Continuous IFF Response Signal Recognition Technology Based on Capsule Network

Yifan Jiang[1], Zhutian Yang[1](✉), Chao Bo[2], and Dongjia Zhang[3]

[1] School of Electronic and Information Engineering, Harbin Institute of Technology, Harbin, China
yangzhutian@hit.edu.cn
[2] Nanjing Institute of Electronic Equipment, Nanjing, China
[3] China Aerospace Science and Industry Corporation Limited, Beijing, China

Abstract. Identification of friend or foe (IFF) system has become an indispensable part in modern war. In order to meet the needs of air target situation control in rapid response operations, it is urgent to find an intelligent IFF signal recognition method. Aiming at the problems of low recognition accuracy and high false alarm rate of continuous IFF signal of single channel multiple air maneuvering targets in low SNR environment, a signal pattern recognition method of continuous IFF signal based on capsule network and attention mechanism in complex environment is proposed by improving signal data set and capsule network model structure. Using the good generalization ability and strong feature interpretation ability of attention mechanism provided by capsule network, the improved method has a certain degree of improvement in the pattern recognition ability of simulated complex signals compared with traditional frame detection method and multilayer convolutional neural network. At the same time, the false alarm rate and the missed alarm rate are significantly improved, which can meet the actual detection requirements.

Keywords: Identification of friend or foe (IFF) · DM-CapsNet · Attention mechanism · Co-frequency interference

1 Introduction

With the development of the times, the demand for rapid response combat air mobile target situation control is increasing, the electromagnetic environment has become more and more complex, and multiple modulation, coding methods, and encryption protocols have begun to be applied to the identification friend or foe system.

There are some common algorithms for IFF signal recognition. As far as multi-channel array received signals are concerned, the ICA algorithm [1] proposed by P·Comon and the Fast-ICA algorithm [2] proposed by Hyvarinen can be applied to the separation of IFF signals. However, the high-order statistics

X. Wang et al. (Eds.): AICON 2021, LNICST 396, pp. 455–468, 2021.
https://doi.org/10.1007/978-3-030-90196-7_39

of IFF signals have been proved to be pseudo-Gaussian, so any signal classification and recognition algorithm based on kurtosis is not robust. At present, the projection algorithm (PA) proposed by Petrochilos [3] is more effective. This algorithm can realize effective signal recognition in the case of different mode signals. However, when the signal is polluted by noise, the sorting effect of the algorithm is poor, and the amplitude of the sorted signal is quite different from the original signal, which is easy to cause decoding errors.

As far as the single channel received signal is concerned, because each mode response signal is on the same carrier frequency, it can only use the pulse frame information for detection, and the feature is relatively single, so there are few recognition algorithms related to machine learning. At present, sliding window method is commonly used [4]. After combining correlation detection and pulse PRI sequence analysis, this method also has the ability to detect the response signals of new mode signals such as Mode S and Mode 5. However, this method only uses frame information in real-time signal detection, because the signal amplitude changes too much, there will be false alarm and missing alarm. At the same time, because of the need to detect multiple modes of the framework, and mode 5 synchronization pulse also need to carry out MSK demodulation and correlation detection, it takes too long.

On the other hand, through the actual test, it is found that although the traditional convolutional neural network and the general capsule network have good detection performance in the verification set, they also have a high false alarm rate in the continuous signal detection, causing serious interference to the whole recognition system. In order to solve the above problems, this paper proposes a common channel automatic pattern recognition method (DM-CapsNet) based on capsule network. This method directly processes the intermediate frequency signal, uses supervised learning method, and directly takes the simulated low signal-to-noise ratio mode signals as the training set. When verifying the measured signals, the classification results are obtained by threshold determination.

2 Problem Statement

2.1 IFF Signal Model

The Mark X mode response signal includes 16 information codes, and the sequence of the signals is F1, C1, A1, C2, A2, C4, A4, X, B1, D1, B2, D2, B4, D4 and F2 in the order of time, where the last one is SPI pulse. The format of the response signal is shown in Fig. 1. Each code has two states, that is, with pulse or without pulse, with pulse as '1' state and without pulse as '0' state. F1 and F2 are called frame pulses with a time interval of 20.3 us $\pm$ 0.1 us. They are the flag pulses of the response signal, which are always in the '1' state, and X is the spare bit, which is always in the '0' state. SPI is a special location identification code, which will not be used in general. The time interval between the pulse at each code point and the front edge of the F1 pulse is an integer multiple of 1.45 us. The SPI pulse is at 4.35 us after the F2 pulse with an allowable tolerance of $\pm$0.1 us, and the width of each pulse is 0.45 us.

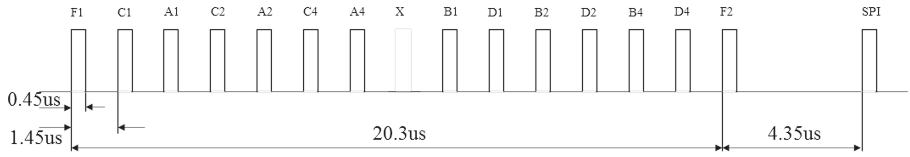


Fig. 1. Mark X series reply pulse.

Mark XII system is the second generation of IFF enemy identification systems, which adds a secure mode (Mode 4) on the original basis, and its response signal format is shown in Fig. 2. The mode signal is a response pulse group composed of three pulse widths of 0.45 us $\pm$ 0.1 us, and the pulse interval is 1.75 us $\pm$ 0.1 us. The responder starts to respond after receiving the query pulse P4 and fixing the delay 202 us. The response pulse group starts at the delay $t_x = (202 + 4N)$ us (N = 0,1,2...15).

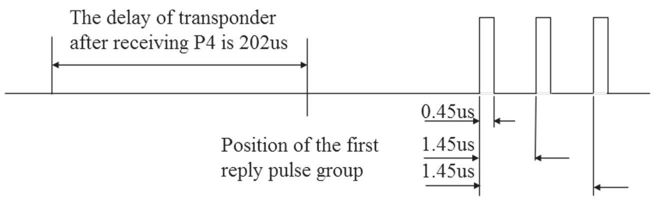


Fig. 2. Mode 4 response signal.

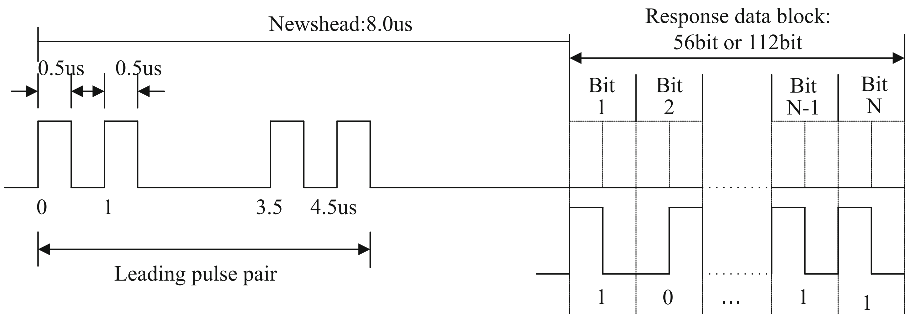


Fig. 3. Mode S response signal.

The inquiry signal of Mode S is added with the address of the inquiry plane [5], which can be used for roll call inquiry. The signal format of Mode S downlink response data link is shown in Fig. 3.

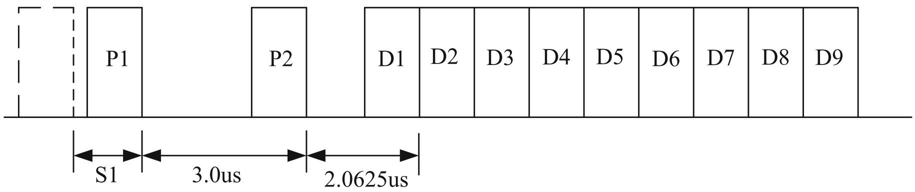


Fig. 4. Mode 5 level 1 signal diagram.

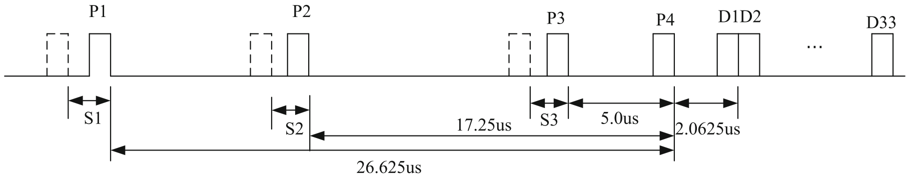


Fig. 5. Mode 5 level 2 signal diagram.

The mode signal consists of four leading pulses and response data pulse blocks. The pulse width of the leading pulse is 0.5 us, and pulse amplitude modulation is adopted. At the same time the duration of the response data pulse group is 56 us or 112 us (containing 56 or 112 bit data pulse), which uses binary pulse position modulation, pulse occurs in the first half segment represents '1', a pulse occurs in the second half for '0', and the last 24 bits are the combined parity and address field. The interval between the first preamble pulse and the following three preamble pulses of the Mode S response is 1 us, 3.5 us, and 4.5 us, respectively.

The waveform of level 1 reply signal of mode5 is shown in Fig. 4. It is composed of two synchronous header pulses (P1, P2) and nine data pulses (D1–D9), and the modulation mode is MSK. The response signal format of level 2 type is shown in Fig. 5. Due to space limitation, more information about mode 5 can be found in literatures [6,7].

3 Basic Principle of Algorithm

According to the characteristics of the measured IFF signal, it is found that the amplitude of the received IFF signal fluctuates greatly and the signal-to-noise ratio is low. Because the frequency band of IFF system is 1030 MHz and 1090 MHz, it is not easy to be interfered by other types of radar signals, so the noise is basically composed of Gaussian noise and system noise. Combined with the actual situation, this paper assumes that the system only contains Gaussian white noise, the intermediate frequency of the signal is 70 MHz, the sampling rate is 240 MHz, and the SNR of the simulation data set is between −10 dB and 0 dB. Considering the length of the synchronous pulse group and the actual length of the received signal, it is found that the length of the network input signal is about 30 us, which can ensure the recognition of all mode signals. Therefore, this paper sets the length of the data set as 8000 points.

Single signal is used as sample in training set and verification set. In order to meet the needs of real-time test, a long-time complex signal sample needs to be designed. This part will be explained in detail in the algorithm section. When simulating the actual test environment, only using these five mode signals as the training set and simple capsule network structure is not enough to meet the actual needs, so we need to improve the existing samples.

3.1 Dual Task Capsule Network Structure of Attention Mechanism (DM-CapsNet)

Assuming that the neural network can only intercept longer signals for classification and recognition, the input of the network at this time is 35 us, and it happens that there is a shorter mode signal (such as Mode 4, about 4 us) followed by a longer one. Mode signal (such as Mode S, at least 64 us), because the Mode S signal with more features accounts for too much of the input of the entire network, the first signal is directly ignored because of too few features, leading to missed alarms. On the other hand, when the network input is too short, it is impossible to extract all the characteristic information of the long signal. In order to avoid these situations, this paper designs a new dual-task constrained capsule neural network (DM-CapsNet).

Overall Framework
In this part, we will describe how DM-CapsNet identifies different patterns of continuous IFF signals.

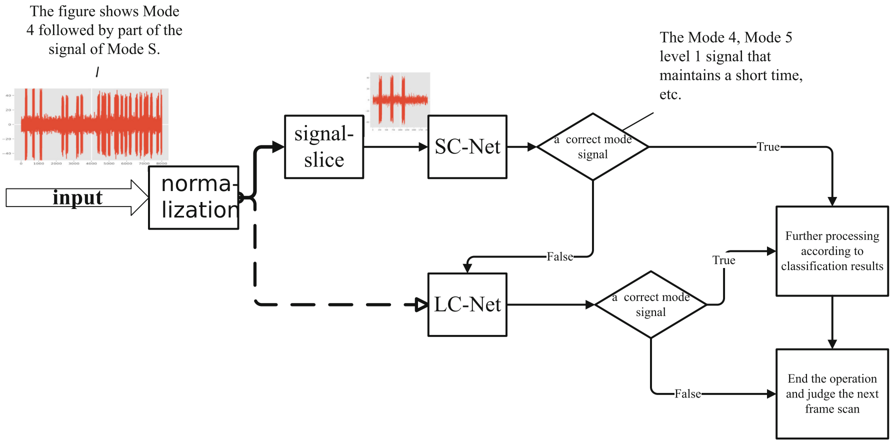


Fig. 6. Overall framework of the system.

The overall approach of the method is shown in Fig. 6. Starting from any position of IFF signal, we first intercept a longer signal, then intercept a shorter signal from the starting position of the longer signal, and finally use the signal characteristics of the intercepted short signal to dynamically normalize the amplitude of the input signal. Then, this short signal is applied to the attention capsule neural network of the simplified parameters of the short signal (SC-Net) for the first classification, and if the short characteristic signal is detected, the detection result is used for further processing.

When it is detected that the signal does not belong to short feature signal, the normalized input will be reclassified using a complete long capsule network (LC-net) suitable for long signal. When it is detected that the signal belongs to long feature signal, it will use the detection results and use different signal processing methods for further interpretation and other operations. Finally, the signal recognition process of this frame ends, and the scanning frame moves forward to continue the process.

Network Model

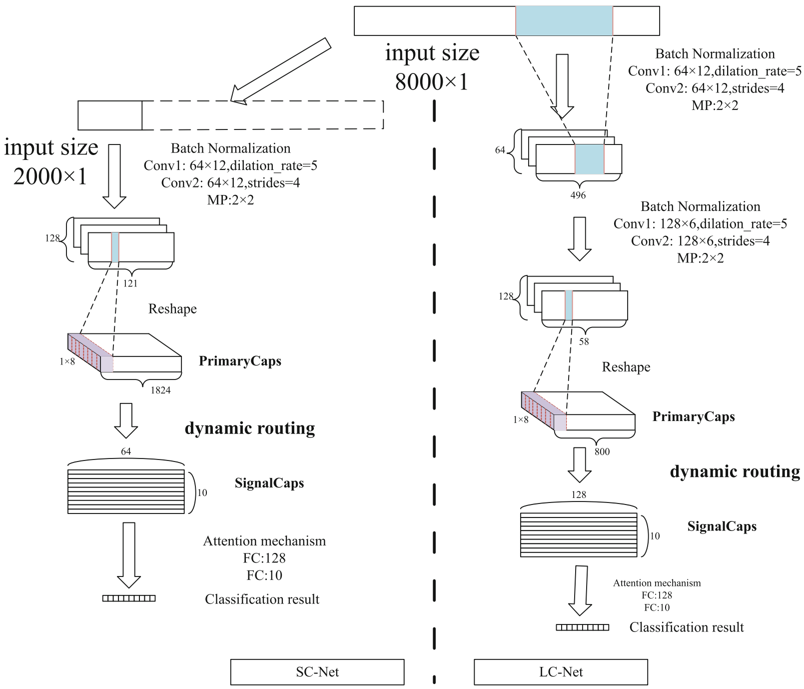


Fig. 7. DM-CapsNet network model.

The overall network structure model is shown in Fig. 7. The feature extraction network is divided into two parts: long signal (LC-Net) and short signal (SC-Net). The reason why it is not set as a convolutional network with shared weights, so as to improve the speed of training and the generalization ability of the network, is because the short signal itself can extract fewer features, if the same convolutional layer is used as the long signal Compared with the convolutional layer that uses non-weight sharing but fewer network parameters, although this can extract more detailed feature maps, the amount of calculation required for the convolution operation is larger, so the convolution with shared weights is used The network to extract features here does more harm than good. At the same time, because relatively short signals require higher accuracy of the extracted features when recognizing, the capsule input layer of SC-Net has more capsule neurons than the input layer of LC-Net.

Dynamic Routing Algorithm

Each capsule neuron is actually a set of directional scalar neurons, and its output is a multidimensional vector, so it can be used to represent some attribute information of entities [8–10]. Capsule neural network is divided into 4 steps, namely matrix transformation, input weighting, weighted summation and nonlinear transformation. The process is formulated as:

$$\hat{u}_{ji} = W_{ij} u_i \tag{1}$$

The first step is as shown in Eq. (1). The purpose of this step is to transform the spatial relationship between the low-level features and the high-level features and other important relationships through the matrix.

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k (b_{ij})} \tag{2}$$

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \tag{3}$$

c_{ij} in Eq. (2) is calculated by the softmax function, which is the coupling coefficient determined by the dynamic routing process. And the input weight of the second step and the weighted sum of the third step are realized by Eq. (3).

$$v_j = \frac{||s_j||^2}{1 + ||s_j||^2} \frac{s_j}{||s_j||} \tag{4}$$

The fourth step is to perform a nonlinear transformation on A to obtain B. The activation function used is shown in the formula EC. The first part of the formula is to compress, and the second part is to unitize the output vector. Through the above steps, the length of the output vector is between 0 and 1, so as to determine the probability of having a certain feature.

3.2 Attention Mechanism

In the above network model, the convolutional layer is responsible for feature extraction of the signal, and the capsule layer is responsible for compressing the extracted features into multiple vectorized capsule neurons through the mapping relationship between the underlying features and the high-level features.

According to the processing method described in the literature [8], the extracted neurons directly obtain the probabilities of different categories through multiple fully connected layers. It is obvious that because different capsule neurons may contain different attribute characteristics of the signal, not all capsule neurons are helpful for the final classification and recognition, so a method is needed to weight different capsule neurons.

Because some words have a decisive influence on sentence semantics, and the concept of time step is consistent with that of capsule neuron, we can introduce the idea of attention mechanism into the field of signal processing based on capsule network, that is, to pay attention to the output of capsule neuron we are interested in.

The use of the attention mechanism in the capsule network is divided into 3 steps, as follows:

1) Denote the last capsule hidden states as $[h_1, ..., h_N]$, multiply it with each of the remaining capsule neurons, that is, the hidden states, to obtain the current capsule neuron's attention scores, denoted as $e^1 = \left[s_1^T h_1, ..., s_1^T h_N\right]$, finally use softmax to convert the scores into an attention distribution with a probability sum of 1. The formula is $\alpha^1 = softmax\left(e^1\right)$.
2) Use the probability distribution obtained in step 1 to sum all hidden states to get the attention output. The formula is $\alpha_1 = \sum \alpha_i^1 h_i$. Finally, the residual connection of the attention output and the last capsule will give $[\alpha_1; s_1]$.
3) Use the fully connected layer to merge again and output it as the attention module.

3.3 Enhanced Dataset

According to the previous introduction, there are only 5 types of IFF response signals, namely Mark X series, Mode 4, Mode S, Mode 5 level 1 and level 2. At the same time, because these signals are in the same channel and frequency, it is easy to cause false alarms when scanning similar features in real-time detection, and general data enhancement methods such as short-time Fourier transform STFT and discrete wavelet transform commonly used in speech signals DWT and Wigner-Ville distribution (WVD) are of little use here, so we can only consider the data enhancement method of negative sample enhancement from the data set itself.

Because the IFF signal has both long and short, and the length of the data set is fixed, when constructing the negative sample data set, the starting position of the different modes of signals is shielded into negative samples according to the length percentage of 20–80%, and the proportion of positive and negative samples should be in the range of 0.5 and 2.0. By adopting this method of generating negative samples with varying length ratios, the number of negative samples can be adjusted adaptively, and the ratio of positive and negative samples can be adjusted within an appropriate range. Finally, the new data set consists of the original data set and the supplementary negative sample data set, which contains 10 different types of samples.

3.4 Training Parameters and Similarity Calculation

The deep learning model in this paper uses a common index to evaluate and predict the signal recognition, that is, the mean square error (MSE) based on L2 loss, which can be expressed as:

$$MSE = \frac{1}{n}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2. \tag{5}$$

where $\hat{y}_i$ is the prediction result of sample i, and y_i is the ground truth of the corresponding response signal mode.

The reason why MSE is used here is that it is easy to calculate, and the gradient of loss increases with the increase of loss, while the gradient decreases when the loss approaches zero, which makes the result of MSE model more accurate at the end of training.

The network is trained by proper training samples, which can effectively classify the continuous overlapped signals of the same channel. The output of DM-CapsNet is a 10-dimensional vector. $f(x)$ is a threshold decision function, expressed as:

$$f(x) = \begin{cases} 1 \; if \;\; x \in [T, 1] \\ 0 \; if \;\; x \in [0, T) \end{cases} \tag{6}$$

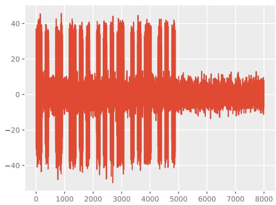
where T is the threshold set manually.

4 Experimental Results and Discussion

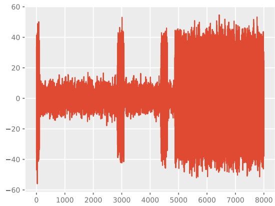
4.1 Experimental Conditions and a Simulation Data Set for Testing

All experiments in this chapter are based on a 16 GB RAM workstation, equipped with 3700X CPU and NVIDIA GeForce GTX 2080Ti GPU. And we use the Keras framework with Tensorflow as the back-end for network construction and performance optimization experiments.

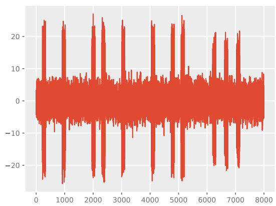
When using other methods to compare with the method proposed in this paper, the test data set used is different from the data set with negative sample data enhancement used in the training process. The samples of test data set usually contain multiple signals of different modes, and the amplitude and position are floating in a certain range. The data set used for training can ensure that the input of each signal has a corresponding pattern, which is very simple to recognize, while the data set used for testing can only scan the entire sample point by point or pulse. At this time, the network input signal will have different types of defects, as shown in Fig. 8(a), (b); or the input contains multiple mode signals, as shown in Fig. 8(c); or the signal amplitude may be missed due to too low signal amplitude, as shown in Fig. 8(d). It can be seen that the identification of the actual test signal is very difficult.



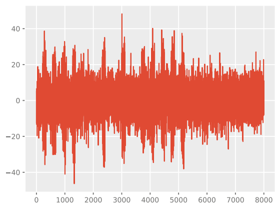
(a) Partial Mode s signal.



(b) Partial Mode 5 signal.



(c) Mark series signal and Mode 4 signal.



(d) Low amplitude signal.

Fig. 8. Possible wrong input test samples.

The related parameters of the simulated test data set are shown in the table below Table 1. Each sample contains 3–8 IFF signals with random mode type. The position of each IFF signal in the sample is random, and the amplitude fluctuates within a certain range. Because the duration of the IFF signal of different modes is different, and the system only collects the signal part when the signal is actually received, there is no guarantee that each sample has the same length. Therefore, the signal length should be dynamically adjusted according to the mode type and the number of signals contained in each sample, so the signal length of the data set is not fixed. A sample in the test data set is shown in Fig. 9.

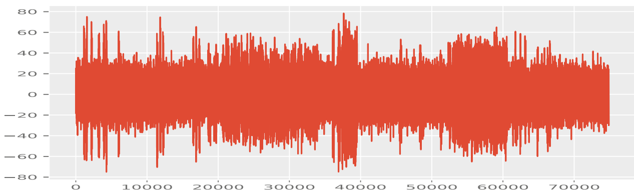


Fig. 9. Sample in test set.

This sample contains a total of 7 signals, starting from the starting point on the left, there are one Mark X series IFF signal, two Mode 4 IFF signals, one Mode S IFF signal, one Mode 5 level1 IFF signal, one Mode 5 level2 IFF signal, and one Mark X series IFF signal.

Table 1. Parameters of test data set.

Signal parameter	Specific values
Number of signal types	5
Center frequency	70 MHz
Sampling rate	240 MHz
Amplitude range of each IFF signal	25–60
SNR of a sample	-9 dB to 9 dB
Points of each sample	$\geqslant$50000
Number of signals per sample	3–8
Sample size	10000

4.2 Experimental Result

In the process of continuous detection of IFF signal, we need to pay attention to the changes of accuracy, instantaneous false alarm rate and missing alarm rate under different SNR. The global false alarm rate is 10^{-4} order of magnitude in most cases because of more non signal detection, which has no practical reference significance. The instantaneous false alarm rate only compares the detection performance of the method at the pulse position, and the instantaneous false alarm rate can make the comparison of results more intuitive. If there is no special explanation in this paper, the false alarm rate refers to the instantaneous false alarm rate.

Compare the loading and testing time of DM-CapsNet, multilayer convolutional neural network (CNN) and traditional detection method sliding window method (SW), as shown in Table 2. Although the sliding window method (SW) does not use the machine learning method, it needs multi-step correlation calculation because of the introduction of mode 5 signal. At the same time, compared with the neural network method, the sliding window method can only use CPU operation, and can not use GPU acceleration technology. It can be seen that the test time of the improved sliding window method is close to that of the multilayer convolution network. Because DM-CapsNet contains LC-Net and SC-Net, the time of DM-CapsNet is twice as long as the former two, but it also shows that the network has the possibility of further improvement.

The accuracy can intuitively understand the classification effect of the network on the test data set or continuous signal, and the recognition result is shown in Fig. 10. It can be seen that because the training data set uses low SNR signals as training samples, the recognition effect of neural network method is much

Table 2. Related parameters of the test data set.

	SW	CNN	DM-CapsNet
Load	–	0.82 s	1.35 s
Test	1481 s	1739 s	3517 s

better than sliding window method for low SNR test data set. However, with the improvement of SNR, the recognition accuracy of neural network method decreases, and the recognition accuracy of sliding window method increases gradually because of the higher requirement of SNR. At the same time, DM-CapsNet has the best recognition accuracy in the range of −9 and 9 dB.

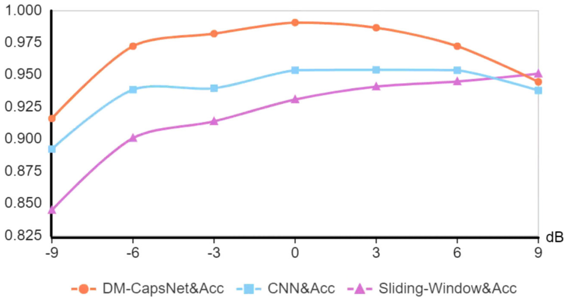
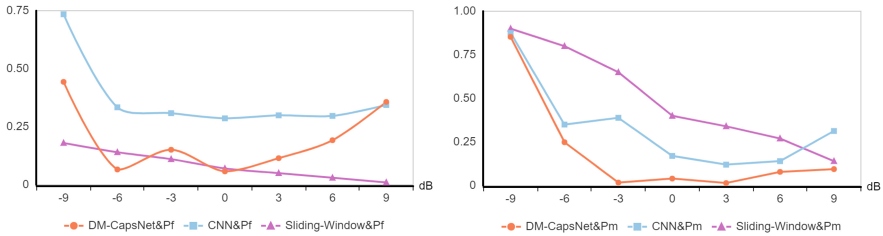


Fig. 10. Accuracy comparison results.

By comparing the results of false alarm rate and missing alarm rate of the three methods in Fig. 11, it can be seen that the false alarm rate of the multi-layer convolution neural network is high, and it can be found that the network has weak comprehensive ability to the characteristics of the signal by synthesizing the recognition accuracy of the method. Although the false alarm rate is low, the false alarm rate is relatively high. On the whole, false alarm rate and missed alarm rate are positively correlated with SNR, which shows that the method is most sensitive to the signal-to-noise ratio and has poor robustness. The average false alarm rate of the proposed dual task attention capsule neural network is less than 15%, and the false alarm rate is the lowest in the range of −9 dB to 9 dB. At the same time, the false alarm rate is less than 5% when it is more than 3 dB. To sum up, the overall recognition effect of the network is the best.

It can be seen from the analysis of the experimental results that: 1) Because the threshold value of sliding window method is fixed, when the signal amplitude is low and the signal-to-noise ratio is low, the situation of missing detection and wrong detection will occur, and the signals must be all in the frame before the method can be detected; 2) Although the multilayer convolution neural network also uses one-dimensional convolution layer to extract the signal features,



(a) False alarm rate comparison. (b) Comparison of Missing Alarm Rate.

Fig. 11. Compare results

because the feature processing ability is not enough, each time in the pulse detection will output a large detection probability, so the false alarm rate will be very high, and when the signal amplitude is too low, it can not be detected, so the detection ability has defects; 3) The method proposed in this paper can use convolution neural network to extract features, and then use capsule network and attention mechanism to further strengthen and weight the features, so that the method always has good recognition effect when the SNR is within a certain range.

5 Conclusion

In this paper, a recognition and classification method of continuous IFF signal based on attention mechanism and capsule neural network is proposed. The capsule network is creatively applied to the advanced feature extraction of IFF signal, and the effect is much better than that of traditional convolution neural network, The network structure of dual task output can deal with the error recognition caused by multiple signals input at the same time, and further improve the recognition performance. At the same time, negative samples are added to the signal training set to enhance the data, which makes the recognition ability of the network better, and the false alarm rate in the test samples decreases significantly, which is of great help to the recognition of continuous signals. Finally, through continuous IFF analog signal experiment, the test results show that DM-CapsNet has good and stable performance in the test set. In the future, we can try to optimize the network model parameters or new routing algorithm to improve the classification and recognition effect.

References

1. Comon, P.: Independent component analysis, a new concept? Signal Process. **36**(3), 287–314 (1994)
2. Hyvarinen, A.: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Networks **10**(3), 626–634 (1999)

3. Tang, B., Cheng, S.Y., Zhang, H.: Separation of Garbled secondary surveillance radar signal based on multichannel array processing. Telecommun. Eng. **54** (2014)
4. Qian, T., Mao, Y.L., Rong, Z.: A study on the analysis and identification of IFF signals. Radar ECM (2008)
5. Zhao, B.: Research on method of interrogator code assignment of mode S secondary surveillance radar based on distance matrix. Radio Eng. (2017)
6. Duan, H., Cheng, Y., Seen, B., He, K., Bai, G.: LFM interference cancellation algorithm based on MDPT-WC for mark XIIA mode 5. In: 2020 IEEE 20th International Conference on Communication Technology (ICCT), pp. 246–252 (2020). https://doi.org/10.1109/ICCT50939.2020.9295892
7. Min, L.: A sorting method of mark XIIA mode 5 signals. Telecommun. Eng. (2019)
8. Sabour, S., Frosst, N., Hinton, G.E.: Dynamic routing between capsules (2017)
9. Hinton, G.E., Sabour, S., Frosst, N.: Matrix capsules with EM routing (2018)
10. Ma, X., Zhong, H., Li, Y., et al.: Forecasting transportation network speed using deep capsule networks with nested LSTM models. IEEE Trans. Intell. Transp. Syst. **PP**(99), 1–12 (2020)
11. Pang, T., Li, Y., Niu, Y., et al.: Periodic pulsed jamming detection method and its performance in wireless communication. J. Terahertz Sci. Electron. Inf. Technol. (2019)
12. Ke, W., Hao, C.: Research on radar main lobe false target jamming feature extraction based on time-frequency domain and fluctuation characteristics. J. Phys. Conf. Ser. **1871**(1) (2021)
13. Rosario, D., Martins, V., Mauricio, B., Edson, B.: Efficiency and scalability of multi-lane capsule networks (MLCN). J. Parallel Distrib. Comput. (2021)