Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6296954, 12 pages
https://doi.org/10.1155/2022/6296954

Research Article

Electromagnetic Signal Intelligent Identification Based on Radio Frequency Fingerprints

Jian Kang,[1] Hui Mu,[1] Hui Ren,[1] Jicheng Jia,[2] Lin Qi,[2] and Zherui Zhang[2]

[1]*Telecommunication Research from Beijing Institute of Astronautical System Engineering, Beijing 100076, China*
[2]*College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China*

Correspondence should be addressed to Zherui Zhang; zhangzherui@hrbeu.edu.cn

Received 24 March 2022; Accepted 6 May 2022; Published 7 June 2022

Academic Editor: Mingqian Liu

Due to the open nature of WIFI connection, it is exposing its private information to the attackers. Traditional WIFI security methods are no longer able to meet the current security needs, and more and more wireless-side physical layer security solutions provide solutions, among which RF fingerprinting is an endogenous security technology with potential. Constructing an effective and accurate method to identify WIFI devices that steal information is a difficulty that today's society needs to face. The main problem is not only that the recognition accuracy is difficult to improve but also the problem of data shortage. In this paper, we first construct a large-scale WIFI real-world measurement dataset. Next, we use PSD and bispectrum features, as well as complex ResNet schemes for WIFI device identification experiments, and compare and analyze them from multiple perspectives. The experimental results show that the proposed algorithm can achieve up to 97% recognition accuracy among 100 devices. Moreover, when the SNR is 0 dB, the complex ResNet method can still achieve 78% recognition accuracy among 100 devices. Finally, this paper summarizes the experimental analysis of the measured dataset and discusses the open issues related to this area.

1. Introduction

According to IoT analytics, the Internet of Things (IoT) market is expanding at an accelerated rate and the number of connected IoT devices is increasing rapidly. In 2021, IoT analytics expects the number of connected IoT devices to grow by 9 percent to 12.3 billion active IoT devices worldwide. This number is expected to reach 30 billion by 2025. As shown in Figure 1, while IoT technology brings convenience to life, it also poses many security risks and its security is becoming increasingly prominent.

As a standard for high-throughput wireless connectivity, the IEEE 802.11 standard, or WIFI standard, is highly vulnerable to various attacks [1–3]. There are many methods for WIFI device identification. However, all these schemes have certain vulnerabilities and the standard IEEE 802.11 authentication method may encounter various attacks such as man-in-the-middle attacks, deauthentication flooding attacks, MAC address replication, and other attacks. Such security vulnerabilities can lead to serious system risks, so it becomes especially important to perform secure and accurate WIFI device identification.

In addition to the methods mentioned above, the identification and authentication of wireless devices can also be performed using Radio Frequency Fingerprinting (RFF) technology [4]. The concept of "RF fingerprinting" was first proposed in 2003, and this technology is a physical layer method to enhance the security of wireless networks by extracting the subtle features of radiation source devices to build device fingerprints similar to those in biology, thus completing device identification. The process of RF fingerprint extraction and identification is carried out in its physical layer, which can work alone as well as assist network-side security mechanisms to improve security for communication networks.

The basic characteristics of the physical layer of the device are only related to the hardware characteristics of the device itself, and the characteristics of the physical layer can accurately locate each machine in a large number of devices [5]. Although such hardware differences can be

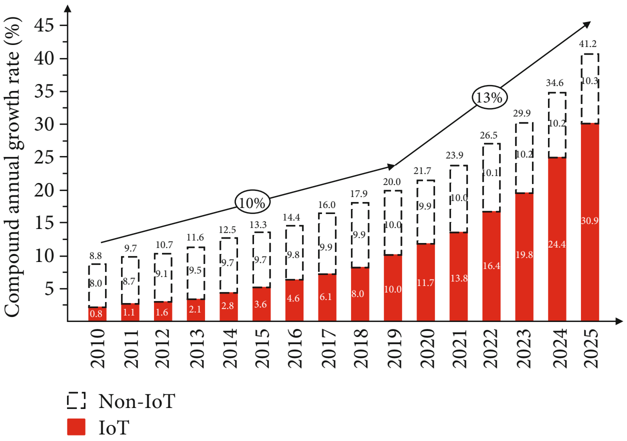


Figure 1: Trends in the number of IoT device connections.

reduced by more precise manufacturing and quality control means, they are usually not feasible in practical environments because they lead to significantly higher production costs. Physical layer-based security protection mechanisms have been extensively studied in recent years. They play an important role in network intrusion detection [6], fault diagnosis [7], target tracking [8, 9], radar detection [10, 11], etc.

RF fingerprinting technology has had important applications in radiation source identification systems, which can be used for secure identification and authentication of WIFI devices based on physical layer information. Fingerprint features are generated due to certain hardware "defects" introduced during the device manufacturing process, which are reflected in the communication signal and are similar to fingerprints in biology, which are also unique and difficult to be copied [5].

The individual identification of communication devices based on RF fingerprints is of great importance in many fields. In the military field, the use of RF fingerprint extraction and identification technology can identify important radar, communication, navigation, data chain and other radiation source individuals, grasp the identity and attributes of the user, and make strategic adjustments by tracking and monitoring enemy equipment, so as to grasp the initiative of military operations in the complex battlefield environment. In the civilian field, RF fingerprinting technology has important applications in wireless network security, equipment fault diagnosis, and quality management. Especially in the field of security of communication networks [12], the security problems brought by wireless communication networks due to their inherent openness need urgent attention [13], and most of the traditional methods for securing wireless networks are realized through authentication based on the physical layer above. Since RF fingerprinting is an inherent property of the physical layer of wireless devices, it is not easily tampered with and can effectively improve wireless network security, and RF fingerprinting technology can detect malicious network attacks, which is of great significance to enhance wireless network security.

In this paper, we address the problem of accurate identification of physical layer based on RF fingerprinting for a large number of WIFI devices working in real environment. It is planned to collect in batches for more than 100 WIFI devices in multiple channel environments to build the dataset. And multiple methods are used for identification testing. Therefore, the difficulty of this paper is that the number of devices to be identified is large and the recognition effect is difficult to improve due to the complex and diverse channel environment, but this paper is also more practical.

This paper will introduce three algorithms to solve the above problem. The main contributions of this paper are summarized as follows:

(1) We propose one of the largest WIFI real-world datasets available. This includes 100 WIFI devices of IEEE 802.11b standard, collected in various scenarios, including darkroom and laboratory, and various channel environments, including LoS and NLoS, using several spectrometer devices for real measurements. A total of 500 GB of WIFI signal dataset was constructed

(2) We designed two algorithms based on feature engineering, namely, power spectrum analysis and dual spectrum analysis, and conducted experiments on WIFI device identification. The results show that the two algorithms based on feature engineering can achieve more than 92% recognition accuracy on 100 devices. The performance, advantages, and shortcomings of the algorithms are compared from several perspectives. The experimental results are also fully analyzed

(3) We designed a deep learning-based algorithm for WIFI device recognition using a deep complex residual network. The results show that the algorithm can achieve more than 97% recognition accuracy on 100 devices. The experimental results are analyzed and compared with the two algorithms based on feature engineering

The rest of the paper is organized as follows. In Section 2, the work on RF fingerprinting is presented. In Section 3, the WIFI device recognition system of this paper is introduced and the acquisition process of the large-scale real-world dataset used in this paper is presented. The WIFI device recognition algorithm based on feature engineering and deep learning will be introduced in Section 4. And in Section 5, simulation experiments are conducted to compare and analyze the results from multiple perspectives. Finally, a summary of the full work is presented in Section 6, where the experimental results are summarized and analyzed.

2. Related Work

The RF fingerprint feature extraction and device identification system is shown in Figure 2.

Generally speaking, the radiation source identification system based on RF fingerprint technology is composed as follows: data acquisition part, preprocessing part, fingerprint feature calculation and extraction part, and classification and

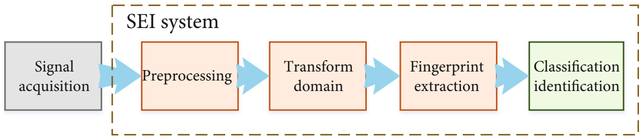


Figure 2: Block diagram of radiation source equipment classification and identification system.

identification part. Data acquisition part is the use of acquisition equipment for wireless or wired signal acquisition, usually using oscilloscopes, spectrometers, Universal Software Radio Peripheral (USRP), Software Defined Radio (SDR), and other equipment and platforms, etc. The performance and structure of the front-end equipment usually have an impact on the RF fingerprinting system, and the performance of the system and the performance, mobility, and portability of the acquisition equipment need to be considered. The preprocessing part is to perform data cleaning, interception, noise reduction, and other operations on the collected device signal data and to unify the multiple signals that may be collected by multiple devices, so that they can be applied to the calculation and extraction of fingerprint features. Fingerprint feature calculation and extraction is the most important part of the system, mainly through one or more signal processing algorithms to carry out transform domain analysis of the signal, so as to extract subtle features such as unintentional modulation that may otherwise be submerged in the modulated signal, find subtle differences in the signal that can be used to distinguish multiple devices, and output them as the device's RF fingerprint features. Finally, the classification and recognition part often uses generative or nongenerative classifiers, etc. The obtained fingerprint features are used to train the classifier, or they are fed to the already trained classifier to get the final recognition result, so that the RF fingerprint-based radiation source device recognition is achieved. Besides, there are data enhancement, data dimensionality reduction, and other components, which can be added or deleted as appropriate.

The transient and steady-state signal segments are considered separately for RF fingerprinting studies. Among them, the transient signal segment is a signal segment with a short duration of gradual power increase generated by the device at the moment of switching on and off. Since it is only related to the hardware circuit of the device and contains rich information about the subtle characteristics of the radiation source device, the transient signal is a suitable signal to be used for RF fingerprinting technology research [14, 15]. However, it is difficult to capture the signal segment quickly and accurately in a real system. The steady-state signal is the signal segment of the transmitter RF signal that operates stably at rated power, and the steady-state signal lasts longer and is easier to obtain than the transient signal [16]. The radiation source signal received by the receiver is shown in Figure 3, and it can be seen that the signal collected by the receiver is extremely rich, including the channel noise segment, transient signal segment, and steady-state signal segment.

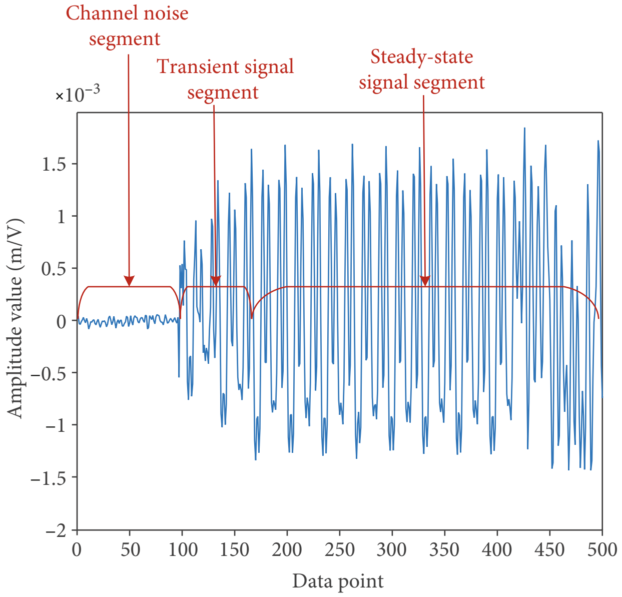


Figure 3: IEEE 802.11a wireless signal.

For the problem of individual identification of WIFI devices based on RF fingerprinting studied in this paper, the domestic and international research is summarized into the following five categories of methods.

2.1. RF Fingerprinting Based on Location Information. Location-based RF fingerprinting systems are built on features such as radio signal strength (RSS), channel state information (CSI), and channel frequency response (CFR), which are signal segments that contain location information about the target device. Therefore, these systems are designed to use the unique location information of the device for device identification.

Generally speaking, such systems mainly utilize the special structure of IEEE 802.11 standard frames of WIFI to extract the corresponding fingerprint features and perform device identification by processing data segments such as short training field (STF), long training field (LTF), and pilot frequencies.

In 2018, Li et al. used an FPGA- (Field Programmable Gate Array-) based platform to achieve 85% recognition rate in three devices using features, such as CSI, RSS, and received frequency bias extracted from the leading signal, and verified the algorithm and platform proposed in the paper in real-time communication with commercial WIFI devices in validity [17]. In 2019, the literature [18] proposed a RFF feature extraction method for the differential phase of the guide frequency (DPoP) and the amplitude of the quotient (AoQ) using the guide frequency, STF, and LTF of OFDM frames and used deep neural networks (DNN) to classify and identify a total of 55 WIFI devices for five devices, achieving 95% identification accuracy when the SNR was higher than 40 dB.

2.2. RF Fingerprinting Based on Transmission Signal. RF fingerprinting technology systems based on signal statistical features are built on features extracted from transient and steady-state signals and presynchronization codes. These systems use the uniqueness of a fixed segment of all RF signal packets sent by the authenticated device for device identification. The feature extraction algorithms used are generally commonly used signal processing algorithms with advantages such as reproducibility and ease of implementation.

In 2019, Yang et al. collected the overall burst signal of a transmitter from power on to data communication to power off and proposed a classification scheme based on sparse representation (SRC), whose results indicated that the identification performance using multiple signal segments as RF fingerprints outperformed the results using only one signal segment [19]. In 2020, the literature [20] proposed the use of fractional order Fourier transform, power spectrum, and bispectral analysis of feature extraction method to achieve multiangle feature extraction, which can still achieve better recognition performance on 10 intercom devices even with low SNR.

2.3. RF Fingerprinting Based on RF Modulation Error. The RF fingerprinting system based on signal modulation errors is similar to the previous type of scheme, using transient and steady-state signal segments for feature extraction. The difference is that this part of the features has a strong physical meaning and is an unintentional modulation feature of the transmitter.

In 2020, it is also possible to use the constellation diagram feature of the signal for the extraction of fine features, the principle of which lies in the fact that communication standards specify that the signal modulation can have a certain range of errors, and this error is the feature space in the field of RF fingerprinting technology, thus enabling the RF fingerprinting of devices [21].

2.4. RF Fingerprinting Based on Physical Modeling. Mathematical modeling of key hardware components in wireless communication systems is also an extremely important task. The physical model obtained by modeling can also be used as a unique feature to distinguish different radiation source devices.

Such systems are mainly used to perform device identification by mathematical modeling of the overall system or a single device, and the obtained model will be used. The literature [22] allows the extraction of the nonlinear components of the radiation source device and the analysis of the signal in the frequency domain for the identification of the radiation source device. The literature [23], on the other hand, directly models the transmitter link as a whole and fits the model using a kernel regression model to obtain an RF fingerprint of the radiation source called FID, which can achieve 99% identification accuracy on 33 devices.

2.5. RF Fingerprinting Based on Deep Learning. Deep learning has achieved good results in a variety of classification domains, and the use of deep learning for transmitter identification is a feasible solution. Deep learning models directly use the RF signals sent by the device and can be trained by training a designed deep network model to classify the device [24–27].

In 2018, literature [28] used convolutional neural network (CNN) to classify wireless signals to identify IoT devices with 92.29% recognition accuracy on seven ZigBee devices and high channel robustness. In 2019, literature [29] proposed a neural network-based radiation source identification algorithm that can be executed on resource-constrained IoT devices and can achieve better good recognition performance, which provides a reference for the grounded application of the algorithm. In 2019, and for the small and zero-sample problem in deep learning, a ZSL framework for signal recognition and reconstructed convolutional neural network (SR2CNN) is proposed in the literature [30] to solve the relevant recognition problem in this case. Appropriate combinations of cross-entropy loss, central loss, and reconstruction loss as well as a suitable distance metric space are introduced to learn the representation of the signal semantic feature space so that the semantic features have a larger minimum interclass distance than the maximum intraclass distance. In 2021, the literature [31] used the actual collected ADS-B signals to construct a dataset that provides an in-depth study of the performance of the deep learning model and compares it with recognition benchmarks using machine learning and deep learning methods. An open-ended discussion is also provided.

To address the challenge that the existing models have large number of parameters and are difficult to deploy. Pruning methods have also been proposed in the literature that omit the less important convolutional filters and achieve equal or higher classification accuracies [32, 33]. In addition to this, there are distributed sensor systems that use incremental learning to solve the RF fingerprint identification problem [34].

The above analysis reveals that most of the existing algorithms can achieve better results in a certain scale of radiation source devices; however, they are more susceptible to noise, channel environment, etc. Some algorithms can achieve considerable results in a larger number of devices; however, their algorithms are more complex and more difficult to implement, which is not easy to carry out practical applications and deployments. Therefore, the current RF fingerprinting technology focuses on the need to solve the problem of achieving effective identification of large-scale radiation source devices easily and quickly in a variety of complex channel environments, which is the main research direction of this paper.

3. Dataset Construction

For the content to be studied in this paper, the dataset is constructed as follows. The data collection object is 100 IEEE 802.11b standard 2.4GHz WIFI network card modules. The selected 2.4GHz WIFI module model is ESP8266. The WIFI module is set to IEEE 802.11b WIFI standard communication mode through configuration, sending Beacon frame signal every 100 ms, and RF signal bandwidth 20 MHz. The module is a complete module, including WIFI chip, external circuit, and PCB antenna. In this paper, a spectrometer is used for wireless acquisition of signal power triggers, and the FSW26 spectrometer from Rohde & Schwarz is selected. This spectrometer provides multiple sampling rates for the experiments.

Signal data acquisition and dataset construction are carried out using the abovementioned equipment. Wireless acquisition of LoS, multipath, and NLoS multipath channels in the laboratory and microwave darkroom is carried out using FSW26 spectrometers. Signal acquisition is mainly performed for 100 2.4 GHz WIFI modules, 100 5 GHz WIFI modules for management frame Beacon frames, and I/Q signal acquisition using 40 M, 80 M, and 160 M sampling rates in channel 1 and 6 transmission channels.

Data acquisition scenarios cover many scenarios, such as laboratory LoS, LoS multipath, NLoS multipath scenarios, and darkroom LoS scenarios, including the acquisition of WIFI signal data in scenarios with different receiving devices, different transmission channel numbers, and different channel environments.

The data acquisition platform is built using MATLAB and python-based platforms on the computer side, and the RF signal is automatically acquired from WIFI devices by using a network cable to connect the spectrum meter for control and setting the acquisition parameters, etc.

The data acquisition environment is set as follows. Among them, the laboratory LoS environment is the WIFI module, and the spectrum meter antenna is about 1 m apart; the laboratory NLoS environment is the WIFI module, and the spectrum meter antenna is about 2 meters apart, and there is a baffle in the direct path.

In several acquisition scenarios, the obtained WIFI signal data are stored in Excel xlsx file format using the data acquisition platform to record the I/Q two-way amplitude values of RF signals. The final dataset obtained is 500 GB in total.

4. System Implementation

This paper will study the performance of two commonly used signal processing algorithms, power spectral density analysis and bispectral analysis, in the field of RF fingerprinting technology, analyze their basic theory, and study their feasibility in the field of RF fingerprinting technology for WIFI devices. Besides, this paper will study the performance of the residual neural network model technique applied to WIFI device identification, analyze its basic theory, and introduce its complex processing form to study its feasibility in the field of WIFI device identification.

In this paper, we use the individual identification framework shown in Figure 1 to design the WIFI device identification system shown in Figure 4. The system includes a signal acquisition module, which uses a spectrum meter for power-triggered acquisition of WIFI signals. After that, the collected signals are intercepted as well as normalized. After that, the RFF calculation and extraction are performed, and this paper uses two methods of feature engineering and deep learning models for RFF extraction. Next, the system will build the RFF library for the training of the classifier. Finally, the output of the device recognition results is performed.

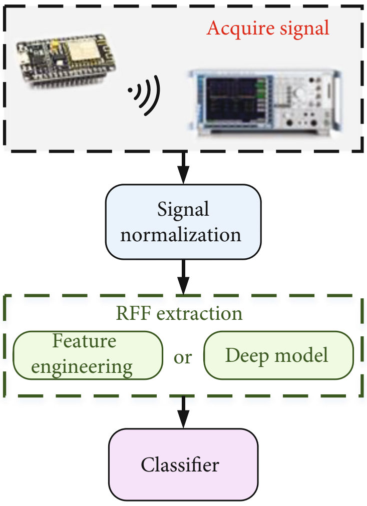


Figure 4: Block diagram of WIFI device identification process.

4.1. PSD-Based WIFI Device Recognition Algorithm.

Power spectral density (PSD) analysis is a probabilistic statistical characterization of a random signal, which refers to the "power" (mean square value) per unit frequency band, also known as energy spectrum, power spectrum, spectral density, etc. Suppose there is a smooth random signal such as then the PSD is shown in

$$P(\omega) = |S(\omega)|^2, \tag{1}$$

where $S(\omega)$ is the Fourier transform of the smooth random signal $s(t)$, i.e.

$$S(\omega) = \int_{-\infty}^{+\infty} s(t) \cdot e^{-j\omega t} dt. \tag{2}$$

According to the Wiener-Khinchin theorem, the autocorrelation function of a smooth random signal and the power spectral density present a Fourier transform relationship. That is, the autocorrelation function of a signal can be used for its power spectral density estimation, which is also a classical method for power spectral density estimation.

Now, assume that the received signal is $s[n]\{n = 0, 1, 2, 3, \cdots, N-1\}$, where N is the number of received signal points, then the autocorrelation function $\widehat{R}_{ss}[\tau]$ of signal $s[n$] can be expressed as

$$\widehat{R}_{ss}[\tau] = \frac{1}{N} \sum_{i=0}^{N-|\tau|-1} s[i]s^*[i+\tau]. \tag{3}$$

Since the received signal $s[n]$ is a smooth random process signal, its autocorrelation function is only related to the time interval τ, where $s^*[i]$ denotes the complex conjugate of signal $s[n]$.

After obtaining the autocorrelation function $\widehat{R}_{ss}[\tau]$, the power spectral density estimate $\widehat{P}(\omega)$ of signal $s[n]$ is shown in

$$\widehat{P}(\omega) = P_N(\omega) = \sum_{\tau=-(N-1)}^{N-1} \widehat{R}_{ss}[\tau] \cdot e^{-j\tau\omega T}. \quad (4)$$

It can be seen that the power spectral density estimate obtained indirectly using the autocorrelation function method is based on the average of the N-point signal $s[n]$, which is a limitation of using finite-length signals for power spectral density estimation. Therefore, in practical applications, the method of adding a window function is also used to make the power spectral density estimation more accurate, and the operation of adding a window is shown in

$$\widehat{P}(\omega) = \sum_{j=-\infty}^{\infty} W(j)\widehat{R}_{ss}[\tau]. \quad (5)$$

Among them, $W(j)$ is the window function, which can be selected as a series of window functions such as rectangular window and Hamming window.

Of course, different types of window functions can lead to different values of the final estimated power spectral density, which requires a comprehensive consideration of the estimated resolution, frequency band, and other parameters of interest, so as to select the appropriate window function.

Based on the above theoretical analysis and derivation, Algorithm 1 demonstrates the PSD-based RFF calculation and extraction designed in this paper.

After extracting the PSD RF fingerprint feature of the signal, it can be substituted into the framework of the RF fingerprint-based identification system shown in Figure 2 to realize the RF fingerprint-based WIFI device identification.

4.2. Bispectrum-Based WIFI Device Recognition Algorithm. High-order spectral analysis has good antinoise and time-varying properties because it can accurately extract phase and amplitude information, which can provide a better analysis basis in the complex and changing communication environment nowadays. High-order spectral analysis can perform multidimensional feature analysis of amplitude and phase information, while bispectral analysis can achieve the effect of suppressing Gaussian noise and is widely used in various signal processing analysis fields.

First of all, higher-order spectral analysis is a higher-order analytical extension of the power spectral density estimation of a signal, in other words, higher-order spectral analysis also satisfies the Wiener-Synchon theorem. Therefore, it is obtained that the order spectrum of a signal is related to its order cumulative quantity by the order Fourier transform. Then, the higher-order spectral analysis of a continuous random variable proceeds as follows.

(1) First determine the received signal $s[n]\{n = 0, 1, 2, 3, \cdots, N-1\}$, which has a sampling frequency of f_s and N_0 frequency sampling points in the bispectral domain with an interval of $\triangle_0 = f_s/N_0$

(2) Dividing the received signal $s[n]$ into K segments, each with M points, and centralizing each subsignal segment

(3) The DFT coefficients of each subsignal segment are calculated as shown in

$$S_i[m] = \frac{1}{M}\sum_{n=0}^{M-1} s_i[n]e^{-j2\pi nm/M} \quad (6)$$

where $m = 0, 1, \cdots, M/2$, $i = 0, 1, \cdots, K$, and $s_i[n]$ are the i th subsignal segment

(4) Next, the bispectral estimate of the ith subsignal segment $s_i[n]$ can be calculated

(5) Finally, the bispectral estimates of the K signal segment $s_i[n]$ are combined and averaged to obtain the bispectral estimate $B(\omega_1, \omega_2)$ of signal $s[n]$

$$\begin{gathered} B(\omega_1, \omega_2) = \frac{1}{K}\sum_{K}^{i=1} B_i(\omega_1, \omega_2), \\ \omega_1 = \left(\frac{2\pi f_s}{N_0}\right)m_1, \omega_2 = \left(\frac{2\pi f_s}{N_0}\right)m_2 \end{gathered} \quad (7)$$

After the above analysis, it has been possible to obtain the bispectrum analysis of the signal and realize the fast and simple bispectrum estimation by the simplification operation. However, the final obtained bispectrum is a two-dimensional complex image, and in the field of RF fingerprint technology, the processing of the image usually makes the efficiency of the operation reduced, and a dimensionality reduction method is needed to realize the conversion of the bispectrum analysis from a two-dimensional image to a one-dimensional vector. Here, we introduce the partial integration bispectrum analysis method, which integrates the two-dimensional bispectrum image to obtain the one-dimensional vector value through different integration paths to realize the information dimensionality reduction.

Radially Integrated Bispectra (RIB): The integration path is a straight line passing through the origin of coordinates, and the obtained integration value is $\mathrm{RIB}(\alpha)$, as shown in

$$\mathrm{RIB}(\alpha) = \int_{0^+}^{1/(1+\alpha)} B(\omega_1, \alpha\omega_1)d\omega_1. \quad (8)$$

This RIB analysis is time-shift invariant as well as phase-invariant.

Axially Integrated Bispectra (AIB): The integration path is a straight line parallel to the coordinate axis, and the

```
Input: Signal, $s[n]$; signal points, $N$; FFT points, $L$; signal bandwidth, $B_s$; sampling frequency, $f_s$;
Output: PSD RF fingerprint features, $RFF_{PSD}$;
The received signal $s[n]$ is normalized to obtain $s'[n]$.
The PSD of $s'[n]$ is obtained according to Equation (5).
    $\widehat{P}(\omega) = \sum_{j=-\infty}^{\infty} W(j)\widehat{R}_{ss}[\tau]$
The $\widehat{P}(\omega)$ subsegment of the intercepted validity signal is used as the PSD RF fingerprint feature
    $RFF_{PSD} = P[L/2 - L \cdot B_s/2 \cdot f_s, (L/2) + (L \cdot B/2 \cdot f)]$
return $RFF_{PSD}$.
```

Algorithm 1: RFF calculation and extraction based on PSD analysis.

obtained integration value is $\mathrm{AIB}(\omega)$, as shown in

$$\mathrm{AIB}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} B(\omega_1, \omega_2) d\omega_2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} B(\omega_1, \omega_2) d\omega_1. \tag{9}$$

The AIB analysis is time-shift invariant as well as scale-stretch invariant.

Circularly Integrated Bispectra (CIB): The integration path is a circle with the origin of coordinates as the center, and the obtained integration value is $\mathrm{CIB}(\alpha)$, as shown in

$$\mathrm{CIB}(\alpha) = \oint B(\alpha, \theta) d\theta. \tag{10}$$

CIB analysis is time-shift invariant as well as scale-stretch invariant.

Square Integrated Bispectra (SIB): The integration path is a rectangle centered at the origin of coordinates, and the obtained integration value is $\mathrm{SIB}(\omega)$, as shown in

$$\mathrm{SIB}(\omega) = \oint_{S_l} B(\omega_1, \omega_2) d\omega_1 d\omega_2. \tag{11}$$

The SIB analysis is time-shift invariant, phase-invariant, and scale-stretch invariant. Based on the above theoretical analysis and derivation, Algorithm 2 demonstrates the computation and extraction of bispectrum-based RFF designed in this paper.

The block diagram of the system based on the integral bispectrum method is the same as that of the power spectrum-based method and is not repeated here. In this paper, SIB is used as the feature extraction algorithm of choice.

4.3. Complex-ResNet-Based WIFI Device Recognition Algorithm. With the rise of data-driven algorithms, intelligent detection and recognition technologies have been greatly developed. Deep neural networks can learn different levels of features from data in an autonomous way, which has great advantages in the face of structured information and massive data. In this paper, we will study the performance of residual neural network model technology applied to WIFI device recognition, analyze its basic theory, and introduce its complex processing form to investigate its feasibility in the field of WIFI device recognition.

Wireless RF signals propagate in free space in the form of electromagnetic waves, and what is generally obtained at the receiving end is the complex form of the signal, with instantaneous amplitude and phase information. In signal processing, the real and imaginary parts of the complex RF signal are taken as the in-phase and quadrature signal taps, respectively. In the real number processing neural network model, the amplitude information of the signal is generally taken as the input data for processing. However, the phase information of the signal is also a very important RF signal feature, so it is not enough to use the real number neural network model for EM signal processing, but the model input and related processing of the neural network need to be modified to adapt to the wireless RF complex signal processing in the field of RF fingerprinting technology. Here, this subsection introduces the relevant processing of the complex neural network model and its difference from the real number model and constructs the complex ResNet model for WIFI device identification.

Complex convolution layer: The processing of the complex convolution layer is also performed by the convolution kernel, unlike the real convolution kernel, and the complex convolution kernel based on the complex convolution is used here.

Assuming that the input data is a complex signal vector $\vec{s} = \vec{x}_I + i \cdot \vec{x}_Q$, the convolution operation is performed through the complex weight matrix of the complex convolution kernel as shown in

$$\mathbf{W} \otimes \vec{s} = \left(\mathbf{A} \otimes \vec{x}_I - \mathbf{B} \otimes \vec{x}_Q\right) + i \cdot \left(\mathbf{B} \otimes \vec{x}_I + \mathbf{A} \otimes \vec{x}_Q\right), \tag{12}$$

where the complex weight matrix is $\mathbf{W} = \mathbf{A} + i \cdot \mathbf{B}$, and $\vec{x}_I$, $\vec{x}_Q$ is a vector of real numbers, and $\mathbf{A}$, $\mathbf{B}$ is a matrix of real numbers.

Complex fully-connected layer: The fully-connected layer of the traditional real neural network model is used to recognize the probabilistic output, and its input is the real feature map. In the complex neural network model, it is modified to a complex fully connected layer to maximize the use of the output features of the complex convolutional layer. The processing of the complex fully connected layer is based on the real fully connected layer, with the difference that four

```
Input: Signal, $s[n]$; signal points, $N$; FFT points, $L$;
Output: Bispectrum RF fingerprint features, $RFF_{BS}$;
The received signal $s[n]$ is normalized to obtain $s'[n]$.
The bispectrum of $s'[n]$ is obtained according to Equation (7).
    $B(\omega_1, \omega_2) = 1/K\sum_{K}^{i=1} B_i(\omega_1, \omega_2)$
Calculate the SIB of bispectrum estimation is used as the bispectrum RF fingerprint feature
    $RFF_{BS} = \oint_{S_l} B(\omega_1, \omega_2) d\omega_1 d\omega_2$
returnRFF_BS.
```

Algorithm 2: RFF calculation and extraction based on bispectrum analysis.

Table 1: Complex-ResNet and parameter setting.

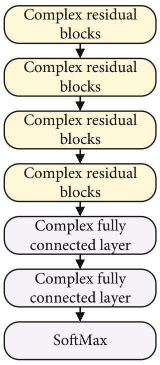
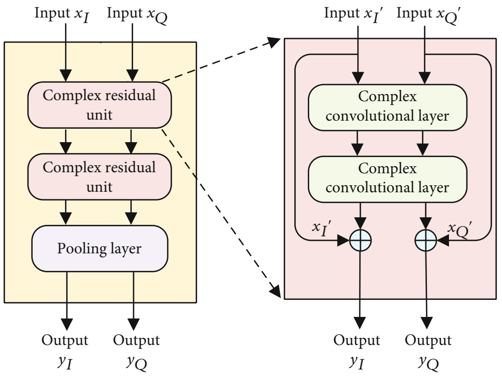
Model structure	Active layer	Output size	Number of parameters
	Complex residuals module	(None,512,32)	10432
	Complex residuals module	(None,256,32)	10912
	Complex residuals module	(None,128,32)	10912
	Complex residuals module	(None,64,32)	10912
	Complex fully connected layer	(None,256)	262400
	Complex fully connected layer	(None,256)	33024
	SoftMax layer	(None,200)	25800

Table 2: Complex residual model and parameter setting.

Complex residual block structure	Active layer	Kernel parameters	Output size	Number of parameters
	Complex convolutional layers	(16,1)	(None,1024,32)	64
	Complex convolutional layers	(16,5)	(None,1024,32)	2592
	Complex convolutional layers	(16,5)	(None,1024,32)	2592
	Complex convolutional layers	(16,5)	(None,1024,32)	2592
	Complex convolutional layers	(16,5)	(None,1024,32)	2592
	Pooling layer	2	(None,512,32)	0

product operations are required to achieve a more accurate probability estimation. The essence of the operation of the complex fully-connected layer is also the convolution operation of the tensor, so its calculation formula is the same as Equation (12) of the complex convolution layer, with the difference that its weights are two real values instead of a real matrix.

After the above complex adaptation of the convolutional and fully connected layers, the complex ResNet model can be built. The structure and parameters of the complex ResNet model constructed in this paper are shown in Table 1.

Among them, the relevant parameters of the complex residual module are shown in Table 2.

5. Results and Analysis

This section conducts relevant experiments using the dataset introduced in Section 3 to verify the performance of the RF fingerprinting system discussed in this paper. The relevant

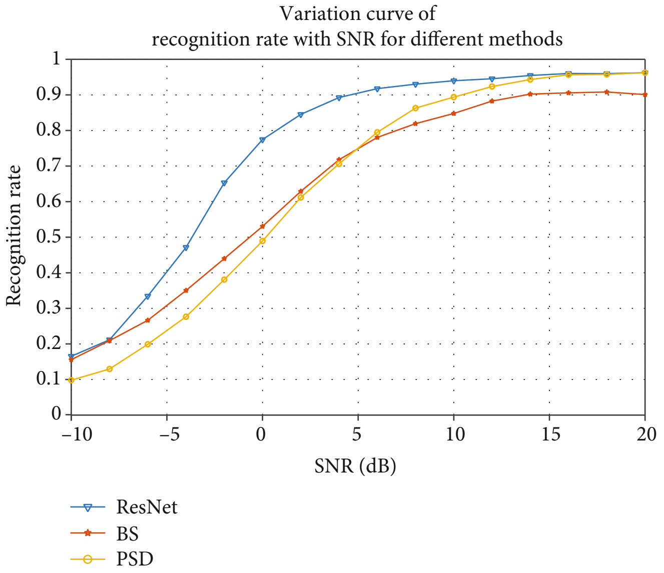


Figure 5: Recognition accuracy with SNR curve.

Table 3: Performance comparison of feature engineering methods in LoS and NLoS scenarios.

Method	Channel environment	Recognition accuracy	Signal length
PSD	LoS	96.4%	4096 points
BS	LoS	89.3%	4096 points
ResNet	LoS	96.7%	1024 points
PSD	NLoS	89.1%	4096 points
BS	NLoS	86.9%	4096 points
ResNet	NLoS	92.6%	1024 points

experimental analysis is performed specifically according to the data acquisition environment, including the analysis of the system's resistance performance to Gaussian white noise with different Signal-to-Noise Ratio (SNR) levels, the analysis of the adaptability to the channel environment, and the analysis of the user capacity of the features.

5.1. System Noise Immunity Analysis. The actual communication system is bound to be affected by various noises in the communication, and these noises will not only affect the communication quality but also cause the effects of feature diffusion and feature degradation in the field of RF fingerprint technology, which makes the performance of the identification system degrade. Therefore, in the test of the system, the noise immunity performance test is a very important test, and the experimental analysis of the noise immunity performance of the RF fingerprint features proposed in this paper is as follows.

In order to test the noise immunity performance of the RF fingerprint system, the dataset introduced in Section 3 is selected for the relevant experiments in this section, with a signal sampling rate of 40 M and a data acquisition environment of a laboratory LoS channel environment containing 100 classes of targets with 100 sample signal data for each class of targets. The ratio of training set to test set of KNN classifier is adjusted to 4 : 1, where the K value is set to 5. The noise interference environment is simulated by adding Gaussian white noise with different SNR levels to this dataset, and the testing of WIFI device recognition system is carried out, and its recognition accuracy under different SNR can reflect the noise immunity performance of the system.

The SNR variation of the signal after noise addition ranges from -10 dB: 2: 20 dB, with a total of 16 levels. The accuracy of the final system recognition with SNR variation curve is shown in Figure 5.

As can be seen from Figure 5, when the SNR is between -10 dB and 5 dB, the BS feature has a maximum improvement of about 8% compared with the PSD feature, and the recognition accuracy increases from 10% to 18.5% at -10 dB. When the SNR is higher than 5 dB, the PSD features have more performance, and the PSD features can reach 95% accuracy when the SNR is 20 dB.

In addition, the complex ResNet can obtain better performance compared to the feature engineering method. For example, when the SNR is 0 dB, the complex ResNet can achieve 78% recognition accuracy, while the feature engineering method can only achieve 50% recognition accuracy.

5.2. System Environmental Adaptability Analysis. In addition to noise, the complex and variable channel environment is also a major problem affecting the wireless signal quality. Also, the harsh channel environment can affect the stability of RF fingerprint features, and the drift of features can lead to poor recognition. In WIFI practical application scenarios, this paper considers laboratory indoor scenarios, which mainly include indoor LoS channels and NLoS channels. The experimental analysis of the environmental adaptation of the RF fingerprint features proposed in this paper is as follows.

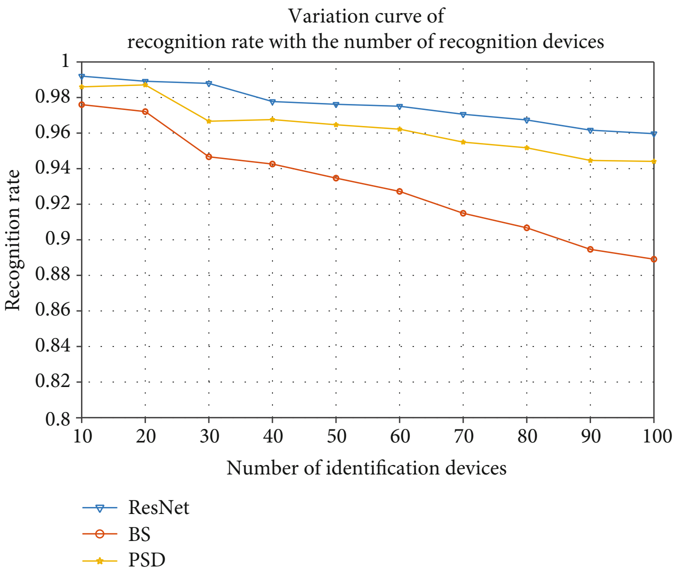


Figure 6: System recognition accuracy curve with the number of devices to be recognized.

In order to test the environmental adaptability of the RF fingerprint system, the dataset introduced in Section 3 is selected for relevant experiments in this section, both of which have a signal sampling rate of 40 M, and the data acquisition environments are laboratory LoS and laboratory NLoS channel environments, respectively, containing 100 classes of targets with 100 sample signal data for each class of targets. The ratio of training set and test set of KNN classifier is adjusted to 4 : 1, where the K value is set to 5. The experiment of WIFI device recognition based on RF fingerprint is conducted to test the recognition accuracy of the system under different channel environments.

The deep learning model constructed in this paper is used to conduct WIFI device recognition experiments in two channel environments using both PSD and BS features, and the results obtained are shown in Table 3.

The introduction of deep learning models such as ResNet will increase the environmental adaptability of the system. As can be seen from Table 3, deep learning methods have 1%-3% improvement in recognition accuracy over feature engineering methods under LoS and NLoS channel conditions. Complex-ResNet can achieve 92.6% recognition accuracy in NLoS scenarios, which has the best performance. And using 1024 points for recognition, it has higher efficiency in practical applications. This indicates that the deep learning method is more capable of automatically extracting signal features, and its recognition is better.

5.3. System User Capacity Analysis. After proposing an identification or device authentication system, its user capacity is also one of its important parameters. The user capacity can measure the maximum number of devices to be identified or authenticated that the system can accommodate to be online at the same time. When the number of online devices exceeds the system capacity, it will make the system unable to work properly and lead to system problems such as identification error and authentication failure. The experimental analysis of the user capacity of the RF fingerprint feature proposed in this paper is as follows.

In order to perform user capacity testing of the RF fingerprinting system, the dataset introduced in Section III is selected for relevant experiments in this section, with a signal sampling rate of 40 M and a data acquisition environment of a laboratory LoS channel environment containing 100 classes of targets with 100 sample signal data for each class of targets. WIFI device recognition experiments are conducted on this dataset to test the recognition accuracy of the system in a large number of device scenarios and to analyze the trends.

Feature engineering-based RF fingerprint feature extraction recognition tests are performed using four features. The final system recognition accuracy variation curve with the number of devices to be recognized is shown in Figure 6.

Analysis of Figure 6 shows that for a user capacity size with 96% recognition accuracy, the PSD feature can reach about 60 devices and the BS feature can reach about 25 devices. For a user capacity size with 90% recognition accuracy, the PSD feature can reach more than 100 devices and the BS feature can reach about 85 devices. And ResNet and other deep learning models can all reach more than 100 devices. This shows that deep learning methods have a greater potential. Compared with the article [35], the channel we use is more complex. Although the recognition accuracy decreases with the increase of the number of recognitions, it still remains at a high level, which shows the superiority of the method we use.

6. Conclusion

RF fingerprinting technology, as one of the important development directions to improve the security performance of wireless devices, has been widely studied in the past two decades. In this paper, we first construct a large-scale measured WIFI signal dataset with the characteristics of a large number of devices, full channel scenarios, and a large data scale. And two methods based on feature engineering and deep learning are proposed for individual identification of WIFI devices. The comprehensive performance simulation results described in this paper can achieve better recognition accuracy both in the darkroom and in the laboratory, both in LoS environment and in NLoS environment, and after adding Gaussian white noise with a certain SNR.

Feature engineering-based RF fingerprint recognition methods are fast in model training, flexible and convenient, and easy to deploy. However, the traditional feature-based WIFI signal recognition methods are increasingly difficult to cope with the new characteristics of big data due to their need for a large amount of expert knowledge and limited data processing capacity. The deep learning-based methods show better performance in EM signal recognition due to their powerful nonlinear fitting ability and end-to-end learning mode.

However, through the research and analysis in this paper, it is found that there are still some problems in the field of RF fingerprint-based individual identification technology, such as the lack of a mathematical model for the forward derivation of the RF fingerprint generation process, i.e., the inability to quantify the differences in RF fingerprint features between devices; the poor adaptability and low extraction efficiency of RF fingerprints; and the privacy issues and possible attacks. We will conduct research and analysis in our future work.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work is supported by the Fundamental Research Funds for the Central Universities (3072021CF0801).

References

[1] I. Ahmad, T. Kumar, M. Liyanage, J. Okwuibe, M. Ylianttila, and A. Gurtov, "Overview of 5G security challenges and solutions," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 36–43, 2018.

[2] X. S. Ji, K. Z. Huang, L. Jin et al., "Overview of 5G security technology," *Science China-Information Sciences*, vol. 61, no. 8, pp. 107–131, 2018.

[3] M. Wang, Y. Lin, Q. Tian, and G. Si, "Transfer learning promotes 6G wireless communications: recent advances and future challenges," *IEEE Transactions on Reliability*, vol. 70, no. 2, pp. 790–807, 2021.

[4] Z. Bao, Y. Lin, S. Zhang, Z. Li, and S. Mao, "Threat of Adversarial Attacks on DL-Based IoT Device Identification," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 9012–9024, 2022.

[5] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: modeling and validation," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2091–2106, 2016.

[6] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir, "Fingerprinting Wi-Fi devices using software defined radios," in *9th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*, pp. 3–13, Darmstadt Germany, 2016.

[7] Q. Tian, J. C. Jia, and C. B. Hou, "Research on fingerprint identification of wireless devices based on information fusion," *Mobile Networks & Applications*, vol. 25, no. 6, pp. 2359–2366, 2020.

[8] R. Duan, Z. Li, and Y. Yin, "Improvement of LANDMARC indoor positioning algorithm," *International Journal of Performability Engineering*, vol. 16, no. 3, pp. 446–453, 2020.

[9] M. Liu, C. Liu, M. Li, Y. Chen, S. Zheng, and N. Zhao, "Intelligent passive detection of aerial target in space-air-ground integrated networks," *China Communications*, vol. 19, no. 1, pp. 52–63, 2022.

[10] X.-H. Ru, Z. Liu, W.-L. Jiang, and Z.-T. Huang, "Recognition performance analysis of instantaneous phase and its transformed features for radar emitter identification," *Iet Radar Sonar and Navigation*, vol. 10, no. 5, pp. 945–952, 2016.

[11] T. D. Ridder and R. M. Narayanan, "Radar detection performability under graceful degradation," *International Journal of Performability Engineering*, vol. 17, no. 8, pp. 666–675, 2021.

[12] A. C. Polak, S. Dolatshahi, and D. L. Goeckel, "Identifying wireless users via transmitter imperfections," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 7, pp. 1469–1479, 2011.

[13] J. W. Zhang, F. G. Wang, O. A. Dobre, and Z. D. Zhong, "Specific emitter identification via Hilbert-Huang transform in single-hop and relaying scenarios," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1192–1205, 2016.

[14] J. Hall, M. Barbeau, and E. Kranakis, "Detection of transient in radio frequency fingerprinting using signal phase," *Wireless and Optical Communications*, pp. 13–18, 2003.

[15] B. Danev, D. Zanetti, and S. Capkun, "On physical-layer identification of wireless devices," *ACM Computing Surveys*, vol. 45, no. 1, pp. 1–29, 2012.

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

[17] X. Li, J. Liu, B. Ding, Z. Li, H. Wu, and T. Wang, "TickSEC: a novel reconfigurable platform for WIFI physical layer security," in *2018 International Conference on Networking and Network Applications (NaNA)*, pp. 237–244, Xi'an, China, 2018.

[18] G. Li, J. Yu, Y. Xing, and A. Hu, "Location-invariant physical layer identification approach for WIFI devices," *IEEE Access*, vol. 7, pp. 106974–106986, 2019.

[19] K. Yang, J. Kang, J. Jang, and H. Lee, "Multimodal sparse representation-based classification scheme for RF fingerprinting," *IEEE Communications Letters*, vol. 23, no. 5, pp. 867–870, 2019.

[20] Y. Lin, J. Jia, S. Wang, B. Ge, and S. Mao, "Wireless device identification based on radio frequency fingerprint features,"

in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, 2020.

[21] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1091–1095, 2020.

[22] A. C. Polak and D. L. Goeckel, "RF fingerprinting of users who actively mask their identities with artificial distortion," in *2011 Conference record of the forty fifth Asilomar conference on signals, Systems and Computers (ASILOMAR)*, pp. 270–274, Pacific Grove, CA, USA, 2011.

[23] T. Zheng, Z. Sun, and K. Ren, "FID: function modeling-based data-independent and channel-robust physical-layer identification," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, pp. 199–207, Paris, France, 2019.

[24] Y. Lin, H. Zhao, X. Ma, Y. Tu, and M. Wang, "Adversarial attacks in modulation recognition with convolutional neural networks," *IEEE Transactions on Reliability*, vol. 70, no. 1, pp. 389–401, 2021.

[25] Y. Lin, Y. Tu, Z. Dou, L. Chen, and S. Mao, "Contour stella image and deep learning for signal recognition in the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 34–46, 2021.

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

[27] M. Liu, Z. Liu, W. Lu, Y. Chen, X. Gao, and N. Zhao, "Distributed few-shot learning for intelligent recognition of communication jamming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 395–405, 2022.

[28] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.

[29] J. M. McGinthy, L. J. Wong, and A. J. Michaels, "Groundwork for neural network-based specific emitter identification authentication for IoT," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6429–6440, 2019.

[30] Y. Dong, X. Jiang, H. Zhou, Y. Lin, and Q. Shi, "SR2CNN: zero-shot learning for signal recognition," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2316–2329, 2021.

[31] Y. Tu, Y. Lin, H. Zha et al., "Large-scale real-world radio signal recognition with deep learning," *Chinese Journal of Aeronautics*, vol. 16, pp. 1–14, 2021.

[32] Y. Lin, Y. Tu, and Z. Dou, "An Improved Neural Network Pruning Technology for Automatic Modulation Classification in Edge Devices," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5703–5706, 2020.

[33] S. Zhang, L. Yun, T. Ya, and S. Mao, "Electromagnetic signal modulation recognition technology based on lightweight deep neural network," *Journal on Communications*, vol. 41, no. 11, pp. 12–21, 2020.

[34] M. Liu, J. Wang, N. Zhao, Y. Chen, H. Song, and R. Yu, "Radio frequency fingerprint collaborative intelligent identification using incremental learning," *IEEE Transactions on Network Science and Engineering*, pp. 1–11, 2021.

[35] 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 Communications Letters*, vol. 11, no. 3, pp. 543–547, 2022.