Meta-RFF: Few-Shot Open-Set Incremental Learning for RF Fingerprint Recognition via Multi-phase Meta Task Adaptation

Taotao Li[1], Zhenyu Wen[1,2](✉), Yuting Jiang[1], Jian Wang[1], Jie Su[1], Zhen Hong[1](✉), Li Yu[1], and Shibo He[3]

[1] The Institute of Cyberspace Security, and the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China,
{zhenyuwen, zhong1983}@zjut.edu.cn

[2] University of Science and Technology of China, Hefei, China

[3] College of Control Science and Technology, Zhejiang University, Hangzhou, China

Abstract. RF fingerprinting (RFF) is a non-encrypted authentication technique that provides an additional layer of security for wireless devices, which has a promising application. However, existing RFF recognition techniques that rely on deep learning (DL) are usually available with limited equipment. In actual application scenarios, new wireless devices are constantly appearing, such as unknown drones that appear suddenly in the sky. In such cases, the RF monitoring system should equip the ability to discover the unknown device (i.e., open-set recognition (OSR)) and use captured few samples of new devices incrementally updating knowledge of the system. This requirement brings two challenges: 1) incremental updates from few-shot samples are prone to lead to catastrophic forgetting and over-fitting problems; 2) constructing a reliable OSR mechanism for new devices with few-shot samples is difficult. To tackle this challenge, for the first time, we propose a novel few-shot open-set incremental learning (FSOSIL) framework via meta-learning for RFF recognition (*Meta-RFF*). The core idea of *Meta-RFF* is to simulate few-shot RF signal incremental learning by constructing many pseudo-FSOSIL tasks. In particular, to strengthen the OSR capability, we further design RF feature augmentation and open space learning modules. The algorithm is validated on the large-scale aircraft recognition dataset (namely ADS-B), which shows that the close-set accuracy and open-set AUROC of the new class improve the performance by about 10-20% compared to other algorithms with 1-shot. And in 10 increments, our algorithm possesses a lower performance decay rate (about 3%).

Keywords: Deep Learning, RF Fingerprints, Few-Shot Incremental Learning, Open-set Recognition.

1 Introduction

In recent years, the increasing number of wireless devices has led to the openness of the wireless space [9]. However, it is difficult to guarantee data confidentiality, leaving wireless systems open to potential security threats. For example, in the automatic dependent surveillance-broadcast (ADS-B), the aircraft identify information is also easily imitated and tampered with, resulting in potential security risks [17]. To solve the above problems, non-imitated RF fingerprinting (RFF)

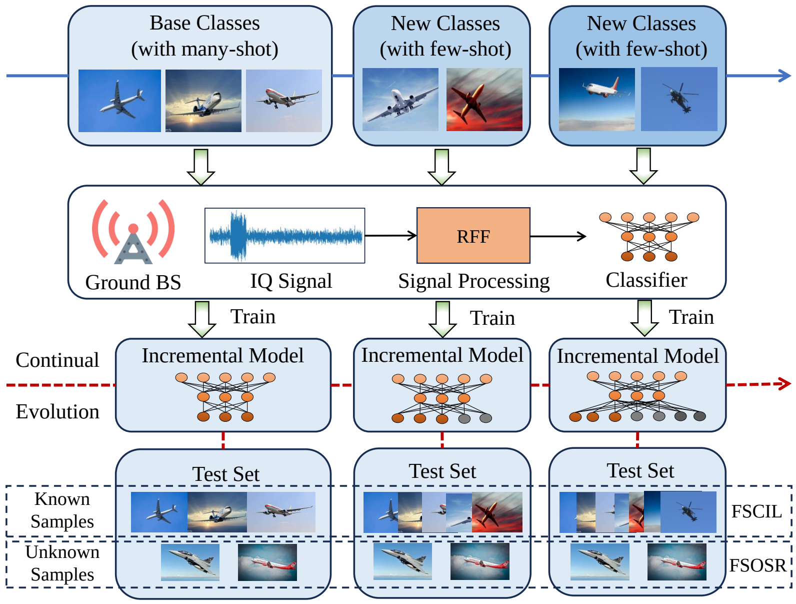


Fig. 1. Aircraft classifier system deployed in ground BS not only needs to accurately recognize known aircraft (base classes) but also needs to recognize unknown aircraft (new classes). This classifier is equipped with continuous few-shot class incremental learning capability to achieve co-recognition of new and old classes. At the same time, the few-shot OSR capability also needs to be improved in the continuous increment.

techniques have been developed to identify devices by exploiting the hardware imperfections of the transmitter like I-Q imbalance, nonlinear distortion of power amplifier (PA), loop filter variations, etc[12]. Nowadays, RFF technologies are widely applied in military and civilian fields, such as IoT device authentication [13], spectrum monitoring [1], and aviation management [8], etc.

RFF technologies can be divided into traditional and deep learning-based (DL-based) methods. The traditional approaches use expert knowledge to extract statistical, spectral, and transient features of the signal, which are labor-intensive and time-consuming [12]. Compared with them, the DL-based approaches can automatically learn RFF features to identify the devices. Relevant deep-learning methods include convolutional neural networks (CNNs)[19], residual network (ResNet)[9], long short-term memory (LSTM)[4], etc. However, the above methods are usually available with limited equipment and do not consider the following deployment challenges.

Autonomous radio monitoring system. Fig. 1 shows that the aircraft identification system (AIS) deployed ground base station (BS) will utilize the DL model for aircraft classification by extracting RF fingerprints (RFF) of IQ signal. The recognition system must be equipped to meet various practical environmental demands, including (1) the capacity for open-set recognition (OSR) to identify sudden appearances of unfamiliar aerial objects [8], (2) the ability to update and recognize new categories using a limited number of samples from

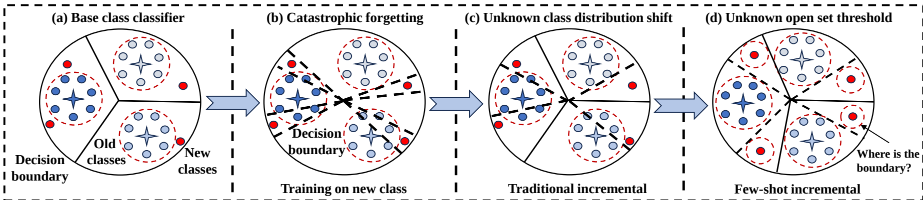


Fig. 2. Existing DL technology bottlenecks: (a)The decision boundary (solid line) of the base class classifier can accurately classify old classes, but not classify the new classes (red); (b)By training the classifier directly on the new class, the model is biased towards fitting the new class (dotted line), forgetting the knowledge of the old class; (c)Uncertain new class distributions are shifted to the vicinity of old classes, resulting in limited classification performance; (d)Difficulty in determining the OSR boundary (red dotted line) for the new class due to its few-shot limitation.

these newly emerged classes [7], and (3) the flexibility to continually evolve, accommodating an increasing range of categories and adapting to unforeseen environments in the future.

However, building such a generalized evolutionary model that combines few-shot class incremental learning (FSCIL), few-shot open-set (FSOSR), and multi-stage continuous increment needs to overcome the following challenges.

– *Catastrophic Forgetting and Overfitting.* As shown in Fig. 2b, direct training on new classes with few-shot inevitably leads to the model forgetting the knowledge of the old classes [6] and over-fitting effects on the limited data [14, 16].
– *Unknown Class Distribution Shift.* As shown in Fig. 2c, unknown new class distributions may be shifted in the vicinity of old classes, leading to limited classification performance of traditional incremental models [6, 15].
– *Unknown Open Set Boundary.* As shown in Fig.2d, the lack of samples makes us unable to see the true class distribution, which hinders the accurate estimation of the open set boundaries.

In addition, the multi-stage continuous incremental process aggravates the above challenges and leads to a drastic decrease on the efficiency of existing incremental algorithms[6], iCARL[15], OSR[2], and FSCIL [22]. To address the above practical challenges, we have rethought human learning patterns that derive from long-term adaptation to complex environments, i.e., learning how to learn in meta-learning [3]. Motivated by this, we expect the DL model to become a generalized evolution model with both FSCIL, OSR, and lifelong learning capabilities like humans through continuous environment simulation.

In this paper, we propose a meta-learning-based few-shot open-set incremental learning (FSOSIL) for RF fingerprint recognition (*Meta-RFF*) framework. We first propose a signal augmentation scheme for few-shot samples. Second, based on the meta-learning idea, we define the FSOSIL meta-tasks and sampled a large number of pseudo-tasks from the training set to realize environment simulation. Further, we utilize the meta-learning technique for multi-task training to make the neural network adaptive in such an environment. Finally, to solve the

FSOSR problem, we incorporate open loss in the meta-task. Our contributions can be summarized as follows:

- To the best of our knowledge, we are the first to formulate an FSOSIL framework for RFF recognition and realize a generalized RFF continuous learning system.
- We define the FSOSIL meta-tasks and design a multi-task training mechanism-based prototype network.
- We propose a soft orthogonalization loss and an open loss to achieve automatic calibration of prototype points and OSR.

2 Problem Definition

Few-Shot Open-Set Incremental Learning for RFF Recognition. In a traditional DL-based RFF recognition setting, the RF receiver (Rx) will get a series of signals $x_(n)$ from a set of RF transmitters ($Tx = \{Tx_1, Tx_2, ..., Tx_k\}$). We define the base session (0-th session) dataset consisting of k transmitters as $\mathcal{D}^0 = \{x_i, y_i\}_{i=0}^{n_0}$ with sufficient instances. Then we define the emerging unknown source signals training sets as $\{\mathcal{D}^1, ..., \mathcal{D}^b\}$ with limited instances of known classes, i.e., $\mathcal{D}^b = \{(x_i, y_i)|y_i \in \{\mathbb{Y}_1, ..., \mathbb{Y}_b\}\}_{i=1}^{n^b}$. At the same time, there are open sets of unknown classes in the environment, i.e., $\mathcal{D}^o = \{(x_i, y_i)|y_i \in \mathbb{Y}_o\}_{i=1}^{n^o}$. The $\mathbb{Y}_b$ is the label space of task b, and $\mathbb{Y}_b \cap \mathbb{Y}_o = \emptyset$. Then the n^b and n^o denotes the number of samples in $\mathcal{D}^b$ and $\mathcal{D}^o$, respectively. When facing a new dataset $\mathcal{D}^b$, a model should learn new classes while maintaining performance on old classes and rejecting unknown classes. The process can be formalized as the minimization of the expected risk overall seen and unseen classes:

$$\min_{\theta} \mathbb{E}_{\{x_i, y_i\} \sim \{\mathcal{D}^0, ..., \mathcal{D}^b, \mathcal{D}^o\}} [\ell(f_\theta^b(x_i; \mathcal{D}^b, \psi_{b-1}, W_{b-1}), y_i)]. \tag{1}$$

The model $f_\theta^0(x)$ comprises a embedding function $\psi(\cdot) : x \to \mathbb{R}^d$ and a linear classifier $W_0 = \{w_i\}_{i=0}^{i=|\mathbb{Y}_0|}$, i.e., $f_\theta^0(x) = W_0^T \psi(x)$ and $\mathbb{R}^d$ denotes the d-dimensional feature space. By Eq. 1, the model $f_\theta^{b-1}(\bullet)$ should construct the new model based on the new dataset $\mathcal{D}^b$ and the current model W_{b-1}, ψ_{b-1}. Then in real-world testing, we expect that the newly constructed model $f_\theta^b(\bullet)$ to minimize the loss over all known and unknown test dataset.

3 Methodology

Fig. 3 shows the pipeline of *Meta-RFF* framework for solving the FSOSIL task. Specifically, the proposed *Meta-RFF* framework can be separated into three stages: **1) Feature Pre-training stage, 2)Meta-Task sampling stage, and 3) Meta-Incremental training stage**. In the feature pre-training stage, the feature embedding networks and classifier weights (i.e., prototype points) are obtained using base session data. Subsequently, we will construct few-shot incremental recognition scenarios and perform task sampling. In the training phase, a prototype network is utilized for few-shot learning, and a Transformer is used to calibrate the distribution of new and old class prototype points in an orthogonality prototype space. Finally, an open loss is optimized to achieve few-shot open-set recognition.

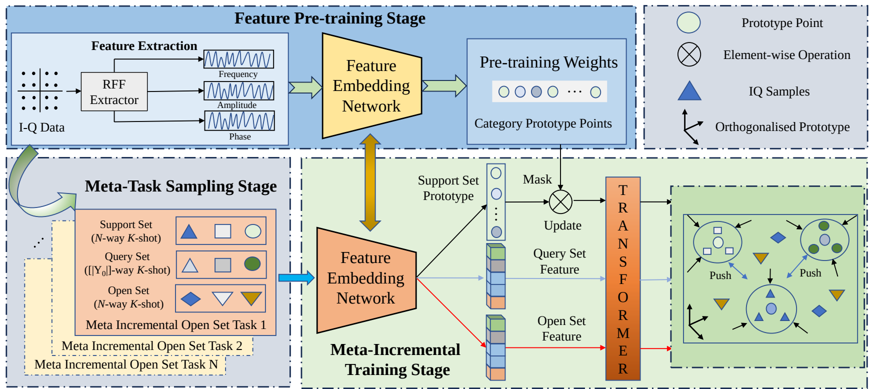


Fig. 3. The workflow of *Meta-RFF* framework.

3.1 Feature Pre-training Stage

Prior incremental learning approaches have demonstrated that fine-tuning the network with new class data from a subsequent session can result in catastrophic orgetting [6], wherein previously acquired knowledge is lost and overfitting transpires with the introduction of new data. Recent advances in incremental learning [21, 20] suggest that decoupling the feature embedding network from the classifier can largely reduce the effect of catastrophic forgetting. Thus, we follow the previous work and additionally employ the classic prototype network [16] (widely used under the few-shot learning scenario) to alleviate the catastrophic forgetting and overfitting problems under the few-shot sample condition.

Feature Extraction.. To improve the recognition efficiency of I-Q data with few shots, we need to extract more signal modal information. In wireless communication, IQ signals are usually defined by using amplitude, frequency, and phase [9]. Therefore, we follow the setting of Zheng et al. [23] and mainly extract the instantaneous amplitude $A(n)$, instantaneous phase $\varphi(n)$, and instantaneous frequency $F(n)$ information of the signal. The relevant calculations are shown below

$$\begin{aligned}
&A(n) = \sqrt{x_I(n)^2 + x_Q(n)^2},\\
&\varphi(n) \propto \arctan(x_Q(n)/x_I(n)),\\
&F(n) = \varphi(n) - \varphi(n-1), n = 1, 2, ..., N-1,\\
&\mathcal{F}(n) = F(n) - \tfrac{1}{N}\textstyle\sum_{n=1}^{N} F(n),
\end{aligned} \tag{2}$$

where $\mathcal{F}(n)$ denotes the centered instantaneous frequency, and the details of $\varphi(n)$ refer to [23]. For easier representation, we redefine the sample symbols as

$$x(n) = concat\{x_{IQ}(n), A(n), \varphi(n), \mathcal{F}(n)\}. \tag{3}$$

Model Pre-training.. Specifically, we first train the feature embedding network $\psi(\bullet)$ and classifier W_0 in the base session. The classifier weights w_i are represented by the average embedding of each new class c_i (i.e., the class prototype or the most representative feature of the class). The class prototype w_j in

$\mathcal{D}^0$ can be calculated by:

$$w_j = \frac{1}{|\mathcal{D}^0|} \sum_{i=1}^{|\mathcal{D}^0|} I(y_i = j)\psi(x_i), \tag{4}$$

where $I(\cdot)$ denotes the indicator function. With the class prototype points w_i, we can calculate the class probability p_j for each base session sample as

$$p(y = j|x; \psi_0) = \frac{\exp(sim(\psi(x), w_j))}{\sum_{j \in Y_0} \exp(sim(\psi(x), w_j))}, \tag{5}$$

where $sim(\bullet)$ denotes the cosine similarity function. The Eq. 5 suggests that the similarity of the sample to the class prototype point determines the sample prediction category. Finally, we using the cross-entropy loss function [22] to perform with training on the base session samples:

$$\ell_{ce} = -\frac{1}{|\mathcal{D}^0|} \sum\nolimits_i \sum\nolimits_{j=1}^{Y_0} y_{ij} \log(p_{ij}). \tag{6}$$

During the base session pre-training, a network with robust feature extraction capabilities is obtained by leveraging a substantial amount of available sample data. Then, when a new task arrives, the parameters of the feature embedding network are frozen to prevent knowledge forgetting, and the prototype points of the new class are computed to update the classifier.

3.2 Multi-phase Meta Task Sampling

The generalization ability of the learned features largely affects the performance of the incremental session. Under the FSOSIL task, this effect is magnified, as the model must generalize to new classes with a limited number while maintaining the ability to reject unknown classes. However, the model does not have direct access to new class data and unknown class data in the incoming incremental sessions, making it challenging to evaluate the generalization ability of learned features for future tasks. Thus, motivated by the idea of meta-learning, we propose to sample a large number of "fake" FSOSIL tasks from the base session data to simulate the procedure of real FSOSIL tasks. The sampled "fake" FSOSIL tasks aim to provide a way for the neural network to learn generalizable embeddings.

To make the "fake" FSOSIL tasks share the same data format as the 'real' FSOSIL task, we first divide the base dataset $\mathcal{D}^0$ into three non-overlapping sets: $T = \{S, Q^*, Q^o | C^s, C^*, C^o\}$, where $S = \{x_i, y_i\}_{i=1}^{NK}$ denotes support set with N-way K-shot and the label is C^s, $Q^* = \{x_i, y_i\}_{i=1}^{|Y_0|K}$ denotes the many-shot query set with $|Y_0|$-way K-shot and the label is C^*, Q^o denotes the open set with N-way K-shot. Also, the label spaces of Q^o and Q^* are non-overlapping, i.e., $C^* \cap C^o = \emptyset$, $C^s \subset C^*$. After the "fake" FSOSIL task sampling, the training objects can be formulated as the minimization of the empirical risk on the m-th task T_m:

$$\min_\theta \mathbb{E}_{\{x_i, y_i\} \sim \{Q_m^s, Q_m^o, Q_m^*\}} [\ell((x_i; S_m, \psi_{m-1}, W_{m-1}), y_i)]. \tag{7}$$

3.3 Meta-Training for FSOSIL

To facilitate the training process of the established "fake" FSOSIL tasks, we propose an optimization strategy that simultaneously trains the network to extract robust features for both incremental recognition and OSR tasks.

Optimization for few-shot incremental recognition.. During the incremental session, the classification model will continuously receive new session tasks, which require the network to generalize across both new and old class samples. To emulate the incremental session, we utilize the support set S, and the all class query set Q^* from the base session to represent new class incremental samples, and all class samples, respectively. During the optimization process, the network first generates new prototypes c_i for the few-shot support set in the incoming task using Eq. 5, and then directly employs the pre-trained old class weights w_i to recognize the old classes. Thus, the classifier weights W_{T_m} for task m will be updated by:

$$W_{T_m} = \begin{cases} c_i, \forall i \in C^s \\ w_i, \forall i \notin C^s, i \in C^* \end{cases} . \tag{8}$$

By jointly constructing new prototype points and optimizing through the similarity-based cross-entropy loss function, we can obtain the optimization function for the incremental session:

$$L_{T_m} = \ell_{ce}(sim(f(Q_m^*), W)). \tag{9}$$

Finally, based on idea of meta-learning, we need to optimize for a large number of "fake" few-shot incremental tasks at the same time. So that the neural network can learn how to adapt to this context. This meta-learning optimization loss can be shown by the following

$$\min_{\theta} \frac{1}{T_m} \sum\nolimits_{m=1}^{T_m} L_{T_m}(Q_m|S_m, f_\theta, W_m). \tag{10}$$

Meta-calibration module.. The incremental optimization is based on the many-shot old classes, which are tailored to depict old class features. To calibrate the semantic gap between new and old class prototypes (i.e., updating the relative spatial distribution), we design a transformer-based calibration module capable of extracting inductive bias during meta-training and generalizing to subsequent incremental sessions.

We define the adaptation function $\mathcal{T}(\cdot)$ by using the transformer to carry out the calibration process. The transformer utilizes a triplet of information (query $\mathcal{Q}$, key $\mathcal{K}$, and value $\mathcal{V}$) and learns through the attention mechanism. First, the query sample ($\psi(x_m) \in \mathbb{R}^{d\times|\mathcal{M}|}, m \in \mathcal{M}$) is projected linearly using weights W_q, W_k, and W_v respectively. Next, the attention coefficients α_{qk} are computed using $\mathcal{Q}$, $\mathcal{K}$, and softmax functions. Finally, the attention coefficients are applied to weight $\mathcal{V}$, obtaining the final attention result in $\psi(x_m)'$. In our framework, we drop query samples into $\mathcal{T}(\cdot)$ for adaptive optimization along with prototype points, i.e.,

$$\mathcal{Q} = \mathcal{K} = \mathcal{V} = [w, c]. \tag{11}$$

The calibrated prototype point are then defined as $(\hat{w}, \hat{c}) = \mathcal{T}(w, c)$.
Optimization for few-shot open-set recognition.. While the aforementioned approach enables the recognition of seen classes, Eq. 9 fails to effectively constrain the open space, resulting in limited open-set recognition capabilities. To address this issue, we propose simulating the open-set scenario in the meta-task to optimize open-set recognition ability.

The main idea of OSR is to control inter-class distances, i.e., the new class prototype adapts its open space distribution concerning the old class prototypes. We introduce an orthogonalization loss in the inner product space that all prototypes are orthogonalized vectors the cosine similarity is 0, i.e., $||w_i||^T ||w_j|| = 0, i \neq j, \forall i, j \in C$. Then, to make the prototype points of the new class orthogonal to the old class, we impose the orthogonalization loss ℓ_{or} in the meta task, i.e.,

$$\begin{aligned} &\ell_{or}(\hat{W}) = M \odot ||\hat{W}||^T ||\hat{W}||, \\ &M_{ij} = \begin{cases} 0, i = j \\ 1, i \neq j \end{cases}, \end{aligned} \tag{12}$$

where $M \in \mathbb{R}^{|Y_0| \times |Y_0|}$ denotes the mask matrix, $\odot$ denotes the element-wise multiplication, $|| \bullet ||$ denotes the ℓ_2-norm.

After controlling inter-class distances, we need to further control intra-class distances. We use the idea of clustering samples with their class prototype points to progressively reduce the differences. This process can be implemented by the following equation

$$\ell_d(f_\theta(x(n)), \hat{W}) = \ell_e(f_\theta(x(n)), \hat{W}) - \ell_c(f_\theta(x(n)), \hat{W}), \tag{13}$$

where ℓ_e denotes the euclidean distance, ℓ_c denotes the cosine similarity.

In summary, we combine the Eq. 9, 12, 13 and the final meta-task loss is

$$L_{T_m} = \ell_{ce}(f(Q_m^*, \hat{W})) + \alpha \ell_{or}(\hat{W}) + \beta \ell_d(f_\theta(Q_m^*), \hat{W}), \tag{14}$$

where α and β denote the hyper-parameters.

4 Evaluation

4.1 Experiment Setup

Experimental Environment. The *Meta-RFF* algorithm is implemented by using Pytorch 1.10.0 and NVIDIA GeForce RTX3090 GPUs. In addition, we use 1D ResNet-18 as the backbone network both for our algorithms and other baseline algorithms.
Dataset. We use real-world ADS-B signals to verify FSOCIL methods for wireless device identification. For the FSOCIL task setting, we extracted IQ samples for a total of 893 aircraft in the dataset. Aircraft categories ranging from 0-237 possess many shots, which can be used as the base session dataset. 200-437 range has relatively few samples, which can be used as the few-shot incremental session. The samples in the range of 437-893 are scarce and not easy to train, and this paper serves as an open unknown class. This data set is publicly available at [11].

Models' setup and comparison baselines. To ensure the fairness of the comparison, we use a uniform network architecture for all baseline methods. In addition, we use standard transformer [18] as a calibration module and adaptive generator. To evaluate the effectiveness of our proposed *Meta-RFF* on the FSOSIL task, we first compare to current state-of-the-art FSOSR algorithms, *e.g.,* Softmax [5], ARPL [2] and PEELER [10]. Besides, we also compare to current state-of-the-art FSCIL algorithms: iCaRL [15], CASTLE [20], CEC [22] and LIMIT[24].

4.2 Experimental Results

Comparison of Closed-Set Performance among Baselines.. As shown in Table 1, we compare the accuracy of closed-set data with mainstream algorithms. We divided the test dataset into base set data and incremental data. In traditional supervised learning algorithms (Softmax, ARPL), we can see that the lack of incremental learning capability of the model produces a catastrophic forgetting problem, with test accuracies for the base set and incremental classes of data almost approaching 0. In addition, although the classical incremental learning algorithm iCaRL guarantees a certain base set accuracy, it cannot adapt to few-shot incremental environments and its performance on incremental data is poor. In contrast, the rest of the algorithms that impose a few-shot incremental mechanism not only solve the catastrophic forgetting problem but also achieve good results on incremental datasets. However, we can see that the proposed *Meta-RFF* achieves state-of-the-art performance on both base and incremental data. The above results show that our *Meta-RFF* can significantly improve few-shot incremental performance by imposing RF data augmentation, multi-stage meta-task training, and soft orthogonalization loss.

Table 1. Comparison with the state-of-the-art on ADS-B dataset. "*" denotes the use of a fine-tuning operation in the algorithm, "†" denotes the use of CATSTLE in the algorithm, and PD denotes the performance dropping rate.

Methods	**Test set**	**Accuracy in each Session with 1-shot (%) ↑**											**PD↓**
		0	1	2	3	4	5	6	7	8	9	10	
Softmax*	Base	98.76	81.44	9.09	8.75	4.53	0.48	1.21	2.57	1.00	0.57	0.09	-98.67%
	Incremental	-	0.00	0.00	0.53	1.35	0.35	0.19	0.20	0.59	1.48	0.54	None
ARPL*	Base	97.32	90.90	27.62	47.03	19.08	27.32	19.79	14.08	9.10	5.42	6.09	-91.23%
	Incremental	-	0.00	0.00	0.02	0.36	0.24	0.43	0.69	1.05	1.27	0.71	None
PEELER†	Base	99.20	93.90	93.82	93.72	93.62	93.53	88.13	88.07	88.02	70.35	70.28	-28.92%
	Incremental	-	49.52	42.82	41.90	40.03	38.29	36.20	35.36	34.66	27.87	26.40	-23.12%
iCaRL	Base	99.86	91.43	90.19	88.89	87.98	86.57	84.43	83.00	80.34	78.15	76.79	-23.07%
	Incremental	-	1.51	11.56	5.16	6.28	5.70	6.02	3.83	4.08	3.78	4.35	None
CATSTLE	Base	99.32	98.32	98.25	98.23	98.20	98.18	98.12	98.10	98.08	97.99	97.96	-1.36%
	Incremental	-	55.38	50.30	46.23	44.30	42.35	42.73	42.01	41.27	40.40	40.66	-14.72%
CEC	Base	98.76	97.04	96.94	96.90	96.85	96.77	96.56	96.52	96.45	96.38	96.34	-2.42%
	Incremental	-	58.33	49.77	44.83	44.10	42.95	43.26	42.35	41.98	41.84	42.42	-15.91%
LIMIT	Base	99.96	98.63	98.15	97.61	97.17	96.84	96.52	96.12	95.88	95.49	95.23	-4.73%
	Incremental	-	70.08	67.73	63.82	59.04	57.58	57.95	55.35	54.60	52.92	52.63	-17.45%
Meta-RFF (1-shot)	Base	99.86	**99.82**	**99.79**	**99.78**	**99.77**	**99.76**	**99.76**	**99.70**	**99.68**	**99.67**	**99.64**	**-0.22%**
	Incremental	-	**80.17**	**79.17**	**78.20**	**78.21**	**76.42**	**76.62**	**76.98**	**77.70**	**76.92**	**77.03**	**-3.14%**
Meta-RFF (5-shot)	Base	99.86	**99.83**	**99.76**	**99.75**	**99.73**	**99.72**	**99.70**	**99.65**	**99.60**	**99.55**	**99.54**	**-0.32%**
	Incremental	-	**94.65**	**92.89**	**91.10**	**90.88**	**89.88**	**89.95**	**91.54**	**91.49**	**91.22**	**91.57**	**-3.08%**

Table 2. Comparison of few-shot incremental open-set recognition performance and state-of-the-art algorithms. "*" denotes the use of a fine-tuning operation in the algorithm, "†" denotes the use of CATSTLE in the algorithm.

Methods	FSOSR	FSCIL	AUROC in each session with 1-shot ↑										
			0	1	2	3	4	5	6	7	8	9	10
Softmax*	-	-	0.9546	0.8697	0.6543	0.6116	0.5398	0.5426	0.5881	0.5375	0.4555	0.5569	0.5475
ARPL*	✓	-	0.9564	0.9054	0.7164	0.7593	0.6765	0.6698	0.6848	0.6413	0.6392	0.6057	0.5795
PEELER†	✓	-	0.9414	0.9341	0.9278	0.9229	0.9185	0.9141	0.8918	0.8878	0.8837	0.8330	0.8287
iCaRL	-	✓	0.9090	0.8490	0.8250	0.7940	0.7802	0.7642	0.7422	0.7321	0.7221	0.7071	0.7011
CATSTLE	-	✓	0.7328	0.7308	0.7488	0.7497	0.7629	0.7653	0.7730	0.7807	0.7816	0.7767	0.7652
CEC	-	✓	0.7319	0.7265	0.7940	0.8010	0.8060	0.8073	0.8098	0.8166	0.8165	0.7990	0.7842
LIMIT	-	✓	0.9787	0.9666	0.9534	0.9409	0.9298	0.9174	0.9122	0.9015	0.8984	0.8892	0.8834
Meta-RFF	✓	✓	**0.9965**	**0.9921**	**0.9867**	**0.9803**	**0.9755**	**0.9715**	**0.9672**	**0.9645**	**0.9622**	**0.9561**	**0.9529**

Comparison of Open-Set Performance among OSR Baselines.. Table 2 shows the AUROC performance for the few-shot incremental open-set recognition task, where higher AUROC means better open-set recognition performance. We can see that our algorithm not only achieves state-of-the-art performance over multiple incremental sessions but also possesses a low-performance decay rate. This result is mainly because we impose open-set loss in multi-stage meta-task training, which reduces the inter-class differences. In addition, we normalize the distance between prototype points using soft orthogonalization loss, which allows new prototype points to be automatically separated from the old ones and adjust the distribution during a continuous incremental process. This mechanism allows our algorithm to adapt to multiple incremental processes, yielding a low AUROC decay rate.

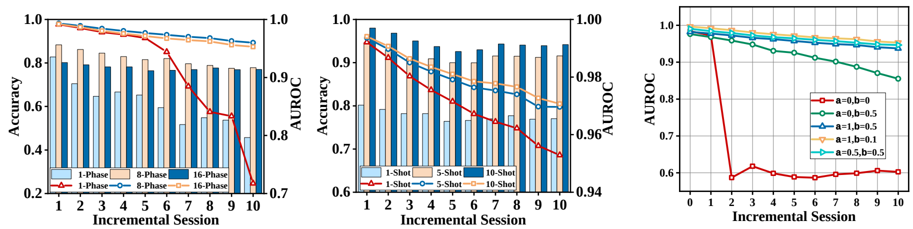
4.3 Ablation Study

Effects of the number of phases and shots.. Fig. 4a illustrates the performance impact of the number of phases on FSCIL as well as OSR in a continuous incremental session. In this figure, the line graph represents AUROC and the bar graph represents accuracy. We can observe that under 1-phase, the model's classification accuracy and open-set performance decline significantly in incremental sessions, which implies that adding more meta-task phases improves the model's ability. Similarly, in Fig. 4b, increasing the sample size of the incremental class will be able to further improve the model classification ability and open-set recognition ability.

Effect of hyper-parameters.. Fig. 4c evaluates the effect of parameters α and β in the model regularization term on OSR, where α denotes the prototype orthogonalization loss and β denotes the open loss in Eq. 14. When $\alpha = 0, \beta = 0$, we can observe clearly that the AUROC of OSR performance drops drastically in continuous incremental sessions. The AUROC of the model improves significantly when we apply just $\beta = 0.5$, indicating that reducing the distance between the samples and the prototype points can reduce the area of overlap with the unknown samples. Our model achieved optimal AUROC performance when we parameterized both regularization terms α, β. In particular, modifying the parameter values of α and β produces AUROC performance that is close.

5 Conclusion and Future Work

In this paper, we propose a continuous evolutionary algorithm for RFF recognition models that implements FSOSIL tasks. Our algorithm incorporates a meta-



(a) (b) (c)

Fig. 4. Accuracy and AUROC performance comparison of different parameters. (a)Effect of training phases. (b)Effect of shots. (c)Effect of hyper-parameters.

learning technique that utilizes many FSOSIL tasks constructed and performs meta-training to adapt to the FSOSIL scenarios gradually. In particular, to improve the performance of open-set recognition during a continuous incremental process, we introduce open loss as well as an adaptive open-set threshold generation technique based on reciprocal points. In the future, for the RFF recognition model, we will further construct the signal large model to realize a more general recognition system.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grants 62072408 and 62073292, the Zhejiang Provincial Science Fund for Distinguished Young Scholars under Grant LR24F020004, Zhejiang Provincial Natural Science Foundation of Major Program (Youth Original Project) under Grant LDQ24F020001, China Postdoctoral Science Foundation under Grant 2023M743403, and NSFC under grant U23A20326.

References

1. Bhatti, F.A., Khan, M.J., Selim, A., Paisana, F.: Shared spectrum monitoring using deep learning. IEEE Transactions on Cognitive Communications and Networking **7**(4), 1171–1185 (2021)
2. Chen, G., Peng, P., Wang, X., Tian, Y.: Adversarial reciprocal points learning for open set recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(11), 8065–8081 (2021)
3. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: International conference on machine learning. pp. 1126–1135 (2017)
4. He, B., Wang, F.: Cooperative specific emitter identification via multiple distorted receivers. IEEE Transactions on Information Forensics and Security **15**, 3791–3806 (2020)
5. Hendrycks, D., Gimpel, K.: A baseline for detecting misclassified and out-of-distribution examples in neural networks. arXiv preprint arXiv:1610.02136 (2016)
6. Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al.: Overcoming catastrophic forgetting in neural networks. Proceedings of the national academy of sciences **114**(13), 3521–3526 (2017)
7. Li, H., Tang, Y., Lin, D., Gao, Y., Cao, J.: A survey of few-shot learning for radio frequency fingerprint identification. In: Artificial Intelligence for Communications and Networks: Third EAI International Conference. pp. 433–443 (2021)

8. Li, T., Hong, Z., Cai, Q., Yu, L., Wen, Z., Yang, R.: Bissiam: Bispectrum siamese network based contrastive learning for uav anomaly detection. IEEE Transactions on Knowledge and Data Engineering (2021)
9. Li, T., Wen, Z., Long, Y., Hong, Z., Zheng, S., Yu, L., Chen, B., Yang, X., Shao, L.: The importance of expert knowledge for automatic modulation open set recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(11), 13730–13748 (2023)
10. Liu, B., Kang, H., Li, H., Hua, G., Vasconcelos, N.: Few-shot open-set recognition using meta-learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8798–8807 (2020)
11. Liu, Y., Wang, J., Niu, S., Song, H.: Ads-b signals records for non-cryptographic identification and incremental learning. (2021). https://doi.org/10.21227/1bxc-ke87, https://dx.doi.org/10.21227/1bxc-ke87
12. Polak, A.C., Goeckel, D.L.: Wireless device identification based on rf oscillator imperfections. IEEE Transactions on Information Forensics and Security **10**(12), 2492–2501 (2015)
13. Rajendran, S., Sun, Z.: Rf impairment model-based iot physical-layer identification for enhanced domain generalization. IEEE Transactions on Information Forensics and Security **17**, 1285–1299 (2022)
14. Ravi, S., Larochelle, H.: Optimization as a model for few-shot learning. In: International conference on learning representations (2016)
15. Rebuffi, S.A., Kolesnikov, A., Sperl, G., Lampert, C.H.: icarl: Incremental classifier and representation learning. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 2001–2010 (2017)
16. Snell, J., Swersky, K., Zemel, R.: Prototypical networks for few-shot learning. Advances in neural information processing systems **30** (2017)
17. Subramani, J., Maria, A., Neelakandan, R.B., Rajasekaran, A.S.: Efficient anonymous authentication scheme for automatic dependent surveillance-broadcast system with batch verification. IET Communications **15**(9), 1187–1197 (2021)
18. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
19. Wu, Z., Wang, F., He, B.: Specific emitter identification via contrastive learning. IEEE Communications Letters **27**(4), 1160–1164 (2023)
20. Ye, H.J., Hu, H., Zhan, D.C.: Learning adaptive classifiers synthesis for generalized few-shot learning. International Journal of Computer Vision **129**, 1930–1953 (2021)
21. Zhang, C., Cai, Y., Lin, G., Shen, C.: Deepemd: Few-shot image classification with differentiable earth mover's distance and structured classifiers. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 12203–12213 (2020)
22. Zhang, C., Song, N., Lin, G., Zheng, Y., Pan, P., Xu, Y.: Few-shot incremental learning with continually evolved classifiers. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 12455–12464 (2021)
23. Zheng, S., Zhou, X., Zhang, L., Qi, P., Qiu, K., Zhu, J., Yang, X.: Toward next-generation signal intelligence: A hybrid knowledge and data-driven deep learning framework for radio signal classification. IEEE Transactions on Cognitive Communications and Networking **9**(3), 564–579 (2023)
24. Zhou, D.W., Ye, H.J., Ma, L., Xie, D., Pu, S., Zhan, D.C.: Few-shot class-incremental learning by sampling multi-phase tasks. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(11), 12816–12831 (2022)