Cureus Journal of Computer Science

Open Access Original Article

Radio Frequency Fingerprinting Method Based on Squeeze-and-Excitation Attention Multisource Domain Adversarial Network

Isack Bulugu [1]

1. Electronics and Telecommunications Engineering, University of Dar es Salaam, Dar es Salaam, TZA

Corresponding author: Isack Bulugu, bulugu@udsm.ac.tz

Received 02/16/2025
Review began 03/10/2025
Review ended 05/19/2025
Published 05/23/2025

DOI: https://doi.org/10.7759/s44389-025-03341-w

Abstract

The radio frequency (RF) fingerprint uses the hardware features of the RF front-end as identifiers to identify the device. To solve the problem of existing RF fingerprinting methods ignoring the interference of receiver hardware features, which results in poor generalization of the model to different receiver devices, an RF fingerprinting method based on squeeze-and-excitation (SE) attention multisource domain adversarial network is proposed. In this method, adversarial training is carried out using multiple source domain labeled data and a small amount of target domain labeled data to extract features that are not related to the receiver domain. The SE attention mechanism is integrated to enhance the model's learning ability for the transmitter's RF fingerprint features. Combined with a very small amount of labeled data in the target domain, the model parameters are fine-tuned to further improve the recognition performance of the transmitter. Experimental results on the WiSig public dataset show that the proposed method can effectively identify the sender device in the cross-receiver scenario, with an average accuracy of 83.3%. The average accuracy can be further improved to 93.4% after fine-tuning with a small amount of labeled data.

Categories: Network Security, IoT Security and Privacy, Machine Learning (ML)
Keywords: rf fingerprinting, multisource domain confrontation, deep learning, physical layer security, se attention mechanism

Introduction

With the rapid development of Internet of Things (IoT), wireless devices have become inseparable from people's lives. However, the broadcast characteristics of wireless transmission enable malicious users to access the network by impersonating legitimate user identities, so device authentication is essential to protect the security of IoT [1]. Owing to the limited computing resources of network terminal equipment, designing a lightweight authentication scheme for IoT devices is urgent [2].

In recent years, the rise of radio frequency fingerprint identification (RFFI) technology has provided a new idea for lightweight security authentication of IoT terminals [3,4]. The radio frequency (RF) fingerprint uses the tolerance effect generated by the hardware production process of wireless devices to build a unique RF fingerprint for each device to achieve effective identification of the identity of the docking device, which has been successfully applied to network intrusion detection [5], aircraft identification [6], and other civil and military fields. Compared with traditional key encryption technology, RF fingerprinting technology has two advantages. First, RF fingerprint recognition takes the inherent characteristics of RF devices of wireless devices as an authentication element, without any requirement for the computing power of the device, and avoids the problem of safe storage of keys. Second, with the rapid development and widespread application of wireless technologies such as Wi-Fi, LoRa, and ZigBee, wireless communication modules have become indispensable parts of most digital terminals. Therefore, RF fingerprint recognition technology has strong universality and compatibility, and it can be quickly applied in IoT scenarios without adding additional functional modules to the device.

The existing research on RF fingerprinting can be divided into methods based on expert feature engineering and methods based on deep learning [7]. The methods based on expert feature engineering need to manually extract signal-related features, such as the carrier frequency offset (CFO) [8], I/Q offset [9], phase error [9], Hilbert spectrum [10], etc., and then use the classifier for device identification. However, methods based on expert feature engineering are highly dependent on signal feature extraction algorithms, and in some cases, accurate estimation of individual features is challenging.

In recent years, with the successful application of deep learning in various fields, neural network models have gradually become mainstream RF fingerprinting models [11]. Riyaz et al. [12] proposed the use of a convolutional neural network (CNN) as an RF fingerprint feature classifier and experimentally proved that CNNs have higher device recognition accuracy than traditional machine learning algorithms such as support vector machines and logistic regression. Considering that wireless signals themselves are

represented by complex numbers, researchers from the University of California [13] designed a special complex neural network to perform deep feature learning and device identification on the original sampled signals and classified 19 Wi-Fi devices, obtaining more than 80% classification accuracy. Shen et al. [14] used spectrograms to represent the fine-grained time–frequency characteristics of LoRa signals and reported that the spectrogram GCNN model was the best model for LoRa device classification, with an accuracy of 96.40%, the lowest complexity, and the shortest training time. Experiments have shown that the CFO will drift, resulting in classification error, and that CFO compensation is an effective mitigation measure. To solve the problem of RF fingerprint recognition under the condition of a low signal-to-noise ratio, Du et al. [15] proposed a residual network algorithm based on channel attention and used a dynamic threshold method to suppress noise components. Gu et al. [16] proposed a novel dual-attention convolution module, which can learn both channel attention and spatial attention to adjust the RF fingerprint features, thereby enhancing the feature learning ability of the convolutional layer, and the experimental results show that the RF fingerprinting performance can be improved by 1.5% on average. Jiang et al. [17] fused the features of the spectrogram and the differential constellation trajectory map, added the channel attention force mechanism to improve the feature extraction ability, and optimized and reduced the spike neural network. The power consumption of the RFFI system was used by Weng et al. [18] to separate the signal samples with different pulse waveform distributions, feed them into the recognition network by using prior information, extract and merge the feature maps contained in multiple data blocks, and design a spatial attention mechanism for the low-dimensional discrete signals, which enhances the pertinence of model learning and achieves a recognition accuracy of 98.20%.

Recent developments in RF fingerprinting have also leveraged transformer architectures and advanced domain adaptation strategies. For instance, Du et al. [15] introduced a residual network under noisy conditions, and Gu et al. [16] proposed a dual-attention CNN architecture that combines channel and spatial attention to improve fingerprint discrimination. Furthermore, methods such as domain-adversarial neural networks (DANNs) and maximum mean discrepancy have been employed in general domain adaptation but are rarely applied in RF scenarios. To benchmark the proposed model fairly, comparisons are included with DANN and a dual-attention CNN variant (DA-CNN) under consistent experimental settings using the WiSig dataset.

However, all of the above studies assume that the same receiver is used in the training and testing phases and how to make the model learn more robust and discriminative RF fingerprinting to improve the recognition accuracy of the model without considering the influence of receiver hardware interference on the RF fingerprinting model. Zhang et al. [19] considered the interference of receiver RF devices on RFFIs, systematically modeled the damage of transmitters and receivers in narrowband systems, and carried out comprehensive experiments and simulations to verify the impact on RFFIs. Hanna et al. [20] constructed a large-scale Wi-Fi dataset and used the CNN model to train the data of multiple receivers and transmitters, which improved the generalization ability of the model to identify transmitters on different receivers; however, this method requires a large amount of receiver data for training and is not universal.

To solve the above problems, this paper proposes an RF fingerprinting method based on a squeeze-and-excitation (SE) attention multisource domain anti-immunity network, which aims to reduce the dependence of the recognition model on known receivers and improve its robustness in the cross-receiver scenario of the IoT. On the basis of multisource domain adaptive learning [21] and unsupervised adversarial domain adaptive thinking [22], adversarial training is carried out via labeled data from multiple source domain receivers and a small amount of unlabeled data from target domain receivers to extract the receiver domain invariance characteristics. Moreover, the SE attention mechanism is added to the feature extractor part of the model to improve the feature learning ability of the model. Furthermore, only a small amount of labeled data from the target domain receiver needs to be introduced to fine-tune the model parameters, which can greatly improve the recognition performance of the model in the target domain receiver. The experimental results on the WiSig public dataset show that the average accuracy of the RF fingerprint recognition model proposed in this paper can reach 83.3% in the cross-receiving airport scenario, and the average accuracy can be further increased to 93% after fine-tuning the model with a small number of labeled samples added to the target domain receiver.

Materials And Methods

The IoT system often has a plurality of terminal devices (transmitters) and gateway devices (receivers), and RFFI aims to identify the IoT terminal devices in the wireless network by analyzing the physical layer signals received by the IoT gateway. To obtain a model with strong generalization performance and improve its recognition performance in the cross-receiver scenario of the IoT, this paper proposes an RF fingerprinting method based on an SE attention multisource domain adversarial network, which mainly includes three stages: a training stage, a fine-tuning stage, and a recognition stage, as shown in Figure *1*.

1) Training phase: The data of multisource receivers and a small number of target domain receivers are jointly input into the neural network model for training, and the neural network is guided to learn features that are not related to the receiver to improve the recognition performance of the model in the target domain receiver.

2) Fine-tuning phase: A small amount of labeled data from the receiver in the target domain is used to fine-tune the model parameters. Fine-tuning is an optional operation that can be used when the network model does not perform well on the receiver in the target domain.

3) Identification phase: The trained network model is used to perform RF fingerprinting on the receiver test set signal in the target domain, and the classification and recognition results are obtained.

In the above three stages, the data preprocessing operations are carried out on the datasets, including channel equalization and power normalization. Channel equalization operations are used to mitigate the impact of wireless channels on signals to eliminate channel-induced distortion and interference. The use of power normalization prevents the neural network from identifying the device on the basis of the received power.

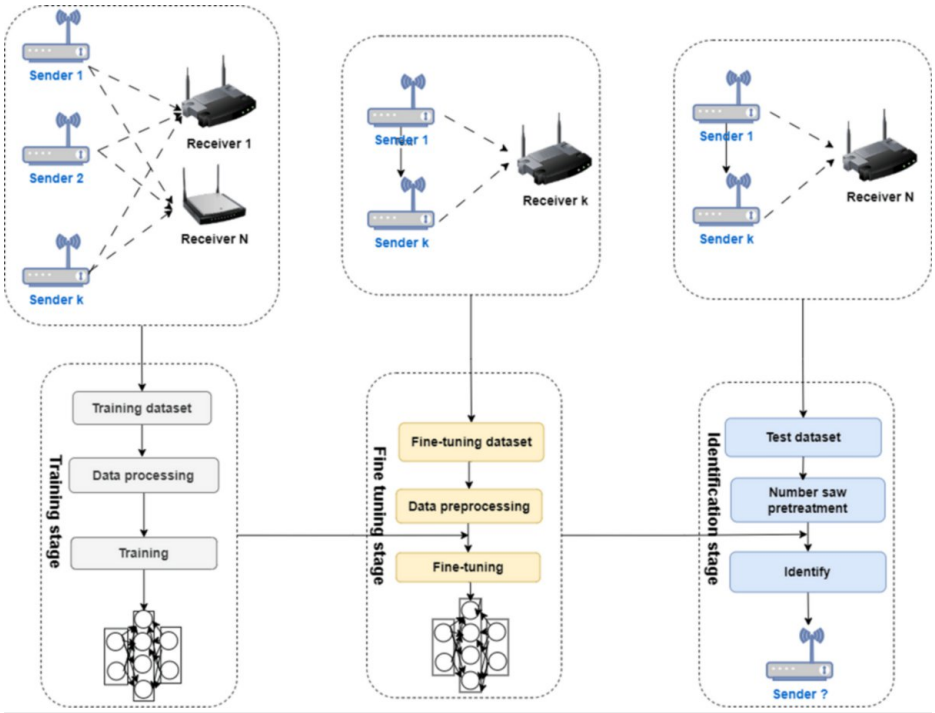

FIGURE 1: Overall process

Proposed RF fingerprinting model

In the field of deep learning, implementing cross-domain tasks is a challenging task [23]. However, in the RF fingerprint classification and identification task, this method requires the target domain receiver to provide high-quality labeled data. IoT systems often have multiple gateway nodes, and each gateway is unlikely to collect wireless signals from all end devices. To solve this difficulty, this paper draws on the ideas of multisource domain adaptive learning and unsupervised adversarial domain adaptation and proposes an RF fingerprinting model based on an SE attention multisource domain adversarial network, which uses the labeled data of multiple source domain receivers and a small amount of unlabeled data of the target domain receiver to jointly train the neural network model and extracts the shared features between the source domain and the target domain through adversarial training to improve the sender recognition performance of the model in the target domain. The overall structure of the network model is shown in Figure 2, which mainly includes three parts: a feature extractor, sender classifier, and receiver domain classifier. The specific functions of each part are as follows.

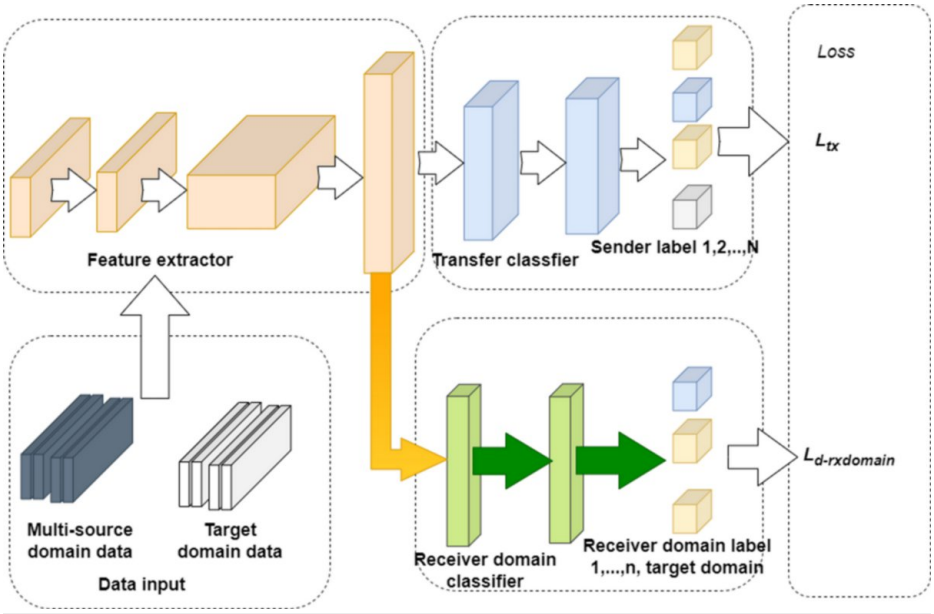

FIGURE 2: Structure of the network model

Feature Extractor

This module is one of the core components of the multisource domain adversarial network, and its main function is to extract useful feature representations from the input data, which are passed to the sender classifier and receiver domain classifier for training at the same time. In this paper, the SE attention mechanism is added as a feature extractor in the residual network.

As a deep neural network, the residual network can effectively alleviate the problems of gradient vanishing and gradient explosion in deep neural networks by introducing residual connections, allowing the training of deeper network layers and reducing the number of network parameters [24].

However, in the complex electromagnetic environment, the RF fingerprint features of the transmitter are weak, and it is difficult for the network to extract effective features from the signal to identify the transmitter device, whereas studies in the field of computer vision have shown that the attention mechanism can effectively enhance the feature learning ability of the deep neural network, such as the SE attention mechanism [25], convolutional block attention module (CBAM) [26], and coordinate attention (CA) mechanism [27]. In computer vision tasks, since most of the data objects to be processed are image data, the attention mechanism can be regarded as a dynamic selection process, which is implemented by adaptively weighting features according to the importance of the input data. However, in the RF fingerprinting task, the processed data object is an I/Q signal sample, and the dimension of the data is 256 × 2, where 256 represents the time series length of the signal and 2 represents the in-phase(I) and quadrature(Q) dimensions of the signal. It is difficult to measure the importance of RF features in the spatial dimension of the signal because all I/Q signal samples contain RF features due to hardware defects, and the time series of the signals are correlated. However, not all RF features have the same weight for the recognition results, and the SE attention mechanism can selectively emphasize the information-rich channel features and suppress the channel features that are not very useful, and the number of parameters is relatively small. Considering the limited computing resources of IoT devices, the SE attention mechanism is introduced in this paper, and the overall process of the SE attention mechanism is shown in Figure 3.

The specific working mechanism is composed of the following three steps:

1) Squeeze: For any input characteristic tensor $F\varepsilon C \times H \times W$, the channel dimension of the input eigentensor is reduced to 1 through the global average pooling operation; that is, an eigentensor of size $1 \times 1 \times C$ is obtained.

2) Excitation: The compressed eigentensor is nonlinearly transformed through a two-layer fully connected network. First, the layer 1 fully connected network is activated via the *ReLU* function, and then the layer 2 fully connected network is activated via the sigmoid activation function. These two layers of fully connected networks act as incentives to map global information to the weights of each channel.

The rationale behind integrating the SE attention mechanism lies in its ability to adaptively recalibrate channel-level features based on their global contextual importance. In RF fingerprinting tasks, not all extracted signal features contribute equally to transmitter discrimination - some frequency components or modulation artifacts may be more distinctive due to hardware imperfections. The SE module performs a global average pooling across each feature map channel to capture a global descriptor, then passes this through a bottleneck of two fully connected layers with non-linear activations to generate per-channel weights. These weights emphasize channels that are more informative for classification while suppressing less relevant ones. Unlike spatial attention mechanisms that require complex spatial modeling, the SE approach maintains low computational overhead, making it well-suited for RF signal analysis where channel-wise variations often reflect meaningful device-specific characteristics, especially in the I/Q time series representation. This mechanism enhances the model's capacity to focus on transmitter-distinctive RF features even under domain shift introduced by receiver variability.

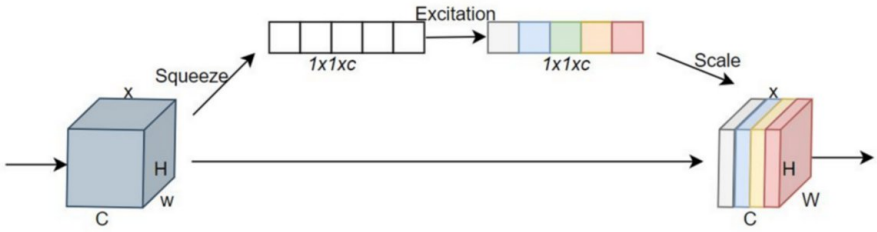

FIGURE 3: Schematic diagram of squeeze-and-excitation attention

The residual element of the SE attention mechanism is shown in Figure 4, where $C \times H \times W$ represents the input characteristics of the residual element, $F(x)$ is the output obtained after residual learning, y is the output of the last whole residual element, batch normalization (BN) is the batch normalization layer, and SE is the attention mechanism module.

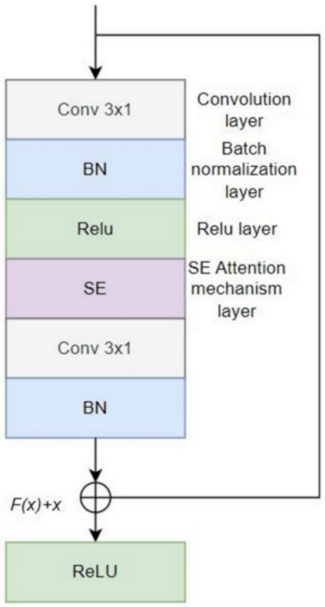

FIGURE 4: Attention residual block

BN, Batch Normalization; SE, Squeeze-and-Excitation

The network structure of the whole feature extractor is listed in Table *1*. First, the preprocessed I/Q signal samples were input into the model for training, with a size of 256 × 2. The first convolutional layer is composed of 64 convolutional kernels, each of which is 3 × 1 in size, in which each convolutional kernel learns the changes in three samples in the I/Q dimensions over time to generate 64 different feature maps. Second, two residual layers and two pooling layers with an SE attention mechanism are added to extract deeper RF features, which enables the network to pay more attention to the effective RF features in the I/Q signal samples and suppress the features that are not useful to improve the model's ability to identify and select important features. Third, the second convolutional layer consists of 64 convolutional kernels of size 3 × 2, and each convolutional kernel learns the changes in three activation values in the I/Q dimensions of the eigengraph obtained via the above convolution operation. Moreover, each convolution operation is followed by a nonlinear *ReLU* activation function and a batch normalization layer BN. The *ReLU* performs a predetermined nonlinear transformation for each neuron of the convolutional output, enhancing the expressive ability of the network model. BN normalizes the data of each layer of the network model, which makes the data distribution more stable and helps to improve the training efficiency and generalization ability of the network. Then, the deeper RF features are extracted again through two residual layers with the SE attention mechanism added. Finally, the global average pooling layer is used to reduce the spatial dimension of the feature map to 1 × 1 so that the network can capture the global feature information.

Layer	Output dimension	Parameters
Input	(1,256,2)	Input:(1,256,2)
Conv2d+BN+Relu	(64,256,2)	Conv2d:3×1,64, stride 1, padding (1,0)
Res Block 1	(64,256,2)	3×1,643×1,64×1, stride 1, padding (1,0)
MaxPool2d	(64,128,2)	2×1 max pool
Res Block 2	(128,128,2)	3×1,1283×1,128×1, stride 1, padding (1,0)
MaxPool2d	(128,64,2)	2×1 max pool
Conv2d+BN+Relu	(64,64,1)	Conv2d:3×2,64, stride 1, padding (1,0)
Res Block3	(128,64,1)	3×1,1283×1,128×1, stride 1, padding (1,0)
MaxPool2d	(128,32,1)	2×1 max pool
Res block4	(128,32,1)	3×1,1283×1,128×1, stride 1, padding (1,0)
Average_pool	(128,1,1)	AdaptiveAvgpool2d

TABLE 1: Network structure of the feature extractor

Transmitter Classifier

The transmitter classifier is the final task classifier used to identify specific transmitters. In this paper, a fully connected layer with *Softmax* is used to classify the transmitters. In the training process, the features learned by the feature extractor are input into the fully connected layer for classification, and the specific formula of the function is as follows:

$$\hat{p}_i = \frac{e^{z_i}}{\sum_{j=1}^{N} Z_i}$$

where Z_i is the output of the i^{th} neuron before the Softmax function is activated, *N* is the total number of sender classes, and is the predicted probability of the class. In this work, cross-entropy is used as the sender classifier loss L_{tx} , and the specific calculation formula is as follows:

$$L_{tx} = -\sum_{j=1}^{N} p_i log\,(\hat{p})$$

where p_i is the true label of the sample, $\hat{p} = [p_0, ..., p_{N-1}]$ is a probability distribution, and *N* is the total number of classes. During the training period, the goal of this paper is to find the feature extractor parameters that minimize L_{tx} to ensure the performance of the sender classifier and obtain better

classification recognition results.

Receiver Domain Classifier

The receiver domain classifier is only used during the training phase, and after the training is completed, the receiver domain classifier is removed because the goal of RFFI is to classify and identify the transmitter device. In the training stage, the task of the receiver domain classifier is to identify the receiver domain label of the sample, the relationship between the receiver domain classifier and the feature extractor is adversarial learning, and the receiver domain classifier cannot distinguish which receiver the sample comes from after adversarial training. The receiver domain classifier needs to improve its own performance and classify samples from different receiver domains according to the features that are not related to the receiver learned from the feature extractor. Moreover, the feature extractor also needs to improve its own performance and learn features that are more irrelevant to the receiver domain to confuse the receiver domain classifier.

Since the relationship between the receiver domain classifier and the feature extractor is adversarial learning, a gradient reversal layer (GRL) $R_\lambda(x)$ is added between the model feature extractor and the receiver domain classifier, and the gradient is negated after the GRL in the backpropagation process to realize adversarial learning between the receiver domain classifier and the feature extractor. The specific expression is as follows:

$$R_\lambda(x) = x$$

$$\frac{dR_\lambda(x)}{dx} = -\lambda I$$

where I is an identity matrix, and λ is an adjustable parameter with a value range of (0,1).

In this work, cross-entropy is used as the receiver domain classifier to lose $L_{d-rxdomain}$, and the specific expression is as follows:

$$L_{d-rxdomain} = -\sum_{j=1}^{n} q_i log(\hat{q}_i)$$

where q_i is the real label of the sample; $\hat{q}_i = [q_0, \ldots, q_{n-1}]$ is a probability distribution; and *n* is the total number of receiver domain categories.

Objective Function

The final objective function of the RF fingerprint recognition model based on the SE attention multisource domain adversarial network proposed in this paper is composed of the loss of the sender classifier and the loss of the receiver domain classifier, and the loss function of the whole network is as follows:

$$LOSS = L_{tx} + \lambda L_{d-rxdomain}$$

$$\lambda = \frac{2}{1 + e^{-\lambda \cdot p}} - 1$$

The parameter λ is an adjustable parameter with a value range of (0,1]; p represents the ratio of the number of current iterations to the total number of iterations; and λ is a constant, set to 10.

In this section, an RF fingerprinting model based on an SE attention multisource domain adversarial network is proposed. In the training of the network model, a labeled source domain dataset, a small number of unlabeled target domain datasets, and receiver domain labels of the source and target domain datasets are used as inputs. In the training process, the model continuously minimizes the loss of the sender classifier for the labeled source domain dataset and maximizes the loss of the receiver domain classifier for the dataset from the source domain and the target domain so that the sender classifier and the receiver domain classifier are opposed to each other during the training process, and finally, the two losses reach an equilibrium state. Through this domain adversarial learning method, the model narrows the feature difference between the source domain data and the target domain data so that the features learned by the feature extractor not only have sender classification recognizability but also receiver domain invariance to achieve receiver domain migration.

On the basis of the above model, this paper proposes a method to fine-tune the model parameters by using

a small amount of labeled data from the receiver in the target domain to reduce the learning cost of the receiver in the target domain as much as possible and improve the recognition performance of the model in the receiver in the target domain. The trained SE attention multisource domain adversarial network model may perform well on most target domain receivers. However, it may not have high recognition accuracy for a very small number of target domain receivers. This problem can be mitigated by fine-tuning. Fine-tuning refers to the process of adapting a pretrained model to a specific task by further training the parameters of the model. Specifically, it is necessary to collect a small amount of labeled data and store them in the fine-tuning dataset using the target domain receiver with poor recognition performance and then use the trained SE attention multisource domain adversarial network model to initialize the network parameters, freeze all the parameters of the feature extractor part, and use the fine-tuning dataset to fine-tune the parameters of the sender classifier part with a low learning rate. By freezing the parameters, the number of training parameters of the model can be reduced, the risk of overfitting can be reduced, and the speed of model training can be accelerated. Through the fine-tuning operation, the model can be guided to learn the data distribution of the receiver in the target domain and improve the recognition performance of the model in the receiver in the target domain.

Results

Dataset and experimental environment

This experiment uses an HP desktop, Windows 10 operating system, Intel(R) Core (TM) i5-10500 CPU64-bit processor, 8 GB of internal memory, and GPU acceleration. The GPU acceleration hardware is an NVIDIA RTX2080, the compilation environment is Python 3.8.7, and the PyTorch version is 1.11.0. The stochastic gradient descent (SGD) method was used to train the model. The initial learning rate was set to 0.0001, the fine-tuned learning rate was 0.00001, the maximum number of training rounds was 40, and the batch size was 32.

The dataset used in this experiment is a subset of the open-source dataset WiSig [21], *ManyRx* , which consists of signals collected by 10 Wi-Fi devices (transmitters) and 32 USRP devices (receivers) in 4 days (within one month), with a sampling frequency of 25 MHz, in which the number of samples corresponding to each pair of transmitters and receivers per day is 200. To reduce the impact of channel changes, only one day's data from the *ManyRx* dataset were used. According to the ratio of 9:1 , the source domain dataset is randomly divided into a training set and a test set, the target domain dataset randomly selects 30 samples of each transmitter device for training, and the remaining samples are tested.

Evaluation indicators

To more intuitively represent the detection performance of the model, a common confusion matrix in binary classification is constructed, as listed in Table *2*. In this paper, the four indicators of *accuracy*, *precision*, *recall* and *F1* are selected as the overall evaluation indicators of the model, and the specific formula is as follows.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

	Predicted positive sample	Predicted negative sample
Actual positive sample	True positive (TP)	False negative (FN)
Actual negative sample	False positive (FP)	True negative (TN)

TABLE 2: Binary confusion matrix

Analysis of experimental results

Comparison of the Performances of Different Methods

Since the existing RF fingerprinting studies are all single-receiver RF fingerprinting studies and few cross-receiver RF fingerprinting studies exist, this paper compares the SE attention multisource domain adversarial network model with the following three methods.

1) RF fingerprinting based on traditional deep learning models

The traditional RF fingerprinting methods based on deep learning assume that the same receiver device is used in the training and testing stages, so the traditional deep learning model only needs to complete the sender classification task of a single receiver. In this paper, the most advanced model [17] in this field is selected, which adds channel and spatial attention mechanisms on the basis of the CNN. For narrative convenience, this model is abbreviated as C-CNN.

2) RF fingerprinting based on multisource domain learning

To improve the generalization performance of the model on different receivers, Xie et al. [21] use a CNN model to train the data of multiple source domain receivers. For narrative convenience, this model is abbreviated as MR-CNN.

3) RF fingerprinting based on adversarial learning

Inspired by the multisource DANN model, only multiple source domain receiver labeled data can be used for adversarial training without adding the unlabeled data of the receiver in the target domain. In this case, the receiver domain classifier only classifies the data of multiple source domain receivers. The loss function of the network is as follows:

$$LOSS = L_{tx} + \lambda L_{d-rx}$$

During network training, the loss of the sender classifier is continuously minimized for labeled data from multiple source domains, and the loss of the receiver classifier is continuously maximized. Through this adversarial training, the feature extractor extracts the feature information that is related to the transmitter but not related to the receiver to achieve the purpose of migration across different receiver domains. For narrative convenience, this model is abbreviated as AD-CNN.

The method in this paper performs well on most receiver devices, but it may not perform well on individual receiver devices, such as receivers 8-10. In this work, the model presented in this paper is fine-tuned via 20 labeled samples of each pair of transmitter and receiver devices, and the experimental results are shown in Figure 5. The experimental results show that the average accuracy of the fine-tuned model from receivers 6-20 is 93.4%, which is 10% higher than that before fine-tuning.

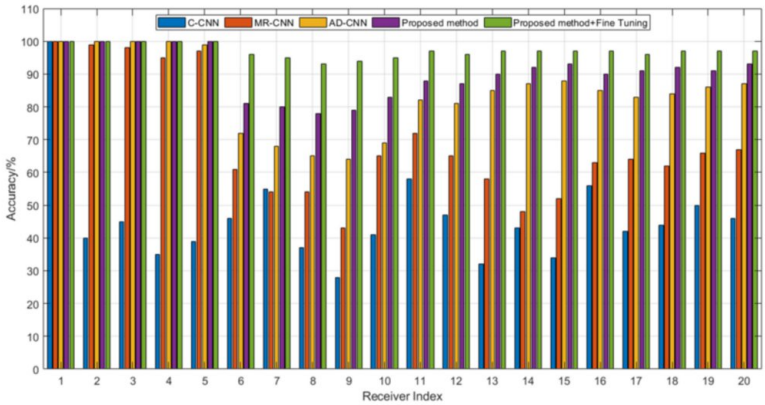

FIGURE 5: Cross-receiver identification results of different models

Expanded performance comparison with additional baselines

In addition to C-CNN, MR-CNN, and AD-CNN, the proposed model evaluated against two additional baselines:

DANN: A domain-adversarial neural network adapted to RF fingerprinting, trained using labeled source domain data and unlabeled target domain data.

DA-CNN: A dual-attention convolutional neural network incorporating both channel and spatial attention, trained in a supervised setting with source domain receivers only.

Table 3 summarizes the average accuracy of each model over the target domain receivers (6-20).

Model	Accuracy (%)
C-CNN	43.3
MR-CNN	62.5
AD-CNN	75.5
DANN	77.3
DA-CNN	78.4
Proposed (SE-MDANN)	83.3
Proposed (SE-MDANN, fine-tuned)	93.4

TABLE 3: Average cross-receiver identification accuracy of various models (Receivers 6–20)

DA-CNN: Dual-Attention Convolutional Neural Network; DANN: Domain Adversarial Neural Network; SE-MDANN, Squeeze-and-Excitation attention Multisource Domain Adversarial Neural Network

Ablation experimental analysis

To verify the effectiveness of the partial fusion of the SE attention mechanism of the model feature extractor in this paper, the following ablation experiments were carried out on the feature extractor part of the model, and the other parameters of the model were consistent except for the changes. Considering the difference in the distributions of receiver data in different target domains and to avoid randomness, experiments are carried out on receiver 6-10 devices.

1) *ResNet* network. On the basis of the method in this paper, the SE attention mechanism module is reduced, and the ResNet network is used for feature learning and classification recognition, aiming to explore the advantages of the SE attention mechanism in terms of performance.

2) *ResNet+CBAM* attention mechanism. Instead of using the SE attention mechanism, the CBAM attention mechanism is introduced into the network model, which aims to explore the advantages of the SE attention mechanism over the CBAM attention mechanism in terms of performance. In the CBAM attention mechanism, a convolutional kernel with a size of 3 × 1 is used to extract spatial attention features, and the other parameters are consistent with those of the SE attention mechanism.

3) *ResNet+CA* attention mechanism. Instead of using the SE attention mechanism, the network model introduces the CA attention mechanism, which aims to explore the performance advantages of the SE attention mechanism over the CA attention mechanism. The other model parameters are consistent with those of the SE attention mechanism.

The experimental results are listed in Table 3. As shown in the table, the *ResNet+SE* attention mechanism proposed in this paper has the highest accuracy among the five target domain receivers, and the training time is the shortest and the best performance compared with those of the other two attention mechanisms. Compared with that of the *ResNet* method, the average accuracy of the proposed method is 4.26% higher, and the experimental results show that the model performance of the *ResNet+SE* attention mechanism as the feature extractor is better than that of the *ResNet* network as the feature extractor. In addition, the accuracy of the *ResNet+CBAM* attention mechanism and the *ResNet+CA* attention mechanism is greater than that of the *ResNet* network on the receiver devices of the five target domains, which indicates that the CBAM and CA attention mechanisms have a certain effect on the performance of RF fingerprinting tasks. However, the *ResNet+SE* attention mechanism is better than the above two attention mechanisms are, which shows that the SE attention mechanism can make the model pay more attention to important information and features, thereby improving the recognition accuracy. The experimental results demonstrate the effectiveness and accuracy of the proposed method.

Feature extractor	Accuracy rate (%)					
	Receiver 6	Receiver 7	Receiver 8	Receiver 9	Receiver 10	Training takes time/s
Resnet	77.2	76.0	74.3	74.2	75.9	225.73
Resnet+CBAM	77.5	76.4	74.8	74.7	76.2	355.33
Resnet+CA	78.0	76.5	75.0	75.2	76.8	352.04
Resnet+SE	82.0	81.5	77.4	78.6	79.4	247.60

TABLE 4: Results of the ablation experiment

CBAM, Convolutional Block Attention Module; CA, Coordinate Attention; SE, Squeeze-and-Excitation

Effect of the number of multisource domain receivers

For the method proposed in this paper, a network model with different numbers of receivers in the source domain is trained and tested on receivers 14 that are not included in the training set, and the results are shown in Figure 6. The experimental results show that as the number of receivers in the source domain increases, *accuracy*, *precision*, *recall,* and *F1* value of the model improves. After the number of receivers in the source domain reaches 5, the accuracies of the model and other indicators improve slightly. Therefore, in the subsequent experiments, the model is trained by default with five receivers as the source domain receiver data.

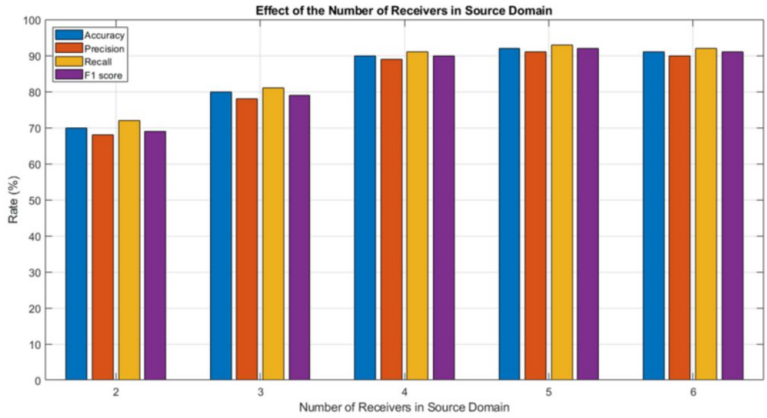

FIGURE 6: Effect of the number of receivers in the source domain

Influence of the number of unlabeled training samples in the target domain

In view of the SE attention multisource DANN model proposed in this paper, the influence of the number of unlabeled samples in the target domain on the performance of the model is further studied. In this work, 10, 20, 30, 40, and 50 unlabeled samples of 10 transmitters in the target domain are randomly sampled, and the results are shown in Figure 7. The experimental results show that the accuracy of the model increases gradually with increasing number of unlabeled samples, and the accuracy of the model tends to be stable when the number of unlabeled samples in the target domain reaches 30. This finding indicates that adding unlabeled samples in the target domain during model training can improve the model recognition performance. In this paper, 30 unlabeled samples of the target domain are used to train the model by default.

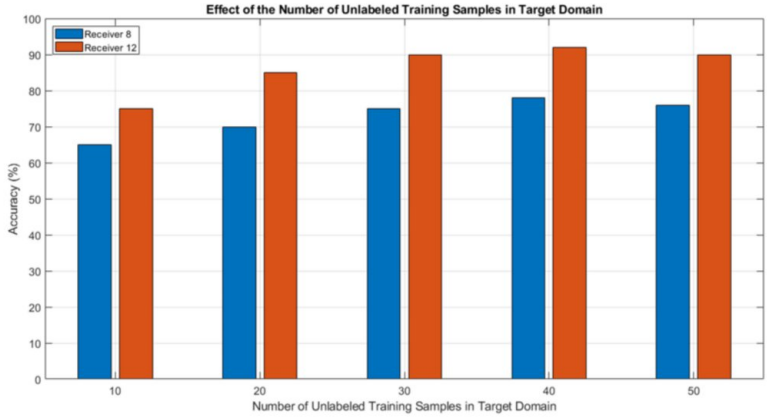

FIGURE 7: Effect of the number of unlabeled training samples in the target domain

Effect of the number of labeled fine-tuned samples in the target domain

In view of the fine-tuning method proposed in this paper, the influence of the number of labeled samples used for fine-tuning is further studied. Fine-tuning experiments are carried out via labeled samples of receiver 9 and receiver 13, and their fine-tuned performance is subsequently evaluated. In this work, 10, 20, 30, 40, 50, and 60 labeled samples were randomly sampled from 10 transmitters in the target domain, and the results are shown in Figure *8*. The experimental results show that the accuracy of the model increases with increasing number of labeled samples of the two receivers. It can also be observed that the accuracy of the labeled samples does not improve significantly after the number of samples reaches 40. The experimental results show that the recognition accuracy can be improved to 99% by using labeled samples to fine-tune the model parameters.

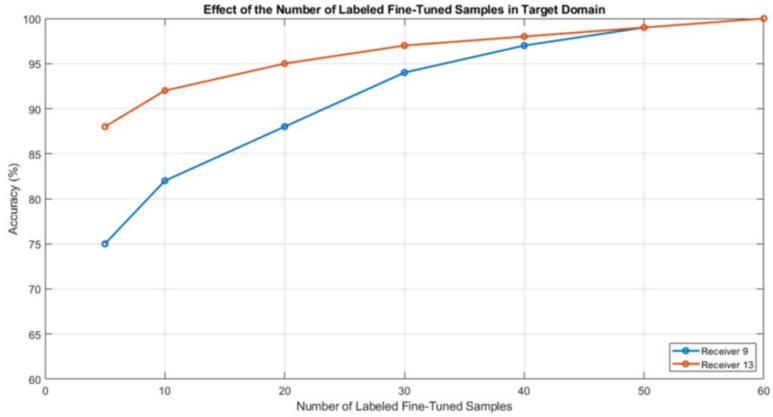

FIGURE 8: Effect of the number of labeled fine-tuned samples in the target domain

To further investigate the effectiveness and efficiency of the fine-tuning strategy, marginal improvement is analyzed in accuracy as the number of labeled samples per transmitter in the target domain increased from 10 to 60. As shown in Figure *8*, model accuracy improves steadily from 10 to 40 samples, with the most significant gains observed between 10 and 30 samples (e.g., from ~87% to ~92%). However, beyond 40 labeled samples, the accuracy gain begins to plateau, with only minor increases observed when moving from 40 to 60 samples (e.g., 93.1% to 93.4%). This trend suggests that the fine-tuning process exhibits diminishing returns past a certain threshold. Therefore, in practical deployments where labeled data is costly or limited, using approximately 30-40 labeled samples per transmitter strikes a favorable balance between recognition performance and annotation effort.

Computational overhead and feasibility of SE attention

In practical IoT deployments, computational cost is a critical factor when selecting deep learning models for real-time or resource-constrained environments. While the proposed SE attention mechanism enhances the model's ability to emphasize relevant transmitter features, it is important to assess its computational implications.

To evaluate this, proposed model's training time is compared with number of parameters of four configurations: baseline ResNet, ResNet with CBAM, ResNet with CA, and ResNet with SE attention (proposed model). As shown in Table *3*, the ResNet+SE configuration introduces a modest increase in training time (from 225.73 seconds to 247.60 seconds on average, approximately a 9.7% increase), while maintaining a lower parameter overhead compared to CBAM and CA. Despite this slight increase in computational cost, the SE-based model achieved the highest average accuracy across target receivers (82.8%), outperforming ResNet by 4.26%, ResNet+CBAM by 2.96%, and ResNet+CA by 2.2%.

These findings indicate that SE attention offers a favorable trade-off between performance and efficiency. Its lightweight structure-consisting primarily of global average pooling and two small fully connected layers-makes it particularly well-suited for IoT scenarios with constrained processing and memory resources.

Discussion

To verify the effectiveness of the model proposed, this paper compares the SE attention force multisource domain adversarial network model with the above three methods in cross-receiver scenarios. Owing to the differences in the hardware characteristics of different receivers, the support of the model proposed in this paper for diverse receivers should be verified, and randomness should be avoided. In this work, 20 receiver devices are selected for experimentation. For method 1, only receiver 1 is used as the source domain dataset, and receivers 2-20 are used as the target domain dataset. For methods 2 and 3 and the method in this paper, receivers 1-5 are used as the source domain dataset, and receivers 6-20 are used as the target domain dataset. The experimental results are shown in Figure *5*.

The experimental results show that the proposed method has the best performance compared with the above three methods, and the average accuracy can reach 83.3% on the target domain receiver 6-20. When different receivers 2-20 are used for testing, the accuracy is greatly reduced, and the average accuracy is only 43.3%. The average accuracy of the MR-CNN model can reach 62.5% for receivers 6-20 because of the use of data from multiple receivers for training, and its generalizability is better than that of the C-CNN model. The average accuracy of the AD-CNN model for receivers 6-20 is 75.5%, which is 32.4% and 13% higher than those of the C-CNN model and MR-CNN model, respectively. Compared with those of the C-CNN model, MR-CNN model and AD-CNN model, the average accuracy of the model proposed in this paper is improved by 40%, 20.8% and 7.8%, respectively. The inclusion of DANN and DA-CNN models in the experiments provides further validation of the proposed method. As shown in Table *4*, proposed model (SE-MDANN) consistently outperforms both recent domain adaptation and attention-based models. Notably, the improvement over DANN (by 6%) and DA-CNN (by 4.9%) demonstrates the added benefit of combining multisource adversarial learning with the SE attention mechanism. Fine-tuning with only 20 labeled samples per transmitter in the target domain further boosts accuracy to 93.4%, highlighting the adaptability and practicality of the approach in real-world IoT scenarios with limited data availability. The experimental results show that the proposed method has better generalizability and higher accuracy than the existing methods do, which indicates the effectiveness of the proposed method.

Conclusions

In view of the fact that the existing research on RF fingerprinting ignores the influence of the change of receiver hardware characteristics, resulting in the poor generalization ability and serious reduction of the accuracy of the model on different receiver devices, this paper proposes an RF fingerprinting method based on SE attention multi-source domain adversarial network, which reduces the dependence of the recognition model on the known receiver under the premise of ensuring good recognition performance. Furthermore, the SE attention multi-source domain anti-immune network model can be fine-tuned by using a small number of labeled samples of the receiver in the target domain to further improve the transmitter recognition performance of the network model. Experimental results show that the RF fingerprinting method based on SE attention multi-source domain adversarial network proposed in this paper has better recognition performance in cross-receiver scenarios, with an average accuracy of 83.3%, and the fine-tuning method can further improve the average accuracy to 93.4%. The machine captures the transmitter RF fingerprint and applies it to any receiver.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Isack Bulugu

Acquisition, analysis, or interpretation of data: Isack Bulugu

Drafting of the manuscript: Isack Bulugu

Critical review of the manuscript for important intellectual content: Isack Bulugu

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Xu Q, Zheng R, Saad W, Han Z: Device fingerprinting in wireless networks: Challenges and opportunities. IEEE Communications Surveys & Tutorials. 2016, 18:94-104. 10.1109/comst.2015.2476338
2. Zou Y, Zhu J, Wang X, Hanzo L: A survey on wireless security: Technical challenges, recent advances, and future trends. Proceedings of the IEEE. 2016, 104:1727-1765. 10.1109/jproc.2016.2558521
3. Jagannath A, Jagannath J, Kumar PSPV: A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges. Computer Networks. 2022, 219:109455. 10.1016/j.comnet.2022.109455
4. Soltanieh N, Norouzi Y, Yang Y, Karmakar NC: A review of radio frequency fingerprinting techniques. IEEE Journal of Radio Frequency Identification. 2020, 4:222-233. 10.1109/jrfid.2020.2968369
5. Wang W, Gan L: Radio frequency fingerprinting improved by statistical noise reduction. IEEE Transactions on Cognitive Communications and Networking. 2022, 8:1444-1452. 10.1109/tccn.2022.3171792
6. Nicolussi A, Tanner S, Wattenhofer R: Aircraft fingerprinting using deep learning. 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands. 2021, 740-744. 10.23919/Eusipco47968.2020.9287691
7. Özkiper Zİ, Turgut Z, Atmaca T, Aydın MA: Fingerprint liveness detection using deep learning. 2022 9th International Conference on Future Internet of Things and Cloud (FiCloud), Rome, Italy. 2022, 129-135. 10.1109/FiCloud57274.2022.00025
8. Liu P, Yang P, Song WZ, Yan Y, Li XY: Real-time identification of rogue WiFi connections using environment-independent physical features. IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, Paris, France. 2019, 190-198. 10.1109/INFOCOM.2019.8737455
9. Brik V, Banerjee S, Gruteser M, Oh S: Wireless device identification with radiometric signatures. MobiCom '08: Proceedings of the 14th ACM international conference on Mobile computing and networking. 2008, 116-127. 10.1145/1409944.1409959
10. Satija U, Trivedi N, Biswal G, Ramkumar B: Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios. IEEE Transactions on Information Forensics and Security. 2018, 14:581-591. 10.1109/tifs.2018.2855665
11. Jian T, Rendon BC, Ojuba E, Soltani N, Wang Z, Sankhe K: Deep learning for RF fingerprinting: A massive experimental study. IEEE Internet of Things Magazine. 2020, 3:50-57. 10.1109/iotm.0001.1900065
12. Riyaz S, Sankhe K, Ioannidis S, Chowdhury K: Deep learning convolutional neural networks for radio identification. IEEE Communications Magazine. 2018, 56:146-152. 10.1109/mcom.2018.1800153
13. Gopalakrishnan S, Cekic M, Madhow U: Robust wireless fingerprinting via complex-valued neural networks. 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA. 2019, 1-6. 10.1109/GLOBECOM38437.2019.9013154
14. Shen G, Zhang J, Marshall A, Peng L, Wang X: Radio frequency fingerprint identification for LoRa using deep learning. IEEE Journal on Selected Areas in Communications. 2021, 39:2604-2616. 10.1109/jsac.2021.3087250
15. Du Z, Liu D, Zhang J, Lin D, Gao Y, Cao J: RSEN-RFF: Deep learning-based RF fingerprint recognition in noisy environment. Artificial Intelligence for Communications and Networks. AICON 2021. Wang X, Wong KK, Chen S, Liu M (ed): Springer, Cham; 2021. 396:87-99. 10.1007/978-3-030-90196-7_9
16. Gu H, Su L, Zhang W, Ran C: Attention is needed for RF fingerprinting. IEEE Access. 2023, 11:87316-87329. 10.1109/access.2023.3305533
17. Jiang Q, Sha J: The use of SNN for ultralow-power RF fingerprinting identification with attention mechanisms in VDES-SAT. IEEE Internet of Things Journal. 2023, 10:15594-15603. 10.1109/jiot.2023.3264715
18. Weng L, Peng J, Li J, Zhu Y: Message structure aided attentional convolution network for RF device fingerprinting. 2020 IEEE/CIC International Conference on Communications in China (ICCC), Chongqing, China. 2020, 495-500. 10.1109/ICCC49849.2020.9238868
19. Zhang J, Woods R, Sandell M, Valkama M, Marshall A, Cavallaro J: Radio frequency fingerprint identification for narrowband systems, modelling and classification. IEEE Transactions on Information Forensics and Security. 2021, 16:3974-3987. 10.1109/tifs.2021.3088008
20. Hanna S, Karunaratne S, Cabric D: WiSig: A large-scale WiFi signal dataset for receiver and channel agnostic RF fingerprinting. IEEE Access. 2022, 10:22808-22818. 10.1109/access.2022.3154790
21. Xie Q, Dai Z, Du Y, Hovy E, Neubig G: Controllable invariance through adversarial feature learning. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017, 585-596.

22. Ganin Y, Lempitsky V: Unsupervised domain adaptation by backpropagation. ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning. 2015, 37:1180-1189.
23. Kumar P, Suresh S: DeepTransHAR: a novel clustering-based transfer learning approach for recognizing the cross-domain human activities using GRUs (Gated Recurrent Units) Networks. Internet of Things. 2023, 21:100681. 10.1016/j.iot.2023.100681
24. Guo J, Fu H, Pan B, Kang R: Recent progress of residual stress measurement methods: A review. Chinese Journal of Aeronautics. 2021, 34:54-78. 10.1016/j.cja.2019.10.010
25. Hu J, Shen L, Sun G: Squeeze-and-excitation networks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA. 2018, 7132-7141. 10.1109/CVPR.2018.00745
26. Woo S, Park J, Lee JY, Kweon IS: CBAM: Convolutional block attention module. Computer Vision - ECCV 2018. Ferrari V, Hebert M, Sminchisescu C, Weiss Y (ed): Springer, Cham; 2018. 11211:3-19. 10.1007/978-3-030-01234-2_1
27. Hou Q, Zhou D, Feng J: Coordinate attention for efficient mobile network design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA. 2021, 13708-1371. 10.1109/CVPR46437.2021.01350