Extraction and Recognition of Fingerprint Characteristic of Mobile Terminal's Transient Signal

Fucai Luo, Fei Wu, Hongfa Li, Ting Li, Jindong He and Qian Chen

State Grid Fujian Electric Power Company, 350000, China.
Email: Fucai Luo. 519594474@qq.com

Abstract. The fingerprint characteristic of mobile terminal's signal is unique and it can be used to identify the source of the signal. In all the characteristics, the characteristic of transient signals has been more favoured because of the greater diversity. However, the duration of transient signal is extremely short and difficult to accurately detect. Therefore, in order to successfully obtain the fingerprint characteristic, most of the research results are based on the laboratory environment. In this paper, through the methods of differential constellation trajectory and neural network, we realize the extraction and recognition of fingerprint characteristic of mobile phone's transient signal in the real environment. Meanwhile, by controlling the distance between the terminal and the base station, we also studied the recognition of fingerprint characteristic under different noise conditions.

1. Introduction

With the development of smart phones, wireless communication plays an increasingly important role in people's lives. Especially since this year, 5G will gradually become the new mainstream communication standard. However, no matter how the communication standard is updated, GSM has a status that cannot be ignored in the history of communication. In this paper, we will mainly consider the fingerprint characteristic of GSM signal, especially the GSM 900M frequency band (Figure 1 is the corresponding band diagram).

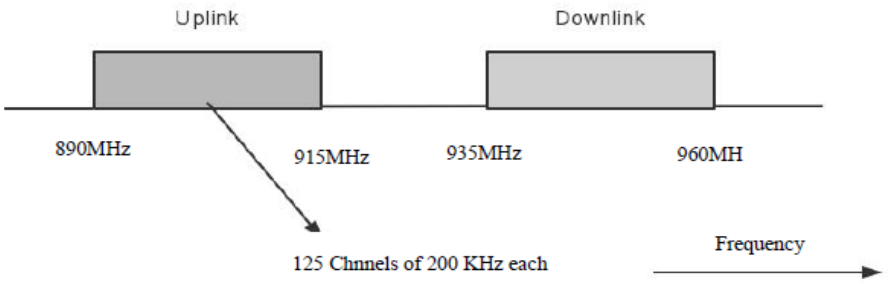

Figure 1. GSM 900M band diagram

Since the first day of wireless communication, electronic attack and defence at the information level has followed. With the development of technologies such as mobile payment, wireless network terminals are increasingly vulnerable to various information intrusion methods: WIFI phishing attacks, data sniffing attacks, etc. to overcome these problems, people usually need to perform various encryption and authentication on the transmitted information. In this process, people usually need to complete operations such as key generation, key agreement, and key distribution. This will cause a lot

of waste on communication resources, and will bring a series of security problems. The uniqueness of the fingerprint characteristic of mobile terminal's signal provides the possibility to solve the above problems.

RF fingerprint characteristic is the characteristic vector or vector set extracted from the received wireless signal that uniquely identifies the transmitter individual and embodying the hardware characteristics of the wireless device transmitter. The characteristic has a certain stability; the characteristics are the same when the detection conditions are the same. Which means if the conditions of the two communication terminals do not change, the fingerprint characteristic of them will not change. If the adversary joins this communication, the characteristic will change because the conditions of communication have changed.

In 2003, Hall et.al. firstly propose RF fingerprinting in the identification of wireless device terminals [1]. RF fingerprinting uses the characteristic information extracted by the transient part of the mobile terminal transmitter signal to uniquely identify the terminal, and builds an anti-intrusion detection system that recognizes the RF fingerprint of IEEE 802.11b devices and successfully countered attacks such as MAC address cloning [2]. The characteristic is unique when the terminal is powered on, and divide terminal fingerprint characteristic recognition process into three stages: the transient starting point detection stage, the signal characteristic extraction stage and the signal characteristic recognition stage [3]. This concept is also applied to the field of ZigBee, Bluetooth and GSM terminals [4-9].

According to the timing of signal acquisition, the RF fingerprinting can be divided into transient RF fingerprint features and steady-state RF fingerprint features. Transient RF fingerprint refers to the characteristics of electromagnetic waves when the transmitter is turned on and off. This signal is determined by the hardware itself. Steady-state characteristics are characteristics of the signal itself, such as frequency offset, amplitude error, phase error, IQ offset, IQ imbalance, etc. Transient characteristic duration is very short, need to determine the starting point, difficult to collect. Steady-state characteristic has lower requirements for acquisition equipment, easy to collect. However, the steady-state characteristic may change, have lower difference, and it is not possible to extract identifiable characteristic in a signal. Therefore, transient characteristic better reflect terminal differences. And in transient signal-based RF fingerprinting, detecting the starting point of a transient signal is a critical step.

In order to more clearly depict the characteristics of transient signals, a method based on differential constellation trajectory map is used for RF fingerprint characteristic extraction [10]. In [10], the USRP equipment is used to draw the differential constellation using the signals generated by QPSK, BPSK, MSK, etc. The corresponding graphics are extracted, and the relevant characteristic are extracted to verify the reliability and practicability of the proposed method. It is also confirmed that in the device recognition process, the wireless terminals can be identified without extracting the a priori information of the transmitter.

The rest of the article is structured as follows: Section 2 gives the introduction of fingerprint characteristic extraction method, the differential constellation trajectory. Section 3 gives the introduction of fingerprint characteristic recognition method, the neural network and machine learning. Section 4 gives the specific experimental methods and experimental results.

2. Differential Constellation Trajectory

In digital communication, the constellation trajectory is directly processed by processing the sampled signal in the complex plane. It can provide a convenient way to study the relationship between the I/Q signals and visually represent the relationship between the signals.

When using oversampling to collect the radio frequency fingerprint characteristic, the constellation trajectory can be obtained by drawing the constellation map through the determined sampling points. This is because the nonlinear response of the amplifier and the interference factors from the filter are expressed in the constellation trajectory map, and the received signal characteristics are more closely distinguished in the transmitter signal transmission.

Under normal circumstances, there will exists errors such as frequency offset between the transmitter and the receiver, which will lead to instability of the constellation trajectory map. For a giving transmit signal

$$\boldsymbol{S}(t) = \boldsymbol{X}(t)e^{-2\pi jtf_{ct1}} \tag{1}$$

Then, in the ideal experimental case, the receiving signal is

$$\boldsymbol{Y}(t) = \boldsymbol{S}(t)e^{2\pi jt+\varphi f_{ct2}} \tag{2}$$

If there is a deviation between the receiving end and the transmitting signal, for example, f_{ct1} and f_{ct2} have a deviation Δf, then

$$\boldsymbol{Y}(t) = \boldsymbol{X}(t)e^{2\pi jt+\varphi\Delta f} \tag{3}$$

Obviously, the constellation trajectory map is a time-dependent function. It will rotate over time t, and makes us difficult to compare different images through the constellation. To overcome this problem, we need to differentially process the data under the premise of retaining the frequency offset. Then

$$\boldsymbol{D}(t) = \boldsymbol{Y}(t) * \boldsymbol{Y}^*(t+n) = \boldsymbol{X}(t) \cdot \boldsymbol{X}(t+n)e^{-2\pi j\theta n} \tag{4}$$

Although the image after the difference still rotates, the rotation is a stable value and is related to the frequency offset, so that we can better observe the characteristics of the frequency offset.

3. Neural Network and Machine Learning

Since the 1980s, neural networks have developed very rapidly. In our experiment, we mainly used 3 different neural network schemes. We will introduce them separately.

3.1. BP Neural Network

BP neural network is currently the most widely used neural network architecture for its ability to learn complex multidimensional mappings. It is a nonlinear dynamic systeme. It is made up of many simple processing units, and the larger the number, the better the recognition effect.

The main feature of BP neural network is that the signal is forward propagating and the error is back propagating. The model of BP neural network is shown in the figure 2 below:

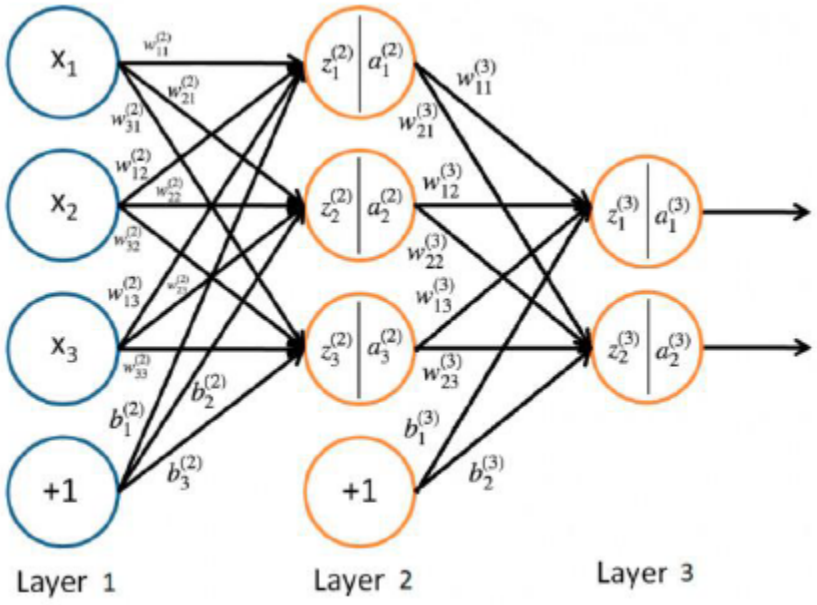

Figure 2. 3-layer BP neural network

The process of BP neural learning network is mainly divided into two stages, the first stage to the forward propagation of the output layer; the second stage is the error return to the input layer back propagation. The input to the neuron includes the sum of its deviation and the weighted input. The output of a neuron depends on the input of the neuron and its transfer function. The network implements the calculation by mapping the input values to the output values. The specific mapping problem to be performed determines the number of inputs and the number of outputs from the network. Therefore, we need to choose the best BP neural network structure according to the type of problem.

3.2. ALEXNET Neural Network

Compared with BP neural networks, ALEXNET neural networks are more mature and more complex. AlEXNET has five layers of convolution, three layers of fully connected network (see Figure 3). At the same time, ALEXNET can use two GPUs for calculation at the same time, which greatly improves the computational efficiency.

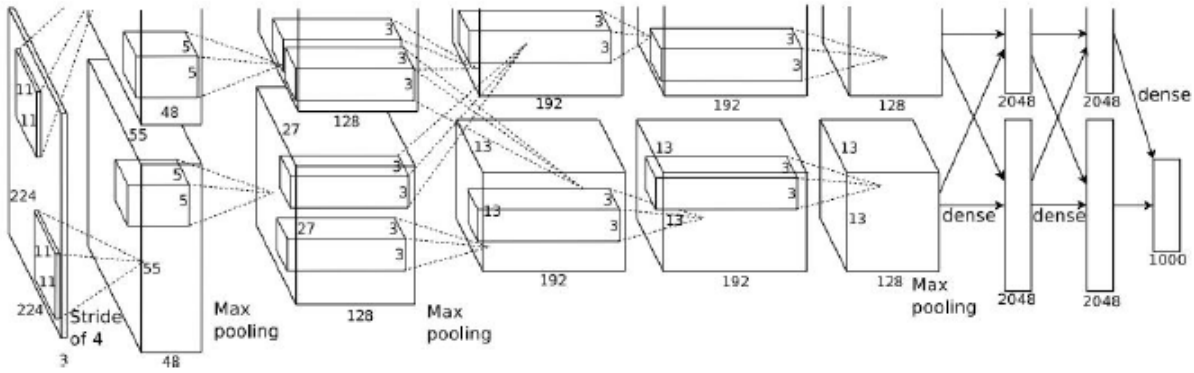

Figure 3. ALEXNET neural network training structure

3.3. Support Vector Machines

Support Vector Machine (SVM) is a new machine learning algorithm based on statistical learning theory with the goal of minimizing structural risk. It is mainly composed of 3 parts: Preprocessor, SVM classifier, and decision system. The preprocessor is to process the data into the form of SVM. The SVM classifier is the core component of the SVM system and needs to be used to train data training to achieve performance. Finally, the judgment equations in the decision system are pained to determine the category in which the object is located.

SVM was originally designed for binary classification problems. It need to construct a suitable multi-class classifier when dealing with multiple types of problems. In this paper, we combine our SVM classifier with the histogram of oriented gradients (HOG) feature to complete our experiments.

In an image, the appearance and the shape of the object can be described by gradient of the shape and the colour. Therefore, we can divide the image into individual cell elements, and analyze every individual cell element. For easy calculations, we can also normalize the date in each cell. Normalization can also achieve better results for changes in the color of image. Then, we can process the graphical gradients and features between cell and cell. According to all these results, we can draw the direction histogram to form the characteristic of each image for comparative recognition. Figure 4 gives the implementation process of HOG.

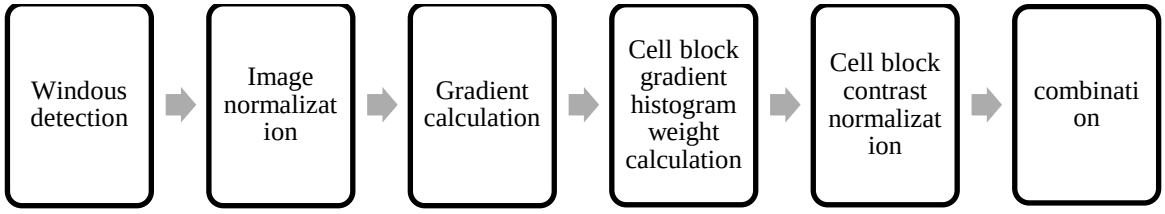

Figure 4. The implementation process of HOG

4. Our Experiments

The experimental equipment we experimented with includes hardware devices and software devices. On the software device, we chose GNU Radio. We use it to write data processing modules, and integrated the modules to form a complete wireless communication system model. On the hardware devices, we use Hack Rf and USRP devices as our transmitter and receiver, respectively. The Hack Rf (see Figure 5) is a software radio external device that mainly use to transmit and receive signals, and to mix and sample the signals. The USRP (Figure 6) is also a general-purpose software radio peripheral. It handles high-speed general-purpose operations such as up-and-down conversion, interpolation, and extraction.

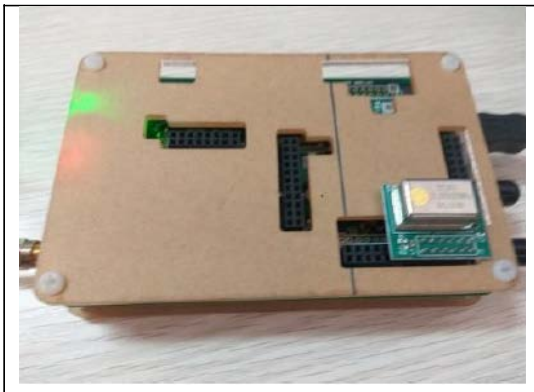

Figure 5. The Hack Rf

Figure 6. The USRP

After we obtain the signals by the Hack RF. We first perform GMSK modulation on the received GSM signal. Then, we analyse the modulation results by means of spectrum analysis, frequency finding and filtering. So as to facilitate our SDCCH detection. With SDCCH, we can extract the transient characteristics of the signal, and then draw a differential constellation trace. Finally, by the neural networks and machine learning algorithms, we identify and classify the signals. Figure 7 gives the process of our experiment.

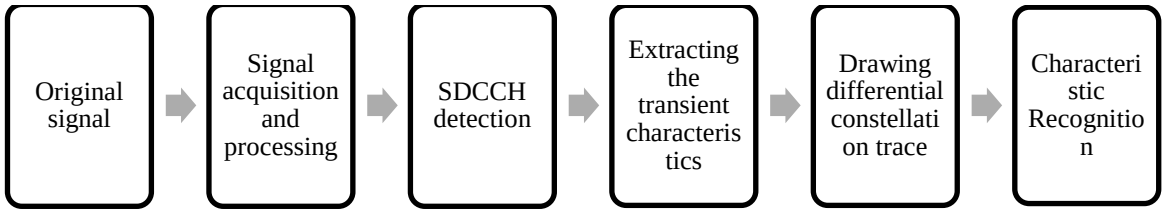

Figure 7. The process of fingerprint characteristic of signal

4.1. Characteristic Extraction

Before processing the acquired data, we use the GNUradio device to map the differential constellation traces of the GMSK modulated signals under ideal conditions. After the GMSK modulation of the transmitted signal, it is received by the USRP and then processed directly by GNUradio, and the interference caused by noise and multipath can be almost ignored. Set a delay amount during the

receiving process, and delay and differentially calculate the signal each time. Then a differential constellation trajectory map in an ideal state can be obtained.

For the single GSM signal that has been acquired, we use Matlab to draw the differential constellation for processing. Then we select the appropriate differential value to directly process the data differentially. By this method, a differential constellation trajectory of single GSM signal with obvious features can be obtained.

In order to better reflect the differential constellation, we performed a visualization process to mesh the differential constellation trajectory map, by giving red, green and blue colors according to the density. The higher the density, the closer the color is to red. Note that, when drawing differential constellation trajectories, there will be different results with different interval difference. Figure 8 gives some example of differential constellation trajectory maps with different interval difference.

When choosing the appropriate differential interval, the differential constellation trajectory map will have a better contour, and it will also have a better shape on the curve near the center inside the contour. When the difference interval is large, the image contour will form a wide range of distortion, and it will not exhibit a good range of polymerization. The center curve will become blurred and cannot be a good recognition standard. When the difference interval is too small, it can be found that the constellation trajectory maps no longer exhibits a circular range that can be clearly observed, but the range appears to decrease, and the aggregation is too obvious. Furthermore, When the difference interval is close to 1, the image becomes a single line and no longer provides any feature information. Therefore, appropriate differential spacing is a great help for getting good fingerprint features.

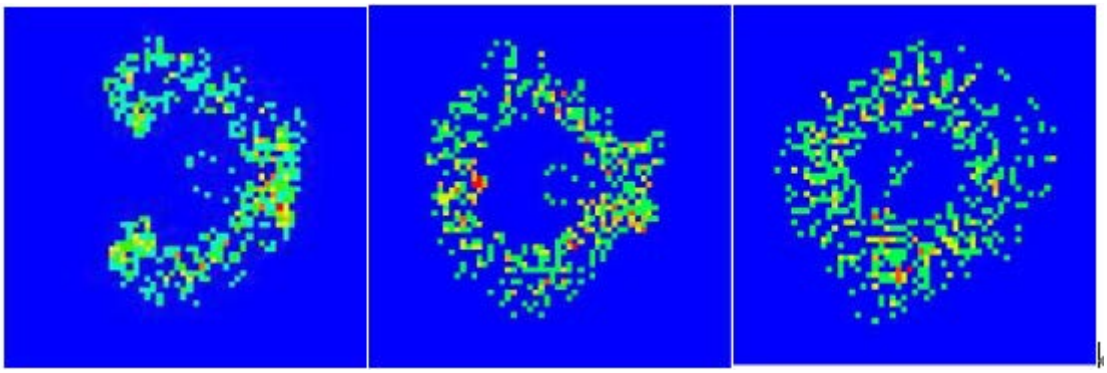

Figure 8. Differential constellation trajectory maps with different interval difference (From left to right, followed by smaller intervals, appropriate intervals, larger intervals)

From the differential constellation, we can observe both steady-state and transient characteristics. Due to the slot structure of GSM, the transient characteristics of the transmitter at the guard interval are fully reflected in the curve near the center of the contour of the differential constellation trajectory map. It is generated due to the differential data of the fading data of the guard interval, and the differential data of the fading data of the partial guard interval and the header bits of the frame, which represent the characteristics of the transmitter at the transmission interval of the base station. Compared with the traditional wavelet transform and instantaneous frequency, envelope and other methods, the differential constellation trajectory map directly shows these data in a curve, which is more intuitive.

After the differential constellation trajectory map is drawn, it is easy to find that in the differential constellation trajectory map of different devices, there will be a certain difference in the contour curve. And this is the key to the fingerprint recognition of mobile devices. Through the difference between the contour and the curve within the contour, we can effectively classify different terminal devices.

In addition, we found that there are still differences in the differential constellation traces of different terminal signals over longer distances. This means that the differential constellation trajectory can still be used as an important feature to identify the terminal signal under certain interference conditions.

4.2. Characteristic Recognition

In this chapter, we will realize the characteristic recognition of terminal signals from different angles through three different methods. They are BP neural network, ALEXNET neural network and Support Vector Machine, respectively.

4.2.1. BP Neural Network. This experiment uses the neural network tool system nntool that comes with Matlab to realize the construction of BP network. We only need to input the length of the vector and the number of output classifications, and then we can get a good neural network after training.

First, we perform binary recognition on the obtained 41*41 differential constellation trajectory map and convert the image into a matrix. Then, we input the obtained matrix into the BP network training box as the basic, and train and tested autonomously another part of the samples.

Here, we set up three layers of hidden layers, which are 25, 12, and 1 neuron, as shown in the figure 9. It can be seen that the BP neural network responds very well to our data, and only with a short training layer number, we have obtained considerable classification results, as shown in the figure 10.

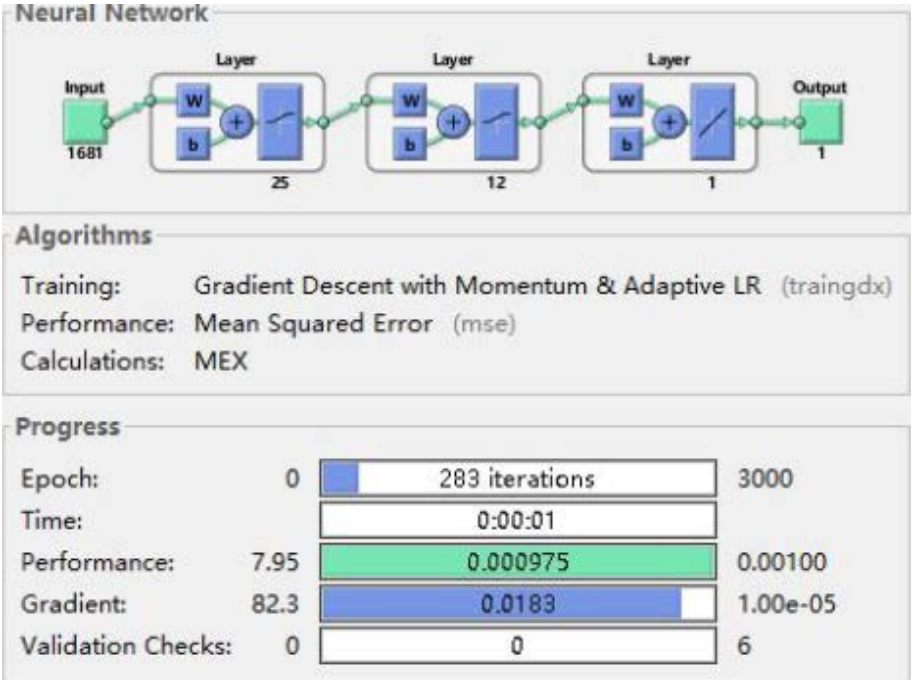

Figure 9. BP Neural Network

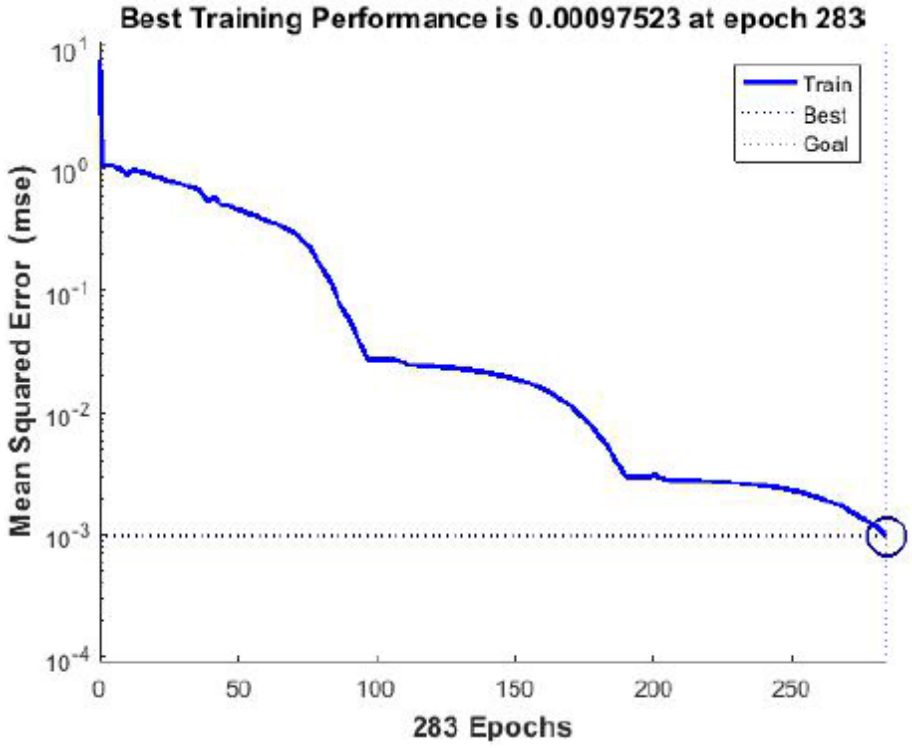

Figure 10. Training performance of BP

We also process the RF fingerprint of the terminal equipment tested in a noisy environment, and identify the single picture when the trained model has been deployed to determine whether it can successfully match the data in the sample library. Although the vast majority of samples have been identified. Unfortunately, one sample was misidentified. The reason is because BP needs to convert the original differential constellation trace map into binary data when inputting data. This can cause some data distortion and even loss of critical data.

4.2.2. ALEXNET neural network. ALEXNET can directly process the differential constellation trace map (41*41*3) without binarization. This means that the image can be fully recognized for recognition without losing some of the key information that can be avoided.
First, we input the differential constellation trace map drawn by the close-range terminal GSM data into the ALEXNET neural network system. Training parameters are shown in Table 1.

Table 1. Training parameters of AVERNET

Training Parameter	Value
Training Sample Set	200
Minimum Batch Size for Each Training Iteration	30
Maximum Number of Training Cycles	30
Initial Learning Rate	0.0001
Network Verification Frequency (Iterations)	30
Data for Verification during Training	Auto

The training results (as shown in Figure 11) show that the training speed is very fast, and all test samples can be correctly classified, and excellent experimental results are obtained (Root mean square error (RMSE) and Loss Value Function (LOSS) are both reduced to a very low range, as shown in Figure 12).

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:00	13.28%	3.0845	1.0000e-04
1	50	00:00:02	64.84%	1.0945	1.0000e-04
2	100	00:00:03	74.22%	0.7260	1.0000e-04
3	150	00:00:05	83.59%	0.4741	1.0000e-04
4	200	00:00:06	91.41%	0.3089	1.0000e-04
5	250	00:00:08	92.97%	0.2333	1.0000e-04
6	300	00:00:09	97.66%	0.1539	1.0000e-04
7	350	00:00:11	97.66%	0.1317	1.0000e-04
7	400	00:00:12	96.09%	0.0944	1.0000e-04
8	450	00:00:14	98.44%	0.0662	1.0000e-04
9	500	00:00:15	99.22%	0.0458	1.0000e-04
10	550	00:00:17	100.00%	0.0545	1.0000e-04
11	600	00:00:18	99.22%	0.0660	1.0000e-04
12	650	00:00:19	100.00%	0.0338	1.0000e-04

Figure 11. The training results of AVERNET

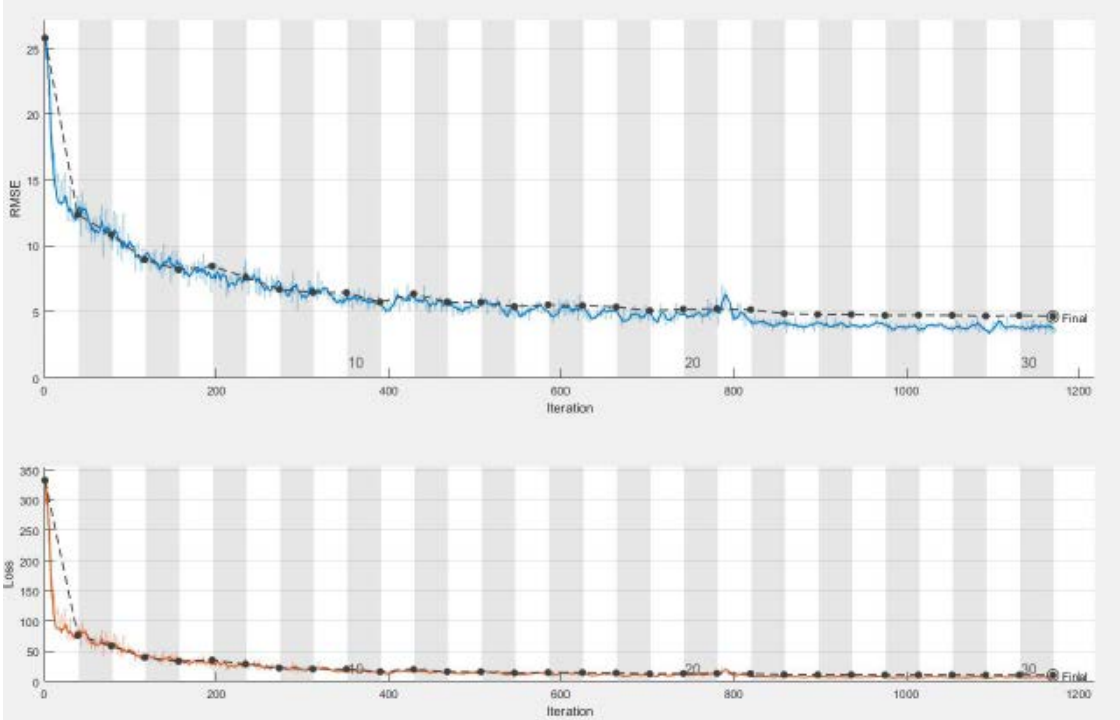

Figure 12. The RMSE and LOSS of training

After that, we imported the trained neural network and input the differential constellation trace map drawn under the loud noise into the SOFTMAX with the connection layer, which obtained a good recognition effect. The results show that the recognition of the constellation map in the ALEXNET training network has achieved a recognition rate of 98.5%.

4.2.3. Support Vector Machine.

As mentioned earlier, the SVM system is more about the binary classification problem. In order to use the SVM system, we took the fitcecoc function carried by MATLAB to complete our experiment. We input the extracted HOG features and the required classification labels through fitcecoc. Then, we can classify them with SVM, create sample sets, and predictively identify them by predict function.

As with the ALEXNET experiment, we used SVM for machine learning processing with a sample set of 200 different constellation trajectory maps and tested the test set. The results show that the recognition of the constellation map in the SVM has also achieved a recognition rate of minimum 85%.

5. Conclusion

In this paper, by analysing the terminal GSM transient signal, the differential constellation trajectory representing the signal fingerprint characteristics are obtained. Then, according to these fingerprint characteristics, the characteristic recognition of the wireless terminal device is realized by means of neural network and machine learning. Our experiments show that BP neural network, AVERNET neural network and SVM have excellent recognition ability for differential constellation trajectory feature recognition under reasonable training conditions. Moreover, the recognition of AVERNET neural network and SVM is more accurate than BP neural network.

6. References

[1] 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

[2] Guyue Li, Jiabao Yu, Yuexiu Xing, and Aiqun Hu. Location invariant physical layer identification approach for WiFi devices. *IEEE Access*, pages 1–1, 2019.

[3] Linning Peng, Aiqun Hu, Junqing Zhang, Yu Jiang, Jiabao Yu, and Yan Yan. Design of a hybrid RF fngerprint extraction and device classifcation scheme. IEEE Internet Things J., 6(1):349–360, Feb. 2019.

[4] Yuexiu Xing, Aiqun Hu, Junqing Zhang, Linning Peng, and Guyue Li. On radio frequency fngerprint identifcation for DSSS systems in low SNR scenarios. *IEEE Commun. Lett.*, 22(11):2326–2329, Nov. 2018.

[5] J. Hall, M. Barbeau, and E. Kranakis, “Detecting rogue devices in Bluetooth networks using radio frequency fingerprinting,” in *Proc. IASTED Int. Conf. Commun. Comput. Netw. (CCN)*, Lima, Peru, Oct. 2006, pp. 108–113.

[6] A. M. Ali, E. Uzundurukan, and A. Kara, “Assessment of features and classifiers for bluetooth RF fingerprinting,” IEEE Access, vol. 7, pp. 50 524–50 535, 2019.

[7] Jiabao Yu, Aiqun Hu, Guyue Li, and Lin Ning Peng. A robust RF fngerprinting approach using multi-sampling convolutional neural network. *IEEE Internet Things J.*, pages 1–1, 2019.

[8] Xinyu Zhou, Aiqun Hu, Guyue Li, Linning Peng, Yuexiu Xing, and Jiabao Yu. Design of a robust rf fngerprint generation and classifcation scheme for practical device identifcation. In *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, pages 1–9, Washington, DC, USA, Jun. 2019.

[9] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, “Comparison of techniques for radiometric identification based on deep convolutional neural networks,” *Electron. Lett.*, vol. 55, no. 2, pp. 90–92, Jan. 2019.

[10] Linning Peng, Aiqun Hu, Yu Jiang, Yan Yan, and Changming Zhu. A differential constellation trace fgure based device identifcation method for ZigBee nodes. In *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, pages 1–6, Yangzhou, China, Oct. 2016.