Journal of Systems Engineering and Electronics
Vol. 34, No. 5, October 2023, pp.1182 – 1190

Radar emitter signal recognition method based on improved collaborative semi-supervised learning

JIN Tao and ZHANG Xindong*

College of Electronic Science and Engineering, Jilin University, Changchun 130012, China

Abstract: Rare labeled data are difficult to recognize by using conventional methods in the process of radar emitter recognition. To solve this problem, an optimized cooperative semi-supervised learning radar emitter recognition method based on a small amount of labeled data is developed. First, a small amount of labeled data are randomly sampled by using the bootstrap method, loss functions for three common deep learning networks are improved, the uniform distribution and cross-entropy function are combined to reduce the overconfidence of softmax classification. Subsequently, the dataset obtained after sampling is adopted to train three improved networks so as to build the initial model. In addition, the unlabeled data are preliminarily screened through dynamic time warping (DTW) and then input into the initial model trained previously for judgment. If the judgment results of two or more networks are consistent, the unlabeled data are labeled and put into the labeled data set. Lastly, the three network models are input into the labeled dataset for training, and the final model is built. As revealed by the simulation results, the semi-supervised learning method adopted in this paper is capable of exploiting a small amount of labeled data and basically achieving the accuracy of labeled data recognition.

Keywords: emitter signal identification, time series, bootstrap, semi supervised learning, cross entropy function, homogenization, dynamic time warping (DTW).

DOI: 10.23919/JSEE.2023.000126

1. Introduction

Radar radiation source identification [1−3] is an important component and key step in electronic warfare, and after completing signal interception and pulse parameter analysis, the signal needs to be further analyzed to obtain the radiation source model or even individual information. To be able to better the radar radiation source signal, many scholars introduced machine learning methods. Meng et al. proposed a radar radiation source identification method based on time-frequency image texture features [4]; Cao et al. proposed wavelet invariant moment based for radiation source identification [5]; Li et al. used stacked sparse self-encoder to extract features and identify the signal [6]; Kong et al. used convolutional neural network (CNN) for radar signal waveform identification [7]. All of the abover-mentioned have achieved good results.

However, the above methods work well on the premise that the radar radiation source variety is relatively single and limited, and as technology continues to advance, the modern battlefield electromagnetic environment is increasingly intricate and complex, with relatively few available samples of unknown radiation source signals, which can lead to overfitting of machine learning models. In recent years, semi-supervised learning methods [8−12] have gradually been emphasized by researchers. This approach can use the information in a small number of labeled samples to discover patterns from a large number of unlabeled samples for classification recognition.

In this context, this paper proposes a semi-supervised classification algorithm that can act directly on a small amount of labeled and a large amount of unlabeled radiation source magnitude sequence data. The risk of classification being too absolute is reduced by adding homogenization to the loss function, which effectively improves the recognition accuracy of semi-supervised learning. Further optimization of the unlabeled dataset by dynamic time warping (DTW) technique [13−16] greatly improves the recognition speed while ensuring the recognition accuracy, which can well balance the recognition accuracy and recognition speed.

In this paper, the first three sections mainly introduce the theoretical foundation. Section 1 introduces the concept and types of semi-supervised learning, Section 2 introduces the concept and improvement methods of the loss function, and Section 3 introduces the DTW technique. Section 4 details the semi-supervised networks used in this paper, including the basic structure of the networks and the parameters used. Section 5 focuses on designing simulation experiments for the semi-supervised networks and analyzing the experimental results.

Manuscript received November 16, 2021.
*Corresponding author.

2. Semi-supervised learning

2.1 Introduction of semi-supervised learning

Semi-supervised learning is a machine learning approach between supervised and unsupervised learning [17,18], which can learn using both labeled and unlabeled data. In many practical problems, only a small amount of labeled data is available because it is costly and impractical to label all data. For example, in daily electronic countermeasures reconnaissance, labeling all intercepted radiation source signals is basically impossible to achieve. While, on the contrary, unlabeled radiation source signals are easily available. To make better use of unlabeled data to achieve better recognition results, semi-supervised learning techniques have emerged.

2.2 Classification of semi-supervised learning

The main semi-supervised learning approaches are as follows.

(i) Self-training: first train a classifier using a small amount of sample data, then use the classifier to predict unlabeled data, select samples with high confidence to be added to the labeled data, and finally retrain the classifier using the updated labeled data set, repeating the above two procedures until the target is reached.

(ii) Generative model [19]: an algorithm implemented based on the clustering assumption, which first clusters the unlabeled data using a clustering algorithm (usually using a Gaussian mixture model), then estimates the component parameters in the Gaussian mixture model using the expectation maximum algorithm, and finally determines the class of each component using a small amount of labeled data.

(iii) Low-density partitioning algorithms [20]: combining support vector machines with clustering assumptions so that the classification boundary bypasses data-dense regions.

(iv) Graph-based semi-supervised algorithm [21−24]: it is assumed that all data (both labeled and unlabeled) and the interrelationships between all data can be represented in the form of an undirected graph, where the nodes of the graph represent data sample points and the similarity relationships between samples are represented as edges. The graph-based semi-supervised algorithm is to make the results of the labeled sample points conform to the flow-type assumption.

(v) Divergence-based semi-supervised algorithm [25,26]: two different classifiers are trained on x_1 and x_2 (two views of the labeled data) using the labeled data, and then the unlabeled samples are predicted by using the two classifiers, and the unlabeled samples with higher confidence in the classification results of the classifiers and their prediction labels are put into the labeled data set for the next training, and so on iteratively.

Considering the advantages and disadvantages of the five semi-supervised learning approaches, this paper selects the divergence-based semi-supervised algorithm as the radar radiation source sequence data recognition, and improves on the original algorithm to make it better classify and recognize the radiation source sequence data.

3. Loss function and its improvement

3.1 Loss function concept and selection

The loss function is used to evaluate the degree of difference between the true value and the predicted value of a machine learning model, and generally speaking, the better the model performance, the better the loss function. In general, loss functions can be divided into two types: empirical risk loss function and structural risk loss function, the difference is that the structural risk loss function is obtained by adding a regular term to the empirical risk loss function.

For a multi-classification problem like radar radiation source identification, the cross-entropy function is usually used as the loss function, and the cross-entropy function is essentially a log-likelihood function. The standard form of the cross-entropy function is

$$C = -\frac{1}{n}\sum_{x=1}^{n}[y\ln a + (1-y)\ln(1-a)] \tag{1}$$

where x denotes the input radiation source sample, y denotes the label of the radiation source, a denotes the predicted output, and n denotes the total number of radiation source data samples. For a multi-classification problem like radiation source identification, the loss function takes the form of

$$\text{loss} = -\frac{1}{n}\sum_{i=1}^{n} y_i \ln a_i \tag{2}$$

where i denotes the input different radiation source samples, y_i denotes the labels of different radiation sources, and a_i denotes the predicted different outputs.

3.2 Customized complex loss function

There are also some problems when the cross-entropy function solves the multi-classification problem: when the results of classification are output using softmax, the classification results are too absolute, and the results are also non-zero or 1 when the noise is classified, which may lead to model overfitting, and this risk is further extended especially when applied to the semi-supervised classification in this paper. In this paper, the loss function is modified to circumvent this risk.

First, a uniform distribution is constructed, the cross-entropy of the predicted values with this uniform distribution is calculated, and then this cross-entropy is combined with the original cross-entropy function. This

allows the cross-entropy function to fit the uniform distribution appropriately while fitting the one-hot distribution, thus mitigating the classification of softmax for results that are too absolute and reducing the risk of overfitting.

4. DTW

In the process of time series data analysis, to determine the relationship between two sets of time series data, DTW is often used to calculate the similarity between two-time series. Any data that can be converted into linear series, including time series converted from graphics, video, and audio, can be calculated and analyzed by DTW.

Therefore, this paper uses the DTW method to judge the similarity between unlabeled radiation source amplitude sequence data and a small amount of labeled amplitude sequence data, to filter out the unlabeled data with higher similarity to the labeled data, thus achieving the effect of reducing the data volume while ensuring the accuracy. The following conditions must be met to achieve DTW.

(i) Each index of the labeled radiation source magnitude sequence must match one or more indices of the other unlabeled magnitude, and vice versa.

(ii) The first index in the labeled radiation source amplitude sequence must match the first index in the unlabeled amplitude sequence (which may not be a unique match).

(iii) The last index of the labeled radiation source amplitude sequence must match the last index of the unlabeled amplitude sequence (which may not be the only match).

(iv) The index mapping from the labeled radial source amplitude sequence to the other unlabeled amplitude sequences must be monotonically increasing and vice versa.

5. Improved collaborative semi-supervised network model

5.1 Basic structure of the network

In this paper, based on the original collaborative semi-supervised algorithm, we optimize the cross-entropy function by adding uniform distribution to reduce the error caused by too absolute classification, and use the DTW method to filter the unsupervised data set to balance the recognition accuracy and the number of unsupervised data, the main structure of the model is shown in Fig. 1.

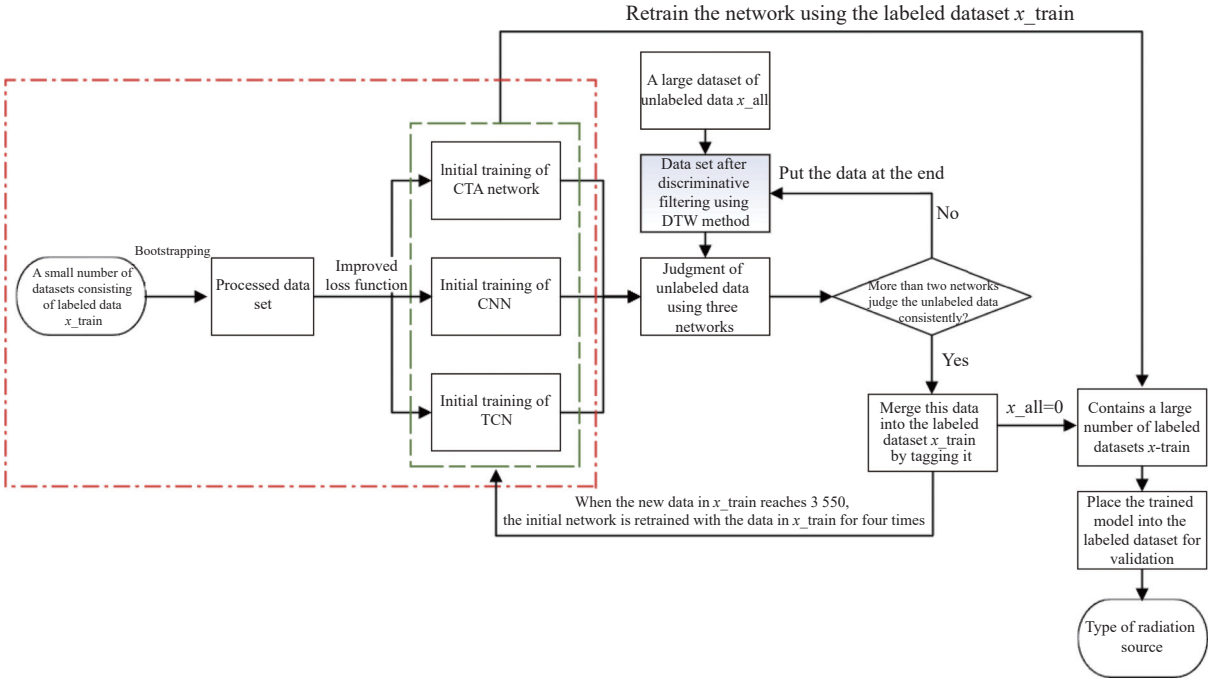

Fig. 1 Semi-supervised framework used in this paper

The structure of this paper consists of two main parts: (i) initial training of three common network models with a small number of labeled datasets; (ii) updating the labeled database with the initially trained network models and unlabeled data and further training the network models.

The first part is the initial training of the network. Firstly, the labeled dataset is sampled three times by using the bootstrap method to obtain three different validation and test sets. The three common classification net-

works are: one-dimensional (1D) CNN [27,28] (1DCNN), temporal convolutional network [29,30] (TCN), one-dimensional convolutional network + temporal convolutional network + attention layer (CTA). The loss functions of the three networks are added to the uniform distribution, and three different training and testing sets obtained by the bootstrap method are used to initially train the three classification networks.

The second part is to optimize the networks by using the three initial training models and the unlabeled data. First, the unlabeled data are divided into groups of every 256, and then the distances to the labeled data are calculated by grouping them using the DTW method, and the closer ones are selected to construct a new unlabeled data set. Then, the data of the new unlabeled dataset are taken out one by one and predicted by one of the three initially trained network models, and if two or more network models predict the unlabeled data consistently, the data are labeled with the predicted labels into the labeled dataset, otherwise, the data are put into the end of the data set. When the data added to the labeled dataset reaches 3 550, the network is retrained by using the labeled dataset and the trained network is used to continue predicting the unlabeled data until the unlabeled dataset is empty.

5.2 Parameter setting

The three basic classification networks used in this paper are shown in Fig. 2, and the basic parameters of the networks are as follows.

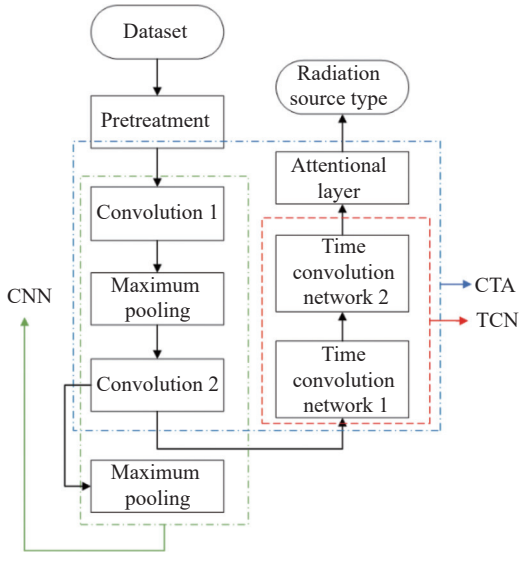

Fig. 2 Three classification networks used in this paper

CNN: The number of convolutional kernels of Convolutional layer 1 is 32, the time-domain window length of convolutional kernels is 5, and the window size of the maximum pooling layer is 2; the number of convolutional kernels of convolutional layer 2 is 16, and the time-domain window length of convolutional kernels is 5.

TCN: The time domain window length of the convolutional kernel of the temporal convolutional network 1 is 5 and the number of convolutional kernels is 32. The time domain window length of the convolutional kernel of the temporal convolutional network 2 is 5 and the number of convolutional kernels is 16. The expansion ratio of the expansion convolution is set to 2.

CTA: The CTA network is obtained by combining a CNN network with a maximum pooling layer removed and a TCN network with parameters using the two network parameters mentioned above.

The fully connected layer coefficients of the three classification networks are set to 8; the activation functions of the three classification networks all use the Leaky ReLU activation function.

5.3 Network summary

In order to be able to describe the semi-supervised classification algorithm used in this paper more clearly and intuitively, the main contents of the algorithm are summarized in Table 1.

Table 1 Main contents of the algorithm

Attribute	Content
Algorithm name	Improved cooperative semi-supervised learning classification algorithm
Classification object	Eight kinds of radar radiation source sequence data
Algorithm objectives	Use a small amount of labeled data and a large amount of unlabeled data to achieve the effect of supervised recognition
Major innovations	Improving the activation function to improve the recognition accuracy, while reducing the network time consumption without reducing the classification accuracy by DTW technique
Classification networks	CNN, TCN, CTA networks
Algorithm flow	The similarity between labeled and unlabeled data is calculated in groups for filtering, and then the initially trained classification network model is used to judge each unlabeled data

6. Simulations and result analysis

6.1 Simulation parameter setting

There are eight kinds of radar radiation source signals to be classified and identified in this paper, which are frequency modulated continuous wave (FMCW), multi-phase code (Frank, P1, P2, P3, P4), binary phase shift keying (BPSK), and Costas. the carrier frequency range is 1−1.2 kHz, except Costas, the sampling frequency of the other seven signals is 7 kHz. The parameters of each signal are shown in Table 2. According to the actual signal

characteristics, Matlab simulation is used to generate the corresponding labeled and unlabeled data with signal to noise ratios ranging from −20 dB to 10 dB, with an interval of 2 dB and a length of 1 024 for each sample. Experimental test platform uses: Intel(R) Core (TM) i7-8550U, NVIDIA GeForce RTX 1050. The neural network model is built on the TensorFlow platform using Keras 2.3.0, PyCharm 2018.3.6 x64, and Python 3.6 compiler.

Table 2 Main parameters of signal

Signal	Main parameter	Value
BPSK	Barker code digits	{7,11,13}
Multi-phase code	Number of code bits	{36,64}
FMCW	Sampling frequency/kHz	{15,17}
	Modulation period/ms	{50,25,35}
	Modulation bandwidth/kHz	{0.25,0.35,0.5}
Costas	Frequency sequences/kHz	{[4 7 1 6 5 2 3], [2 6 3 8 7 5 1]}

6.2 Simulation experiments and analysis

Experiment 1 When using the cross-entropy loss function in combination with softmax, it does have a better classification effect for the supervised radiation source classification and identification network. However, the problem that softmax output results are more absolute is further amplified in the semi-supervised network. The addition of homogenization in the loss function will reduce this absolute effect. The comparison of loss, the accuracy of the semi-supervised classification network with homogenization and the semi-supervised network with initial cross-entropy loss function under the same data set (supervised data set is 128, unsupervised data number is 51 200) are shown in Table 3 and Table 4, where Accuracy 1 and Loss 1 represent the initial classification results for a small amount of supervised data, and Accuracy 2 and Loss 2 represent the classification results after the semi-supervised network.

Table 3 Loss function without adding uniform distribution

Network	Loss 1	Accuracy 1	Loss 2	Accuracy 2
CTA	1.766 5	0.731 6	0.691 5	0.848 0
CNN	1.258 6	0.742 0	1.298 7	0.827 4
TCN	1.494 1	0.784 6	1.857 4	0.850 1

Table 4 Loss function added to uniform distribution

Network	Loss 1	Accuracy 1	Loss 2	Accuracy 2
CTA	1.065 7	0.804 6	0.661 5	0.933 9
CNN	1.047 7	0.826 1	0.686 6	0.927 1
TCN	0.949 3	0.857 8	0.661 6	0.934 4

As can be seen in Table 2 and Table 3, the cross-entropy loss function without uniform distribution in it is less effective when the initial model obtained by training with a small amount of labeled data when there is too little labeled data. The problem of too absolute classification of the cross-entropy function is further amplified when using semi-supervised networks for classification prediction of unlabeled data, resulting in excessive loss and lower accuracy. The table also shows that adding homogenization to the cross-entropy function can more obviously improve the classification recognition effect of the semi-supervised network, which can improve the recognition accuracy by about 10%.

Experiment 2 To investigate the effect of the number of labeled data on the semi-supervised network, 64, 128, and 256 labeled data are selected for training in the same network and the same number of unlabeled data, respectively, and the accuracy of the resulting model under the same validation set is shown in Table 5, where the odd rows indicate the initial classification results for a small amount of supervised data, and the even rows indicate the classification results after the semi-supervised network.

Table 5 Identification results of different labeled data under the same conditions

Number of labeled data	CTA	CNN	TCN
64	0.653 2	0.752 0	0.752 3
	0.835 6	0.832 3	0.836 1
128	0.804 6	0.826 1	0.857 8
	0.933 9	0.927 1	0.934 4
256	0.852 7	0.884 6	0.903 0
	0.939 7	0.934 5	0.940 3

It is obvious from the table that when the number of labeled samples is 64, the number of labeled samples is too small and the initial model obtained after the initial training of the network is difficult to meet the requirements of the semi-supervised network, so the training results are poor even with a large amount of unlabeled data. When the number of labeled data is 256, although the initial and final training results are relatively good, it can be seen from the table that for some easier to train networks such as the TCN, these initial 256 labeled data can basically meet certain requirements and do not require further optimization of the semi-supervised network. Therefore, 128 labeled data are selected as the labeled data set in this paper.

To verify the effect of unlabeled data on the recogni-

tion effect, labeled data are fixed at 128, 25 600, 51 200, 76 800, and 102 400. Unlabeled data are selected to be put into the semi-supervised network used in this paper for training, and the accuracy of the resulting model under the same validation set is shown in Table 6.

Table 6 Recognition effects of three networks with different amounts of unlabeced data

Number of labeled data	CTA	CNN	TCN
25 600	0.924 7	0.923 0	0.932 0
51 200	0.933 9	0.927 1	0.934 4
76 800	0.936 4	0.932 5	0.936 2
102 400	0.937 9	0.931 5	0.936 8

As can be seen from Table 5, when the unlabeled data grows from 25 600 to 51 200, the recognition accuracy can be improved more obviously for the CTA network which has a more complex model. For the TCN and CNN which have a simpler structure and are easier to train, the accuracy will be improved to some extent, but the improvement is not obvious enough.

When the unlabeled data grows from 51 200 to 76 800, the overall accuracy improvement of the three networks is not as big as that when the unlabeled data grows from 25 600 to 51 200. When the unlabeled data set grows from 76 800 to 102 400, the increase of unlabeled data does not significantly increase the recognition accuracy, and even some networks that are easier to train, such as the TCN, may cause the accuracy improvement due to overfitting. The accuracy decreases due to problems such as overfitting. This shows that there is a limit to the improvement of the network by unlabeled data, and the accuracy cannot be improved by continuously increasing the unlabeled data.

The training of unsupervised data is slow and time-consuming, so how to balance the amount of unsupervised data and the recognition accuracy is the main concern of this paper. In this paper, a DTW method is used to initially filter the unlabeled data to reduce the training time without decreasing the recognition accuracy.

Experiment 3 To investigate the role of DTW method in this paper for balancing the number of unlabeled data and recognition accuracy, the number of labeled data is fixed to 128, and the initial unlabeled data are taken to be 51 200 and 76 800 respectively.

The recognition accuracy of the three classification networks with 51 200 unlabeled data is shown in Table 7.

Table 7 Results of 51 200 data filtered by DTW

Network	Initial	80%	70%	60%	50%
CTA	0.933 9	0.934 6	0.932 9	0.932 5	0.928 9
CNN	0.927 1	0.931 2	0.927 0	0.928 8	0.925 0
TCN	0.934 4	0.936 1	0.933 6	0.936 1	0.933 8

From Table 7, it can be seen that the recognition accuracy does not decrease when the selection ratio is 80%, but increases. When the selection ratio continues to be reduced to 60%−70%, the recognition accuracy does not decrease due to the reduction of the data set, although 30%−40% of the data is removed at this time. When the selection ratio decreases to 50%, due to the removal of more data, it will have a greater impact on the recognition accuracy of networks with relatively complex structures like CTA, and less impact on networks with simpler structures like CNN and TCN, which also verifies the conclusion of Experiment 2.

In order to verify the adaptability of this method, 76 800 unlabeled data are selected for the same experiments described above as shown in Table 8.

Table 8 Results of 76 800 data filtered by DTW method

Network	Initial	80%	70%	60%	50%
CTA	0.936 4	0.937 7	0.935 5	0.937 3	0.933 9
CNN	0.932 5	0.922 5	0.933 9	0.931 5	0.930 4
TCN	0.936 2	0.937 3	0.935 5	0.936 0	0.935 3

As can be seen from Table 8, when the total amount of unlabeled data is 76 800, the conclusions obtained are basically the same as those obtained in Table 7, the only difference is that when the selected proportion is 50% and 80%, the decrease and increase in accuracy in Table 7 is more obvious than that in Table 8, which is due to the larger initial data set in Table 8. From the conclusion of Experiment 2, it is clear that the more the amount of unlabeled data is, the lower the gain it receives from data growth. The results in Table 7 and Table 8 illustrate that using the DTW algorithm, the unlabeled data set into sex screening and selecting the appropriate proportion of data that are closer to the labeled data set can be a good balance between unlabeled data and recognition accuracy.

Experiment 4 To investigate the advantages of the improved collaborative semi-supervised method over the common semi-supervised learning methods, two common semi-supervised learning methods are selected for comparison with the improved collaborative semi-supervised learning method used in this paper. And, to better illustrate the effect of semi-supervised learning, a network model obtained by direct training with a small

amount of labeled data is added as a control group. A total of 128 labeled data and 51 200 unlabeled data are used for all four models, and the improved collaborative semi-supervised model used in this paper selects 60% of the unlabeled data using a time-warping algorithm. To facilitate comparison, the classification networks all use the CTA network model. The classification accuracy and loss of the four models are shown in Table 9, and the recognition accuracies of the four models at different signal to noise ratios are shown in Fig. 3. In Fig. 3 CTA1 denotes the network model with a small amount of labeled data trained directly, CTA2 denotes the self-training semi-supervised learning network model, CTA3 denotes the collaborative semi-supervised network model, and CTA4 denotes the improved collaborative semi-supervised model used in this paper.

Table 9 Loss and accuracy of the four models

Network	Loss	Accuracy
CTA1	1.766 5	0.731 6
CTA2	0.916 0	0.790 0
CTA3	0.767 3	0.846 4
CTA4	0.661 5	0.932 5

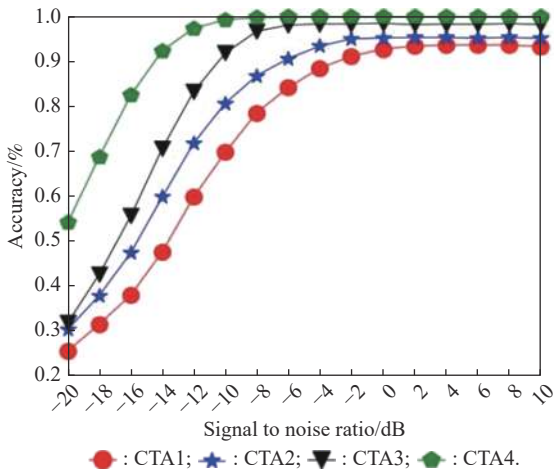

Fig. 3 Recognition ability of four models with different signal to noise ratios

As can be seen from Table 9, the improved collaborative semi-supervised network model used in this paper is much better than both the common self-training semi-supervised learning network model and the collaborative semi-supervised learning model in terms of both loss and accuracy. Comparing the commonly used collaborative semi-supervised network model and this paper's model, we can see that the network model used in this paper can improve the recognition accuracy by about 10% based on the collaborative semi-supervised learning network model. At the same time, the training speed is significantly improved compared with the common collaborative semi-supervised model because 40% of the unlabeled data are removed by the DTW technique.

As can be seen from Fig. 3, when the signal to noise ratio is low, the network model used in this paper still has good recognition effect, and can maintain more than 90% recognition accuracy even in the environment of −14 dB signal to noise ratio, which is sufficient to meet daily needs and far better than several other network models.

Experiment 5 Through the analysis of the above experimental results, 128 labeled data and 51 200 unlabeled data are selected, and 60% of these unlabeled data are used to train three common network models by using the DTW algorithm. The recognition ability of the network models obtained by direct training with a small amount of labeled data at different signal to noise ratios and the recognition ability of the network models obtained after semi-supervised learning at different signal to noise ratios are shown in Fig. 4. In Fig. 4, CTA, TCN, and CNN denote the models obtained from initial training with a small amount of labeled data, and CTA1, TCN1, and CNN1 denote the network models obtained after semi-supervised learning.

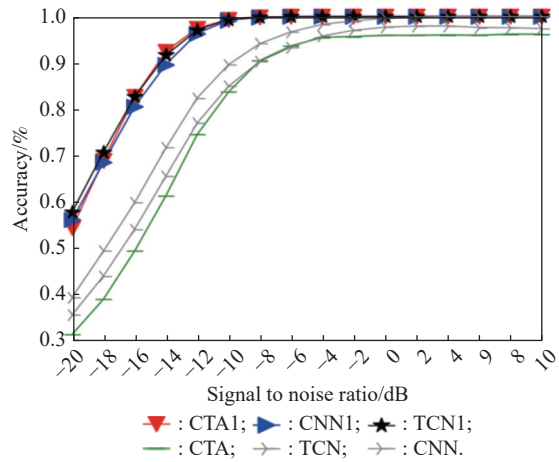

Fig. 4 Recognition ability of three models with different signal to noise ratios

As can be seen in Fig. 4, the network model obtained after semi-supervised learning has a significant improvement in the recognition of radiation source signals compared with the model obtained from the initial training. In particular, at low signal to noise ratios, the network model obtained after semi-supervised learning can improve the recognition accuracy by up to 20%.

To further investigate the recognition ability of the network model obtained after semi-supervised learning for

eight radiation source signals, the CTA network model trained in this paper is used to recognize eight radiation source signals at different signal to noise ratios. The results are shown in Fig. 5.

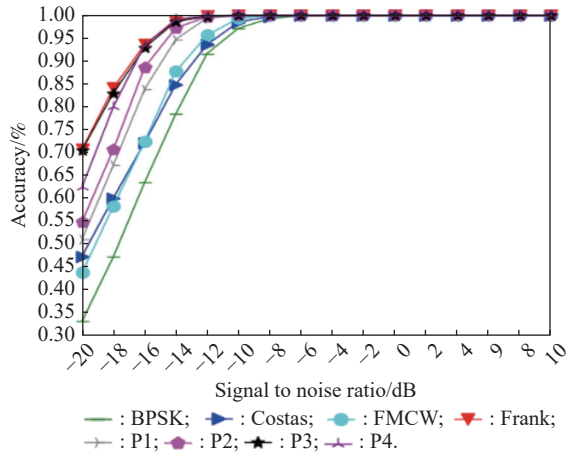

Fig. 5 Recognition ability of CTA model for eight signals at different signal to noise ratios

As can be seen from Fig. 5 where the CTA network models used in this paper, when the signal to noise ratio is greater than −8 dB, the classification accuracy of the three models for the eight signals reaches 100%. Even when the signal to noise ratio is at −14 dB, the classification accuracy of the three classification models for the eight signals remains above 80%, which can still meet the basic requirements for the classification and identification of radar radiation source signals under harsh conditions.

It can also be seen from Fig. 5 that, with the lower signal to noise ratio, the recognition effect of CTA network model for BPSK signals becomes the worst; for Frank and P3 signals, the recognition effect is the best, and when the signal to noise ratio is at −20 dB, it still has an accuracy rate of over 70%. For the other five signals, except for the FMCW signal and Costas signal, the accuracy of the other three signals can still maintain above 50% at −20 dB.

The overall confusion matrix for the eight signals is shown in Fig. 6. It can be seen from the figure that most of the signals can be accurately classified, and a clear diagonal line can be clearly seen. In the case of eight signals at −20−10 dB signal to noise ratio, except for BPSK signal, the classification accuracy of other signals can basically reach more than 90%, which can identify the radiation source signal more accurately. The model has the best recognition effect for Frank signal and P3 signal, and the recognition accuracy can reach about 96.5%, and the recognition effect for BPSK and Costas signal is the worst.

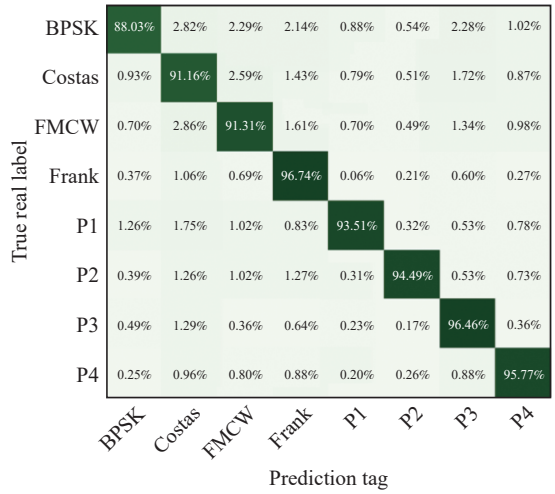

Fig. 6 Confusion matrix of eight signals

7. Conclusions

To solve the problem of poor recognition when there is less labeled data, an improved collaborative semi-supervised learning radar radiation source recognition method based on a small amount of labeled data is proposed in this paper. The loss function is improved on the original collaborative semi-supervised method to reduce the risk of too absolute classification. Meanwhile, the DTW technique is used to filter the unlabeled data to ensure the classification accuracy and reduce the classification time at the same time. The comparison through simulation experiments shows that, for eight kinds of radiation source signals, the semi-supervised network in this paper can classify and recognize radar radiation source signals more accurately, and basically can achieve the recognition effect of supervised network.

References

[1] WANG S Q, BAI J, HUANG X Y, et al. Analysis of radar emitter signal sorting and recognition model structure. Procedia Computer Science, 2019, 154: 500–503.

[2] MAN P, DING C B, REN W J, et al. A nonlinear fingerprint-level radar simulation modeling method for specific emitter identification. Electronics, 2021, 10(9): 1030.

[3] CAO R, CAO J W, MEI J P, et al. Radar emitter identification with bispectrum and hierarchical extreme learning machine. Multimedia Tools and Applications, 2019, 78(20): 28953–28970.

[4] MENG F J, TANG H, WANG Y Z, et al. Radar radiation source signal recognition based on texture characteristics of time-frequency images. Journal of Missile and Guidance, 2017, 37(3): 152–156. (in Chinese)

[5] CAO X H, WANG L X, SHU X Y. Radar emitter signal recognition based on wavelet moment invariants. Computer Engineering and Applications, 2020, 56(19): 269–272.

[6] LI Y B, GE J, LIN Y, et, al. Radar emitter signal recognition based on multi-scale wavelet entropy and feature weighting. Journal of Central South University, 2014, 21(11):

4254–4260.
[7] KONG S H, KIM M, HOANG L M. Automatic LPI radar waveform recognition using CNN. IEEE Access, 2018, 6: 4207–4219.
[8] VAN ENGELEN J E, HOOS H H. A survey on semi-supervised learning. Machine Learning, 2020, 109(2): 373–440.
[9] OUALI Y, HUDELOT C, TAMI M. An overview of deep semi-supervised learning. http://arxiv.org/abs/2006.05278.
[10] FENG Y T, WANG G L, LIU Z P, et al. An unknown radar emitter identification method based on semi-supervised and transfer learning. Algorithms, 2019, 12(12): 271.
[11] MOSTAJABI M, WANG C M, RANJAN D, et al. High-resolution radar dataset for semi-supervised learning of dynamic objects. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 100−101.
[12] YING F, XING W. Radar signal recognition based on modified semi-supervised SVM algorithm. Proc. of the 2nd Advanced Information Technology, Electronic and Automation Control Conference, 2017: 2336–2340.
[13] MULLER M. Information retrieval for music and motion. Berlin: Springer, 2007: 69−84.
[14] PERMANASARI Y, HARAHAP E H, ALI E P. Speech recognition using dynamic time warping (DTW). Journal of Physics: Conference Series, 2019, 1366(1): 012091.
[15] BERNDT D J, CLIFFORD J. Using dynamic time warping to find patterns in time series. Proc. of the 3rd International Conference on Knowledge Discovery and Data Mining, 1994: 359−370.
[16] SENIN P. Dynamic time warping algorithm review. http://api.semanticscholar.org/CorpusID:16629907.
[17] BARLOW H B. Unsupervised learning. Neural Computation, 1989, 1(3): 295–311.
[18] GHAHRAMANI Z. Unsupervised learning. Proc. of the Summer School on Machine Learning, 2003: 72–112.
[19] SHAHAM T R, DEKEL T, MICHAELI T. Singan: learning a generative model from a single natural image. Proc. of the IEEE/CVF International Conference on Computer Vision, 2019: 4570−4580.
[20] LU X J. Automatic segmentation and structural characterization of low density fibreboards. Image Analysis & Stereology, 2013, 32(1): 13–25.
[21] CHONG Y W, DING Y, YAN Q, et al. Graph-based semi-supervised learning: a review. Neurocomputing, 2020, 408: 216–230.
[22] DUNLOP M M, SLEPCEV D, STUART A M, et al. Large data and zero noise limits of graph-based semi-supervised learning algorithms. Applied and Computational Harmonic Analysis, 2020, 49(2): 655–697.
[23] SUBRAMANYA A, TALUKDAR P P. Synthesis lectures on artificial intelligence and machine learning. Cham: Springer, 2014.
[24] ZHUANG C, MA Q. Dual graph convolutional networks for graph-based semi-supervised classification. Proc. of the World Wide Web Conference, 2018: 499−508.
[25] ZHAI X L, ZHOU Z H, TIN C. Semi-supervised learning for ECG classification without patient-specific labeled data. Expert Systems with Applications, 2020, 158: 113411.
[26] SUN Y G, XU J Q, WU H, et al. Deep learning based semi-supervised control for vertical security of maglev vehicle with guaranteed bounded airgap. IEEE Trans. on Intelligent Transportation Systems, 2021, 22(7): 4431−4442.
[27] LU G, WANG Y B, YANG H Y, et al. One-dimensional convolutional neural networks for acoustic waste sorting. Journal of Cleaner Production, 2020, 271: 122393.
[28] HAN A G, BYRA M, HEBA E, et al. Noninvasive diagnosis of nonalcoholic fatty liver disease and quantification of liver fat with radiofrequency ultrasound data using one-dimensional convolutional neural networks. Radiology, 2020, 295(2): 342–350.
[29] FARHA Y A, GALL J. MS-TCN: multi-stage temporal convolutional network for action segmentation. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3575−3584.
[30] KOK C, JAHMUNAH V, OH S L, et al. Automated prediction of sepsis using temporal convolutional network. Computers in Biology and Medicine, 2020, 127: 103957.

Biographies

JIN Tao was born in 1997. He received his Master's degree from Jilin University in integrated circuit engineering. His research interests are deep learning and radar signal processing.
E-mail: jintao019@163.com

ZHANG Xindong was born in 1970. She received her Ph.D. degree from Jilin University in microelectronics and solid-electronics. She is a professor in Jilin University. Her research interest is radar signal processing.
E-mail: xindong@jlu.edu.cn