A Comprehensive Survey on Self-Supervised Learning for Specific Emitter Identification

Chao Liu, *Graduate Student Member, IEEE*, Guan Gui, *Fellow, IEEE*, Yu Wang, *Member, IEEE*, Tomoaki Ohtsuki, *Senior Member, IEEE*, Dusit Niyato, *Fellow, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—The rapid proliferation of the Internet of Things (IoT) has intensified the need for strong authentication mechanisms to ensure the integrity and reliability of connected devices. Recent advancements in Deep Learning (DL)-based Specific Emitter Identification (SEI) have demonstrated significant potential in leveraging unique Radio Frequency Fingerprints (RFF) for accurate device identification and authentication. However, the efficacy of these DL-based SEI methods is critically dependent on the availability of extensive labeled datasets, which are often scarce and expensive to obtain in practical applications. To address this limitation, Self-Supervised Learning (SSL) becomes a promising solution, capable of harnessing unlabeled data to learn effective representations. Furthermore, current surveys and reviews on SEI are generally summarized from a high-level perspective, lacking a detailed discussion of SEI methods under label-limited scenarios. This article comprehensively surveys SSL-based SEI, including its motivation, definition, paradigms, related work, challenges, and future direction combined with large models. To help readers quickly engage with this field, this paper also undertakes two specific efforts: collecting and organizing currently available open-source datasets with download links and comparing various SSL-based SEI methods with related codes.**

***Index Terms*—Specific emitter identification, radio frequency fingerprinting, deep learning, self-supervised learning.**

I. Introduction

A. Background

THE INTERNET of Things (IoT) has revolutionized the way devices interact with each other and with the environment, creating a vast interconnected network that spans various industries, including healthcare to smart cities, industrial automation, and home automation [1], [2], [3]. This rapid expansion of IoT has led to an estimated 40 billion connected devices by 2033 [4]. Such an extensive network poses significant challenges, as each device represents a potential entry point for malicious attacks. In addition, due to the openness of wireless links in IoT and the varying resource capabilities across the IoT devices, they are highly vulnerable to intrusion and malicious attacks. Many IoT devices are deployed in highly sensitive environments such as military bases, manufacturing facilities, and smart grids, offering malicious actors potential platforms to launch attacks that can result in significant damage [5].

Most existing authentication schemes in wireless communication systems rely on upper-layer authentication mechanisms, such as conventional cryptography-based algorithms. However, their limitations have become increasingly apparent [6]. With the rapid increase in computational power, the security guarantees of cryptography-based algorithms are being increasingly undermined. These algorithms are also vulnerable to replay and Denial-of-Service (DoS) attacks such as jamming [7]. Cryptography-based algorithms often introduce significant communication overhead and system complexity due to upper-layer operations such as encryption and decryption. These challenges make them less suitable for large-scale, delay-sensitive, and power-constrained IoT deployments.

Wireless security protocols in IoT often contain exploitable weaknesses, making them highly susceptible to forgery and insider attacks when malicious users acquire security credentials from legitimate users and integrate them into the networks. Specific Emitter Identification (SEI), a passive Physical Layer Authentication (PLA) scheme leveraging the Radio Frequency Fingerprint (RFF) characteristics embedded in each device, can effectively address the issue and has shown great potential in both the military and civilian scenarios [8], [9]. RFFs provide robust security due to device-specific hardware variations introduced during manufacturing, making them difficult to clone and thus highly reliable for device identification. As illustrated in Fig. 1, SEI plays a crucial role in various IoT domains, including smart transportation, smart homes, and smart industry. Each of these domains can leverage SEI for tasks including physical-layer-based device authentication and anomaly detection. While enhancing the overall security of IoT networks by preventing unauthorized access, SEI is also

Received 12 January 2025; revised 31 May 2025; accepted 6 July 2025. Date of publication 11 July 2025; date of current version 2 January 2026. This work was supported in part by the Natural Science Foundation of China under Grant 62471247, Grant 62401281, and Grant 62472019; in part by the Key Project of the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province under Grant 22KJA510002; and in part by Japan Science and Technology Agency (JST) Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE) program under Grant JPMJAP2326. *(Corresponding authors: Guan Gui; Yu Wang.)*

Chao Liu, Guan Gui, and Yu Wang are with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: 1221014001@njupt.edu.cn; guiguan@njupt.edu.cn; yuwang@njupt.edu.cn).

Tomoaki Ohtsuki is with the Department of Information and Computer Science, Keio University, Yokohama 223-8522, Kanagawa, Japan (e-mail: ohtsuki@keio.jp).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/COMST.2025.3588171

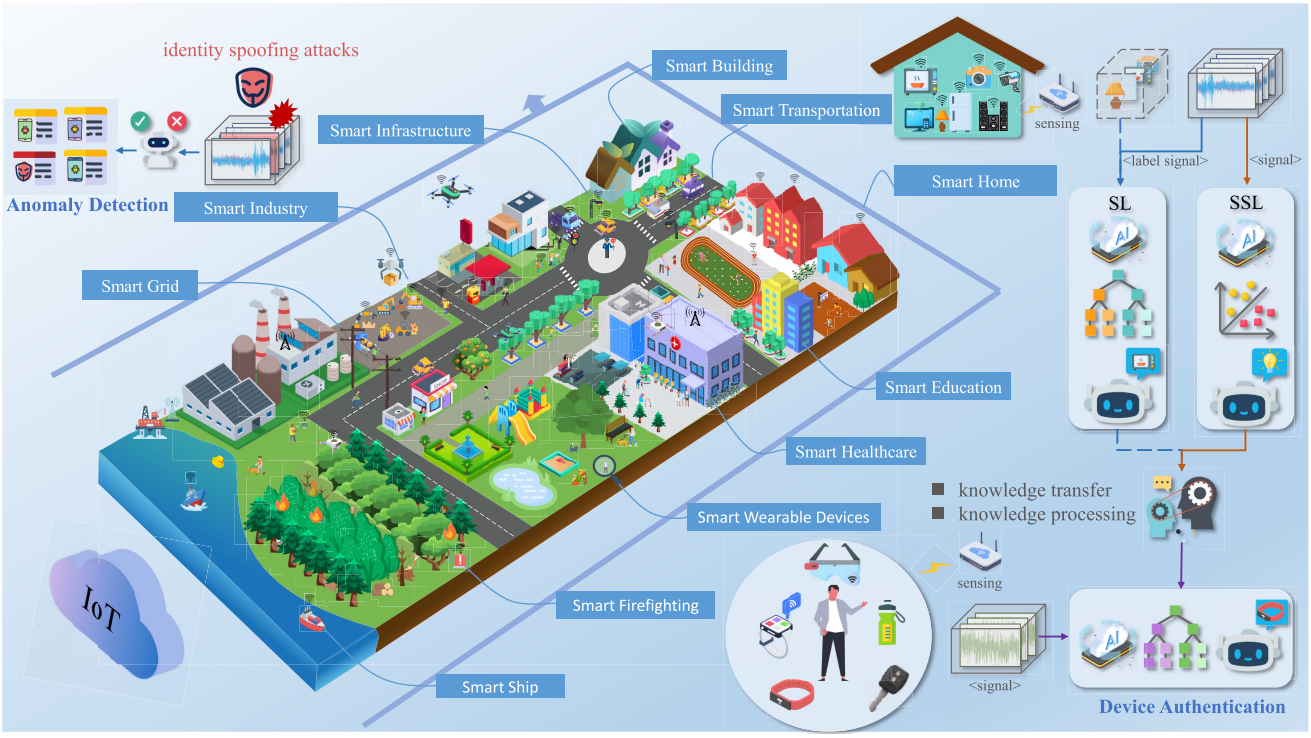


Fig. 1. Application domains for SEI in IoT. With the proliferation of IoT devices for smart transportation, smart home, smart industry, etc., SEI is emerging as a promising technology, especially in device authentication and anomaly detection tasks. Each IoT sub-scenario can use the technology with local data to train a local physical layer device authentication model, or it can be used to assist other IoT sub-scenarios through knowledge transfer and processing.

non-intrusive to the device itself and requires no hardware modifications.

B. Research Motivation

A typical SEI system consists of signal collection, preprocessing, RFF extraction and identification steps [10], with the latter two being the core components of the process. Traditional RFF extraction methods primarily rely on statistical techniques to manually design features, which are then classified using Machine Learning (ML) algorithms. However, such handcrafted features often suffer from suboptimal performance, limited adaptability to changing environments, and poor scalability in the face of growing IoT device diversity. In recent years, Deep Learning (DL) has demonstrated strong capabilities in data analysis and feature extraction, and has been widely applied in wireless communication technologies [11]. DL-based, data-driven methods eliminate the need for manual feature engineering and enable the extraction of more effective and adaptable RFF representations, showing great promise in SEI applications.

However, most high-performing DL-based SEI methods rely on supervised learning (SL), which requires large amounts of high-quality annotated data that is often labor-intensive and costly. This labeling overhead becomes particularly challenging in certain IoT scenarios where labeled data is impractical or infeasible to obtain. While some studies have explored semi-supervised or few-shot learning to mitigate this issue, there remains a strong need for fully label-free approaches that can leverage large volumes of unlabeled data to learn effective RFF features for SEI. Self-supervised learning (SSL), a subset of unsupervised learning (USL) that constructs pretext task to learn useful representations without labels, has achieved remarkable success in fields such as Computer Vision (CV) [12] and natural language processing [13]. A detailed comparison between SL and SSL is provided in Table II. The label-independent nature and advantages of the SSL make it highly promising for SEI applications. Therefore, reviewing the latest SSL-based SEI methods is essential to raise awareness, consolidate recent advances, and offer guidance for future research.

C. Survey Methodology

To ensure comprehensive coverage of existing studies on SSL applied to SEI, we adopted a structured literature review approach. The review process involved the following steps:

- Databases Searched: IEEE Xplore, ACM Digital Library, SpringerLink, and Google Scholar were primarily used.
- Search Keywords: Keywords included combinations such as “self-supervised learning”, “SSL”, “specific emitter identification”, “SEI”, “radio frequency fingerprinting”, and “unsupervised learning”.
- Time Range: We focused on literature published between 2016 and 2025, with emphasis on the most recent developments in SSL and SEI.
- Inclusion Criteria: Studies that (1) proposed or applied SSL techniques to SEI or related RF tasks, (2) introduced novel SSL paradigms relevant to time-series or signal data, or (3) benchmarked SSL models on open RF datasets were included.

TABLE I
LIST OF ACRONYMS USED IN THE PAPER

Acronyms	Definitions	Acronyms	Definitions
ACGAN	Auxiliary Classifier Generative Adversarial Network	InfoNCE	Noise-Contrastive Estimation
AE	Autoencoder	InfoGAN	Information Maximized Generative Adversarial Network
AMC	Automatic Modulation Classification	IoT	Internet of Things
AMAE	Asymmetric Masked Auto-Encoder	KT	Knowledge Transfer
ARI	Adjusted Rand Index	KNN	Nearest Neighbor Classifier
AWGN	Additive White Gaussian Noise	LMMD	Local Maximum Mean Discrepancy
BPS	Bit-Pulse Selection	LLM	Large Language Model
BYOL	Bootstrap Your Own Latent	MAE	Masked Autoencoder
CFO	Carrier Frequency Offset	MAML	Model-Agnostic Meta-Learning
CGAN	Conditional Generative Adversarial Networks	ML	Machine Learning
CMC	Contrastive Multiview Coding	MoCo	Momentum Contrast
CNN	Convolutional Neural Network	MSE	Mean Squared Error
CPC	Contrastive Predictive Coding	NCE	Noise-Contrastive Estimation
CSSL	Contrastive Self-Supervised Learning	NMI	Normalized Mutual Information
DAC	Digital-to-Analog Converter	PAC	Probably Approximately Correct
DAE	Denoising Autoencoder	PA	Power Amplifier
DB	Davies-Bouldin	PIRL	Pretext-Invariant Representation Learning
DCTF	Constellation Trace Figure	PLA	Physical Layer Authentication
DL	Deep Learning	PSSL	Predictive Self-Supervised Learning
DoS	Denial-of-Service	PSD	Power Spectral Density
EMD	Empirical Mode Decomposition	Q	Quadrature-Phase
ERM	Empirical Risk Minimization	ResNet	Residual Network
GAI	Generative Artificial Intelligence	RF	Radio Frequency
GAN	Generative Adversarial Network	RFF	Radio Frequency Fingerprint
GSSL	Generative Self-Supervised Learning	SEI	Specific Emitter Identification
HCTF	Heat Constellation Trace Figure	SimCLR	Simple Framework for Contrastive Learning
I	In-Phase	SimSiam	Simple Siamese
IMF	Intrinsic Mode Function	SL	Supervised Learning
SSL	Self-Supervised Learning	SVM	Support Vector Machine
SCSC	Signal Contrastive Self-Supervised Clustering	SVD	Singular Value Decomposition
SC	Silhouette Coefficient	SwAV	Swapping Assignments Between Multiple Views of the Same Image
SNR	Signal-To-Noise Ratio	VAT	Virtual Adversarial Training
USL	Unsupervised Learning	VicReg	Variance-Invariance-Covariance Regularization

TABLE II
THE COMPARISON BETWEEN SL AND SSL

Comparison	SL	SSL
Description	Labelling is required and the goal is to learn the mapping relationship between inputs and outputs in order to make accurate predictions.	Learning feature representations of data through pretext tasks without external labeling.
Advantages	• Accurate results: SL can learn from accurate training data, so it can get more accurate results. • Efficient: Due to the provision of detailed labeling information, SL requires less data and is more efficient. • Wide range of applicability: SL has a wide range of uses and can be applied to a variety of fields, such as open-set identification, anomaly detection, etc.	• Ideal for large datasets: SSL can discover patterns, structures, and features from large amounts of unlabelled data, which is very useful for large datasets. • Does not require labeled data: SSL does not require labeled data, so there are no problems associated with providing incorrectly labeled data, reducing costs. • New domains can be explored: Since no labeled data is required, new domains can be explored using SSL to discover new patterns and trends.
Disadvantages	• High cost of labeled data: Labelling data usually requires expert experience, which is time-consuming and costly. • High data requirements: SL requires high-quality data, because wrong data can mislead model training and obtain wrong results.	• Bias exists in learning objective: As the data is not labeled, the model can not learn the desired representation explicitly. • Sensitive to parameter setting and algorithm selection: SSL is more complex and involves important issues such as parameter setting and pretext task design, which require more experience.

- Exclusion Criteria: Works that only dealt with fully supervised SEI methods and purely theoretical SSL papers unrelated to wireless applications were excluded.

Through this process, over 80 relevant papers were initially identified, and approximately 30 of the most representative and technically sound studies were selected for detailed discussion in this review.

D. Existing Surveys: Contributions & Limitations

As authentication security in the IoT garners increasing attention, several surveys and reviews on PLA technologies, including RFF, have been published. Table III summarizes the key contributions and limitations of these papers that have compiled literature on PLA techniques, particularly focusing on RFF and SEI [6], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. These surveys and reviews describe various application scenarios, including IoT, wireless network, global navigation satellite systems, industry wireless communications, etc. Table III lists the associations and differences between this survey and others.

It can be observed that most reviews organize and introduce PLA techniques from a broader perspective. For example, Xie et al. [6] categorized existing PLA schemes into passive and active types, where RFF-based and channel-feature-based techniques are discussed under passive schemes. Similarly, works by Al-Garadi et al. [15] and Angueira et al. [19] focus on IoT and industrial security applications, providing a high-level overview of PLA but lacking emphasis on SEI techniques. Several other reviews, such as those by Soltanieh et al. [16], Wang et al. [17], and Xie et al. [21], delve deeper into RFF or SEI from signal processing or DL perspectives. However, their discussions are often limited to conventional supervised learning techniques or assume the availability of abundant labeled data.

Table III summarizes these existing efforts, highlighting both their contributions and limitations. A common gap among them is the lack of attention to label-efficient learning settings, particularly in SEI. While Jagannath et al. [18] and Tyler et al. [20] explore open challenges and dataset issues, they stop short of categorizing or analyzing SEI approaches under label scarcity. Despite the increasing relevance of SSL in representation learning, no existing survey, based on our review, has systematically examined its application to SEI. The potential of SSL in SEI, especially for scenarios with limited or no labels, remains largely underexplored in the current literature. This gap motivates our work to provide a focused and structured review of SSL-based SEI, aiming to inform future research in this promising direction.

Fig. 2 illustrates the taxonomy of PLA schemes, which can broadly be categorized into passive and active PLA. Among these, SEI is a major subset, further divided into conventional SEI, DL-based SEI, and the emerging SSL-based SEI approaches. This survey emphasizes SSL-based SEI, distinguishing itself from other surveys by offering a deep and targeted discussion on this specific area. Compared with conventional methods and DL approaches, SSL-based SEI offers the potential to reduce reliance on labeled data while enhancing generalization capabilities, making it highly suitable for dynamic and large-scale IoT environments.

E. Our Survey: Organization & Contributions

This survey addresses a gap in the detailed analysis of SSL-based SEI by covering the latest research and providing readers with a comprehensive introduction to its fundamental concepts and taxonomy. The overall structure of the paper is outlined in Fig. 3, which illustrates the content of each section, the corresponding problems addressed, and the contributions made. Section II introduces the system model, detailing the process from signal generation with RFF at the transmitter to feature extraction and SEI at the receiver. It also summarizes representative methods for both well-labeled and label-limited scenarios. Section III presents an analysis and taxonomy of SL, USL, and SSL. It elaborates on the unified paradigms of SSL and explains the role of SSL in downstream tasks from both feature and application perspectives. Section IV lists currently available open-source SEI datasets with direct download links. It also categorizes existing SSL-based SEI methods based on the paradigms described in Section III. Section V discusses key research challenges and introduces an outlook for SSL-based SEI in the next era, considering the integration with large models. Section VI compares the current state-of-the-art SSL-based SEI methods with some classical SSL methods under unified experimental conditions and the related code.[1] will be publicized. Section VII discusses the implications of SSL-base SEI for both research and practical deployment. Finally, Section VIII concludes the paper. A list of frequently used acronyms is summarized in Table I.

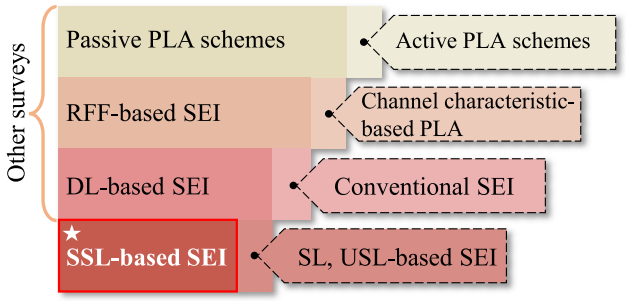


Fig. 2. The focused PLA schemes in this paper. This survey focuses on SSL-based SEI, providing more in-depth and specifically targeted research on the issue than other PLA surveys and reviews.

The main contributions of this survey are summarized as follows.

- We present a comprehensive and up-to-date survey of SSL applications in SEI, clearly outlining the background, motivations, and problem formulation, thus enabling readers to quickly grasp the core concepts of this emerging field.
- We categorize the existing SSL-based SEI schemes into three unified paradigms and perform a holistic investigation of the related literature according to the paradigms, which provides a clear roadmap for new researchers.
- We perform a comprehensive collection of existing open-source datasets for SEI with download links and provide a case study for several advanced SSL-based SEI algorithms. These resources aim to support fast prototyping and experimentation.
- We identify the limitations of current approaches and discuss future directions, particularly the integration of

[1] https://github.com/LIUC-000/SSL-SEI_Survey

TABLE III
EXISTING SURVEYS AND REVIEWS ON PLA TECHNIQUES

Reference	Topic	Contributions	Limitations
Xu et al. [14], 2016	RFF in wireless networks	A systematic review of fingerprint algorithms, including both white-list-based and unsupervised learning approaches, and introduces a taxonomy of features that can be used for fingerprinting.	Limited unsupervised SEI methods discussed. Lacks SSL methods.
Al-Garadi et al. [15], 2020	ML, DL, IoT security	A comprehensive survey of ML methods and recent advances in DL methods that can be used to develop enhanced security methods for IoT systems. Various IoT security threats and IoT attack surfaces are discussed.	RFF is missing from the PLA schemes discussed.
Soltanieh et al. [16], 2020	RFF	Theoretically analyzes the common transient RFF extraction techniques and discusses some approaches using the modulated part of the signal.	Limited discussion on steady state-based RFF extraction and identification.
Xie et al. [6], 2020	PLA in wireless communications	A comprehensive survey on the PLA from passive and active schemes.	Lacks detailed introduction for DL-based SEI.
Wang et al. [17], 2021	RFF in global navigation satellite systems	A comprehensive survey of the RFF methods applicable in the context of global navigation satellite systems for spoofing mitigation.	Lacks exploration of the latest DL-based RFF schemes.
Jagannath et al. [18], 2022	RFF, signal intelligence	A systematic review of RFF, including background, applications, relevant DL algorithms, literature review spanning the past two decades, discussion on datasets, and potential research avenues.	Lacks discussion of SEI in scenarios with insufficient labeled samples. The summarized open-source datasets are not comprehensive enough.
Angueira et al. [19], 2022	Physical layer techniques for wireless communications in the industry	A comprehensive survey of security aspects of industrial wireless communications, which analyzes representative references from industry, academia, and standardization bodies.	Lacks a deeper discussion and detailed classification of RFF schemes.
Tyler et al. [20], 2023	SEI	A review of SEI publications from the perspective of its use as a practical, effective, and usable IoT security mechanism. Considers operating conditions, SEI threats, SEI at scale, publicly available data sets, and SEI considerations.	Lacks discussion of SEI in scenarios with insufficient labeled samples.
Xie et al. [21], 2023	RFF in IoT	A comprehensive survey on RFF identification which covers signal preprocessing, current schemes of RFF feature extraction, and further processing methods. Refines the framework of RFF identification from the perspective of closed-set and open-set problems.	Lacks discussion of SEI in scenarios with insufficient labeled samples. The summarized open-source datasets are not comprehensive enough.
Zhang et al. [22], 2023	RFF in IoT	Reviews both the state-of-the-art in engineered feature-based RFF identification protocol design and advances in recent DL-based protocols, as well as a hybrid protocol that combines their advantages.	Lacks a detailed classification of DL-based RFF identification algorithms.
Abbas et al. [23], 2023	RFF in IoT	A systematic literature review of RFF focuses on exploring the commonly used RFF approaches, feature extraction and filtration techniques, and classification algorithms used in device identification.	Lacks categorization according to different application scenarios.

SSL with large models, offering insights for researchers interested in advancing this area.

II. RADIO FREQUENCY FINGERPRINT EXTRACTION FOUNDATIONS: MODELS & TECHNIQUES

A. System Model

The system model, from the generation of signals with RFF at the transmitter to the extraction of RFF and SEI at the receiver, is shown in Fig. 4. During the signal processing phase at the transmitter, the binary bits sequence from the data source is first converted into complex transmit symbols by the digital modulation block.

The symbols are then transformed into the analog baseband signal, $x(t)$, through a Digital-to-Analog Converter (DAC). Following up-conversion and power amplification, $x(t)$ is converted to a modulated signal, $s(t)$, transmitted from the antenna and received by the receiver through the wireless channel. The received radio frequency (RF) signal $r(t)$ can be expressed as

$$r(t) = \mathcal{F}(x(t)) * h(t) + n(t), \tag{1}$$

where $h(\cdot)$ represents the channel response, and $n(t)$ indicates Additive White Gaussian Noise (AWGN). "$*$" represents the convolution operation. $\mathcal{F}(\cdot)$ encompasses the overall effects

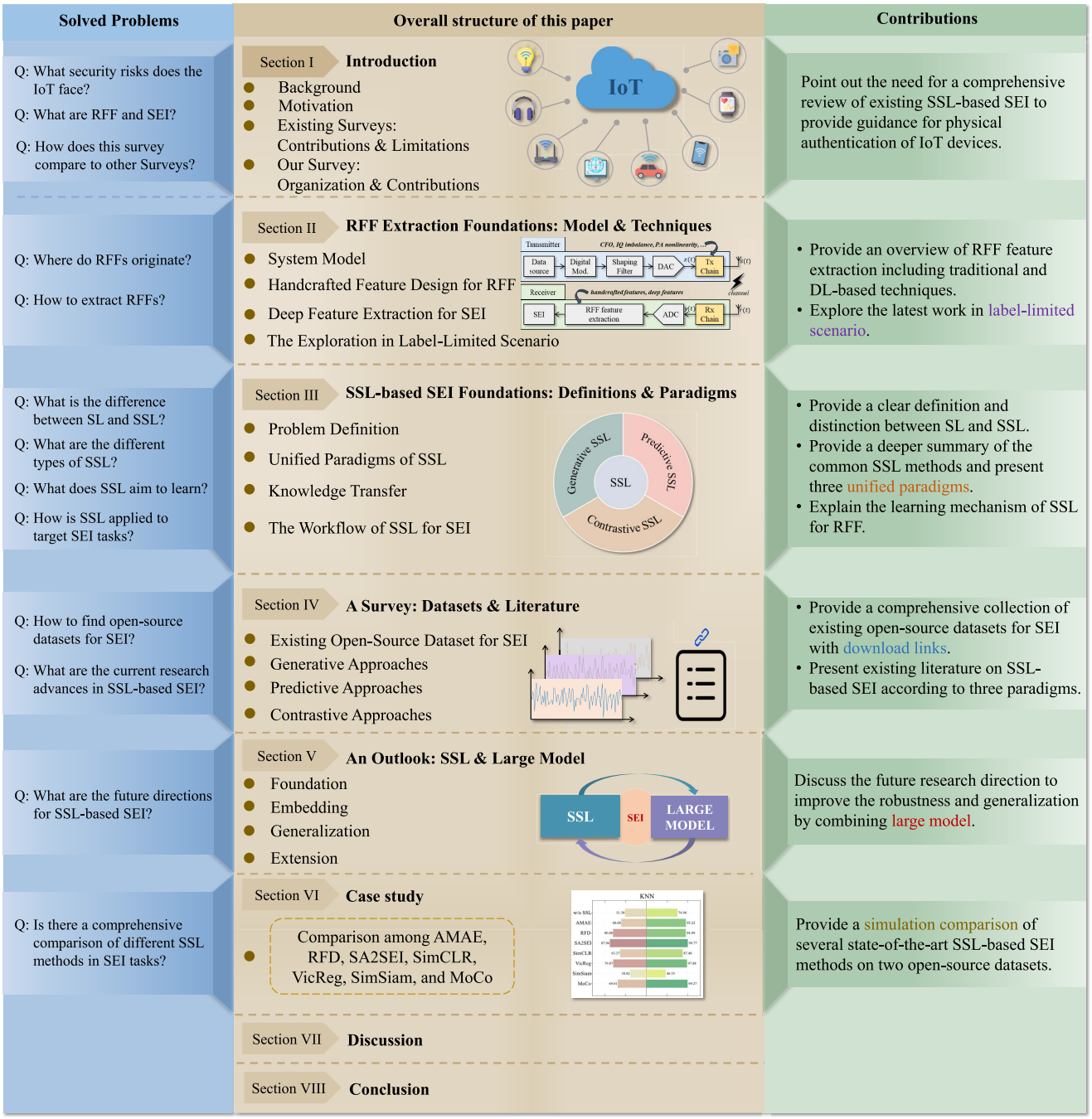


Fig. 3. General framework of the solved problems, overall structure of this paper, and contributions.

of the transmitter chain, including intentional modulation that generates symbol information and unintentional modulation that generates RFF features. RFF features arise from the hardware imperfections in the transmitter chain's electronic modules, such as Carrier Frequency Offset (CFO), phase noise, In-phase (I) and Quadrature-phase (Q) imbalance, Power Amplifier (PA) nonlinearity, and antenna patterns. These imperfections cannot be eliminated and subtly affect the transmitted signal's waveform. Due to the uniqueness and stealthiness of RFF, it plays a crucial role in SEI tasks. Therefore, effectively extracting high-quality RFF features remains a significant and promising research direction.

B. Handcrafted Feature Design for RFF

1) Transient Features: Transient signals are the brief signals produced when communication devices initiate or cease signal transmission. These signals encapsulate the characteristics of device circuit activation and deactivation, often harboring rich RFF information. Therefore, early SEI work involved experts designing and extracting features from transient signals based on experience. Choe et al. [24] employed a Daubechies-4 wavelet to characterize transient signals over time, achieving the authentication of unknown transmitters through an artificial neural network. Hall et al. [25] firstly extracted the transient portion of the signal and then

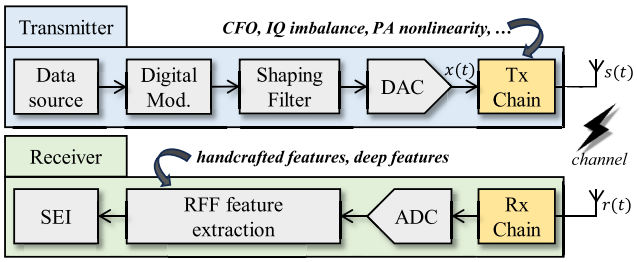


Fig. 4. The system model, from the generation of signals with RFF at the transmitter to the extraction of RFF and SEI at the receiver.

extracted the amplitude, frequency and phase components as the transient features to authenticate Bluetooth devices. Then Rasmussen and Capkun et al. [26] used the variance of the transient amplitude to categorize wireless sensor devices, achieving a 70% accuracy rate in authentication. Suski II et al. [27] enhanced authentication accuracy in regions with low Signal-to-Noise Ratio (SNR) by leveraging the Power Spectral Density (PSD) of a preamble signal during the transient phase as a distinguishing feature. Köse et al. [28] extracted the energy spectrum of the transient amplitude envelopes to differentiate eight distinct devices. Tu et al. [29] utilized the statistical metrics and wavelet features of turn-on transient signals as the RFF of transmitters. Ali et al. [30] constructed the RFF from transient signals based on 13 features across time and frequency domains using the Hilbert-Huang spectrum. Although RFFs designed from transient signals are of high quality, the challenges in collecting transient signals and the complexity in processing and analysis have hindered the widespread application of SEI.

2) Steady-State Features: Compared to transient signals, steady-state signals are easier to capture due to their longer duration and higher stability during communication processes, making steady-state RFF feature extraction a more popular research direction. For example, Zhao et al. [31] proposed a robust physical layer authentication algorithm based on phase noise fingerprinting. In the authentication scheme [32], the CFO of the received signal is initially estimated and used for authentication through consistent comparison. Later, the CFO was modeled as an autoregressive random process by analyzing the variable CFO values at different communication times [33]. Dolatshahi et al. [34] leveraged the nonlinearity differences in the input/output characteristics of power amplifiers, modeled through the Volterra series, and designed algorithms to identify wireless users based on these differences. Polak et al. [35] modeled the RFF with imperfections of PA and DAC to identify commercially used RF transmitters, achieving high performance even with short data records and relatively low SNR. Hao et al. [36] utilized device-dependent IQ imbalance as RFF and combined the collaboration of multiple receivers for wireless authentication.

In summary, traditional handcrafted RFF feature design typically relies on physical signal characteristics such as signal strength, phase, and frequency. These methods are comparatively straightforward to understand and implement, generally do not demand extensive computational resources or datasets, and are practical in resource-constrained environments. However, these methods depend on predefined features designed by expert experience, potentially failing to capture all subtle differences in signals, thereby limiting their performance.

C. Deep Feature Extraction for SEI

DL transforms RFF by shifting from handcrafted feature design to automated, data-driven feature extraction. This approach enhances SEI accuracy and adaptability in dynamic environments, setting a new standard in secure wireless communication. SEI based on DL (DL-SEI) is an end-to-end framework, which can be categorized into raw signal-based and transformed signal-based approaches, depending on the form of the input signals to the framework.

1) Raw Signal-Based DL-SEI: Using raw RF signals in IQ format as the input for training and identification maintains the integrity of all original signal characteristics. Merchant et al. [9] demonstrated the effectiveness of Convolutional Neural Network (CNN) for RFF extraction, achieving high identification and verification on 7 ZigBee devices using raw steady-state IQ signals. Yu et al. [37] proposed a multisampling CNN, which extracts multiscale features via multiple downsampling transformations from raw IQ signals. Sankhe et al. [38] combined raw and feature-engineered IQ samples, achieving 80-95% SEI accuracy on 100+ WiFi devices and 16 X310 USRP radios in a static environment, and 99% accuracy after mitigating wireless channel impact by intentionally inserting and learning transmitter impairment effect. Zha et al. [39] proposed a novel complex Fourier neural operator embedded with a time and frequency domain attention mechanism to learn features from different domain perspectives. Zhang et al. [40] identified performance degradation in transmitter-variable modulation scenarios and proposed a variable-modulation SEI framework to reduce data distribution distinctions among different modulation types by constructing ideal modulation signals. Zhang et al. [41] explored a multisource heterogeneous SEI method using real-world multisource heterogeneous RF signals, comprising a multi-channel RFF feature extraction module, an attention-based RFF fusion module, and an automatic classifier, demonstrating its advantages in noisy environments. Although raw signal-based methods offer simplicity, preservation of original characteristics, and suitability for real-time applications, they are more susceptible to noise and interference in the raw signal, potentially affecting identification accuracy.

2) Transformed Signal-Based DL-SEI: This method involves processing raw RF signals through various transformations or feature extraction techniques before training or identification, such as spectral analysis, wavelet transforms, or statistical feature extraction. Ding et al. [42] transformed received steady-state signals into bispectrum and used a supervised dimensionality reduction method to obtain the dimension-reduced bispectrum as the input for CNN. Pan et al. [43] performed the Hilbert-Huang transform on the received signals to obtain the Hilbert spectrum images as input for a deep Residual Network (ResNet), finding

that the extracted RFFs are more effective and robust than the handcrafted ones. Peng et al. [44] used Differential Constellation Trace Figure (DCTF) of signals as input for CNN to extract RFF features, achieving high identification accuracy and low complexity without prior information. Peng et al. [45] used the Heat Constellation Trace Figure (HCTF) to avoid the issues with the length and synchronization of RF signals, achieving high SEI performance with a slice integration cooperation. To overcome the wireless channel distortion on the received signal, Shen et al. [46] proposed a channel-independent spectrogram to mitigate the channel effects in the time-frequency domain while preserving RFF characteristics, combining it with deep metric learning to train a channel-robust RFF extractor. Li et al. [47] transformed the signal into graph tensors and captured the connections among various sampling points as well as the relationships among different fingerprint features at a single sampling point within the graph tensor. Although transformed signal-based methods can be less sensitive to noise and signal interference, their effectiveness heavily relies on the chosen signal transformation or feature extraction technique, which requires high-quality expert experience. Additionally, the transformation process might result in the loss of RFF information.

D. The Exploration in Label-Limited Scenario

Although DL continues to redefine the boundaries of SEI through its ability to extract complex patterns from large datasets, it remains heavily dependent on the availability of extensive labeled datasets for training. This dependence poses significant challenges in practical applications, where such labeled signals are scarce or costly to obtain. Consequently, researchers have begun to explore SEI under data-limited or label-limited scenarios, such as few-shot and semi-supervised learning, with the goal of developing methodologies that can achieve robust SEI using partially labeled datasets. These approaches represent a critical shift towards making SEI more feasible and effective in environments constrained by limited data availability. For example, Yang et al. [48] introduced a strategy for SEI by employing Model-Agnostic Meta-Learning (MAML), achieving high precision with only a small set of labeled training examples. Xie et al. [49] proposed a semi-supervised SEI algorithm based on bispectrum analysis and Virtual Adversarial Training (VAT), which trains the CNN in an adversarial manner with a small amount of labeled data and exhibits excellent generalization capability. Fu et al. [50] proposed a Metric-Adversarial Training (MAT) algorithm combining deep metric learning and VAT to learn RFF features from an insufficiently labeled training dataset and a large unlabeled training dataset. X. Fu further innovated by introducing pseudo labels into semi-supervised metric learning to extract the discriminative semantic features of signals, improving SEI performance [51]. Cai et al. [52] employed data augmentation techniques, including noise disturbance and time-delay transformation at the data level, along with VAT at the model level, to enhance the small sample dataset and improve the SEI model's generalization and robustness.

Amidst the evolving landscape of SEI, the reliance on labeled datasets poses a significant barrier, especially in dynamic or non-cooperative environments where labels are not available. This challenge has spurred growing interest in alternative approaches that can circumvent the dependence on labels, thus facilitating the emergence of unsupervised and self-supervised SEI methodologies. These innovative strategies hold the potential to unlock new capabilities in SEI by leveraging unlabeled data, offering a promising avenue for enhancing identification capabilities in scenarios where annotated samples are unavailable. In this paper, we focus on the application of SSL in SEI.

III. SSL-Based SEI Foundations: Definitions & Paradigms

Given the promising potential of SSL in SEI, this section provides an in-depth investigation of SSL-based SEI techniques. Specifically, Section III-A formally defines the SSL-based SEI problem and compares it with SL-based and USL-based SEI. Section III-B presents a comprehensive overview of existing SSL-based SEI approaches, categorizing them into three representative paradigms. Finally, Sections III-C and III-D discuss the applications of SSL-based SEI from the perspectives of underlying mechanisms and practical implementation workflows, respectively.

A. Problem Definition

1) Supervised Learning With Ground Truth Label: SL is a data-driven learning scheme within DL, aimed at understanding and extracting intrinsic characteristics from the data. Let $\mathcal{X}$ represent the domain set of objects for which we wish to obtain target labels, and $\mathcal{D}$ be the probability distribution over $\mathcal{X}$. $\mathcal{Y}$ indicates the set of ground truth labels for $\mathcal{X}$. For example, in SEI, $\mathbf{x}_i \in \mathcal{X}$ represents a signal received from a specific transmitter, and $\mathrm{y}_i \in \mathcal{Y}$ represents the label of the individual transmitter. Due to the unknown nature of $\mathcal{D}$ in real-world scenarios, a learner A_{sl} has access to only a finite sequence of pairs $S = \{(\mathbf{x}_1, \mathrm{y}_1) \cdots (\mathbf{x}_m, \mathrm{y}_m)\}$, often referred to as the training dataset. $\mathbf{x}_i$ in S are sampled from $\mathcal{D}$, and y_i is obtained by manual labeling based on expert experience. Assuming there exists a perfectly correct labeling function f that can label all $\mathbf{x}_i$ as y_i. The learner A_{sl} needs to output a mapping rule $h : \mathcal{X} \to \mathcal{Y}$ to fit f as closely as possible. In DL, h consists of network structure c and network parameters θ. Typically, c is designed to be fixed, and the space of different θ, which constitutes a set of hypotheses, is defined as $\mathcal{H}$, so $h \in \mathcal{H}$. The optimal mapping rule $A_{sl}(\mathcal{D})$ based on expected risk minimization can be approximated by the Empirical Risk Minimization (ERM) on the S as $A_{sl}(S)$. The approximation process is expressed as

$$A_{sl}(\mathcal{D}) = \arg\min_{h_{c,\theta} \in \mathcal{H}} \mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}} \mathcal{L}_{sl}\big(h_{c,\theta}(\mathbf{x}_i), f(\mathbf{x}_i)\big) \tag{2a}$$

$$\overset{\text{fix c}}{\approx} \arg\min_{h_\theta \in \mathcal{H}} \mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}} \mathcal{L}_{sl}(h_\theta(\mathbf{x}_i), f(\mathbf{x}_i)) \tag{2b}$$

$$\overset{\text{ERM}}{\approx} \arg\min_{h_\theta \in \mathcal{H}} \mathbb{E}_{(\mathbf{x}_i, \mathrm{y}_i) \sim S} \mathcal{L}_{sl}(h_\theta(\mathbf{x}_i), \mathrm{y}_i) \tag{2c}$$

$$= A_{sl}(S), \tag{2d}$$

where $\mathcal{L}_{sl}$ is the supervised loss, such as cross-entropy, which measures the difference between the output of h and the

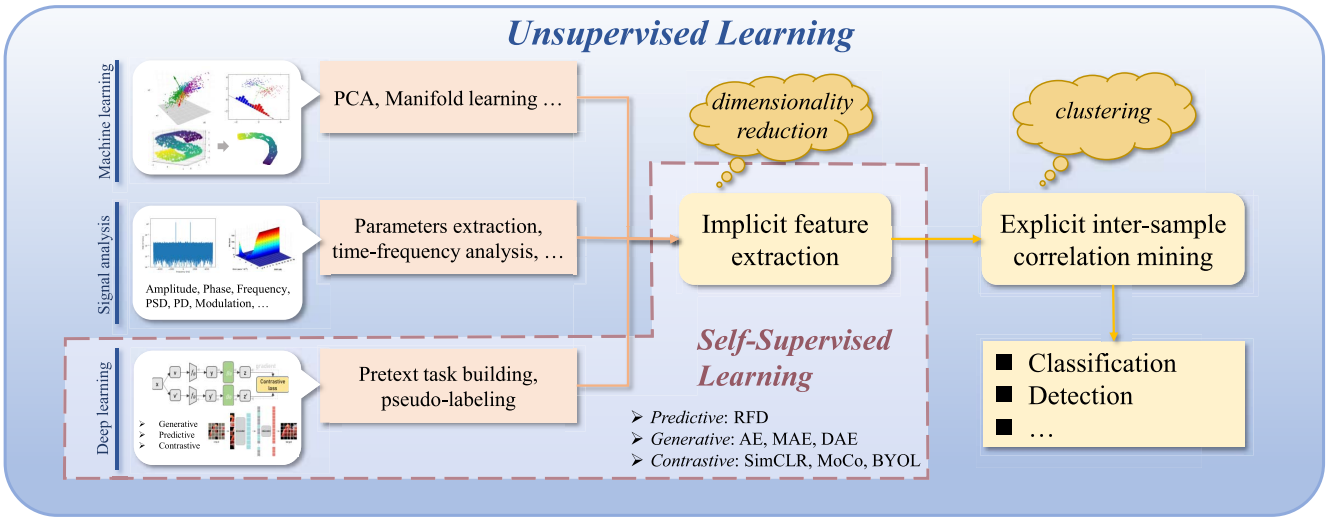


Fig. 5. The correlation between USL and SSL.

ground truth label. According to the Probably Approximately Correct (PAC) theory, if the sample size exceeds a certain value, the difference between the empirical risk obtained through ERM and the expected risk will be less than a certain threshold. Therefore, the larger S, the more closely $A_{sl}(S)$ can approximate $A_{sl}(\mathcal{D})$. However, in SL, the larger labeled datasets mean more work on label annotation, which can be costly and time-consuming. Non-cooperative scenarios in signal recognition, where many acquired signals cannot be manually labeled, pose a significant challenge to SL and promote the exploration of USL and SSL.

2) Self-Supervised Learning With Pseudo Label: SSL can be regarded as a specific case of USL, and the association of USL and SSL in the field of signal recognition is shown in the upper part of Fig. 5. Due to the unavailability of labels, USL can only start from the data itself, extracting implicit features of the signals through methods such as ML, signal analysis, and DL. Based on these features, USL mines explicit inter-sample correlations, such as clustering, ultimately completing signal classification or detection tasks. In particular, the term "implicit features" refers not only to the output of the neural network but also to all basic information extracted from the input signal, such as the vectors after dimensionality reduction, the amplitude and phase of the signal, the Power Spectral Density (PSD), the type of modulation, etc. SSL, as a type of USL method, trains the network using SL-like operations by building a pretext task based on deep neural networks and designing pseudo labels. Given a designed set of pseudo labels $\tilde{\mathcal{Y}}$, where $\tilde{\mathrm{y}}_i \in \tilde{\mathcal{Y}}$ represents some intrinsic structure or information of $\mathbf{x}_i$, the SSL learner A_{ssl} aims to output a mapping rule $\tilde{h} : \mathcal{X} \to \tilde{\mathcal{Y}}$ to fit pseudo labeling function $\tilde{f}$. Unlike in SL, where the function f is unobtainable, in SSL, the function $\tilde{f}$ is artificially designed and known. Formally, let $S|_{\mathbf{x}} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ represent the instances of the training dataset, the approximated SSL learner A_{ssl} can be obtained from

$$A_{ssl}(\mathcal{D}) = \underset{\tilde{h}_{\tilde{c},\tilde{\theta}} \in \tilde{\mathcal{H}}}{\arg\min}\, \mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}} \mathcal{L}_{ssl}\left(\tilde{h}_{\tilde{c},\tilde{\theta}}(\mathbf{x}_i), \tilde{f}(\mathbf{x}_i)\right) \tag{3a}$$

$$\overset{\text{fix } \tilde{c}}{\approx} \underset{\tilde{h}_{\tilde{\theta}} \in \tilde{\mathcal{H}}}{\arg\min}\, \mathbb{E}_{\mathbf{x}_i \sim \mathcal{D}} \mathcal{L}_{ssl}\left(\tilde{h}_{\tilde{\theta}}(\mathbf{x}_i), \tilde{f}(\mathbf{x}_i)\right) \tag{3b}$$

$$\overset{\text{ERM}}{\approx} \underset{\tilde{h}_{\tilde{\theta}} \in \tilde{\mathcal{H}}}{\arg\min}\, \mathbb{E}_{\mathbf{x}_i \sim S|_{\mathbf{x}}} \mathcal{L}_{ssl}\left(\tilde{h}_{\tilde{\theta}}(\mathbf{x}_i), \tilde{f}(\mathbf{x}_i)\right) \tag{3c}$$

$$= A_{ssl}(S|_{\mathbf{x}}). \tag{3d}$$

Although both SSL and SL train networks in a supervised manner, comparing Eq. (2b) and Eq. (3b), the distinctions can be observed from three perspectives: data, model, and loss. The data in SL is goal-oriented, including instances and ground truth labels y_i that specify the goal required by the target task. In contrast, the data in SSL is task-oriented, with pseudo labels $\tilde{f}(\mathbf{x}_i)$ generated according to the type of pretext task. Different pretext tasks design different $\tilde{f}$, resulting in pseudo labels that are highly relevant to the pretext tasks but different from the ground truth labels. Similarly, the model structure and loss also vary significantly depending on the differences in pretext tasks in SSL. In the following content, we classify and discuss SSL in detail based on pretext tasks.

B. Unified Paradigms of SSL

According to the type of pretext task, common SSL methods can be divided into three paradigms: Generative SSL (GSSL), Predictive SSL (PSSL), and Contrastive SSL (CSSL). The training schemes of these different paradigms are shown in Fig. 6.

1) GSSL: As shown in part (a) of Fig. 6, the family of GSSL methods is characterized by prompting the mapping rule $\tilde{h}_{gssl}$ to restore distorted input to its original form. The distortion function is defined as $d(\cdot)$ and mainly includes noise, loss of information, data point corruption, etc. For GSSL, the loss in Eq. (3b) can be written as follows:

$$\mathcal{L}_{gssl}\left(\tilde{h}_{\tilde{\theta}}(\mathbf{x}), \tilde{f}(\mathbf{x})\right) = \mathcal{L}_{sim}\left(\tilde{h}_{gssl}(d(\mathbf{x})), \mathbf{x}\right). \tag{4}$$

From the data perspective, the input of GSSL is the distorted instance $d(\mathbf{x})$, and the pseudo label is the original instance $\mathbf{x}$. From the perspective of the model, $\tilde{h}_{gssl}$ typically manifests

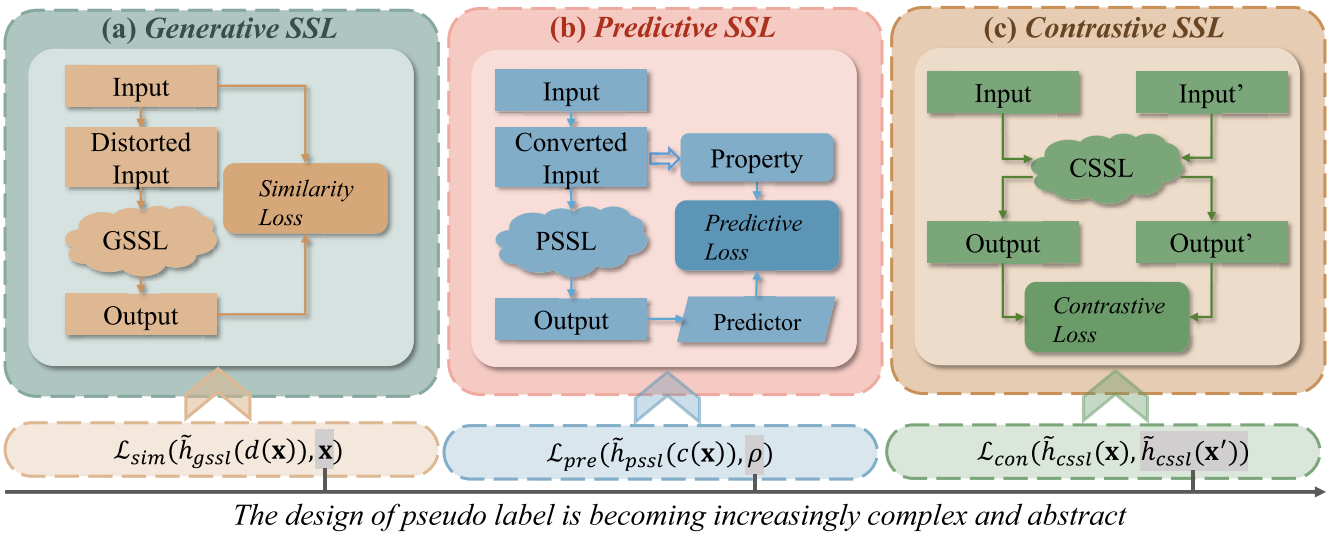


Fig. 6. Categorization of SSL: generative, predictive and contrastive paradigms.

as an encoder-decoder structure, where the encoder performs feature extraction on the distorted instances, and the decoder is motivated to generate the original instances based on these features. From the perspective of loss, $\mathcal{L}_{sim}$ is a similarity loss that measures the similarity between the generated instances and the original instances, such as the Mean Squared Error (MSE).

2) PSSL: As shown in part (b) of Fig. 6, the family of PSSL methods is characterized by prompting the mapping rule $\tilde{h}_{pssl}$ to predict the property of the converted instance. The conversion functions $c(\cdot)$ may include sequence shuffling, geometric transformation, and real-false samples, with the corresponding properties ρ being the correct sequence, type of transformation, and sample legitimacy, respectively. For PSSL, the loss in Eq. (3b) can be written as follows:

$$\mathcal{L}_{pssl}\left(\tilde{h}_{\tilde{\theta}}(\mathbf{x}), \tilde{f}(\mathbf{x})\right) = \mathcal{L}_{pre}\left(\tilde{h}_{pssl}(c(\mathbf{x})), \rho\right). \tag{5}$$

From the perspective of data, the input of PSSL is the converted instance $c(\mathbf{x})$, where $c(\cdot)$ is sampled from a finite set of conversions $\{c_1(\cdot), \ldots, c_K(\cdot)\}$. Each conversion corresponds to a property, so the property $\rho \in \{\rho_1, \ldots \rho_K\}$ is designed as the pseudo label of $c(\mathbf{x})$. From the perspective of the model, $\tilde{h}_{pssl}$ is a standard supervised classification structure, including an encoder and a classifier. The classifier accepts features from the encoder and outputs category confidence belonging to ρ_i via softmax. From the perspective of loss, $\mathcal{L}_{pre}$ is a supervised loss used to measure the degree of inconsistency between the model's predicted property and the actual property, such as the cross-entropy [53].

3) CSSL: As shown in part (c) of Fig. 6, the family of CSSL methods is characterized by prompting the mapping rule $\tilde{h}_{cssl}$ to learn transformation-invariant representations embedded in different views of one instance. The views after transformation might be different modalities of the instance or variants of the instance obtained through $d(\cdot)$ or $c(\cdot)$. For CSSL, the loss in Eq. (3b) for CSSL can be rewritten as follows:

$$\mathcal{L}_{cssl}\left(\tilde{h}_{\tilde{\theta}}(\mathbf{x}), \tilde{f}(\mathbf{x})\right) = \mathcal{L}_{con}\left(\tilde{h}_{cssl}(\mathbf{x}).\tilde{h}_{cssl}(\mathbf{x}')\right). \tag{6}$$

From the perspective of data, the input of CSSL is the instance itself (which may also be transformed in some works), and the pseudo label is the output of CSSL for the different view $\mathbf{x}'$ of $\mathbf{x}$, which is an embedding vector in a high-dimensional representation space. From the perspective of the model, in some works, $\tilde{h}_{cssl}$ consists of two separate networks for the embedding vector representations of $\mathbf{x}$ and $\mathbf{x}'$. In other words, these two networks are structurally independent but share interconnected parameters, and in yet other works, $\mathbf{x}$ and $\mathbf{x}'$ share the same network. From the perspective of loss, $\mathcal{L}_{con}$ is a contrastive loss used to measure the degree of difference between a pair of embedding vectors, thereby bringing them closer together. Common self-supervised contrastive losses include mean MSE, cosine similarity, Noise Contrastive Estimation (InfoNCE), etc.

C. Knowledge Transfer

The mapping rule $\tilde{h}$ learned from SSL consists of the network structure $\tilde{c}$ and network parameters θ, where $\tilde{c}$ can be divided into a feature extractor $\tilde{q}$ and a pretext task-related head $\tilde{p}$. The full description of $\tilde{h}(\cdot)$ is expressed as

$$\tilde{h}(\cdot) = \tilde{h}_{\tilde{c},\tilde{\theta}}(\cdot) = \tilde{p}_{\tilde{\theta}}\left(\tilde{q}_{\tilde{\theta}}(\cdot)\right), \tag{7}$$

where $\tilde{q}$ learns the intrinsic representations of the input signal, achieving data compression and automatic feature extraction, and $\tilde{p}$ analyzes the correlation between features to complete the pretext task. Due to the differences between the pre-task and the target task, $\tilde{p}$ cannot be directly applied to the target task. However, $\tilde{q}$ has task universality, so the most common method is to replace $\tilde{q}$ with a head g related to the target task. Then the mapping rule of the target task can be written as

$$t(\cdot) = t_{c,\theta}(\cdot) = g_\theta(q_\theta(\cdot)), \ \tilde{q}_{\tilde{\theta}} \to q_\theta. \tag{8}$$

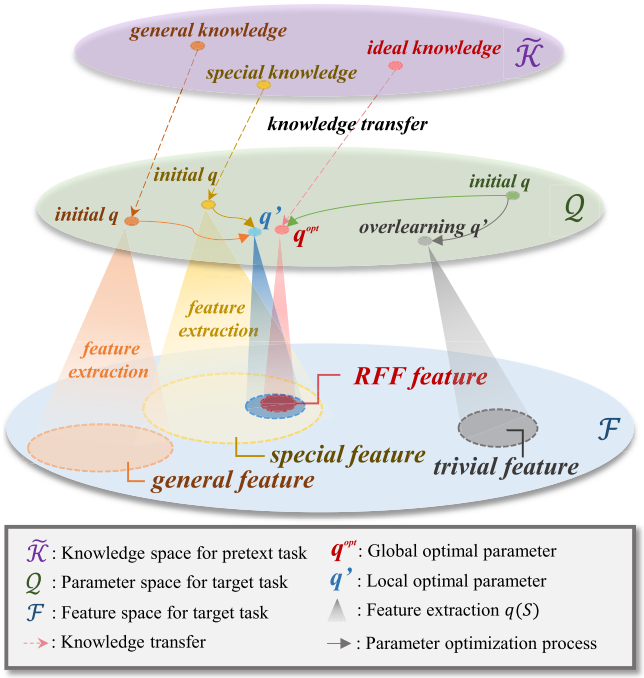


Fig. 7. The feature representation before and after KT.

Typically, the structure of the feature extractor for the pretext task remains the same as that for the target task. For the target task, $\tilde{q}_{\tilde{\theta}}$ can be regarded as a form of knowledge, and the process of $\tilde{q}_{\tilde{\theta}} \rightarrow q_\theta$ is referred to as Knowledge Transfer (KT). Different knowledge has varying effects on the target task, which can be reflected in the feature space $\mathcal{F}$ of the signal. As shown in Fig. 7, $\mathcal{F}$ represents all the features contained in the target instance $\mathcal{X}$. The goal of SEI is to extract as many RFF features as possible while reducing other irrelevant features. Key RFF features include oscillator imperfection, IQ imbalance, and power amplifier nonlinearity, among others. If a certain knowledge, after being transferred to the target task, results in q^{opt} that can perfectly extract RFF features, such knowledge is referred to as "ideal knowledge". However, this is difficult to achieve with pretext task-oriented SSL. The next best scenario is to acquire a "special knowledge" that is highly relevant to the target task. After KT, it can extract redundant features that contain RFF and other features, such as modulation features and wireless technique features, among others. SSL may also acquire a "general knowledge", which, after KT, can extract high-level features in the signal weakly associated with the RFF, such as denoising features and time-frequency features.

After KT, the target task can continue to train the parameters θ of q, obtaining the local optimum q' close to q^{opt}. Compared to random parameter initialization, KT can significantly reduce the likelihood of q being optimized to an over-learning state, which can only extract trivial features, such as amplitude features and channel characteristics, etc. In conclusion, a well-designed pretext task allows SSL to learn higher-quality knowledge from unlabeled data and optimize the extraction of RFF features in the target SEI task.

D. The Workflow of SSL for SEI

In the non-cooperative scenario, although the emitter identity of received signals is unknown, SSL can also learn knowledge from the unlabeled data by constructing a pretext task and transferring the knowledge to the target SEI task. However, due to the diversity of knowledge, the transferred knowledge may not fully match the target task. Therefore, it can be adjusted through a small amount of labeled data in the target task, significantly saving the resources consumed by manual labeling. This workflow of SSL for SEI is illustrated in Fig. 8. In the signal collection stage, the receiver collects signals from different emitters and obtains the unlabeled signal dataset $S|_{\mathbf{x}}$. Then the pseudo labeling function $\tilde{f}$, the network structure $\tilde{h}_{\tilde{c}}$, and the loss $\mathcal{L}_{ssl}$ should be designed by an expert to construct a pretext task. Based on the designed pretext task and $S|_{\mathbf{x}}$, SSL can learn a knowledge $\tilde{q}_{\tilde{\theta}}$ by training the network parameters $\tilde{h}_{\tilde{\theta}}$. For the signal dataset collected from target emitters, which can be different from the emitters of $S|_{\mathbf{x}}$, a small amount of labeling based on expert experience can be conducted to obtain the labeled dataset S. After KT $\tilde{q}_{\tilde{\theta}} \rightarrow q_\theta$ and building target task-orient head g_θ, the parameters θ can be adjusted via SL. Specifically, to prevent overfitting, some parameters in q_θ can be frozen, and the frozen parameters do not participate in parameter updates during SL. Finally, in the SEI stage, the newly collected signals can be classified or detected through the trained q_θ.

In this section, we clearly give the definition and common paradigms of SSL-based SEI, which can be summarized as a process of mining potential information in unlabeled signal data by constructing pseudo-labels and designing pretext tasks. We also explain why the potential information mined by SSL is useful and how to utilize this information on the target SEI task and signal data.

IV. A Survey: Datasets & Literature

Based on the problem definition, paradigms, and workflow of SSL on the SEI task presented in Section III, in this section, we collect existing open-source SEI datasets and comprehensively introduce the SSL-based SEI literature.

A. Existing Open-Source Dataset for SEI

In the quest to advance the field of SEI through SSL, the availability and diversity of datasets play a crucial role. However, unlike the well-established large-scale open-source datasets in the field of CV, open-source dataset resources for SEI are fragmented and lack uniformity. This discrepancy arises from the diversity of emitter types, the variety of signal samples, and the complexity of channel environments. Although few studies, such as those mentioned in [76], have attempted to catalog these datasets, their coverage remains insufficiently comprehensive. Recognizing this, we have meticulously compiled a comprehensive collection of open-source datasets, as shown in Table IV, which lists the key parameters, the type of emitter, the number of emitters, and a direct link to the download page. This makes it easy for researchers to quickly access and download the information about the dataset. Moreover, the table also provides some

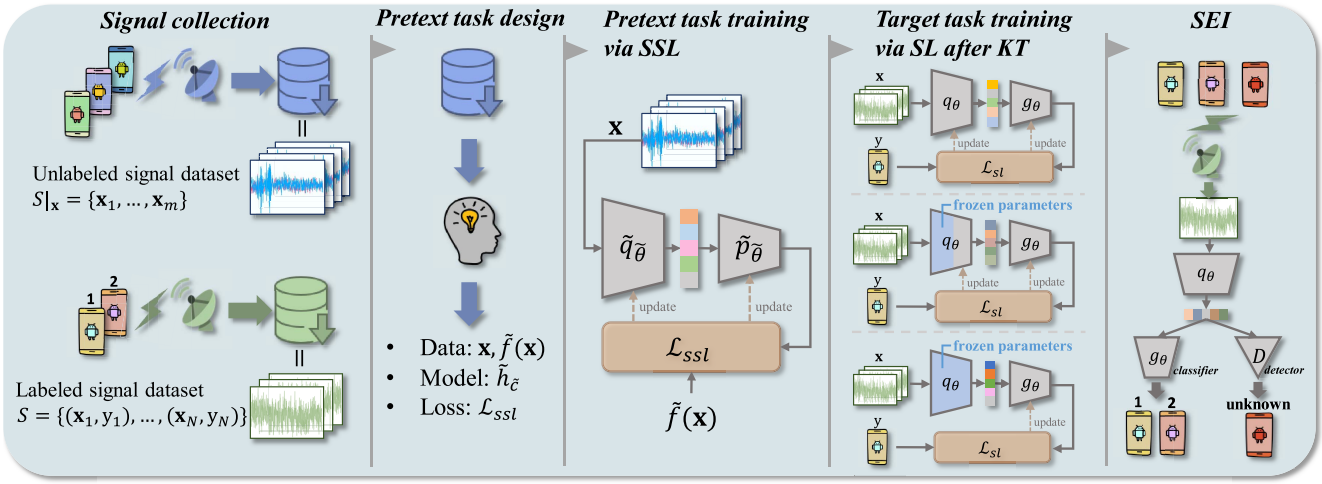


Fig. 8. The self-supervised workflow starts with an unlabeled source data set and a labeled target data set.

details about the type or environment of the collected signals. For dataset [60], it provides Bluetooth signals sampled at different sample rates, which can be used to study the impact of variable sampling rates on SEI performance and further provide solutions, offering flexibility for the input of the SEI system. Similarly, dataset [61] collects WiFi signals from different transmitter-receiver distances and can be used to study the effect of distance on SEI. Dataset [73] contains 150GB of WiFi transmissions of 15 Pycom devices captured over 3 consecutive days in both wired and wireless connections and across 4 different locations, providing rich signal acquisition scenarios and essential data support for the robust SEI studies in complex electromagnetic environments. We hope this table can assist researchers in the field in quickly obtaining effective, rich, and high-quality open-source datasets, thereby accelerating the research progress in the SEI field.

B. Generative Approaches

1) Motivation: Recently, Generative Artificial Intelligence (GAI) has emerged as a significant research topic in network optimization [77], physical layer communication security [78], [79], space-air-ground integrated networks [80] and semantic communications [81], [82]. In the SEI domain, to provide appropriate pretext tasks to assist models in understanding and capturing the intrinsic information of data, the most direct approach is to design self-supervised tasks based on common tasks, such as dimensionality reduction, denoising, and repair. These concepts, serving as motivation, have led to the development of GSSL, which aims to perform transformation and generation of samples at the sample point level according to task requirements.

2) Autoencoding-Based Methods: Autoencoder (AE), one of the basic schemes of GSSL, was first introduced in [83] for pre-training neural networks. In 2006, Hinton and Salakhutdinov [84] introduced a nonlinear, adaptive multi-layer encoder to transform high-dimensional data into low-dimensional encodings, and a corresponding decoder restores the low-dimensional encoding back to the original high-dimensional data, forming an entire AE framework. Then AE became a popular dimensionality reduction scheme, and multiple variants emerged to be applied to SSL. Unlike AE, in which both the pseudo label and input are original samples, the inputs of these variants are distorted samples. For example, Denoising Autoencoder (DAE) [85] takes noisy samples as input to the encoder and trains the decoder to generate noise-free samples, thereby compelling the encoder to learn robust features. In the CV field, there are some similar methods, such as generating color images from black-white images [86] and generating high resolution images from low resolution images [87]. However, it is difficult to apply these methods to other fields. In 2022, Masked Autoencoder (MAE) [88] was proposed, which adds masks at random positions on the input samples and recovers the masked content through the encoder-decoder structure, demonstrating robust performance. Then Shi et al. [89] built an MAE for self-supervised reconstruction tasks using unlabeled signal samples transformed into images by Gramian Angular Field (GAF), and the GAF-MAE outperforms other well-known models on the public dataset RadioML.2016.10a for Automatic Modulation Classification (AMC) task.

Inspired by [88], Huang et al. [90] proposed an unsupervised pre-training method based on MAE for RFF learning. At the time dimension of IQ signals, K. Huang designed three types of masking patterns: block-wise masking, which removes a consecutive segment; grid-wise masking, which removes a segment of a certain length periodically; and random masking, which removes random segments. At the channel dimension, K. Huang designed two types of masking patterns: channel-aligned masking, which masks I and Q signals at the same positions, and channel-separated masking, which masks I and Q signals independently. The experimental results show that a block-wise channel-aligned masking pattern best suits RFF learning. Based on the above work, Yao et al. [91] proposed an Asymmetric Masked Auto-Encoder (AMAE) to pretrain the RFF extractor, which is finetuned in the downstream SEI task with few-shot samples. Specifically, in AMAE, the decoder and encoder structures are asymmetric, and its performance on

TABLE IV
THE OPEN-SOURCE DATASET FOR SEI

Reference	Protocol/device	Frequency	Sample rate	Categories	Access	Details
[54]	USRP N2932 SDR	433MHz	5MS/s	21	link	different signal type and transmission complexity
[55]	GNSS	Multiple	50MS/s	6	link	from sky satellites and Spectracom simulator
[56]	ADS-B	1090MHz	10MS/s	100+	link	can be used for incremental learning
[57]	ADS-B	1090MHz	50MS/s	100+	link	long signal format, short signal format
[58]	Drones RF activity	40MHz	2.4GS/s	3	link	include background and different flight modes
[59]	Drones remote controllers	2.4GHz	20GS/s	17	link	from eight different manufacturers
[60]	Bluetooth from smartphone	2.4GHz	5-20GS/s	27	link	include signals under different sample rates
[61]	WiFi from USRP X310	2.4GHz	5MS/s	16	link	include different transmitter-receiver distance
[62]	DJI M100 UAV	900MHz	5MS/s	4	link	contain custom RFFs
[63]	USRP X310	2.685GHz	5MS/s, 7.68MS/s	4	link	collected under different days
[64]	DJI M100 drones	2.4GHz	10MS/s	7	link	include different transmitter-receiver distance
[65]	ZigBee	2505MHz	10MS/s	60	link	indoor/outdoor, LOS/NLOS
[66]	NI N210 and NI X310 SDRs	2.432GHz	20MS/s	20	link	wild/anechoic chamber, wired/wireless, wireless with different antennas/same antenna
[67]	VSG60A with different PAs	2.4GHz	10MS/s	9	link	different PAs with 8 modulation types
[68]	WiFi modules	2462MHz	25MS/s	174	link	41 receivers over 4 captures spanning a month
[69]	Bluetooth and WiFi emitter	2.414GHz	66.67MS/s	10	link	different transmitter-receiver distance and days
[70]	Bluetooth from IoT emitters	2.414GHz	2MS/s	10	link	different transmitter-receiver distance and days
[47]	LoRa devices	868.1MHz	1MS/s	60	link	3 rooms, 6 locations, stationary/moving
[71]	LoRa from Pycom devices	915MHz	1MS/s	25	link	wired/wireless, indoor/outdoor, 5 days, 4 distances, 3 locations, 2 receivers
[72]	WiFi from Pycom devices	2.412GHz	25MS/s	50	link	indoor/outdoor, 5 days
[73]	WiFi from Pycom devices	2.412GHz	45MS/s	15	link	wired/wireless, 4 locations, 3 days
[74]	XSRP	800MHz	–	10	link	LOS/NLOS
[75]	LTE devices	–	30.72MS/s	10	link	different locations, routes and days

Readers can select the appropriate open-source dataset according to your own task scenarios and quickly reach the download address by clicking on the "link". Please refer to the paper in the "Reference" or the description of the data in the download address for the detailed use of each data.

downstream few-shot SEI tasks significantly surpasses that of MAE for the open-source datasets [61] and [46]. Zhu et al. [92] combined stacked convolutional DAE and contrastive representation learning as a multi-task SSL framework for vehicle classification based on carrier-free ultrawideband radar signal.

3) GAN-Based Methods: In 2014, the introduction of the Generative Adversarial Network (GAN) [93] sparked a revolution in the field of DL, garnering acclaim from both the academic and industrial sectors and leading to numerous technological breakthroughs. A GAN is composed of two parts: a generator that creates data samples, and a discriminator that distinguishes between genuine and generated samples, improving the quality of the generated data through this adversarial process. For the generator in GAN, its input is noise, and the pseudo label is the distribution of the original samples. In particular, the similarity between the sample distribution generated by the generator and the original sample distribution is measured by the discriminator. With the development of GAN, improved versions have subsequently been developed and applied to SSL. Conditional Generative Adversarial Networks (CGAN) [94] modify the traditional GAN framework by incorporating label information to condition the generation process, enabling the creation of targeted data samples. Auxiliary Classifier GANs (ACGAN) [95] extend CGAN by adding an auxiliary classifier on top of the discriminator to further guide the generation process with more explicit class label information, enhancing the control over the categories of the generated data.

Since GANs do not require explicit probabilistic assumptions, they perform well on complex signal data distributions and have a better ability to generate flexible, varied, and high-quality signal data than other models. There have been many applications of GAN, and its versions have been improved for SEI. Gong et al. [96] proposed an unsupervised SEI

framework in non-cooperative scenarios, which combines the advantages of information-maximized GAN (InfoGAN), manual RFFs, and the knowledge of the wireless propagation channel. Specifically, the gray histogram is embedded into the proposed framework to enhance individual discriminability, and the InfoGAN integrates a structured multimodal latent vector that incorporates the priori statistical characteristics of wireless propagation channels. To overcome the degradation of RFFs due to multipath fading and enhance the performance of SEI, Fadul et al. [97] proposed a channel equalization approach, which leverages a CGAN to learn the distribution of multipath channel effects, estimate and correct their impact on RFF. The performance is evaluated in a Rayleigh fading channel environment with degrading SNR, involving up to 32 IoT devices.

C. Predictive Approaches

1) Motivation: Although the pretext tasks and pseudo labels in GSSL are easy to understand and design, its pixel-level, fine-grained generative framework presents high training complexity and instability. Subsequently, some methods choose or design certain properties of the samples as pseudo-labels, resulting in PSSL. These properties include the state of the samples, component scales, spatial structure, and temporal association, among others. Based on these pseudo-labels, which are similar to class labels in SL, PSSL can be trained with a less complex loss function, such as cross-entropy loss [53], similar to SL.

2) Approaches: Instance discrimination [98] is one of the classic PSSL schemes, which treats each instance as its own class and trains the model to discriminate between different instances, hence the problem becomes a supervised classification problem in which the number of categories is the same as the size of the dataset. Besides, some studies split one sample into multiple patches, use the spatial context properties between patches as pseudo-labels, and predict these pseudo-labels through a neural network as the pretext task for SSL. For example, Doersch et al. [99] randomly selects two patches from each image to form a pair and trains a CNN to predict the position of the second patch relative to the first. In this setup, the two patches are inputs, and their relative position serves as the pseudo labels. Since the relative positions are finite, this task can be regarded as a supervised closed-set classification task during training. Similarly, Noroozi and Favaro [100] used a jigsaw puzzle game as the pretext task for SSL, dividing an image into nine patches, shuffling them according to a specific strategy, and then learning both a feature mapping of object parts as well as their correct spatial arrangement. Misra et al. [101] proposed a sequential verification task for SSL by extracting a tuple of frames from a video, and asking whether the frames are in the correct temporal order.

Inspired by the rotation prediction-based SSL [102], which trains deep neural networks to recognize geometric transformations applied to images and learns to represent the images effectively through the identification of rotations, Xu et al. [103] proposed a few-shot SEI method using rotation feature decoupling (RFD) as the pretext task. For an RF signal $\mathbf{x} = (I, Q)$, I and Q are two channels of $\mathbf{x}$, corresponding to the real and imaginary parts, respectively. The input of RFD is the rotated $\mathbf{x}$ based on phase, which can be represented as

$$\mathbf{x}^r = \begin{bmatrix} \cos\ \sigma & -\sin\ \sigma \\ \sin\ \sigma & \cos\ \sigma \end{bmatrix} \begin{bmatrix} I \\ Q \end{bmatrix}, \tag{9}$$

where σ is the angle of rotation, chosen from $\{0, \pi/2, \pi, 3\pi/2\}$. Xu et al. decomposed the feature of $\mathbf{x}^r$ into two parts: rotation-relevant features and rotation-irrelevant features. Consequently, the rotation-relevant features can be classified into four classes corresponding to the rotation angles by a rotation classifier. Rotation-irrelevant features are obtained by reducing the inter-feature differences of samples with various rotation angles. To prevent a collapsed solution, where the encoder outputs the same rotation-irrelevant feature for any input sample, Xu et al. introduced an instance classification task into the optimization process of rotation-irrelevant features. In the downstream few-shot SEI task, the encoder trained by RFD was finetuned on the open-source Wi-Fi dataset [61] and the simulation results showed up to 92.90% identification accuracy with only 30 training samples per category, exceeding the benchmark performance without pre-training by 25%. Based on the work, Xu et al. [104] proposed a new few-shot SEI approach using phase shift prediction and decoupling, which is an alternative to [103] by obtaining the time-frequency features of the phase-shifted IQ signal through the Short-Time Fourier Transform.

D. Contrastive Approaches

1) Motivation: The aforementioned instance discrimination framework provides a powerful approach for SSL, enabling models to learn complex feature representations without explicit labels by distinguishing between different instances. However, since the number of categories equals the number of samples, the softmax operation used to compute class probabilities becomes increasingly costly as the size of the dataset grows. It can only be estimated by other methods such as the Noise-Contrastive Estimation (NCE) [105] as used in [103]. Another problem within the instance discrimination framework is the lack of intraclass variability, which does not enable the model to learn the inherent variation within each class [106]. To address the problem, the CSSL scheme has emerged, synergizing extensive data augmentation and instance discrimination. In CSSL, approaches based on negative examples and those not based on negative examples are two main strategies.

2) Negative Sample-Based Methods: Contrastive learning involves choosing an anchor, a positive sample (same class as the anchor), and a negative sample (different class from the anchor), with the goal of minimizing the distance between the anchor and the positive sample and maximizing the distance between the anchor and the negative sample. However, positive and negative samples cannot be chosen directly in an SSL task because the corresponding labels are not accessible. To address this problem, CSSL generates distinct views of an instance through data augmentation. In early negative example-based CSSL, views from the same instance are regarded as positive examples for an anchor sample,

while views from different instances within the batch are considered negative examples. The optimization direction of these methods involves increasing view difficulty and adding more negative examples. However, obtaining more negative examples requires a larger batch size, which is limited by GPU memory constraints. One possible approach is to add a memory bank [107] to store previously encoded samples and use them as negative examples when calculating loss. For example, Contrastive Multiview Coding (CMC) [108] introduced a memory bank to CSSL in the CV domain, aiming to maximize mutual information between different views of the same scene. Nevertheless, memory bank-based schemes face the problem of inconsistency between stored encodings, calculated by a previous encoder, and updated encodings, affecting objective optimization.

He et al. [109] introduced the Momentum Contrast (MoCo) model at the end of 2019, which continues the idea of a memory bank and uses momentum to update the encoder parameters, solving the problem of inconsistent encoding of old and new candidate samples. Davaslioglu et al. [110] applied a MoCo-v3 [111] SSL framework to motivate RF signal representation learning for AMC, utilizing five RF waveform augmentations: DC shift, time shift, amplitude scale, zero-masking, and Gaussian noise. Chen proposed a Simple Framework for Contrastive Learning (SimCLR) [112], which does not use the memory bank but improves augmentation methods and adds a nonlinear projection following the encoder. The nonlinear projection removes the augmentation-related information within the extracted features from the encoder. Liu et al. [113] applied CSSL to a semi-supervised AMC task, using unlabeled data with the SimCLR to train an intrinsic feature extractor and a small amount of labeled data to train a modulation classifier. Liu et al. [114] considered the frequency domain characteristics of the signals and improved the SimCLR based on time-domain and frequency-domain data augmentation, including low-pass filtering, amplitude perturbation, phase perturbation, and phase shift, achieving high accuracy on human activity recognition tasks. Bocus et al. [115] integrated SimCLR and CMC to improve activity recognition performance using WiFi signals captured by synchronized receivers deployed in different positions. In addition, there are some other negative example-based CSSL methods include Contrastive Predictive Coding (CPC) [116] and Pretext-Invariant Representation Learning (PIRL) [117].

Most CSSL approaches for SEI are inspired by the above schemes. For example, Zhao et al. [118] combined the complex-valued neural network with the MoCo framework for SEI, employing phase rotation, random cropping, and jitter as data augmentations. Liu et al. [119] converted emitter signals to constellation figures as inputs for SimCLR and used CV-based data augmentations, including random color distortion, horizontal flip and vertical flip. Wu et al. [120] employed SimCLR in a semi-supervised SEI problem, where SimCLR works as an unsupervised contrastive learning model to train a feature extractor on unlabeled signals, followed by cross-entropy and supervised contrastive learning to train an emitter classifier. Given the success of CSSL, Hao et al. [121] proposed a Signal Constrastive Self-Supervised Clustering (SCSC) algorithm for SEI, using a Bit-Pulse Selection (BPS) strategy and several signal data augmentation methods to obtain positive and negative instance pairs. This method extracts hierarchical semantic features via a designed 1-D fingerprint pyramid feature extractor that combines instance-level and emitter-level contrastive learning to maximize agreements. Specifically, the data augmentation methods for BPS include segment switching, amplitude jitter, time skew, and random noise.

3) Positive Sample-Based Methods: Although the diversity and number of negative samples can significantly improve the generalization ability of CSSL, especially on large-scale datasets, certain scenarios make it difficult to obtain a large number of negative samples and consume vast computational resources. Additionally, the selection of negative samples is highly dependent on the particular task and data type, potentially leading to poor learning outcomes if chosen improperly. To address these challenges, Caron et al. [122] proposed Swapping Assignments Between Multiple Views of the Same Image (SwAV), which integrates online clustering and multi-view data augmentation into CSSL without the need for negative examples. SwAV utilizes cluster representations as a proxy to negate the need for traditional negative samples, offering a more computationally efficient and broadly capturing SSL method. Similarly, Bootstrap Your Own Latent (BYOL) [123] eliminated negative examples and proposed an asymmetrical architecture with a momentum update strategy for the target encoder, similar to MoCo. Chen and He [124] demonstrated that the most critical component in BYOL is the stop gradient operation. Based on this, they proposed an effective CSSL framework, Simple Siamese (SimSiam), which does not require negative example pairs, large batches, or momentum encoders. Similar to Barlow twins [125], Variance-Invariance-Covariance Regularization (VicReg) [126] is a joint embedding architecture with variance, invariance and covariance regularization. In this method, the distance between two embeddings from the same sample is minimized, the variance of each embedding variable over a batch is maintained above a threshold, and the covariance between pairs of embedding variables over a batch is driven to zero.

Building on the work [118] and BYOL, Zhao et al. [127] proposed an improved BYOL scheme for SEI that incorporates three data augmentation methods: rotation, random cropping and jitter. The simulation results on downstream SEI tasks show that it significantly outperforms MoCo across various sample sizes and Signal-to-Noise Ratios (SNR). Zha et al. [128] considered the impacts of receivers on RFFs, viewing this as a data augmentation method, and utilized the SimSiam network to learn and eliminate receiver-agnostic features. Subsequently, they optimized the model with Local Maximum Mean Discrepancy (LMMD) regularization to prevent distribution shifting. Inspired by VAT [129], Liu et al. [130] proposed an SEI framework based on SSL and Adversarial Augmentation (SA2SEI), which integrates data-driven augmentations such as rotation and flipping, as well as model-driven augmentation, specifically feature adversarial augmentation. This framework achieved high accuracy in the downstream few-shot SEI task.

TABLE V
SUMMARY OF WORK FOR SSL-BASED SEI

SSL paradigm	Reference	Algorithm	Pros & Cons
GSSL	[96]	InfoGAN	● An unsupervised DL SEI framework using RFF embedding, to perform identification in noncooperative scenarios. ☺ Incorporate a priori statistical characteristics of the wireless propagation channels to further improve the GAN quality. ☹ Need to estimate suitable hyperparameters for the chosen Nakagami-m distribution.
	[90]	MAE	● An unsupervised pre-training method based on MAE for RFF learning. ☺ Improve generalization capability of learned RFFs. ☹ ResNet used in this letter may not be the most appropriate architecture for MAE pre-training due to its weakness of learning long-range dependencies.
	[91]	MAE	● A few-shot SEI using asymmetric MAE to solve the few-shot problem. ☺ Better feature extraction performance, and stronger generalization ability in the target domain. ☹ Require relatively longer training time than normal MAE.
	[97]	CGAN	● A work presents and analyzes two DL approaches to enhance IoT security using SEI under Rayleigh fading and degrading SNR. ☺ Combine label embedding with CGAN to efficiently learn each emitter's conditional feature distribution. ☹ CGAN shows weak scalability due to the number of preamble reconstructions and classification decisions increasing linearly with the number of emitters.
PSSL	[103]	Rotation-Prediction	● A few-shot SEI method using Rotation Feature Decoupling using an unlabeled auxiliary dataset. ☺ Guide the encoder to learn decoupled rotation features from auxiliary dataset. ☹ Limited performance advantage with large sample sizes.
	[104]	Rotation-Prediction	● An advanced few-shot SEI approach using Phase Shift Prediction and Decoupling. ☺ The use of Short-Time Fourier Transform further enhances performance. ☹ Relatively limited robustness on sample quality and noise.
CSSL	[118]	MoCo	● A complex-valued SSL-based method for few-shot SEI. ☺ Combine the complex-valued neural network with SSL and propose three data augmentation methods based on communication signals: phase rotation, random cropping, and jitter. ☹ More parameters and longer training time.
	[119]	SimCLR	● An SEI method based on self-supervised contrast learning. ☺ Choose constellation trace figures as the input data, which suppresses the interference of the symbol information. ☹ Relatively limited performance advantage at low SNR.
	[127]	BYOL	● An SEI model based on improved BYOL SSL. ☺ Reduce the computational resources by removing negative samples. ☹ The data enhancement approach is relatively simple.
	[121]	SCSC	● A signal contrastive self-supervised clustering method for unsupervised SEI applications. ☺ Propose a BPS strategy and several signal data augmentation methods, and SCSC includes an unknown class rejection strategy to adapt to open environments. ☹ Need suitable signal data enhancement methods and their combinations on different datasets and tasks.
	[120]	SimCLR	● A contrastive learning identifier using two-stage semi-supervised training scheme to identify the unknown emitter when the labeled data is insufficient. ☺ Combine SSL and contrastive learning to improve the discrimination of the features. ☹ The data enhancement approach is relatively simple.
	[130]	BYOL	● A few-shot SEI method based on SSL and adversarial augmentation. ☺ The adversarial augmentation can obtain flexible and dynamic data transformation on the feature space to enable SSL process. ☹ Relatively limited robustness on sample quality.
	[128]	SimSiam	● A novel scheme for cross-receiver RFF identification. ☺ Combine contrastive learning and local maximum mean discrepancy regularization to capture the fine-grained RFF information while maintaining the distribution alignment. ☹ Relatively high computational complexity.

Tips: ● describe the methods. ☺ and ☹ represent pros and cons, respectively.

In this section, we comprehensively collect available open-source SEI datasets and list the related work based on the SSL and SEI according to the three paradigms discussed in the previous section. The comparison of different signal recognition algorithms based on SSL is summarized in Table V.

V. AN OUTLOOK: SSL & LARGE MODEL

As SSL continues to evolve, its integration with large models presents a promising direction for future research and applications, especially in the field of SEI, where the abundance of device types, the large number of devices, the variety of signal styles, and the complexity of the electromagnetic environment are critical factors that cannot be ignored. While SSL on small models and small datasets is beneficial for specific small tasks, it faces several limitations. Small models with fewer parameters have limited expressive power and struggle to capture complex patterns in signals. Additionally, small datasets lack diversity and coverage, leading to overfitting and poor generalization, preventing the model from performing well across different tasks and scenarios.

In contrast, combining SSL with large models and large datasets offers significant advantages. Large models, with their extensive parameters and deep structures, possess strong

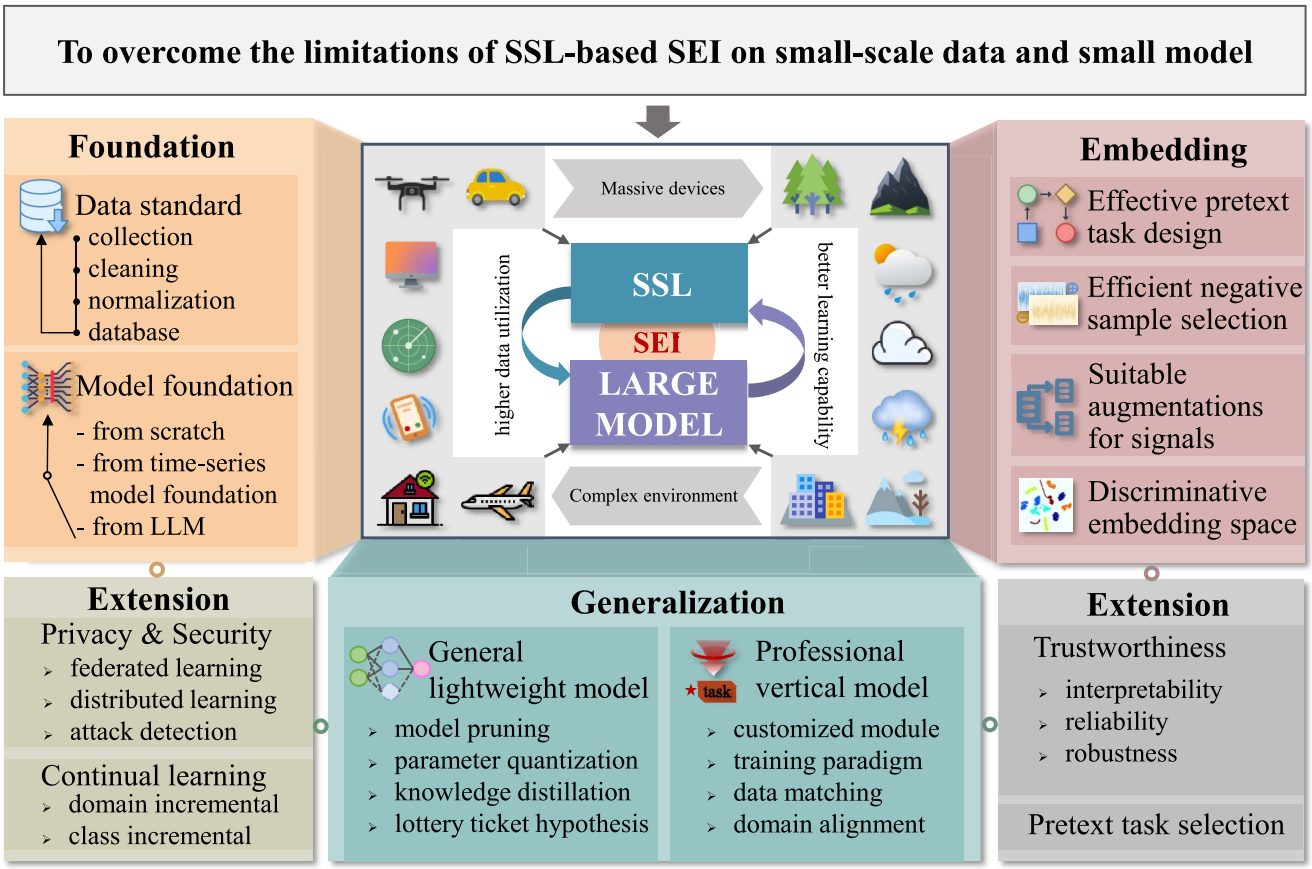


Fig. 9. The future work of SSL-based SEI combined with large models.

feature extraction and pattern recognition capabilities, enabling a better understanding and handling of complex data. Large datasets provide rich diversity and coverage, allowing the model to learn a wide range of features during the pre-training phase, thus enhancing generalization and adaptability. This combination not only enables efficient learning from unlabeled data but also excels in subsequent supervised tasks, reducing the reliance on large amounts of labeled data.

In this section, we will explore potential problems and future works of combining SSL with large models, focusing on three key future directions and two extensions, as shown in Fig. 9.

A. Foundation Construction

RFF features are simultaneously affected by a variety of factors, such as operating parameters, receiver, transceiver distance, reception time, and noise. Most existing works consider signal variations under only a few such conditions. However, training a large model foundation requires vast signal data encompassing a wide range of cross-domain variations. This calls for a standardized dataset that covers diverse dimensions, including cross-receiver [133], cross-time [134], cross-channel condition [46], cross-SNR [135], cross-protocol configuration [136], cross-modulation [40], cross-distance [137], and cross-sample rate [60]. Based on such a standard, signal preprocessing procedures, such as signal collection, cleaning, and normalization, should be employed to construct a robust and scalable signal database suitable for large-model training.

Once the standardized signal database is built, a large foundation model can be trained from scratch. This approach provides maximum flexibility and customization but is computationally expensive. An alternative is to fine-tune an existing time-series foundation model, such as TimeGPT [138], Lag-llama [139], or MOMENT [140]), which can significantly reduce training costs while leveraging strong temporal representation capabilities. Although the data used to train these models differs from RF signals, their excellent time-series analysis and feature representation capabilities can still be beneficial for SEI tasks. However, the availability of open-source time series models is limited, whereas Large Language Models (LLM) are comparatively more mature. Therefore, leveraging some conversion or alignment methods [141], [142], [143], [144] between time-series data and text data is also a promising option, as it can address the numerical characteristic differences between LLM text data and RF signal data, enabling the construction of a large model foundation for SEI using a LLM foundation.

While most research on large models for SEI focuses on IoT devices [145], the underlying idea of learning device-specific fingerprints using SSL can be extended to a broader range of wireless systems. In particular, foundation models trained on IoT transmissions may generalize to cellular base stations, UAV communication links, and satellite networks.

Although these systems differ in hardware configurations, channel characteristics, and signal protocols, the core objective of identifying device-specific RFF features remains consistent. Domain adaptation and fine-tuning techniques can help adjust the model to account for such domain shifts. Investigating the cross-domain generalizability of foundation models is therefore a valuable future direction to enhance the scalability and applicability of SSL-based SEI across diverse wireless platforms.

B. Embedding Design

Large models have shown tremendous potential in various tasks, while the successful training of these models on vast amounts of signal data hinges on the quality of feature embeddings, especially through SSL, where no real labels guide the training objectives. Feature embeddings transform raw signal data into more informative representations, which are crucial for capturing the underlying patterns, structures, and RFF within the data. Inadequate feature embedding methods can lead to suboptimal model performance, as the model may struggle to extract meaningful representations from complex and high-dimensional signal data. Therefore, developing and employing high-quality feature embedding schemes according to the characteristics of signal data is paramount for leveraging the full potential of large models.

Effective pretext tasks are the basis of SSL, guiding the model to learn meaningful representations from unlabeled data. The three SSL paradigms summarized in this paper can serve as references for the selection of pretext tasks, or innovative pretext tasks can be designed according to the structural characteristics of large models and the diversity of large datasets, to obtain richer feature embeddings. Additionally, key research areas for combining SSL with large models in the future include how to construct positive and negative sample pairs to avoid collapse during training, how to develop augmentation methods suitable for signal data to prevent the destruction of RFF information, and how to construct discriminative embedding spaces to learn the associations and distinctions between features.

C. Downstream Generalization

As we look to the future of combining SSL with large models, enhancing their downstream universality is a crucial area of focus. Due to the substantial size and computational demands of large models, many devices are unable to support their deployment. To address this, various model compression techniques should be employed to create general lightweight models that maintain high performance while being more resource-efficient. In addition, while large models perform well on many tasks, there are specialized use cases, such as specific types of emitters or unique signal types and rare channel environments, where further customization is needed.

Compressing large models through pruning [146], quantization [147], knowledge distillation [148], and lottery ticket hypothesis [149] is an important research direction in the future, which can create efficient, general lightweight models suitable for a wide range of devices. Simultaneously, developing specialized vertical models by adding customized modules, choosing appropriate training paradigms, matching business data with large databases, and aligning the domains between large models and vertical models ensures that the model can effectively handle unique and specialized tasks.

D. Extension Work

The aforementioned foundation and generalization directions lead to some extended work that needs exploration. Firstly, the development and application of large models must prioritize security and controllability, particularly regarding data privacy and data security. In terms of privacy, building large signal databases involves collecting vast amounts of signal data, which may be distributed across different locations and could contain sensitive information. Centralizing all data in a single data center could lead to serious privacy leakage issues. To address this, integrating large models with federated learning [150] and distributed learning [151] could be an effective solution. In terms of security, given the vulnerability of large model training data to attacks [152] or poisoning [153], SEI large models need capabilities for attack detection and defense to safeguard the downstream fine-tuning and deployment of large models. Secondly, as signal collection continues and the types of emitters and scenarios change, an important future direction is enabling models to learn new knowledge without forgetting old information, to avoid catastrophic forgetting and achieve continual learning [154]. This can be pursued primarily through domain incremental learning and class incremental learning [155], [156].

The aforementioned embedding and generalization directions also lead to some extended work that needs exploration. Firstly, the trustworthiness assessment of SEI large models is critical to their widespread deployment and application. Trustworthiness assessment should encompass the evaluation of interpretability, reliability, and robustness of SEI large models. This involves comprehensively analyzing diverse measurement scenarios, establishing a system of trustworthiness metrics, and developing a trustworthiness assessment platform. Secondly, an essential gap exists between pretext and downstream tasks. Researchers create sophisticated pretext tasks to assist models in learning crucial dataset features that can be transferred to other tasks. However, this approach occasionally fails to achieve its goals. Additionally, the process of selecting pretext tasks often appears overly heuristic and complex, lacking consistent patterns to follow [157]. Therefore, for a specific downstream task, it is important to be clear about the properties desired in the embedding representations and to select effective pretext task base on those properties [106].

In this section, we focus on the combination of SSL and large models in the future of SEI, which is expected to significantly enhance system accuracy and robustness. SSL reduces reliance on labeled data by extracting features from unlabeled data, while large models excel at processing complex time-series signals. This integration is particularly crucial in areas such as secure communications, where it can enable more efficient device identification, network security monitoring,

TABLE VI
THE SIMULATION PARAMETERS

Simulation parameters	Pretext task		Downstream task	
Dimension of samples	Wi-Fi	2×4800	Wi-Fi	2×4800
	LoRa	2×8192	LoRa	2×8192
Number of categories	Wi-Fi	10	Wi-Fi	6
	LoRa	20	LoRa	10
Number of sample per category	Wi-Fi	2000	Wi-Fi	20
	LoRa	500	LoRa	20
Ratio between training and validation samples	8:2		1:0	
Batch size	128		64	
Maximun epoches	300		100	
Learning rate	0.001		0∼0.001	

TABLE VII
THE NETWORK STRUCTURES IN SSL METHODS

Network	Structure	Number of layers
Encoder	Complex Conv1D + ReLU + BatchNorm1D + MaxPool1D	× 9
	Flatten + LazyLinear(1024)	×1
Projector	LazyLinear(512) + ReLU	×1
	LazyLinear(128)	×1
Predictor	LazyLinear(512) + ReLU	×1
	LazyLinear(128)	×1
Decoder	LazyLinear(2*N) + reshape	×1
Angle Classifier	LazyLinear(4)	×1

and privacy protection, advancing SEI technology in dynamic and complex wireless environments.

VI. CASE STUDY

In this section, to compare the RFF extraction capabilities of different SSL methods, we construct various downstream SEI tasks for several classic and latest SSL methods to present and analyze the identification performance under different SNRs and using different downstream classifiers. The features of the downstream dataset will be extracted to evaluate the RFF quality. The goal of this section is to provide a comprehensive evaluation of how these SSL methods perform in diverse scenarios, shedding light on their robustness and generalization capabilities, and to help readers quickly choose the suitable approach to solve the target task.

A. Dataset, Simulation Environment, and Parameters

The open source Wi-Fi dataset [61] and LoRa dataset [46] are selected from Table IV due to their public availability and widespread use in existing SEI literature. They cover distinct signal characteristics and deployment scenarios, making them suitable benchmarks for evaluating SSL-based SEI performance. The DL models in these methods are based on PyTorch, and the simulation platform is GTX 3090Ti. To highlight the gain of SSL in downstream tasks, we randomly selected 30 groups of different samples as $S|_{\mathbf{x}}$ for Monte Carlo simulations, and each simulation is tested on the same test dataset. The average result of simulations is applied as the performance indicator of these methods. The detailed simulation parameters are listed in Table VI.

B. Comparison Methods and Network Structures

According to the aforementioned three SSL paradigms, three of the latest SSL-based SEI methods, including AMAE [91], RFD [103] and SA2SEI [130] and several common SSL methods, including SimCLR, VicReg [126], SimSiam [128] and MoCo [118], are selected in this paper for evaluation. Specifically, the MoCo here represents MoCo-v2, whose encoder is followed by a projector such as SimCLR. For a fairer comparison, the required network structures of these SSL methods are standardized as shown in Table VII, where the Encoder is the indispensable feature extractor for all methods and is composed of nine complex convolution layers and one linear layer. The feature dimension is unified to 1024. The Projector is present in all methods except AMAE and RFD, while the Predictor is used in SA2SEI and SimSiam. The decoder is an important component in AMAE for recovering samples, and the angle classifier is used in RFD to predict the rotation angle of samples based on the extracted features. In addition, the methods except AMAE and RFD are contrastive SSL, where data augmentation plays an important role. In this section, the data augmentations for these contrastive SSL methods are unified to include rotation, flipping, and adversarial augmentation, similar to those used in SA2SEI.

C. Robustness Analysis: Under Different SNRs

In practical application scenarios, signals are subject to varying degrees of noise interference in the channel environment, which significantly impacts the SEI system. To analyze the noise robustness of these SSL methods, we first conduct self-supervised pre-training on noise-free pretext data to obtain the feature extractor (encoder). Then, noise with different SNR levels is added to the downstream dataset, and the features are extracted for classification. To ensure a fair and controlled evaluation of feature robustness under varying SNRs, we adopt a Support Vector Machine (SVM) with the linear kernel as the downstream classifier. This setting eliminates the influence of classifier complexity and focuses solely on the quality of the learned features. The simulation accuracy is shown in Fig. 10 and Fig. 11, where the gray line represents the result of traditional feature extraction and classification without applying SSL. Specifically, the signal is decomposed into multiple Intrinsic Mode Functions (IMF) by the Empirical Mode Decomposition (EMD) algorithm. The time-frequency distribution of the signal is obtained by using IMF reconstruction, and then the feature vector is derived by applying the Singular Value Decomposition (SVD) process.

From Fig. 10 and Fig. 11, it can be seen that AMAE, RFD, and SA2SEI exhibit good and stable performance on both downstream datasets compared to other methods, and their accuracies are close to each other in high SNR scenarios. VicReg shows a performance advantage under low SNR, with its accuracy even exceeding that of the three aforementioned methods on the Wi-Fi dataset and surpassing AMAE on

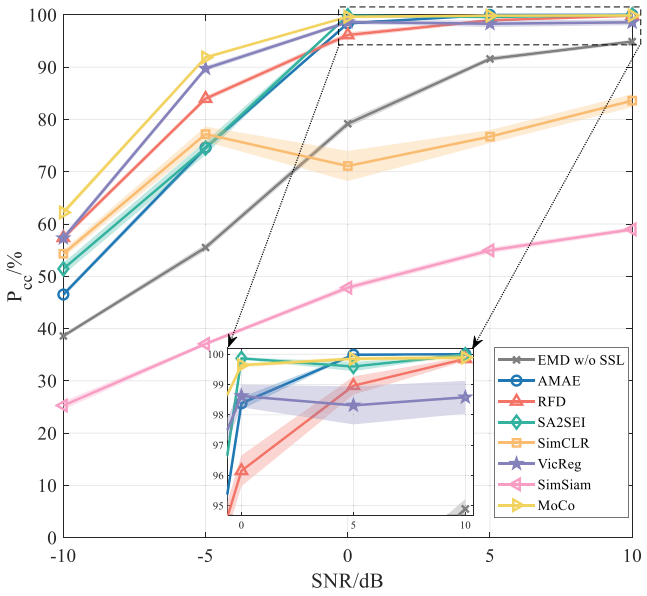


Fig. 10. The identification performance under different SNRs on the Wi-Fi dataset. The darker line shows the average accuracy of 30 simulations, with each simulation using different samples. The shaded areas represent the 95% confidence intervals, indicating the variability and reliability of the results.

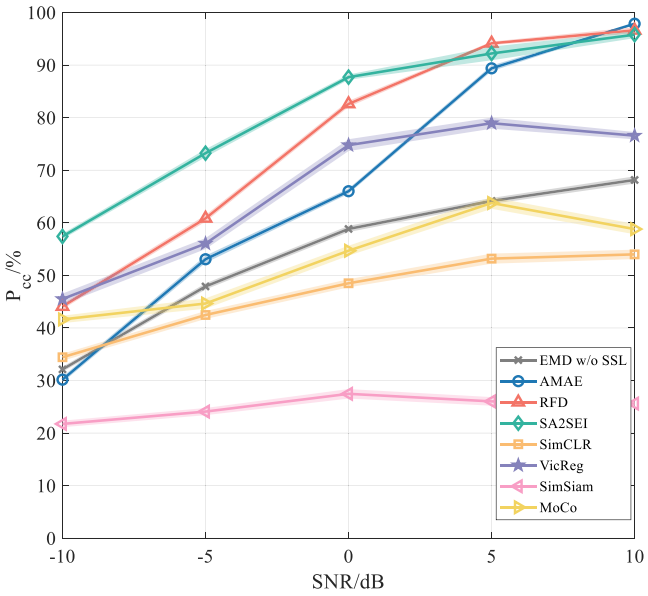


Fig. 11. The identification performance under different SNRs on the LoRa dataset. The darker line shows the average accuracy of 30 simulations, with each simulation using different samples. The shaded areas represent the 95% confidence intervals, indicating the variability and reliability of the results.

the LoRa dataset. However, as the SNR increases, VicReg's advantage gradually decreases; its accuracy plateaus, and its stability is not as good as the above three methods, as indicated by the confidence intervals. Due to the batch size limitation, SimCLR lacks sufficient negative samples and is unable to extract high-quality RFF features, performing even worse than traditional features extracted by EMD under high SNRs. Conversely, MoCo benefits from the queue for storing negative samples and performs best on the Wi-Fi dataset. However, on the smaller LoRa dataset, with only 500 samples per category in the pretext dataset, as shown in Table VI, MoCo performs poorly because of its insufficient negative sample sequence, slightly outperforming SimCLR. Although SimSiam is effective and prominent in the field of CV, it performs poorly on the two SEI datasets and is unable to extract effective RFF features.

D. Generalization Analysis: Using Different Classifiers

The SVM with a linear kernel used in the above robustness analysis is a common ML classifier. In addition to SVM, other common ML classifiers include the Nearest Neighbor Classifier (KNN), decision tree, random forest, Gaussian naive Bayes, and SVM with ploy and RBF kernels. These classifiers have their advantages in different data classification domains. To thoroughly analyze the SEI performance of different SSL methods in various downstream scenarios, this subsection evaluates the generalization performance of these SSL methods with different downstream ML classifiers and a DL classifier. The simulation results on Wi-Fi and LoRa datasets are shown in Fig. 12. The results w/o SSL for different machine learning classifiers indicate the performance of traditional features extracted by EMD, while the results w/o SSL for the DL classifier represent the performance of raw signal samples classified by a neural network, in which the feature extraction network has the same structure as the Encoder in Tab. VII, and the classifier is a fully-connected layer with the number of output units equal to the number of categories.

As shown in Fig. 12, in addition to SVM with a linear kernel, other ML and DL classifiers can also effectively classify the extracted features, demonstrating the generalization capability of these SSL methods. It can also be observed that the overall performance on the Wi-Fi dataset is higher than that on the LoRa dataset, especially when deep neural networks are used to classify the signal samples. This suggests that the quality of the RFFs in this Wi-Fi dataset is better than that in the LoRa dataset. Furthermore, the prior knowledge brought by SSL and KT can help classifiers improve their identification performance and avoid the overfitting problem caused by training from scratch.

E. Feature Quality Analysis: Clustering

As shown in Fig. 7, the features extracted by the feature extractor obtained through self-supervised training contain not only RFF information but also some redundant information, which affects the performance of SEI. In the supervised downstream tasks, the classifier is label-oriented and can decouple and remove this redundant information during the training process. In contrast, in unsupervised downstream tasks, the quality of the features directly affects the SEI performance. In this subsection, the features of the downstream dataset will be extracted and clustered to evaluate the feature quality. The evaluation metrics for the clustering results include external metrics such as the Adjusted Rand Index (ARI) [158] and Normalized Mutual Information (NMI) [159], which require the participation of real labels, and internal metrics such as the Silhouette Coefficient (SC) [160] and Davies-Bouldin (DB)

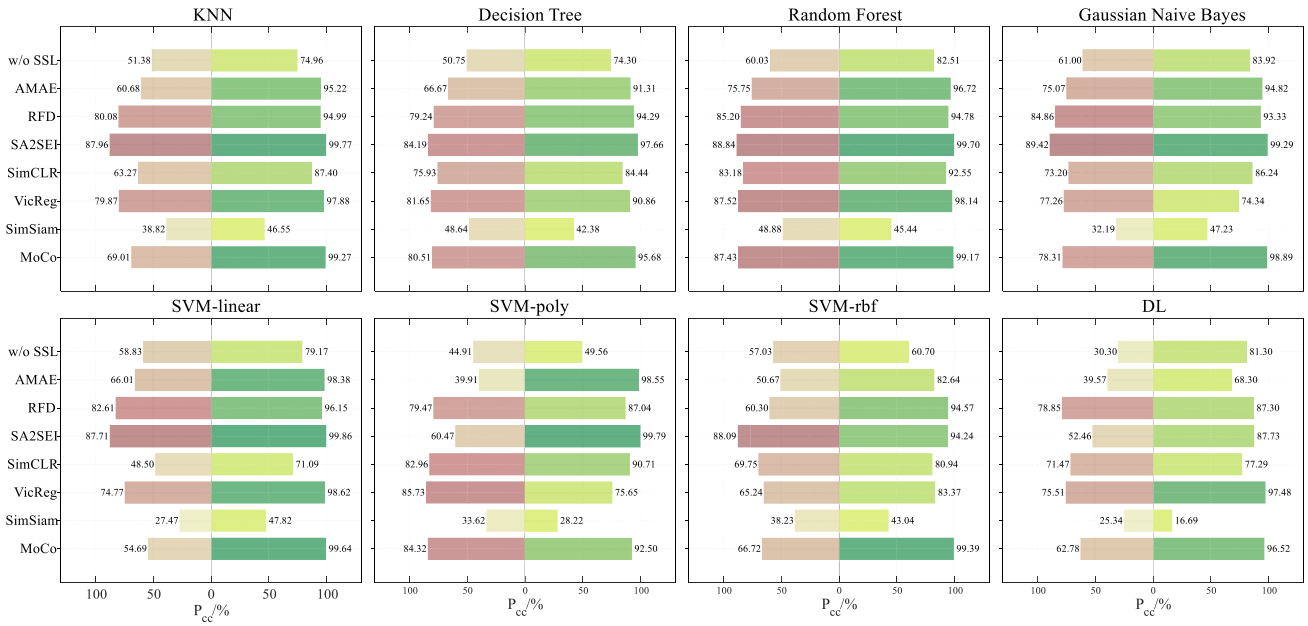


Fig. 12. The identification performance using different downstream classifiers under SNR=0. For each graph, the left pink bars and the right green bars represent the downstream identification accuracy on the LoRa dataset and the Wi-Fi dataset, respectively. Each result is the average accuracy of 30 simulations. Classifiers are implemented using the official Scikit-learn package, with $n\_estimators = 100$ for Random Forest and $n\_neighbors$ in KNN selected via cross-validation.

index [161], which do not require real labels. Specifically, ARI measures the similarity between two clustering results, and NMI quantifies the mutual dependence between the predicted and true cluster assignments. SC evaluates the cohesion and separation of clusters, indicating how similar an object is to its own cluster compared to other clusters. DB assesses cluster separation and compactness, with lower values indicating better clustering quality.

The clustering results are shown in Table VIII, suggesting that the features extracted by AMAE, RFD, and SA2SEI are of high quality and maintain effective clustering performance across different datasets. The high ARI and NMI values indicate that these methods can accurately reflect the true label distribution in the data. The strong performance in internal metrics indicates that the features extracted by these methods have good cohesion and separation within the cluster structure. In contrast, SimCLR and MoCo perform well on the LoRa dataset but lack consistency on the Wi-Fi dataset, particularly with higher DB values indicating poorer clustering quality on Wi-Fi. This inconsistency suggests that these methods might have variations in feature extraction quality across different datasets, leading to ineffective separation of data points.

In this section, we compare several SSL-based SEI methods based on two open-source datasets under the same simulation conditions. These methods are evaluated from three perspectives: robustness, generalization, and feature quality, respectively. From the results, it can be seen that the performance advantages of different methods vary on different datasets and evaluation metrics. For a specific task, a suitable method needs to be selected according to the attributes and conditions of the task.

TABLE VIII
THE CLUSTERING PERFORMANCE ON THE DOWNSTREAM DATASET

Methods	External Metric		Internal Metric	
	ARI	NMI	SC	DB
AMAE	**1.000**|0.777	**1.000**|0.902	0.834|0.740	0.231|0.459
RFD	0.870|0.779	0.953|0.903	0.860|0.806	0.221|0.313
SA2SEI	0.998|**1.000**	0.998|**1.000**	**0.869**|**0.899**	**0.176**|**0.137**
SimCLR	0.870|0.361	0.953|0.580	0.860|0.178	0.221|2.099
VicReg	0.295|0.380	0.541|0.581	0.347|0.338	1.226|1.263
SimSiam	0.151|0.367	0.341|0.536	0.368|0.530	0.893|0.591
MoCo	0.227|**1.000**	0.563|**1.000**	0.323|0.639	1.080|0.671

Note: The values to the left and right of "|" are clustering metrics on the LoRa and Wi-Fi datasets, respectively, and the best results are highlighted.

VII. DISCUSSION

The integration of SSL into SEI presents important implications for both research and practical deployment in wireless communication systems, while also introducing potential security risks such as model vulnerabilities and privacy leakage that must be carefully addressed.

A. Implications for Research

From a research perspective, SSL opens new avenues for tackling long-standing challenges in SEI, such as data scarcity, domain generalization, and robustness to noise or channel variations. Our review provides a state-of-the-art survey on the application of SSL on SEI tasks, performs a comprehensive collection of existing open-source datasets with download

links, and provides a case study for several advanced SSL-based SEI algorithms. The open-source datasets and algorithm codes will be useful for readers who want to conduct research quickly. Moreover, the modularity of SSL paradigms allows for flexible design of pretext tasks, which can be customized for various wireless environments.

B. Implications for Practice

In practical terms, SSL-based SEI models have the potential to significantly enhance the scalability and adaptability of wireless security systems. By reducing the dependency on labeled data, they make it feasible to deploy SEI in dynamic environments where manual labeling is infeasible. Applications include network access control, spectrum surveillance, anomaly detection, and PLA in IoT, UAV, and satellite systems. However, practical deployment also requires addressing challenges such as real-time inference efficiency, robustness to adversarial perturbations, and regulatory compliance.

C. Potential Risks

The advancement of SSL-based SEI also brings emerging challenges related to privacy and ethical concerns. For instance, powerful foundation models trained on large-scale wireless data may be vulnerable to model inversion attacks, where malicious entities infer sensitive characteristics of the original signals or devices. Additionally, SSL models could be susceptible to adversarial perturbations [162], [163], which might compromise the integrity of identification results or be exploited to evade detection. From an ethical standpoint, since SSL does not rely on labels, the use of SSL to pretraining without user consent may raise serious privacy implications, especially in civilian or commercial wireless networks. While SEI can contribute to network security and spectrum management, its deployment must consider legal and ethical boundaries. Future research should thus incorporate robust defenses against adversarial threats and establish clear governance mechanisms to ensure that SSL-based SEI systems are both technically effective and socially responsible.

VIII. Conclusion

DL-based SEI has become a promising PLA approach for enhancing IoT security. The integration of SSL offers a powerful means to reduce reliance on labeled data, making SEI techniques more scalable and adaptable in real-world, label-scarce scenarios.

This paper addresses the lack of dedicated reviews on SSL-based SEI and aims to provide a comprehensive and structured survey of the current state of the field. We provide a comparison of supervised, unsupervised, and SSL methods within the SEI context, and further divide SSL into generative, predictive, and contrastive paradigms. Representative works are categorized and analyzed accordingly. In addition, we have analyzed the challenges facing current SSL-based SEI research and discussed future research directions in conjunction with large models. We also offer practical support to researchers by listing open-source datasets, comparing SSL-based SEI methods, and releasing code to facilitate further development. Nevertheless, SSL paradigms are rapidly evolving, and many emerging techniques such as prompt-based learning or hybrid paradigms have not yet been systematically explored within the SEI domain. These limitations suggest the need for continuous updates to benchmarks and more extensive evaluations across diverse wireless environments. We hope this survey will inspire more attention to SSL in wireless system security and serve as a foundation for future research and innovation in this promising direction.

References

[1] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.

[2] D. C. Nguyen et al., "6G Internet of Things: A comprehensive survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, Jan. 2022.

[3] M. Vaezi et al., "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1117–1174, 2nd Quart., 2022.

[4] "Global IoT connections forecast to reach 40 billion in 2033." 2024. Transform Insights. [Online]. Available: https://transformainsights.com/news/iot-connections-40-billion-2033

[5] S. Arisdakessian, O. A. Wahab, A. Mourad, H. Otrok, and M. Guizani, "A survey on IoT intrusion detection: Federated learning, game theory, social psychology, and explainable AI as future directions," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4059–4092, Mar. 2023.

[6] 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., 2020.

[7] S. Shrivastava et al., "A survey on security issues in cognitive radio based cooperative sensing," *IET Commun.*, vol. 15, no. 7, pp. 1751–8628, Feb. 2021.

[8] P. Li, "Research on radar signal recognition based on automatic machine learning," *Neural. Comput. Appl.*, vol. 32, no. 7, pp. 1959–1969, Sep. 2019.

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

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

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

[12] L. Jing and Y. Tian, "Self-supervised visual feature learning with deep neural networks: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 11, pp. 4037–4058, Nov. 2021.

[13] A. Mohamed et al., "Self-supervised speech representation learning: A review," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 6, pp. 1179–1210, Oct. 2022.

[14] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 94–104, 1st Quart., 2015.

[15] 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.

[16] N. Soltanieh, Y. Norouzi, Y. Yang, and N. C. Karmakar, "A review of radio frequency fingerprinting techniques," *IEEE J. Radio Freq. Identif.*, vol. 4, no. 3, pp. 222–233, Sep. 2020.

[17] W. Wang, I. Sanchez, G. Caparra, A. McKeown, T. Whitworth, and E. Lohan, "A survey of spoofer detection techniques via radio frequency fingerprinting with focus on the GNSS pre-correlation sampled data," *Sensors*, vol. 21, no. 9, p. 3012, Apr. 2021.

[18] A. Jagannath, J. Jagannath, and P. Kumar, "A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep Learning, and open challenges," *Comput. Netw.*, vol. 219, Jan. 2022, Art. no. 109455.

[19] 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.

[20] J. Tyler, M. Fadul, and D. Reising, "Considerations, advances, and challenges associated with the use of specific emitter identification in the security of internet of things deployments: A survey," *Information*, vol. 14, no. 9, p. 479, Jul. 2023.

[21] L. Xie, L. Peng, J. Zhang, and A. Hu, "Radio frequency fingerprint identification for Internet of Things: A survey," *Secur. Saf.*, vol. 3, Sep. 2023, Art. no. 2023022.

[22] J. Zhang, G. Shen, W. Saad, and K. Chowdhury, "Radio frequency fingerprint identification for device authentication in the Internet of Things," *IEEE Commun. Mag.*, vol. 61, no. 10, pp. 110–115, Oct. 2023.

[23] S. Abbas et al., "Radio frequency fingerprinting techniques for device identification: A survey," *Int. J. Inf. Secur.*, vol. 23, pp. 1389–1427, Dec. 2023.

[24] H. C. Choe, C. E. Poole, A. M. Yu, and H. H. Szu, "Novel identification of intercepted signals from unknown radio transmitters," in *Proc. Int. Soc. Opt. Eng.*, 1995, pp. 504–517.

[25] J. Hall, M. Barbeau, and E. Kranakis, "Detecting rogue devices in Bluetooth networks using radio frequency fingerprinting," in *Proc. 3rd IASTED Int. Conf. Commun. Comput. Netw. (CCN)*, Lima, Peru, 2006, pp. 4–6.

[26] K. B. Rasmussen and S. Capkun, "Implications of radio fingerprinting on the security of sensor networks," in *Proc. 3rd Int. Conf. Secur. Privacy Commun. Netw. Workshops (SecureComm)*, Nice, France, 2007, pp. 331–340.

[27] W. C. Suski II, M. A. Temple, M. J. Mendenhall, and R. F. Mills, "Using spectral fingerprints to improve wireless network security," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, New Orleans, LO, USA, 2008, pp. 1–5.

[28] M. Köse, S. Taşcıoğlu, and Z. Telatar, "RF fingerprinting of IoT devices based on transient energy spectrum," *IEEE Access*, vol. 7, pp. 18715–18726, 2019.

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

[30] A. M. Ali, E. Uzundurukan, and A. Kara, "Assessment of features and classifiers for Bluetooth RF fingerprinting," *IEEE Access*, vol. 7, pp. 50524–50535, 2019.

[31] C. Zhao, M. Huang, L. Huang, X. Du, and M. Guizani, "A robust authentication scheme based on physical-layer phase noise fingerprint for emerging wireless networks," *Comput. Netw.*, vol. 128, pp. 164–171, Dec. 2017.

[32] W. Hou, X. Wang, and J.-Y. Chouinard, "Physical layer authentication in OFDM systems based on hypothesis testing of CFO estimates," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, ON, Canada, 2012, pp. 3559–3563.

[33] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1658–1667, May 2014.

[34] S. Dolatshahi, A. Polak, and D. L. Goeckel, "Identification of wireless users via power amplifier imperfections," in *Proc. Conf. Rec. 44th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Pacific Grove, CA, USA, 2010, pp. 1553–1557.

[35] A. C. Polak, S. Dolatshahi, and D. L. Goeckel, "Identifying wireless users via transmitter imperfections," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 7, pp. 1469–1479, Aug. 2011.

[36] P. Hao, X. Wang, and A. Behnad, "Performance enhancement of I/Q imbalance based wireless device authentication through collaboration of multiple receivers," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, 2014, pp. 939–944.

[37] J. Yu, A. Hu, G. Li, and L. Peng, "A robust RF fingerprinting approach using multisampling convolutional neural network," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6786–6799, Aug. 2019.

[38] K. Sankhe et al., "No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 1, pp. 165–178, Mar. 2020.

[39] 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.

[40] X. Zhang, T. Li, P. Gong, X. Zha, and R. Liu, "Variable-modulation specific emitter identification with domain adaptation," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 380–395, 2023.

[41] Y. Zhang, Q. Zhang, H. Zhao, Y. Lin, G. Gui, and H. Sari, "Multisource heterogeneous specific emitter identification using attention mechanism-based RFF fusion method," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2639–2650, 2024.

[42] 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.

[43] 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.

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

[45] 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.

[46] 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.

[47] H. Li, Y. Liao, W. Wang, J. Hui, J. Liu, and X. Liu, "A novel time-domain graph tensor attention network for specific emitter identification," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.

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

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

[50] 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.*, London, U.K., 2022, pp. 1–5.

[51] 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.

[52] Z. Cai, W. Ma, X. Wang, H. Wang, and Z. Feng, "The performance analysis of time series data augmentation technology for small sample communication device recognition," *IEEE Trans. Rel.*, vol. 72, no. 2, pp. 574–585, Jun. 2023.

[53] A. Mao, M. Mohri, and Y. Zhong, "Cross-entropy loss functions: Theoretical analysis and applications," in *Proc. 40th Int. Conf. Mach. Learn. (ICML)*, Honolulu, Hawaii, USA, 2023, pp. 23803–23828.

[54] C. Morin, L. Cardoso, J. Hoydis, J.-M. Gorce, and T. Vial, "Transmitter classification with supervised deep learning," in *Proc. Int. Conf. Cogn. Radio Orient. Wireless Netw.*, 2019, pp. 73–86.

[55] R. Morales-Ferre, W. Wang, A. Sanz-Abia, and E.-S. Lohan, "Identifying GNSS signals based on their radio frequency (RF) features–A dataset with GNSS raw signals based on roof antennas and spectracom generator," *Data*, vol. 5, no. 1, p. 18, Feb. 2020.

[56] Y. Liu, J. Wang, S. Niu, and H. Song, 2021, "ADS-B signals records for non-cryptographic identification and incremental learning," Dataset, IEEE Dataport. [Online]. Available: https://dx.doi.org/10.21227/1bxc-ke87.

[57] Y. Tu et al., "Large-scale real-world radio signal recognition with deep learning," *Chinese J. Aeronaut.*, vol. 35, no. 9, pp. 35–48, Sep. 2022.

[58] M. S. Allahham, M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "DroneRF dataset: A dataset of drones for RF-based detection, classification and identification," *Data Brief*, vol. 26, Oct. 2019, Art. no. 104313.

[59] M. Ezuma, F. Erden, C. Kumar Anjinappa, O. Ozdemir, and I. Guvenc, "Detection and classification of UAVs using RF fingerprints in the presence of Wi-Fi and Bluetooth interference," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 60–76, 2020.

[60] E. Uzundurukan, Y. Dalveren, and A. Kara, "A database for the radio frequency fingerprinting of Bluetooth devices," *Data*, vol. 5, no. 2, p. 55, May 2020.

[61] 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. (ICCC)*, Paris, France, 2019, pp. 370–378.

[62] S. Mohanti, N. Soltani, K. Sankhe, D. Jaisinghani, M. Di Felice, and K. Chowdhury, "AirID: Injecting a custom RF fingerprint for enhanced UAV identification using deep learning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, 2020, pp. 1–6.

[63] G. Reus-Muns, D. Jaisinghani, K. Sankhe, and K. R. Chowdhury, "Trust in 5G open RANs through machine learning: RF fingerprinting on the POWDER PAWR platform," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, 2020, pp. 1–6.

[64] N. Soltani, G. Reus-Muns, B. Salehi, J. Dy, S. Ioannidis, and K. Chowdhury, "RF fingerprinting unmanned aerial vehicles with non-standard transmitter waveforms," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15518–15531, Dec. 2020.

[65] J. Shi, L. Peng, H. Fu, and A. Hu, "Robust RF fingerprint extraction based on cyclic shift characteristic," *IEEE Internet Things J.*, vol. 10, no. 21, pp. 19218–19233, Nov. 2023.

[66] A. Al-Shawabka et al., "Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, 2020, pp. 646–655.

[67] H. Liu, C. Hao, Y. Peng, Y. Wang, T. Ohtsuki, and G. Gui, "An effective radio frequency signal classification method based on multi-task learning mechanism," in *Proc. IEEE 96th Veh. Technol. Conf.*, London, U.K., 2022, pp. 1–5.

[68] S. Hanna, S. Karunaratne, and D. Cabric, "WiSig: A large-scale WiFi signal dataset for receiver and channel agnostic RF fingerprinting," *IEEE Access*, vol. 10, pp. 22808–22818, 2022.

[69] A. Jagannath, Z. Kane, and J. Jagannath, "RF fingerprinting needs attention: Multi-task approach for real-world WiFi and Bluetooth," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Rio de Janeiro, Brazil, 2022, pp. 4607–4612.

[70] A. Jagannath and J. Jagannath, "Embedding-assisted attentional deep learning for real-world RF fingerprinting of Bluetooth," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 4, pp. 940–949, Aug. 2023.

[71] A. Elmaghbub and B. Hamdaoui, "LoRa device fingerprinting in the wild: Disclosing RF data-driven fingerprint sensitivity to deployment variability," *IEEE Access*, vol. 9, pp. 142893–142909, 2021.

[72] A. Elmaghbub, B. Hamdaoui, and W.-K. Wong, "ADL-ID: Adversarial disentanglement learning for wireless device fingerprinting temporal domain adaptation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, 2023, pp. 6199–6204.

[73] A. Elmaghbub and B. Hamdaoui, "EPS: Distinguishable IQ data representation for domain-adaptation learning of device fingerprints," 2023, *arXiv:2308.04467*.

[74] S. Zhang et al., "A real-world radio frequency signal dataset based on LTE system and variable channels," 2022, *arXiv:2205.12577*.

[75] X. Yang and D. Li, "LED-RFF: LTE DMRS-based channel robust radio frequency fingerprint identification scheme," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 1855–1869, 2024.

[76] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song, "Machine learning for the detection and identification of Internet of Things devices: A survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 298–320, Jan. 2022.

[77] H. Du et al., "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2611–2646, 4th Quart., 2024, doi: 10.1109/COMST.2024.3400011.

[78] C. Zhao et al., "Enhancing physical layer communication security through generative AI with a mixture of experts," 2024, *arXiv:2405.04198*.

[79] C. Zhao et al., "Generative AI for secure physical layer communications: A survey," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 3–26, Feb. 2025, doi: 10.1109/TCCN.2024.3438379.

[80] R. Zhang et al., "Generative AI for space-air-ground integrated networks," 2024, *arXiv:2311.06523*.

[81] G. Liu et al., "Semantic communications for artificial intelligence generated content (AIGC) toward effective content creation," *IEEE Netw.*, vol. 38, no. 5, pp. 295–303, Sep. 2024, doi: 10.1109/MNET.2024.3352917.

[82] H. Du et al., "Generative AI-aided joint training-free secure semantic communications via multi-modal prompts," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2024, pp. 12896–12900.

[83] D. H. Ballard, "Modular learning in neural networks," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 1987, pp. 279–284.

[84] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[85] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, New York, NY, USA, 2008, pp. 1096–1103.

[86] R. Zhang, P. Isola, and A. Efros, "Colorful image colorization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 649–666.

[87] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, 2017, pp. 105–114.

[88] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, 2022, pp. 15979–15988.

[89] Y. Shi, H. Xu, Y. Zhang, Z. Qi and D. Wang, "GAF-MAE: A self-supervised automatic modulation classification method based on Gramian angular field and masked autoencoder," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 94–106, Feb. 2024.

[90] 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.

[91] Z. Yao et al., "Few-shot specific emitter identification using asymmetric masked auto-encoder," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2657–2661, Oct. 2023.

[92] Y. Zhu, S. Chen, X. Li, S. Zhang, and L. Zhu, "Multi-task self-supervised learning for vehicle classification based on carrier-free UWB radars," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[93] I. J. Goodfellow et al., "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Cambridge, MA, USA, 2014, pp. 2672–2680.

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

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

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

[97] M. K. M. Fadul, D. R. Reising, L. P. Weerasena, T. D. Loveless, M. Sartipi, and J. H. Tyler, "Improving RF-DNA fingerprinting performance in an indoor multipath environment using semi-supervised learning," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 3194–3209, 2024.

[98] A. Dosovitskiy, P. Fischer, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative unsupervised feature learning with exemplar convolutional neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1734–1747, Sep. 2016.

[99] C. Doersch, A. Gupta, and A. A. Efros, "Unsupervised visual representation learning by context prediction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, 2015, pp. 1422–1430.

[100] M. Noroozi and P. Favaro, "Unsupervised learning of visual representations by solving jigsaw puzzles," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 69–84.

[101] I. Misra, C. Zitnick, and M. Hebert, "Shuffle and learn: Unsupervised learning using temporal order verification," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 527–544.

[102] S. Gidaris, P. Singh, and N. Komodakis, "Unsupervised representation learning by predicting image rotations," 2018, *arXiv:1803.07728*.

[103] L. Xu et al., "Few-shot specific emitter identification method using rotation feature decoupling for secure 6G," in *Proc. IEEE 23rd Int. Conf. Commun. Technol. (ICCT)*, Wuxi, China, 2023, pp. 490–494.

[104] L. Xu et al., "Enhanced few-shot specific emitter identification via phase shift prediction and decoupling," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 145–155, Feb. 2025, doi: 10.1109/TCCN.2024.3435886.

[105] M. Gutmann and A. Hyvärinen, "Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics," *J. Mach. Learn. Res.*, vol. 13, pp. 307–361, Feb. 2012.

[106] L. Ericsson, H. Gouk, C. C. Loy, and T. M. Hospedales, "Self-supervised representation learning: Introduction, advances, and challenges," *IEEE Signal Process. Mag.*, vol. 39, no. 3, pp. 42–62, May 2022.

[107] Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, "Unsupervised feature learning via non-parametric instance discrimination," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 3733–3742.

[108] Y. Tian, D. Krishnan, and P. Isola, "Contrastive multiview coding," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Glasgow, U.K., 2020, pp. 776–794.

[109] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, 2020, pp. 9726–9735.

[110] K. Davaslioglu, S. Boztaş, M. C. Ertem, Y. E. Sagduyu, and E. Ayanoglu, "Self-supervised RF signal representation learning for nextG signal classification with deep learning," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 65–69, Jan. 2023.

[111] X. Chen, S. Xie, and K. He, "An empirical study of training self-supervised vision transformers," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Montreal, QC, Canada, 2021, pp. 9620–9629.

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

[113] 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)*, San Diego, CA, USA, 2021, pp. 777–782.

[114] D. Liu, T. Wang, S. Liu, R. Wang, S. Yao, and T. Abdelzaher, "Contrastive self-supervised representation learning for sensing signals from the time-frequency perspective," in *Proc. Int. Conf. Comput. Commun. Netw. (ICCCN)*, Athens, Greece, 2021, pp. 1–10.

[115] M. J. Bocus, H.-S. Lau, R. McConville, R. J. Piechocki, and R. Santos-Rodriguez, "Self-supervised WiFi-based activity recognition," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, 2022, pp. 552–557.

[116] A. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2018, *arXiv:1807.03748*.

[117] I. Misra and L. Maaten, "Self-supervised learning of pretext-invariant representations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, 2020, pp. 6706–6716.

[118] D. Zhao, J. Yang, H. Liu, and K. Huang, "A complex-valued self-supervised learning-based method for specific emitter identification," *Entropy*, vol. 24, no. 7, p. 851, Jun. 2022.

[119] B. Liu et al., "Specific emitter identification based on self-supervised contrast learning," *Electronics*, vol. 11, no. 18, p. 2907, Sep. 2022.

[120] Z. Wu, F. Wang, and B. He, "Specific emitter identification via contrastive learning," *IEEE Commun. Lett.*, vol. 27, no. 4, pp. 1160–1164, Apr. 2023.

[121] X. Hao, Z. Feng, R. Liu, S. Yang, L. Jiao, and R. Luo, "Contrastive self-supervised clustering for specific emitter identification," *IEEE Internet Things J.*, vol. 10, no. 23, pp. 20803–20818, Dec. 2023.

[122] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, "Unsupervised learning of visual features by contrasting cluster assignments," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Red Hook, NY, USA, 2020, pp. 9912–9924.

[123] J. Grill et al., "Bootstrap your own latent-a new approach to self-supervised learning," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Red Hook, NY, USA, 2020, pp. 21271–21284.

[124] X. Chen and K. He, "Exploring simple Siamese representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, 2021, pp. 15745–15753.

[125] J. Zbontar, L. Jing, I. Misra, Y. LeCun, and S. Deny, "Barlow twins: Self-supervised learning via redundancy reduction," in *Proc. 38th Int. Conf. Mach. Learn. (ICML)*, 2021, pp. 12310–12320.

[126] A. Bardes, J. Ponce, and Y. LeCun, "VICReg: Variance-invariance covariance regularization for self-supervised learning," in *Proc. 10th Int. Conf. Learn. Represent. (ICLR)*, 2022, pp. 1–12.

[127] D. Zhao, J. Yang, H. Liu, and K. Huang, "Specific emitter identification model based on improved BYOL self-supervised learning," *Electronics*, vol. 11, no. 21, p. 3485, Oct. 2022.

[128] X. Zha, T. Li, Z. Qiu, and F. Li, "Cross-receiver radio frequency fingerprint identification based on contrastive learning and subdomain adaptation," *IEEE Signal Process. Lett.*, vol. 30, pp. 70–74, Feb. 2023.

[129] T. Miyato, S.-I. 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.

[130] C. Liu et al., "Overcoming data limitations: A few-shot specific emitter identification method using self-supervised learning and adversarial augmentation," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 500–513, 2024.

[131] X. Yun and X. Zhou, "Exploring self-supervised learning for radio signal recognition," in *Proc. IEEE Int. Conf. High Perform. Comput. Commun. (HPCC)*, Haikou, China, 2021, pp. 2425–2430.

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

[133] Y. Zheng, W. Ying, S. Hong, and L. Wang, "A method for cross-receiver specific emitter identification based on CBAM-CNN-BDA," in *Proc. IEEE 4th Int. Conf. Civil Aviat. Safety Inf. Technol. (ICCASIT)*, Dali, China, 2022, pp. 1320–1324.

[134] J. Liu, J. Li, J. Wang, and H. Huang, "Specific emitter identification at different time based on multi-domain migration," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Chiang Mai, Thailand, 2022, pp. 917–922.

[135] G. Shen, J. Zhang, A. Marshall, M. Valkama, and J. R. Cavallaro, "Toward length-versatile and noise-robust radio frequency fingerprint identification," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 2355–2367, 2023.

[136] M. Zeng, Y. Yao, H. Liu, Y. Hu, and H. Yang, "A specific emitter identification system design for crossing signal modes in the air traffic control radar beacon system and wireless devices," *Sensors*, vol. 23, no. 20, p. 8576, Oct. 2023.

[137] S. Mackey, T. Zhao, X. Wang, and S. Mao, "Cross-domain adaptation for RF fingerprinting using prototypical networks," in *Proc. 20th ACM Conf. Embed. Netw. Sens. Syst. (SenSys)*, New York, NY, USA, 2022, pp. 812–813.

[138] A. Garza, C. Challu, and M. Mergenthaler-Canseco, "TimeGPT-1," 2023, *arXiv:2310.03589*.

[139] K. Rasul et al., "Lag-Llama: Towards foundation models for time series forecasting," 2023, *arXiv:2310.08278*

[140] M. Goswami, K. Szafer, A. Choudhry, Y. Cai, S. Li, and A. Dubrawski, "MOMENT: A family of open time-series foundation models," 2024, *arXiv:2402.03885*.

[141] J. Lin, E. Keogh, L. Wei, and S. Lonardi, "Experiencing SAX: A novel symbolic representation of time series," *Data Min. Knowl. Discov.*, vol. 15, pp. 107–144, Apr. 2007.

[142] H. Xue and F. Salim, "PromptCast: A new prompt-based learning paradigm for time series forecasting," 2023, *arXiv:2210.08964*.

[143] Z. Borsos et al., "AudioLM: A language modeling approach to audio generation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 31, pp. 2523–2533, Jun. 2023.

[144] Z. Wang and H. Ji, "Open vocabulary electroencephalography-to-text decoding and zero-shot sentiment classification," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 2022, pp. 5350–5358.

[145] Y. Liu, N. Gao, X. Li, and S. Jin, "Large language model enabled lightweight RFFI for 6G edge intelligence," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Milan, Italy, 2025, pp. 1–6.

[146] M. Tao et al., "Resource-constrained specific emitter identification using end-to-end sparse feature selection," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, 2023, pp. 6067–6072.

[147] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of Things," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 142–153, Jan. 2021.

[148] 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.

[149] T. Chen et al., "The lottery tickets hypothesis for supervised and self-supervised pre-training in computer vision models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, TN, USA, 2021, pp. 16301–16311.

[150] A. Saeed, F. D. Salim, T. Ozcelebi, and J. Lukkien, "Federated self-supervised learning of multisensor representations for embedded intelligence," *IEEE Internet Things J.*, vol. 8, no. 2, pp. 1030–1040, Jan. 2021.

[151] X. Fu, G. Gui, Y. Wang, H. Gacanin, and F. Adachi, "Automatic modulation classification based on decentralized learning and ensemble learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7942–7946, Jul. 2022.

[152] W. Li et al., "SLPA: Single-line pixel attack on specific emitter identification using time-frequency spectrogram," *IEEE Trans. Veh. Technol.*, *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15763–15767, Oct. 2024, doi: 10.1109/TVT.2024.3405547.

[153] Z. Tang et al., "RF domain backdoor attack on signal classification via stealthy trigger," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 11765–11780, Dec. 2024, doi: 10.1109/TMC.2024.3404341.

[154] S. Purushwalkam, P. Morgado, and A. Gupta, "The challenges of continuous self-supervised learning," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 702–721.

[155] M. Liu, J. Wang, N. Zhao, Y. Chen, H. Song, and F. R. Yu, "Radio frequency fingerprint collaborative intelligent identification using incremental learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 5, pp. 3222–3233, Sep./Oct. 2022.
[156] D. Li, J. Qi, S. Hong, P. Deng, and H. Sun, "A class-incremental approach with self-training and prototype augmentation for specific emitter identification," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 1714–1727, 2023.
[157] A. Jaiswal, A. Babu, M. Zadeh, D. Banerjee, and F. Makedon, "A survey on contrastive self-supervised learning," *Technologies*, vol. 9, no. 1, p. 2, Dec. 2020.
[158] J. M. Santos and M. Embrechts, "On the use of the adjusted rand index as a metric for evaluating supervised classification," in *Proc. Int. Conf. Artif. Neural Netw.*, Limassol, Cyprus, 2009, pp. 175–184.
[159] T. O. Kvålseth, "On normalized mutual information: Measure derivations and properties," *Entropy*, vol. 19, no. 11, p. 631, 2017.
[160] P. J. Rousseeuw. "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987.
[161] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 1, no. 2, pp. 224–227, Apr. 1979.
[162] L. Papangelo, M. Pistilli, S. Sciancalepore, G. Oligeri, G. Piro, and G. Boggia, "Adversarial machine learning for image-based radio frequency fingerprinting: Attacks and defenses," *IEEE Commun. Mag.*, vol. 62, no. 11, pp. 108–113, Nov. 2024.
[163] Z. Yao, Y. Wang, G. Gui, S. Mao, and X. Wang, "New frontier of communication security on radio frequency fingerprints concealment," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 156–163, Oct. 2024.

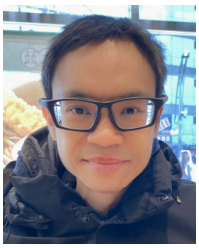
Chao Liu (Graduate Student Member, IEEE) received the B.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2021, where he is currently pursuing the Ph.D. degree in communication engineering. He has published over five papers in peer-reviewed IEEE journals/conferences, including one ESI highly-cited paper and one hot paper. His research interests include deep learning and radio signal identification.

Guan Gui (Fellow, IEEE) received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. From 2009 to 2014, he joined Tohoku University as a Research Assistant and a Postdoctoral Research Fellow. From 2014 to 2015, he was an Assistant Professor with Akita Prefectural University, Akita, Japan. Since 2015, he has been a Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China. His recent research interests include intelligence sensing and recognition, intelligent signal processing, and physical layer security. He has published over 200 IEEE Journal/Conference papers and won several best paper awards, such as ICC 2017, ICC 2014, and VTC 2014-Spring. He received the IEEE Communications Society Heinrich Hertz Award in 2021, the Clarivate Analytics Highly Cited Researcher in Cross-Field from 2021 to 2024, the Member and Global Activities Contributions Award in 2018, the Top Editor Award of IEEE Transactions on Vehicular Technology in 2019. Since 2022, he has been a Distinguished Lecturer of the IEEE Vehicular Technology Society. He serves or serves on the editorial boards for several journals, such as IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE INTERNET OF THINGS JOURNAL, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. In addition, he served as the IEEE VTS Ad Hoc Committee Member in AI Wireless, the Executive Chair of IEEE ICCT 2023 and Chair of VTC 2021-Fall, and the Vice Chair of WCNC 2021.

Yu Wang (Member, IEEE) received the Ph.D. degree in signal and information processing from the Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2023, where he has been a Principal-Appointed Professor since 2023. He has published more than 100 articles in peer-reviewed IEEE journals/conferences. His research interests include deep learning, optimization, and its application in wireless communications. He received six best paper awards.

Tomoaki Otsuki (Ohtsuki) (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1990, 1992, and 1994, respectively.

From 1994 to 1995, he was a Postdoctoral Fellow and a Visiting Researcher of Electrical Engineering with Keio University. From 1993 to 1995, he was a Special Researcher with the Fellowships of the Japan Society for the Promotion of Science for Japanese Junior Scientists. From 1995 to 2005, he was with the Science University of Tokyo. In 2005, he joined Keio University where he is currently a Professor. From 1998 to 1999, he was with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley. He has published more than 315 journal papers and 540 international conference papers. He is engaged in research on wireless communications, optical communications, signal processing, and information theory. He is a recipient of the 1997 Inoue Research Award for Young Scientist, the 1997 Hiroshi Ando Memorial Young Engineering Award, the Ericsson Young Scientist Award 2000, the 2002 Funai Information and Science Award for Young Scientist, the IEEE the 1st Asia–Pacific Young Researcher Award 2001, the 5th International Communication Foundation Research Award, the 2011 IEEE SPCE Outstanding Service Award, the 27th TELECOM System Technology Award, the ETRI Journal's 2012 Best Reviewer Award, the 9th International Conference on Communications and Networking in China 2014 (CHINACOM 2014) Best Paper Award, the 2020 Yagami Award, the 26th Asia–Pacific Conference on Communications Best Paper Award, the International Conference on Internet of Things, Communication and Intelligent Technology 2024 Best Paper Award, and the 2024 6th International Conference on Robotics, Intelligent Control and Artificial Intelligence Best Paper Award. He served as the Chair for IEEE Communications Society, Signal Processing for Communications and Electronics Technical Committee. He served as a Technical Editor for the *IEEE Wireless Communications Magazine* and an Editor for *Physical Communications* (Elsevier). He is currently serving as an Area Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and an Editor for the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS. He is also serving as the IEEE Communications Society, and the Asia–Pacific Board Director. He has served as the General-Co Chair, the Symposium Co-Chair, and the TPC Co-Chair of many conferences, including IEEE GLOBECOM 2008, SPC, IEEE ICC 2011, CTS, IEEE GLOBECOM 2012, SPC, IEEE ICC 2020, SPC, IEEE APWCS, IEEE SPAWC, and IEEE VTC. He gave tutorials and keynote speeches at many international conferences, including IEEE VTC, IEEE PIMRC, and IEEE WCNC. He was Vice President and a President of the Communications Society of the IEICE, also he was a Distinguished Lecturer of the IEEE. He is a Fellow of the IEICE and Asia–Pacific Artificial Intelligence Association, and a member of the Engineering Academy of Japan.



Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkuts Institute of Technology Ladkrabang, Thailand, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada. He is a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests are in the areas of mobile generative AI, edge general intelligence, quantum computing and networking, and incentive mechanism design.

Xuemin Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990.

He is a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on network resource management, wireless network security, Internet of Things, AI for networks, and vehicular networks. He received the West Lake Friendship Award from Zhejiang Province in 2023, the President's Excellence in Research from the University of Waterloo in 2022, the Canadian Award for Telecommunications Research from the Canadian Society of Information Theory in 2021, the R.A. Fessenden Award in 2019 from IEEE, Canada, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015, the Education Award in 2017 from the IEEE Communications Society, and the Technical Recognition Award from Wireless Communications Technical Committee in 2019, and the AHSN Technical Committee in 2013. He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier's Research Excellence Award in 2003 from the Province of Ontario, Canada. He is the Past President of the IEEE Communications Society. He was the Vice President for Technical and Educational Activities, a Vice President for Publications, a Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a Member of IEEE Fellow Selection Committee of the ComSoc. He served as the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and *IET Communications*. He is a Registered Professional Engineer of ON, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, an International Fellow of the Engineering Academy of Japan, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.