Journal of Nonlinear and Convex Analysis
Volume 26, Number 7, 2025, 2319–2333

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

INTERFERENCE SIGNAL INCREMENTAL RECOGNITION BASED ON DOUBLE AUXILIARY LEARNING

SHAOGANG DAI AND YINZHI WANG*

Abstract. In the realm of communication interference recognition, non-incremental approaches dominate the field. Aiming at incremental recognition, a double auxiliary learning (DAL) based interference incremental recognition algorithm (named DAL-NME) is proposed. A feature extracting network(FENet) based on DAL, which contains two feature extractors and four classifiers, is designed to obtain the interference signal feature vector. The feature extractor adopts the DenseNet structure to better extract signal features. The four classifiers are two main classifiers, two auxiliary classifiers for both new and old samples, which are used to better train the network and balance the learning of both new and old interference signals. The Nearest-Mean-of-Exemplars (NME) criterion is used to construct a memory set and classify the communication interference signals. Experimental results show that the DAL-NME method excels in incremental recognition of communication interference signals, having a 5% higher recognition accuracy compared to the commonly used WA algorithm at a JNR of 2dB in both random incremental stages and various scenarios.

1. Introduction

As the application of wireless devices continues to expand rapidly, the electromagnetic environment becomes increasingly intricate, leading to a rise in the variety of interference signals. Accurately identifying interference signals has attracted widespread attention of researchers to ensure the quality and reliability of communications. Traditional methods for interference identification primarily involve feature extraction and classifiers. However, feature extraction is heavily reliant on human expertise and may not comprehensively capture the entirety of interference signal features. Deep learning has powerful feature extraction capability. Therefore, more and more deep learning-based interference identification algorithms have been proposed [13, 15, 17], and their recognition performance is usually better than traditional methods. However, the majority of current deep learning-based interference signal recognition algorithms are non-incremental. Given that new interference signals may persistently emerge in real-world scenarios, retraining the networks of these non-incremental algorithms leads to repetitive and complex tasks. Incremental

---

2020 *Mathematics Subject Classification.* 94-05, 94A13.

Key words and phrases. interference recognition, incremental learning, feature extractor, auxiliary classifier, feature vector.

This work was supported by the National Natural Science Foundations of China under grant nos. U19B2016 and Key Lab of Data Storage and Transmission Technology of Zhejiang Province, Hangzhou Dianzi University.

*Corresponding author.

learning can identify newly emerging interference signals while retaining the recognition capability of old interference signals. Therefore, incremental recognition of interference signals holds significant practical value.

To address the issue of catastrophic forgetting of old knowledge in incremental learning, scholars in the field of image processing have conducted an amount of research from three aspects. The first aspect is the data, that data memory or data generation is used to alleviate the problem of forgetting old knowledge. Rebuffi et al. [16] learned the stored representative old knowledge while learning new tasks. Bang et al. [1] assessed the model's uncertainty regarding the samples through data augmentation. Utilizing this insight, they selectively sampled old samples that significantly influenced model training and leveraged them as memory sets to enhance overall model performance. Iscen et al. [9] directly stored the sample feature representation extracted by the model, reducing storage load. Zhu et al. [22] employed a generative model to dynamically generate old class data, but the generative model itself will deteriorate performance as the number of tasks increases, so it is unsuitable for long-term learning. The second aspect is the parameter, where regularization methods are commonly employed to mitigate the issue of forgetting old knowledge. Kirkpatrick et al. [11] proposed the method of elastic weight consolidation (EWC), which introduced a new quadratic penalty term in the loss function to diminish the impact of new knowledge on important weight parameters in old class knowledge. Zenke et al. [20] proposed synaptic intelligence (SI), which assessed the importance of network parameters by using their contribution to the reduction of model loss. Chaudhry et al. [2] combined SI and EWC to estimate the importance of network parameters. The third aspect is the model, in which different methods are used to alleviate the problem of forgetting old knowledge. Li et al. [12] proposed a learning without forgetting (LwF) algorithm, which used knowledge distillation [6] to modify the cross entropy loss, by which a new network is trained, thereby retaining some of the knowledge of the old network. Hussain et al. [8] cloned and shared the fully connected layer of the network to retain the knowledge of the old class. Zhao et al. [21] used weight align (WA) to alleviate the imbalance problem between new and old samples. Douillard et al. [4] used an efficient spatial-based distillation-loss applied throughout the model, yielding favorable outcomes in small task increments. Kang et al. [10] used a knowledge distillation method of adaptive feature consolidation to adjust the weight of each feature map to minimize the upper bound of loss increases.

Among the studies on one-dimensional signal recognition, researchers have explored incremental learning methods focusing on signal quantity [5, 18] and signal class [3, 14, 23]. Guo et al. [5] proposed the Incremental Small Ball Large Margin (IncSSLM) algorithm, which can learn the compact boundary between its own communication signal and known interference. It can update the classifier model in real time, eliminating the need for extensive memory storage of interference signal samples, saving a lot of training time, and realizing incremental learning on the signal quantity. Yang et al. [18] proposed a deep integrated siamese network of performing incremental learning which is based on signal quantity. However, [5,18] cannot solve the class-incremental learning of a large number of classes,which contain new class labels. There are a few algorithms for class-incremental learning of signals. Most methods rely on fine-tuning the model for incremental learning, but this approach often leads to significant catastrophic forgetting issues, making it unsuitable for long-term incremental recognition. Montes et al. [14] evaluated the performance of non-incremental method, knowledge distillation method , and bias correction method which is based on the same network in incremental learning of modulated signals. Experimental results show that knowledge distillation and bias correction methods exhibit similar performance in the incremental learning of modulated signals and outperform non-incremental learning methods. Zhao et al. [23] proposed an incremental learning algorithm for modulated signals based on sample playback, but each incremental stage requires the generation of old samples, and the algorithm complexity is high. Chen et al. [3] proposed an open set incremental recognition method based on prototype learning. This method achieves open set recognition of signals while performing incremental recognition. Nonetheless, its performance in incremental recognition falls short. The performance of the identification of interference types will affect the anti-jamming effect of the communication [19]. Hence, we propose a double auxiliary learning based interference signal incremental recognition algorithm from the aspect of the balance between the consolidation of old knowledge and the learning of new knowledge. Two auxiliary learning networks are used to strengthen the extracting and learning capability of both new and old knowledge of interference signals, thereby improving the classification performance of the model. Additionally, we utilize the Nearest-Mean-of-Exemplars (NME) criterion to enhance interference classification accuracy. The contributions of this paper are mainly as follows:

•The new feature extraction network (FENet) is designed by using DenseNet module and four classifiers which include two main classifiers and auxiliary classifiers for both new and old samples.

•The Nearest-Mean-of-Exemplars (NME) criterion is used to construct a memory set and classify the communication interference signals.

•We conduct comprehensive experiments by using the real communication interference signals. The experimental results of 30 signal types show that the proposed method has a better recognition accuracy than other methods in the incremental learning.

2. Algorithm design

The purpose of incremental recognition of communication interference signals is to continuously identify new classes appearing in the electromagnetic environment while retaining the ability to recognize old class samples. Assume that $T = \{T_1, T_2, \dots\}$ is the training task sequence. Among them, the training of i-th class set of the t-th subtask T_t is $D^{i,t} = \{(x_1^{i,t}, y_1^{i,t}), (x_2^{i,t}, y_2^{i,t}), \dots, (x_{n_t}^{i,t}, y_{n_t}^{i,t})\}$, n_t is the number of samples, $x_j^{i,t}$ is the j-th sampling sequence of the i-th class of communication interference signal, which is the IQ dual-channel sampling signal, and $y_j^{i,t}$ is the label of the j-th sampling sequence of the i-th communication interference signal. There is no intersection between the sample categories of different subtasks. To keep the model's memory for old samples, a memory set G^{t-1} is usually constructed and saved from some old samples of the former t-1 subtasks, and used to train the model for the t-th subtask together with the new samples. We regard the first subtask as the pre-training stage of the model, and the remaining subtasks as the incremental stage of the model.

The framework of the proposed DAL-NME algorithm is shown in Figure 1. It consists of three parts: a feature extracting network based on double auxiliary learning, a distance classifier, and a memory set. The cross entropy loss and distillation loss are used to train the feature extractor in the FENet based on DAL. The Nearest-Mean-of-Exemplars (NME) criterion is used to perform incremental phase recognition and construct the memory samples for subsequent incremental learning.

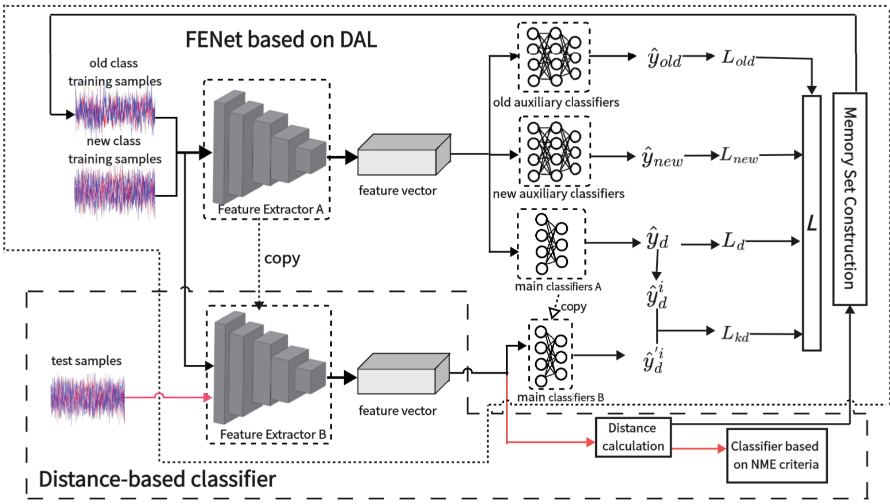


FIGURE 1. DAL-NME Algorithm Framework

2.1. **FENet based on Double Learning.** The DAL-based FENet consists of two feature extractors, two auxiliary classifiers, and two main classifiers. The network input is the IQ signal sample of the new class in the incremental stage t and the IQ signal sample of the old class from the memory set constructed in stage t-1. The feature vector of the communication interference signal is extracted by the feature extractor, followed by a network loss calculation to update the parameters of the FENet. DenseNet has fewer parameters and computational complexity, and its performance is better than ResNet which solves the gradient vanishing and network degradation problems in CNN, thus the feature extractor adopts the DenseNet architecture. The basic structure of the DenseNet network is shown in Figure 2. The number of channels is increased by gr when the input features pass through this structure. The feature extractor is shown in Figure 3. It consists of a convolutional layer, a max pooling layer, 15 DenseBlocks, 3 Transition Layers, and a global average pooling layer. DenseBlock consists of two convolutional layers and a channel attention mechanism SENet [7]; Transition Layer consists of a convolutional layer and an average pooling layer, which is used to reduce the dimension of the extracted features. gr is the growth rate of the DenseBlock structure, cr is the compression rate of the Transition Layer, which are set to gr=32 and cr=0.5, respectively. 'in' is the number of channels in the previous module. All one-dimensional convolutional layers have a batch normalization (BN) layer and a Rectified Linear Unit (ReLu) layer. {7Conv1d,64,/2} indicates a one-dimensional convolution with a kernel size of 7, an output channel number of 64, and a stride of 2. {1dMaxPool,3,/2} indicates a one-dimensional max pooling layer with a pooling window size of 3 and a stride of 2. { $\times$ 4} represents the module with four duplicate pieces. The structure of two feature extractors is the same. The main classifier is directly composed of a fully connected layer and a softmax layer. The structure of the two main classifiers is the same. The auxiliary classifier is composed of two fully connected layers and a softmax layer. In the pre-training stage, only the main classifier A is used, while in the incremental stage, two main classifier, a new sample auxiliary classifier and an old sample auxiliary classifier are used. In the new incremental stage, the parameters of the feature extractor B and the main classifier B are copied from the feature extractor A and the main classifier A in the previous incremental stage, respectively, and they will remain unchanged and not participate in the gradient update of the network. The specific settings of hyperparameters are shown in Table 1.

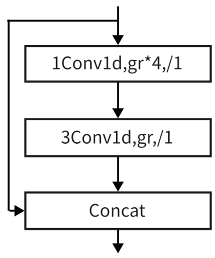


FIGURE 2. The Basic Structure of the DenseNet Network

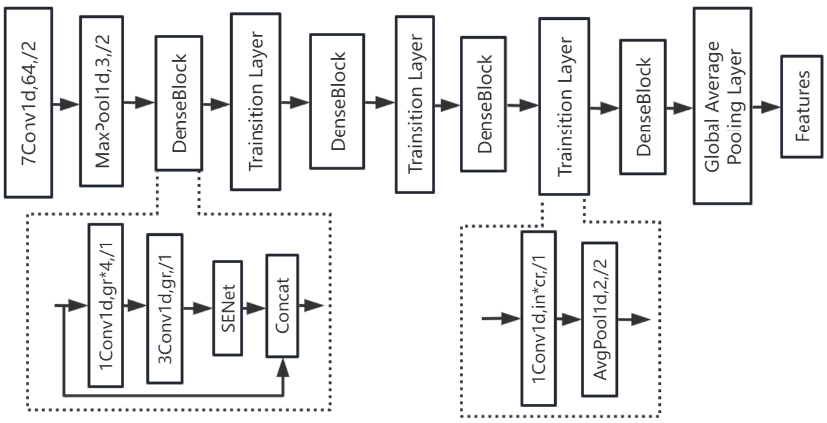


FIGURE 3. Feature Extractor

2.2. **Design of loss function.** Assume that the number of old class samples in the incremental stage is denoted as OS, the number of new class samples as NS, and the total number of classifications as M, then M=OS+NS. For the main classifier,

TABLE 1. Parameters of Different Communication Interference Signals

The name of the structure	Network parameters
7Conv1d, 64/2	Size:7, channels:64, stride:2
1Conv1d, gr*4,/1	Size:1, channels:gr*4, stride:1
3Conv1d, gr,/1	Size:3, channels:gr, stride:1
Maxpool1d, 3, /2	Size:3, stride:2
Avgpool1d, 2, /2	Size:2, stride:2
Growth rate(gr)	32
1Conv1d, in*cr,/1	Size:1, channels:in*cr, stride:1
Compression rate(cr)	0.5

the number of classifications is set to M, that is, all samples are classified. For the old sample auxiliary classifier, the number of classifications is set to OS+1, the old class samples are classified normally, and all new class samples are classified into one category. For the new sample auxiliary classifier, the number of classifications is set to NS+1, the new class samples are classified normally, and all old class samples are classified into one category. $\mathbf{y}_d$, $\mathbf{y}_{old}$ and $\mathbf{y}_{new}$ are the classification labels corresponding to the main classifier A, the old sample auxiliary classifier, and the new sample auxiliary classifier, respectively, $\hat{\mathbf{y}}_d$, $\hat{\mathbf{y}}_{old}$ and $\hat{\mathbf{y}}_{new}$ are the outputs of each classifier. And their cross entropy classification losses are recorded as L_d, L_{old} and L_{new} respectively, as shown below:

$$L_d = \mathbf{y}_d log(\hat{\mathbf{y}}_d), \tag{2.1}$$

$$L_{old} = \mathbf{y}_{old} log(\hat{\mathbf{y}}_{old}), \tag{2.2}$$

$$L_{new} = \mathbf{y}_{new} log(\hat{\mathbf{y}}_{new}), \tag{2.3}$$

$$L_{kd} = \sum_{i=1}^{OS} \hat{y}_d^{'i} log(\hat{y}_d^i). \tag{2.4}$$

In order to enhance the memory of old samples, we use the distillation loss [6], denoted as L_{kd}

Among them, $\hat{y}_d^{'i}$ represents the predicted probability of the $i-th$ class obtained by using the feature extractor B and main classifier B while $\hat{y}_d^i$ represents the predicted probability of the $i-th$ class obtained by using the feature extractor A and main classifier A. In order to keep the balance between the network's consolidation of old knowledge and the absorption of new knowledge, the regularization coefficient $\lambda_{old} = \frac{OS}{OS+NS}$ is set. It can be seen that as the incremental stage progresses, the proportion of the old class samples gradually increases, and λ_{old} also gradually increases, so that the network can better retain the knowledge of the old samples. Therefore, the total loss function L of the incremental stage is:

$$L = \lambda_{old}(L_{old} + L_{kd}) + (1 - \lambda_{old})(L_{new} + L_d). \tag{2.5}$$

2.3. **Nearest-Mean-of-Exemplars (NME) Classification Criterion.** The ultimate goal of the auxiliary classifiers and the main classifiers is to optimize the feature extractor within the network and facilitate accurate classification of both new and old samples. Therefore, the algorithm in this paper adopts the Nearest-Mean-of-Exemplars (NME) classification criterion to classify both new and old samples. In the $t-th$ incremental stage, the feature vector extracted by FENet based on DAL for the sample $(x_k^{i,t}, y_k^{i,t})$ is $\mathbf{z}_k^{i,t}$, i is the category label of the sample, k is the sequence label of the sample, I is the number of old sample categories in this incremental stage, and J is the number of samples in the $l-th$ class. The distance between the extracted feature vector $\mathbf{z}_k^{i,t}$ and the mean of the feature vector of the $l-th$ class is denoted as $d_k^{i,t}(l)$, and it is calculated as:

$$d_k^{i,t}(l) = \left\| \mathbf{z}_k^{i,t} - \frac{1}{J}\sum_{j=1}^{J} \mathbf{z}_j^{l,t} \right\| \qquad i = 1, 2, \ldots, I. \tag{2.6}$$

$l_{min} = \underset{l}{argmin} d_k^{i,t}(l)$ is the predicted category, and when $i = l_{min}$, it is the correct classification.

2.4. **Construction of Memory Set.** To prevent the escalation of network complexity caused by the continual accumulation of memory samples throughout the incremental learning process, the number of samples of the memory set is fixed, which is set K, and the selection process follows the following rules. From formula (2.6), we can get that the distance $d_k^{i,t}$ between the feature vector $\mathbf{z}_k^{i,t}$ of the sample and the mean of the feature vector of the $i-th$ class sample $(x_j^{i,t}, y_j^{i,t})$ is:

$$d_k^{i,t}(i) = \left\| \mathbf{z}_k^{i,t} - \frac{1}{J}\sum_{j=1}^{J} \mathbf{z}_j^{i,t} \right\| \qquad i = 1, 2, \ldots, I. \tag{2.7}$$

The samples corresponding to the smallest front ξ^t values in $d_k^{i,t}(k = 1, 2, \ldots, J)$ are selected as samples of the memory set G^t of the $i-th$ communication interference signal in the increment stage $t+1$, and ξ^t is the number of saved samples of each category after the increment stage t.

3. Algorithm simulation

All methods use momentum SGD with momentum of 0.9 and learning rate of 0.01. The weight decay coefficient for all optimizers is set to 1e-4, and the fixed step decay learning rate method is used, with 100 iterations, a step size of 30, and a learning rate decay coefficient of 0.1. The average of the recognition accuracy of the algorithm under each jamming noise ratio (JNR) in different incremental stages is denoted as the task-avg-accuracy, and the average of the recognition accuracy in different incremental stages under each JNR is denoted as the JNR-avg-accuracy. The recognition accuracy, task-avg-accuracy and JNR-avg-accuracy are used to evaluate the algorithm performance. The pseudocode of the algorithm is given in Algorithm 1. And the specific settings of the training parameters are shown in Table 2.

To test the performance of the proposed DAL-NME algorithm for the interference signal incremental recognition, a wireless signal transceiver system is built. It consists of a signal generator, a signal receiver, a transceiver antenna and a laptop, as shown in Figure 4. The signal generator uses SMBV100A from R&S to generate

Algorithm 1 The training of DAL-NME Algorithm.

Require: Input x_k^i are the k-th training interference signals of i-th class.
Ensure: The losses of FENet and the mean of the feature vector $\mathbf{m}_i$ of the i-th class.

for i = 1,2,...,I **do**
  for k = 1,2,...,K **do**
    $\mathbf{z}_k^i = Feature\ ExtractorA(x_k^i)$
    $\hat{\mathbf{y}}_{old} = old\ classfication(\mathbf{z}_k^i)$
    $\hat{\mathbf{y}}_{new} = new\ classfication(\mathbf{z}_k^i)$
    $\hat{\mathbf{y}}_d = main\ classfication(\mathbf{z}_k^i)$
    L_d, L_{old}, L_{new}, and L_{kd} is according to Eq.2.1, Eq.2.2, Eq.2.3 and Eq.2.4
  end for
end for
for i = 1,2,...,I **do**
  The mean of the feature vector of the i-th class is: $\mathbf{m}_i = \left\| \frac{1}{K} \sum_{k=1}^{K} \mathbf{z}_k^i \right\|_2$
end for
The testing of DAL-NME Algorithm

Require: For each class i, input x_k^i are the k-th testing interference signals of i-th class.
Ensure: The result of classification.

for i = 1,2,...,I **do**
  for k = 1,2,...,K **do**
    $\mathbf{z}_k^i = Feature\ ExtractorB(x_k^i)$
    $d_i^k(l) = \left\| \mathbf{z}_k^i - \mathbf{m}_l \right\|_2$
    $l_{min} = \underset{l}{argmin}\ d_k^i(l)$ is the predicted category
  end for
end for

Table 2. Training Parameters

Parameters	Value
Momentum	0.9
Initial learning rate	0.01
Decay learning rate	1e-4
Iterations	100
Learning rate decline cycle	30
Learning rate decay coefficient	0.1
Optimizer	Momentum SGD

the communication interference signal required for incremental identification; the transceiver antenna can transmit and receive the signal frequency range between 30-500MHz; TuWR-G69DDCe receiver is used to receive the communication interference signal transmitted wirelessly; the WiNRADiO G69DDC software developed by Radixon is installed on the laptop to collect the communication interference signal data required for the experiments. The collected communication interference signals include 7 interference signals and 23 modulation signals. The interference signals are single-tone interference (ST), multi-tone interference (MT), linear frequency sweep interference (LFS), secondary frequency sweep interference (SFS), noise FM modulation interference (NFM), frequency hopping interference (FH) and periodic Gaussian pulse interference (PGP), and the modulation signals are BPSK, QPSK, 8PSK, OQPSK, 16PSK, 32PSK, 2FSK, 4FSK, 8FSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, 16APSK, 32APSK, 64APSK, 128APSK, 256APSK, 4PAM, 8PAM, 16PAM and MSK. Orthogonal sampling is used to obtain the in-phase (I) component samples and the quadrature (Q) component samples. The number of sampling points is 1024, and the specific parameters are shown in Table 3.

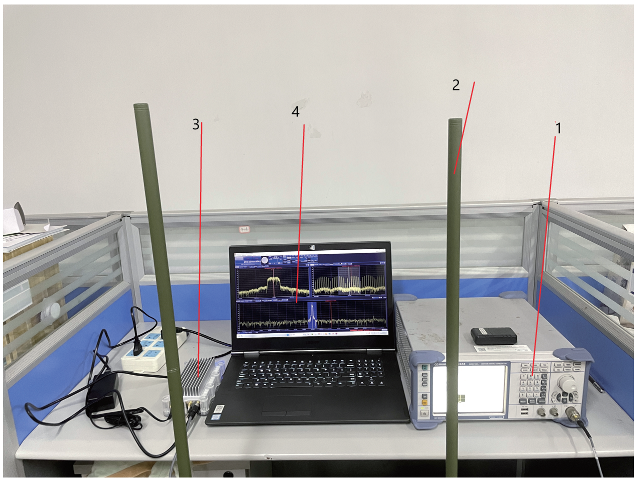


FIGURE 4. Wireless Signal Transceiver System

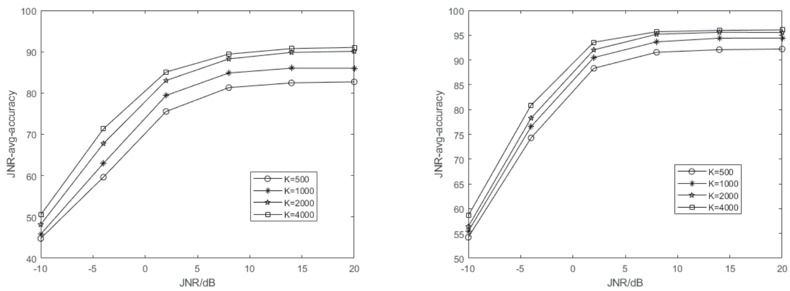
TABLE 3. Parameters of Different Communication Interference Signals

Interference Type	Interference signal	Modulation Signal
Special Parameters	MT: 2∼10 audio numbers. NF: The frequency modulation factor is 0.125∼0.933. FH: The frequency hopping times are 20. PGP: The pulse period is 16 to 64 sampling points.	Information sequence: 01 bit sequence. Shaping filter: root raised cosine filter with a roll-off factor of 0.2 to 0.7. Normalized frequency offset: -0.2 to 0.2.
Unified parameters	Phase offset: $0 \sim \pi$. Interference-to-noise ratio: -10∼20dB, training set interval 2dB, test set interval 6dB. 8 times oversampling. The number of training samples for each category and each interference-to-noise ratio is 100, and the number of test samples is 500.	

In order to analyze the performance of the proposed DAL-NME algorithm, two sets of incremental sequences are set, as shown in Table 4. Each incremental sequence carries on two incremental processes.

Table 4. Parameters of Different Communication Interference Signals

Incremental sequence 1	ST, LFM, PGP, FH, QPSK, BPSK, NFM, SFS, 8PSK, OQPSK, MT, 256QAM, 16PSK, 32PSK, 2FSK, 16APSK, 128APSK, 8PAM, 4FSK, 8FSK, 16QAM, 32APSK, 256APSK, 16PAM, 32QAM, 64QAM, 128QAM, 64APSK, 4PAM, MSK
Incremental sequence 2	32QAM, 64QAM, 128QAM, 64APSK, 4PAM, MSK, 4FSK, 8FSK, 16QAM, 32APSK, 256APSK, 16PAM, NFM, SFS, 8PSK, OQPSK, MT, 256QAM, ST, LFM, PGP, FH, QPSK, BPSK, 16PSK, 32PSK, 2FSK, 16APSK, 128APSK, 8PAM
Incremental scene 1	Incremental sequence 1, 6 interference signals are added in each increment stage
Incremental scene 2	Incremental sequence 1, 3 interference signals are added in each increment stage
Incremental scene 3	Incremental sequence 2, 6 interference signals are added in each increment stage
Incremental scene 4	Incremental sequence 2, 3 interference signals are added in each increment stage



(a) Incremental scene 1 (b) Incremental scene 2

Figure 5. JNR-avg-accuracy of Incremental Recognition of DAL-NME Algorithm under Different Memory Sets

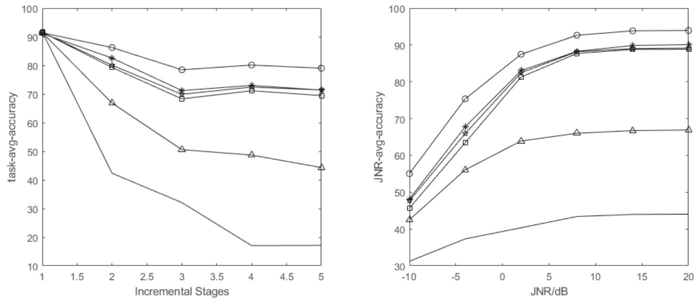
3.1. **Impact of Memory Set Size on DAL-NME Algorithm.** The size of the memory set will directly affect the performance of incremental learning. In order to investigate the impact of the memory set size on the performance of the DAL-NME algorithm, the fixed memory set size K is set to 500, 1000, 2000 and 4000. The DAL-NME algorithm performs incremental recognition in incremental scenes 1 and 3. The JNR-avg-accuracy of the DAL-NME algorithm is shown in Figure 5.

As can be seen from Figure 5, in the two incremental scenes, as the size of the memory set increases, the JNR-avg-accuracy under each JNR is improved. When the size of the memory set increases from 500 to 1000 and from 1000 to 2000, the JNR-avg-accuracy is greatly improved; when the memory set size increases from 2000 to 4000, the JNR-avg-accuracy under low JNR ratio increases greatly, while under high JNR it increases less. Since a large memory set will increase computational complexity and storage complexity, the fixed memory set size K=2000 is chosen to compromise between algorithm performance and complexity.

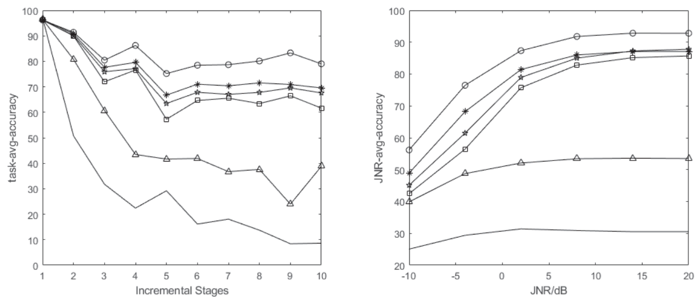
3.2. **Incremental Learning Performance Analysis.** The DAL-NME algorithm without the double auxiliary classifier is denoted as NAL-NME. The LwF [12], WA [13], joint, TOCIL [21], NAL-NME and DAL-NME algorithms are used to incrementally identify communication interference signals. Additionally, joint means that all old samples are retained for joint training during the incremental process. All algorithms use the same feature extractor. Except for the DAL-NME algorithm, all other algorithms only have the main classifier. The LwF algorithm does not use the memory set, and the other algorithms all use the memory set. In the four incremental scenes, the task-avg-accuracy and JNR-avg-accuracy of the six algorithms for incremental recognition are shown in Figure 6.

As can be seen from Figure 6: (1) As JNR rises, the JNR-avg-accuracy of all algorithms in each incremental scene exhibits a gradual increase, eventually stabilizing after JNR reaches 8dB. As the incremental stage increases, the task-avg-accuracy reduces and fluctuates. This is because that the increase of the number of categories results in higher degree of difficulty of the classification. The occasional upward trend is because the recognition of some interference signals is relatively easy. As can be seen from Figure 6(b), in the incremental stage 3 {NFM, SFS, 8PSK} combination recognition is more difficult, while in the incremental stage 4 {OQPSK, MT, 256QAM} combination is relatively easy to recognize, so the recognition accuracy rate increases instead of decreasing. (2) The joint algorithm consistently achieves the highest task-avg-accuracy and JNR-avg-accuracy across each incremental scene. This is because the joint algorithm uses all samples to train the network. The joint algorithm is generally considered to be the upper limit of the performance of incremental learning. (3) Incorporating memory sets can significantly enhance the recognition accuracy of communication interference signals during the incremental process. The utilization of memory sets in conjunction with the joint, WA, TOCIL, NAL-NME, and DAL-NME algorithms has demonstrated superior performance compared to the LwF algorithm without memory sets. Therefore, it is necessary to add some memory sets in the incremental recognition process of communication interference signals. (4) In each incremental scene, the recognition performance of the proposed DAL-NME algorithm outshines that of the WA, TOCIL and LwF algorithms. The TOCIL algorithm greatly reduces the closed-set recognition ability while ensuring the open-set recognition performance. (5) The recognition accuracy of the DAL-NME algorithm surpasses that of the NAL-NME algorithm, highlighting the efficacy of the dual auxiliary classifier.

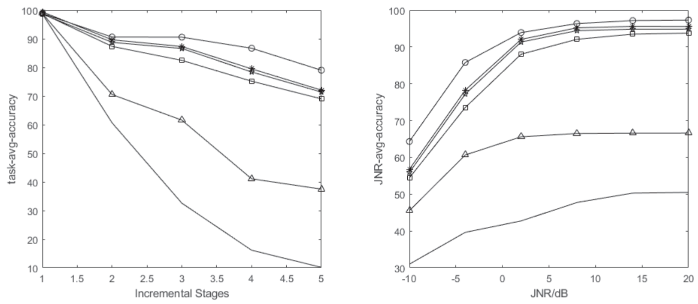
3.3. **Recognition Performance in the Incremental Stage.** In order to evaluate the performance of various algorithms in the incremental process, the third incremental stage of scene 1, the fifth incremental stage of scene 2, the fourth incremental stage of scene 3, and the sixth incremental stage of scene 4 were randomly selected. Figure 7 shows the relationship curves between the recognition accuracy and JNR of the six algorithms in the four scenes.



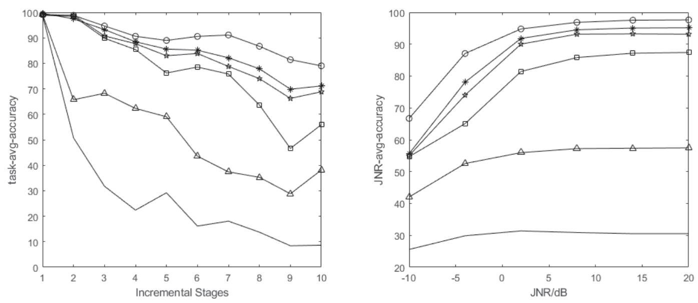
(a) Incremental scene 1



(b) Incremental scene 2



(c) Incremental scene 3



(d) Incremental scene 4

JIONT DAL-NME NAL-NME WA LwF TOCIL

FIGURE 6. Task-avg-accuracy and JNR-avg-accuracy in Four Incremental Scenes

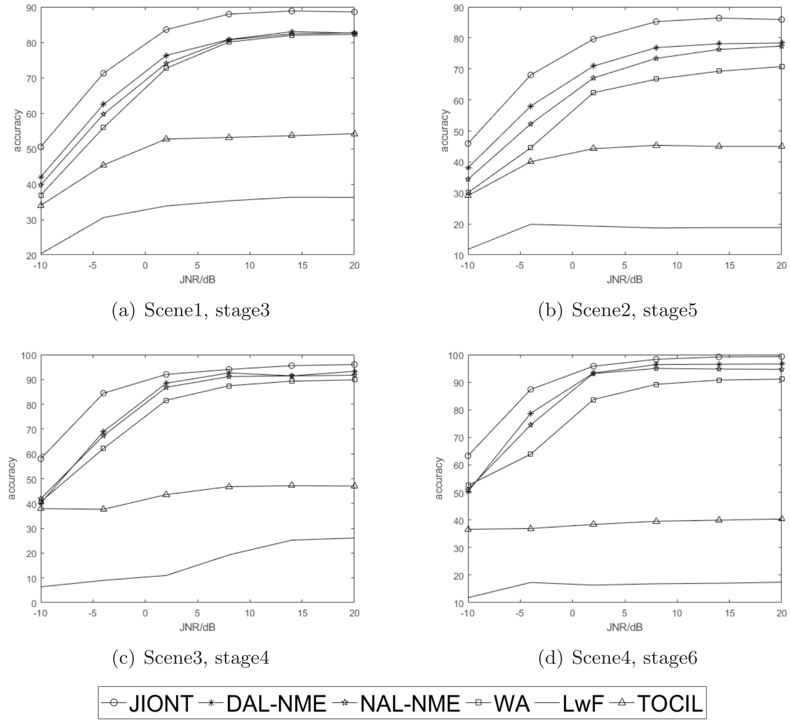


FIGURE 7. Recognition Accuracy of the Algorithm in the Incremental Stage

As can be seen from Figure 7: (1) In the random incremental stage of each scene, the joint algorithm has the best recognition performance, and the recognition accuracy of the six algorithm increases with the increase of JNR, and tends to be stable after JNR exceeds 8dB. (2) Except for the recognition performance of the DAL-NME algorithm in the incremental stage 6 of scene 4 at JNR of -10dB, which is slightly lower than that of the WA algorithm, the recognition accuracy of the DAL-NME algorithm in the random incremental stage of other scenes is higher than that of the WA, TOCIL, LwF and NAL-NME algorithms. In the random incremental stage of each scene, when JNR is 2dB, the accuracy of DAL-NME algorithm is 5% and 25% higher than that of WA and TOCIL, respectively. (3) The LwF algorithm, lacking a memory set, almost loses its ability to recognize old samples as the incremental stage increases. This indicates that the use of a memory set can alleviate catastrophic forgetting.

4. Conclusion

An incremental learning algorithm DAL-NME suitable for communication interference signals has been proposed. The designed feature extracting network is composed of two feature extractors, two main classifiers, and two auxiliary classifiers. The designed distance-based classifier uses the feature extractor and the NME criterion to test communication interference signals. Samples with more obvious features are selected as memory sets by using the NME criterion. The WA, TOCIL, joint, LwF, NAL-NME and DAL-NME algorithms are simulated on 30 actual sampled communication interference signals. The experimental results show that the DAL-NME algorithm has a high recognition effect under various JNR, and it can better keep the balance between the retention of old knowledge and the learning of new knowledge in the incremental process. This algorithm can better identify newly emerging interference signals with labels while maintaining superior recognition capabilities for existing interference signals. However, new class of interference signals in reality may be unlabeled in some cases, so the problem of unlabeled new class signals in incremental learning should be considered in the future.

References

[1] J. Bang, H. Kim, Y. Yoo, J. W. Ha and J. Choi, *Continual learning with a memory of diverse samples*, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2021, pp. 8218–8227.

[2] A. Chaudhry, P. K. Dokania, T. Ajanthan and P. H. Torr, *Riemannian walk for incremental learning: Understanding forgetting and intransigence*, in: Proceedings of the European Conference on Computer Vision (ECCV), IEEE, 2018, pp. 532–547.

[3] X. Chen, *Open set recognition of communication interference signals based on deep learning*, Hangzhou Dianzi University, 2023.

[4] A. Douillard, M. Cord, C. Ollion, T. Robert and E. Valle, *Podnet: Pooled outputs distillation for small-tasks incremental learning*, in: Proceedings of 16th European conference, part XX 16, Springer International Publishing, 2020, pp. 86–102.

[5] Y. Guo, J. Meng, Y. Li, S. Ge, J. Xing and H. Wu, *Incremental small sphere and large margin for online recognition of communication jamming*, Applied Intelligence **50** (2020), 3429–3440.

[6] G. Hinton, *Distilling the knowledge in a neural network*, ArXiv Preprint ArXiv:1503.02531, 2015.

[7] J. Hu, L. Shen and G. Sun, *Squeeze-and-excitation networks*, in: Proceedings of 2018 IEEE/CVF the Conference on Computer Vision and Pattern Recognition, IEEE, (2018), pp. 7132–7141.

[8] M. A. Hussain , S. A. Huang and T. H. Tsai, *Learning with sharing: An edge-optimized incremental learning method for deep neural networks*, IEEE Transactions on Emerging Topics in Computing **2** (2022), 461–473.

[9] A. Iscen, J. Zhang, S. Lazebnik and C. Schmid, *Memory-efficient incremental learning through feature adaptation*, in: Proceedings of 16th European Conference, Part XVI 16, Springer International Publishing, 2020, pp. 699–715.

[10] M. Kang , J. Park and B. Han, *Class-incremental learning by knowledge distillation with adaptive feature consolidation*, in: Proceedings of the Conference on Computer Sision and Pattern Recognition, IEEE, 2022, pp. 16071–16080.

[11] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu and R. Hadsell, *Overcoming catastrophic forgetting in neural networks*, Proceedings of the National Academy of Sciences. **13** (2017), 3521–3526.

[12] Z. Li and D. Hoiem *Learning without forgetting*, IEEE Transactions on Pattern Analysis and Machine Intelligence **12** (2017), 2935-2947.

[13] Q. Liu and W. Zhang, *Deep learning and recognition of radar jamming based on CNN*, in: Proceedings of 12th International Symposium on Computational Intelligence and Design (ISCID), IEEE, vol. 1, 2019, pp. 208–212.

[14] C. Montes , T. Morehouse and R. Zhou, *Class-Incremental learning for baseband modulation classification: A Comparison*, in: Proceedings of International Wireless Communications and Mobile Computing (IWCMC), IEEE, 2024, pp. 1370–1375

[15] Q. Qu, S. Wei, S. Liu , J. Liang and J.Shi, *JRNet: Jamming recognition networks for radar compound suppression jamming signals*, IEEE Transactions on Vehicular Technology **12** (2020), 15035–15045.

[16] S. A. Rebuffi, A. Kolesnikov, G. Sperl and C. H. Lampert, *Incremental classifier and representation learning*, in: Proceedings of the Conference on Computer Vision and Pattern Recognition Systems Control Lett(CVPR), Institute of Science and Technology, 2024, pp. 5533–5542.

[17] G. Shao and Y.Chen, *Convolutional neural network-based radar jamming signal classification with sufficient and limited samples*, IEEE Access **8** (2020), 80588–80598.

[18] C. Yang and S. Yang, *Deep ensemble siamese network for incremental signal classification*, in: Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2021, pp. 3200–3204.

[19] F. Q. Yao, *Communication anti-jamming engineering and practice*, Beijing Publishing House of Electronics Industry, (2008), 1–8.

[20] F. Zenke, B. Poole and S. Ganguli, *Continual learning through synaptic intelligence*, in: Proceedings of International Conference on Machine Learning, PMLR, 2017, pp. 3987–3995.

[21] B. Zhao, X. Xiao, G. Gan, B. Zhang and S Xia, *Maintaining discrimination and fairness in class incremental learning*, in; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, (2020), pp. 13208–13217.

[22] Y. Zhao, S. Chen, T. Chen, W. Shen, S. Zheng, Z. Zhao and X. Yang, *Incremental learning of radio modulation classification based on sample recall*, China Communications **7** (2023), 258–272.

[23] K. Zhu, Y. Cao, W. Zhai, J. Cheng and Z. J. Zha, *Self-promoted prototype refinement for few-shot class-incremental learning*, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, (2021), 6801-6810.

Manuscript received September 30, 2024
revised December 17, 2024

Shaogang Dai
School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, 310000, China
E-mail address: 15381139055@163.com

Yinzhi Wang
School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, 310000, China
E-mail address: 222080135@hdu.edu.cn