Research Article

Automatic Modulation Classification Based on Deep Learning for Software-Defined Radio

Peng Wu,[1] Bei Sun [ID],[1] Shaojing Su,[1] Junyu Wei,[1,2] Jinhui Zhao,[3] and Xudong Wen[4]

[1]*College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China*
[2]*College of Electric and Information Engineering, Hunan University of Technology, Zhuzhou 412007, China*
[3]*Teaching and Research Support Center, National University of Defense Technology, Changsha 410073, China*
[4]*93920 Unit of the Chinese People's Liberation Army, Hanzhong, China*

Correspondence should be addressed to Bei Sun; sunbei08@nudt.edu.cn

With the development of artificial intelligence technology, deep learning has been applied to automatic modulation classification (AMC) and achieved very good results. In this paper, we introduced an improved deep neural architecture for implementing radio signal identification tasks, which is an important facet of constructing the spectrum-sensing capability required by software-defined radio. The architecture of the proposed network is based on the Inception-ResNet network by changing the several kernel sizes and the repeated times of modules to adapt to modulation classification. The modules in the proposed architecture are repeated more times to increase the depth of neural network and the model's ability to learn features. The modules in the proposed network combine the advantages of Inception network and ResNet, which have faster convergence rate and larger receptive field. The proposed network is proved to have excellent performance for modulation classification through the experiment in this paper. The experiment shows that the classification accuracy of the proposed method is highest with the varying SNR among the six methods and it peaks at 93.76% when the SNR is 14 dB, which is 6 percent higher than that of LSTM and 13 percent higher than that of MentorNet, Inception, and ResNet purely. Besides, the average accuracy from 0 to 18 dB of the proposed method is 3 percent higher than that of GAN network. It will provide a new idea for modulation classification aiming at distraction time signal.

1. Introduction

With the rapid development of communication technology, the wireless communication environment is becoming more and more complex and communication signals with various types of modulation are becoming more diverse and complex [1]. Automatic modulation classification (AMC) plays an important role in modern wireless communication [2]. The signal analysis and processing can be carried out only when the modulation of the signal is recognized [3]. It finds applications in various commercial and military areas. For example, recognition of the modulation type is used in software-defined radio (SDR) in order to adapt to various communication systems without requiring control overhead in the shortest time [4, 5]. Under such conditions, advanced automatic modulation classification techniques are required.

It is also essential for identifying the source of received wireless signals [6, 7].

At present, AMC can be divided into two categories: likelihood-based (LB) and feature-based (FB) [8]. The LB modulation classifier recognizes the modulation of signal by comparing the likelihood function value of received signal within the known modulation pool [6]. It has been used for modulation classification in multiple channel environment with high accuracy [9]. While it needs some parameters to be known in advance, such as carrier frequency, code rate, and channel parameters, it becomes very complex when unknown parameters are introduced. Thus, it is difficult to design a system for the acquisition of signal. Some researchers have studied the way to simplify the likelihood function, which will lead to information absence and inaccurate results [10]. As LB method is sensitive to parameter

estimation deviations or model mismatches, it is not applicable in many practical communication scenarios.

For FB method, features of the received signal are extracted and the modulation of signal can be identified by either comparing features with threshold values or feeding feature to pattern recognizer [11, 12]. In many traditional pattern recognition methods, it is necessary to extract features of signal manually, such as instantaneous statistics, high-order statistics, time-frequency characteristics, asynchronous delay sampling characteristics, etc. [13]. Then, these features are used as input of the classifier, such as decision tree and support vector machine [14]. Although it is simple with less computation, it shows poor performance for non-linear problems. Apart from this, features selected manually may not reflect the characteristics of signals with different modulation and improper feature selection will reduce the classification recognition accuracy of the classifier.

In recent years, great progress has been made in artificial intelligence and a single computer chip's computing power has been greatly improved, which promotes deep learning algorithms to be widely used in modulation classification [15, 16]. It solves the core problem of how to automatically select and extract the features of samples. Besides, it realizes the combination of simple features into more efficient and complex features to achieve classification recognition [17]. In addition, deep neural networks have a multi-layer structure, which can better extract features of signal avoiding the tedious manual selection of data features [18]. At present, CNN, Google's Inception, and residual network (ResNet) have been used in modulation classification with good results.

CNN-based modulation classification method was proposed and it can be used for the sampling sequence of the intermediate frequency signal directly [19]. Rajendran et al. proposed that recurrent neural network (RNN) can also be used for modulation classification of the sampling sequence of the intermediate frequency signal [20]. Hu et al. studied the effect of different noise on the modulation classification by using RNN [21].

Generated adversarial network (GAN) was proposed for data argument with RGB three-channel constellation in modulation classification filed [22]. Wang et al. established a 2-level convolutional neural network (CNN) architecture to distinguish 16QAM from 64QAM. In the second CNN architecture, 16QAM and 64QAM constellation diagrams were used as input to obtain a higher recognition rate [23]. However, this method relies on the constellation diagram and can only identify the baseband signal. The accuracy is very low when modulated signals are sampled on the radio frequency directly. A convolutional long short-term deep neural network (CLDNN) was proposed by combining the architectures of CNN and long short-term memory (LSTM) into a deep neural network, which takes advantage of CNNs, LSTMs, and conventional deep neural network architectures and enables the learning of long-term dependencies [24]. A ResNet architecture was proposed to distinguish between different modulation types by adding the bypass connections and showed good performance [17]. Shengliang P proposed

two convolutional neural network- (CNN-) based DL models, AlexNet and GoogLeNet [25]. Cheong et al. extended the previous work on automatic modulation classification (AMC) by using deep neural networks (DNNs) and evaluated the performance of these architectures on signals [26]. Xie et al. proposed two algorithms, including M2M4-aided algorithm and multi-label DL based algorithm, to combat the varying SNR [27]. Peng et al. converted the raw modulated signals into images that had a grid-like topology and fed them to CNN for network training [28].

The dataset used in this paper was released on Proceedings of the GNU Radio Conference in 2016 for applications of machine learning (ML) to the Radio Signal Processing domain [29]. Li et al. used GAN network with data argument for modulation classification testing on this dataset and the highest accuracy was no more than 90% [30]. LSTM, Inception, and residual network were also used for modulation classification on this dataset, while the highest accuracy among them was around 88%. Miao Du proposed a new network structure called fully dense neural network (FDNN) for the automatic modulation classification and the average accuracy with FDNN is 89.6% from 0 to 18 dB [31]. Yuan Zeng proposed CNN architecture with spectrogram images as signal representation and achieved good recognition accuracy in 2019, while the average accuracy from 0 to 18 dB of it was less than 90% [32]. Wu et al. proposed convolutional neural network and multi-feature fusion for automatic modulation classification. The overall classification accuracy of it is less than 93% [33]. In addition, Zhang et al. proposed an automatic digital modulation classification network based on curriculum learning and the overall classification accuracy is less than 90% [34]. The highest accuracy of the proposed method in this paper is 93.76% at 14 dB and the average accuracy from 0 to 18 dB with the proposed method is 93.04%, which has excellent performance.

This paper is organized as follows: Section 2 introduces the related work. Section 3 formulates the basic principle of the ResNet. Section 4 gives the basic principle of Inception network. Section 5 details the proposed method in this paper. Section 6 presents the results of simulations and experiments to support the theoretical analysis.

2. Related Work

2.1. Basic Principle of Signal Modulation. At present, digital modulation technology has been used widely in wireless communication. Although signal processing techniques such as modulation and demodulation of digital modulation are more difficult than analog modulation and the signal processing system is more complex for digital modulation, digital modulation technology has strong anti-interference and it is easy to use modern digital signal processing technology to process and analyze the signal. Therefore, digital modulation signals are widely used in practice. The modulation classification of digital modulation signals is studied in this paper.

A general expression for the received baseband complex envelope is

$$r(t) = s(t; \mathbf{u_i}) + n(t), \quad (1)$$

where $s(t; \mathbf{u_i})$ denotes baseband complex envelope of the received signal with no noise, and $n(t)$ is noise. $s(t; \mathbf{u_i})$ can be expressed as

$$s(t; \mathbf{u_i}) = a_i e^{j2\pi\Delta f t} e^{j\theta} \sum_{k=1}^{K} e^{j\phi_k} s_k^{(i)} g(t - (k-1)T - \varepsilon T), \quad 0 \le t \le \mathrm{KT}, \quad (2)$$

where a_i represents the amplitude of signal, Δf denotes the offset of carrier frequency, θ is the time-invariant carrier phase, ϕ_k is the phase jitter, T is the symbol period, ε is the normalized epoch for time offset between the transmitter and signal receiver, $g(t) = P_{\text{pulse}}(t) \otimes h(t)$ is the composite effect of the residual channel with $h(t)$ denoting the channel impulse response and $\otimes$ denoting mathematical convolution, and $P_{\text{pulse}}(t)$ is the transmit pulse shape. In equation (2), $\mathbf{u_i}$ is the multidimensional vector that includes the deterministic unknown signal or channel parameters for the ith modulation type. The goal of modulation classification is to recognize the modulation type i from the received signal $r(t)$ [30]. The Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK), Phase Shift Keying (PSK), and Quadrature Phase Shift Keying (QPSK) are commonly used modulation types.

2.2. Deep Neural Network for Classification Modulation. Convolutional layers are a common element in all state-of-the-art deep neural networks. A convolutional layer usually consists of convolutional filters. The size of convolution kernels is typically very small, such as 1×1 through 5×5 sizes. The transfer function for a standard convolutional layer [6] is given below:

$$y_i = f\left(b_j + \sum_i k_{ij} * x_i\right), \quad (3)$$

where y_i is the output feature map for the ith filter, b and k represent learned bias and filter weight parameters, x_i represents the input activations, $*$ denotes the convolution operation, and $f(\cdot\cdot)$ denotes a (typically non-linear) activation function [35].

A visible trend in neural networks for classification task is building deeper networks to learn more complex functions and hierarchical feature relationships. Deep networks enable more complex functions to be learned more readily from raw data.

Typically, applying deep neural networks to solve modulation classification is a matter of

(i) Designing a network architecture

(ii) Training the network to select weights which minimizes loss

(iii) Validating and testing in practice to solve problem

To this end, we use various machine learning classifiers based on deep neural network architectures, where a training dataset is used to train the network, and then the classification accuracy is computed over the classification output for a testing dataset.

Figure 1 illustrates a diagram of deep learning for modulation classification. Generally speaking, the classification method mainly includes three parts: signal preprocessing, feature extraction, and evaluation. The deep neural network can automatically combine simple basic features into more complex features gradually, achieve effective feature expression of data samples, and maintain high recognition accuracy. In this paper, a supervised recognition method is used. Firstly, a large number of labeled samples are used to train the deep neural network, and then the trained model is used to recognize the unknown samples.

3. Residual Network

Although neural networks with more layers have better learning ability, the degeneration occurs sometimes, which leads to low accuracy. An effective approach so far, which won ImageNet 2015, is residual networks. A residual network adds one layer's output to the output of the layer two layers deeper, as shown in Figure 2. Vanishing gradients are resolved by normalization techniques that have been widely adopted and that network depth is instead limited by training complexity of deep networks which can be simplified with residual functions [36].

ResNet improves the original network structure by adding connection identity mapping. The learning function is defined as $F(x) + x$ and the input is added to the learned function. By adopting this method, the parameters and calculation amount of the network will not increase, and the training speed and recognition accuracy of the model will be significantly improved.

4. Inception Network

The inception architecture is one successful approach to increasing the depth of network and learning ability. This network consists of repeated inception modules. As shown in Figure 3, each inception module contains four parallel paths with the output being the concatenation of the four parallel outputs. The first path is a bank of 1×1 convolutions that forward along selected information. The 1×1 convolutions are a form of selective highway networks that simply pass information forward with no transformation. The second and third paths are 1×1 convolutions followed by a bank of 3×3 and 5×5 convolutions to make the network have larger receptive field. The last parallel path is a 3×3 pooling layer followed by 1×1 convolutions. Intermediate
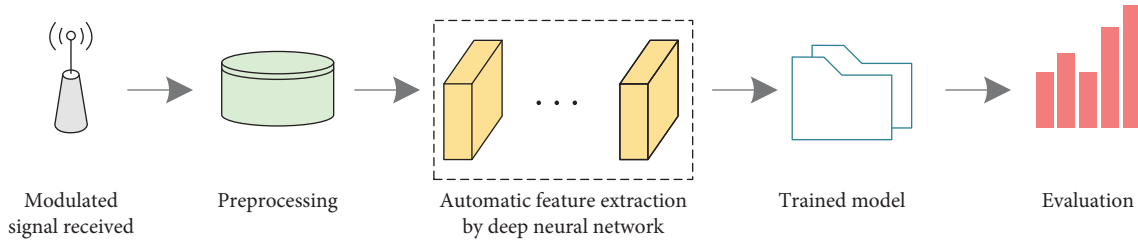
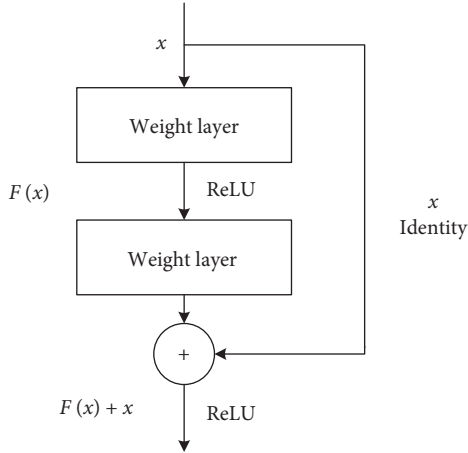FIGURE 1: Flow diagram of deep learning for modulation classification.

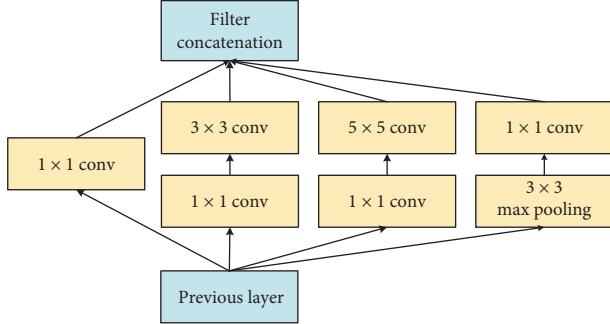FIGURE 2: Residual network diagram.

FIGURE 3: Inception unit diagram.

inception modules in the network are connected to Softmax classifiers that contribute to the network's global loss for training. These classifiers are believed to help in increasing the model's ability to learn features [37].

5. Proposed Method

5.1. The Structure of the Proposed Network. The architecture of the proposed network is based on the Inception-ResNet network by changing the several kernel sizes and the repeated times of modules to adapt to modulation classification for software-defined radio [38]. The architecture of the proposed network consists of stem network module, three Inception-ResNet modules, and reduction modules. This network combines the advantages of the Inception module and ResNet module, which has faster convergence rate. The network is proved to have excellent performance

for modulation classification through the experiment in this paper.

For the residual versions of the Inception networks, cheaper Inception blocks than the original Inception are used. Each Inception block is followed by filter-expansion layer (1×1 convolution without activation) which is used for scaling up the dimensionality of the filter bank before the addition to match the depth of the input. This is needed to compensate for the dimensionality reduction induced by the Inception block.

Figure 4(a) shows the overall schema of the proposed network; the input dimension of input data is $1 \times 2 \times 128$ which is set in the dataset for modulated signal. There are three Inception-ResNet modules in Figure 4(a); the structure of each Inception-ResNet module is shown in Figures 5(a)–5(c), respectively. The module A is repeated 10 times. 20 times and 10 times are repeated for module B and module C. In order to make residual operation feasible, 1×1 convolution is added to match the depth of the network in each Inception-ResNet module. The schema for the stem of the proposed method is shown in Figure 4(b), where there are 6 convolutional layers and 1 max pooling layer.

5.2. Loss Function. For a multi-class classification task such as modulation recognition, the objective function is often categorical cross-entropy. Categorical cross-entropy (equation (4)) is a measure of difference between two probability distributions. For deep learning classification tasks, the probability distribution is usually a Softmax (equation (5)) of the output of the classifier network which is then converted to one-hot encoding for classification purposes. The error is calculated in what is known as the forward-pass and weights are adjusted using the chain rule to find error contribution for each parameter in what is known as the backward-pass. This kind of network output layer, optimization, and loss function have been used very successfully for multi-class vision tasks such as object recognition on the ImageNet dataset [39].

$$H(p, q) = -\sum_x p(x) \log q(x). \tag{4}$$

In equation (4), $p(x)$ and $q(x)$ are two probabilities and $H(p, q)$ is a value to measure the difference of two probabilities.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}, \quad \text{for } j = 1, \ldots, K z_j. \tag{5}$$
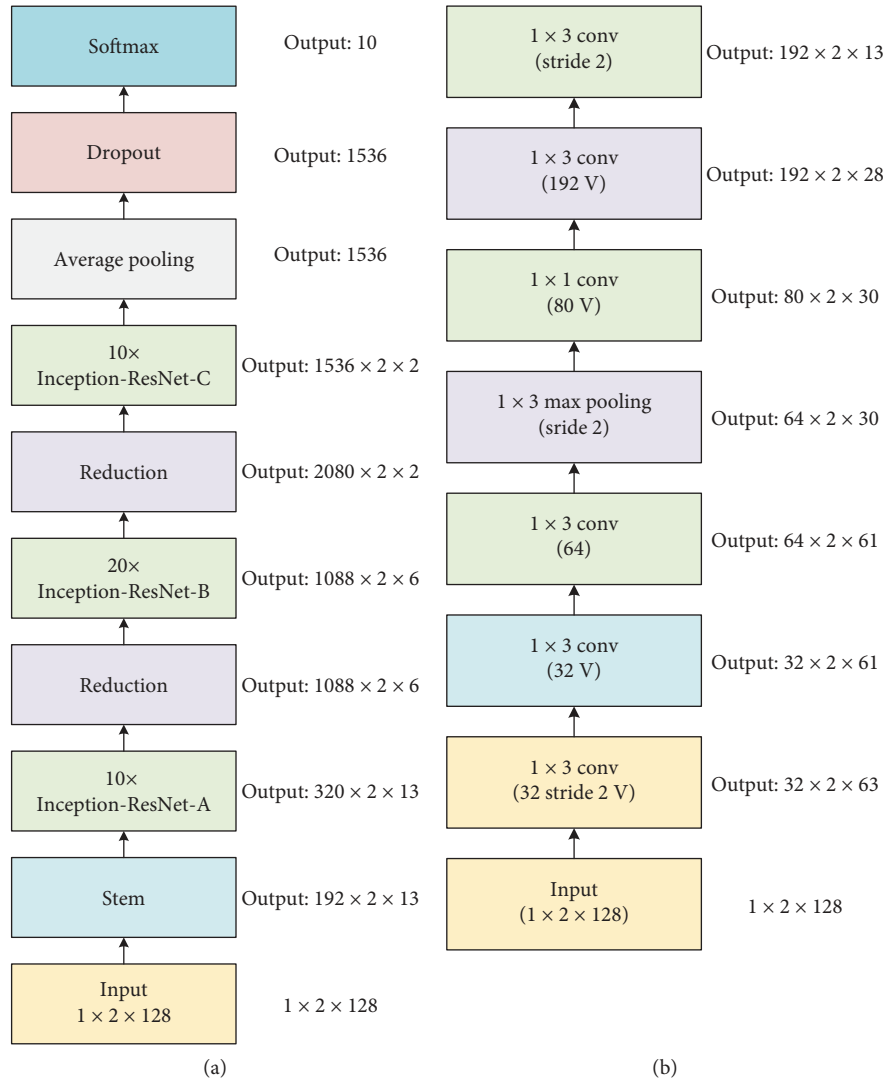
FIGURE 4: The overall schema and stem of the proposed network. (a) The overall schema of proposed network. (b) The schema for stem of the proposed network.

In equation (5), the standard exponential function is applied to each element of the input vector z and these values are normalized by dividing by the sum of all these exponentials; this normalization ensures that the sum of the components of the output vector $\sigma(z)$ is 1.

6. Experiment and Discussion

6.1. Dataset. We use the RadioML2016.10b dataset as a basis for evaluating the modulation recognition task. The dataset adopted in this paper was first released at the 6th Annual GNU Radio Conference in 2016. The dataset allows machine learning researchers with new ideas to dive directly into an important technical area without the need for collecting or generating new datasets and allows for direct comparison to efficacy of prior work. The dataset is generated with GNU Radio, consisting of digital and analog modulations at varying signal-to-noise ratios. Details about the generation of this dataset can be found in [30]. Figure 6 shows a high-level framework of the data generation.

The time segments were sampled randomly from the output stream of each simulation and stored in an output vector commonly used for Keras, Theano, and TensorFlow. The dataset uses a 4D real float 32 vector, taking the form Nexamples $\times$ Nchannels $\times$ Dim1 $\times$ Dim2, where each example consists of 128 complex floating point time samples. Nchannels = 1, a representation which commonly is used for RGBA values in imagery, Dim1 = 2 holding I and Q channels, and Dim2 = 128.

6.2. Training Process. The dataset is divided into two parts: 960000 samples are used for training the deep neural network and 240000 samples for validation. All models and training are done with the Keras with TensorFlow as a deep learning library using a TITAN RTX 24G GPU. The Adam optimizer was used for all architectures, and the loss function was the categorical cross-entropy function. We also used ReLU activation functions for all layers, except the last dense layer where we used Softmax activation functions. We used a minimum batch size of 128 and
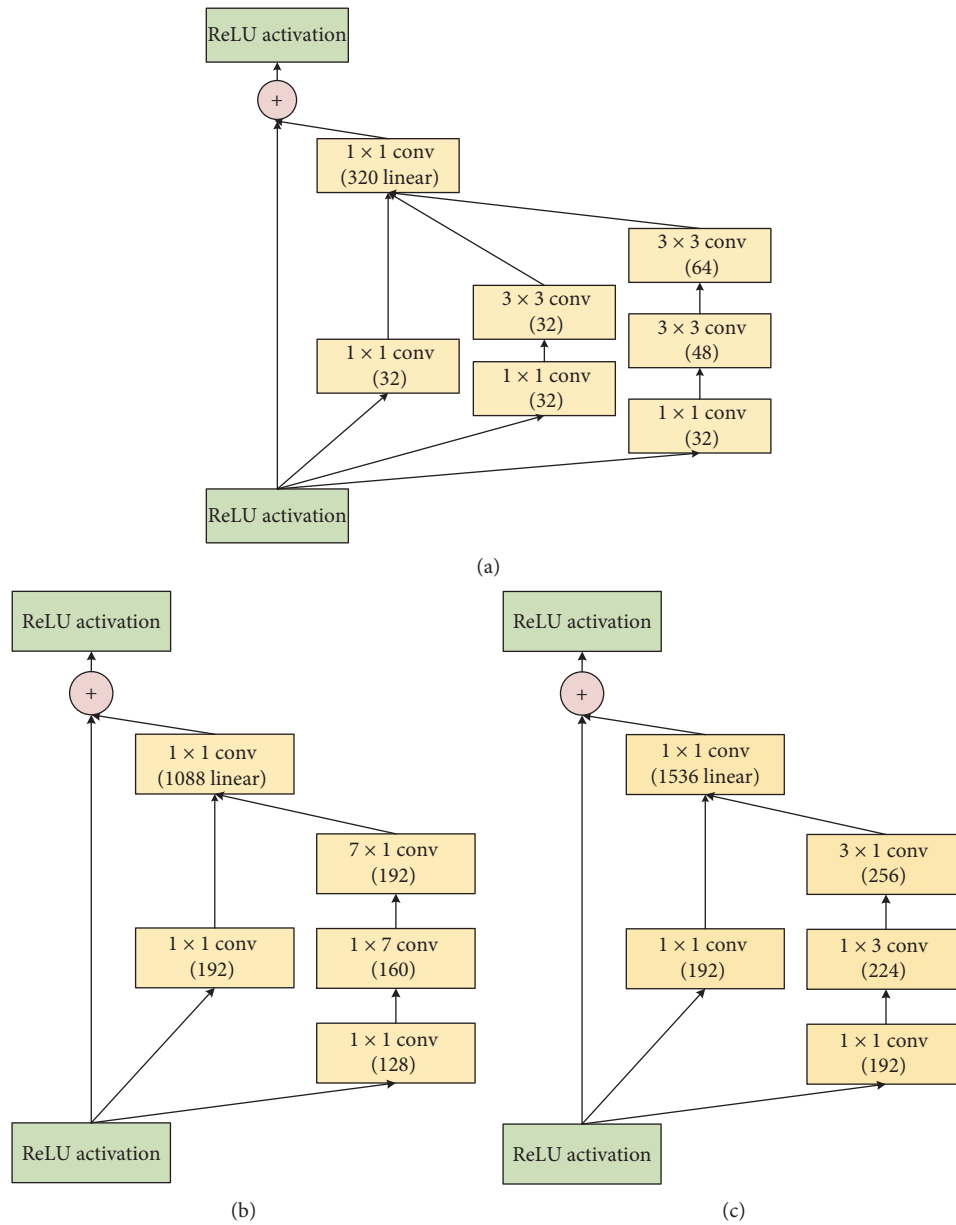
(a)

(b)

(c)

FIGURE 5: The schema for the three Inception-ResNet modules of the proposed network. (a) The schema for Inception-ResNet module (A). (b) The schema for Inception-ResNet module (B). (c) The schema for Inception-ResNet module C.
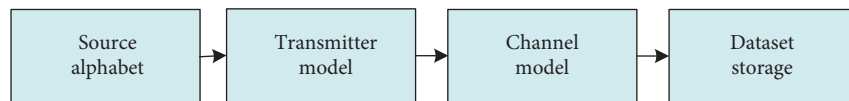

FIGURE 6: A frame of data generation.

a learning rate of 0.01 and 0.001. Figure 7 shows the flow diagram of our experiment.

6.3. *Trick during the Process.* In order to get better performance and avoid overfitting, we introduce several tricks in training.

(a) Batch Normalization

Normalize the input data between −1 and 1. This trick can be related to

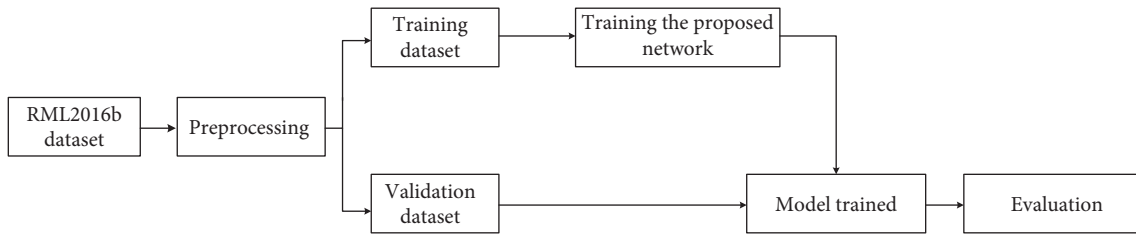$$x' = \frac{x - \text{mean}(x)}{\max(x) - \min(x)}, \quad (6)$$

FIGURE 7: The flow diagram of our experiment.

where x is the input data for the network and x' is normalized data. This trick can accelerate training and prevent overfitting.

(b) Early stopping

A compromise is to train on the training dataset but to stop training at the point when performance on a validation dataset starts to degrade. This simple, effective, and widely used approach to training neural networks is called early stopping. Stopping the training of a neural network early before it has overfit the training dataset can reduce overfitting and improve the generalization of deep neural networks.

The challenge of training a neural network long enough to learn the mapping, but not so long that it overfits the training data.
Model performance on a holdout validation dataset can be monitored during training and training can be stopped when generalization error starts to increase.
The use of early stopping requires the selection of a performance measure to monitor, a trigger to stop training, and a selection of the model weights to use.

In this paper, we stop training process when the validation loss does not decrease within 10 epochs. As shown in Figure 8, the validation loss is 0.8820 at the 8th epoch and the validation loss is more than 0.8820 in the following 10 epochs. In this connection, we adopt the parameters at the 8th epoch to evaluate the performance of the model.

(c) Dropout

Assuming that the probability of discarding neurons in each layer is p, then the remaining probability is $1 - p$. This trick can prevent the network from overfitting to a certain extent. In the dropout module in this paper, the remaining probability is 0.8.

6.4. Result and Discussion

6.4.1. Baseline. In this paper, the performance of the proposed method is compared with five methods including CNN, ResNet, Inception, LSTM, and one of the state-of-the-art methods MentorNet in [34]. The optimizer adopted in this paper is Adam. The default value of Adam optimizer (beta1 = 0.9, beta2 = 0.999, epsilon = $1e - 8$, decay = 0) is set. For each method, the performances under different
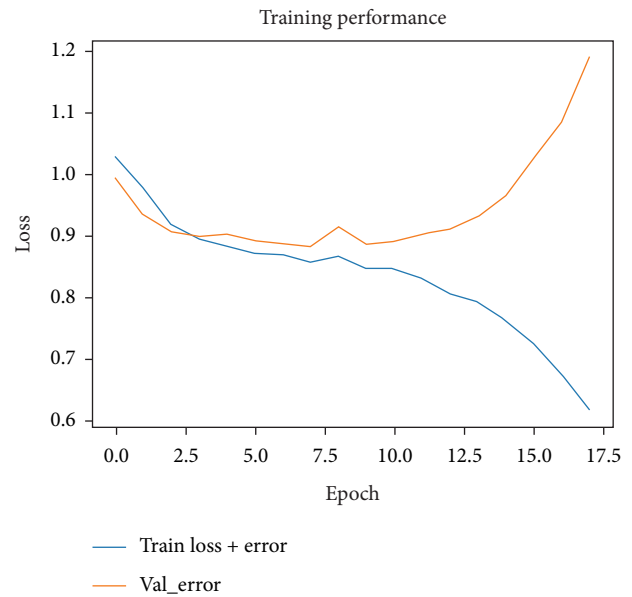

FIGURE 8: The results of training and validation loss in our experiment.

parameters are compared in order to get higher classification accuracy. Here, the influence of different learning rates and batch sizes is considered. The initial learning rate (lr) is initialized to 0.001 and 0.01, and the batch size is initialized to 128, 512, and 1024. It is detected that too high or too low learning rates and batch sizes could yield homogeneous prediction and low accuracy. Figure 9 illustrates the overall classification accuracy of the six methods from −20 dB to 18 dB with different parameters.

As shown in Figure 9, overall classification accuracy of the proposed method and five other methods from −20 dB to 18 dB with different parameters is compared. In this experiment, the highest classification accuracy in Figure 9 for each method is adopted as baseline. The baseline result is gotten from CNN, pure ResNet, pure Inception, LSTM, and MentorNet network on the RML2016b dataset. The accuracy of the baseline and the proposed method on the validation is shown in Table 1.

6.4.2. Overall Classification Accuracy. Figure 10 shows the overall classification accuracy of CNN, LSTM, Inception, ResNet, MentorNet, and the proposed method on the validation dataset from −20 dB to 18 dB. It can be seen that the
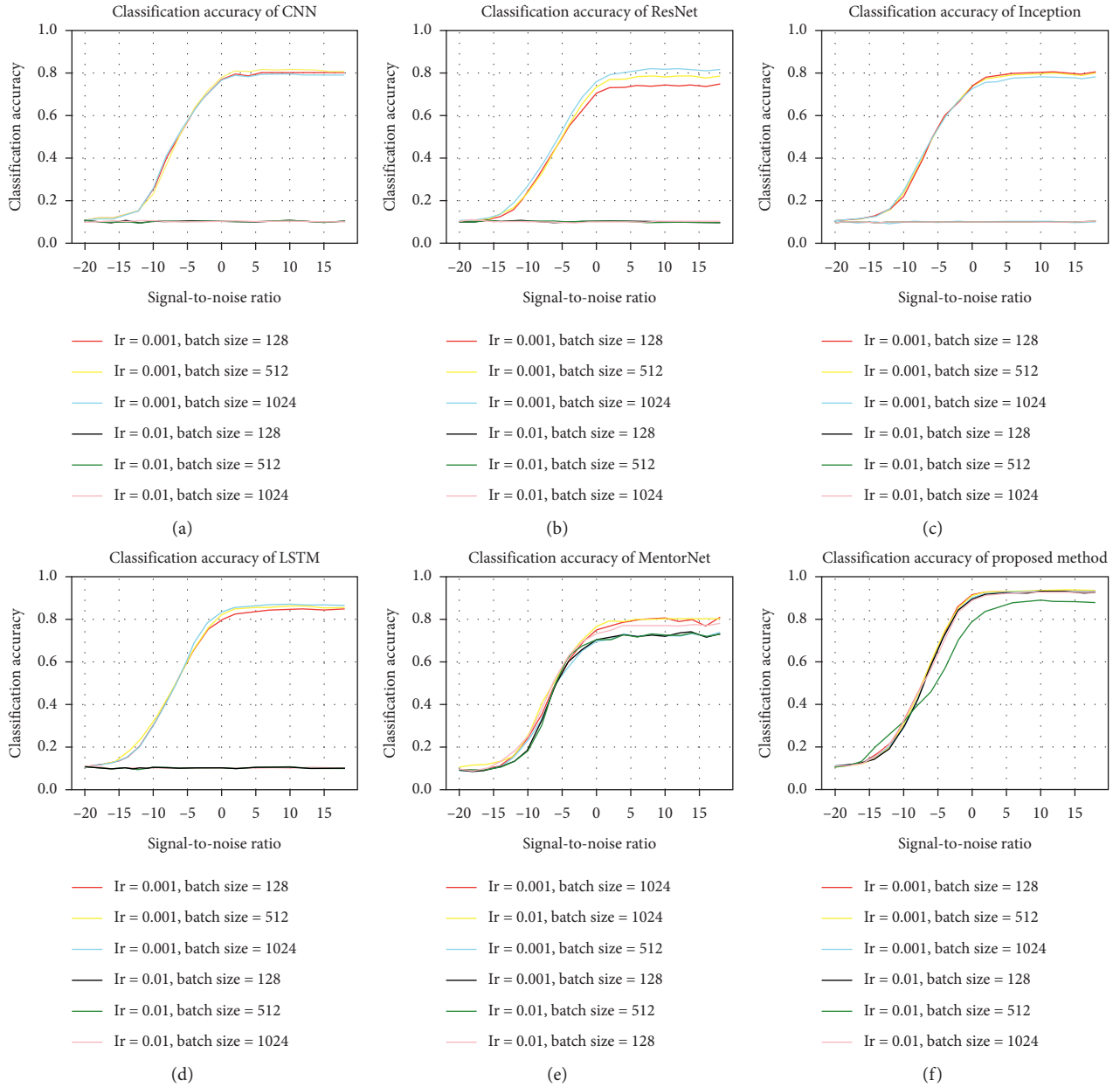
FIGURE 9: The overall classification accuracy of the six methods from −20 dB to 18 dB with different parameters. (a) The overall classification accuracy of CNN with different parameters. (b) The overall classification accuracy of ResNet with different parameters. (c) The overall classification accuracy of Inception with different parameters. (d) The overall classification accuracy of LSTM with different parameters. (e) The overall classification accuracy of MentorNet with different parameters. (f) The overall classification accuracy of the proposed method with different parameters.

TABLE 1: Test accuracy of baseline and the proposed method.

SNR (dB)	CNN (%)	ResNet (%)	Inception (%)	LSTM (%)	MentorNet (%)	Proposed method (%)
−10	25.71	26.92	23.61	29.56	24.86	**30.40**
−4	54.32	48.53	50.86	58.46	62.45	**74.30**
0	77.64	75.41	73.65	83.19	76.36	**90.81**
4	80.53	80.25	78.87	85.96	78.59	**92.84**
10	81.61	81.69	79.91	87.09	80.16	**93.29**

classification accuracy of the six methods increases gradually and remains stable with the increase of SNR. The classification accuracy of the proposed method is highest with the

varying SNR among the six methods. The classification accuracy of the proposed method reaches 90% at 0 dB, while the accuracy of the other methods is less than 90%. The
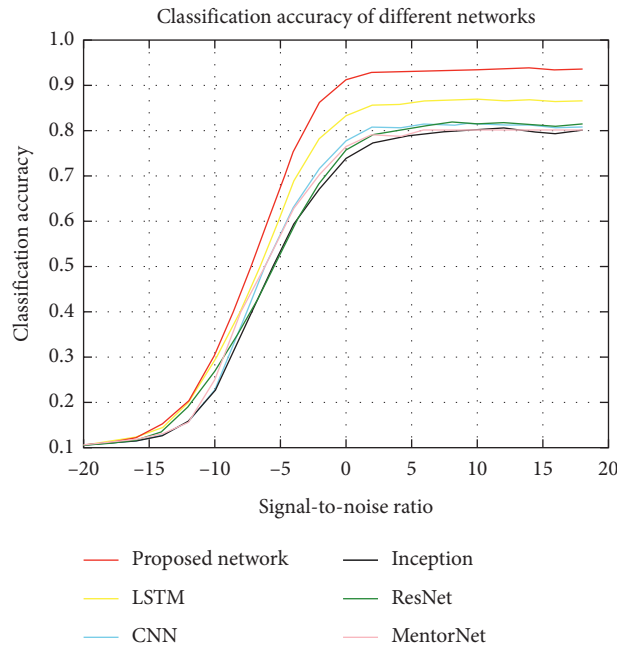
FIGURE 10: The overall classification accuracy of the six methods.

classification accuracy of the proposed method peaks at 93.76% when the SNR is 14 dB and it is 6 percent higher than that of LSTM and 13 percent higher than that of MentorNet, Inception, and ResNet purely. The performance of the proposed method is verified by this experiment.

6.4.3. Classification Accuracy of Each Kind of Modulated Signal.

Figure 11 shows the normalized classification of each kind of modulated signal including 8PSK, AM-DSB, BPSK, CPFSK, GFSK, PAM4, QAM16, QAM64, QPSK, and WBFM by the six methods when the SNR is 18 dB. In Figure 11, the vertical column represents the true label of the modulated signal and the horizontal row represents the predicted label gotten from the deep neural network. All data is normalized. In this connection, the data on the diagonal is the classification accuracy. For example, in Figure 11(a), there are two non-zero values in the top row, that is, 0.95 and 0.05, which indicates that the classification accuracy is 95% for 8PSK modulated signal by the proposed method and there are 5% 8PSK signals that are recognized as QPSK signals. In this connection, it can be found that the classification accuracy of 8PSK, BPSK, CPFSK, GFSK, PAM4, QAM16, QAM64, QPSK, and WBFM gotten from the proposed method is highest among the six methods when the SNR is 18 dB.

From Figure 11, we can get the classification accuracy of each modulated signal gotten from the six methods when the SNR is 18 dB. In this context, we can get the classification accuracy of each modulated signal using the proposed method, LSTM, CNN, Inception, and ResNet from −20 dB to 18 dB, which is shown in Figure 12.

Figure 12 indicates the classification accuracy of each modulated signal from −20 dB to 18 dB using the proposed method, LSTM, CNN, Inception, ResNet, and MentorNet

networks. It can be seen that the classification accuracy of each modulation type using the six methods shows a general upward trend gradually and fluctuates within a narrow range with the increase of SNR. The classification accuracy of 8PSK, BPSK, CPFSK, GFSK, PAM4, QAM16, QAM64, and QPSK is more than 95% using the proposed method, which has the highest number of modulation types whose classification accuracy is more than 95% among the six methods. We can also see that the classification accuracy of GFSK and QAM64 is close to 100% using the proposed method, while that of the other four methods is less than 80%. In addition, the classification accuracy of WBFM is close to 100% using the proposed method, while that of the other four methods is less than 40%. We can also see that the classification accuracy of AM-DSB increases and then declines from −20 dB to 18 dB using the proposed method, which shows the proposed method is not well used to recognize AM-DSB modulation type. Overall, the proposed method has the best performance in classifying the modulated signals for the RML2016b dataset among the six methods adopted in this paper.

6.4.4. Computation Complexity.

Table 2 shows the computation complexity of different methods. Total parameters indicate the number of parameters in different models. Training and inference time refer to the cost of time in training and validation process. FLOPs (floating point of operations) can also represent the complexity of deep neural network. From the table, we can see that the number of total parameters in the proposed network is highest among the six methods. The training time, inference time, and FLOPs are also higher than other methods. In this connection, it will take more time to use the proposed method to get the trained model.
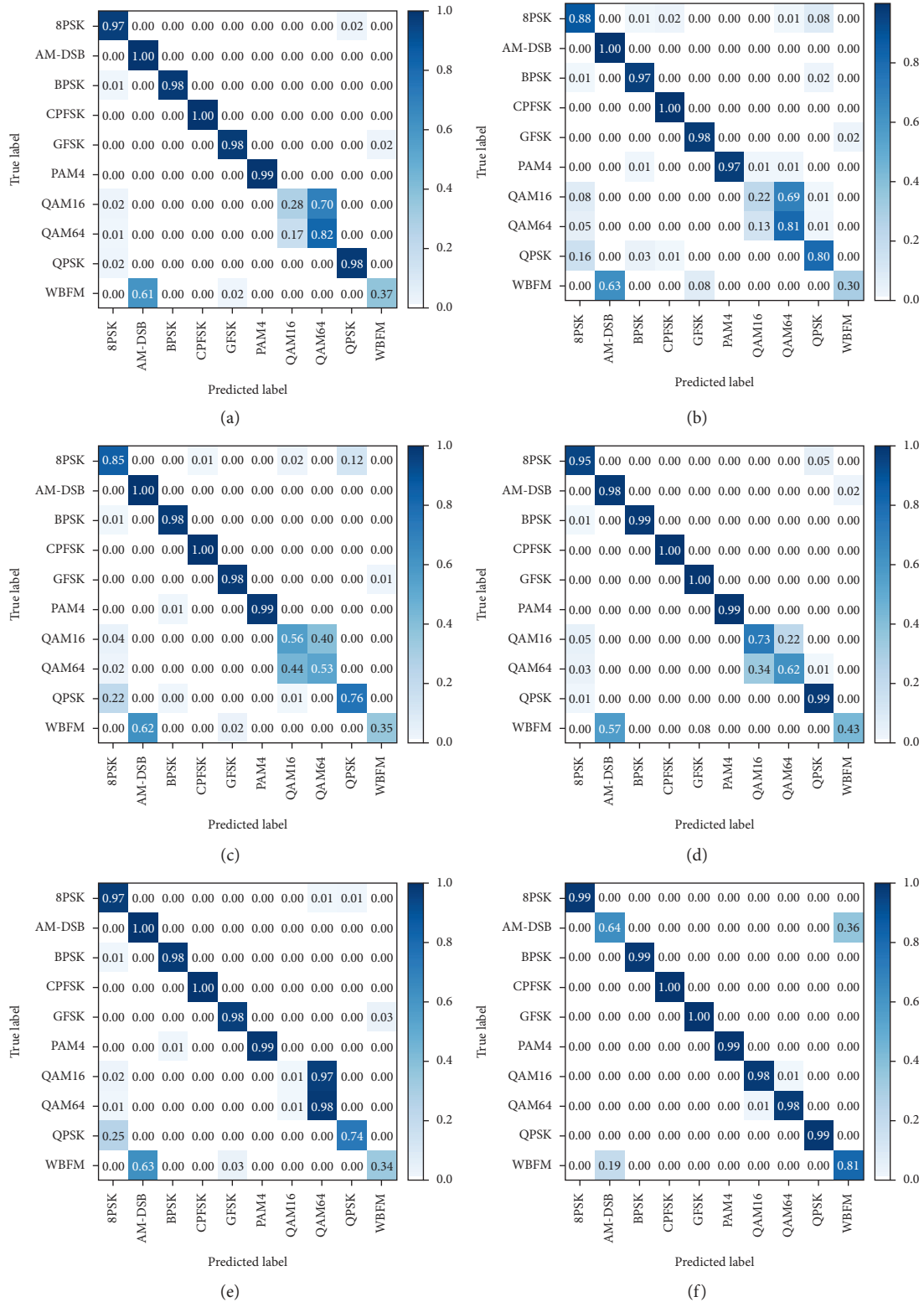
Figure 11: The confusion matrix of the six methods when the SNR is 18 dB. (a) Normalized confusion matrix of CNN (SNR = 18 dB). (b) Normalized confusion matrix of Inception (SNR = 18 dB). (c) Normalized confusion matrix of LSTM (SNR = 18 dB). (d) Normalized confusion matrix of ResNet (SNR = 18 dB). (e) Normalized confusion matrix of MentorNet (SNR = 18 dB). (f) Normalized confusion matrix of the proposed method (SNR = 18 dB).
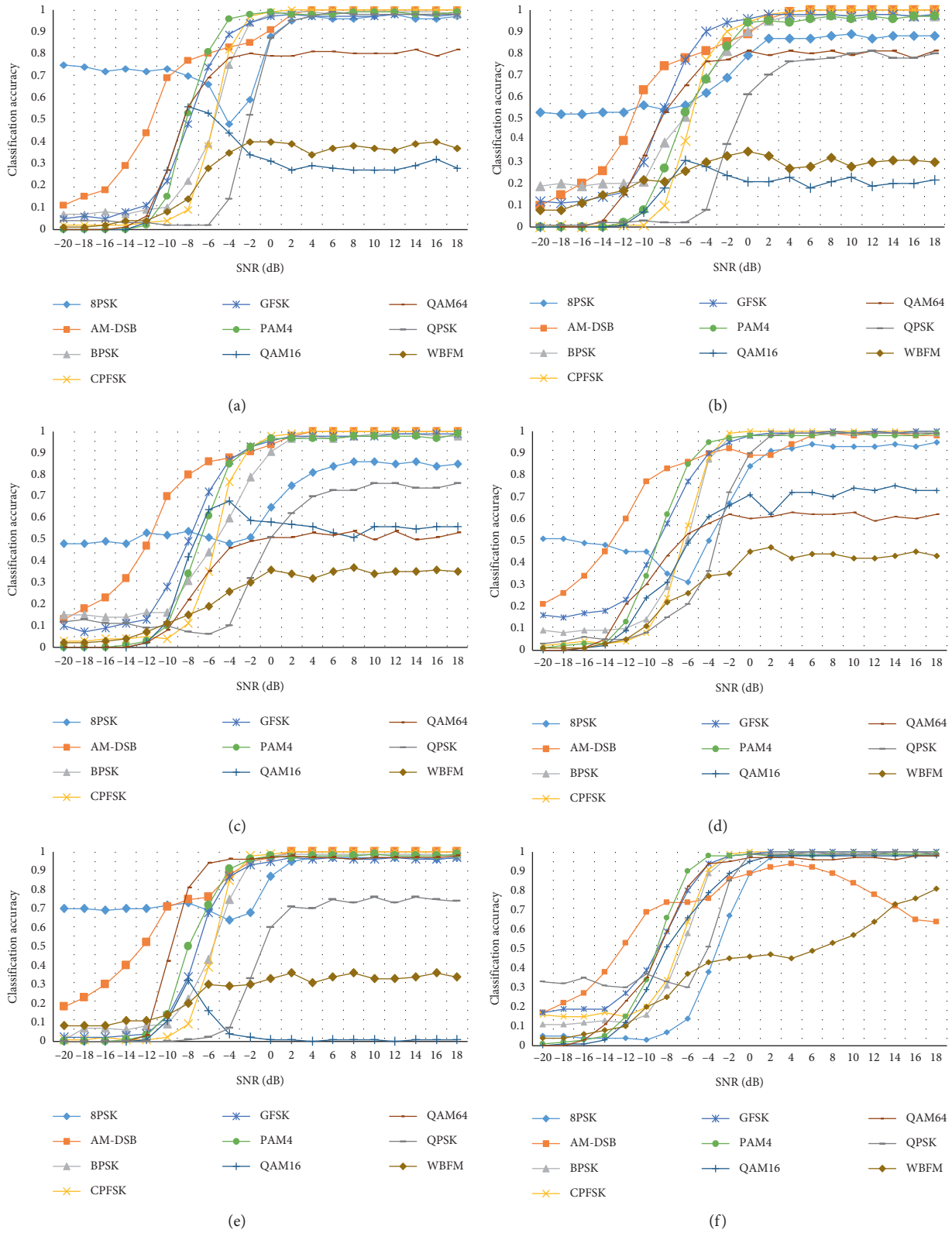
FIGURE 12: The classification accuracy of each modulated signal using the six methods from −20 dB to 18 dB. (a) The classification accuracy of each modulated signal using CNN. (b) The classification accuracy of each modulated signal using Inception. (c) The classification accuracy of each modulated signal using LSTM. (d) The classification accuracy of each modulated signal using ResNet. (e) The classification accuracy of each modulated signal using MentorNet. (f) The classification accuracy of each modulated signal using the proposed method.

TABLE 2: Computation complexity comparison between the six methods.

	CNN	ResNet	Inception	LSTM	MentorNet	Proposed method
Total parameters	2830170	3424976	10142726	846376	3327400	50278762
Training time(s)/epoch	44.40	48.90	38.16	91.75	45.54	590.18
Inference time(s)	2.57	2.44	2.19	5.73	3.07	46.94
FLOPs	8490523	10274951	30428197	2539146	1664319	150897325

7. Conclusion

In this paper, we introduced an improved deep neural architecture for implementing radio signal identification tasks, which is an important facet of constructing the spectrum-sensing capability required by software-defined radio. The goal was to achieve feature extraction by learning from original sampled signals on training dataset and evaluate the performance on validation dataset. We also compared the classification accuracy of the proposed method with that of LSTM, CNN, Inception, ResNet, and one of the state-of-the-art methods MentorNet. The experiment shows that the classification accuracy of the proposed method is highest with the varying SNR among the six methods and it peaks at 93.76% when the SNR is 14 dB and is 6 percent higher than that of LSTM and 13 percent higher than that of MentorNet, Inception, and ResNet purely. The proposed method has excellent performance for modulation classification. We believe that the significance of this paper far exceeds the modulation recognition task itself and it will offer a new idea for modulation classification.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

S. S. contributed to data curation; P.W. and X.W. contributed to methodology; J.Z. contributed to resources; B.S. contributed to software; J.W. contributed to supervision.

Acknowledgments

References

[1] G. Q. Yang, "Modulation classification based on extensible neural networks," *Mathematical Problems in Engineering*, vol. 9, p. 6416019, 2017.

[2] F. Wang, S. Huang, H. Wang, and C. Yang, "Automatic modulation classification exploiting hybrid machine learning network," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–14, 2018.

[3] Y. Wang, H. Zhang, Z. Sang, L. Xu, C. Cao, and T. A. Gulliver, "Modulation classification of underwater communication with deep learning network," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1–12, 2019.

[4] B. Kim, J. Kim, H. Chae, D. Yoon, and J. W. Choi, "Deep neural network-based automatic modulation classification technique," in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, South Korea, October 2016.

[5] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radio: state-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[6] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137–156, 2007.

[7] D. Zhang, W. Ding, B. Zhang et al., "Automatic modulation classification based on deep learning for unmanned aerial vehicles," *Sensors*, vol. 18, no. 3, p. 924, 2018.

[8] Y. Kumar, M. Sheoran, G. Jajoo, and S. K. Yadav, "Automatic modulation classification based on constellation density using deep learning," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1275–1278, 2020.

[9] J. Zhang, D. Cabric, F. Wang, and Z. Zhong, "Cooperative modulation classification for multipath fading channels via expectation-maximization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6698–6711, 2017.

[10] E. Nachmani, Y. Bachar, E. Marciano, D. Burshtein, and Y. Be'ery, "Near maximum likelihood decoding with deep learning," 2018, https://arxiv.org/abs/1801.02726.

[11] A. Afan and Y. Y. Fan, "Unsupervised feature learning and automatic modulation classification using deep learning model," *Physical Communication*, vol. 25, pp. 75–84, 2017.

[12] T. J. O'Shea, N. West, and M. Vondal, "Semi-supervised radio signal identification," in *Proceedings of the IEEE International Conference on Advanced Communication Technology*, Bongpyeong, South Korea, February 2017.

[13] M. Abu-Romoh, A. Aboutaleb, and Z. Rezki, "Automatic modulation classification using moments and likelihood maximization," *IEEE Communications Letters*, vol. 22, no. 5, pp. 938–941, 2018.

[14] X. Sun, S. Su, Z. Zuo, X. Guo, and X. Tan, "Modulation classification using compressed sensing and decision tree-support vector machine in cognitive radio system," *Sensors*, vol. 20, no. 5, p. 1438, 2020.

[15] Z. Zheng, T. Huang, H. Zhang, S. Sun, J. Wen, and P. Wang, "Towards a resource migration method in cloud computing based on node failure rule," *Journal of Intelligent & Fuzzy Systems*, vol. 31, no. 5, pp. 2611–2618, 2016.

[16] X. Shi, Z. Zheng, Y. Zhou et al., "Graph processing on GPUs: a survey," *ACM Computing Surveys*, vol. 50, 2017.

[17] T. O'Shea, T. James, T. Roy, and T. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, pp. 168–179, 2018.

[18] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: opportunities and challenges," *China Communications*, vol. 14, no. 11, pp. 92–111, 2017.

[19] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proceedings of the International Conference on Engineering Applications of Neural Networks*, Aberdeen, UK, September 2016.

[20] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433–445, 2018.

[21] S. Hu, Y. Pei, and P. P. Liang, "Robust modulation classification under uncertain noise condition using recurrent neural network," in *Proceedings of the IEEE Global Communications Conference*, Abu Dhabi, United Arab Emirates, December 2018.

[22] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15713–15722, 2018.

[23] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4074–4077, 2019.

[24] T. N. Sainath, O. Vinyals, A. W. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in *Proceedings of the IEEE International Conference on Acoustics*, Brisbane, Australia, April 2015.

[25] P. Shengliang, J. Hanyu, W. Huaxia, A. Hathal, Z. Yu, and S. M. Mazrouei, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, pp. 718–727, 2019.

[26] P. S. Cheong, M. Camelo, and S. Latré, "Evaluating deep neural networks to classify modulated and coded radio signals," in *Proceedings of the 13th EAI International Conference, CROWNCOM*, Ghent, Belgium, April 2018.

[27] X. Xie, Y. Nil, S. Peng, and Y. D. Yao, "Deep learning based automatic modulation classification for varying snr environment," in *Proceedings of the 2019 28th Wireless and Optical Communication Conference (WOCC)*, Beijing, China, May 2019.

[28] S. Peng, H. Jiang, H. Wang, H. Alwageed, and Y. D. Yao, "Modulation classification using convolutional neural network based deep learning model," in *Proceedings of the 2017 26th Wireless and Optical Communication Conference (WOCC)*, IEEE, Newark, NJ, USA, April 2017.

[29] T. J. O'Shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proceedings of the GNU Radio Conference*, Boulder, Colorado, September 2016.

[30] M. Li, O. Li, G. Liu, and C. Zhang, "Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks," *Sensors*, vol. 18, no. 11, p. 3913, 2018.

[31] M. Du, Q. Yu, and S. Fei, "Fully dense neural network for the automatic modulation recognition," 2019, https://arxiv.org/abs/1912.03449.

[32] Y. Zeng, M. Zhang, F. Han, Y. Gong, and J. Zhang, "Spectrum analysis and convolutional neural network for automatic modulation recognition," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 929–932, 2019.

[33] H. Wu, Y. Li, L. Zhou, and J. Meng, "Convolutional neural network and multi-feature fusion for automatic modulation classification," *Electronics Letters*, vol. 55, no. 16, pp. 895–897, 2019.

[34] M. Zhang, Z. Yu, H. Wang, H. Qin, W. Zhao, and Y. Liu, "Automatic digital modulation classification based on curriculum learning," *Applied Sciences*, vol. 9, no. 10, p. 2171, 2019.

[35] S. Ramjee, S. Ju, and D. Yang, "Fast deep learning for automatic modulation classification," 2019, https://arxiv.org/abs/1901.05850.

[36] N. E. West and T. J. O'Shea, "Deep architectures for modulation recognition," in *Proceedings of the IEEE International Symposium on Dynamic Spectrum Access Networks*, Piscataway, NJ, USA, March 2017.

[37] I. Laina, C. Rupprecht, and V. Belagiannis, "Deeper depth prediction with fully convolutional residual networks," 2016, https://arxiv.org/abs/1606.00373.

[38] A. A. Litonjua and D. R. Gold, "Asthma and obesity: common early-life influences in the inception of disease," *Journal of Allergy and Clinical Immunology*, vol. 121, no. 5, pp. 1075–1084, 2008.

[39] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," 2016, https://arxiv.org/abs/1602.07261.