


## Article

# Modulation Recognition of Communication Signal Based on Convolutional Neural Network

Kaiyuan Jiang, Xvan Qin, Jiawei Zhang and Aili Wang \* 

Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application, Harbin University of Science and Technology, Harbin 150080, China; jiangkaiyuan@hrbust.edu.cn (K.J.); 1920610086@stu.hrbust.edu.cn (X.Q.); 1820600029@stu.hrbust.edu.cn (J.Z.)

\* Correspondence: aili925@hrbust.edu.cn

**Abstract:** In the noncooperation communication scenario, digital signal modulation recognition will help people to identify the communication targets and have better management over them. To solve problems such as high complexity, low accuracy and cumbersome manual extraction of features by traditional machine learning algorithms, a kind of communication signal modulation recognition model based on convolution neural network (CNN) is proposed. In this paper, a convolution neural network combines bidirectional long short-term memory (BiLSTM) with a symmetrical structure to successively extract the frequency domain features and timing features of signals and then assigns importance weights based on the attention mechanism to complete the recognition task. Seven typical digital modulation schemes including 2ASK, 4ASK, 4FSK, BPSK, QPSK, 8PSK and 64QAM are used in the simulation test, and the results show that, compared with the classical machine learning algorithm, the proposed algorithm has higher recognition accuracy at low SNR, which confirmed that the proposed modulation recognition method is effective in noncooperation communication systems.



**Citation:** Jiang, K.; Qin, X.; Zhang, J.; Wang, A. Modulation Recognition of Communication Signal Based on Convolutional Neural Network. *Symmetry* **2021**, *13*, 2302. <https://doi.org/10.3390/sym13122302>

**Keywords:** modulation recognition; convolutional neural network; bidirectional long- and short-term memory network; attention mechanism

Academic Editors: Adam Glowacz and José Carlos R. Alcantud

Received: 2 October 2021

Accepted: 25 November 2021

Published: 2 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Modulation recognition is a technology to judge the modulation mode of a received signal when the content of modulation information is unknown. It is widely used in radio signal monitoring, electronic countermeasures, intelligent communication and other fields. In a real environment, due to the interference of noncooperative communication and background noise, some features of the received signal will be blurred, which will affect the recognition result. How to obtain higher recognition accuracy of the modulation mode under low SNR is an important research topic. Automatic modulation recognition is the main modulation recognition method in wireless communication.

In 1969, Alarabi published an article entitled “Automatic Classification of Modulation Types by Pattern Recognition Technology”, in which he proposed for the first time that the recognition of signal modulation mode is essentially a pattern recognition problem [1]. At present, the methods of automatic modulation recognition can be divided into three categories: modulation recognition based on hypothesis testing, modulation recognition based on feature extraction and modulation recognition based on deep learning. The research status of these three methods at home and abroad is analyzed as follows.

Based on the recognition method of the modulation mode of the hypothesis test, the maximum likelihood function and the most appropriate decision threshold of the signal are derived according to the signal characteristics by using the theory of probability and hypothesis test, and then the statistical quantity of the signal to be tested is compared with the decision threshold, so as to realize the recognition of the modulation mode of the signal. In 1988, Kim and Polydoros proposed a recognition method based on the average

likelihood ratio test (ALRT) for the first time to identify the modulation pattern of BPSK and QPSK [2]. In 1997, Schreyogg et al. proposed a robust method for asynchronous modulation recognition without a priori knowledge about modulation and transmission parameters and proved the robustness of the classifier modulation types on ASK, BPSK, QPSK, 2FSK, MSK and CW simulation signals [3]. In 2000, Panagiotou et al. proposed a recognition method based on a hybrid likelihood ratio test (HLRT). This method combines the above two methods and absorbs the advantages of the two methods to improve the recognition effect [4]. The recognition method of the modulation mode based on hypothesis testing requires a lot of prior information such as mean value and variance of modulation signal, which is difficult to obtain accurately in noncooperative communication. In addition, the recognition accuracy is low at low SNR because of the great influence of noise. Therefore, the traditional recognition method is gradually being replaced by new methods.

The method of modulation recognition based on feature extraction is to extract the most representative and reflective features in the time domain or frequency domain of signals of different modulation types, so as to accurately identify signals of different modulation types. Nandi proposed a recognition method based on instantaneous statistical features and used the decision tree as a classifier to identify 13 kinds of communication signals. When the SNR is 10 dB, the recognition accuracy reaches more than 90% [5]. Tan proposed a recognition method based on time-domain characteristics. Feature parameters are composed into feature vectors, and the classifier selects random forest, which can realize automatic recognition of six modulation modes of low-order digital signals such as 2ASK and 2FSK [6]. Swami used fourth-order cumulants to identify MPSK and MQAM signals [7]. Wang combined the fourth-order and sixth-order cumulants and used the support vector machine (SVM) as the classifier to realize the modulation identification of MPSK [8]. Vladimir proposed a recognition method based on normalized sixth-order cumulants. This method can identify BPSK, QPSK, 16QAM and 64QAM efficiently [9].

With the popularity of big data, deep learning has become a research hotspot in the field of artificial intelligence. With its powerful ability to automatically extract features, it has been widely used in computer vision and natural language processing. As a result, more and more researchers are applying deep learning to their field of study. In recent years, deep learning has been gradually introduced into the research of digital signal modulation recognition. Automatic modulation recognition based on deep learning does not require prior information. It fully utilizes the autonomic learning mechanism of the neural network to take the original signal or the converted signal data as the input of the neural network to train the label. Then the characteristics of the nonlinear data can be extracted through the nonlinear function in the network. Then the features are sent to the output layer to realize the recognition of the modulation mode of the communication signal.

Tu proposed a modulation recognition method for digital signals based on deep autoencoder networks. According to different modulation signals with different cyclic spectrum characteristics and small wave characteristics, the original characteristics of signals are extracted to complete the identification of the unknown signal modulation mode [10]. O'Shea takes the baseband complex signal as input directly and uses CNN to extract the signal characteristics and identify the modulation methods of 11 signals. When SNR is 0 dB, the recognition accuracy reaches 80% [11]. Peng proposed a modulation recognition method based on deep neural network (DNN), which preprocesses the signal to be detected, generates the constellation map and takes the constellation map as the input of the neural network. Using the trained network, the model can recognize the modulated signal. When SNR is greater than 4 dB, the recognition accuracy reaches over 95% [12]. Hou proposed a modulation recognition method for communication signals based on deep learning, which solves the end-to-end signal recognition problem by designing a deep neural network model, thus simplifying the tedious process of artificial feature extraction [13]. Peng proposed to represent data in the form of a grid topology and combine it with CNN to complete the recognition of modulation signals [14].

Wang proposed a convolution neural network recognition algorithm based on the constellation to identify the modulation modes of different signals [15]. Xie proposed a new modulation identification method. First, high-order cumulant features are extracted. On this basis, features of signals of different modulation types can be further extracted, and features are taken as the input of the DNN to improve the recognition accuracy [16]. Tang proposed a modulation recognition method for communication signals based on generating antagonistic networks. Firstly, the auxiliary classifier is used to generate adversarial networks (ACGANs) to extend the data. The classic model AlexNet is then used as a classifier. Compared with the recognition accuracy of the original data set, the expanded data set has a significant improvement in accuracy [17]. Tu used pruning technology to reduce the convolution parameter in CNN and the number of floating-point operations per second and used this modulation method of CNN to identify digital signals. Compared with the original CNN, this lightweight CNN can reduce the training time to 33–35% on the basis of maintaining the recognition accuracy [18]. Shi proposed a particle swarm optimization algorithm for the problem that DNN falls into local minimum. The number of hidden layer nodes of DNN was optimized to extract the characteristics of the signal to be detected as the network input, which could improve the recognition accuracy under low SNR [19]. Liu [20] proposed a deep complex network that cascades the bidirectional long short-term memory network (DCN-BiLSTM) for automatic modulation recognition, and its recognition rate for the 11 modulation signals can reach 90% when SNR exceeds 4 dB.

In this paper, a novel modulation recognition method of the digital signal based on CNN is proposed, called CNN + BiLSTM + attention (C-BiLSTM-A). The algorithm includes two steps of feature extraction and recognition. In the feature extraction operation, the convolution neural network was first used to extract the frequency-domain features of the signal, and then the feature was taken as input. BiLSTM was used to extract the timing features of the signal. Finally, the feature parameters were weighted and summed through the attention mechanism, and a SoftMax classifier was used to identify the modulation mode of the signal.

## 2. Algorithm Design

The structure of the C-BiLSTM-A network model proposed in this paper is shown in Figure 1. The digital signal data set was used as the input of C-BiLSTM-A. Firstly, a one-dimensional convolution of two layers was selected to extract the signal's coarse features, then the activation function ReLU was used to map the features to the nonlinear space, and then the nonlinear features were mapped to the sample label space through a full connection layer. After that, the signal features were input into BiLSTM, and the two-way gating structure in BiLSTM was used to filter the features and retain part of the feature information to further extract the timing features of digital signals. Due to the different feature information carried by each short subsequence in a long sequence, its importance is also different. Therefore, an attention mechanism was introduced after BiLSTM, and the attention value of different short sequences was obtained by calculating the attention distribution of the input feature information and weighted sum. Finally, it was input to the full connection layer and combined with the classifier to recognize the modulation mode of the digital signal correctly.

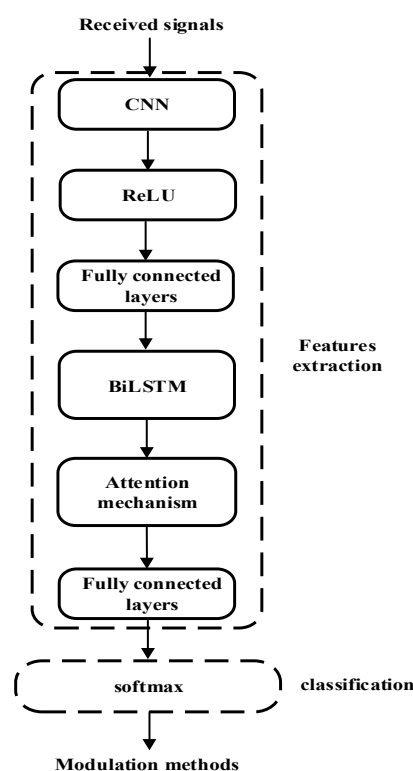


Figure 1. Algorithm block diagram of C-BiLSTM-A.

### 3. Bi-Directional Long Short-Term Memory

At present, the deep learning structure is mainly applied in image and natural processing, and there is less network structure in the application of sequence signal classification. The data processed in this paper belong to one-dimensional time-series signals. On the basis of referring to the convolutional neural network designed by Li, the two-dimensional convolutional layer was adjusted to the one-dimensional convolutional layer [21]. The linear rectifying function ReLU was selected as the activation function, which can effectively alleviate the problem of gradient disappearance in the network and thus make the deep learning model of training more stable.

It was found that digital modulation signals can be regarded as a kind of unit sequence, while the long short-term memory network (LSTM), a powerful sequential signal processing structure, belongs to a special kind of recurrent neural networks (RNNs) [22]. Compared with RNN, LSTM can well express the long-term dependent information in the input. Therefore, LSTM can perform better in longer sequences [23,24]. Whereas RNN has only one transitive state  $h_t$ , LSTM has two transitive states: one cell state,  $c_t$ , and one hidden state,  $h_t$ . The cell state,  $c_t$ , provides time dependence and time characteristics for the input data, and the LSTM realizes long-term control through the cell state. The unit state realizes the long-term control function, mainly through three kinds of gate structures: forgetting gate, inputting gate and outputting gate. The structure of the LSTM is shown in Figure 2.

The positions of the three gate structures in Figure 2 correspond to forgetting gate,  $f$ ; inputting gate,  $i$ ; and outputting gate,  $o$ , respectively, and the candidate value,  $g$ , adds information to the cell state. The element state can filter and update the input timing signals well and give it time dependence, which is beneficial to the classification and prediction of timing signals.

For the structure in Figure 2, the learnable weights of the LSTM layer are input weights,  $w$ ; recurrent weights,  $R$ ; and bias,  $b$ . The matrices are the series of input weights, recursive

weights, and biases for each component, respectively. The connection of these matrices is shown in Equation (1):

$$\begin{cases} f_t = \sigma(w_f x_t + R_f h_{t-1} + b_f) \\ i_t = \sigma(w_i x_t + R_i h_{t-1} + b_i) \\ o_t = \sigma(w_o x_t + R_o h_{t-1} + b_o) \\ g_t = \sigma_c(w_g x_t + R_g h_{t-1} + b_g) \end{cases} \quad (1)$$

where  $x_t$  represents current input,  $h_{t-1}$  represents the output of the previous hidden state,  $\sigma$  is activation function sigmoid,  $\tanh$  represents another activation  $\sigma_c$ .

$$w = \begin{bmatrix} w_f \\ w_i \\ w_g \\ w_o \end{bmatrix}, R = \begin{bmatrix} R_f \\ R_i \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_i \\ b_g \\ b_o \end{bmatrix} \quad (2)$$

The expression of unit state about time,  $t$ , is as follows:

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g_t \quad (3)$$

where  $\otimes$  represents the Hadamard product (vector product of elements). The hidden state expression at  $t$  is expressed as follows:

$$h_t = o_t \otimes \sigma_c(c_t) \quad (4)$$

In general, LSTM layer functions use the hyperbolic tangent function ( $\tanh$ ) as the state activation function.

The specific working principle is as follows: the forgetting gate belongs to the forgetting stage, which is mainly about the selective forgetting of the input passed in by the previous node. Specifically,  $f_t$  controls which parts of the output,  $c_{t-1}$ , for the previous state need to be retained and which parts need to be forgotten. Its expression is as follows:

$$f_t = \sigma_g(\omega_f x_t + R_f h_{t-1} + b_f) \quad (5)$$

where  $f_t$  refers to forgetting gating, is the last output information and also the input data information at the current moment.

The input gate belongs to the input stage, which is to restrict the input information of this stage. Here, it is mainly to process the input,  $x_t$ , at the current moment. Specifically, the current input is controlled through input gating  $i_t$  ( $i$  stands for information). Its expression is as follows:

$$i_t = \sigma_g(\omega_i x_t + R_i h_{t-1} + b_i) \quad (6)$$

The output gate belongs to the output stage. In this stage, it is determined which feature information will be the output of the current state, which is specifically controlled by output gating,  $o_t$ . Its expression is as follows:

$$o_t = \sigma_g(\omega_o x_t + R_o h_{t-1} + b_o) \quad (7)$$

The role of the element state in Figure 2 is to make the information of the past moment run directly along the chain, with only a few linear interactions. Its expression is as follows:

$$g_t = \sigma_c(\omega_g x_t + R_g h_{t-1} + b_g) \quad (8)$$

Similar to the LSTM calculation process, BiLSTM adds the reverse operation on top of it, which can be understood as reversing the input sequence and calculating the output in the way of LSTM again. The final result is the stack of the results of forward LSTM and reverse LSTM [25]. The structure of BiLSTM is shown in Figure 3. The forgetting gate selectively forgets the input passed by the previous node. The input gate is to learn

new information to replace the forgotten information. Here, it is mainly to process the input at the current time. The output gate determines which feature information is output as the current unit state. The bidirectional long-term and short-term memory network adds reverse operation on the basis of LSTM. It can be understood that the input sequence is reversed and calculated again in the way of LSTM. The final result is the stacking of forward LSTM and reverse LSTM, which is a symmetrical structure.

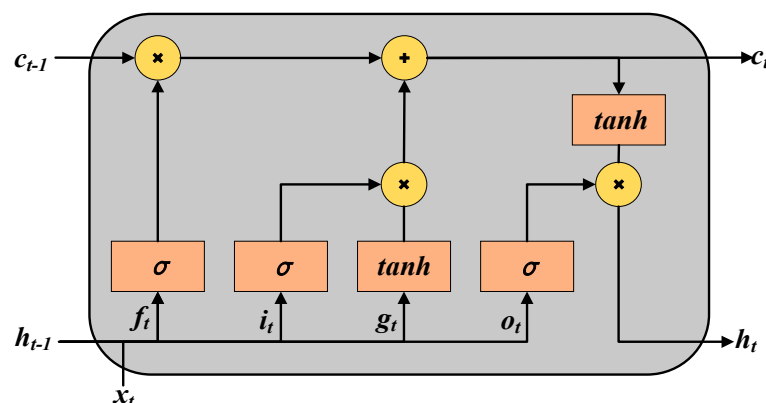


Figure 2. LSTM internal unit structure.

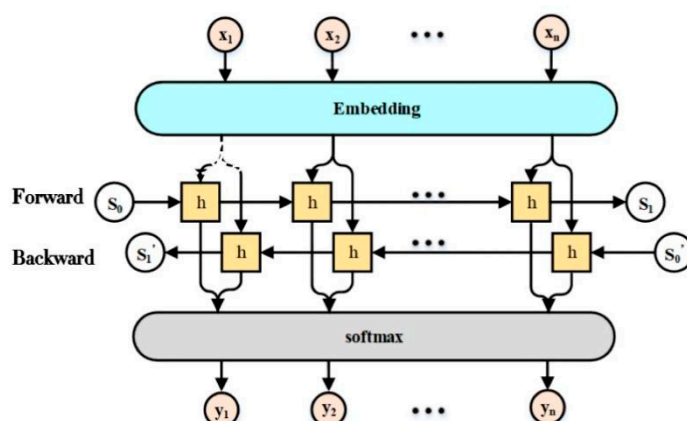


Figure 3. Internal structure of BiLSTM.

#### 4. Attention Mechanism

The mechanism of attention stems from the study of human vision. The attention mechanism can be expressed as a resource allocation model. At some point, the attention will focus on one part of the target, while ignoring the rest. Using mathematical description is a weighted change of the target data. Its core objective is also to select the information that is more critical to the current mission objective from the numerous data [26,27]. Because the importance degree of different short subsequences in the long time series varies, and the important salient features usually contain more information, it needs to analyze the importance degree of the attention mechanism introduced into the target sequence.

The attention mechanism can be divided into hard attention and soft attention mechanisms according to the method of selecting task-related information from the input information. Hard attention refers to the selection of information at a certain position in the input sequence as input, such as the random selection of a piece of information or the information with the highest probability of occurrence. Soft attention refers to the choice of information in the input sequence, it means that all pieces of information are used as input. Since soft attention is more common in practical applications, this paper chooses it as the attention mechanism weighted method. Input information  $[x_0, x_1, \dots, x_n]$  is represented as  $n + 1$  groups, and here as a group of input information  $x_i, i = 0, 1, \dots, n$ . From the



perspective of computing resources, it is not necessary to input all the information into the computing network but rather to select some information that is more relevant to the task. The computational process of the attention mechanism can be expressed as two steps: firstly, the attention distribution is calculated on all input information. Then the input information is weighted according to the distribution of attention. The specific calculation process is divided into the following three steps:

The first step is to introduce different functions and calculation mechanisms. According to *query* and input, the correlation of them is calculated, and the formula is expressed as follows:

$$\text{similarity}(\text{query}, x_i) = \frac{\text{query} \cdot x_i}{\|\text{query}\| \cdot \|x_i\|} \quad (9)$$

The second step is to introduce a similar calculation method, *softmax*, to numerically transform the score in the first step, which can not only transform the original score into the probability distribution with the sum of all element weights equal to 1 through normalization but also highlight the weight of important elements in the sequence through the internal mechanism. The formula is shown as follows:

$$a_i = \text{softmax}(\text{simil}_i) = \frac{e^{\text{simil}_i}}{\sum_{i=0}^n e^{\text{simil}_i}} \quad (10)$$

$a_i$  represents the distribution of attention,  $\text{simil}_i$  refers to  $\text{similarity}(\text{query}, x_i)$ , which represents the attention rating mechanism.

The third step is to calculate the weight coefficients according to the second step, the input,  $x_i$ , can be a weighted sum to get the attention value. The formula is represented as follows:

$$\text{Attention} = \sum_{i=0}^N a_i \cdot x_i \quad (11)$$

The attention value can complete the classification prediction task. In view of the characteristics of the digital modulation signal itself, a new deep learning model is proposed in this paper. The corresponding network structure of the proposed C-BiLSTM-A model is shown in Figure 4.

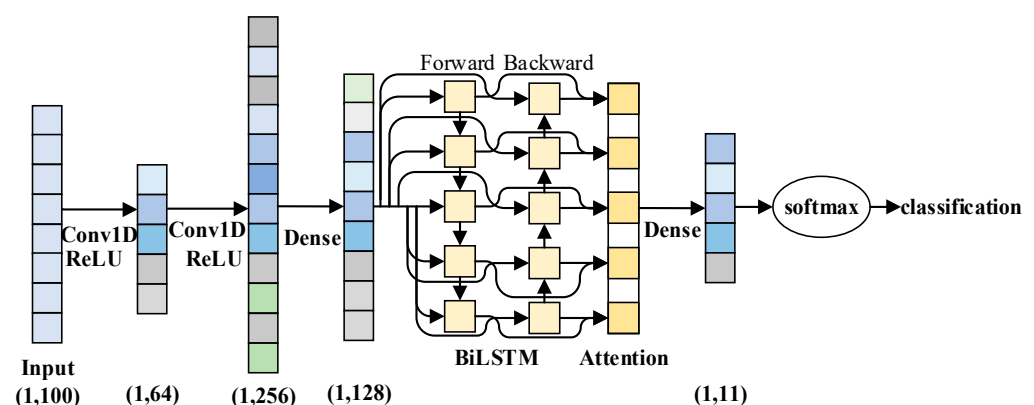


Figure 4. The structure of C-BiLSTM-A.

In this paper, firstly, CNN was used on a training data set to extract spatial features. Then, BiLSTM was introduced to further extract the temporal features of signals and to train and learn the two types of features, which can improve the recognition accuracy. Although BiLSTM is suitable for the classification and prediction of sequences, considering that the signals to be identified are all long sequences, the importance of different short subsequences in long sequences is different, so only using BiLSTM cannot accurately

distinguish the differences of each subsequence. Therefore, the attention mechanism was introduced after BiLSTM to effectively distinguish the importance by calculating the attention distribution of different short subsequences in long sequences. This method can further improve the recognition accuracy of digital signal modulation at low SNR.

## 5. Experimental Results and Analysis

### 5.1. Experimental Environment and Super Parameter Selection

In order to verify the universality and effectiveness of the proposed algorithm, this paper used MATLAB simulation to get the seven signal sets of 2ASK, 4ASK, 4FSK, BPSK, QPSK, 8PSK and 64QAM. Each signal is obtained by pulse shaping of baseband signal, shaping filtering, carrier modulation and Gaussian noise simulation. Meanwhile, the TensorFlow-1.8.0 + Keras-2.2.4 framework was set up to train the CNN + BiLSTM + attention (C-BiLSTM-A) network. The data set required for the experiment was divided into a training set and a test set in a ratio of 9:1 and a k-Fold cross-validation method was used to evaluate machine learning mode.

Since the length of the signal data used for training was not long, the step size of the convolution kernel was set to 2, and the size of the convolution kernel is set to  $1 \times 1$  and  $1 \times 3$ . The spatial features of digital signals were extracted using a convolution neural network. The cross-entropy loss function was selected as the loss function, and the formula is as follows:

$$L = - \sum_{k=1}^7 \log(p(k))q(k) \quad (12)$$

where  $p(k)$  represents the probability distribution of real SNR and  $q(k)$  represents the probability distribution of SNR predicted by training model. The cross-entropy loss function can measure the similarity between two distributions, which is often used in multi classification problems. The optimizer was used to reduce the value of loss function to update hidden layer parameters. Adam was selected as the optimizer, which has a series of advantages, such as less memory requirement, simple implementation and high efficiency. The learning rate was set to a fixed value of 0.001, and dropout was set to 0.1. In this way, the of over fitting problem in the full connection layer due to too many parameters was avoided.

### 5.2. Recognition Performance Comparison

This paper adopts the C-BiLSTM-A algorithm to identify the modulation mode of the digital signal. Meanwhile, CNN, C-BiLSTM, SVM [28], random forest (RF) [29] and KNN [30] were taken as comparison algorithms. The evaluation index required by the experiment was the recognition accuracy of seven typical signals under different SNRs (−20–18 dB). The experimental results are shown in Figure 5 below.

Figure 5 shows the recognition rates of seven signals at different SNRs, respectively. As can be seen from the figure, the recognition rate of each digital signal increases with the increase of the SNR. The recognition rate of the algorithm presented in this paper was obviously higher than that of other comparison algorithms. When the SNR was 0 dB, the recognition accuracy of the proposed algorithm was 3–10% higher than that of other comparison algorithms.

Figure 5b shows the recognition accuracy of signal 4ASK at different signal-to-noise ratios. It can be seen from the figure that the recognition accuracy of the C-BiLSTM-A model increased rapidly from SNR = −14 dB. When SNR = 2 dB, the recognition accuracy gradually reached a stable level and approached 100%, especially under low SNR, the recognition accuracy was better than that in other comparison algorithm models. It can be seen from Figure 5d that the recognition rate of the C-BiLSTM-A model for BPSK increased rapidly from SNR = −6 dB. When SNR = 6 dB, the recognition rate was close to 90% until SNR = 10 dB, which tended to be stable, while other comparison algorithms gradually reached a stable state when SNR = 14 dB.



Tables 1–7 show seven kinds of signal recognition rates obtained by applying different modulation recognition algorithms to the results, including C-BiLSTM-A, C-BiLSTM, CNN, random forest, KNN and SVM. It can be seen from the table that the C-BiLSTM-A algorithm had the highest recognition accuracy for seven kinds of signals.

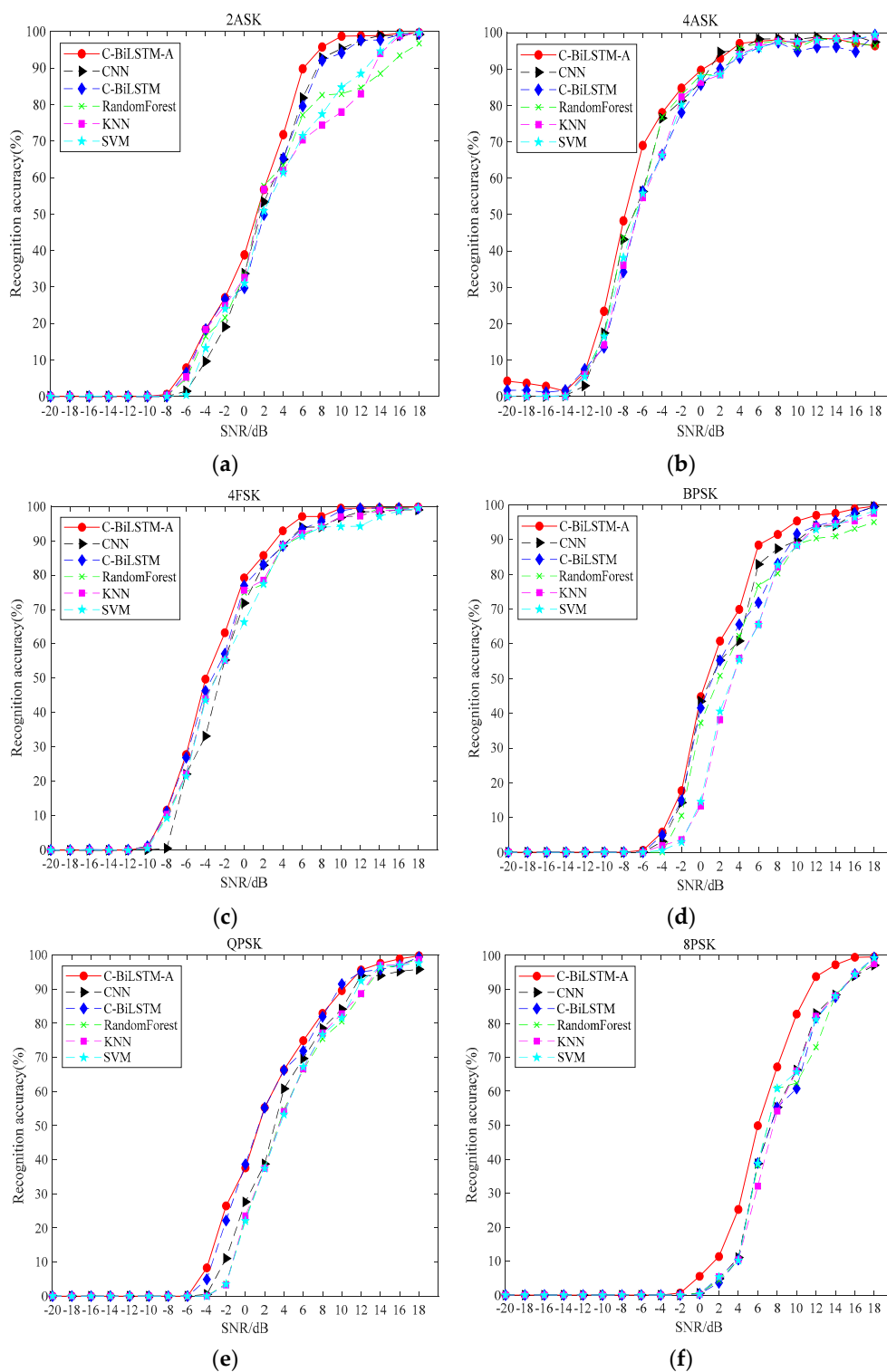
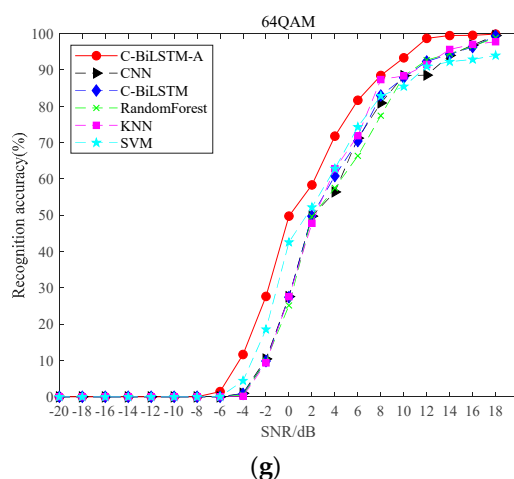


Figure 5. Cont.



**Figure 5.** Signal recognition rate comparison by different methods. (a) 2ASK, (b) 4ASK, (c) 4FSK, (d) BPSK, (e) QPSK, (f) 8PSK, (g) 64QAM.

**Table 1.** Recognition rate comparison of BPSK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0	0	0	0	0	0
−8	0	0	0	0	0	0
−6	0.005528	0	0	0	0	0
−4	0.058043	0.030956	0.049198	0	0.019348	0.005528
−2	0.176893	0.142620	0.149254	0.105030	0.035931	0.029851
0	0.447761	0.434494	0.415699	0.372029	0.133223	0.145384
2	0.608071	0.552792	0.552239	0.508015	0.381426	0.405749
4	0.699281	0.608071	0.655611	0.622443	0.55832	0.553897
6	0.884467	0.829187	0.718629	0.767828	0.655611	0.655058
8	0.914870	0.873964	0.830846	0.802653	0.821448	0.826976
10	0.953566	0.897181	0.916528	0.886678	0.883914	0.883361
12	0.970149	0.938640	0.939746	0.903814	0.939193	0.929795
14	0.975677	0.939746	0.953566	0.909895	0.946379	0.941404
16	0.988944	0.975677	0.976783	0.931454	0.954118	0.964069
18	0.996130	0.995025	0.99392	0.950249	0.975677	0.982311

**Table 2.** Recognition rate comparison of QPSK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0	0	0	0	0	0
−8	0	0	0	0	0	0
−6	0	0	0	0	0	0
−4	0.082919	0.005528	0.049198	0	0	0
−2	0.264787	0.110558	0.221669	0.036484	0.032615	0.034826
0	0.375898	0.276396	0.386401	0.224433	0.235489	0.220564
2	0.552792	0.387507	0.552239	0.378662	0.374793	0.37424
4	0.663350	0.608071	0.662797	0.546711	0.541183	0.532891
6	0.749033	0.696517	0.718629	0.66335	0.666114	0.672195
8	0.829187	0.784964	0.819237	0.754561	0.771697	0.766722
10	0.895522	0.841349	0.91487	0.804865	0.828082	0.81592
12	0.956329	0.939746	0.950802	0.883914	0.886678	0.923715
14	0.975677	0.939746	0.957435	0.959646	0.970702	0.965174
16	0.988944	0.95246	0.970149	0.970702	0.970702	0.970149
18	0.997789	0.957988	0.99392	0.988944	0.986733	0.976230

**Table 3.** Recognition rate comparison of 8PSK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0	0	0	0	0	0
−8	0	0	0	0	0	0
−6	0	0	0	0	0	0
−4	0	0	0	0	0	0
−2	0.006081	0	0	0	0	0
0	0.055279	0.005528	0.003870	0.004975	0.002211	0.003317
2	0.113322	0.048093	0.035932	0.042565	0.054174	0.053068
4	0.252073	0.110558	0.104478	0.103372	0.103925	0.102266
6	0.498618	0.386954	0.388613	0.381426	0.320619	0.387507
8	0.671642	0.552792	0.552239	0.608071	0.541736	0.608071
10	0.826976	0.662797	0.608071	0.624102	0.660586	0.656716
12	0.937535	0.829187	0.816473	0.729685	0.820896	0.811498
14	0.972360	0.883914	0.876728	0.882808	0.880597	0.880044
16	0.994472	0.939746	0.945274	0.941404	0.943615	0.944168
18	0.995578	0.970702	0.993920	0.988391	0.975677	0.992261

**Table 4.** Recognition rate comparison of 2FSK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0	0	0	0	0	0
−8	0.005528	0	0	0.002211	0.003317	0
−6	0.077944	0.014373	0.064677	0.048093	0.053068	0.004422
−4	0.184080	0.096186	0.183527	0.164732	0.183527	0.132670
−2	0.270315	0.190713	0.266446	0.215589	0.248756	0.239912
0	0.388060	0.336650	0.297402	0.331675	0.326147	0.309563
2	0.567164	0.532891	0.498065	0.577114	0.566611	0.508568
4	0.717523	0.649530	0.652294	0.634052	0.618574	0.613599
6	0.897734	0.818132	0.795467	0.771144	0.703151	0.714207
8	0.957435	0.927584	0.920398	0.825871	0.743505	0.773908
10	0.987286	0.953013	0.941404	0.829187	0.779436	0.847430
12	0.987839	0.976230	0.97623	0.845771	0.829187	0.884467
14	0.988944	0.988944	0.976783	0.884467	0.939746	0.945826
16	0.995025	0.988944	0.991155	0.934218	0.988944	0.994472
18	0.996683	0.991708	0.993920	0.967385	0.995578	0.996130

**Table 5.** Recognition rate comparison of 4ASK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0.233831	0.174129	0.134881	0.165837	0.141515	0.164732
−8	0.482587	0.43173	0.342178	0.436705	0.360973	0.381426
−6	0.689884	0.563295	0.563847	0.552792	0.546711	0.55832
−4	0.779989	0.765064	0.663903	0.767828	0.663350	0.662797
−2	0.847430	0.813156	0.780542	0.829187	0.823107	0.801548
0	0.896628	0.862908	0.856827	0.883914	0.864013	0.877833
2	0.929243	0.946379	0.90105	0.884467	0.884467	0.885019
4	0.970149	0.95246	0.930901	0.957435	0.939746	0.939193
6	0.976230	0.982311	0.960199	0.970149	0.963516	0.957988
8	0.980652	0.983416	0.973466	0.977888	0.975677	0.974572
10	0.969044	0.981758	0.949143	0.957988	0.975677	0.974572
12	0.984522	0.987839	0.960199	0.983416	0.978441	0.978441
14	0.982863	0.976783	0.960752	0.980100	0.982863	0.982311
16	0.970702	0.990050	0.947485	0.970702	0.983969	0.982311
18	0.964069	0.974019	0.993920	0.964069	0.989497	0.993919

**Table 6.** Recognition rate comparison of 4FSK.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0.008292	0.001106	0.011609	0.012161	0.010503	0.005528
−8	0.116086	0.005528	0.112769	0.099502	0.104478	0.093975
−6	0.278054	0.221117	0.269210	0.220564	0.221669	0.215036
−4	0.497512	0.331675	0.463239	0.442233	0.441128	0.435600
−2	0.632394	0.552792	0.571034	0.552792	0.552239	0.553344
0	0.791598	0.718629	0.769486	0.749033	0.756219	0.66335
2	0.856827	0.828635	0.830293	0.773908	0.784964	0.774461
4	0.928690	0.884467	0.884467	0.884467	0.886125	0.883914
6	0.970149	0.939193	0.933665	0.914870	0.920951	0.913765
8	0.970149	0.939746	0.956329	0.939746	0.940851	0.939193
10	0.994472	0.967385	0.987286	0.970149	0.969596	0.941404
12	0.994472	0.983969	0.994472	0.970702	0.971808	0.941957
14	0.995578	0.983969	0.994472	0.994472	0.988391	0.970702
16	0.996683	0.989497	0.995025	0.994472	0.988944	0.98618
18	0.997789	0.989497	0.99392	0.995578	0.995025	0.994472

**Table 7.** Recognition rate comparison of 64QAM.

SNR (dB)	C-BiLSTM-A	CNN	C-BiLSTM	RF	KNN	SVM
−10	0	0	0	0	0	0
−8	0	0	0	0	0	0
−6	0.014373	0	0	0	0	0.000553
−4	0.116639	0.009950	0.008292	0.004975	0.001658	0.044223
−2	0.276396	0.104478	0.098397	0.098950	0.093422	0.185738
0	0.497512	0.275843	0.274185	0.251520	0.275843	0.425650
2	0.583195	0.497512	0.497512	0.496407	0.477612	0.521835
4	0.717523	0.563847	0.608071	0.574903	0.627971	0.627971
6	0.816473	0.711996	0.704256	0.66335	0.718629	0.742952
8	0.884467	0.808734	0.829187	0.773908	0.873411	0.828082
10	0.933112	0.883914	0.878386	0.883361	0.883361	0.854063
12	0.986733	0.885572	0.923162	0.927032	0.914870	0.909342
14	0.994472	0.939746	0.943062	0.944721	0.956329	0.922609
16	0.995025	0.968491	0.962963	0.965174	0.970149	0.928690
18	0.998342	0.994472	0.993920	0.988391	0.977336	0.939746

Table 1 shows the recognition accuracy of BPSK. It can be seen from Table 1 that the recognition accuracy of the C-BiLSTM-A model gradually increased from SNR = −6 dB, reaching a recognition accuracy of 0.914870 when SNR = 8 dB, and CNN and C-BiLSTM only reached a recognition accuracy of 0.873964 and 0.830846 when SNR = 8 dB. The recognition accuracy of C-BiLSTM-A was more than 4% higher than that of CNN and C-BiLSTM. Compared with the three traditional machine learning algorithms, SVM had the highest recognition rate of 0.826976 when SNR = 8 dB. The results show that the recognition effect of the C-BiLSTM-A model was better than that of other comparison algorithm models.

Table 2 shows the recognition accuracy of QPSK. From Table 2, it can be seen that the recognition accuracy of the C-BiLSTM-A model increased rapidly from SNR = −4 dB, which was significantly better than that of other models. When SNR = 8 dB, it reached a recognition accuracy of 0.829187, while CNN and C-BiLSTM only reached a recognition accuracy of 0.784964 and 0.819237 when SNR = 8 dB. C-BiLSTM-A had a higher recognition rate than CNN and C-BiLSTM. It shows that the recognition effect of the C-BiLSTM-A model on QPSK was better than that of other comparison algorithm models.

Table 7 shows the recognition accuracy of 64QAM. It can be seen from Table 7 that the recognition accuracy of C-BiLSTM-A model increased rapidly from SNR = −4 dB, which was significantly better than that of other models. When SNR = 10 dB, it reached a recognition accuracy of 0.933112 and tended to be stable. When SNR = 10 dB, CNN and C-BiLSTM

only achieved a recognition accuracy of 0.883914 and 0.878386. The recognition accuracy of C-BiLSTM-A was 5% and 6% higher than that of CNN and C-BiLSTM. Compared with the three traditional machine learning algorithms, the highest recognition rate of the SVM model was 0.883361 when SNR = 10 dB. The results show that the recognition effect of the C-BiLSTM-A model on 64QAM was better than that of other comparison algorithm models.

It can be seen from the table that the recognition accuracy of the method proposed in this paper was higher than that of other comparison algorithms at low SNR.

The confusion matrix of various algorithm models obtained by experiments at SNR = 0 dB is shown in Figure 6. The rightmost column of the chart shows the percentage of all samples belonging to each category correctly classified and incorrectly classified, which is called precision. The row at the bottom of the chart shows the percentage of all samples for each category correctly classified and incorrectly classified, which is called recall rate. The lower right cell of the chart shows the overall accuracy. As can be seen from Figure 6, the confusion matrix can visually show the misclassification of samples by different algorithms. It can be seen from the confusion matrix that the error rate of the proposed algorithm was smaller than that of other comparison algorithms. The C-BiLSTM-A model easily confused the [QPSK, 8PSK] modulation types, which means that the recognition effect of this model on this group of modulation types was general. In terms of accuracy, the recognition accuracy of 4ASK, 4FSK, BPSK and 64QAM is high and reaches more than 90%, and the recognition accuracy on QPSK, 2ASK and 8PSK was relatively general and also reaches more than 70%. In terms of recall rate, the recall rate of 4ASK has reached more than 90%, the recall rate of 2ASK has also reached 70%, and the effect on BPSK is relatively general. However, in general, the deep learning algorithm designed in this chapter had a very good overall recognition effect in the recognition of digital signal modulation.

C-BiLSTM-A:0dB												
Output Class	2ASK	4ASK	2FSK	4FSK	BPSK	QPSK	8PSK	MSK	16QAM	32QAM	64QAM	
	702 3.5%	23 0.1%	87 0.4%	86 0.4%	150 0.8%	243 1.2%	227 1.1%	215 1.1%	85 0.4%	67 0.3%	84 0.4%	35.7%
	7 0.0%	1622 8.2%	137 0.7%	94 0.5%	0 0.0%	0 0.0%	0 0.0%	13 0.1%	60 0.3%	59 0.3%	25 0.1%	80.4%
	34 0.2%	125 0.6%	932 4.7%	743 3.7%	0 0.0%	2 0.0%	4 0.0%	31 0.2%	6 0.0%	2 0.0%	14 0.1%	49.2%
	13 0.1%	10 0.1%	560 2.8%	810 4.1%	4 0.0%	1 0.0%	3 0.0%	63 0.3%	19 0.1%	6 0.0%	23 0.1%	53.6%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	810 4.1%	1 0.0%	2 0.0%	0 0.0%	0 0.0%	0 0.0%	19 0.1%	97.4%
	339 1.7%	1 0.0%	1 0.0%	2 0.0%	336 1.7%	680 3.4%	668 3.4%	277 1.4%	315 1.6%	257 1.3%	196 1.0%	22.1%
	2 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.0%	5 0.0%	100 0.5%	0 0.0%	7 0.0%	6 0.0%	31 0.2%	65.4%
	695 3.5%	15 0.1%	87 0.4%	73 0.4%	476 2.4%	829 4.2%	775 3.9%	1188 6.0%	394 2.0%	319 1.6%	265 1.3%	23.2%
	5 0.0%	13 0.1%	0 0.0%	0 0.0%	6 0.0%	14 0.1%	8 0.0%	1 0.0%	762 3.8%	98 0.5%	37 0.2%	80.7%
	5 0.0%	0 0.0%	2 0.0%	0 0.0%	17 0.1%	24 0.1%	11 0.1%	6 0.0%	107 0.5%	846 4.3%	215 1.1%	68.6%
	7 0.0%	0 0.0%	3 0.0%	1 0.0%	8 0.0%	10 0.1%	11 0.1%	15 0.1%	54 0.3%	149 0.7%	900 4.5%	77.7%
	38.8% 61.2%	89.7% 10.3%	51.5% 48.5%	44.8% 55.2%	44.8% 55.2%	37.6% 62.4%	5.5% 94.5%	65.7% 34.3%	42.1% 57.9%	46.8% 53.2%	49.8% 50.2%	47.0% 53.0%
Target Class												

Figure 6. Confusion matrix of C-BiLSTM-A.

## 6. Conclusions

This paper proposed the C-BiLSTM-A algorithm, which improves the recognition accuracy of digital signal modulation by combining CNN, BiLSTM and the attention mechanism. A lot of experimental results show that the improved C-BiLSTM-A model proposed in this chapter has achieved good recognition results under the modulation types of 2ASK, 4ASK, 4FSK, BPSK, QPSK, 8PSK and 64QAM, and the recognition accuracy is about 5% higher than that of the comparison algorithm under low signal-to-noise ratio. It shows that adding BiLSTM and attention mechanism to CNN is very helpful to improve the recognition effect of the model. At the same time, it also proves that the deep learning model performs better in universality in the field of digital signal modulation recognition than the traditional machine learning methods. There are still many aspects to be improved in this experiment, for example, the network structure and super parameters still have some room for improvement, and the follow-up work will continue to try and improve them.

**Author Contributions:** Conceptualization, writing—review and editing, A.W. and K.J.; methodology, software and validation, X.Q. and J.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Natural Science Foundation of China under grant no. NSFC-61671190.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Haibin Wu for his valuable comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Alarabi, A.; Alkishiwi, O.A.S. Modulation Classification Based on Statistical Features and Artificial Neural Network. In Proceedings of the 2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA, Tripoli, Libya, 25–27 May 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 748–751.
2. Kim, K.; Polydoros, A. Digital Modulation Classification: The BPSK versus QPSK case. In Proceedings of the Military Communications Conference, MILCOM 88, San Diego, CA, USA, 23–26 October 1988; IEEE: Piscataway, NJ, USA, 2003; Volume 2, pp. 431–436.
3. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [\[CrossRef\]](#)
4. Panagiotou, P.; Anastasopoulos, A.; Polydoros, A. Likelihood ratio tests for modulation classification. In Proceedings of the Likelihood Ratio Tests for Modulation Classification. 21st Century Military Communications, Architectures and Technologies for Information Superiority, Los Angeles, CA, USA, 22–25 October 2000; IEEE: Piscataway, NJ, USA, 2002; Volume 2, pp. 670–674.
5. Nandi, A.; Azzouz, E. Algorithms for automatic modulation recognition of communication signals. *IEEE Trans. Commun.* **1998**, *46*, 431–436. [\[CrossRef\]](#)
6. Tan, Z.; Shi, J.; Hu, J. Low-Order Digital Modulation Recognition Algorithm based on Random Forest. *Commun. Technol.* **2018**, *51*, 527–532.
7. Swami, A.; Sadler, B.M. Hierarchical digital modulation classification using cumulants. *IEEE Trans. Commun.* **2000**, *48*, 416–429. [\[CrossRef\]](#)
8. Wang, L.-X.; Ren, Y.-J.; Zhang, R.-H. Algorithm of digital modulation recognition based on support vector machines. In Proceedings of the 2009 International Conference on Machine Learning and Cybernetics, Baoding, China, 12–15 July 2009; IEEE: Piscataway, NJ, USA, 2009; Volume 2, pp. 980–983.
9. Orlic, V.; Dukic, M.L. Automatic modulation classification algorithm using higher-order cumulants under real-world channel conditions. *IEEE Commun. Lett.* **2009**, *13*, 917–919. [\[CrossRef\]](#)
10. Ya, T.; Lin, Y.; Wang, H. Modulation Recognition of Digital Signal Based on Deep Auto-Encoder Network. In Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), Prague, Czech Republic, 25–29 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 256–260.
11. O’Shea, T.J.; Corgan, J.; Clancy, T.C. Convolutional Radio Modulation Recognition Networks. In Proceedings of the Communications in Computer and Information Science, Aberdeen, UK, 2–5 September 2016; Springer: Cham, Switzerland, 2016; pp. 213–226.



12. Peng, C.; Diao, W.; Du, Z. Digital Modulation Recognition Based on Deep Convolutional Neural Network. *Comput. Measurem. Control* **2018**, *26*, 222–226.
13. Hou, T.; Zheng, Y. Communication Signal Modulation Recognition Based on Deep Learning. *Radio Eng.* **2019**, *49*, 796–800.
14. Peng, S.; Jiang, H.; Wang, H.; Alwageed, H.; Zhou, Y.; Sebdani, M.M.; Yao, Y.-D. Modulation Classification Based on Signal Constellation Diagrams and Deep Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 718–727. [[CrossRef](#)] [[PubMed](#)]
15. Wang, Y.; Liu, M.; Yang, J.; Gui, G. Data-Driven Deep Learning for Automatic Modulation Recognition in Cognitive Radios. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4074–4077. [[CrossRef](#)]
16. Xie, W.; Hu, S.; Liao, J.; Yu, C.; Zhu, P.; Peng, X. Deep Learning in Digital Modulation Recognition Using High Order Cumulants. *IEEE Access* **2019**, *7*, 63760–63766. [[CrossRef](#)]
17. Tang, B.; Tu, Y.; Zhang, S.; Lin, Y. Digital Signal Modulation Classification with Data Augmentation Using Generative Adversarial Nets in Cognitive Radio Networks. *IEEE Access* **2018**, *6*, 15713–15722. [[CrossRef](#)]
18. Tu, Y.; Lin, Y. Deep Neural Network Compression Technique towards Efficient Digital Signal Modulation Recognition in Edge Device. *IEEE Access* **2019**, *7*, 58113–58119. [[CrossRef](#)]
19. Shi, W.; Liu, D.; Cheng, X.; Li, Y.; Zhao, Y. Particle Swarm Optimization-Based on Deep Neural Network for Digital Modulation Recognition. *IEEE Access* **2019**, *7*, 104591–104600. [[CrossRef](#)]
20. Liu, K.; Gao, W.; Huang, Q. Automatic Modulation Recognition Based on a DCN-BiLSTM Network. *Sensors* **2021**, *21*, 1577. [[CrossRef](#)] [[PubMed](#)]
21. Li, G.; RangZhuoma, C.; Zhijie, C.; Chen, D. Tibetan Voice Activity Detection Based on One-Dimensional Convolutional Neural Network. In Proceedings of the 2021 3rd International Conference on Natural Language Processing (ICNLP), Beijing, China, 26–28 March 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 129–133.
22. Yu, J.; Yi, Z.; Zhou, J. Continuous Attractors of Lotka–Volterra Recurrent Neural Networks with Infinite Neurons. *IEEE Trans. Neural Netw.* **2010**, *21*, 1690–1695. [[CrossRef](#)] [[PubMed](#)]
23. Palangi, H.; Deng, L.; Shen, Y.; Gao, J.; He, X.; Chen, J.; Song, X.; Ward, R. Deep Sentence Embedding Using Long Short-Term Memory Networks: Analysis and Application to Information Retrieval. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 694–707. [[CrossRef](#)]
24. Kim, K.; Kim, D.K.; Noh, J.; Kim, M. Stable Forecasting of Environmental Time Series via Long Short Term Memory Recurrent Neural Network. *IEEE Access* **2018**, *6*, 75216–75228. [[CrossRef](#)]
25. Wang, J.; Zhang, J.; Wang, X. Bilateral LSTM: A Two-Dimensional Long Short-Term Memory Model with Multiply Memory Units for Short-Term Cycle Time Forecasting in Re-entrant Manufacturing Systems. *IEEE Trans. Ind. Inform.* **2017**, *14*, 748–758. [[CrossRef](#)]
26. Jiao, S.; Wang, J.; Hu, G.; Pan, Z.; Du, L.; Zhang, J. Joint Attention Mechanism for Person Re-Identification. *IEEE Access* **2019**, *7*, 90497–90506. [[CrossRef](#)]
27. Cheng, K.; Yue, Y.; Song, Z. Sentiment Classification Based on Part-of-Speech and Self-Attention Mechanism. *IEEE Access* **2020**, *8*, 16387–16396. [[CrossRef](#)]
28. Wu, J.; Yang, H. Linear Regression-Based Efficient SVM Learning for Large-Scale Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 2357–2369. [[CrossRef](#)] [[PubMed](#)]
29. Yuan, H.; Fang, M.; Zhu, X. Hierarchical Sampling for Multi-Instance Ensemble Learning. *IEEE Trans. Knowl. Data Eng.* **2012**, *25*, 2900–2905. [[CrossRef](#)]
30. Zhang, S.; Li, X.; Zong, M.; Zhu, X.; Wang, R. Efficient kNN Classification with Different Numbers of Nearest Neighbors. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 1774–1785. [[CrossRef](#)] [[PubMed](#)]