Adaptive Decision Boundary for Few-Shot Class-Incremental Learning

Linhao Li[1*], Yongzhang Tan[1*], Siyuan Yang[2], Hao Cheng[1†], Yongfeng Dong[1], Liang Yang[1]

[1]School of Artificial Intelligence and Data Science, Hebei University of Technology, China
[2]College of Computing and Data Science, Nanyang Technological University, Singapore
{lilinhao, chenghao, dongyf}@hebut.edu.cn, siyuan.yang@ntu.edu.sg, 202232805041@stu.hebut.edu.cn, yangliang@vip.qq.com

Abstract

Few-Shot Class-Incremental Learning (FSCIL) aims to continuously learn new classes from a limited set of training samples without forgetting knowledge of previously learned classes. Conventional FSCIL methods typically build a robust feature extractor during the base training session with abundant training samples and subsequently freeze this extractor, only fine-tuning the classifier in subsequent incremental phases. However, current strategies primarily focus on preventing catastrophic forgetting, considering only the relationship between novel and base classes, without paying attention to the specific decision spaces of each class. To address this challenge, we propose a plug-and-play *Adaptive Decision Boundary Strategy* (ADBS), which is compatible with most FSCIL methods. Specifically, we assign a specific decision boundary to each class and adaptively adjust these boundaries during training to optimally refine the decision spaces for the classes in each session. Furthermore, to amplify the distinctiveness between classes, we employ a novel inter-class constraint loss that optimizes the decision boundaries and prototypes for each class. Extensive experiments on three benchmarks, namely CIFAR100, miniImageNet, and CUB200, demonstrate that incorporating our ADBS method with existing FSCIL techniques significantly improves performance, achieving overall state-of-the-art results.

Code — https://github.com/Yongzhang-Tan/ADBS

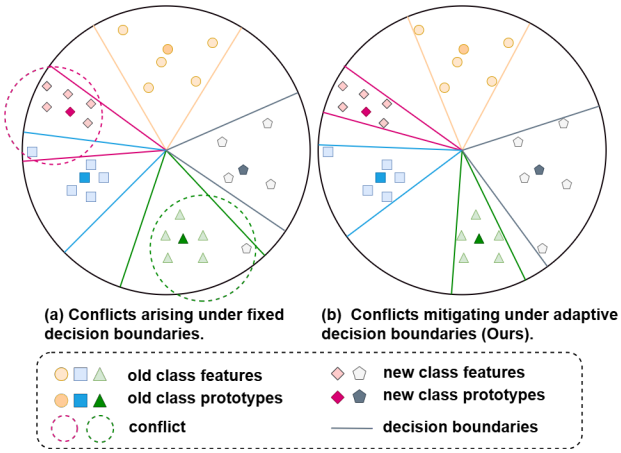


Figure 1: Illustration of FSCIL classification with (a) Fixed and (b) Our proposed adaptive decision boundaries. A fixed decision boundary strategy often struggles to reserve adequate space for new classes at each incremental stage, resulting in space conflicts between old and new classes. In contrast, our proposed adaptive decision boundary strategy can effectively alleviate this issue by adjusting the decision boundaries of both old and new classes.

Introduction

To date, deep Convolutional Neural Networks (CNNs) have achieved significant advancements in the field of computer vision, primarily using models trained on static and pre-collected large-scale annotated datasets (Deng et al. 2009; He et al. 2016). However, the practical challenge of gradually integrating data from new classes into a model initially trained on existing classes introduces significant obstacles, known as Class Incremental Learning (CIL) (Hou et al. 2019; Li and Hoiem 2017; Rebuffi et al. 2017; Cheng et al. 2025). Unlike standard classification tasks, CIL requires handling new classes and restricted access to prior task data in incremental sessions. Consequently, simply updating the model with new class data may lead to overfitting on these new classes and cause significant performance drops on previously learned classes, a phenomenon known as catastrophic forgetting (McCloskey and Cohen 1989; Masana et al. 2022). Nonetheless, CIL requires sufficient training data from novel classes. In many applications, data collection and labeling can be prohibitively expensive, posing challenges for implementing CIL in real-world scenarios. To address these challenges, Few-Shot Class-Incremental Learning (FSCIL) is proposed to address the dual challenges of learning new classes from limited examples and integrating this knowledge into an existing model without forgetting previously acquired information (Tao et al. 2020). Critically, the model must maintain a balance between stability and plasticity, preserving previously acquired knowledge while seamlessly integrating new information.

Recent advancements in FSCIL methods (Zhang et al.

*Linhao Li and Yongzhang Tan are co-first authors.
†Corresponding Author.

2021; Cheraghian et al. 2021; Dong et al. 2021; Zhao et al. 2021; Zhou et al. 2022a; Chen, Wang, and Hu 2022; Yang et al. 2023; Fan et al. 2024) have shown remarkable performance in image classification tasks. These methods typically employ a learning pipeline that includes pre-training the model during a data-rich base session, followed by merely constructing a classifier for novel classes on the frozen backbone during incremental sessions. While this approach effectively integrates new classes while retaining previously learned information, it still remains a significant challenge, *i.e.*, conflicts in the decision space between old and novel classes, as depicted in Figure 1(a). When the features of base classes are similar to those of novel classes, classifying novel classes may disrupt the classification of the pre-established base classes. Moreover, in scenarios with a fixed backbone, base classes typically do not reserve space for unknown novel classes, leading to performance degradation. Several approaches (Song et al. 2023; Zhou et al. 2022a) have attempted to address this issue by introducing virtual classes to reserve feature space for novel classes. However, it remains challenging to predict the feature space of novel classes that significantly deviate from base classes using only data augmentation-based strategies, *e.g.*, mixup. Additionally, finding a balance between the space reserved for base and novel classes to avoid compromising base classification tasks remains a significant issue. Previous studies (Peng et al. 2022; Zou et al. 2022; Guo et al. 2023) have attempted to introduce decision boundaries in FSCIL to improve inter-class cohesion. However, these methods, which rely on predetermined hyperparameters to set boundaries, often fail to accurately define precise demarcations. The aforementioned observations prompt us to construct adaptive decision boundaries to accurately predict base classes and reserve more space for incoming novel classes.

In this paper, we propose a novel *Adaptive Decision Boundary Strategy* (ADBS) for FSCIL, which assigns a specific decision boundary to each class that dynamically adjusts based on the training data. To enhance the distinctiveness between classes, we also implement a novel inter-class constraint loss that optimizes the decision boundaries and prototypes for each class. The proposed ADBS significantly improves class separation and effectively reduces conflicts in the feature space between existing and incoming classes, as shown in Figure 1(b). Furthermore, ADBS is a plug-and-play module that can be easily integrated with existing FSCIL frameworks without necessitating modifications to the network architecture.

To summarize, our main contributions are as follows:

- We introduce a plug-and-play *Adaptive Decision Boundary Strategy* (ADBS) designed to mitigate conflicts between base and novel classes in the feature space within FSCIL tasks. Our theoretical analysis verifies that this strategy effectively differentiates class centers and optimizes their boundaries.
- We employ an inter-class constraint to optimize the decision boundaries and prototypes for each class, further enhancing the distinguishability between classes.
- We evaluate the ABDS method over three FSCIL benchmarks: CIFAR100, *mini*ImageNet, and CUB200. Experimental results and visualizations demonstrate that incorporating ADBS with existing FSCIL algorithms, including both the baseline and state-of-the-art methods, consistently enhances performance and achieves state-of-the-art performance.

Related Work

Few-Shot Class-Incremental Learning

The concept of Few-Shot Class-Incremental Learning (FSCIL) was first introduced in TOPIC (Tao et al. 2020), which aims to address the dual challenges of catastrophic forgetting and learning new classes incrementally from a limited number of labeled samples. TOPIC tackled these issues by implementing a *neural gas* (NG) network. Current methods in FSCIL can be categorized into two main groups: one updates the backbone network during incremental sessions (Cheraghian et al. 2021; Dong et al. 2021; Kang et al. 2023; Li et al. 2024), while the others maintain a fixed backbone. The latter group attempts to suppress forgetting old knowledge while adapting smoothly to novel classes using various approaches. Several methods improve model performance by constructing diverse powerful classifiers, such as ETF (Yang et al. 2023) or stochastic classifiers (Kalla and Biswas 2022). Other methods (Zhou et al. 2022a; Song et al. 2023) introduce virtual classes to pre-allocate feature space for novel classes, while additional methods include episodic training (Zhou et al. 2022b; Zhu et al. 2021) or ensemble learning (Ji et al. 2023) to boost the capabilities of backbone. Recent methods (Yang, Liu, and Xu 2021; Liu et al. 2023) consider to employ distribution calibration to adjust the classifier. OrCo (Ahmed, Kukleva, and Schiele 2024) promotes class separation by leveraging feature orthogonality in the representation space and contrastive learning.

Boundary-based Method

Boundary-based methods, which focus on learning optimal decision boundaries, are extensively employed in various vision tasks such as image classification (Chen et al. 2020; Wang et al. 2024), semantic segmentation (Liu et al. 2022), and domain generalization (Dayal et al. 2024). The widespread use and success of these methods in diverse applications highlight their efficacy, showcasing their capability to manage complex visual data with exceptional accuracy and reliability.

Several methods also explore boundary strategies within the FSCIL domain. In prior studies, ALICE (Peng et al. 2022) utilizes angular penalty loss with hyperparameter-defined boundaries to enhance inter-class cohesion, yet these boundaries are not effectively utilized during inference. Conversely, CLOM (Zou et al. 2022) employs hyperparameter-based methods to establish positive and negative boundaries by considering distances between class prototypes, while DBONet (Guo et al. 2023) assumes that data feature vectors follow a spherical Gaussian distribution and employs within-class variance to define boundaries. Although these approaches promote class separation to some

degree, they often struggle to achieve precise boundary accuracy. Additionally, traditional boundary-based methods frequently require modifications to the model architecture, which restricts their widespread application. To overcome these limitations, we introduce a plug-and-play *Adaptive Decision Boundary Strategy* that captures more accurate boundaries and effectively resolves conflicts between new and old classes in the feature space.

Preliminary

Problem Statement

In FSCIL, we conduct a series of continuous learning sessions, each featuring a steady stream of training data represented as $D_{\text{train}} = \{D^t_{\text{train}}\}^T_{t=0}$. Each subset $D^t_{\text{train}} = \{(x_i, y_i)\}^{N_t}_{i=0}$ contains training samples from session t, with x_i and y_i denote the i-th image and its corresponding label, respectively. The initial session, termed the base session, provides a substantial amount of training data. Each subsequent session, referred to as an incremental session, adopts an N-way, K-shot setting, which includes N classes, each with K samples. The label space for the t-th session is denoted by $\mathcal{C}^t$, which is disjoint between different sessions, *i.e.*, $\mathcal{C}^{t_1} \cap \mathcal{C}^{t_2} = \varnothing$ when $t_1 \neq t_2$. The model trained on D^t_{train} should be evaluated on D^t_{test}, which includes all classes encountered up to the t-th session, represented as $\bigcup^t_{i=0} \mathcal{C}^i$.

Base Pretraining and Novel Fine-tuning Strategy

The Base Classes Pretraining and Novel Classes Fine-tuning (BPNF) strategy (Tian et al. 2024) is a common approach in the FSCIL scenario, which involves initial pre-training on abundant base data, followed by fine-tuning to enhance adaptation to novel classes during the incremental phase.

In general, an FSCIL model is decomposed into two components: a feature extractor $f(\cdot)$ and a dynamic classifier with weights W. The output of the model is represented as:

$$\phi(x) = W^\top f(x), \tag{1}$$

where $\phi(x) \in \mathbb{R}^{|\mathcal{C}|\times 1}$, $W \in \mathbb{R}^{d\times|\mathcal{C}|}$, and $f(x) \in \mathbb{R}^{d\times 1}$ with d and $\mathcal{C}$ denotes the output feature dimension and number of classes, respectively.

Specifically, BPNF first leverages sufficient data in the base session to train the model by optimizing the loss for each sample x:

$$\mathcal{L}_{cls}(\phi; x, y) = l(\phi(x), y), \tag{2}$$

where $l(\cdot,\cdot)$ denotes the cross-entropy loss function.

After the base training phase, $f(\cdot)$ is fixed, and the classifier is expanded in each subsequent incremental session: $W = \bigcup^t_{i=0}\{w^i_1, w^i_2, \cdots, w^i_{|\mathcal{C}^i|}\}$, where each term is parameterized by the prototype of the corresponding novel class:

$$w^t_c = \frac{1}{K}\sum^K_{i=1} f(x_i). \tag{3}$$

In t-th session, inference is performed using the Nearest Class Mean (NCM) (Mensink et al. 2013) algorithm to evaluate the accuracy of all encountered classes by predicting the class $\hat{c}^t_x$ with:

$$\hat{c}^t_x = \underset{c,t}{\text{argmax}}\ \text{sim}\left(f(x), w^t_c\right), \tag{4}$$

where $\text{sim}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^\top\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$ represents the cosine similarity between two vectors.

Methodology

In this section, we provide a detailed description of our proposed methodologies, *Adaptive Decision Boundary* and *Inter-class Constraint*. The overview of the entire training pipeline is illustrated in Figure 2.

Adaptive Decision Boundary

To more effectively partition the feature space, we propose the *Adaptive Decision Boundary* (ADB). This approach involves assigning a unique decision boundary to each class and continuously adapting it throughout the training process. This strategy conserves the feature space utilized by the base classes, thus liberating additional feature space for the coming new classes.

Previous research typically assigns a unified decision boundary to all classes instead of implementing individual boundaries for each class. This common practice usually results in a decision boundary that is determined by the class with the most dispersed inter-class features, which can lead to an excessively large boundary for all classes, as illustrated in Figure 2 (a). This practice can clutter the feature space for new classes, resulting in conflicts when new classes are introduced. Drawing inspiration from (Zou et al. 2022), we establish a specific boundary for each class and further propose adaptive adjustments to these boundaries.

In the base session, we initially assign a decision boundary to each base class, formally defined as: $M = \{m^0_1, m^0_2, \ldots, m^0_{|\mathcal{C}^0|}\}$, $M \in \mathbb{R}^{1\times|\mathcal{C}^0|}$. We then incorporate these adaptive boundaries into the model as follows:

$$\phi(x) = (W \cdot M)^\top f(x). \tag{5}$$

Throughout the training process, we apply Eq. 2 to refine both the original model and the boundaries M, thereby facilitating the adaptive adjustment of M towards the optimal decision boundary.

Compared to the less precise boundaries determined by hyperparameters in CLOM (Zou et al. 2022), our method adaptively adjusts the decision boundaries during training. Assuming that clear distinctions already exist between base classes without the use of boundaries M, our method allows the boundaries of classes with smaller feature spaces to adaptively and substantially contract, guided by the loss function described in Eq. 2. Concurrently, the boundaries of other classes will also contract to a degree, as illustrated in Figure 2 (b). This strategy liberates additional space for new classes, effectively reducing the conflicts between existing and incoming classes within the feature space.

In incremental sessions, previous methods typically employ only Eq. 3 to compute prototypes without adjusting boundaries. As a result of continuing to use the static boundaries of the old classes, the new classes fail to adapt to their

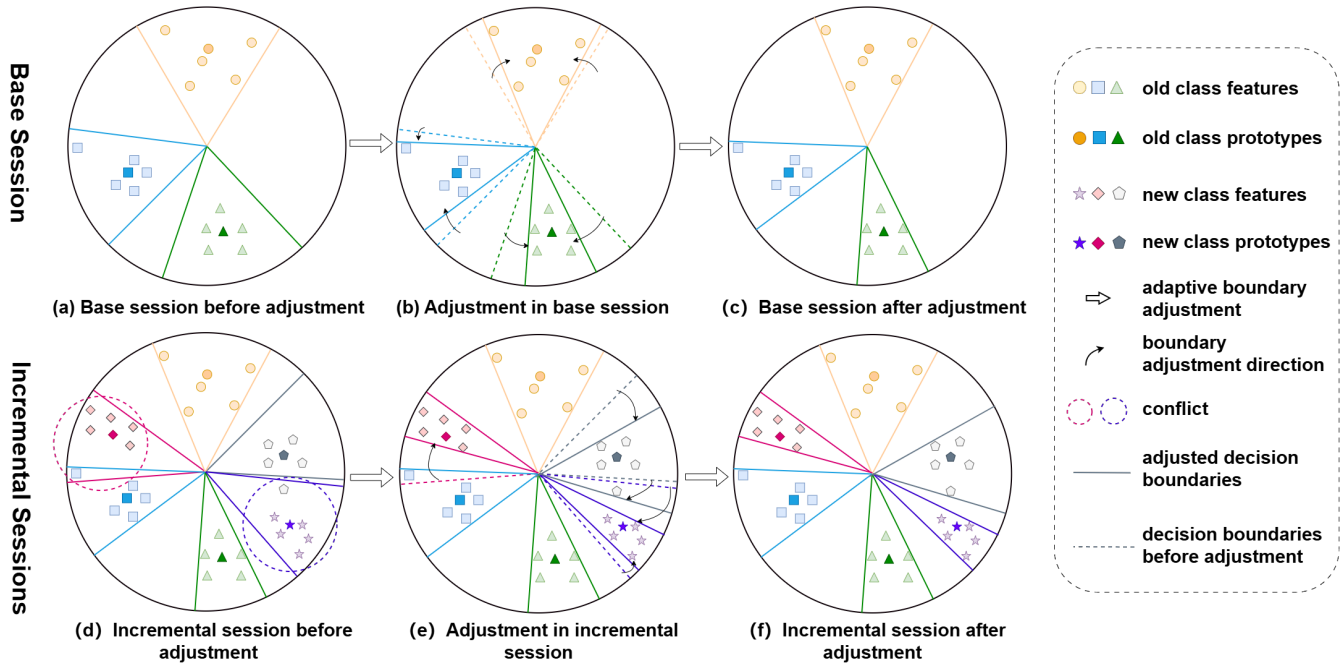


Figure 2: The overall pipeline of our *Adaptive Decision Boundary Strategy* (ADBS). In the base session, we compress the decision boundaries of the base classes to reserve feature space for the upcoming new classes, as depicted in (a)-(c). Subsequently, in the incremental sessions, while maintaining the boundaries of the base classes, we dynamically adjust the boundaries of the new classes to optimize classification performance and compress the boundaries of current classes to allocate feature space for forthcoming new classes, as shown in (d)-(f). Furthermore, we impose Inter-class Constraint (IC) to enhance class distinguishability in each session.

own distribution, leading to suboptimal classification performance and catastrophic forgetting, as illustrated in Figure 2 (d). Therefore, we resolve to continuously fine-tune the boundaries for new classes during these sessions.

In each incremental session t, following the BPNF strategy, we first update the classifier using prototypes of the new classes. Next, we adjust the decision boundaries while keeping the backbone network and the classifier fixed. To accomplish this, we first expand M to $M = \{m_1^0, m_2^0, \cdots, m_{|\mathcal{C}^0|}^0\} \cup \cdots \cup \{m_1^t, m_2^t, \cdots, m_{|\mathcal{C}^t|}^t\}$. Subsequently, we calculate the mean boundary of all old classes to set the initial boundaries for the new classes in the current session. Concurrently, the boundaries of all old classes remain unchanged, while the boundaries for new classes are specifically adjusted using the Eq. 2.

By fine-tuning the boundaries of new classes, we are able to improve the classification performance of these classes while preserving the knowledge of the old classes, as shown in Figure 2 (f). During inference, we utilize the previously trained boundaries. Eq. 4 is reformulated as follows:

$$\hat{c}_x^t = \underset{c,t}{\operatorname{argmax}}\, m_c^t \operatorname{sim}\left(f(x), w_c^t\right). \tag{6}$$

Inter-class Constraint

With the help of adaptive decision boundaries, we can adjust the decision space of existing classes. To further enhance the distinguishability between classes, the decision boundary M should satisfy the following proposition:

Proposition 1 *Given a classifier with weights W, the adaptive boundary strategy could help to better separate classes i, j if the boundary weights m_i and m_j satisfy the following equation:*

$$(1 - m_i) p_i^\top w_i + (m_j - 1) p_i^\top w_j \leq 0, \quad \forall i \neq j, \tag{7}$$

where p_i denotes the prototype of class i. w_i and w_j correspond to the weights of classes i and j in the classifier, respectively.

Proof 1 *Please refer to the* ***supplemental materials***.

Building upon Proposition 1, we introduce a novel Inter-class Constraint (IC) loss to optimize the boundaries M as:

$$\mathcal{L}_{IC} = \sum_{i=1}^{N} \sum_{j=1}^{N} \max(0, (1 - m_i) p_i^\top w_i + (m_j - 1) p_i^\top w_j), \tag{8}$$

where N denotes the total number of classes and p_i is the normalized prototype of i-th class.

IC loss helps optimize the decision boundaries and prototypes for each class, facilitating further separation of the classes and the establishment of clearer decision boundaries.

The overall objective function is defined as:

$$\mathcal{L} = \mathcal{L}_{cls} + \alpha \mathcal{L}_{IC}, \tag{9}$$

where α denotes the weighting parameters of $\mathcal{L}_{IC}$.

Methods	Accuracy in each session (%)									Average
	0	1	2	3	4	5	6	7	8	Acc.
Topic (Tao et al. 2020)	64.10	55.88	47.07	45.16	40.11	36.38	33.96	31.55	29.37	42.62
CEC (Zhang et al. 2021)	73.07	68.88	65.26	61.19	58.09	55.57	53.22	51.34	49.14	59.53
FACT (Zhou et al. 2022a)	74.60	72.09	67.56	63.52	61.38	58.36	56.28	54.24	52.10	62.24
C-FSCIL (Hersche et al. 2022)	77.47	72.40	67.47	63.25	59.84	56.95	54.42	52.47	50.47	61.64
CLOM† (Zou et al. 2022)	74.20	69.83	66.17	62.39	59.26	56.48	54.36	52.16	50.25	60.57
DBONet† (Guo et al. 2023)	77.81	73.62	71.04	66.29	63.52	61.01	58.37	56.89	55.78	64.93
MCNet (Ji et al. 2023)	73.30	69.34	65.72	61.70	58.75	56.44	54.59	53.01	50.72	60.40
SoftNet (Kang et al. 2023)	72.62	67.31	63.05	59.39	56.00	53.23	51.06	48.83	46.63	57.57
WaRP (Kim et al. 2023)	80.31	75.86	71.87	67.58	64.39	61.34	59.15	57.10	54.74	65.82
TEEN (Wang et al. 2023)	74.92	72.65	68.74	65.01	62.01	59.29	57.90	54.76	52.64	63.10
NC-FSCIL (Yang et al. 2023)	82.52	76.82	73.34	69.68	66.19	62.85	60.96	59.02	56.11	67.50
ALFSCIL (Li et al. 2024)	80.75	77.88	72.94	68.79	65.33	62.15	60.02	57.68	55.17	66.75
DyCR (Pan et al. 2024)	75.73	73.29	68.71	64.80	62.11	59.25	56.70	54.56	52.24	63.04
baseline	73.92	67.14	63.71	60.07	57.10	54.85	52.52	50.49	48.60	58.71
baseline+ADBS	**79.93**	**75.22**	**71.11**	**65.99**	**62.46**	**58.38**	**55.96**	**53.72**	**51.15**	**63.77**
	(+6.02)	**(+8.08)**	**(+7.40)**	**(+5.92)**	**(+5.36)**	**(+3.53)**	**(+3.43)**	**(+3.22)**	**(+2.55)**	**(+5.06)**
OrCo* (Ahmed, Kukleva, and Schiele 2024)	79.77	63.29	62.39	60.13	58.76	56.56	55.49	54.19	51.12	60.19
OrCo+ADBS	79.77	**63.46**	61.89	**60.43**	**59.23**	56.32	**55.76**	**54.48**	**51.54**	**60.32**
	(+0.00)	**(+0.17)**	(-0.50)	**(+0.29)**	**(+0.46)**	(-0.25)	**(+0.27)**	**(+0.30)**	**(+0.42)**	**(+0.13)**
ALICE†* (Peng et al. 2022)	80.37	72.34	67.67	63.61	61.11	58.53	57.40	55.43	53.46	63.32
ALICE+ADBS	80.12	**74.11**	**70.51**	**66.72**	**63.90**	**61.25**	**60.00**	**58.07**	**56.00**	**65.63**
	(-0.25)	**(+1.77)**	**(+2.84)**	**(+3.11)**	**(+2.79)**	**(+2.72)**	**(+2.60)**	**(+2.64)**	**(+2.54)**	**(+2.31)**
SAVC* (Song et al. 2023)	78.60	72.95	68.73	64.59	61.41	58.46	56.29	54.40	52.19	63.07
SAVC+ADBS	**85.13**	**80.39**	**77.07**	**72.61**	**69.54**	**66.54**	**64.70**	**62.72**	**60.60**	**71.03**
	(+6.53)	**(+7.43)**	**(+8.34)**	**(+8.03)**	**(+8.13)**	**(+8.08)**	**(+8.41)**	**(+8.32)**	**(+8.41)**	**(+7.96)**

Table 1: Comparison with SOTA methods on CIFAR100 for FSCIL. †: Boundary-based method. *: Reproduced results.

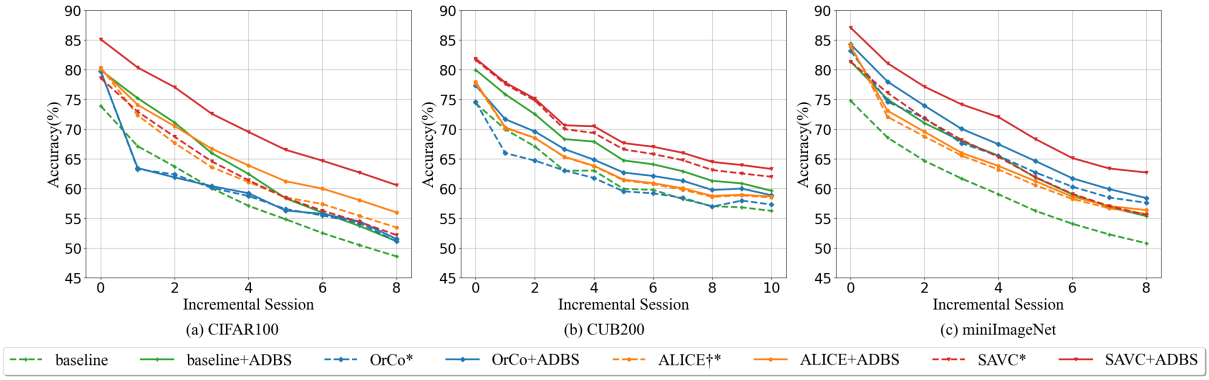


Figure 3: Comparison with different baseline methods on CIFAR100, CUB200, and miniImageNet. The dashed line represents the accuracy of the baseline method, while the solid line depicts the classification performance with our proposed ADBS.

Experiment

Experimental Setup

Dataset Following the benchmark setting (Tao et al. 2020), we evaluate the effectiveness of our proposed ADBS method on three datasets, *i.e.*, CIFAR100 (Krizhevsky 2009), miniImageNet (Deng et al. 2009), and Caltech-UCSD Birds-200-2011 (CUB200) (Wah et al. 2011).

Implementation Details We integrate our approach into four FSCIL methods: the standard FSCIL baseline, SAVC (Song et al. 2023), ALICE (Peng et al. 2022), and OrCo (Ahmed, Kukleva, and Schiele 2024). We implemente these methods using the official codes released by the authors to ensure a fair comparison. [1]

Evaluation Protocol and Metric Following previous FSCIL studies (Tao et al. 2020; Zhang et al. 2021), we report the Top-1 accuracy of current classes after the base session and each incremental session.

[1]More details about datasets and experimental settings are included in the supplementary materials.

Methods	Accuracy in each session (%)									Average
	0	1	2	3	4	5	6	7	8	Acc.
Topic (Tao et al. 2020)	61.31	50.09	45.17	41.13	37.48	35.52	32.19	29.46	24.42	39.64
CEC (Zhang et al. 2021)	72.00	66.83	62.97	59.43	56.70	53.73	51.19	49.24	47.63	57.75
FACT (Zhou et al. 2022a)	72.56	69.63	66.38	62.77	60.60	57.33	54.34	52.16	50.49	60.70
C-FSCIL (Hersche et al. 2022)	76.40	71.14	66.46	63.29	60.42	57.46	54.78	53.11	51.41	61.61
CLOM† (Zou et al. 2022)	73.08	68.09	64.16	60.41	57.41	54.29	51.54	49.37	48.00	58.48
DBONet† (Guo et al. 2023)	74.53	71.55	68.57	65.72	63.08	60.64	57.83	55.21	53.82	63.44
MCNet (Ji et al. 2023)	72.33	67.70	63.50	60.34	57.59	54.70	52.13	50.41	49.08	58.64
SoftNet (Kang et al. 2023)	77.17	70.32	66.15	62.55	59.48	56.46	53.71	51.68	50.24	60.86
WaRP (Kim et al. 2023)	72.99	68.10	64.31	61.30	58.64	56.08	53.40	51.72	50.65	59.69
TEEN (Wang et al. 2023)	73.53	70.55	66.37	63.23	60.53	57.95	55.24	53.44	52.08	61.44
NC-FSCIL (Yang et al. 2023)	84.02	76.80	72.00	67.83	66.35	64.04	61.46	59.54	58.31	67.82
ALFSCIL (Li et al. 2024)	81.27	75.97	70.97	66.53	63.46	59.95	56.93	54.81	53.31	64.80
DyCR (Pan et al. 2024)	73.18	70.16	66.87	63.43	61.18	58.79	55.00	52.87	51.08	61.40
baseline	74.80	68.58	64.69	61.71	59.06	56.26	54.09	52.30	50.81	60.25
baseline+ADBS	**81.40**	**75.03**	**71.03**	**68.00**	**65.56**	**61.87**	**59.04**	**56.87**	**55.38**	**66.02**
	(+6.60)	**(+6.45)**	**(+6.34)**	**(+6.29)**	**(+6.50)**	**(+5.61)**	**(+4.95)**	**(+4.58)**	**(+4.57)**	**(+5.77)**
OrCo* (Ahmed, Kukleva, and Schiele 2024)	83.22	74.60	71.89	67.65	65.53	62.73	60.33	58.51	57.62	66.90
OrCo+ADBS	**84.30**	**78.02**	**73.96**	**70.07**	**67.47**	**64.65**	**61.74**	**59.92**	**58.42**	**68.73**
	(+1.08)	**(+3.42)**	**(+2.07)**	**(+2.41)**	**(+1.95)**	**(+1.92)**	**(+1.41)**	**(+1.41)**	**(+0.80)**	**(+1.83)**
ALICE†* (Peng et al. 2022)	84.23	72.08	68.77	65.57	63.21	60.58	58.26	56.69	55.76	65.02
ALICE+ADBS	83.92	**73.11**	**69.57**	**66.03**	**63.82**	**61.26**	**58.64**	**57.07**	**56.39**	**65.53**
	(-0.32)	**(+1.03)**	**(+0.80)**	**(+0.46)**	**(+0.61)**	**(+0.68)**	**(+0.38)**	**(+0.38)**	**(+0.63)**	**(+0.52)**
SAVC* (Song et al. 2023)	81.38	76.12	71.79	68.21	65.34	61.94	59.13	57.06	55.60	66.29
SAVC+ADBS	**87.10**	**81.11**	**77.16**	**74.19**	**72.05**	**68.34**	**65.14**	**63.39**	**62.72**	**72.36**
	(+5.72)	**(+4.98)**	**(+5.37)**	**(+5.97)**	**(+6.71)**	**(+6.40)**	**(+6.01)**	**(+6.33)**	**(+7.12)**	**(+6.07)**

Table 2: Comparison with SOTA methods on miniImageNet for FSCIL. †: Boundary-based method. *: Reproduced results.

ADB	**IC**	Accuracy in each session (%)									Δ_{last}
		0	1	2	3	4	5	6	7	8	
		74.80	68.58	64.69	61.71	59.06	56.26	54.09	52.30	50.81	-
✓		80.13	74.74	70.21	66.51	63.67	60.62	57.92	55.85	54.67	+3.86
✓	✓	**81.40**	**75.03**	**71.03**	**68.00**	**65.56**	**61.87**	**59.04**	**56.87**	**55.38**	**+4.57**

Table 3: Ablation studies on miniImageNet benchmark. **ADB** and **IC** denote *Adaptive Decision Boundary* and *Inter-class Constraint*, respectively. Δ_{last}: Relative improvements of the last sessions compared to the fixed decision boundary baseline.

Main Results

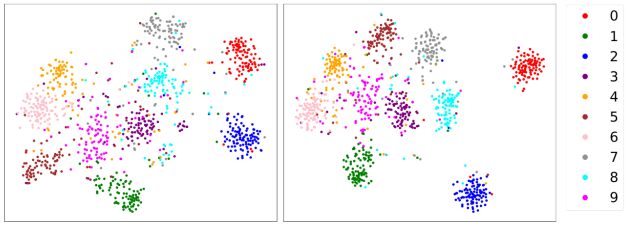
We conduct experimental comparisons on three benchmark datasets [2], *i.e.*, CIFAR100, miniImageNet, and CUB200, shown in Tables. 1, 2, and Figure 3.

Figure 3 shows that our proposed ADBS consistently enhances the performance of all integrated FSCIL methods across all datasets. The results further reveal that integrating SAVC with ADBS (SAVC+ADBS) achieves state-of-the-art performance on all datasets, thereby confirming the effectiveness of the proposed strategy. Furthermore, we observed that for baseline methods and SAVC, which do not optimize for boundaries, our proposed ADBS continuously compresses the boundaries in each session to allocate feature space, effectively alleviating conflicts between new and old classes in the feature space, thereby significantly enhancing performance. In contrast, the ALICE and Orco methods each utilize specific losses (angular penalty loss and center orthogonality loss, respectively) to optimize classification decision boundaries. Consequently, integrating ADBS does not yield significant improvements in base classification performance. Nevertheless, during incremental sessions, ADBS continues to effectively balance the decision spaces between old and new classes without compromising the classification performance of the combined methods, thereby enhancing the overall model performance.

Ablation Study

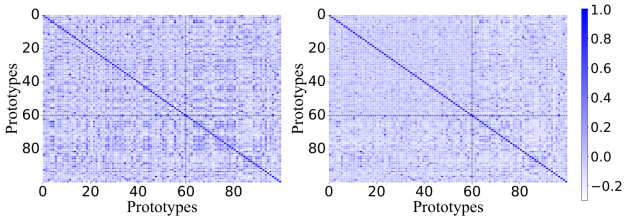
We perform ablation studies to investigate the effectiveness of the key components of the proposed ADBS on the miniImageNet dataset. Built on the BPNF framework, we initially employed only Cross-Entropy loss as the base-

[2]The result table of CUB200 dataset can be found in the supplementary materials.



(a) Fixed decision boundary (b) ADBS(ours)

Figure 4: The t-SNE visualization on the miniImageNet dataset of the feature embeddings. Classes 0-5 denote the base classes, whereas classes 6-9 signify the new classes. Our ADBS, as compared to the fixed decision boundary, demonstrates superior class separability.



(a) Fixed decision boundary (b) ADBS(ours)

Figure 5: Cosine similarity score between classes on CIFAR100. We use black dashed lines to demarcate the regions of base and incremental classes. Compared to the fixed decision boundary, our ADBS creates wider spacing between class prototypes, leading to lighter colors in the figure.

line. Subsequently, we progressively integrated the proposed Adaptive Decision Boundary (ADB) and Inter-class Constraint (IC), observing their impacts on performance. As shown in Table 3, each component contributes to performance improvement. Notably, the ADB strategy significantly enhances classification performance in all sessions, demonstrating that the adaptive strategy effectively adjusts decision boundaries for better space allocation. Additionally, we incorporated an inter-class constraint to increase the distinguishability between classes, which also improved the compactness of embeddings and facilitated FSCIL. [3]

Further Analysis

Visualization of Class Separation We employ T-SNE to visualize the feature space of the miniImageNet dataset, as shown in Figure 4, with six classes randomly selected from the base classes and four from the new classes. As shown in Figure 4 (a), the base classes are densely clustered while the new classes are more scattered, resulting in frequent conflicts within the feature space. For example, the purple class (class 3) displays conflicts with both the cyan (class 8) and magenta (class 9) new classes. In contrast, as shown in Figure 4 (b), our ADBS enhances intra-class cohesion and inter-class separation, clarifying the decision boundaries between classes and effectively mitigating conflicts between old and new classes in the feature space.

[3]More ablation results of CIFAR100 and CUB200 datasets can be found in the supplementary materials.

	CUB200	CIFAR100	miniImageNet
baseline	0.3517	0.8846	0.8944
ADBS	**0.3584**	**0.9497**	**0.9477**
$\Delta_{impro.}$	+0.0067	+0.0651	+0.0533

Table 4: The degree of class separation on three FSCIL benchmarks. $\Delta_{impro.}$: Relative improvements of the overall degree compared to the fixed decision boundary baseline.

Analysis of the Class Separation Degree We first compute the cosine similarity of two class prototypes (consistent with the classifier) to analyze the similarity of the two classes, shown in Figure 5. We observe that a fixed decision boundary strategy causes class prototypes to cluster closely, obscuring boundaries and leading to classification ambiguities. In contrast, our ADBS significantly separates class prototypes, establishing clearer boundaries between classes for more accurate classification. To further explore our effectiveness in enhancing class separation, we quantified the degree of separation D_{cs} for each method by calculating the average cosine distance between the prototypes of any two classes as:

$$D_{cs} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(1 - \text{sim}(p_i, p_j)), \tag{10}$$

where N indicates the number of all classes. Table 4 presents the results of the baseline and our ADBS on three FSCIL benchmarks. The results demonstrate that our ADBS effectively separates different classes, mitigating class boundary conflicts and thereby enhancing performance. It is worth noting that, compared to the other two datasets, CUB200 primarily focuses on fine-grained classification scenarios, where separating different classes is more challenging (*i.e.*, a lower score). However, applying our ADBS still achieves a slight improvement, underscoring its generalizability in complex scenarios.

Conclusion

The conflict between new and old classes within the feature space presents a notable challenge for FSCIL methods. In this paper, we propose a plug-and-play *Adaptive Decision Boundary Strategy* (ADBS) designed specifically to mitigate this issue. ADBS comprises two key components, *i.e.*, an *Adaptive Decision Boundary* (ADB) and an *Inter-class Constraint* (IC). ADB assigns a distinct decision boundary to each class and adaptively adjusts it in each session, while the IC is designed to enhance class distinguishability. Moreover, ADBS offers plug-and-play functionality, enabling seamless integration into various FSCIL methods. Extensive experiments on three benchmarks demonstrate that incorporating ADBS consistently enhances the performance of existing FSCIL methods, highlighting the effectiveness of our proposed strategy.

Acknowledgements

This research was supported by National Natural Science Foundation of China (NSFC) (62376088), Hebei Natural Science Foundation (F2024202047). This research was supported in part by the Wallenberg-NTU Presidential Postdoctoral Fellowship (Award No. 024560-00001).

References

Ahmed, N.; Kukleva, A.; and Schiele, B. 2024. OrCo: Towards Better Generalization via Orthogonality and Contrast for Few-Shot Class-Incremental Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 28762–28771.

Chen, H.; Wang, Y.; and Hu, Q. 2022. Multi-granularity regularized re-balancing for class incremental learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(7): 7263–7277.

Chen, X.; Lan, X.; Sun, F.; and Zheng, N. 2020. A boundary based out-of-distribution classifier for generalized zero-shot learning. In *European conference on computer vision*, 572–588. Springer.

Cheng, H.; Yang, S.; Wang, C.; Zhou, J. T.; Kot, A. C.; and Wen, B. 2025. STSP: Spatial-Temporal Subspace Projection for Video Class-Incremental Learning. In *European Conference on Computer Vision*, 374–391. Springer.

Cheraghian, A.; Rahman, S.; Fang, P.; Roy, S. K.; Petersson, L.; and Harandi, M. 2021. Semantic-aware knowledge distillation for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2534–2543.

Dayal, A.; KB, V.; Cenkeramaddi, L. R.; Mohan, C.; Kumar, A.; and N Balasubramanian, V. 2024. MADG: margin-based adversarial learning for domain generalization. *Advances in Neural Information Processing Systems*, 36.

Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255. Ieee.

Dong, S.; Hong, X.; Tao, X.; Chang, X.; Wei, X.; and Gong, Y. 2021. Few-shot class-incremental learning via relation knowledge distillation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 1255–1263.

Fan, Y.; Wang, Y.; Zhu, P.; and Hu, Q. 2024. Dynamic Sub-graph Distillation for Robust Semi-supervised Continual Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 11927–11935.

Guo, C.; Zhao, Q.; Lyu, S.; Liu, B.; Wang, C.; Chen, L.; and Cheng, G. 2023. Decision boundary optimization for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3501–3511.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Hersche, M.; Karunaratne, G.; Cherubini, G.; Benini, L.; Sebastian, A.; and Rahimi, A. 2022. Constrained few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9057–9067.

Hou, S.; Pan, X.; Loy, C. C.; Wang, Z.; and Lin, D. 2019. Learning a unified classifier incrementally via rebalancing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 831–839.

Ji, Z.; Hou, Z.; Liu, X.; Pang, Y.; and Li, X. 2023. Memorizing complementation network for few-shot class-incremental learning. *IEEE Transactions on Image Processing*, 32: 937–948.

Kalla, J.; and Biswas, S. 2022. S3c: Self-supervised stochastic classifiers for few-shot class-incremental learning. In *European Conference on Computer Vision*, 432–448. Springer.

Kang, H.; Yoon, J.; Madjid, S. R. H.; Hwang, S. J.; and Yoo, C. D. 2023. On the Soft-Subnetwork for Few-Shot Class Incremental Learning. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.

Kim, D.; Han, D.; Seo, J.; and Moon, J. 2023. Warping the Space: Weight Space Rotation for Class-Incremental Few-Shot Learning. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.

Krizhevsky, A. 2009. Learning Multiple Layers of Features from Tiny Images. *Master's thesis, University of Tront*.

Li, J.; Dong, S.; Gong, Y.; He, Y.; and Wei, X. 2024. Analogical Learning-Based Few-Shot Class-Incremental Learning. *IEEE Transactions on Circuits and Systems for Video Technology*.

Li, Z.; and Hoiem, D. 2017. Learning without forgetting. *IEEE transactions on pattern analysis and machine intelligence*, 40(12): 2935–2947.

Liu, B.; Ben Ayed, I.; Galdran, A.; and Dolz, J. 2022. The devil is in the margin: Margin-based label smoothing for network calibration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 80–88.

Liu, B.; Yang, B.; Xie, L.; Wang, R.; Tian, Q.; and Ye, Q. 2023. Learnable distribution calibration for few-shot class-incremental learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10): 12699–12706.

Masana, M.; Liu, X.; Twardowski, B.; Menta, M.; Bagdanov, A. D.; and Van De Weijer, J. 2022. Class-incremental learning: survey and performance evaluation on image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5): 5513–5533.

McCloskey, M.; and Cohen, N. J. 1989. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, 109–165. Elsevier.

Mensink, T.; Verbeek, J.; Perronnin, F.; and Csurka, G. 2013. Distance-based image classification: Generalizing to new classes at near-zero cost. *IEEE transactions on pattern analysis and machine intelligence*, 35(11): 2624–2637.

Pan, Z.; Yu, X.; Zhang, M.; Zhang, W.; and Gao, Y. 2024. DyCR: A Dynamic Clustering and Recovering Network for Few-Shot Class-Incremental Learning. *IEEE Transactions on Neural Networks and Learning Systems*.

Peng, C.; Zhao, K.; Wang, T.; Li, M.; and Lovell, B. C. 2022. Few-shot class-incremental learning from an open-set perspective. In *European Conference on Computer Vision*, 382–397. Springer.

Rebuffi, S.-A.; Kolesnikov, A.; Sperl, G.; and Lampert, C. H. 2017. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2001–2010.

Song, Z.; Zhao, Y.; Shi, Y.; Peng, P.; Yuan, L.; and Tian, Y. 2023. Learning with fantasy: Semantic-aware virtual contrastive constraint for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 24183–24192.

Tao, X.; Hong, X.; Chang, X.; Dong, S.; Wei, X.; and Gong, Y. 2020. Few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12183–12192.

Tian, S.; Li, L.; Li, W.; Ran, H.; Ning, X.; and Tiwari, P. 2024. A survey on few-shot class-incremental learning. *Neural Networks*, 169: 307–324.

Wah, C.; Branson, S.; Welinder, P.; Perona, P.; and Belongie, S. 2011. The caltech-ucsd birds-200-2011 dataset.

Wang, Q.; Zhou, D.; Zhang, Y.; Zhan, D.; and Ye, H. 2023. Few-Shot Class-Incremental Learning via Training-Free Prototype Calibration. In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Wang, Y.; Yao, X.; Zhu, P.; Li, W.; Cao, M.; and Hu, Q. 2024. Integrated Heterogeneous Graph Attention Network for Incomplete Multi-modal Clustering. *International Journal of Computer Vision*, 1–20.

Yang, S.; Liu, L.; and Xu, M. 2021. Free Lunch for Few-shot Learning: Distribution Calibration. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Yang, Y.; Yuan, H.; Li, X.; Lin, Z.; Torr, P. H. S.; and Tao, D. 2023. Neural Collapse Inspired Feature-Classifier Alignment for Few-Shot Class-Incremental Learning. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.

Zhang, C.; Song, N.; Lin, G.; Zheng, Y.; Pan, P.; and Xu, Y. 2021. Few-shot incremental learning with continually evolved classifiers. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12455–12464.

Zhao, H.; Fu, Y.; Kang, M.; Tian, Q.; Wu, F.; and Li, X. 2021. Mgsvf: Multi-grained slow versus fast framework for few-shot class-incremental learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(3): 1576–1588.

Zhou, D.-W.; Wang, F.-Y.; Ye, H.-J.; Ma, L.; Pu, S.; and Zhan, D.-C. 2022a. Forward compatible few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9046–9056.

Zhou, D.-W.; Ye, H.-J.; Ma, L.; Xie, D.; Pu, S.; and Zhan, D.-C. 2022b. Few-shot class-incremental learning by sampling multi-phase tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11): 12816–12831.

Zhu, K.; Cao, Y.; Zhai, W.; Cheng, J.; and Zha, Z.-J. 2021. Self-promoted prototype refinement for few-shot class-incremental learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 6801–6810.

Zou, Y.; Zhang, S.; Li, Y.; and Li, R. 2022. Margin-based few-shot class-incremental learning with class-level overfitting mitigation. *Advances in neural information processing systems*, 35: 27267–27279.