Hindawi
Journal of Sensors
Volume 2022, Article ID 8122815, 9 pages
https://doi.org/10.1155/2022/8122815

Research Article

RFID Scheme for IoT Devices Based on LSTM-CNN

Kaizhi Huang, Xinglu Li, Shaoyu Wang, Zengchao Geng, and Ge Niu

PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China

Correspondence should be addressed to Xinglu Li; lixinglundsc@gmail.com

Received 13 May 2022; Revised 10 October 2022; Accepted 26 October 2022; Published 18 November 2022

Academic Editor: Aijun Yin

As an essential branch of physical layer authentication research, radio frequency identification (RFID) has advantages in achieving lightweight and highly reliable authentication. However, in the Internet of Things (IoT) environment, where a large scale of devices are connected to the network, there is an issue that the difference of the RF fingerprints is less distinct among the same type of devices. To this end, in this paper, we propose an RFID scheme for IoT devices based on long-short term memory and convolutional neural network (LSTM-CNN). This scheme combines the excellent learning ability of LSTM and CNN to perceive the context information and extract the local feature of RF data. Specifically, RF data is first fed into LSTM to obtain long-term dependency features containing temporal information. Then, CNN is designed for secondary feature extraction to enlarge RF differences and further used for device classification. The experiment results on the open RF data set ORACLE indicate that the identification accuracy of the proposed scheme can reach over 99%. Compared with other schemes, the performance is improved by 6%-30%.

1. Introduction

With the rapid development of IoT, a large number of devices are connected to the network through the wireless channel. However, as shown in Figure 1, the broadcast feature of the wireless channel makes it possible for attackers to connect the access point (AP) by impersonating the identity of legitimate nodes. Then, attackers can disrupt the legitimate communication by maliciously monitoring, tampering, or discarding transmission information [1]. The security issues on the wireless network have attracted more and more attention.

RFID is committed to realizing wireless security by exploiting characteristics in the communication process. Specifically, RFID uses the unique features caused by hardware imperfections to mark the identity of devices, thus is expected to achieve lightweight and highly reliable authentication. The source of RF fingerprint in the transmitter is shown in Figure 2, where hardware imperfections, such as harmonic distortion of the digital to analog converter (DAC), direct-current (DC) bias, local oscillator (LO) leakage, I/Q gain imbalance, and power amplifier (PA) nonlinearity, [2] will reflect in the RF waveform and are measurable, so the RF fingerprint can uniquely identify the device. Nevertheless, RF fingerprints are caused by accidents, and there are top limits on the electronic tolerance standards among the same type of devices, resulting in the limited RF feature space and less distinction among massive IoT devices. Therefore, an effective method of feature extraction is of great importance.

The existing RF feature extraction methods are mainly divided into two categories: manual selection [3, 4] and deep learning extraction. Due to the feature space limitation of the same type of devices and the time and labor consuming, manual selection is not able to meet RFID requirements among massive IoT devices. Extraction methods based on deep learning [5–7] provide a new solution to amplify RF feature differences based on its strong feature extraction capability. Merchant et al. [5] conducted synchronization and filtering on the RF baseband signal, and then, calculated the error signal of each transmission. Based on the excellent classification ability of CNN, the error signal is used for training to identify 7 commercial Zigbee devices successfully. Youssef et al. [6] used CNN and support vector machine (SVM) to complete the identification of 12 OFDM devices, respectively. Yu et al. [7] exploited multisampling CNN to realize the identification of 54 IoT devices. However, when CNN is used in RFID, it is usually necessary to transform the RF baseband signal to ensure authentication accuracy, which will increase the complexity of the overall scheme. Another drawback of the above schemes is that CNN tends to exploit the local features but ignores the relevance of RF data in the time dimension and fails to make effective use of the temporal information contained in RF fingerprints. When RF baseband signals are directly used for feature extraction, CNN can only learn partial effective features. As a result, it is difficult to accurately identify massive similar devices in IoT environment.

LSTM is an improved network employing the temporal information of the sequential data. The RF context information in the time dimension is memorized by applying the previous information to calculate the current output. Then, the long-term dependency features hidden in the RF baseband signals are explored to further expand the feature space. Therefore, before the RF data is input into CNN, combining the temporal information perception ability of LSTM is expected to achieve accurate identification of similar IoT devices.

This paper proposes an RFID scheme for massive IoT devices based on LSTM-CNN. Concretely, we preprocess the baseband IQ signal and leverage LSTM as a feature extractor to automatically extract the initial features with temporal information. Then, CNN is connected as a classifier for secondary feature extraction and classification. The simulation results show that the proposed scheme can achieve 99.68% device identification accuracy on the open RF data set ORACLE and has a performance improvement of 6%-30% compared with other common network models (LSTM, CNN, and CNN-LSTM).

The rest of this paper is organized as follows. Section 2 introduces the system model and analyzes the defects of image classification neural networks represented by CNN when it is directly used in RF signal classification. Section 3 elaborates the details of the RFID scheme based on LSTM-CNN proposed in this paper. In Section 4, we explain the source of the dataset and compare the performance of RFID methods based on different network models. Finally, we conclude the paper in Section 5.

2. System Model

Figure 3 shows a general system model of RFID based on deep learning represented by CNN. The baseband IQ signals are forwarded to the AP via the wireless channel. Although the design and manufacture of the integrated circuit are developing, there are still differences [8] in the RF features of different devices at the moment of signal launch. To enhance the recognizability of the RF feature, it is usually necessary to synchronize and filter the baseband IQ signal. Then, CNN uses different sizes of sliding windows (convolution kernels) to perform convolution operations on the input signal samples. In this way, the local features in the samples are extracted. Finally, the local features are combined into global features through the fully connected layer to identify which device the sample comes from.

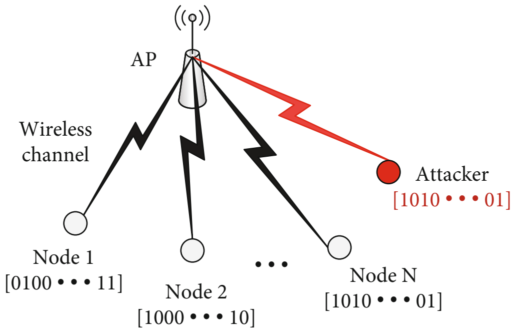


Figure 1: The impersonation attack.

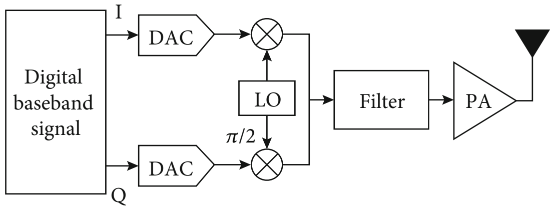


Figure 2: The source of RF fingerprint in the transmitter.

However, in the wireless environment, the collected RF data are commonly long-sequence samples, and hardware imperfections such as carrier frequency offset will affect the whole signal [9]. As we have discussed, the above CNN-based methods will ignore the time correlation of RF signals to a certain extent, resulting in a significant decline in CNN identification accuracy when IQ baseband signal is directly used for training [10]. In order to realize the low-cost and accurate identification of IoT devices, it is necessary to improve the capability of the neural network to perceive the long-term features of the sequences and therefore expand the feature space.

LSTM is a neural network that can perceive the context information of the input data. The network structure is shown in Figure 4, which is sequentially linked by multiple units. LSTM retains the context information of RF data and controls the amount of memory information through "gating." Each unit usually contains three "gates." The "forget gate" selectively forgets the information of the previous node. The "input gate" selectively memorizes the new input information and selectively retains the effective RF characteristics, while the "output gate" controls the response output of the neural network at time t.

$$\begin{aligned} f_t &= \sigma\left(w_f[h_{t-1}, x_t] + b_f\right), \\ i_t &= \sigma\left(w_i[h_{t-1}, x_t] + b_i\right), \\ o_t &= \sigma\left(w_o[h_{t-1}, x_t] + b_o\right). \end{aligned} \tag{1}$$

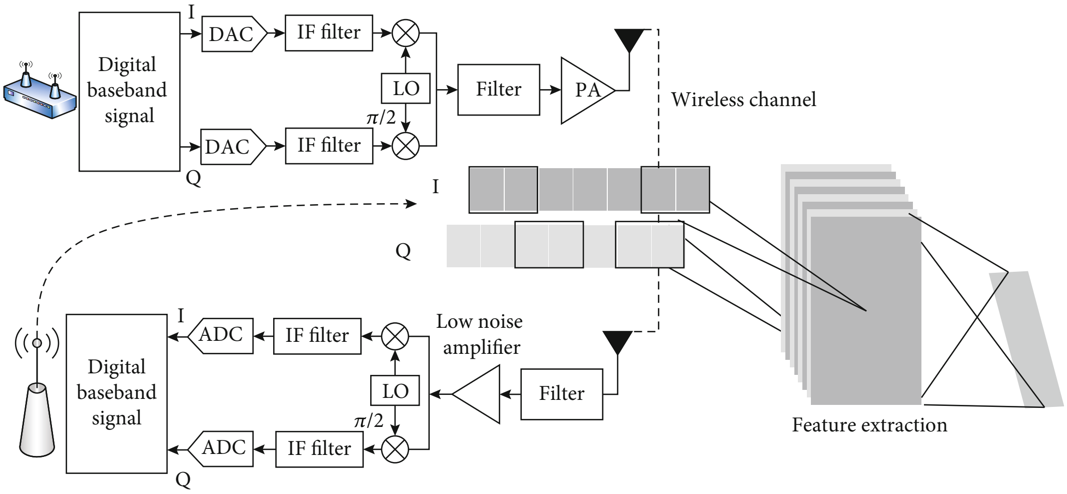


Figure 3: System model.

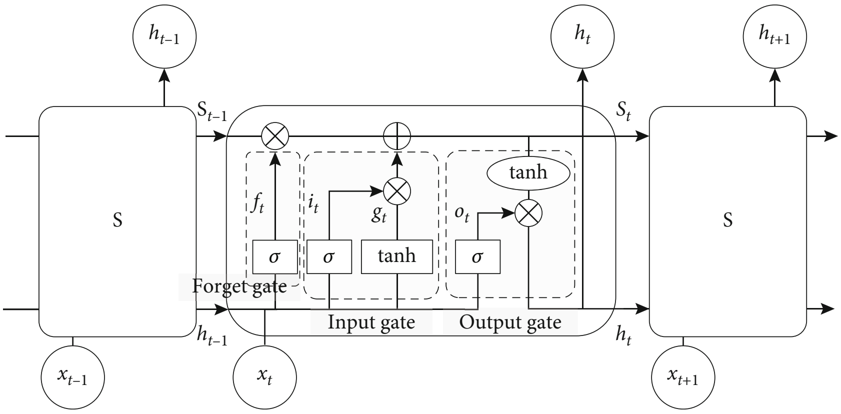


Figure 4: LSTM algorithm structure.

The current transmission state S_t with the results of the input and forget gate applied can be expressed as

$$S_t = f_t * S_{t-1} + i_t * \tanh\left(w_S[h_{t-1}, x_t] + b_S\right),$$
$$h_t = o_t \tanh\left(S_t\right). \tag{2}$$

Through the cooperation of the above three gates, LSTM retains the time information in the RF sequence signal by its "memory" function and optimizes the gradient disappearance problem [11] relying on the "forgetting" function. However, LSTM usually requires higher time complexity to achieve favorable identification accuracy than CNN due to its recurrent structure.

As discussed, we combine the advantages of the above two network models and propose an RFID scheme for IoT devices based on LSTM-CNN. The combination of LSTM and CNN has two significant advantages. On the one hand, the long-term dependence characteristics of RF data can be effectively utilized. On the other hand, the complexity of network training can be reduced.

3. RFID Scheme for Massive IoT Devices Based on LSTM-CNN

RFID scheme for massive IoT devices based on LSTM-CNN is shown in Figure 5, which includes two stages: offline training and online identification. The offline training stage can be divided into two steps: data preprocessing and the establishment of LSTM-CNN authentication model. Specifically, the IQ baseband data is preprocessed to make it suitable for the input format of the neural network, and then it is fed into LSTM-CNN for training. We employ the weights of the deep neural network to map RF features and thus build a learning-based fingerprint database. Therefore, a complete authentication model with a feature database is established. In the online identification stage, the RF data of the unknown node is input into the well-trained LSTM-CNN authentication model. Then, the model detects the similarity between the unknown data and the features stored in the database. The LSTM-CNN model will make the final classification decision according to the similarity detection result, thereby realizing the RFID of the massive IoT devices. The details of the two stages are as follows.

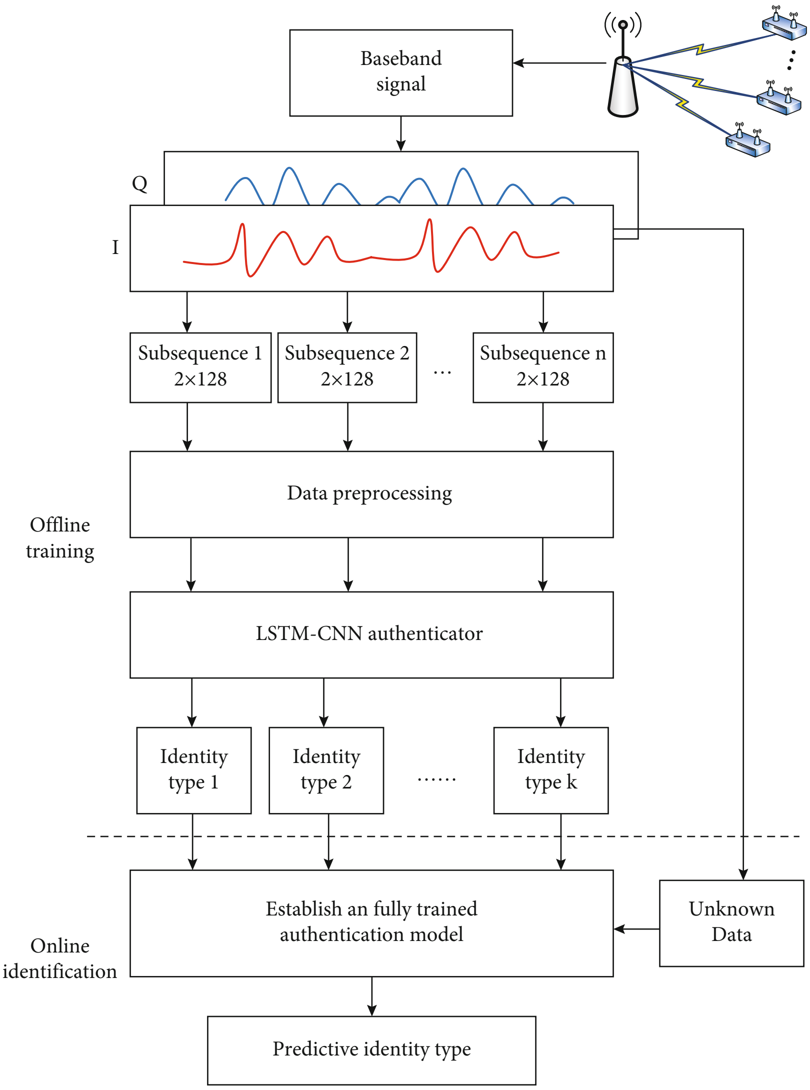


Figure 5: RFID scheme for massive IoT devices based on LSTM-CNN.

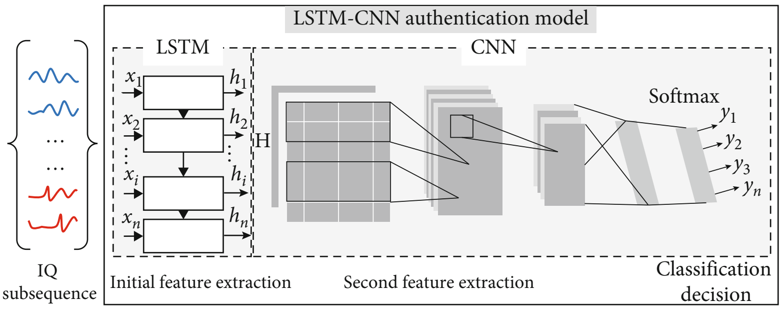


Figure 6: LSTM-CNN authentication model.

3.1. Offline Training

3.1.1. Data Preprocessing. First, we divide the long-baseband IQ signal into subsequences with an equal length of 128. To improve the learning efficiency and accelerate the convergence speed, we conduct the maximum-minimum normalization to linearly transform the original data and limit the data to [0, 1] interval. Then identity labels are added to the data to distinguish the RF information from different IoT devices. Finally, the subsequence set with labelled information is obtained to adapt to the input format of the neural network.

3.1.2. Establishment of LSTM-CNN Authentication Model. As shown in Figure 6, the LSTM-CNN authentication model consists of a pre-LSTM and a post-CNN structure. The

```
Input: RF data set $\omega$.
Offline training stage:
1. Slice $\omega$ to obtain RF signal subsequences set $\omega'$, $\dim_{\omega'}$=[2, 128]
2. Normalize $\omega'$ to get $X$ for the input of the LSTM-CNN model
3. Add label $i$ to the data set, $i \in [0, N]$
4. Training process:
(1) Input $X$ and labels into the LSTM units for eigenvectors set $H$ with temporal information
(2) Input $H$ into the CNN network to obtain the classification result on the training set
(3) Adjust the network model parameters according to loss function to obtain a fully trained LSTM-CNN model
Online identification stage:
for new RF data do:
Get the predictive probability set $P$ with trained authentication model, $P = \{P_1, P_2, \cdots, P_N\}$
if max $\{P_1, P_2, \cdots, P_N\} \geq \alpha$ then
$P_y$ = max $\{P_1, P_2, \cdots, P_N\}$, the label type corresponding to $P_y$ is the classification result
else
Judging that the identity of the unknown node is illegal
Output: identity type of the unknown node.
```

Algorithm 1: RFID algorithm based on LSTM-CNN.

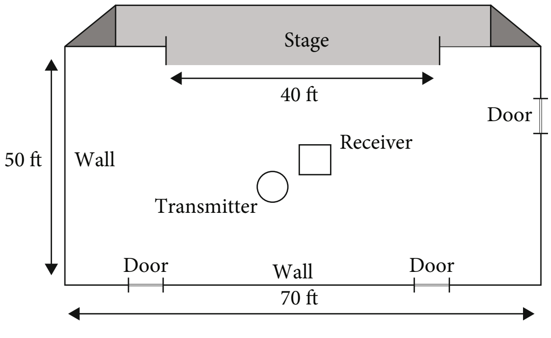


Figure 7: Open RF data source.

establishment of LSTM-CNN authentication model includes three stages: initial feature extraction stage, secondary feature extraction stage, and classification decision stage. The preprocessed IQ subsequence is fed into the pre-LSTM for initial feature extraction. Then, the secondary feature extraction is carried out through the convolution layer, pooling layer, and full connection layer in the post-CNN structure. Finally, the classification decision is made through the softmax function. The specific process is as follows.

(1) Initial Feature Extraction. The preprocessed subsequence set $X = \{x_1, x_2, \cdots, x_n\}$ is fed into the LSTM network and processed by multiple sequentially linked storage units. The information of the previous data remaining in the storage memory is controlled by the "forget gate," while the "input gate" controls new information added at the next moment. The initial feature vector $H = \{h_1, h_2, \cdots, h_n\}$ with time attributes is finally extracted through the "output gate."

(2) Secondary Feature Extraction. The long-term dependency feature H with time information obtained via LSTM is input into CNN for convolution operation. The local feature information is extracted through multiple convolutions to further amplify the difference of RF fingerprints. In order to reduce the complexity of the network and minimize the computation, downsampling is implemented in the pooling layer by max pooling. That is, the maximum feature value in the pooling window is used to replace the network output in this area to achieve dimensionality reduction. After that, all features are combined in the fully connected layer, and the local features are merged into global features.

(3) Classification Decision. The global features are input into the softmax classifier to realize the final decision on device identity. The output of the softmax function is the classification probability, expressed in the form of a vector. Each element value in the vector is in the interval of [0, 1], and the sum of all elements is 1. The softmax function [12] is expressed as follows:

$$P(Y = k | X = x_i) = \frac{e^{x_i}}{\sum_k @^{x_k}}, \tag{3}$$

where k is the device classification type, $k \in [0, N]$, and N is the number of devices. The cross-entropy loss function is used to judge the quality of the model output so as to further optimize the model parameters. The loss function L_i is related to the proportion of correct classification results, and the small value indicates better model performance.

$$L_i = -\log\left(\frac{e^{s_{y_i}}}{\sum_k @^{s_k}}\right). \tag{4}$$

In the training process of LSTM-CNN model, it is necessary to select the appropriate size and number of hidden layers to check whether the model can gradually converge. If the loss function decreases unsteadily or cannot converge to a reasonable interval, the size and number of hidden layers should be increased appropriately to improve the fitting ability of the model. The size of the learning rate will also impact the model convergence process. If the value of

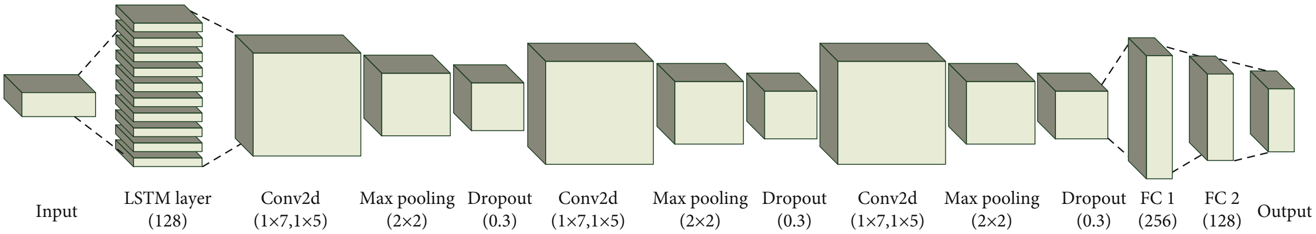


Figure 8: Network parameters.

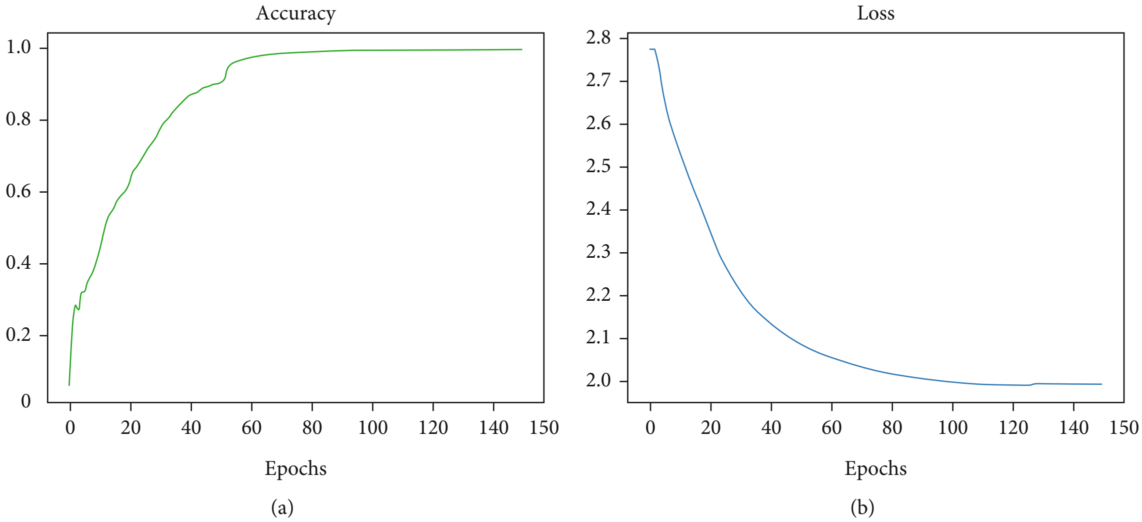


Figure 9: LSTM-CNN performance.

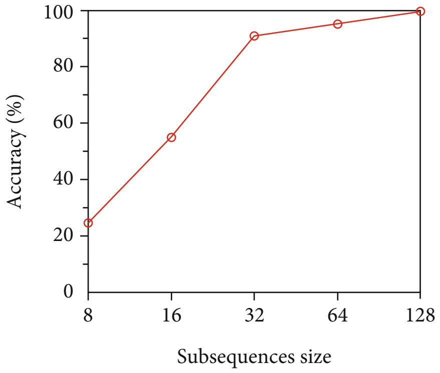


Figure 10: LSTM-CNN's classification accuracy for different subsequences size.

Table 1: Comparison with other popular classification methods.

Schemes	RF data source	Model	Accuracy
Hossein et al. [13]	6 Zigbee devices (MICAz)	DNN	93.9%
		CNN	94%
		LSTM	73%
Wang et al. [14]	8 RF devices embedded in the NRF24LE1E chip	TL-LSTM	81%
Our scheme	16 USRP X310 devices (ORACLE)	LSTM-CNN	99.68%

loss function fluctuates violently in the convergence process, the learning rate is supposed to be reduced. Inversely, the learning rate should be moderately increased if the model converges too slowly. Furthermore, in this paper, the dropout mechanism is introduced to prevent the model from overfitting.

3.2. Online Identification. Once the LSTM-CNN authentication model is established, a probabilistic method is developed to predict the identity of IoT devices based on the fingerprint database and new RF data. The softmax probability vector indicates the similarity between the unknown RF node and the database. When the unknown node connects to the network, the RF baseband signal collected on the AP side is input into the well-trained LSTM-CNN authentication model to detect the similarity with the feature database. The output of the model is expressed as a probability set $P = \{P_1, P_2, \cdots, P_N\}$, where P_i means the similarity between the unknown node and the i-th device in the database.

If $\max \{P_1, P_2, \cdots, P_N\} < \alpha$, that is, the RF characteristics of the unknown node are less similar to the stored devices. Therefore, the unknown node is judged to be an illegitimate device, where α represents the decision threshold of the network model.

Table 2: Comparison of different network models.

Model	Architecture	Learning rate	Accuracy
CNN	Conv1 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv2 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv3 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) FC1 (256, 128) FC2 (128, 16)	0.001	85.41%

LSTM	LSTM(128)
	FC(128, 16)
		0.0001	64.89%

CNN-LSTM	Conv1 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv2 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv3 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) LSTM(256) FC1 (256, 128) FC2 (128, 16)	0.001	93.36%
LSTM-CNN	LSTM(128) Conv1 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv2 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) Conv3 $(1 \times 7, 1 \times 5)$ Maxpooling (2, 2) FC1 (256, 128) FC2 (128, 16)	0.001	99.68%

Otherwise, $P_y = \max \{P_1, P_2, \cdots, P_N\}$, and the corresponding device type in the database is determined to be the identity of the unknown node.

Finally, according to the classification result obtained by the authentication model, the AP decides whether to allow unknown nodes to access the network, thereby ensuring wireless security.

3.3. Overall Procedure. The overall procedure of the RFID scheme for IoT devices based on LSTM-CNN is shown in Algorithm 1.

4. Simulation Implementation

In this section, we conduct experiments on the open RF dataset ORACLE to verify the proposed scheme. Then we analyze the performance of RFID schemes based on CNN, LSTM, and CNN-LSTM to compare with the proposed LSTM-CNN schemes. The source of the dataset, network parameters, and the comparison and analysis of experimental results are introduced in detail below.

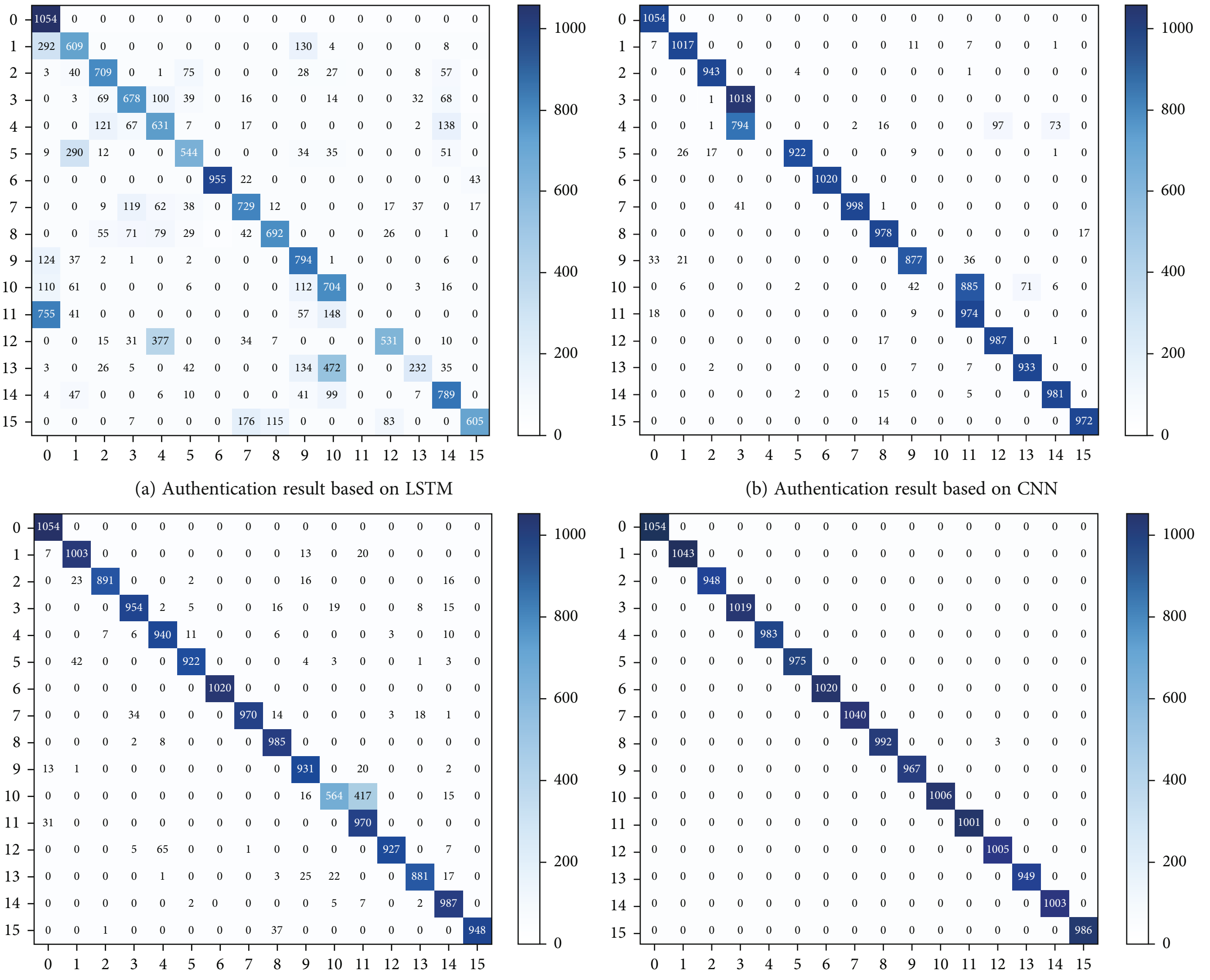
4.1. Source of the Dataset. The dataset we use in this paper comes from the open RF data source ORACLE collected by Sankhe et al. [10]. As shown in Figure 7, the scene of the signal acquisition is an indoor environment with less reflection. At the same time, there are channel fading and multipath effects. The framework for IEEE 802.11a compliant datasets is generated by the MATLAB WLAN toolbox. 16 USRP X310 transmitters and a USRP B210 radio receiver are used for data acquisition. The signal is transmitted at the radio frequency of 2.45 GHz and the sampling rate is 5 ms/s. This paper uses the baseband IQ samples obtained when the distance between communication parties is 2 ft.

4.2. Network Parameters and Results. The LSTM-CNN model parameters are shown in Figure 8: including 1 LSTM layer, 3 convolutional layers, 3 pooling layers, 3 dropout layers, and 2 fully connected layers. The total number of datasets is 80,000, and the input size of a single sample is 2×128. The training set, validation set, and test set are divided in a ratio of $7:1:2$. The RF dataset is fed into the LSTM network with a hidden layer size of 128, while the batch size is 32, then LSTM outputs a feature matrix of $32 \times 128 \times 128$. The convolution kernel size of each convolutional layer is shown in Figure 8. After each convolutional layer, the ReLU activation function is connected to perform the nonlinear transformation, and the transformation result is input to the pooling layer for downsampling. In order to prevent overfitting, the dropout rate of the model is set to 0.3, and the training result is shown in Figure 9. It can be seen that with the increase of epoch, the loss gradually decreases to convergence, and the final test accuracy of the model is 99.68%, which means this scheme can well identify the fingerprint difference from the baseband IQ samples and implement classification of massive similar IoT devices.

Figure 10 shows LSTM-CNN's classification accuracy for different length of input subsequences size, where the amount of information increases as the increasing subsequence length.

4.3. Comparison with Other Popular Classification Methods. As mentioned earlier, with the development of deep learning, the mainstream classification models used in RFID schemes include DNN, CNN, and LSTM. Table 1 shows the comparison between our scheme and other popular schemes, where Hossein et al. [13] considered three different deep learning models, DNN, CNN, and LSTM, to identify 6 similar Zigbee devices. Wang et al. [14] used LSTM to identify 8 RF devices and combined LSTM with transfer learning (TL) to solve the problem of small sample training.

In order to further verify the gain brought by the LSTM-CNN model to RFID, this paper compares the performance of three other network models (CNN, LSTM, and CNN-LSTM) based on the same RF dataset. The parameter settings are shown in Table 2. The dimension of the hidden layer in each model is related to the input dimension. In order to increase the comparability of the experiment, we maximally unify the network parameters of the models. So the dimension of the LSTM is adjusted to 256 to adapt to changes of the network structure in the CNN-LSTM model.



(a) Authentication result based on LSTM

(b) Authentication result based on CNN

(c) Authentication result based on CNN-LSTM

(d) Authentication result based on LSTM-CNN

Figure 11: Authentication results.

The calculation method of the final authentication accuracy is shown in Formula (5), where Y is the actual device category, Y_{pred} is the predicted device category, and num_{test} is the total number of test data.

$$\text{Accuracy} = \frac{\text{sum}\left(Y_{\text{pred}} = Y\right)}{\text{num}_{\text{test}}}. \quad (5)$$

The authentication accuracy of different network models is shown in Table 1. Compared with others, the performance of the LSTM-CNN authentication model is improved by 6%-30%.

When different networks are used as authentication models, the performance on the same RF dataset is shown in Figure 11, where (a), (b), (c), and (d) show the confusion matrix of the authentication results based on LSTM, CNN, CNN-LSTM, and LSTM-CNN, respectively. The confusion matrix is a standard format for accuracy evaluation, in which the rows and columns, respectively, represent the actual and prediction categories. The performance is proportional to the concentration of the value on the diagonal.

It can be seen that when LSTM or CNN works alone, the distribution of the confusion matrix is relatively scattered. Compared with CNN, the distribution of the confusion matrix based on CNN-LSTM is more concentrated and the authentication accuracy is improved. However, the way of prepending CNN may lose some time information, thus fails to achieve the best authentication accuracy. The confusion matrix of LSTM-CNN shows that the authentication results are highly concentrated on the diagonal, which verifies the model we proposed can make good use of the temporal correlation of RF data to expand the feature space.

5. Conclusions

Aiming at the inconspicuous discrimination issue of RF fingerprints among massive devices in IoT environment, this paper proposes an RFID scheme based on LSTM-CNN. Combining the capacity of LSTM on perceiving the context information and the excellent classification advantages of CNN, this scheme makes better use of the long-term dependency of RF signals in the time dimension to achieve the accurate identification of massive IoT devices. Furthermore, this paper compares the performance of RFID schemes based on LSTM, CNN, and CNN-LSTM. The simulation results show that the performance of the LSTM-CNN model is improved by 6%-30% compared with other networks, which means the scheme we proposed can better explore the characteristics of RF signals and achieve high-precision identity authentication.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

No conflict of interest exists in the submission of this manuscript, and the manuscript is approved by all authors for publication.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (61871404).

References

[1] S. Shrestha, E. Irby, R. Thapa, and S. Das, "SoK: a systematic literature review of Bluetooth security threats and *mitigation measures*," in *Emerging Information Security and Applications. EISA 2021*, W. Meng and S. K. Katsikas, Eds., vol. 1403 of Communications in Computer and Information Science, Springer, Cham., 2022.

[2] G. Y. Li, J. B. Yu, and A. Q. Hu, "Research on physical-layer security based on device and channel characteristics," *Journal of Cryptologic Research*, vol. 7, no. 2, pp. 224–248, 2020.

[3] 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.

[4] H. J. Patel, M. A. Temple, and R. O. Baldwin, "Improving zig bee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting," *IEEE Transactions on Reliability*, vol. 64, no. 1, pp. 221–233, 2015.

[5] 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.

[6] K. Youssef, L. Bouchard, K. Haigh, J. Silovsky, B. Thapa, and C. V. Valk, "Machine learning approach to RF transmitter identification," *IEEE Journal of Radio Frequency Identification*, vol. 2, no. 4, pp. 197–205, 2018.

[7] J. Yu, A. Hu, G. Li, and L. Peng, "A multi-sampling convolutional neural network-based RF fingerprinting approach for low-power devices," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, Paris, France, 2019.

[8] J. Hall, M. Barbeau, and E. Kranakis, "Enhancing intrusion detection in wireless networks using radio frequency fingerprinting," *Communications, Internet, and Information Technology*, vol. 1, 2004.

[9] A. Goldsmith, *Wireless Communications*, Cambridge university press, 2012.

[10] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: optimized radio classification through convolutional neural networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 370–378, Paris, France, 2019.

[11] T. Ergen, A. H. Mirza, and S. S. Kozat, "Energy-efficient LSTM networks for online learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 3114–3126, 2020.

[12] A. Joulin, M. Cissé, D. Grangier, and H. Jégou, "Efficient softmax approximation for gpus," in *International conference on machine learning*, pp. 1302–1310, 2017.

[13] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "IoT devices fingerprinting using deep learning," in *2018 IEEE Military Communications Conference (MILCOM)*, pp. 1–9, Los Angeles, CA, USA, 2018.

[14] X. Wang, Y. Zhang, H. Zhang, Y. Li, and X. Wei, "Radio frequency signal identification using transfer learning based on LSTM," *Circuits, Systems, and Signal Processing*, vol. 39, no. 11, pp. 5514–5528, 2020.