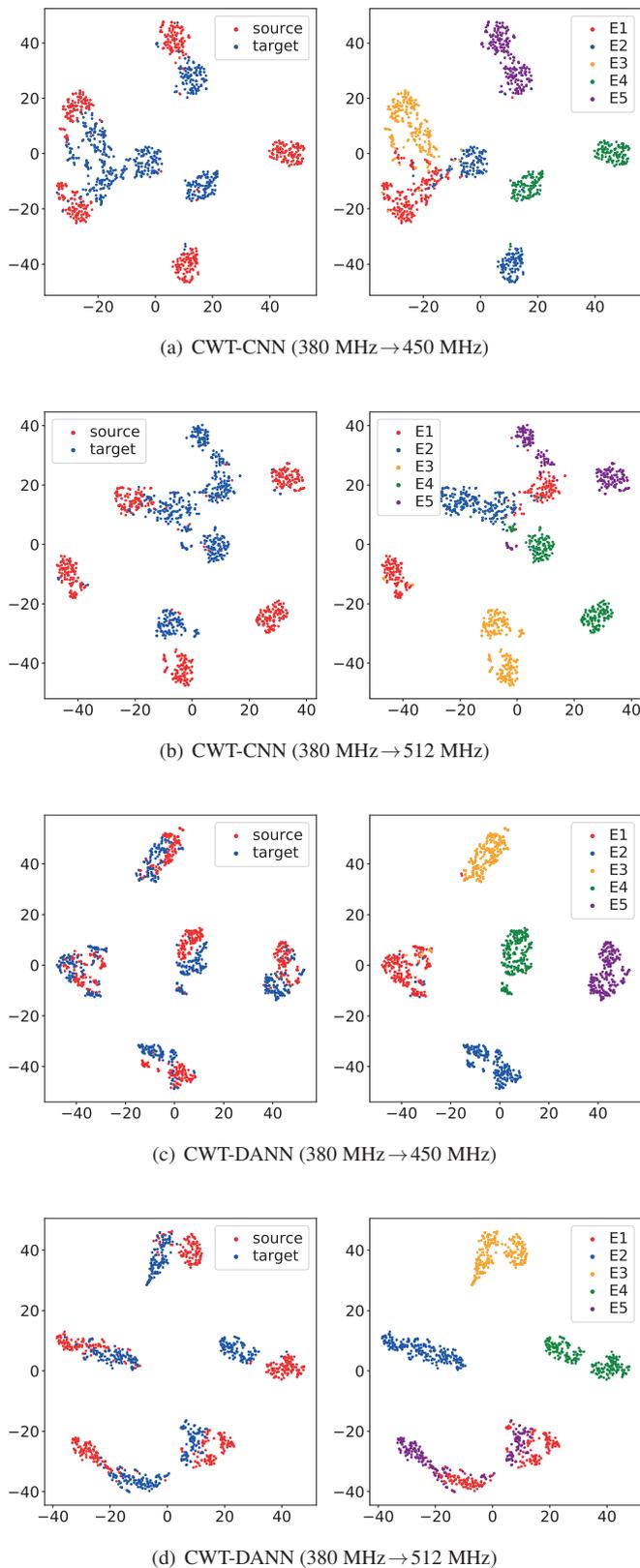


Fig. 5. Visualization of features extracted by the feature extractor. In the left column, red points correspond to source domain data and blue points correspond to target domain data. In the right column, different colors correspond to different emitter identities, denoted as E1–E5. Best viewed in color.

Data were reported as mean  $\pm$  standard error of replicate mean values. The performances of all baseline methods degrade significantly compared with the results in the pre-experiment. This degradation is caused by the domain shifts between training and test data. CWT-DANN performs considerably by utilizing unlabeled target domain data, with improvement of at least 10% over all baseline methods at all conditions, proving the superiority of CWT-DANN. For all methods, the accuracy when transferring between 380 MHz and 512 MHz is evidently lower than other conditions, which may suggest that hardware characteristics of the same emitter are more diverse when the difference between carrier frequencies is larger so that the data distributions are more dissimilar.

**Feature visualization.** We use t-SNE [42] to reduce the dimension of features extracted by the feature extractor to 2 and visualize the transformed features of CWT-CNN and CWT-DANN respectively in Fig. 5. As shown in Fig. 5(a) and Fig. 5(b), CWT-CNN learns features that are discriminative for source domain data. However, due to lack of consideration to the shifts between the source domain and the target domain, features of target domain data extracted by CWT-CNN are distributed away from source domain features. Therefore, CWT-CNN generalizes poorly on target domain data. Unlike CWT-CNN, CWT-DANN aligns the features of source domain data and target domain data by exploiting unlabeled target domain data (Fig. 5(c), Fig. 5(d)). Based on the domain-invariant features, the label classifier trained by the source domain labeled data generalizes preferably on target domain data. However, for 380 MHz → 512 MHz, the features of emitter1 and emitter5 across different domains are misaligned in Fig. 5(d). A possible reason of misalignment is that the source domain data and target domain data are too dissimilar because of the large frequency gap, which violates the assumption of domain similarity.

## 6. Conclusions

In this paper, we propose an approach, namely CWT-DANN, for SEI under varying frequency. Under the assumption that the distributions of training data and test data are different but similar, the approach exploits the unlabeled test data to align the source domain and target domain features, in order to perform effectively on test data by alleviating domain shifts. The results of experiments on real signals verify the effectiveness of the proposed approach, with an improvement of over 10% against baseline approaches under varying frequencies. As we only assume that the training data and test data are different, with no further assumptions of the cause of differences, the proposed approach is promising for other conditions that introduce domain shifts, such as modulation type, bandwidth and channel. The approach provides a novel paradigm for SEI when the training data and test data are not identically distributed, only with additional requirement of test data with no labels, which is practical in reality.

False alignment, a common issue of UDA, which arises when source domain data and target domain data are too dissimilar, may degrade the performance of CWT-DANN. Recently, meth-

ods using pseudo labels, i.e. the estimated labels of target domain data, have shown to be helpful [43]. An implicit assumption in this paper and in conventional UDA is that the label sets of the source domain and the target domain are identical. When this assumption is not satisfied, the performance of CWT-DANN is not guaranteed. A possible way of solving this issue is to develop unsupervised deep domain adaptation methods with relaxed assumption, such as open set domain adaptation [44], partial domain adaptation [45] and universal domain adaptation [46]. The application of these methods for more practical SEI approaches will be our future work.

**Acknowledgements.** The authors would like to acknowledge the support of Anhui Provincial Natural Science Foundation (NO. 1908085MF202) and Independent Scientific Research Program of National University of Defense Science and Technology (NO. ZK18-03-14).

## REFERENCES

- [1] K.I. Talbot, P.R. Duley, and M.H. Hyatt, "Specific emitter identification and verification", *Technol. Rev.* 2003, 113–133, (2003).
- [2] G. Baldini, G. Steri, and R. Giuliani, "Identification of wireless devices from their physical layer radio-frequency fingerprints", in: *Encyclopedia of Information Science and Technology*, pp. 6136–6146, 4th Edition, IGI Global, 2018.
- [3] A.E. Spezio, "Electronic warfare systems", *IEEE Trans. Microw. Theory Tech.* 50(3), 633–644 (2002).
- [4] O. Ureten and N. Serinken, "Wireless security through rf fingerprinting", *Can. J. Electr. Comp. Eng.* 32(1), 27–33 (2007).
- [5] S.U. Rehman, K.W. Sowerby, and C. Coghill, "Radio-frequency fingerprinting for mitigating primary user emulation attack in low-end cognitive radios", *IET Commun.* 8(8), 1274–1284 (2014).
- [6] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures", in: *Proceedings of the 14th ACM international Conference on Mobile Computing and Networking*, San Francisco, USA: ACM, 2008, pp. 116–127.
- [7] Y. Huang, et al., "Radio frequency fingerprint extraction of radio emitter based on i/q imbalance", *Procedia Computer Science* 107, 472–477 (2017).
- [8] L.J. Wong, W.C. Headley, and A.J. Michaels, "Specific emitter identification using convolutional neural network-based iq imbalance estimators", *IEEE Access* 7, 33544–33555 (2019).
- [9] G. López-Risueño, J. Grajal, and A. Sanz-Osorio, "Digital channelized receiver based on time-frequency analysis for signal interception", *IEEE Trans. Aerosp. Electron. Syst.* 41(3), 879–898 (2005).
- [10] C. Bertoncini, K. Rudd, B. Noursain, and M. Hinders, "Wavelet fingerprinting of radio-frequency identification (rfid) tags", *EEE Trans. Ind. Electron.* 59(12), 4843–4850 (2011).
- [11] J. Lundén and V. Koivunen, "Automatic radar waveform recognition", *IEEE J. Sel. Top. Signal Process.* 1(1), 124–136 (2007).
- [12] L. Li, H.B. Ji, and L. Jiang, "Quadratic time–frequency analysis and sequential recognition for specific emitter identification", *IET Signal Process.* 5(6), 568–574 (2011).
- [13] 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.* 8(13), 2404–2412 (2014).
- [14] J. Zhang, F. Wang, Z. Zhong, and O. Dobre, "Novel hilbert spectrum-based specific emitter identification for single-hop and relaying scenarios", in: *2015 IEEE Global Communications Conference (GLOBECOM)*, San Diego, USA, IEEE, 2015, pp. 1–6.
- [15] J. Zhang, F. Wang, O. Dobre, and Z. Zhong, "Specific emitter identification via Hilbert–Huang transform in single-hop and relaying scenarios", *IEEE Trans. Inf. Forensic Secur.* 11(6), 1192–1205 (2016).
- [16] Z. Tang and S. Li, "Steady signal-based fractal method of specific communications emitter sources identification", in: *Wireless Communications, Networking and Applications*, pp. 809–819, Springer, 2016.
- [17] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific emitter identification based on nonlinear dynamical characteristics", *Can. J. Electr. Comp. Eng.* 39(1), 34–41 (2016).
- [18] Y. Jia, S. Zhu, and L. Gan, "Specific emitter identification based on the natural measure", *Entropy* 19(3), 117 (2017).
- [19] J. Dudczyk and A. Kawalec, "Specific emitter identification based on graphical representation of the distribution of radar signal parameters", *Bull. Pol. Acad. Sci. Tech. Sci.* 63(2), 391–396 (2015).
- [20] Y. Zhao, Y. Li, L. Wui, and J. Zhang, "Specific emitter identification using geometric features of frequency drift curve", *Bull. Pol. Acad. Sci. Tech. Sci.* 66(1), 99–108 (2018).
- [21] L. Rybak and J. Dudczyk, "A geometrical divide of data particle in gravitational classification of moons and circles data sets", *Entropy* 22(10), 1088 (2020).
- [22] Q. Wu, et al., "Deep learning based rf fingerprinting for device identification and wireless security", *Electron. Lett.* 54(24), 1405–1407 (2018).
- [23] L. Ding, S. Wang, F. Wang, and W. Zhang, "Specific emitter identification via convolutional neural networks", *IEEE Commun. Lett.* 22(12), 2591–2594 (2018).
- [24] K. Merchant, S. Revay, G. Stantchev, and B. Noursain, "Deep learning for rf device fingerprinting in cognitive communication networks", *IEEE J. Sel. Top. Signal Process.* 12(1), 160–167 (2018).
- [25] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, "Specific emitter identification based on deep residual networks", *IEEE Access* 7, 54425–54434 (2019).
- [26] J. Matuszewski and D. Pietrow, "Recognition of electromagnetic sources with the use of deep neural networks", in *XII Conference on Reconnaissance and Electronic Warfare Systems*, 2019, vol. 11055, pp. 100–114, doi: 10.1117/12.2524536.
- [27] L.J. Wong, W.C. Headley, S. Andrews, R.M. Gerdes, and A.J. Michaels, "Clustering learned cnn features from raw i/q data for emitter identification", in: *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*, Los Angeles, USA, 2018, pp. 26–33.
- [28] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks", *Electron. Lett.* 55(2), 90–92 (2018).
- [29] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation", *IEEE Trans. Inf. Forensic Secur.* 11(9), 2091–2106 (2016).
- [30] S. Andrews, R.M. Gerdes, and M. Li, "Towards physical layer identification of cognitive radio devices", *IEEE Conference on Communications and Network Security (CNS)*, Las Vegas, USA, IEEE, 2017, pp. 1–9.

- [31] I.F. Akyildiz, W.Y. Lee, M.C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey", *Comput. Netw.* 50(13), 2127–2159 (2006).
- [32] S.J. Pan and Q. Yang, "A survey on transfer learning", *IEEE Trans. Knowl. Data Eng.* 22(10), 1345–1359 (2009), doi: 10.1109/TKDE.2009.191.
- [33] Y. Sharaf-Dabbagh and W. Saad, "Transfer learning for device fingerprinting with application to cognitive radio networks", in: *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, China, 2015, pp. 2138–2142.
- [34] M. Wang and W. Deng, "Deep visual domain adaptation: A survey", *Neurocomputing* 312, 135–153 (2018). doi: 10.1016/j.neucom.2018.05.083.
- [35] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation", in: *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015*, Lille, France, 2015, pp. 1180–1189.
- [36] Y. Ganin, et al., "Domain-adversarial training of neural networks", *J. Mach. Learn. Res.* 17(1), 2096–2030 (2016).
- [37] G. Wilson and D.J. Cook, "A survey of unsupervised deep domain adaptation", CoRR, 2018, abs/1812.02849. Available from: <http://arxiv.org/abs/1812.02849>.
- [38] I. Goodfellow, et al., "Generative adversarial nets", in: *Advances in Neural Information Processing Systems*, Montreal, Canada, 2014, pp. 2672–2680.
- [39] U. Satija, N. Trivedi, G. Biswal, and B. Ramkumar, "Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios", *IEEE Trans. Inf. Forensic Secur.* 14(3), 581–591 (2018).
- [40] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation", in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii, 2017, pp. 7167–7176.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", in: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 2016, pp. 770–778.
- [42] L. Maaten and G. Hinton, "Visualizing data using t-sne", *J. Mach. Learn. Res.* 9, 2579–2605 (2008).
- [43] C. Chen, et al., "Progressive feature alignment for unsupervised domain adaptation", in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 627–636.
- [44] P. Panareda-Busto and J. Gall, "Open set domain adaptation", in: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 754–763.
- [45] Z. Cao, M. Long, J. Wang, and M.I. Jordan, "Partial transfer learning with selective adversarial networks", in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, USA, 2018, pp. 2724–2732.
- [46] K. You, M. Long, Z. Cao, J. Wang, and M.I. Jordan, "Universal domain adaptation", in: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, USA, 2019.