Study on Fast-Changing Mixed-Modulation Recognition Based on Neural Network Algorithms

Qingfeng Jing[1]*, Huaxia Wang[1], and Liming Yang[2]
[1]College of Astronautics, Nanjing University of Aeronautics and Astronautics
Nanjing, Jiangsu 210019, China
[email: jing_nuaa@163.com, 424370790@qq.com]
[2]The 504th Research Institution, China Aerospace Science and Technology Corporation
Xi'an, Shanxi, China
[email: happylm@163.com]
*Corresponding author: Qingfeng Jing

Abstract

Modulation recognition (MR) plays a key role in cognitive radar, cognitive radio, and some other civilian and military fields. While existing methods can identify the signal modulation type by extracting the signal characteristics, the quality of feature extraction has a serious impact on the recognition results. In this paper, an end-to-end MR method based on long short-term memory (LSTM) and the gated recurrent unit (GRU) is put forward, which can directly predict the modulation type from a sampled signal. Additionally, the sliding window method is applied to fast-changing mixed-modulation signals for which the signal modulation type changes over time. The recognition accuracy on training datasets in different SNR ranges and the proportion of each modulation method in misclassified samples are analyzed, and it is found to be reasonable to select the evenly-distributed and full range of SNR data as the training data. With the improvement of the SNR, the recognition accuracy increases rapidly. When the length of the training dataset increases, the neural network recognition effect is better. The loss function value of the neural network decreases with the increase of the training dataset length, and then tends to be stable. Moreover, when the fast-changing period is less than 20ms, the error rate is as high as 50%. As the fast-changing period is increased to 30ms, the error rates of the GRU and LSTM neural networks are less than 5%.

Keywords: Neural Network, Fast-Changing, Mixed-Modulation, MR, LSTM, GRU

1. Introduction

Modulation recognition (MR) refers to the judgment of a received signal by a specific algorithm to determine the type of modulation of the signal, which provides a basis for signal demodulation. It is widely used in radio signal monitoring, electronic countermeasures, intelligent communication, and other fields. Fast-changing modulation signals whose modulation type changes rapidly over time play an important role in encrypting information, as they render the intercepted signal difficult to discriminate by intercepting devices, and therefore occupy an important position in military communications.

Two main categories of MR, namely likelihood-based (LB) and feature-based (FB) MR, have been defined in the existing literature. LB MR is a multiple composite hypothesis-testing problem that computes the likelihood ratios of selected received and known signals. Then, a decision is made by comparing this ratio to a threshold [1, 2]. This usually provides an optimal solution in the Bayesian sense, but involves a considerable amount of complexity. However, the LB methods are sensitive to parameter estimates and model mismatches. They also require high computational complexity to obtain a priori information, and are therefore not applicable in many practical communication scenarios. When the SNR is known and signal modeling can be approximated by a Gaussian distribution, the use of maximum likelihood estimation will achieve better results [3]. Existing methods based on likelihood estimation require the knowledge of a variety of a priori information of the signal, including the carrier frequency and code rate of the modulated signal. The acquisition of such information adds great difficulty to the design of the system, so many researchers have investigated how to simplify the likelihood function; however, this can also lead to information loss, resulting in inaccurate classification.

The FB approach is used to find key features of the signal, such as high-order moments or instantaneous features, and then to use machine learning classifiers to achieve MR [4]. Different classifiers will achieve different recognition accuracies. These classifiers include the decision tree(DT), support vector machine(SVM), naive Bayesian (NB), and k-nearest-neighbor (kNN) algorithms [5-9]. If the selected features and classifiers are not suitable, then the classification effect may be extremely poor, i.e., based on the method of artificial feature selection, the generalization ability is weak. A dictionary learning-based automatic modulation classification (AMC) framework has been proposed in which a dictionary is trained using signals with known modulation formats, and the modulation format of the target signal is determined by its sparse representation in the dictionary [10]. Deep learning automatically learns useful features through neural network algorithms. Convolutional neural networks (CNNs) have achieved great success in the field of computer vision. Because any kind of digital amplitude-phase modulation signal can be represented uniquely by a constellation diagram, the MR problem can be converted to an image recognition problem [11]. In addition to CNNs, deep belief networks (DBNs) and deep auto-encoders(DAEs) have also been applied to this field [12]. However, the constellation diagram does not work well at SNR<0dB. The communication signal of the time domain is a type of time-series data that can be processed directly to obtain more complete information. The recurrent neural network(RNN) is very suitable for the analysis of time-series data, and has contributed to great achievements in the fields of natural language processing and speech recognition. Because labeling datasets requires significant labor costs, semi-supervised learning using both unlabeled and labeled data for MR, such as via the use of GAN networks,

has been proposed [13, 14]. Rundong Li et al. [15] proposed a novel MR algorithm for very-high-frequency (VHF) radio signals based on anti-noise processing and a deep sparse filtering CNN (AN-SF-CNN). Y. Du [16] proposed an MR approach for Grassmann constellations based on data clustering, which differs from traditional MR based on classification using a modulation database. Y. Zeng et al. [17] presented a CNN-based MR framework for the detection of radio signals in communication systems, which transforms 1-D radio signals into spectrogram images using short-time discrete Fourier transform. P. Ghasemzadeh [18] analyzed the accuracy of high-order statistics-based (HoS) methods utilizing feature extraction approaches, SVM classifiers, and estimation techniques to determine an optimized framework for different real-time applications. An intra-pulse MR approach for single- and dual-component radar signals was proposed in the work by Z. Qu [19]. It designs and pre-trains a TFI feature extraction network for radar signals based on a CNN. Then, to improve the probability of successful recognition (PSR) of the recognition system in a pulse-overlapping environment, a multi-label classification network based on a deep Q-learning network (DQN) was explored. This paper introduces the application of the long short-term memory (LSTM) and gated recurrent unit (GRU) networks to the field of MR. An end-to-end MR method based on LSTM and GRU is put forward, which can directly predict the modulation type from a sampled signal. Additionally, the sliding window method is applied to the fast-changing mixed-modulation signal for which the signal modulation type changes over time.

The remainder of this paper is organized as follows. Chapter 2 introduces the MR block diagram and neural network algorithms. Chapter 3 introduces the parameters used in this study and the methods by which to choose a proper SNR range and list the optimizer and loss function of the neural network. Chapter 4 introduces the recognition results of single- and mixed-modulation signals. Additionally, the error rate of each type of modulation signal and the influences of the dataset length, SNR, modulation speed, feature visualization, and algorithm complexity are analyzed. Finally, Chapter 5 presents a summary of the full text.

2. Neural Network Algorithms

2.1 The MR block diagram and neural network algorithms

In this paper, four kinds of neural network algorithms, namely FC, RNN, LSTM, and GRU, are used to identify fast-changing mixed-modulation signals.
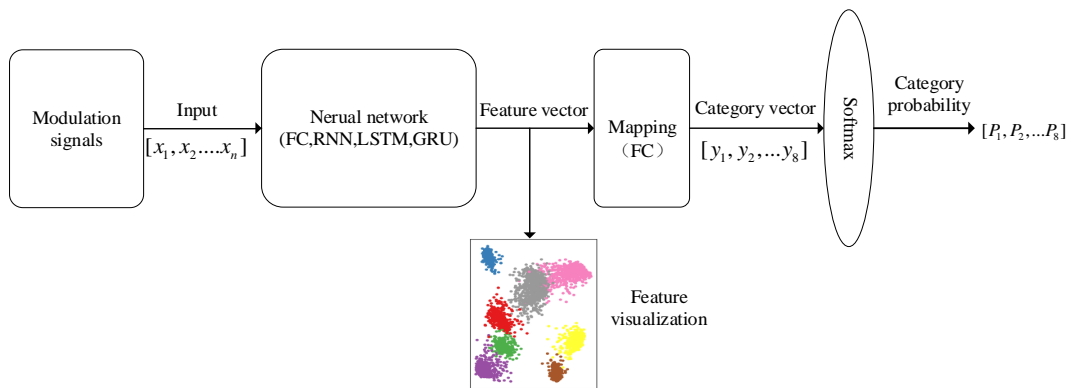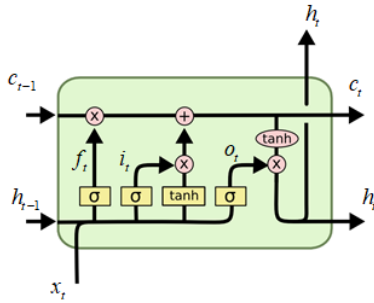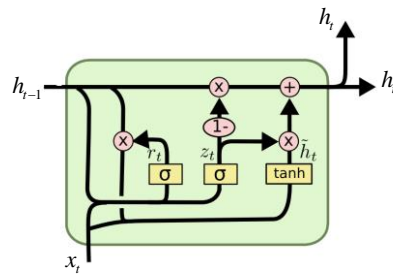

Fig. 1. Modulation recognition block diagram

$$\begin{cases} i_t = sigmoid\left(W_i x_t + U_i h_{t-1} + b_i\right) \\ f_t = sigmoid\left(W_f x_t + U_f h_{t-1} + b_f\right) \\ o_t = sigmoid\left(W_o x_t + U_o h_{t-1} + b_o\right) \\ c_t = f_t \square \ c_{t-1} + i_t \square \ tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \\ h_t = o_t \square \ tanh\left(c_t\right) \end{cases} \tag{1}$$
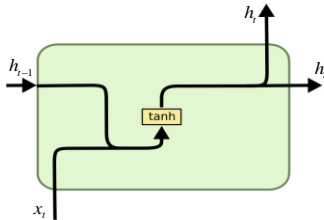
Fig. 2. LSTM structure

$$\begin{cases} z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right) \\ r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right) \\ \tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right) \\ h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{cases} \tag{2}$$
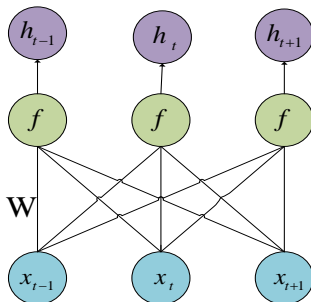

Fig. 3. GRU structure

$$h_t = \tanh(w_h h_{t-1} + w_i x_t) \tag{3}$$

Fig. 4. RNN structure

$$h_t = f(w_1 x_{t-1} + w_2 x_t + w_3 x_{t+1} + b) \tag{4}$$

Fig. 5. FC structure

Fig. 1 is the network frame diagram based on the neural network algorithm, in which the input X is the sampling signal, the output Y is the probability value of 8 signals, and the intermediate neural network is equivalent to a function F_{nn}. The appropriate w and b values are obtained via training to make $Y = F_{nn}(X)$. The specific steps are as follows. First, a 1024-point sampling signal is designed for 64time steps, namely $X(x_1, x_2, x_3.....x_{64})$. Each time step contains 16 sampling points, which are then inputted to the neural network. A neural network equivalent to a feature extractor can obtain the characteristics of the input signal vector. The FC layer plays a role in mapping feature vectors into class vectors, the probability value of each category of the network is then obtained by the softmax function, and the modulation category with the highest probability is the category of the signal recognized by the neural network. F_{nn} includes LSTM, GRU, RNN, and FC, and the specific operation of LSTM is presented in **Fig. 2**, where x_t represents the point at a certain time of the input $X(x_1, x_2, x_3.....x_{64})$ of the neural network, the dimension of which is 16, h_t represents the output of the hidden layer, the dimension of which is 32, and σ represents the sigmoid function. Its equation is as follows:

$$s(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

Additionally, the tanh function is as follows:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{6}$$

The specific calculation formula from x_t to h_t can be obtained by (1), where W, U, and b represent the neural network weight and bias. It can be seen from **Fig. 2** that the value of h_t depends not only on the input x_t for the current moment, but also on the output of the moment h_{t-1}. For the initial value h_0, $h_0 = 0$ is set. **Fig. 3** and **Fig. 4** respectively present the GRU and RNN internal operation diagrams, in which the parameter definitions are the same as those for the LSTM network.

In general, FC, as shown in **Fig. 5**, is a simple feed-forward neural network in which the neurons of each layer do not transmit information, and its structure is relatively simple; in contrast, RNN, LSTM, and GRU can transmit information between neurons in each layer. The output of the next moment depends not only on the input of the current moment, but also on the output of the previous moment. The major differences between these three networks are the calculation methods inside the neurons, and these internal differences are presented in **Fig. 2-4**. The calculation method of RNN is the simplest, that of GRU is the second-simplest, and that of LSTM is the most complex. Because the gating units of LSTM and GRU can solve the problem of gradient disappearance, their effects on long-sequence problems are better than that of RNN.

2.2 Section summary

In this section, the diagram of the modulation identification method and the four neural network algorithms FC, RNN, LSTM, and GRU were introduced. The sampled signal can be directly predicted as the probability of each signal via the neural network, and the signal type corresponding to the probability value of the maximum value is the type of the sampled signal.

3. Neural Network Training Process

3.1 Selection of the training datasets

The parameters used in this study are listed in **Table 1**. Eight kinds of modulation signals were modulated at an SNR range from -20to20 dB and a stride of 2dB.There were 1024 sampling points for each signal, and the total number of signals was 8000. Additionally, the ratio of the training datasets to the validation datasets to the test datasets was 6 to 2 to 2. This ratio is considered appropriate in machine learning.

Table 1. The parameters of the present study

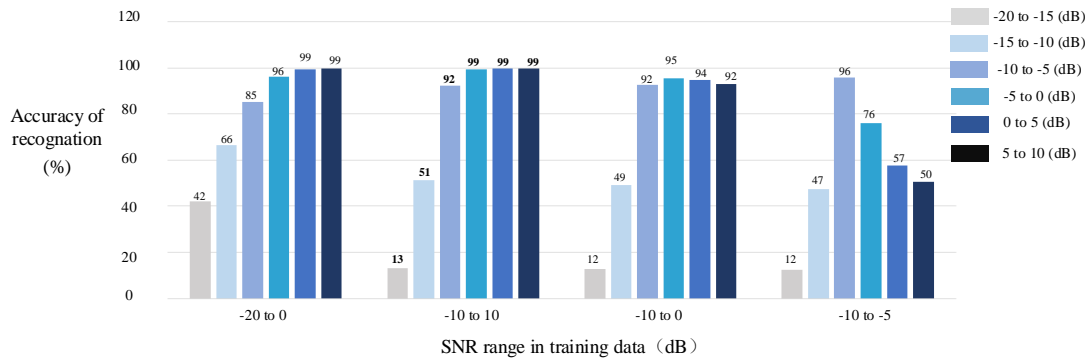| Datasets	Length	SNR	Modulation Type	Sample points	Frequency	Sample frequency
Training	6000	-20 to 20dB	2ASK,4ASK,2FSK MSK,BPSK,QPSK 16QAM,64QAM	1024	10MHz	40MHz
Validation	2000					
Test	2000					

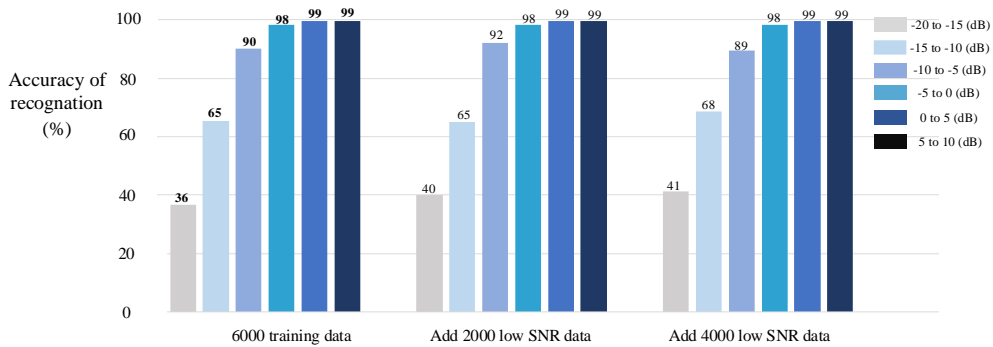**Fig. 6.** Accuracy vs. the SNR range for training data

Fig. 7. Accuracy vs. the length of low-SNR data

Some experiments were conducted to select the SNR interval of the training datasets. **Fig. 6** presents the accuracies of six SNR intervals while training data with different SNRs and keeping the length of the training data constant; additionally, the proportion of each category of SNR data was the same. The recognition effect of training data with SNRs of -10 to 10dB was found to be 7.2% higher at -10dB to -5dB than that of training data with SNRs of -20 to -10dB, and the accuracy was increased by 3% at SNRs of -5 to 0dB. However, in the case of a data with low SNRs of -20 to -15dB, the recognition effect was reduced by about 10%, and the recognition accuracy of data with SNRs of -15 to -10dB was reduced by 15%. Compared with the previous two SNR intervals, the results of training with data with SNRs of -10 to 0dB or -10 to -5dB exhibited significant decreases in the range of 0 to 10dB, but the recognition accuracy of data with SNRs of -10 to 0dB increased. The potential improvement of the recognition accuracy of low-SNR data by training with more -20 to -10dB data required investigation. Thus, related experiments were conducted, as presented in **Fig. 7**, which presents the relationship between the accuracy and the increases of low-SNR (-20 to -10dB) data. It can be seen from the figure that the amount of low-SNR data was increased by 2000 and 4000, but there was no significant improvement in accuracy as compared to the original recognition effect. Based on the results of **Fig. 6** and **Fig. 7**, it is reasonable to select the full range of SNR data as training data, and the data is evenly distributed under the SNR. However, if the SNR interval is increased, then the complexity level of the system will increase. It is reasonable to choose the SNR interval of -10 to 10dB, as the increase from -10 to -20dB does not significantly improve the accuracy, and the accuracy on datasets with SNRs above 10dB was similar to that on datasets with SNRs below10dB.

3.2 Neural network training algorithm

The RMSProp optimization algorithm is adopted in the present study to train the neural network; the weight w and bias b of the neural network must be optimized until the final convergence of the network. The RMSProp algorithm uses the weighted mean of differential squares for the gradient of weight w and bias b. In the iteration process of the t-th round, each formula is as follows:

$$\begin{cases} s_{dw} = \beta s_{dw} + (1-\beta)dw^2 \\ s_{db} = \beta s_{db} + (1-\beta)db^2 \\ w = w - \alpha \dfrac{dw}{\sqrt{s_{dw} + \varepsilon}} \\ b = b - \alpha \dfrac{db}{\sqrt{s_{db} + \varepsilon}} \end{cases} \tag{7}$$

where s_{dw} and s_{db} are the gradient momentum accumulated by the loss function during the previous t-1 iteration and β is an index. The RMSProp algorithm calculates the weighted average of the differential square for the gradient, which increases the convergence speed of the network function. To prevent the denominator from being zero, a small value of ε should be used for smoothing; in this work, the value of ε is taken as 10^{-8}.

3.3 The loss function of the neural network

The advantages and disadvantages of the neural network algorithm primarily depend on the value of the loss function when the network converges. Loss is the value of the cross-entropy loss function, and the smaller the value, the better the network model. Its formula is as follows:

$$L = -\frac{1}{N} * \sum_{N} [y * \ln P + (1-y) * \ln(1-P)] \tag{8}$$

where N represents the length of data, y represents the true one-hot tag vector, and P represents the softmax vector predicted by the neural network. The softmax vector is the ratio of the exponent of the element to the exponential sum of all elements, and is given by

$$P = \frac{e^i}{\sum_{i=1}^{N} e^i} \tag{9}$$

The changes in the loss function values of the validation datasets with the epochs for the four neural network algorithms are presented in **Fig. 8**.
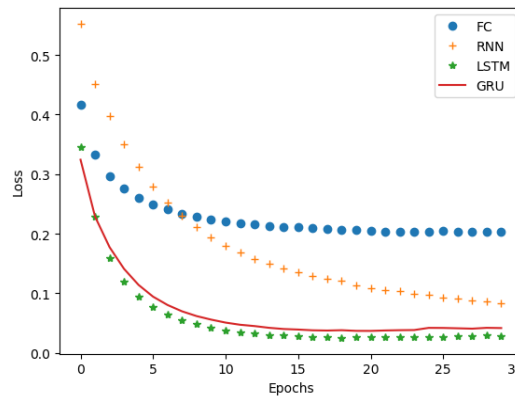

Fig. 8. The losses of the four neural networks vs. epochs

It is evident from **Fig. 8** that with the increase of network training rounds, the loss function values of the four algorithms gradually decreased and eventually converged; additionally, the smaller the loss function, the smaller the gap between the true tag and predicted tag, and the better the neural network algorithm. After the same number of iterations, the performance of the LSTM network was best, that of GRU was the second-best, followed by that of RNN, and that of FC was the worst. Due to structural constraints, RNN cannot perform as well as LSTM and GRU. Its earlier input has very little impact on the network at present, and it is difficult to influence the gradient of the backward propagation when the input is from long ago, i.e., the gradient will disappear. LSTM and GRU solve this problem by building some door structures. They can remember very important information, and the core of the structure is the door unit, such as the "forgetting door," which selectively clears the past memory and updates the new information. The FC cannot capture the correlation of input sequences, so it has a poor effect on signal modulation identification.

Table 2 lists the specific values of several loss functions. In order from greatest to least, the network performance was found to be as follows: LSTM > GRU > RNN > FC. The overall performances of LSTM and GRU were found to be similar; LSTM was only slightly better than GRU, but the training time of GRU was shorter than that of LSTM. Thus, if time is of concern, GRU can be considered for use over LSTM.

Table 2. The loss values of the four neural networks

Networks	FC	RNN	LSTM	GRU
Loss	0.203	0.083	0.028	0.038

3.4 Influence of the length of training datasets on the prediction results of neural network algorithms

The datasets of signals are sometimes limited, especially in an unfamiliar communication environment. The length of a dataset has a certain influence on the prediction results of neural network algorithms. In this study, the length of the training dataset was increased from 1000 to 10000, and the accuracies of the converged neural network model on the test datasets were then observed.
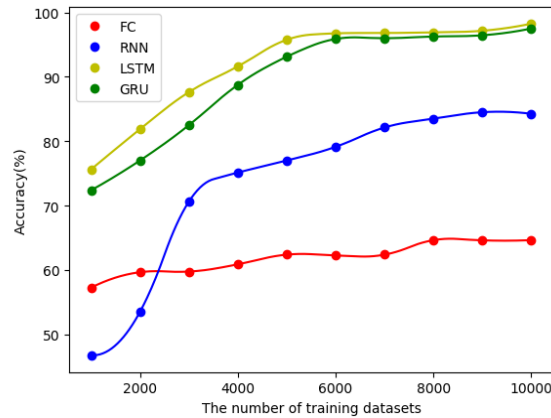

Fig. 9. Recognition accuracy vs. length of the training dataset

It is evident from **Fig. 9** that the greater the length of the training dataset, the better the neural network recognition performance; however, after a certain value, the increase of the length of data did not result in a very effective improvement. This is because when the size of a dataset is too small, the robustness of the network is not good, and the phenomenon of over-fitting is prone to occur; the recognition accuracy of non-participating training data was therefore not high. The increase of the length of the training dataset can enhance the generalization ability of the neural network and ultimately improve the accuracy. However, the accuracy of the FC network could not be improved because its neural network structure is too simple to learn the useful features of time-series data.

3.5 Section summary

In this section, the parameters and the method by which to choose a proper SNR range were introduced. Moreover, the optimizer and loss function of the neural network were presented. The relationships between the loss value and number of epochs of the four neural network

algorithms were compared, and the results confirmed that LSTM and GRU are good at MR. Finally, the influence of the length of the training dataset on the accuracy was analyzed.

4. Prediction results of the four neural networks

4.1 The recognition result of single-modulation signals

After the neural network training was completed, the converged network models were saved and tested with the testing datasets. The confusion matrixes of the four neural networks on the testing datasets are presented in **Fig. 10**.
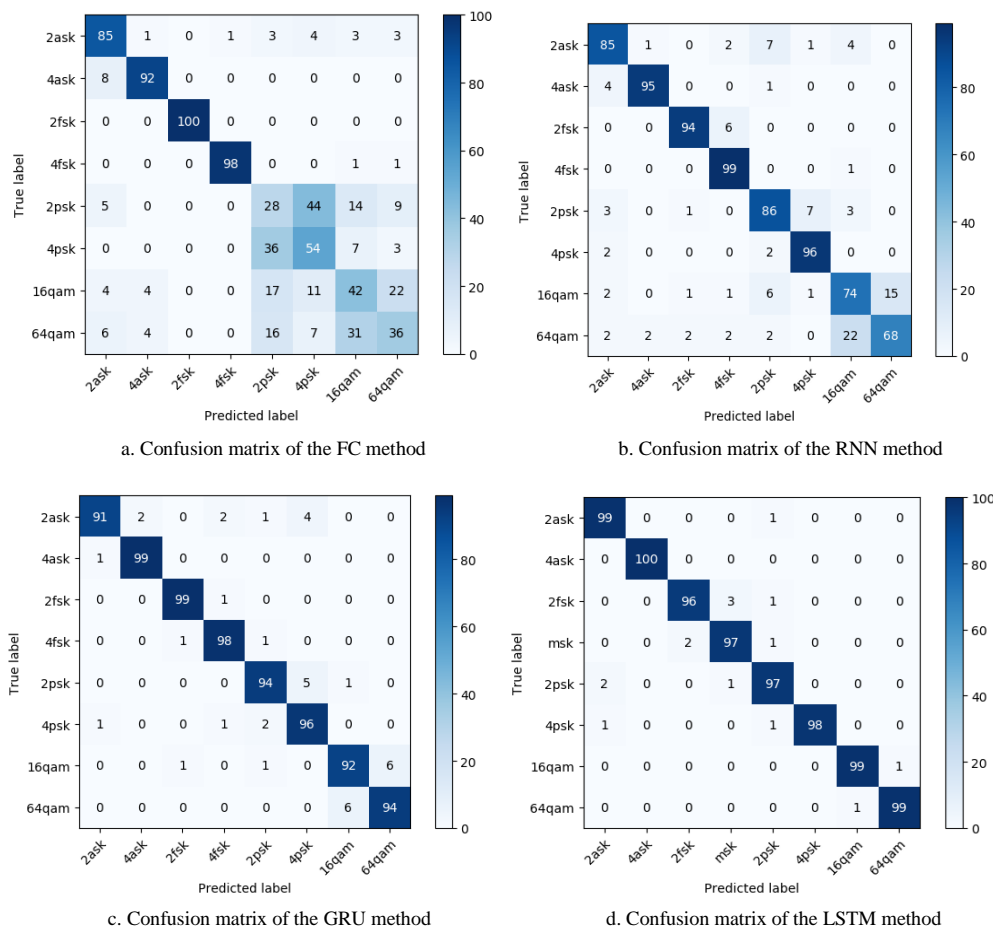

a. Confusion matrix of the FC method

b. Confusion matrix of the RNN method

c. Confusion matrix of the GRU method

d. Confusion matrix of the LSTM method

Fig. 10. Confusion matrixes of the four neural network algorithms

It is clear from **Fig. 10** that the FC network recognition effect was good for the 4ASK, 2FSK, and MSK modulation types, but it had a poor recognition effect for the 2PSK, 4PSK, 16QAM, and 64QAM modulation types. RNN had a poor recognition effect for only the 16QAM and 64QAM modulation types. However, the accuracies of LSTM and GRU for each modulation type were greater than 90%, which corresponded to the value of the loss function. LSTM and GRU have gate structures, which can easily remember long dependency problems. In other words, if there was an earlier important feature, if the forgetting gate or update gate chooses not to rewrite the internal memory unit, then the network will always remember the important

features, which will then continue to have an impact on the present or future. This design also provides a shortcut between states so that the gradient does not disappear too quickly when it is passed back, thus reducing the difficulty of training caused by the vanishing gradient.

Table 3 lists the recognition accuracies of the four network algorithms under the mixed SNR, -10dB, and 10dB conditions. It is evident from Table 3 and **Fig. 6** that the LSTM and GRU networks had a good recognition effect for each type of modulation signal, which also proves that the loss functions of the two networks are small and their recognition accuracies are high. These two networks can be applied to any type of modulation signal. The FC network can better identify the ASK and FSK modulation types, so the network is only suitable for simple modulation signals.

Table 3. The accuracies of the four algorithms

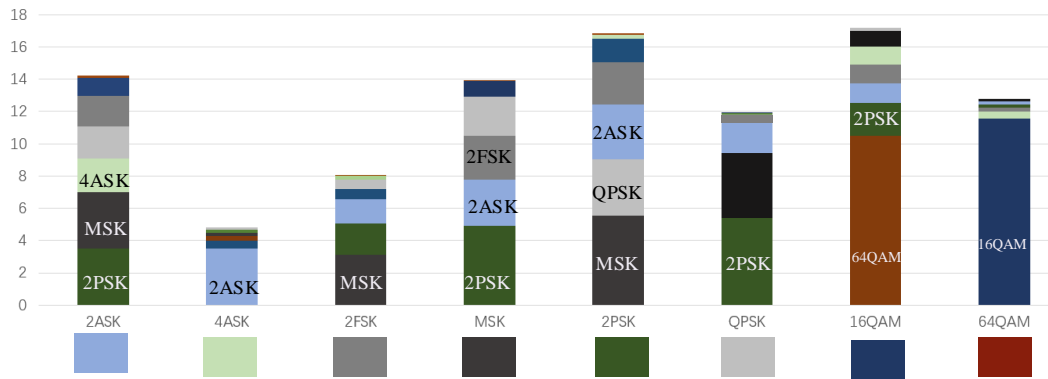|	FC	RNN	LSTM	GRU
Mixed SNR	66.875	86.125	97.75	95.0
-10dB	45.25	49.25	80.5	78.5
10dB	70.25	91.875	99.5	99.0

Fig. 11. The proportion of the error rate of each modulation signal for the LSTM algorithm

Fig. 11 presents the proportions of the eight modulation methods in error classification and the proportion of each modulation method that was misclassified into another modulation category. The height of the histogram represents the proportion of each modulation method. The small columnar blocks represent the proportions that were misclassified into the remaining modulation methods. It is clear that the ratio of 2ASK to 2PSK was the highest, followed by that of MSK to 4ASK, and 4ASK was mainly misclassified as 2ASK. It can also be seen that 2ASK had a higher error ratio than 4ASK, 2PSK had a higher error ratio than QPSK, and 16QAM had a higher error ratio than 64QAM;this is because 2ASK was incorrectly classified as the 2PSK, MSK, 4ASK, and QPSK modulation types. Additionally, 4ASK was mainly misclassified as 2ASK, so the probability of a 2ASK classification error under the influence of noise will be higher than that of 4ASK, and the same is true for the other cases.

4.2 The result of fast-changing mixed-modulation recognition

Fast-changing mixed-modulation recognition means that when transmitting one signal, not only a single modulation method is used; the signal modulation type changes rapidly with time. In this study, the received signals were divided into equal intervals, and each signal interval was used as an input to the neural network. In other words, each modulation period was

predicted to be in a modulation category, and they were then combined. The four neural network models that were previously trained were saved for direct prediction, and the parameters are listed in **Table 4**.

Table 4. Parameters of the simulation signal

Changing speed	SNR	Datasets length	Sampling	Frequency
0.2	0dB	5s	40MHz	10MHz

The parameters of the fast-modulation speed, SNR, feature visualization, and algorithm complexity were respectively analyzed.

4.2.1 Influence of Mixed-modulation Speed on Prediction Results

In the transmission of signals, the faster the modulation type changes, the more difficult it is to be deciphered. The prediction results of the four neural network algorithms for the modulation signals with a 0.2-s variable time period are presented in **Fig. 12**.
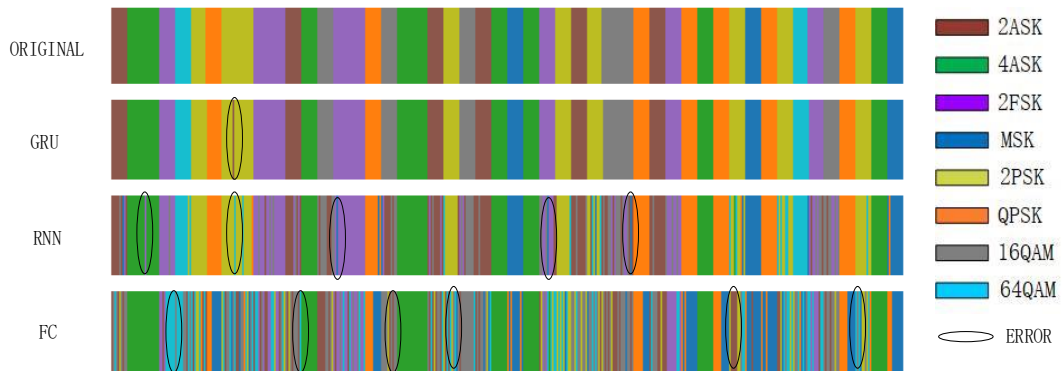

Fig. 12. Recognition effect of mixed modulation (fixed mixed-modulation speed)

As shown in **Fig. 12**, one signal was transmitted with different modulation types at intervals of 0.2 s, and three neural network models were used to predict it. GRU was found to identify the signal modulation type very well, followed by RNN, and finally FC. This is because the recognition performance of GRU for fast-changing mixed-modulation signals is related to the single-MR performance of the neural network algorithm.

The relationships between the error rates of the four algorithms and the fast mixed-modulation period are presented in **Fig. 13**.
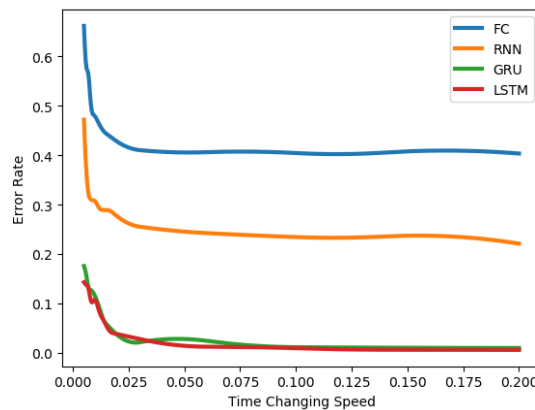

Fig. 13. Error rate vs. modulation speed

When the mixed-modulation period was relatively small, the recognition effect was poor. After the period became greater than 0.02s, the recognition performance tended to be stable, and the error rate gradually became close to the error rate of single-MR.

Similarly the method is also applicable to the case of a variable change period. The result is presented in **Fig. 14**.
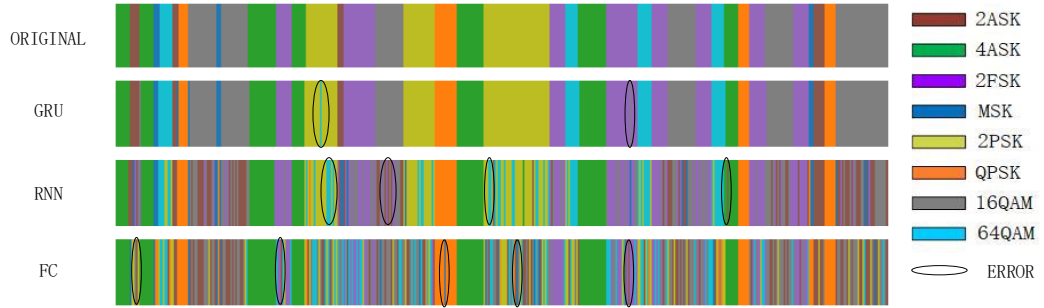

Fig. 14. Recognition effect of modulation signal with a variable mixed-modulation speed
(the speed varied as a multiple of the shortest time of 0.1s)

Table 5 reveals that the recognition effect of the modulation signal with a variable mixed-modulation speed was equivalent to that of the signal with a uniform speed. This is because, although the time interval varied unevenly, a separate prediction of the time signal was used for each segment; thus, as long as the time interval is greater than 0.02 s, there is almost no effect on the recognition accuracy, regardless of the period.

Table 5. The accuracies of the four networks for a variable mixed-modulation speed

Networks	FC	RNN	LSTM	GRU
Acc(%)	69.375	89.875	98.75	97.875

4.2.2 Influence of SNR on Prediction Results of Neural Network Algorithms

In actual communication environments, signals will be mixed with Gaussian noise. It is therefore necessary to analyze the influence of different SNRs on the prediction results of the neural network algorithms. The AWGN function was used to add Gaussian white noise with different SNRs to the original signal. The SNR expression is as follows:

$$SNR_{dB} = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \tag{10}$$

where P_{signal} represents the original signal power and P_{noise} represents the noise signal power.
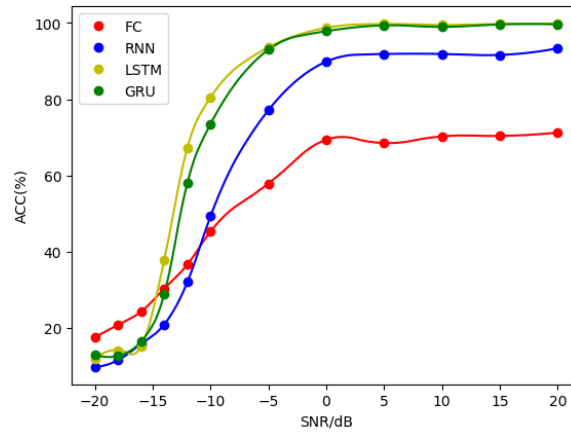
Fig. 15. Recognition accuracy vs. SNR

It is evident from **Fig. 15** that the neural network recognition effect was poor when the SNR was low; however, the neural network recognition performance was greatly improved when the SNR was greater than -10dB, and the recognition effect tended to be stable when the SNR was greater than 0dB.When the SNR was less than -10dB, the signal was basically covered by noise, i.e., it was difficult for the neural network to learn useful features, and only the features of the noise could be learned; thus, the recognition effect was poor. With the increase of the SNR, the neural network could gradually extract some useful features, and the recognition effect steadily improved. When the SNR reached 0dB, the signal strength was sufficient. At this time, the neural network model was less affected by noise and could extract enough useful features from the data for identification. Even if the SNR is improved, the final recognition effect will not change substantially.

4.2.3 Fecture Visualization

A visualization of the output of the neural network is presented in **Fig. 16**.
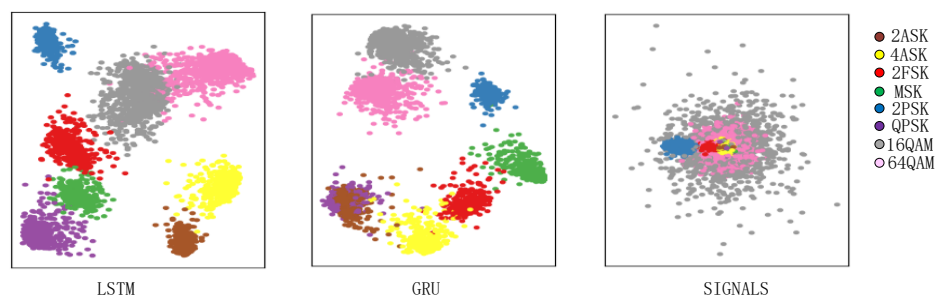

Fig. 16. Feature visualization

The left side of **Fig. 16** is a two-dimensional feature map obtained by reducing the dimension of the 32-dimensional feature outputted by the LSTM network via the PCA algorithm. The middle of the figure is a two-dimensional feature map obtained by reducing the dimension of the 32-dimensional feature outputted by the GRU network, and the right side of the figure was obtained by reducing the dimension of the original signal. It can be seen that each category on the left and middle sides is separated, and those on the right side are mixed

and indistinguishable, which demonstrates that LSTM and GRU can effectively extract and classify the features of the original signal.

4.2.4 Algorithm Complexity Analysis

The training of neural networks takes a certain amount of time. In industry applications, the datasets are often very large, on the order of GB or even TB. For such large datasets, the training time is often very long and the parameters (such as the learning rate, the number of network layers, the number of hidden layer neurons, etc.) must be constantly adjusted to achieve optimal results. Because is necessary to select the network structure as simply and effectively as possible, the complexity of the algorithm must be analyzed.

Table 6 reveals that the number of parameters for FC was the largest, but its training and prediction times were the shortest. This is because, as compared with the RNN network, the FC network does not need to calculate the recursion of the time signal; it just loads all the inputs at a time, and then forwards the values to obtain the outputs. The RNN, LSTM, and GRU networks need to calculate the current neural value each time, and then transfer the current result to the next neuron. From Table 6, it can be seen that the number of parameters for RNN was the least, but its effect was not good because it experienced the problem of gradient disappearance. Based on these findings, LSTM and GRU are proposed to solve this problem. LSTM has one more gate structure than does GRU, so its number of parameters is 1/3 greater than that of GRU and its computing time will be 1/3 longer. If the amount of data is large, the time will not be a small overhead. Therefore, in the case of high accuracy requirements, the LSTM neural network algorithm should be selected, while the GRU neural network algorithm should be selected in cases in which both accuracy and time efficiency are of concern.

Table 6. The complexities of the four neural networks

Neural network	Number of parameters	Time of training for a step	Time of prediction
FC	33064	0.58	0.434
RNN	1832	2.189	0.882
LSTM	6536	6.69	2.343
GRU	4968	4.536	1.917

4.3 Section Summary

In this section, the recognition results of single- and mixed-modulation signals were introduced. Additionally, some influencing factors, including the error rate of each type of modulation signal, the dataset length, the SNR, the modulation speed, feature visualization, and the algorithm complexity, were analyzed. The LSTM and GRU algorithms were found to perform well for MR.

5. Conclusion

In this work, neural network algorithms were successfully applied to the field of mixed-modulation recognition. The experimental results demonstrated that the LSTM and GRU networks exhibited better effects, and that the recognition accuracy of each type of

modulation signal reached 90%. This study also presented the error proportion of each type of modulation and the proportion of each error sample misclassified as other samples. By training on data with different SNR intervals, it was concluded that the SNR interval of the training dataset should be as wide as possible, and the length of SNR data should be set to be similar to avoid overfitting. The influences of multiple parameters, namely the SNR, mixed-modulation speed, dataset length, and algorithm complexity, for fast-changing mixed-modulation recognition were also analyzed. When the SNR was 0 dB, the recognition performance of both the GRU and LSTM networks achieved greater than 95% accuracy. When the mixed-modulation speed was less than 20ms, the error rate was higher, up to 50%. As the mixed-modulation speed increased, the error rates of GRU and LSTM were reduced to less than 5% at a mixed-modulation speed greater than 30ms. The increase in the length of the training dataset can significantly improve the robustness of the neural network, and ultimately leads to better results on the test dataset; therefore, longer datasets are very important for neural network algorithms.

Acknowledgement

References

[1] A. Polydoros and K. Kim, "On the detection and classification of quadrature digital modulations in broad-band noise," *IEEE Trans. Commun.*, vol. 38, no. 8, pp. 1199-1211, Aug. 1990. Article (CrossRef Link)

[2] J. L. Xu, W. Su, and M. Zhou, "Likelihood function-based modulation classification in bandwidth-constrained sensor networks," in *Proc. of 2010 International Conference on Networking, Sensing and Control (ICNSC),* pp. 530-533, Apr. 2010. Article (CrossRef Link)

[3] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189-193, Feb. 2000. Article (CrossRef Link)

[4] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, "An overview of feature-based methods for digital modulation classification," in *Proc. of 2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA),* pp. 1-6, Feb. 2013. Article (CrossRef Link)

[5] J. L. Tan, A. Z. bin Sha'ameri, and Y. M. Chee, "Signal analysis and classification of digital communication signal using higher order time-frequency analysis techniques," in *Proc. of 2010 International Symposium On Information Theory & Its Applications,* pp. 752-756, Oct. 2010. Article (CrossRef Link)

[6] J. Li, L. Qi, and Y. Lin, "Research on modulation identification of digital signals based on deep learning," in *Proc. of 2016 IEEE International Conference on Electronic Information and Communication Technology (ICEICT),* pp. 402-405, Aug. 2016. Article (CrossRef Link)

[7] M. W. Aslam, Z. Zhu, and A. K. Nandi, "Autmomatic Modulation Classification Using Combination of Genetic Programming and KNN," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2742-2750, Aug. 2012. Article (CrossRef Link)

[8] Z. Wu, S. Zhou, Z. Yin, B. Ma, and Z. Yang, "Robust Automatic Modulation Classification Under Varying Noise Conditions," *IEEE Access*, vol. 5, pp. 19733-19741, Aug. 2017. Article (CrossRef Link)

[9]   E. Avci and D. Avci, "Using combination of support vector machines for automatic analog modulation recognition," *Expert Systems with Applications*, vol. 36, no. 2-part-P2, pp. 3956-3964, Mar. 2009. Article (CrossRef Link)

[10] K. Zhang, E. L. Xu, Z. Feng, and P. Zhang, "A Dictionary Learning Based Automatic Modulation Classification Method," *IEEE Access*, vol. 6, pp. 5607-5617, Jan. 2018. Article (CrossRef Link)

[11] M. Zhang, M. Diao, and L. Guo, "Convolutional Neural Networks for Automatic Cognitive Radio Waveform Recognition," *IEEE Access*, vol. 5, pp. 11074-11082, June 2017. Article (CrossRef Link)

[12] A. Ali and F. Yangyu, "Automatic Modulation Classification Using Deep Learning Based on Sparse Autoencoders With Nonnegativity Constraints," *IEEE Signal Processing Letters,* vol. 24, no. 11, pp. 1626-1630, 2017. Article (CrossRef Link)

[13] 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, Mar. 2018. Article (CrossRef Link)

[14] A. Ali and F. Yangyu, "Unsupervised feature learning and automatic modulation classification using deep learning model," *Physical Communication*, vol. 25, pp. 75-84, Dec. 2017. Article (CrossRef Link)

[15] R. Li, L. Li, S. Yang, and S. Li, "Robust Automated VHF Modulation Recognition Based on Deep Convolutional Neural Networks," *IEEE Communications Letters*, vol. 22, no. 5, pp. 946-949, May 2018. Article (CrossRef Link)

[16] Y. Du, G. Zhu, J. Zhang, and K. Huang, "Automatic Recognition of Space-Time Constellations by Learning on the Grassmann Manifold (Extended Version)," *IEEE Transactions on Signal Processing*, vol. 66, no. 22, pp. 6031-6046, Nov. 2018. Article (CrossRef Link)

[17] Y. Zeng, M. Zhang, and F. Han, Y. Gong, and J. Zhang, "Spectrum Analysis and Convolutional Neural Network for Automatic Modulation Recognition," *IEEE Wireless Communication Letters*, vol. 8, no. 3, pp. 929-932, June 2019. Article (CrossRef Link)

[18] P. Ghasemzadeh, S. Banerjee, M. Hempel, and H. Sharif, "Accuracy Analysis of Feature-based Automatic Modulation Classification with Blind Modulation Detection," in *Proc. of 2019 International Conference on Computing, Networking and Communications (ICNC),* pp. 1000-1004, Feb. 2019. Article (CrossRef Link)

[19] Z. Qu, C. Hou, and W. Wang, "Radar Signal Intra-Pulse Modulation Recognition Based on Convolutional Neural Network and Deep Q-Learning Network," *IEEE Access*, vol. 8, pp. 49125-49136, 2020. Article (CrossRef Link)
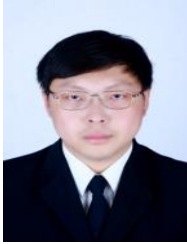
Jing Qingfeng was born in 1981. He received the Master and Ph.D. degree in School of Electronics and Information Engineering, Harbin Institute of Technology in 2005 and 2009, respectively. He has been working as associate professor since 2009 in College of Astronautics, Nanjing University of Aeronautics and Astronautics. He is now in charge of several funded projects. His research interests are digital signal processing, satellite communication，and broadband multi-carrier communication.

Wang Huaxia was born in 1996. He is a Graduate student in the Collegee of Astronautics, Nanjing University of Aeronautics and Astronautics. His research interests are modulation recognition and artificial intelligence (AI).

Yang Liming was born in 1982. He is an undergraduate at Harbin Institute of Technology University of Aerospace. His research interests are modulation and measurement and control communication systems.