An Adversarial Margin Driven Rectified Prototypical Learning Method for Few-shot Incremental Automatic Modulation Recognition

Zhenxi Zhang, Haoyue Tan, Peiru Li, Heng Zhou, *Member, IEEE*, Xiaoran Shi, *Member, IEEE*, Yu Li, Yun Lin, *Senior Member, IEEE*, Feng Zhou, *Member, IEEE*

***Abstract*—Automatic modulation recognition (AMR) has been proven to be a critical component in wireless communication systems. However, the growing diversity of modulation schemes in dynamic environments, coupled with the scarcity of labeled training data, poses significant challenges for few-shot incremental AMR (FSI-AMR) in dynamically changing wireless communication environment. To address this issue, FSI-AMR offers a promising solution. In this work, we introduce an adversarial margin-guided rectified prototypical learning (AMG-RPL) framework, which effectively alleviates catastrophic forgetting of base modulation classes while improving adaptability to novel modulation types. The proposed approach integrates an adversarial margin learning mechanism with dual classifiers during base training, enabling the extraction of transferable modulation-invariant features while retaining discriminative modulation-specific characteristics. Furthermore, a rectified prototype learning strategy is designed to refine class prototypes in incremental sessions, minimizing cognitive bias and facilitating rapid few-shot adaptation. Comprehensive experimental evaluations demonstrate that AMG-RPL outperforms existing advanced methods in FSI-AMR tasks.**

***Index Terms*—Few-shot incremental automatic modulation recognition, Adversarial margin, Rectified prototypical learning.**

I. Introduction

The rapid detection and identification of wireless signals [1] have been extensively applied in cognitive radio domains [2, 3] such as spectrum interference monitoring, dynamic spectrum sensing, and communication resource optimization. In this context, Automatic modulation recognition (AMR) serves as a pivotal component in modern wireless communication systems [4, 5], which is critical for maintaining spectrum awareness in highly dynamic 6G environments [6, 7].

Deep learning (DL) techniques have significantly improved AMR performance by robust feature representation capabilities and advanced end-to-end learning way [8–10]. Most DL-based AMR methods [4, 5, 8] focus on closed-set scenarios, assuming that all modulation types during the test phase are known in the training phase. However, emerging communication systems frequently introduce novel modulation schemes, necessitating retraining of existing models to handle both known and unknown types [11]. However, the closed-set models are unable to recognize unknown modulation types, necessitating research into incremental AMR methods.

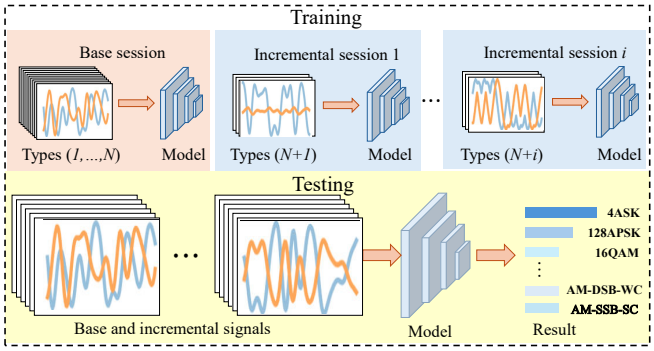

Fig. 1. The flow of FSI-AMR task.

Incremental AMR aims to maintain robust recognition performance for known modulation types while swiftly adapting to unknown types in dynamic wireless environments. A primary issue is catastrophic forgetting of base classes, where the model loses its ability to accurately identify known modulation types as it incorporates new ones, undermining stability. Additionally, when few-shot unknown modulation signals are available, the model tends to overfit, impairing its plasticity to new scenarios. These limitations collectively restrict the practical deployment of few-shot incremental AMR (FSI-AMR) in evolving communication systems, where balancing stability and adaptability is critical. While some incremental learning methods [12, 13] from computer vision have been exploblack, their direct application to radio signal processing often yields inadequate generalization due to the inherent heterogeneity between radio signals and 2D images. This mismatch hinders the model to capture class-common semantic patterns in signal data, ultimately degrading adaptability during few-shot incremental sessions.

To address these challenges, we propose an adversarial margin driven dual-classifier method to foster a equilibrium between stability and plasticity. Our method explicitly enhances transferability to dynamically changed communication system, maintaining adaptability to unknown conditions during

Zhenxi Zhang is with Xidian University, Xi'an 710071, China. Haoyue Tan, Xiaoran Shi, and Yu Li are with the Ministry of Education, Key Laboratory of Electronic Information Counter-measure and Simulation, Xidian University, Xi'an 710071, China. Heng Zhou is with the School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China. Yun Lin is with Harbin Engineering University, Harbin 150009, China. Feng Zhou is with the School of Aerospace Space and Technology, Xidian University, Xi'an 710071, China.

Corresponding author: Feng Zhou, e-mail:fzhou@mail.xidian.edu.cn

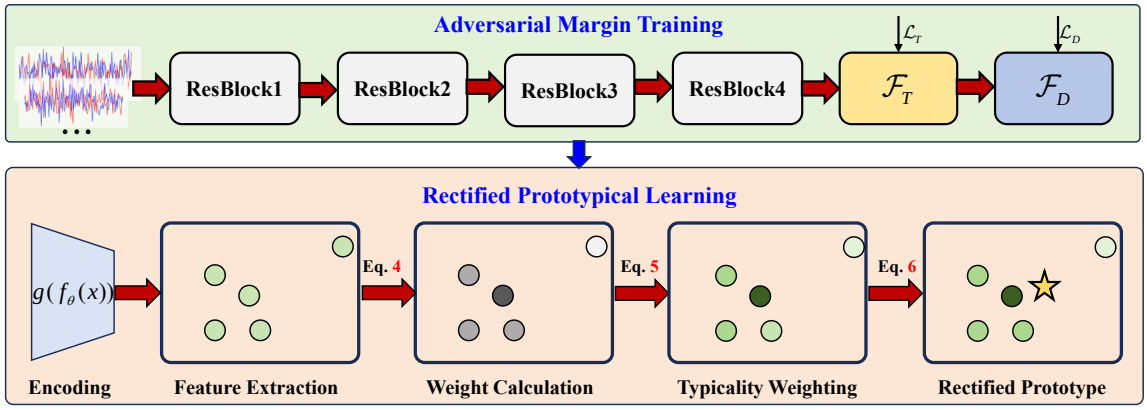

Fig. 2. The scheme of AMG-RPL. The green dots represent the feature vectors of the signals, the gray dots represent the calculated weights, and the yellow star denote the rectified prototype. The darkness of the color represents the weight magnitude.

the base-class training phase. Specifically, a negative margin is designed to strengthen the capacity for learning class-shared modulation features, thereby improving generalization to novel modulation types. Conversely, a positive margin preserves discriminative power for base classes by emphasizing class-specific features.

Moreover, acquiring sufficient labeled samples for novel modulation types is often impractical. Efficient continuous learning is critical for enabling communication systems to rapidly adapt to new modulation schemes with minimal supervision[14]. However, under few-shot conditions, incremental modulation types often suffer from intra-class variability, significantly degrading recognition performance. To reduce the model cognitive bias towards novel modulation types and ensure model forward compatibility, a rectified prototype learning method is designed based on importance re-weighting to correct class prototypes during incremental sessions. In summary, we propose an Adversarial Margin-Guided Rectified Prototypical Learning (AMG-RPL) framework for FSI-AMR in dynamic wireless communication systems. The main contributions include:

• To ensure class-shablack feature extraction for forward transferability and class-specific feature extraction for known discrimination, we design an adversarial learning method with dual classifiers to achieve a balanced trade-off.

• To alleviate the model cognitive bias towards novel modulation types and enhance its adaptability, we propose a rectified prototype learning method based on importance re-weighting to correct class prototypes during incremental sessions.

• AMG-RPL is evaluated on dynamic open scenarios, consisting of 7 base types and 7 novel types. Extensive comparative and ablation experiments demonstrate the superior performance of AMG-RPL.

II. Methodology

A. Problem Definition

In the FSI-AMR task, the training stage is divided into the base session and the few-shot incremental session. Let $\mathcal{D}_s^{\text{train}}$ and $\mathcal{D}_s^{\text{test}}$ denote the training set and the test set, respectively. s is the session index. The base training set $\mathcal{D}_0^{\text{train}}$ is formulated as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{n_0}$, which contains a relatively sufficient number n_0 of communication signals $\mathbf{x}_i$ and their labels $\mathbf{y}_i$ for modulation type. $\mathbf{x}_i$ is a training signal belonging to the class y_i. The base testing set $\mathcal{D}_0^{\text{test}}$ is defined as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{m_0}$. In practical applications of modulation recognition for communication signals, the novel modulation types often emerge incrementally with insufficient labeled instances, i.e., a sequence of datasets $\{\mathcal{D}_1, \cdots, \mathcal{D}_S\}$ will appear. $\mathcal{D}_s = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{NK}, s \geq 0$, where $\mathbf{y}_i \in Y_s$, and Y_s is the label space of the few-shot incremental AMR session s. It satisfies $Y_{s_1} \cap Y_{s_2} = \varnothing$ if $s_1 \neq s_2$. We only have access to the incremental signals in an N-way-K-shot format in the current session. When confronted with a novel signal dataset $\mathcal{D}_s$, the model must learn the novel modulation type while preserving its performance on the previously learned modulation types.

B. Adversarial Dual Classifiers

As shown in Fig. 2, the 1D version of ResNet18 is employed to extract the feature of the I/Q input $x \in \mathbb{R}^{2\times 1024}$. The network architecture retains the core principles of the original ResNet18, but adapts them for 1D input, replacing all 2D convolutions with 1D convolutions. x_i is first processed by a 1D convolutional layer with 16 filters with kernel size of 3, followed by a Batch Normalization (BN) layer. Then, the data is passed through four residual blocks. Each residual block consists of two consecutive 1D convolution layers with kernel sizes of 3 and a stride of 1, followed by batch normalization and ReLU activation. Residual connections are implemented to directly add the input of each block to its output, ensuring gradient flow during back-propagation. To generate a feature embedding of each signal, an adaptive average pooling is used after the final residual block to adjust the output size before passing it to the linear classification layer.

Different from previous works with only one classifier, we introduce two adversarial classifiers as illustrated in Fig. 3, the one is a transferable classifier $\mathcal{F}_T(\cdot)$ with a negative margin for the adaptation of novel modulation types using few-shot samples, the other is a discriminative classifier $\mathcal{F}_D(\cdot)$ with a positive margin for the discrimination of base modulation types in the base session. $\mathcal{F}_T(\cdot)$ performs the AMR task based

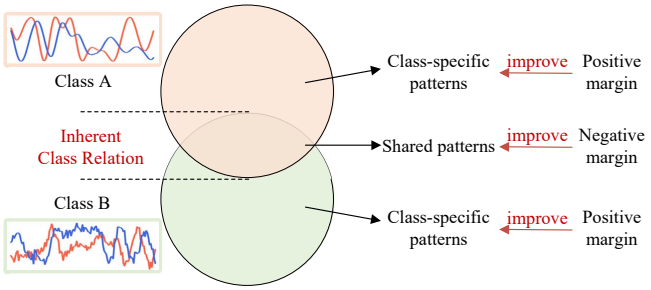

Fig. 3. The flowchart of the proposed adversarial margin-guided learning.

on the extracted feature $f_\theta(x)$. Then, $f_\theta(x)$ is passed through a fully-connected layer $g(\cdot)$ to generate the feature embedding for positive-margin classifier $\mathcal{F}_D(\cdot)$. Thus, $\mathcal{F}_D(\cdot)$ performs AMR task based on $g(f_\theta(x))$.

C. Adversarial Margin-guided Training

A common-used training loss function with an integrated margin for AMR at base session can be formulated as:

$$\mathcal{L}_{CE}\left(\mathbf{x}_i, \mathbf{y}_i\right) = -\log \frac{e^{\tau\left(w_{\mathbf{y}_i} f_\theta(\mathbf{x}_i) - m\right)}}{e^{\tau\left(w_{\mathbf{y}_i} f_\theta(\mathbf{x}_i) - m\right)} + \sum_{\mathbf{y}_j \neq \mathbf{y}_i} e^{\tau w_{\mathbf{y}_j} f_\theta(x_i)}} \tag{1}$$

where $f_\theta(\mathbf{x}_i) \in \mathbb{R}^D$ represents the feature embedding of the input signal by the feature extractor with parameters θ. $w_{y_i} \in \mathbb{R}^D$ denotes the linear classification weight for the true modulation type. $w_{y_j} \in \mathbb{R}^D$ denotes the linear classification weights for other modulation types in base session. m indicates a margin parameter integrated to the loss to adjust the inter-class distribution of different modulation types. τ denotes the temperature parameter which determines the amplified extent to the gap between the largest logit and the other logits. The previous AMR methods [4, 8] usually enable a positive margin between different modulation types. The large margin fosters the intra-class tightness and between-class discrimination for AMR on the close-set scenarios, thus improving the modulation recognition accuracy.

In open communication environment, being too discriminative of base modulation types can harm the model ability to transfer to novel modulation types during the incremental process. By employing a suitable negative margin in the softmax loss, we aim to balance the discriminative ability for learning modulation types in the base session and the transferability to novel modulation types in the incremental session. Thus, we propose an adversarial margin-guided learning method, which introduces a transferable learning loss $\mathcal{L}_T$ and a discriminative learning loss $\mathcal{L}_D$ as Eq. 2.

$$\begin{aligned}
&\mathcal{L}_{Adv}\left(\mathbf{x}_i, \mathbf{y}_i\right) = \mathcal{L}_T\left(\mathbf{x}_i, \mathbf{y}_i\right) + \mathcal{L}_D\left(\mathbf{x}_i, \mathbf{y}_i\right) \\
&= -\log \frac{e^{\tau\left(w_{\mathbf{y}_i}^T f_\theta(\mathbf{x}_i) + m_{neg}\right)}}{e^{\tau\left(w_{\mathbf{y}_i}^T f(\mathbf{x}_i) + m_{neg}\right)} + \sum_{\mathbf{y}_j \neq \mathbf{y}_i} e^{\tau w_j^T f(\mathbf{x}_i)}} \\
&- \lambda \cdot \log \frac{e^{\tau\left(w_{\mathbf{y}_i}^D g(f_\theta(\mathbf{x}_i)) - m_{pos}\right)}}{e^{\tau\left(w_{\mathbf{y}_i}^D g(f_\theta(\mathbf{x}_i)) - m_{pos}\right)} + \sum_{\mathbf{y}_j \neq \mathbf{y}_i} e^{\tau w_j^D g(f_\theta(\mathbf{x}_i))}}
\end{aligned} \tag{2}$$

where $w_y^D \in \mathbb{R}^D$ and $w_y^T \in \mathbb{R}^D$ denote the linear classification weight for $\mathcal{F}_D$ and $\mathcal{F}_T$ corresponding to class y, respectively. m_{neg} denotes the margin for $\mathcal{F}_T$, m_{pos} denotes the margin for $\mathcal{F}_D$. λ is set to 1. The transferable patterns $f(\cdot)$ with negative margin are used to foster the discriminative patterns $g(\cdot)$ and construct the classifier $\mathcal{F}_D$. Since $f(\cdot)$ emphasizes the shablack features among different modulation types, $\mathcal{F}_D$ is endowed with better adaptability and generalization for novel modulation types under few-shot labeling conditions.

D. Incremental Rectified Prototypical Learning

To blackuce the distortion in the feature space when learning a novel task, we calculate and store a prototype $\mu_{t,k}$ for the base modulation type by average calculation on the train set of base session as follows:

$$\mu_{t,k} = \frac{1}{N_{t,k}} \sum_{n=1}^{N_{t,k}} g(f_{\theta_t}\left(\mathbf{x}_{t,n}\right)), t = 0 \tag{3}$$

where $t = 0$ represents the session, k denotes the class index of modulation type. $N_{t,k}$ denotes the total number of signal of class k. n denotes the signal index. For each base modulation type, we do not save any signals. We memorize a class-level prototype to replace the weight of linear layer in $\mathcal{F}_D$ to mitigate the classifier bias.

However, in open-set scenarios, the novel modulation types often have only a very limited number of samples. Such less samples result in biased prototypes for the novel classes. To alleviate this issue, we propose a rectified prototypical learning method to blackuce the adverse effects of outlier samples on the class prototype representation, while alleviating the distortion of incremental modulation types and enhancing the influence of typical samples in the computation of class prototypes. Specifically, we calculate the sum of similarity between each signal and all other signals for each novel modulation type, which is defined as SA_i in Eq. 4.

$$SA_i = \frac{1}{N} \sum_{j=1}^{N} \left\| g(f_\theta\left(\mathbf{x}_i\right)) - g(f_\theta\left(\mathbf{x}_j\right)) \right\|_2 \tag{4}$$

Then, we use SA_i to derive a weight w_{SA}, which ensures that the signals with stronger typicality to obtain a larger weight in the calculation of class prototypes.

$$w_{SA_i} = \frac{\exp(-SA_i)}{\sum_{j=1}^{N} \exp(-SA_j)} \tag{5}$$

Then, the incremental rectified prototype $\mu_{t,k}$ of the modulation type k is calculated by Eq. 6:

$$\mu_{t,k} = \sum_{i=1}^{N} w_{SA_i} \cdot g(f(\mathbf{x}_i)), t \geq 1 \tag{6}$$

The rectified prototypes are served as classification weight for novel modulation type in $\mathcal{F}_D$.

III. EXPERIMENTS

A. Datasets and Evaluation

We validate the performance using the publicly available RML2018.01A dataset [15], which contains various digital and analog modulation schemes. Each IQ signal is represented by a 2×1024 array. The signal-to-noise ratio (SNR) spans from -18dB to 30dB in increments of 6dB. As detailed in Table I, following PASSNet, the base classes comprise 4ASK, 8ASK, 128APSK, 16QAM, QPSK, 8PSK, and OOK, with each modulation type containing 2,867 labeled signals. In contrast, the seven incremental sessions include FM, GMSK, AM-DSB-WC, BPSK, AM-SSB-SC, OQPSK, and 32APSK, where each session uses only 5 signals to train the model in incremental session. The base training set contains a sufficient number of labeled signals. During the test phase, both the base and novel classes are evaluated with 1229 test signals each modulation type. We employ average accuracy Acc_a and harmonic accuracy Acc_h to evaluate FSI-AMR task [11].

B. Training Details

At the base session, we train the AMG-RPL model for 60 epoches with batch size 128 and SGD optimizer,the initial learning rate is 0.002, which is scheduled to decrease periodically using a StepLR scheduler with a step of 20 and a factor of 0.1, helping the model converge more smoothly and effectively. We employ $\mathcal{L}_{Adv}$ to train the network. At the incremental session, we freeze the feature extraction network $f_\theta(\cdot)$, the fully connect layer $g(\cdot)$, and the classifier weights for the base modulation types. After we use the rectified prototype to initialize the classifier weights of the novel modulation types, we use $\mathcal{L}_{CE}$ to fine-tune them. And a SGD optimizer with an initialize learning rate of 0.005 is employed with 10 epochs in each incremental session. All experiments are implemented using PyTorch on a single NVIDIA GeForce RTX 4090 GPU. The model parameter count of AMG-RPL is 0.506 MB and the per-sample inference time is 7.899 ms.

C. Ablation Experiment

To gauge the impact of individual model components on the overall performance and the model stability for different environment. We carry out a series of ablation experiments.

1) Ablation study of different SNR levels: As demonstrated in Table I, as SNR increases, AMG-RPL exhibits an upward trend. When SNR$<$0dB, the power of the noise is greater than the power of the signal. Thus, the signal is overwhelmed by the noise, making it challenging to achieve accurate AMR. Consequently, the features of different modulation types overlap significantly, leading to performance degradation in AMR. In addition, the core innovation of AMG-RPL lies in adversarial margin learning to extract modulation-shared features by negative margin, primarily aimed at mitigating catastrophic forgetting in incremental learning. However, at low SNR, these shared features learned by our method become unclear and ineffective for distinguishability, as they attempt to capture the commonality of severely corrupted signals.

As SNR increases, AMG-RPL demonstrates strong capabilities in mitigating catastrophic forgetting of base classes and rapidly learning novel classes in incremental sessions. Specifically, when the SNR is 6dB, AMG-RPL maintains an average accuracy close to 96% in the first five incremental sessions. However, in the last two sessions, the model plasticity for novel classes decreases, with Acc_h being 88.64% and 75.77% in Session 6 and Session 7, respectively. These findings demonstrate that as the number of incremental sessions increases, the model faces challenges in terms of plasticity for novel classes. When SNR is 12dB, in the incremental Session 7, AMG-RPL achieves an Acc_a of 96.09% and an Acc_h of 84.30%. When SNR is 24dB, Acc_h reaches 87.66%, Acc_a attains to 96.24%. These findings indicate that as SNR increases, our method achieves improved performance in recognizing novel modulation types in incremental sessions.

2) Ablation study of adversarial margin learning: To verify the contribution of the proposed adversarial margin learning method, we perform ablation study as shown in Table II. In the first group, we use the common-used cross-entropy loss for base session training. In the second group, we introduce a positive margin of 0.2 for the base training, written as $\mathcal{L}_D$. Compablack with $\mathcal{L}_{CE}$, the introduction of $\mathcal{L}_D$ improves Acc_a of the base session and ensures a high Acc_h score in Session 1 to Session 6. However, the Acc_h of Session 7 has a large drop, with an Acc_h of only 55.68%. However, when the adversarial $\mathcal{L}_T$ is introduced with m_{neg} of 0.6, the Acc_h for Session 7 improved to 75.77%, indicating an improvement of the model sustained transferability. This confirms the effectiveness of the adversarial margin in maintaining a dynamic balance between the discriminability of base modulation types and the plasticity of incremental modulation types in FSC-AMR task.

3) Ablation study of prototype adaptation method: The bottom part of Table II presents an ablation study on the effect of introducing RPL in incremental learning, and also provides a comparison with the Bayesian updating method for updating prototypes. Specifically, the model without the RPL exhibits a gradual decline in Acc_h as the sessions progress, reaching 90.71% and 74.11% on Acc_h in Session 5 and 7, respectively. While Bayesian updating method achieves relatively high accuracy during initial incremental sessions, our RPL method exhibits superior performance in later sessions with Acc_h of 97.53% and 75.77% in Session 5 and Session 7, respectively. These findings highlight the effectiveness of the RPL method in mitigation of prototype shift in few-shot scenarios and enhancement of rapid FSI-AMR capability.

D. Comparison with Other FSI-AMR methods

As we aim to validate the capability to balance newly acquired knowledge with existing knowledge, we compare AMG-RPL with other FSI-AMR methods under the same experiment setup, including S3C [16], FACT [11], CGSCIL [17], iCARL [12], and AANet [13]. As illustrated in Fig. 4, when SNR falls below 0dB, iCARL yields relatively favorable results, achieving an Acc_a of 26.12% in Session 7, which surpasses other methods. AMG-RPL consistently demonstrates superior performance when SNR is above 0dB. Specifically, in Session 4, AMG-RPL outperforms the second best CEC by 7.41% on 0dB, and outperforms the second best FACT

TABLE I
THE RECOGNITION PERFORMANCE OF AMG-RPL 1-WAY 5-SHOT FSI-AMR TASKS UNDER DIFFERENT SNR LEVELS. WE REPORT Acc_a/Acc_h (%).

SNR/dB	Session 0	Session 1	Session2	Session 3	Session 4	Session 5	Session 6	Session 7
-18	6.18	12.41/6.26	11.15/8.99	9.86/4.74	8.92/8.97	8.11/3.09	7.69/10.65	7.23/8.15
-12	6.89	14.79/7.02	12.93/10.91	11.55/11.95	10.50/11.01	9.73/11.48	8.83/8.65	8.35/9.61
-6	22.68	29.62/27.14	26.09/23.01	24.03/22.71	21.95/22.17	20.47/24.77	18.77/9.47	17.26/15.73
0	66.45	70.05/78.59	70.56/73.68	72.57/75.96	70.15/63.41	66.71/59.20	62.91/42.79	60.05/49.91
6	95.06	96.50/97.72	95.75/95.39	96.07/97.40	96.15/97.01	96.29/97.53	94.86/88.64	91.46/75.77
12	98.70	99.38/99.56	99.45/99.56	99.50/99.56	99.51/99.43	99.10/96.81	98.78/97.38	96.09/84.30
18	98.77	99.39/99.57	99.45/99.53	99.50/99.57	99.39/98.83	99.33/98.91	99.12/98.20	95.69/81.44
24	99.15	99.61/99.70	99.64/99.69	99.67/99.69	99.61/99.19	99.44/98.45	98.94/97.04	96.24/87.66
30	99.21	99.35/99.55	99.32/99.48	99.37/99.47	99.38/99.19	99.25/98.41	99.17/98.71	96.71/86.52

TABLE II
ABLATION STUDY OF ADVERSARIAL MARGIN LEARNING AND RECTIFIED PROTOTYPICAL LEARNING(%)

Function	Session 0	Session 1	Session2	Session 3	Session 4	Session 5	Session 6	Session 7
$\mathcal{L}_{CE}$	90.51	91.50/94.82	89.95/84.68	90.95/94.68	91.01/90.74	91.67/94.59	91.31/82.18	87.77/71.44
$\mathcal{L}_D$	94.17	96.18/97.69	95.78/94.27	95.82/95.79	95.97/97.39	95.17/97.39	94.69/84.07	89.96/55.68
$\mathcal{L}_D + \mathcal{L}_T$	95.06	96.50/97.72	95.75/95.39	96.07/97.40	96.15/97.01	96.29/97.53	94.86/88.64	91.46/75.77
w/o rectified method	95.01	95.91/97.53	95.26/95.20	96.05/97.35	96.68/96.69	95.76/90.71	94.72/82.86	90.94/74.11
Bayesian updating	95.14	96.45/97.85	96.01/96.23	96.84/97.18	96.52/95.03	96.57/97.49	94.63/80.34	90.12/74.53
Our rectified method	95.06	96.50/97.72	95.75/95.39	96.07/97.40	96.15/97.01	96.29/97.53	94.86/88.64	91.46/75.77

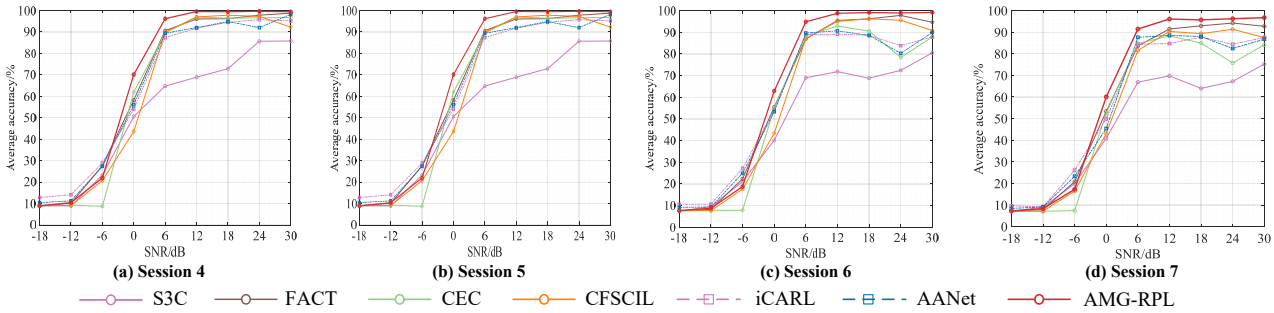

Fig. 4. The comparison with other competitive FSC-AMR methods. We show the FSI-AMR performance in Session 4, 5, 6 and 7 under different SNR levels.

by 5.92% on 6dB. When SNR is greater than or equal to 12dB, AMG-RPL achieves higher recognition accuracy than other methods. The similar experimental observations can also be found in Session 5 and Session 6. In Session 7, AMG-RPL improves the Acc_a performance by 3.86% and 4.60% compared to the second best AANet and FACT on 6dB and 12dB, respectively. These results demonstrate the superiority of AMG-RPL in mitigating catastrophic forgetting of base classes while enhancing the plasticity for novel classes.

IV. CONCLUSION

In this correspondence, we propose an AMG-RPL method to tackle FSI-AMR in dynamic and evolving environments, which effectively mitigates catastrophic forgetting while enhancing the model adaptability toward novel modulation types. By incorporating an adversarial margin learning framework with dual classifiers, AMG-RPL ensures a balanced extraction of class-shablack and class-specific features during the base training phase, thus enabling forward transferability and discriminative power in the incremental sessions. Additionally, the rectified prototype learning method, based on importance re-weighting, plays a key role in blackucing cognitive bias and facilitating fast learning in incremental sessions. Future work will explore the temporal dependencies in I/Q sequences for enhancement of unknown signal rejection under low-SNR.

REFERENCES

[1] Y. Liu, Y. Liu, and C. Yang, "Modulation recognition with graph convolutional network," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 624–627, 2020.
[2] Y. Qu, Z. Lu, and et al, "Contrastive language-signal prediction for automatic modulation recognition," *IEEE Wireless Communications Letters*, vol. 13, no. 11, pp. 3242–3246, 2024.
[3] Y. Lin, Y. Tu, Z. Dou, L. Chen, and S. Mao, "Contour stella image and deep learning for signal recognition in the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 34–46, 2020.
[4] H. Tan, Z. Zhang, Y. Li, X. Shi, and F. Zhou, "Multi-scale feature fusion and distribution similarity network for few-shot automatic modulation classification," *IEEE Signal Processing Letters*, vol. 31, pp. 2890–2894, 2024.
[5] Y. Qu, Z. Lu, R. Zeng, J. Wang, and J. Wang, "Enhancing automatic modulation recognition through robust global feature extraction," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 3, pp. 4192–4207, 2025.
[6] Y. Li, X. Shi, H. Tan, Z. Zhang, X. Yang, and F. Zhou, "Multi-representation domain attentive contrastive learning based unsupervised automatic modulation recognition," *Nature Communications*, vol. 16, no. 1, p. 5951, 2025.
[7] R. Liu, A. Jamalipour, and et al, "6g enabled advanced transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 10 564–10 580, 2024.
[8] Y. Lin, Y. Tu, and Z. Dou, "An improved neural network pruning technology for automatic modulation classification in edge devices," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5703–5706, 2020.
[9] W. Zhang, X. Yang, C. Leng, J. Wang, and S. Mao, "Modulation recognition of underwater acoustic signals using deep hybrid neural networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 5977–5988, 2022.
[10] W. Kong, X. Jiao, Y. Xu, and Q. Yang, "An effective masked transformer model for automatic modulation recognition," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 1, pp. 1–13, 2025.
[11] D.-W. Zhou, F.-Y. Wang, and et al, "Forward compatible few-shot class-incremental learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 9046–9056.
[12] S.-A. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert, "icarl: Incremental classifier and representation learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2001–2010.
[13] Y. Liu, B. Schiele, and Q. Sun, "Adaptive aggregation networks for class-incremental learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2544–2553.
[14] Z. Zhang, Y. Li, Q. Zhai, Y. Li, and M. Gao, "Few-shot learning for fine-grained signal modulation recognition based on foreground segmentation," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2281–2292, 2022.
[15] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.
[16] J. Kalla and S. Biswas, "S3c: Self-supervised stochastic classifiers for few-shot class-incremental learning," in *European Conference on Computer Vision*. Springer, 2022, pp. 432–448.
[17] M. Hersche, G. Karunaratne, A. Rahimi, and et al, "Constrained few-shot class-incremental learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 9057–9067.