Semi-Supervised Specific Emitter Identification via Dual Consistency Regularization

Xue Fu, *Graduate Student Member, IEEE*, Shengnan Shi, *Member, IEEE*, Yu Wang, *Member, IEEE*, Yun Lin, *Member, IEEE*, Guan Gui, *Senior Member, IEEE*, Octavia A. Dobre, *Fellow, IEEE*, and Shiwen Mao, *Fellow, IEEE*

***Abstract*—Deep learning (DL)-based specific emitter identification (SEI) is a potential physical layer authentication technique for Industrial Internet-of-Things (IIoT) Security, which detects the individual emitter according to its unique signal features resulting from transmitter hardware impairments. The success of DL-based SEI often depends on sufficient training samples and the integrity of samples' labels. The extensive deployment of wireless devices generates a huge amount of signals, but signals labeling is quite difficult and expensive with the high demand for expertise. In this article, we present an SEI method based on dual consistency regularization (DCR), which enables feature extraction and identification using a few labeled samples and a large number of unlabeled samples. With the help of pseudo labeling, we leverage consistency between the predicted class distribution of weakly augmented unlabeled training samples and that of strongly augmented training unlabeled samples, and consistency between semantic feature distribution of labeled samples and that of pseudo-labeled samples, which takes the unlabeled samples into account to model parameter tuning for a more accurate emitter identification. Extensive numerical results demonstrate that compared with well-known semi-supervised learning-based SEI methods, our method obtains 99.77% identification accuracy on a WiFi data set and 90.10% identification accuracy on an automatic dependent surveillance-broadcast (ADS-B) data set when only 10% of training samples are labeled, and improves the identification accuracy on the WiFi data set and the ADS-B data set by more than 19.07% and 5.30%, respectively. Our codes are available at https://github.com/lovelymimola/DCR-Based-SemiSEI.**

***Index Terms*—Consistency regularization, deep metric learning, pseudo labeling, semi-supervised learning, specific emitter identification (SEI).**

Manuscript received 13 February 2023; revised 3 April 2023; accepted 27 May 2023. Date of publication 31 May 2023; date of current version 24 October 2023. This work was supported in part by the Key Project of Natural Science Foundation of the Higher Education Institutions of Jiangsu Province under Grant 22KJA510002. *(Corresponding author: Guan Gui.)*

Xue Fu, Shengnan Shi, Yu Wang, and Guan Gui are with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: 1020010415@njupt.edu.cn; dsn@njupt.edu.cn; yuwang@njupt.edu.cn; guiguan@njupt.edu.cn).

Yun Lin is with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150009, China (e-mail: linyun@hrbeu.edu.cn).

Octavia A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1B 3X5, Canada (e-mail: odobre@mun.ca).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Digital Object Identifier 10.1109/JIOT.2023.3281668

I. Introduction

INDUSTRY 4.0 comprises scenarios where the entire production cycle, from its very initial conception phase to the design, rollout, and operation, is accomplished with the use and support of a variety of digital technologies [1]. Communications is the fundamental backbone of this digital environment. However, the open nature of the wireless medium provides more security vulnerabilities and authentication is an important issue in wireless communications [2]. Specific emitter identification (SEI) is a potential physical layer authentication technique for avoiding spoofing attack in Industrial Internet-of-Things (IIoT), which achieves an authentication by comparing the physical-layer features of an unknown emitter with one of the legitimate emitter [3].

Deep learning (DL) has made considerable progress over last few years, and DL methods play an vital role in shifting the security of Internet-of-Things (IoT) systems from merely facilitating secure communication between devices to security-based intelligence systems [4], [5]. Consequently, a large amount of DL-based SEI methods have been proposed for physical layer authentication. Driven by the training samples, the parameters of deep neural network are tuned for minimizing the value of objective function through forward propagation and backward propagation, realizing the mapping operation from sample space to category space for a given testing sample. The existing DL-based SEI methods can be classified into four categories with respect to the input data format: 1) in-phase/quadrature (I/Q)-based; 2) constellation-based; 3) spectrogram-based; and 4) multimodal imformation-based.

Chen et al. [6] used I/Q and inception-residual neural network for 5157 automatic dependent surveillance-broadcast (ADS-B) signals classification and 3143 aircraft communications addressing and reporting system (ACARS) signals classification. Wang et al. [7] used I/Q and deep complex residual network for 20 WiFi network card devices identification. Zha et al. [8] used I/Q and complex Fourier neural network for seven simulated emitters classification. Peng et al. [9] used heat constellation trace figure and deep neural networks for seven power amplifiers classification. Ding et al. [10] used compressed bispectrum and convolutional neural networks for multiple universal software radio peripherals (USRPs) identification. Pan et al. [11] employed Hilbert–Huang spectrum and deep residual network for five simulated emitters classification. Shen et al. [12] used channel independent spectrogram

generated by the short-time fourier transform (STFT) and convolutional neural network for long range radio (LoRa) devices classification. Liu [13] used multimodal information consisting of amplitude, phase, and spectrum asymmetry and convolutional neural network for 20 aircraft transponders classification. Zhao et al. [14] used multimodal information consisting of I/Q, frequency spectrum, amplitude, and phase and deep residual network for 20 aircraft transponders classification.

Sensing and capturing over-the-fly radio signals in the form of I/Q signals can be performed using software definition radio (SDR) platforms. However, I/Q signals annotation is a time-consuming and labor-exhaustive task because an expert is required to identify and annotate each captured signal and the number of unknown signals increases the complexity of the signal annotation task in unknown communication environments [15]. Driven by sufficiently labeled training samples, the above DL-based SEI methods achieve good identification performance, but they will face the risk of declining identification accuracy when the labeled training samples is not sufficient. Therefore, SEI methods in the case of limited-labeled training samples have attracted attention. Transfer learning [16], semi-supervised learning [17], unsupervised learning [18], and data augmentation [19] have been introduced into SEI in the case of limited-labeled training samples, where semi-supervised learning [20], [21], [22] improves the efficiency of deep models by learning from a few labeled training samples and a large number of unlabeled training samples (UTSs).

In this article, we focus on SEI in the case of limited labeled training samples and semi-supervised learning. A Semi-supervised SEI (SemiSEI) method based on dual consistency regularization (DCR) is proposed and referred to as the DCR-based SemiSEI method. We leverage a self-learning framework to propagate label information from the labeled to the UTSs in the form of DCR. Specifically, we adopt confidence thresholds to generate pseudo labels on weakly augmented UTSs and utilize these pseudo labels as annotations for strongly augmented UTSs to enhance the predicted class consistency between weakly augmented UTSs and strongly augmented UTSs; in addition, we leverage deep metric learning to efficiently enhance the semantic feature consistency between labeled and pseudo-labeled training samples which belong to same category. The DCR-based SemiSEI method takes into account both labeled and UTSs in model training and consequently obtains better identification performance.

The main contributions of this article are summarized as follows.

1) We present an SEI method based on DCR that are predicted class consistency and semantic feature consistency, which enables features extraction and identification without the need of sufficiently labeled training samples.
2) We present a data augmentation method for labeled and UTSs to increase the training samples diversity and avoid confirmation bias, while introducing entropy minimization to regularize the predicted class consistency of UTSs with different disturbance and semi-supervised metric learning to regularize the semantic feature consistency of labeled and pseudo-labeled training samples.
3) We evaluate the proposed DCR-based SemiSEI method on a WiFi data set with 16 categories and an ADS-B data set with ten categories. The simulation results show that the proposed method achieves a better identification accuracy and feature discrimination compared with five latest SemiSEI methods.

II. Related Works

Semi-supervised learning can be categorized in deep generative methods, consistency regularization methods, graph-based methods, pseudo-labeling methods, and hybrid methods [23]. To the best of our knowledge, there is little semi-supervised signal recognition method based on graph. Therefore, as shown in Table I, we focus on reviewing the semi-supervised signal recognition methods based on deep generative model, consistency regularization, pseudo-labeling, or hybrid ideas. In addition, we particularly care about how these methods use labeled and UTSs to drive the training process of deep neural networks.

A. Deep Generative Methods

Deep generative methods, such as generative adversarial networks (GANs) [24], convolutional autoencoders (CAEs) [25], and their variants, can learn the data distribution from UTSs. Therefore, SemiSEI methods based on an unsupervised component, such as auxiliary classifier GAN (ACGAN) [26], deep convolution GAN (DCGAN) [27], conditional GAN (CGAN) [28], Triple-GAN [29], improved GAN [31], or CAE and a supervised classification backbone have been proposed in the literatures as follows.

1) ACGAN: Tu et al. [30] presented a semi-supervised ACGAN for automatic modulation classification (AMC), and some training tricks of improved GAN [31], such as feature matching, minibatch discrimination, and classification backbone, were introduced into the ACGAN. The labeled and unlabeled samples were used as real samples to train the ACGAN and the labeled samples participated in the training process of discriminator for enabling its classification capability. However, this framework was effective in small labeled data set but not in larger labeled data set.

2) DCGAN: Similar to [30], Zhou et al. [32] proposed a semi-supervised DCGAN for AMC and SEI, and some training tricks of improved GAN [31], such as one-side label smooth, feature matching, and classification backbone, were introduced into the DCGAN. The unlabeled samples were used as real samples to train the DCGAN and the labeled samples participated in the training process of discriminator for enabling its classification capability. This framework was also effective in small labeled data set, but its effectiveness in larger labeled data set was not been analyzed.

3) CGAN: Similar to [30] and [32], Tan et al. [33] proposed a semi-supervised CGAN for SEI in multiple communication scenarios, and some training tricks of improved

TABLE I
RELATED WORKS

Methods	Reference	Framework	Task	Input Format	Categories	Accuracy
Deep Generative Methods	2018 [30]	GAN	AMC	Contour Stella Image	8	14 dB, 10%, 75.36%
	2019 [34]	GAN and CAE	SEI	I/Q	5	20 dB, 10%, 96.40%
	2020 [32]	GAN	AMC	I/Q	11	12 dB, 5%, 91.80%
			SEI	I/Q	10	Unknown, 20%, 89.80%
	2021 [35]	GAN and CAE	AMC	I/Q	11	4 dB, 4%, 95.20%
			ISR	I/Q	8	4 dB, 15%, 96.10%
	2022 [33]	GAN	SEI	Bispectrum	5	20 dB, 20%, 84.60%
	2019 [15]	CAE	AWTC	I/Q	16	20 dB, 10%, 70.01%
	2020 [36]	CAE	AMC	I/Q	4	0 dB, 5%, 97.50%
Consistency Regularization Methods	2020 [38]	Temporal Ensembling	ISR	I/Q	15	0 dB, 2%, 90.01%
	2021 [40]	Contrastive learning	AMC	I/Q	11	10 dB, 1%, 70.01%
	2022 [42]	UDA	AMC	Multimodal Information	11	10 dB, 1%, 95.01%
	2022 [44]	VAT	SEI	Bispectrum	6	10 dB, 10%, 94.50%
	2022 [46]	VAT	SEI	I/Q	10	30 dB, 10%, 80.01%
Pseudo -Labeling Methods	2017 [48]	Pseudo Label	SEI	STFT Spectrogram	13	Unknown, 10%, 96.01%
	2022 [49]	Pseudo Label	SEI	STFT Spectrogram	7	Unknown, Unknown, 98.01%
	2022 [50]	Pseudo Label	AMC	FFT Spectrogram	11	0 dB, 7%, 93.09%
	2022 [51]	Pseudo Label	AMC	STFT Spectrogram	6	10 dB, 10%, 100.00%
	2022 [53]	Meta Pseudo Label	AMC	STFT Spectrogram	15	Unknown, 1%, 91.25%
Hybrid Methods	2018 [56]	Deep Generative Model & Pseudo-Labeling	AMC	I/Q	11	4 dB, 50%, 91.01%
	2018 [57]	Deep Generative Model & Pseudo-Labeling	AMC	I/Q	11	18 dB, 50%, 84.01%
	2021 [58], [59]	Consistency Regularization & Pseudo-Labeling	AMC	I/Q	11	12 dB, 5%, 93.80%
	2023 [17]	Consistency Regularization & Pseudo-Labeling	SEI	I/Q	10	30 dB, 10%, 83.80%

Tips: In the column of accuracy such as (14dB, 10%, 75.36%), three numbers means that when the signal-to-noise ratio is 14 dB and the number of labeled training samples to the number of all training samples ratio is 10%, the identification accuracy is 75.36%.

GAN [31], such as feature matching and classification backbone, were introduced into the CGAN. The unlabeled samples were used as real samples to train the CGAN and the labeled samples participated in the training process of discriminator for enabling its classification capability.

4) Triple-GAN: Considers that the I/Q signals contain more information, but using I/Q signals with high dimension to train the generator and discriminator brings great difficulty in the training process, Gong et al. [34] introduced an autoencoder into semi-supervised Triple-GAN termed as Quadruple-GAN. The labeled and unlabeled samples were compressed to be feature samples with low dimension using autoencoder, and then the feature samples participated in the collaborative learning of Triple-GAN. Xu et al. [35] presented an improved quadruple-GAN for interference signal recognition (ISR) and AMC, which was an improved version of the method [34] with applications on ISR and AMC, an enhanced collaborative learning, modified objective functions, and lightweight models based on knowledge distillation for inference phase.

5) CAE: Camelo et al. [15] presented a semi-supervised CAE for automatic wireless technology recognition (AWTR), where the encoder and decoder were trained in an unsupervised way using only unlabeled samples; then the frozen encoder together with a classifier were trained in a supervised way using labeled samples; and finally, the de-frozen encoder together with a classifier were fine-tuned in a supervised way using labeled samples. Different from the staged training in [15], Wang et al. [36] designed a multiple-objective function to synchronously train a CAE together with a classifier for AMC, where the labeled samples were fed into encoder and classifier and the unlabeled samples were fed into encoder and decoder to drive the training process.

The above deep generative model-based methods compared with the same frameworks without corresponding generator or decoder have obvious identification performance improvement, especially in small labeled data set. However, GAN-based methods have the disadvantage of training instability and strong probability of model collapse and many training tricks are required to overcome them; CAE incorporating unlabeled samples into model training only through reconstruction objection cannot fully utilize the information embedded in unlabeled samples. The current research trend of semi-supervised signal recognition methods has shifted to consistency regularization, pseudo-labeling or hybrid idea including our proposed SemiSEI method.

B. Consistency Regularization Methods

According to the smoothness assumption or the manifold assumption, the consistency regularization applies the similarity constraints to the final loss functions, which the deep models are encouraged to output similar predictions of perturbed versions of the same signals [37]. Therefore, a simple and efficient SemiSEI method can be provided by combining the consistency regularization with a supervised classification backbone. Consistency regularization combined with advanced data augmentation significantly improves the diversity of samples and features, fully extracts and utilizes

the information of unlabeled samples, and offers superior identification performance.

1) Temporal Ensembling: Huang et al. [38] proposed a temporal ensembling [39]-based semi-supervised ISR method. The labeled samples were fed into convolutional neural network to produce predictions; for these predictions, the cross-entropy loss between the predicted labels and true labels, and the mean square error between the latest predictions and the ensemble predictions were computed. The unlabeled samples were also fed into convolutional neural network to produce predictions, but only the mean square error between the latest predictions and the ensemble predictions were computed.

2) Contrastive Learning: Liu et al. [40] presented a SimCLR [41]-based semi-supervised AMC method. The unlabeled samples were used to train an encoder that preferred mapping such similar inputs (an input rotated with different angles) to nearby locations in the latent space, and the labeled samples were used to train a classifier which took the representations (out of frozen encoder) as input.

3) Unsupervised Domain Adaption: Deng et al. [42] introduced the unsupervised domain adaption [43] into the semi-supervised AMC method. The labeled samples were considered as source domain and the unlabeled samples were considered as target domain.

4) Virtual Adversarial Training: Xie et al. [44] proposed a virtual adversarial training [45]-based SemiSEI method. The labeled and unlabeled samples were used to encourage the neural network to output similar predictions for the perturbed ones, and the labeled samples were used to train the neural network in a supervised way. Fu et al. [46] presented a virtual adversarial training-based SemiSEI method, and introduced deep metric learning into the proposed framework to extract discriminative features.

The above consistency regularization-based methods combined with data augmentation encourages the neural network to output similar predictions for the original sample and the perturbed one, but does not encourage the neural network to output similar predictions for the samples which have same pseudo labels or labels. In this article, we introduce the pseudo-labeling and deep metric learning into consistency regularization to solve this shortcoming.

C. Pseudo-Labeling Methods

Based on the prediction of high-confidence model, pseudo-labeling methods are employed to generate pseudo labels for UTSs, and then use these pseudo-labeled samples to regularize the training process of deep model [23].

1) Pseudo Label: Longi et al. [48] presented an SEI method based on pseudo label. The framework iteratively generated pseudo labels and additionally weighted the importance of the individual samples based on the number of the unlabeled samples and the learning phase. Similar to [48], Yang et al. [49] and Yuan et al. [50] introduced the pseudo label into the SEI. Xu et al. [51] applied the interpolation consistency training [52] in AMC, where the labeled samples were fed into neural network to calculate the cross entropy loss, and the pseudo labels of mixing-unlabeled samples were predicted, and the mixing-unlabeled samples were fed into the neural network to calculated the cross entropy loss.

2) Meta Pseudo Labels: For decreasing the confirmation bias of the pseudo label, Ren et al. [53] introduced meta pseudo labels [54] into SEI. The labeled samples were used to train the teacher model, and then the pseudo labels of unlabeled samples were predicted by teacher model and were used to train the student model, and finally the cross-entropy loss of labeled samples on student model were used to optimize the parameters of the teacher model.

The pseudo label is one popular semi-supervised learning method, but it suffers from confirmation bias [55], where the model overfits to incorrect pseudo labels predicted by owns. Mixing samples is an effective regularization techniques for reducing it [55]. This well matches our approach that drives the training process of deep neural network using mixing-rotating signals and its corresponding pseudo labels generated from rotating signals to reduce the confirmation bias.

D. Hybrid Methods

Hybrid methods are composed of the combination of various methodologies, such as generative model, consistency regularization, pseudo labeling, data augmentation, and other components. The superiority of multiple components than single components were demonstrated in methods [17], [56], [58], [59].

1) Deep Generative Model Combined With Pseudo Labeling: Li et al. [56] designed a DCGAN with a classifier for AMC. For each training round, the DCGAN with classifier was trained in a supervised way using labeled samples, and then the classifier predicted the labels of the unlabeled samples, and those positive samples with high confidence were selected as newly labeled samples for the next training round. Li et al. [57] introduced a conditional variational autoencoder and a spatial transformer network to optimize the training stability of [56].

2) Consistency Regularization Combined With Pseudo Labeling: Dong et al. [58], [59] combined the unsupervised data augmentation with pseudo label for AMC. The labeled samples and pseudo-labeled samples were used to train a neural network in a supervised way, and unlabeled samples were used to train the neural network that preferred mapping such similar inputs (original version and the noise-added version) to similar probabilistic distributions. Fu et al. [17] combined virtual adversarial training with pseudo label for SEI. Different from [46], this method introduced the pseudo labels into deep metric learning, where the improved metric learning operated not only on labeled training samples but also UTSs.

According to our survey on related works, there are few signal recognition methods based on hybrid methodologies; methods [56] and [57] have risks on confirmation bias, and methods [17], and [58], [59] have risks on model generalization because simple data augmentation only propagates the label information to the directly connected neighbors of the samples. In this article, we used a weak data augmentation and a strong data augmentation for the SemiSEI method

because the augmentation diversity leads to more neighbors for each sample and consequently the identification performance is improved.

III. Problem Formulation

A. DL-Based SEI

Let $\mathcal{R}$ and $\mathcal{Y}$ be the sample space and category space, respectively. Given a training data set $D_t = \{(\mathbf{r}_i, y_i)|i = 1, \ldots, N, \mathbf{r} \in \mathcal{R}, y \in \mathcal{Y}\}$, the goal of DL-based SEI problem is to produce a mapping function $f \in \mathcal{F} : \mathcal{R} \to \mathcal{Y}$ and its expected error is minimized approximately, i.e.,

$$\min_{f\in\mathcal{F}} \mathbb{E}_{(\mathbf{r},y)\sim\mathcal{D}_t} \mathcal{L}_s(f(\mathbf{r}), y) \tag{1}$$

where $\mathbb{E}$ represents the computing mathematical expectation, and $\mathcal{L}_s(f(\mathbf{r}), y)$ stands for the loss that compares the prediction $f(\mathbf{r})$ to its ground-truth category. The parameters of deep neural network are tuned by forward propagation and back propagation to minimize the error, so that the deep neural network has the mapping function from sample space to category space. More supervised samples contained in $\mathcal{D}_t$ will bring more constraints on f and then it will bring a good generalization.

B. General SemiSEI

In a SemiSEI problem, the training data set $\mathbf{D}_t$ is composed of labeled training data set $\mathbf{D}_l$ and unlabeled training data set $\mathbf{D}_{ul}$, where their relationship can be described as $\mathbf{D}_t = \mathbf{D}_l \cup \mathbf{D}_{ul}$, $\mathbf{D}_l = \{(\mathbf{r}_l^n, y_l^n)|n = 1, \ldots, L, \mathbf{r}_l \in \mathcal{R}, y \in \mathcal{Y}\}$ consists of L labeled training samples, $\mathbf{D}_{ul} = \{\mathbf{r}_{ul}^m|m = 1, \ldots, N-L, \mathbf{r}_{ul} \in \mathcal{R}\}$ consists of $N-L$ UTSs. The goal of a general SemiSEI is also to produce a mapping function $f \in \mathcal{F} : \mathcal{R} \to \mathcal{Y}$ but approximately minimize the expected error as follows:

$$\min_{f\in\mathcal{F}} \mathbb{E}_{(\mathbf{r}_l,y_l)\sim\mathbf{D}_l} \mathcal{L}_s(f(\mathbf{r}_l), y_l) + \omega \mathbb{E}_{\mathbf{r}_{ul}\sim\mathbf{D}_{ul}} \mathcal{L}_{us}\big(f(\mathbf{r}_{ul}), \mathcal{T}_{\mathbf{r}_{ul}}\big) \tag{2}$$

where $\mathcal{L}_{us}(\mathbf{r}_{ul}, \mathcal{T}_{\mathbf{r}_{ul}})$ is an unsupervised loss that encourages the model to produce predictions which are similar to the target $\mathcal{T}_{\mathbf{r}_{ul}}$ for UTSs, and ω is a weighting scalar for balancing the loss terms.

C. Proposed SemiSEI Method

In general SemiSEI methods, different semi-supervised techniques differ in how they generate the target $\mathcal{T}_{\mathbf{r}_{ul}}$ of (2). In the proposed SemiSEI method, the $\mathcal{L}_{us}$ is predicted class consistency of $\mathbf{D}_{ul}$, which extracts the data distribution embedded in $\mathbf{D}_{ul}$ to tune model parameters for a more accurate emitter identification. To further build the connection between $\mathbf{D}_l$ and $\mathbf{D}_{ul}$, we introduce another consistency regularization to constrain their semantic feature consistency. Therefore, the expected error that is minimized approximately by the proposed SemiSEI method can be formulated as

$$\begin{aligned}\min_{f\in\mathcal{F}} \varepsilon_{em} = \min_{f\in\mathcal{F}} & \mathbb{E}_{(\mathbf{r}_l,y_l)\sim\mathbf{D}_l} \mathcal{L}_s(f(\mathbf{r}_l), y_l) \\ & + \omega_1 \mathbb{E}_{\mathbf{r}_{ul}\sim\mathbf{D}_{ul}} \mathcal{L}_{us}\big(f(\mathbf{r}_{ul}), \mathcal{T}_{\mathbf{r}_{ul}}\big) \\ & + \omega_2 \mathbb{E}_{(\mathbf{r}_*,y_*)\sim\mathbf{D}_t} \mathcal{L}_c\big(f(\mathbf{r}_*), \mathcal{T}_{y_*}\big)\end{aligned} \tag{3}$$

TABLE II
Details of CVNN

Module	Details of CVNN		Number of layers
	for long signals	for short signals	
Features Extractor	Complex Conv1D (64,3) + ReLU + BatchNorm1D + MaxPool1D (2)		×9
	Flatten		×1
	LazyLinear (1024)	LazyLinear (512)	×1
	/	LazyLinear (128)	×1
Classifier	LazyLinear (K)		×1

where $\mathcal{T}_{y_*}$ is the prototypical-semantic features which can be indexed by labels y_*. $\mathcal{L}_c$ constrains not only the semantic features of $\mathbf{D}_l$ but also that of $\mathbf{D}_{ul}$ to be close to their corresponding prototypical-semantic features, where the prototypical-semantic features of $\mathbf{D}_l$ and $\mathbf{D}_{ul}$ can be indexed by their true labels and pseudo labels, respectively.

IV. Proposed DCR-Based SemiSEI Method

In this section, the proposed DCR-based SemiSEI method will be described by four sections that are framework of DCR-based SemiSEI, data augmentation, semantic feature consistency, predicted class consistency, and training and testing procedure.

A. Framework of DCR-Based SemiSEI

The proposed DCR can be combined with various classification backbone networks for the SEI method. In this article, the complex-valued neural network (CVNN) [60] is used as an example of classification backbone network to evaluate the effectiveness of DCR. The structure of CVNN is shown in Table II, where two kinds of CVNN are empirically designed for long signals and short signals. The framework of DCR is depicted in Fig. 1, which contains a training stage and a testing stage.

In the training stage, there is classification backbone loss function designed for labeled training samples. Specifically, the standard cross-entropy loss is utilized to evaluate the gap between true class distribution and predicted class distribution of training samples preprocessed by weak data augmentation. In addition, there are two consistency regularization loss functions: 1) predicted class consistency loss function for UTSs and 2) semantic feature consistency loss function for both labeled and UTSs. Specifically, the predicted class consistency loss function is the cross-entropy for measuring the predicted class distribution gap between UTSs preprocessed by weak and strong data augmentation. It is noted that the pseudo labeling is applied to sharpen the predicted class distribution of weakly augmented UTSs. The semantic feature consistency loss function is the semi-supervised center loss (SSCL) [17], [61] for measuring the gap between semantic feature distribution of training samples preprocessed by weak data augmentation and prototypical-semantic feature distribution. Here, the prototypical-semantic feature distribution is the central feature distribution of a certain class of training samples, and the prototypical-semantic features of these labeled training samples can be indexed by their true labels, while that of these UTSs are indexed by their pseudo labels.

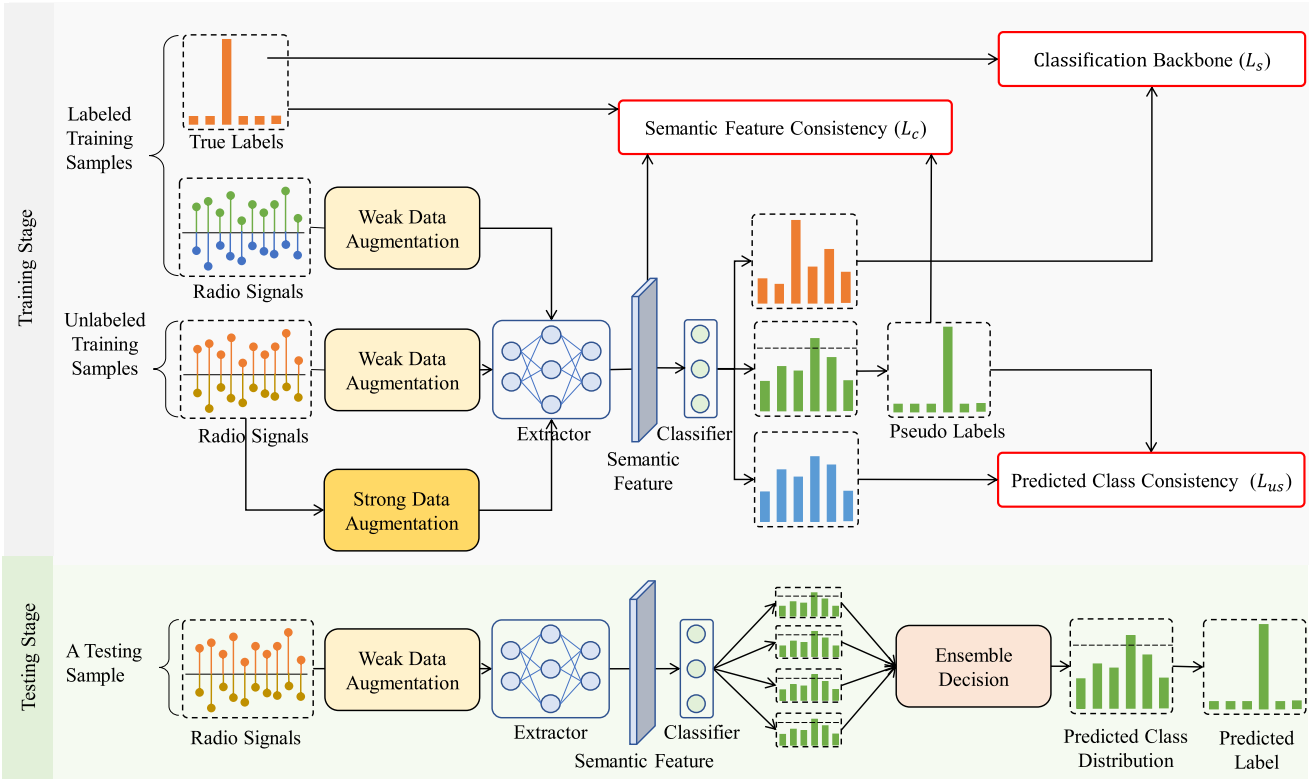

Fig. 1. Framework of the proposed DCR-based SemiSEI method.

In the testing process, radio signals are first preprocessed by weak data augmentation, and then the weakly augmented version of radio signals are fed into the CVNN to predict class distribution, and finally the predicted labels are given by ensemble decision (ED).

B. Data Augmentation

Inspired by literatures [62], [63], the rotation and the combination of rotation and cutout are used as weak data augmentation and strong data augmentation in the DCR, respectively. The weakly augmented symbols of radio signals are obtained by rotating the symbols around its origin, which can be formulated as

$$f_{wda}[\mathbf{r}(i);\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \mathcal{I} \\ \mathcal{Q} \end{bmatrix}, i = 1, 2, \ldots, I \quad (4)$$

where $\mathcal{I}$ and $\mathcal{Q}$ are the in-phase and quadrature components of the symbols of radio signals, respectively; θ is the angle of rotation and $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$; I is the number of sampling points of radio signals. To obtain the radio signals with a stronger perturbation, the rotated signals are further processed by cutout, which can be expressed as

$$f_{sda}[\mathbf{r}(i);\theta] = f_{wda}[\mathbf{r}(i);\theta] \odot \mathbf{z}(i) \quad (5)$$

$$\mathbf{z}(i) = \begin{cases} 1, & 1 \le i < s \\ 0, & s \le i \le s + l - 1 \\ 1, & s + l - 1 < i \le I \end{cases} \quad (6)$$

where $\{\mathbf{z}(i)|i = 1, 2, \ldots, I\}$ is a mask sequence consisting of one and zero, and $\odot$ indicates the elementwise product, and the number of one to the number of zero ratio η obeys the Beta$(1, 1)$ distribution. To obtain the mask sequence based on the ratio η, the cutting region on the rotated signals is designed as $l = I\sqrt{1-\eta}$, and then the starting index s of the cutting region is randomly sampled according to the uniform distribution as

$$s \sim \lceil U(0, I - l + 1) \rceil. \quad (7)$$

To simplify following description, we use $\ddot{\mathbf{r}}_* = \{f_{wda}[\mathbf{r}(i);\theta]|i = 1, 2, \ldots, I\}$ and $\tilde{\mathbf{r}}_* = \{f_{sda}[\mathbf{r}(i);\theta]|i = 1, 2, \ldots, I\}$ to represent a weakly augmented signal and a strongly augmented signal, respectively; we use $\ddot{\mathbf{R}}_* = \{\ddot{\mathbf{r}}_*^b|b = 1, 2, \ldots, B_*\}$ and $\tilde{\mathbf{R}}_* = \{\tilde{\mathbf{r}}_*^b|b = 1, 2, \ldots, B_*\}$ to denote a batch of weakly augmented signals and a batch of strongly augmented signals, respectively. The illustrations of the weak data augmentation and strong data augmentation are shown in Fig. 2.

C. Predicted Class Consistency

Each unlabel training sample (UTS) is weakly augmented and strongly augmented, and the pseudo label of $\ddot{\mathbf{r}}_{ul}$ and predicted class distribution of $\tilde{\mathbf{r}}_{ul}$ are constrained to be consistent by entropy minimization [64]. Specifically, first, the pseudo label of each UTS is obtained by

$$\mathbf{q}_{ul}^b = f_c\left(f_e\left(\ddot{\mathbf{r}}_{ul}^b\right)\right) \quad (8)$$

$$\hat{y}_{ul}^b = \arg\max\left(\mathbf{q}_{ul}^b\right) \quad (9)$$

where $\ddot{\mathbf{r}}_{ul}^b$ represents the bth sample of $\ddot{\mathbf{R}}_{ul} = \{\ddot{\mathbf{r}}_{ul}^b|b = 1, 2, \ldots, \mu B\}$, f_e is the feature extractor, f_c is the classifier,

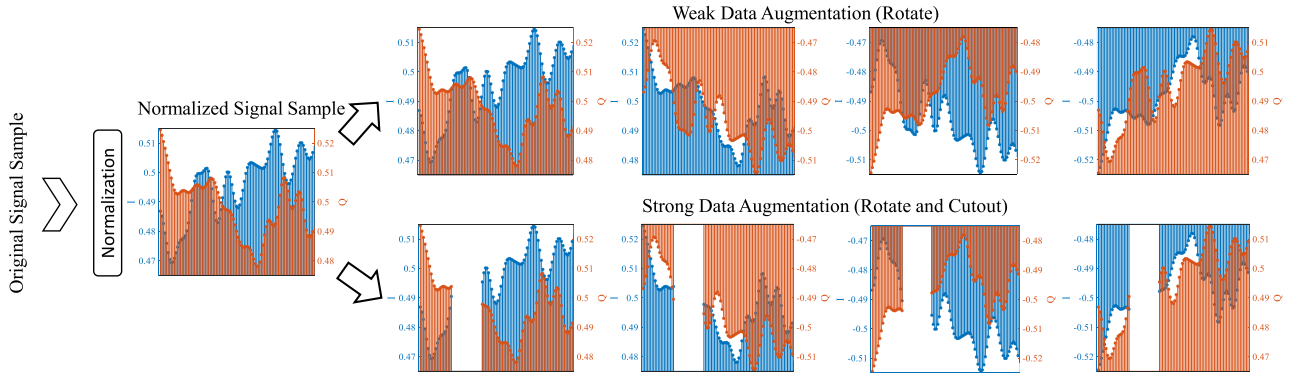

Fig. 2. Weak data augmentation and strong data augmentation in the DCR-based SemiSEI.

and $\hat{y}_{ul}^b$ is the pseudo label; second, the pseudo label is used as the ideal distribution of model's output for $\tilde{\mathbf{r}}_{ul}^b$, i.e.,

$$\mathcal{L}_u = \frac{1}{\mu B}\sum_{b=1}^{\mu B} 1\left(\max\left(\mathbf{q}_{ul}^b\right) > \tau\right) H\left(\hat{y}_{ul}^b, f_c\left(f_e\left(\tilde{\mathbf{r}}_{ul}^b\right)\right)\right) \quad (10)$$

where τ represents a threshold above which a pseudo label is retained, and H stands for the cross-entropy loss.

D. Semantic Feature Consistency

The prototypical-semantic feature distribution is first learned from the $\ddot{\mathbf{R}}_l = \{\ddot{\mathbf{r}}_l^b | b = 1, 2, \ldots, B\}$, and meanwhile the semantic feature distribution of $\ddot{\mathbf{R}}_l$ is also constrained to be closed to the prototypical-semantic feature distribution. When the prototypical-semantic feature reaches certain representativeness and stability, $\ddot{\mathbf{R}}_{ul}$ participates in the tuning of prototypical-semantic feature distribution, and the semantic feature distribution of $\ddot{\mathbf{R}}_{ul}$ is encouraged to be consistent with the prototypical-semantic feature distribution. To achieve this, we introduce the SSCL [17], [61], i.e.,

$$\mathcal{L}_c = \frac{1}{2B}\sum_{b=1}^{B}\left\|f_e\left(\ddot{\mathbf{r}}_l^b\right) - \mathbf{c}_{y_l^b}\right\|_2^2 + \alpha(t)\frac{1}{2\mu B}\sum_{b=1}^{\mu B} 1\left(\max\left(\mathbf{q}_{ul}^b\right) > \tau\right)\left\|f_e\left(\ddot{\mathbf{r}}_{ul}^b\right) - \mathbf{c}_{\hat{y}_{ul}^b}\right\|_2^2 \quad (11)$$

$$\alpha(t) = \begin{cases} 0, & t < 100 \\ 1, & 100 \le t < T \end{cases} \quad (12)$$

where $\mathbf{c}_{y_l^b}$ and $\mathbf{c}_{\hat{y}_{ul}^b}$ are the prototypical-semantic features of category y_l^b and category $\hat{y}_{ul}^b$, respectively. For each category of emitter devices, the prototypical-semantic features are learned. $\alpha(t)$ is a coefficient balancing the terms, which is expected to help the optimization process when increasing $\alpha(t)$. Different from MAT-CL [17], the SSCL encourages the model to learn an invariant representation under weak data augmentation and further build the connection between labeled and UTSs in this article. Specifically, the data augmentation of DCR produces stronger sample diversity than MAT-CL, and DCR employs semantic feature constraints on both labeled and UTSs preprocessed by weak data augmentation, while MAT-CL only applies semantic feature constraints into labeled and UTSs without any data augmentation.

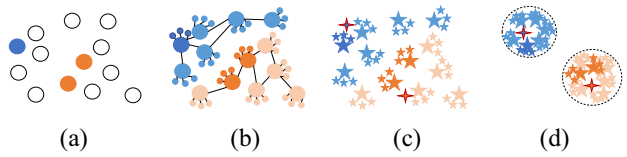

Fig. 3. Semantic feature consistency, where (a) and (b) is the training samples and weakly augmented training samples in sample space; (c) and (d) is the corresponding semantic features in the feature space. Blue and orange nodes are labeled training samples, white nodes are UTSs whose labels have not been determined, and light blue nodes and light orange nodes are unlabeled nodes whose labels are correctly determined, litter nodes are weakly augmented training samples, the stars are the semantic features of the corresponding samples, and cross symbols are the prototypical-semantic features of the corresponding categories.

As shown in Fig. 3, first, the label information is propagated to the directly connected neighbors of labeled training samples by performing weak data augmentation on labeled training samples, and the label information traverses to the connected neighbors of the UTSs by performing weak and strong data augmentation on UTSs as shown in Fig. 3(b); second, the semantic features of labeled and pseudo-labeled training samples which belong to the same category are constrained to approach the prototypical-semantic features of the corresponding category as shown in Fig. 3(c) and (d). The semantic space is more accurately divided into K subspaces, where K represents the number of categories of emitters, and consequently the identification performance is improved.

E. Training and Testing Procedure

Based on the above design, the objective loss function for optimizing the CVNN in the training stage can be expressed as

$$\mathcal{L} = \mathcal{L}_s + \omega_1\mathcal{L}_{us} + \omega_2\mathcal{L}_c. \quad (13)$$

We describe the training procedure of DCR in Algorithm 1. In addition, in the testing stage, the predicted label of a testing

Algorithm 1: Training Procedure of the DCR-Based SemiSEI Method

Require:
- T: Number of training iterations;
- C: Number of batches;
- $\mathbf{W}_m$, $\mathbf{W}_c$: Parameters of CVNN and prototypical-semantic features of semi-supervised center loss, respectively;
- lr_m, lr_c: Learning rate of CVNN and semi-supervised center loss, respectively;
- $\mathbf{z}_l^{a,b}$, $\mathbf{z}_{ul}^{a,b}$, $\mathbf{z}_{ul}^{A,b}$: Semantic feature distribution of a weakly-augmented labeled training sample, a weakly-augmented unlabeled training sample and a strongly-augmented unlabeled training sample, respectively;
- $\mathbf{q}_l^{a,b}$, $\mathbf{q}_{ul}^{a,b}$, $\mathbf{q}_{ul}^{A,b}$: Predicted class distribution of a weakly-augmented labeled training sample, a weakly-augmented unlabeled training sample and a strongly-augmented unlabeled training sample, respectively;

for $t = 1$ *to* T **do**
  for $c = 1$ *to* C **do**
    Randomly Sampling a batch of labeled training samples $\mathbf{R}_l = \{(\mathbf{r}_l^b, \mathrm{y}_l^b)|b = 1, 2, \ldots, B\}$.
    Randomly Sampling a batch of unlabeled training samples $\mathbf{R}_{ul} = \{(\mathbf{r}_{ul}^b)|b = 1, 2, \ldots, \mu B\}$.
    Forward propagation:
    Extracting the semantic features:
    $\mathbf{Z}_l^a, \mathbf{Z}_{ul}^a, \mathbf{Z}_{ul}^A = f_e(\mathbf{W}_m, \mathbf{W}_c; \ddot{\mathbf{R}}_l, \ddot{\mathbf{R}}_{ul}, \ddot{\mathbf{R}}_{ul})$;
    Predicting the class distribution:
    $\mathbf{Q}_l^a, \mathbf{Q}_{ul}^a, \mathbf{Q}_{ul}^A = f_c(\mathbf{W}_m; \mathbf{Z}_l^a, \mathbf{Z}_{ul}^a, \mathbf{Z}_{ul}^A)$;
    Obtaining the artificial labels:
    for $b = 1$ *to* μB **do**
      $\hat{\mathrm{y}}_{ul}^b = \arg\max(\mathbf{q}_{ul}^{a,b})$
    end
    Calculating the sub-objective loss:
    $\mathcal{L}_s = \frac{1}{B}\sum_{b=1}^{B} H\left(\mathrm{y}_l^b, \mathbf{q}_l^{a,b}\right)$
    $\mathcal{L}_{us} = \frac{1}{\mu B}\sum_{b=1}^{\mu B} 1\left(\max\left(\mathbf{q}_{ul}^{a,b}\right) > \tau\right) H\left(\hat{\mathrm{y}}_{ul}^b, \mathbf{q}_{ul}^{A,b}\right)$.
    $\mathcal{L}_c = \frac{1}{2B}\sum_{b=1}^{B}\left\|\mathbf{z}_l^{a,b} - \mathbf{c}_{\mathrm{y}_l^b}\right\|_2^2 +$
    $\alpha(t)\frac{1}{2\mu B}\sum_{b=1}^{\mu B} 1\left(\max\left(\mathbf{q}_{ul}^{a,b}\right) > \tau\right)\left\|\mathbf{z}_{ul}^{a,b} - \mathbf{c}_{\hat{\mathrm{y}}_{ul}^b}\right\|_2^2$
    Calculating the objective loss:
    $\mathcal{L} = \mathcal{L}_s + \omega_1\mathcal{L}_{us} + \omega_2\mathcal{L}_c$
    Backward propagation:
    Updating the parameters of CVNN:
    $\mathbf{W}_m \leftarrow \mathrm{Adam}(\nabla_{\mathbf{W}_m}, \mathcal{L}, lr_m, \mathbf{W}_m)$;
    Updating the prototypical-semantic features:
    $\mathbf{W}_c \leftarrow \mathrm{Adam}(\nabla_{\mathbf{W}_c}, \mathcal{L}, lr_c, \mathbf{W}_c)$
  end
end

sample $\mathbf{r}_t$ is obtained by

$$\arg\max_{\theta} \sum f_c\{f_e[f_{wda}(\mathbf{r}_t; \theta)]\}. \tag{14}$$

TABLE III
DETAILS OF DATA SET AND SIMULATED PARAMETERS

	Wi-Fi [66]	ADS-B [65]
Number of Categories	16	10
Length of Each Sample	6,000	4,800
Number of Training Sample	3,080	3,080
Number of Testing Sample	16,004	1,000
Ratio	5%, 10%, 20%, 50%, 100%	
Framework	PyTorch [67] (v1.10.2 with Python 3.6.13)	
Optimizer	Adam [68]	
Learning Rate lr_m and lr_c	0.001	
Weighting Scalar ω_1	1	
Weighting Scalar ω_2	0.001	
Number of iterations	600	
Batch Size	32	
Flatform	NVIDIA GeForce RTX 3090 GPU	

V. EXPERIMENTAL SETUP AND RESULTS

A. Simulation Parameters

The details of simulated parameters are shown in Table III, where the ratio indicates how many of training samples are labeled. In addition, the validating samples are randomly selected from 30% of the labeled training samples.

B. Comparative Methods

To demonstrate the effectiveness of our DCR-based SemiSEI method, we choose some representative signal recognition methods corresponding to the above categories of semi-supervised learning as the comparative methods, namely, DRCN [36], Triple-GAN [34], and SimMIM [69], [70] (deep generative methods), and SSRCNN [59], MAT-CL, and MAT-PA [17] (hybrid methods composed of pseudo labeling and consistency regularization). We also compare the proposed DCR-based SemiSEI method with the classification backbone CVNN [60] which is merely trained using labeled training samples.

C. Identification Performance

Table IV shows the identification accuracy of the proposed DCR-based SemiSEI method and comparative methods. It is worth noting that some values of identification accuracy may be the same, but the identification results are actually different. We observe that as the number of labeled training samples increases, the identification performance of all methods improves because the more labeled training samples, the more information about samples distribution the model has and the better decision boundaries are. As the number of labeled training samples decreases, the superiority of the proposed method in identification performance can be observed. Specifically, when the ratio is 10%, the identification accuracy of the proposed method can achieve 99.77% for the WiFi data set and 90.10% for the ADS-B data set. Our method has 71.13% and 15.60% higher classification accuracy than the classification backbone CVNN on the WiFi data set and the ADS-B data set, respectively. In addition, our method has 19.07% and 5.30% higher classification accuracy than MAT-CL, which has the highest classification accuracy in these comparative methods on WiFi data set and ADS-B data set, respectively.

TABLE IV
IDENTIFICATION ACCURACY OF THE PROPOSED DCR-BASED SEMISEI METHOD AND COMPARATIVE METHODS UNDER THE WIFI DATA SET AND THE ADS-B DATA SET

Methods	Wi-Fi					ADS-B				
	5%	**10%**	**20%**	**50%**	**100%**	**5%**	**10%**	**20%**	**50%**	**100%**
CVNN [60]	20.47%	28.64%	69.78%	97.14%	99.46%	60.50%	74.50%	92.70%	97.70%	99.20%
DRCN [36]	21.94%	47.51%	76.18%	98.99%	99.64%	54.20%	72.40%	93.60%	97.10%	98.90%
SSRCNN [59]	19.33%	38.09%	99.25%	99.75%	99.76%	49.30%	79.30%	91.00%	97.50%	99.10%
Triple-GAN [34]	27.57%	37.27%	72.88%	99.06%	99.63%	45.10%	61.10%	90.90%	97.10%	99.10%
SimMIM [69], [70]	31.71%	49.59%	75.76%	96.01%	99.41%	65.90%	77.90%	92.90%	97.60%	98.80%
MAT-CL [17]	27.26%	80.70%	99.76%	99.79%	99.79%	70.06%	83.80%	95.00%	**99.10%**	**99.40%**
MAT-PA [17]	28.82%	54.96%	98.18%	99.77%	99.77%	74.00%	84.80%	93.90%	97.30%	99.30%
DCR (Proposed)	**98.99%**	**99.77%**	**99.77%**	**99.86%**	**99.86%**	**79.70%**	**90.10%**	**96.70%**	98.80%	99.00%

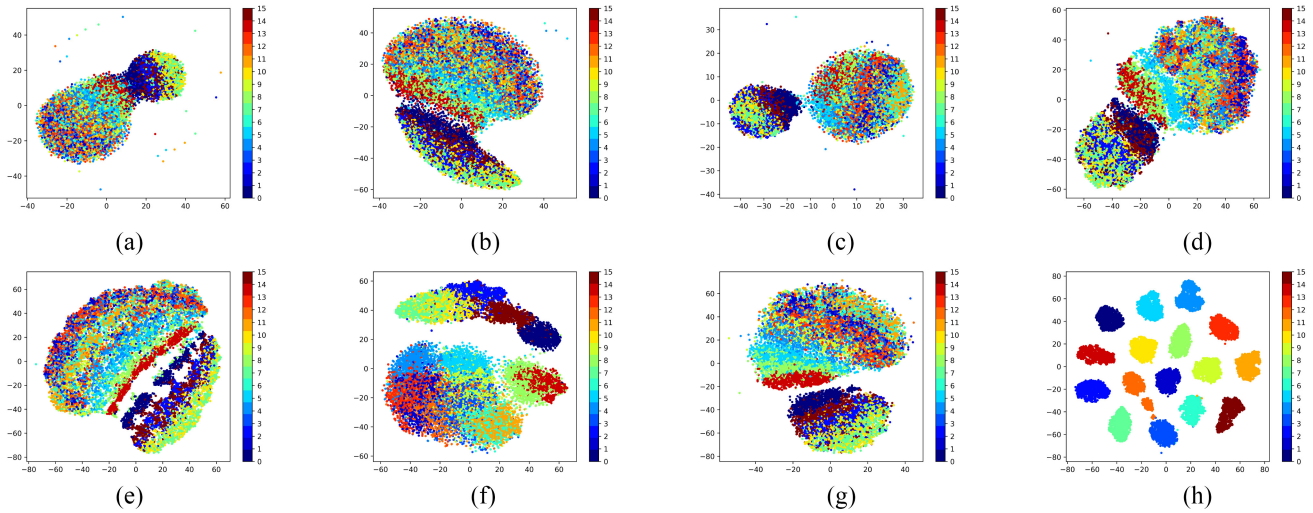

Fig. 4. Visualization of semantic feature of different SEI methods when the data set is WiFi and the number of labeled training samples to the number of all training samples ratio is 5%, where the silhouette coefficient of CVNN, DRCN, SSRCNN, TripleGAN, SimMIM, MAT-CL, MAT-PA, and DCR is −0.0039, −0.0130, −0.0050, −0.0005, −0.0178, 0.0279, 0.0030, and 0.5969, respectively. (a) CVNN [60]. (b) DRCN [36]. (c) SSRCNN [59]. (d) Triple-GAN [34]. (e) SimMIM [69], [70]. (f) MAT-CL [17]. (g) MAT-PA [17]. (h) DCR (proposed).

However, we also observe that the identification accuracy of DCR is not always better than that of comparative methods under the ADS-B data set. The DCR has 1.10% lower identification accuracy than MAT-CL when the ratio is 50%, and 0.20% lower identification accuracy than the classification backbone CVNN when the ratio is 100%. The explanation is that when the number of labeled samples is sufficient for supervised learning, unlabeled samples may bring additional perturbations to degrade the model identification performance, especially as the perturbations increase with the data augmentation diversity of the unlabeled samples. Therefore, semi-supervised learning and data augmentation is not recommended when labeled training samples are sufficient.

D. Visualization of Semantic Feature

T-distributed stochastic neighbor embedding (t-SNE) [71] visualizes high-dimensional data by giving each data point a location in a two or 3-D map. It is able to capture the local structure of high-dimensional data while also revealing the global structure, such as the presence of clusters at multiple scales. Because of these advantages, t-SNE has been widely used in visualization of feature extracted by neural networks.

In this article, the dimension of extracted semantic features of long signals (WiFi) and short signals (ADS-B) is 1024 and 128, respectively. We use t-SNE to compress the dimensional of extracted semantic feature from 1024 or 128 to 2 and visualize it. As shown in Figs. 4 and 5, the proposed method enables a lower density of data points at the decision boundary and makes the classification problem easier to solve for testing samples. When the ratio is 5%, the silhouette coefficient of DCR is 0.5969 and 0.2344 on the WiFi data set and the ADS-B data set, respectively. DCR has 0.5690 and 0.0210 higher silhouette coefficient than MAT-CL, which has the highest silhouette coefficient in comparative methods. The superiority of feature discrimination of DCR can be observed clearly, especially in the WiFi data set. According to the cluster assumption, DCR can get a better identification performance.

E. Ablation Experiment

There are two important factors, SSCL and ED, in the proposed DCR-based SemiSEI method. We analyze the identification performance when any factor of DCR is ablated, to demonstrate the necessity of each factor for the success of DCR. Moreover, it is worth noting that under appropriate assumptions or conditions the UTS can improve the learning performance, but some empirical studies [72], [73] have demonstrated that UTS may result in even worse performance than models learned only with labeled training samples.

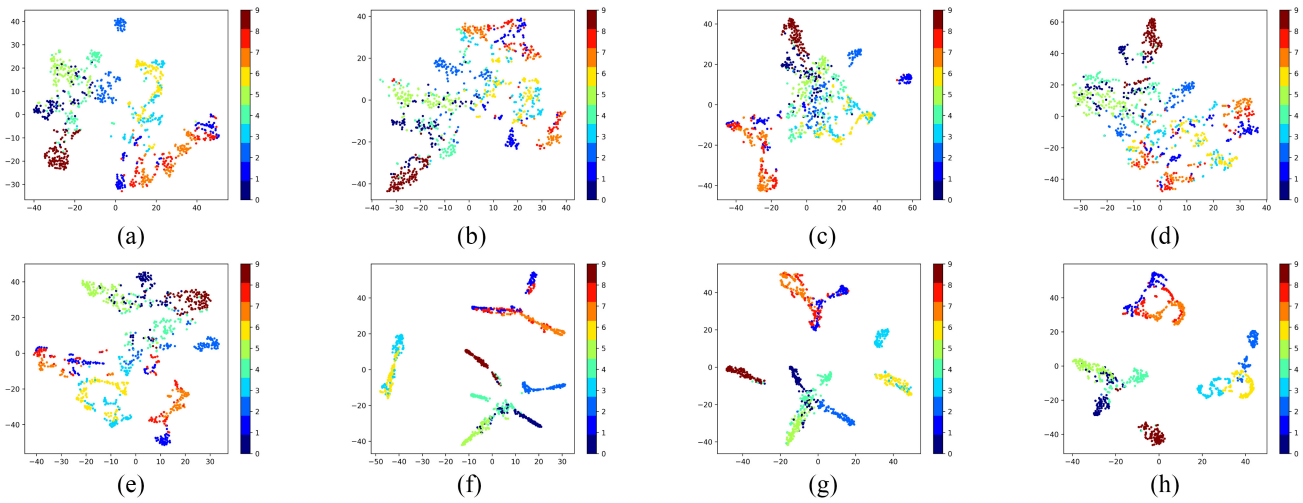

Fig. 5. Visualization of semantic feature of different SEI methods when the data set is ADS-B and the number of labeled training samples to the number of all training samples ratio is 5%, where the silhouette coefficient of CVNN, DRCN, SSRCNN, TripleGAN, SimMIM, MAT-CL, MAT-PA, and DCR is 0.1000, 0.0427, −0.0042, 0.0105, 0.1072, 0.2134, 0.1853, and 0.2344, respectively. (a) CVNN [60]. (b) DRCN [36]. (c) SSRCNN [59]. (d) Triple-GAN [34]. (e) SimMIM [69], [70]. (f) MAT-CL [17]. (g) MAT-PA [17]. (h) DCR (proposed).

TABLE V
ABLATION ANALYSIS OF THE PROPOSED DCR-BASED SEMISEI METHOD UNDER THE WIFI DATA SET AND THE ADS-B DATA SET

Methods	SSCL	UTS	ED	Wi-Fi			ADS-B		
				5%	10%	20%	5%	10%	20%
DCR w/o SSCL		✓	✓	98.01%	99.76%	99.77%	75.80%	89.70%	95.50%
DCR w/o UTS	✓		✓	55.22%	85.00%	97.89%	58.90%	79.30%	93.90%
DCR w/o ED	✓	✓		98.61%	99.77%	99.77%	78.20%	87.00%	94.60%
DCR (Proposed)	✓	✓	✓	**98.99%**	**99.77%**	**99.77%**	**79.70%**	**90.10%**	**96.70%**

Tips: w/o is an abbreviation for without.

Hence, we also demonstrate that the UTS also has a positive effect on the success of DCR.

Table V shows the identification performance when any factor of DCR is ablated. We observe that all factors are important to DCR. When any factor is ablated, the identification accuracy of the DCR decreases. Especially when the UTS is ablated, the identification accuracy of the DCR significantly decreases. This indicates that our proposed DCR can learn the valuable knowledge of UTS that is beneficial to model training.

To further demonstrate that the proposed DCR efficiently utilizes the UTSs data, we analyze the identification accuracy of CVNN whose training process are driven or enhanced by rotation, cutout, center loss, and their combination in a supervised way, and compare the proposed DCR with it as shown in Table VI. In Table VI, the values * in parentheses indicate that the identification accuracy of model with sum of cross-entropy loss and center loss as objective function is *% higher or lower than that of model with cross-entropy loss as objective function. It can be observed that clear superiority of DCR on identification accuracy, although data augmentation brings gains of identification accuracy compared with none, where none means that there is no data augmentation.

In addition, from Table VI, it can be observed that rotation and cutout combined with center loss-based SEI method has a significant decrease in identification accuracy compared to rotation and cutout-based SEI method. Therefore, it is inappropriate to define prototypical-semantic features and encourage the semantic features of strongly augmented unlabeled samples originated from UTS to be close to them. This is the one of the reasons that DCR why not synchronously constrain the semantic features of UTSs which are preprocessed by strong data augmentation to be close to their corresponding prototypical-semantic features (referred to as DCR-S). Another reason is that the prototypical-semantic features are initialized by weakly augmented labeled training samples, and it is intuitively natural to encourage the semantic features of labeled training samples and UTSs which are preprocessed by weak augmentation to be close to their corresponding prototypical-semantic features. Compared with weak augmentation, strong augmentation imposes more complex perturbations to the UTSs. If strongly augmented UTSs are introduced into optimization of prototypical-semantic features, they may bring negative noise to prototypical-semantic features, resulting in oscillations of prototypical-semantic features and even destroying the features extractor. To support this idea, we analyze identification performance of DCR and that of DCR-S as shown in Fig. 6. It can be observed that DCR-S does not bring significant performance gain, but in some semi-supervised scenarios DCR-S actually bring about identification accuracy decline.

F. Analysis of Data Augmentation

Similar to [74], we define a graph where each node corresponds to a sample $\mathbf{r} \in \mathcal{R}$ and an edge $(\hat{\mathbf{r}}, \mathbf{r})$ exists in the graph if and only if the data augmentation operation can be

TABLE VI
IDENTIFICATION ACCURACY OF THE PROPOSED DCR-BASED SEMISEI METHOD AND DATA AUGMENTATION-BASED SEI METHODS UNDER THE WIFI DATA SET AND THE ADS-B DATA SET

Methods	Wi-Fi			ADS-B		
	5%	**10%**	**20%**	**5%**	**10%**	**20%**
None	20.47% (↑0.34%)	28.64% (↑5.68%)	69.78% (↑14.49%)	60.50% (↓0.10%)	74.50% (↑6.20%)	92.70% (↑1.70%)
Rotation	41.09% (↑1.43%)	74.78% (↑7.37%)	97.44% (↓1.93%)	64.90% (↓9.00%)	74.60% (↓0.30%)	85.60% (↑2.20%)
Cutout	24.91% (↑18.77%)	73.22% (↑0.87%)	95.28% (↑1.42%)	67.70% (↓2.90%)	82.10% (↓1.80%)	88.30% (↑4.40%)
Rotation and Cutout	74.36% (↓5.04%)	98.16% (↓3.04%)	99.73% (↑0.01%)	68.10% (↓13.00%)	79.40% (↓1.90%)	92.00% (↓3.60%)
DCR (Proposed)	**98.99%**	**99.77%**	**99.77%**	**79.70%**	**90.10%**	**96.70%**

Note: The values * in parentheses indicate that the identification accuracy of model with sum of cross-entropy loss and center loss as objective function is *% higher or lower than that of model with cross-entropy loss as objective function.

TABLE VII
IDENTIFICATION ACCURACY OF THE PROPOSED DCR-BASED SEMISEI METHOD WITH DIFFERENT DATA AUGMENTATION

DA	Weak Data Augmentation	Strong Data Augmentation	Wi-Fi			ADS-B		
			5%	**10%**	**20%**	**5%**	**10%**	**20%**
V1	Rotation	Rotation and Cutout	**98.99%**	**99.77%**	**99.77%**	79.70%	**90.10%**	**96.70%**
V2	None	Rotation and Cutout	60.17%	99.77%	99.77%	**82.30%**	88.40%	92.50%
V3	Rotation and Cutout	Rotation	82.56%	91.07%	95.78%	50.90%	58.60%	74.60%
V4	None	Rotation	6.32%	64.70%	97.99%	65.90%	67.40%	93.10%

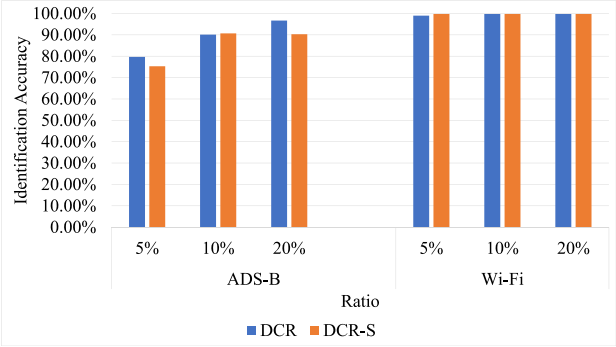

Fig. 6. Identification accuracy of DCR and that of DCR-S.

reserved. For a K-category classification problem, that graph has K components (disconnected subgraphs) when all samples with each category can be traversed by the augmentation operation. Otherwise, the graph will have more than K components. In fact, the quality of the augmentation operation always decides the number of components, and ideal augmentation should be able to reach all other samples of the same category given a starting instance. To find out a better augmentation operation, we analyze the different combination of weak data augmentation and strong data augmentation as shown in Table VII.

In the DA V2 and DA V4, none means that there is no data augmentation and the original data is used to calculate the pseudo label for the strongly augmented one. Without any perturbation on data or model, the model which learn from the pseudo labels may have the risk on confirmation bias. It can be observed that the DA V2 and DA V4 have a lower identification accuracy than DA V1. The predicted class distribution of the weakly augmented data will provide pseudo labels for strongly augmented data. Consequently, the weakly augmented data should maintain high similarity to the original data and the strongly augmented data have partial similarity to the original data. It is intuitively understood that rotation and cutout make more powerful disturbances on original data than rotation because cutout erases part of the data, and we can observe that the identification accuracy of DA V1 is higher than that of DA V3. According to this analysis, the DA V1 is the optimal solution for DCR.

VI. CONCLUSION

In this article, we presented a semi-supervised learning framework referred to as DCR for SEI. Only few training samples were required to be annotated and cooperated with a large amount of UTSs to participate in the training process. In addition to the weak data augmentation and standard cross-entropy loss function for labeled training samples (first term of the objective loss function), there are weak data augmentation, strong data augmentation, and predicted class consistency loss function for UTSs (second term of the objective loss function) and weak data augmentation and semantic feature consistency loss function for both labeled and UTSs (third term of the objective loss function).

Due to the first term, the label information is propagated to the directly connected neighbors of labeled training samples. Due to the second term, the label information is traversed to the connected neighbors of the UTSs. Due to the third term, the semantic features of labeled and pseudo-labeled training samples which belong to the same category are constrained to approach the prototypical-semantic features of the corresponding category, and the semantic space is more accurately divided into K subspaces for K categories of emitters, and consequently the emitters identification accuracy is significantly improved.

The proposed DCR-base SemiSEI method was validated on the WiFi data set with 16 categories and the ADS-B data set with ten categories. The simulation results showed that as the number of labeled training samples decreases, the superiority of the proposed method in identification performance could be observed. Specifically, when the ratio was 10%, the proposed method could achieve 99.77% identification accuracy for the WiFi data set and 90.10% identification accuracy for the ADS-B data set. Compared with the state-of-the-art methods,

the proposed method improved the identification accuracy on the WiFi data set and the ADS-B data set by more than 19.07% and 5.30%, respectively.

REFERENCES

[1] P. Angueira et al., "A survey of physical layer techniques for secure wireless communications in industry," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 810–838, 2nd Quart., 2022.

[2] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2021.

[3] J. Zhang, F. Wang, O. A. Dobre, and Z. Zhong, "Specific emitter identification via Hilbert-Huang transform in single-hop and relaying scenarios," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 1192–1205, 2016.

[4] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.

[5] R. Zhao, Y. Wang, Z. Xue, T. Ohtsuki, B. Adebisi, and G. Gui, "Semi-supervised federated learning based intrusion detection method for Internet of Things," *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8645–8657, May 2023, doi: 10.1109/JIOT.2022.3175918.

[6] S. Chen, S. Zheng, L. Yang, and X. Yang, "Deep learning for large-scale real-world ACARS and ADS-B radio signal classification," *IEEE Access*, vol. 7, pp. 89256–89264, 2019.

[7] S. Wang, H. Jiang, X. Fang, Y. Ying, J. Li, and B. Zhang, "Radio frequency fingerprint identification based on deep complex residual network," *IEEE Access*, vol. 8, pp. 204417–204424, 2020.

[8] X. Zha, H. Chen, T. Li, Z. Qiu, and Y. Feng, "Specific emitter identification based on complex Fourier neural network," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 592–596, Mar. 2022.

[9] 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.

[10] 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.

[11] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, "Specific emitter identification based on deep residual networks," *IEEE Access*, vol. 7, pp. 54425–54434, 2019.

[12] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro, "Towards scalable and channel-robust radio frequency fingerprint identification for LoRa," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 774–787, 2022.

[13] Z.-M. Liu, "Multi-feature fusion for specific emitter identification via deep ensemble learning," *Digital Signal Process.*, vol. 110, Mar. 2021, Art. no. 102939.

[14] Y. Zhao, X. Wang, Z. Lin, and Z. Huang, "Multi-classifier fusion for open-set specific emitter identification," *Remote Sens.*, vol. 14, no. 9, p. 2226, 2022.

[15] M. Camelo et al., "A semi-supervised learning approach towards automatic wireless technology recognition," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw.*, 2019, pp. 1–10.

[16] Y. Wang, G. Gui, Y. Lin, H.-C. Wu, C. Yuen, and F. Adachi, "Few-shot specific emitter identification via deep metric ensemble learning," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 24980–24994, Dec. 2022.

[17] X. Fu et al., "Semi-supervised specific emitter identification method using metric-adversarial training," *IEEE Internet Things J.*, vol. 10, no. 12, pp. 10778–10789, Jun. 2023, doi: 10.1109/JIOT.2023.3240242.

[18] H. Feng, X. Yan, K. Jiang, X. Zhao, and B. Tang, "Contrastive psudo-supervised classification for intra-pulse modulation of radar emitter signals using data augmentation," 2022, *arXiv:2210.06973*.

[19] C. Wang et al., "Few-shot specific emitter identification via hybrid data augmentation and deep metric learning," 2022, *arXiv:2212.00252*.

[20] K. Chen, L. Yao, D. Zhang, X. Wang, X. Chang, and F. Nie, "A semisupervised recurrent convolutional attention model for human activity recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1747–1756, May 2020.

[21] M. Luo, X. Chang, L. Nie, Y. Yang, A. G. Hauptmann, and Q. Zheng, "An adaptive semisupervised feature analysis for video semantic recognition," *IEEE Trans. Cybern.*, vol. 48, no. 2, pp. 648–660, Feb. 2018.

[22] E. Yu, J. Sun, J. Li, X. Chang, X.-H. Han, and A. G. Hauptmann, "Adaptive semi-supervised feature selection for cross-modal retrieval," *IEEE Trans. Multimedia*, vol. 21, no. 5, pp. 1276–1288, May 2019.

[23] X. Yang, Z. Song, I. King, and Z. Xu, "A survey on deep semi-supervised learning," *IEEE Trans. Knowl. Data Eng.*, early access, Nov. 8, 2022, doi: 10.1109/TKDE.2022.3220219.

[24] I. J. Goodfellow et al., "Generative adversarial nets," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Nov. 2020.

[25] J. Masci, U. Meier, D. Ciresan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proc. Int. Conf. Artif. Neural Netw.*, 2011, pp. 52–59.

[26] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, pp. 2642–2651.

[27] A. Radford, L. Metz, and S.Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2016, *arXiv:1511.06434*.

[28] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[29] Z. Gan, L. Chen, and W. Wang, "Triangle generative adversarial networks," in *Proc. 31st Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 5247–5256.

[30] Y. Tu, Y. Lin, J. Wang, and J. Kim, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *Comput., Mater. Continua*, vol. 55, no. 2, pp. 243–254, 2018.

[31] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," in *Proc. 30th Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2016, pp. 2234–2242.

[32] H. Zhou, L. Jiao, S. Zheng, L. Yang, W. Shen, and X. Yang, "Generative adversarial network-based electromagnetic signal classification: A semi-supervised learning framework," *China Commun.*, vol. 17, no. 10, pp. 157–169, Oct. 2020.

[33] K. Tan, W. Yan, L. Zhang, Q. Ling, and C. Xu, "Semi-supervised specific emitter identification based on bispectrum feature extraction CGAN in multiple communication scenarios," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 1, pp. 292–310, Feb. 2023.

[34] J. Gong, X. Xu, Y. Qin, and W. Dong, "A generative adversarial network based framework for specific emitter characterization and identification," in *Proc. 11th Int. Conf. Wireless Commun. Signal Process.*, 2019, pp. 1–6.

[35] X. Xu, T. Jiang, J. Gong, H. Xu, and X. Qin, "WLAN interference signal recognition using an improved quadruple generative adversarial network," *Digit. Signal Process.*, vol. 117, Oct. 2021, Art. no. 103188.

[36] Y. Wang, G. Gui, H. Gacanin, T. Ohtsuki, H. Sari, and F. Adachi, "Transfer learning for semi-supervised automatic modulation classification in ZF-MIMO systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 2, pp. 231–239, Jun. 2020.

[37] P. Bachman, O. Alsharif, and D. Precup, "Learning with pseudo-ensembles," in *Proc. 27th Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2014, pp. 3365–3373.

[38] J. Huang, M. L. Huang, P. H. Tan, Z. Chen, and S. Sun, "Semi-supervised deep learning based wireless interference identification for IIoT networks," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, 2020, pp. 1–5.

[39] S. Laine and T. Aila, "Temporal ensembling for semi-supervised learning," 2017, *arXiv:1610.02242*.

[40] D. Liu, P. Wang, T. Wang, and T. Abdelzaher, "Self-contrastive learning based semi-supervised radio modulation classification," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, 2021, pp. 777–782.

[41] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 1597–1607.

[42] W. Deng, X. Wang, Z. Huang, and Q. Xu, "Modulation classifier: A few-shot learning semi-supervised method based on multimodal information and domain adversarial network," *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 576–580, Feb. 2023.

[43] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1–35, 2016.

[44] C. Xie, L. Zhang, and Z. Zhong, "Virtual adversarial training-based semisupervised specific emitter identification," *Wireless Commun. Mobile Comput.*, vol. 2022, Jan. 2022, Art. no. 6309958.

[45] T. Miyato, S. Maeda, M. Koyama, and S. Ishii, "Virtual adversarial training: A regularization method for supervised and semi-supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1979–1993, Aug. 2019.

[46] X. Fu, Y. Wang, Y. Lin, G. Gui, H. Gacanin, and F. Adachi, "A novel semi-supervised learning framework for specific emitter identification," in *Proc. IEEE 96th Veh. Technol. Conf.*, 2022, pp. 1–5.
[47] D. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Proc. Int. Conf. Mach. Learn. Workshop Challenges Represent. Learn. (ICML)*, 2013, pp. 1–6.
[48] K. Longi, T. Pulkkinen, and A. Klami, "Semi-supervised convolutional neural networks for identifying Wi-Fi interference sources," in *Proc. 9th Asian Conf. Mach. Learn.*, 2017, pp. 391–406.
[49] H. Yang, H. Zhang, H. Wang, and Z. Guo, "A novel approach for unlabeled samples in radiation source identification," *J. Syst. Eng. Electron.*, vol. 33, no. 2, pp. 354–359, Apr. 2022.
[50] S. Yuan, P. Li, B. Wu, X. Li, and J. Wang, "Semi-supervised classification for intra-pulse modulation of radar emitter signals using convolutional neural network," *Remote Sens.*, vol. 14, no. 9, p. 2059, Apr. 2022.
[51] Z. Xu, X. Wu, D. Gao, and W. Su, "Blind modulation recognition of UWA signals with semi-supervised learning," in *Proc. IEEE Int. Conf. Signal Process. Commun. Comput. (ICSPCC)*, 2022, pp. 1–5.
[52] V. Verma et al., "Interpolation consistency training for semi-supervised learning," *Neural Netw.*, vol. 145, pp. 90–106, Jan. 2022.
[53] Z. Ren, P. Ren, and T. Zhang, "Deep RF device fingerprinting by semi-supervised learning with meta pseudo time-frequency labels," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2022, pp. 2369–2374.
[54] H. Pham, Z. Dai, Q. Xie, and Q. V. Le, "Meta pseudo labels," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021, pp. 11557–11568.
[55] E. Arazo, D. Ortego, P. Albert, N. E. Connor, and K. McGuinness, "Pseudo-labeling and confirmation bias in deep semi-supervised learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2020, pp. 1–8.
[56] M. Li, G. Liu, S. Li, and Y. Wu, "Radio classify generative adversarial networks: A semi-supervised method for modulation recognition," in *Proc. IEEE Int. Conf. Commun. Technol. (ICCT)*, 2018, pp. 669–672.
[57] M. Li, O. Li, G. Liu, and C. Zhang, "Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks," *Sensors*, vol. 18, no. 11, p. 3913, Nov. 2018.
[58] Y. Dong, X. Jiang, L. Cheng, and Q. Shi, "Semi-supervised learning for signal recognition with sparsity and robust promotion," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2021, pp. 1–6.
[59] Y. Dong, X. Jiang, L. Cheng, and Q. Shi, "SSRCNN: A semi-supervised learning framework for signal recognition," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 3, pp. 780–789, Sep. 2021.
[60] Y. Wang, G. Gui, H. Gacanin, T. Ohtsuki, O. A. Dobre, and H. V. Poor, "An efficient 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.
[61] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 499–515.
[62] L. Huang, W. Pan, Y. Zhang, L. Qian, N. Gao, and Y. Wu, "Data augmentation for deep learning-based radio modulation classification," *IEEE Access*, vol. 8, pp. 1498–1506, 2020.
[63] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," 2017, *arXiv:1708.04552*.
[64] Y. Grandvalet and Y. Bengio, "Semi-supervised learning by entropy minimization," in *Proc. 17th Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2004, pp. 529–536.
[65] Y. Tu et al., "Large-scale real-world radio signal recognition with deep learning," *Chin. J. Aeronaut.*, vol. 35, no. 9, pp. 35–48, Sep. 2022.
[66] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: Optimized radio classification through convolutional neural networks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2019, pp. 370–378.
[67] A. Paszke et al., "Automatic differentiation in PyTorch," in *Proc. Conf. Adv. Neural Inf. Process. Syst. Workshop Autodiff Submission*, 2017, pp. 1–4.
[68] D. P. Kinga and J. B. Adam, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[69] Z. Xie et al., "SimMIM: A simple framework for masked image modeling," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 9643–9653.
[70] K. Huang, J. Yang, H. Liu, and P. Hu, "Deep learning of radio frequency fingerprints from limited samples by masked autoencoding," *IEEE Wireless Commun. Lett.*, early access, Jun. 21, 2022, doi: 10.1109/LWC.2022.3184674.
[71] L. Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.
[72] A. Singh, R. D. Nowak, and X. Zhu, "Unlabeled data: Now it helps, now it doesn't," in *Proc. 21st Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2008, pp. 1513–1520.
[73] T. Yang and C. E. Priebe, "The effect of model misspecification on semi-supervised classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 2093–2103, Oct. 2011.
[74] Q. Xie, Z. Dai, E. Hovy, T. Luong, and Q. Le, "Unsupervised data augmentation for consistency training," in *Proc. 33rd Conf. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2020, pp. 6256–6268.