# CVPR

## 2024

### OrCo: Towards Better Generalization via Orthogonality and Contrast for Few-Shot Class-Incremental Learning

### OrCo：通过正交性和对比性提升小样本类别增量学习的泛化能力

Few-Shot Class-Incremental Learning (FSCIL) intro-duces a paradigm in which the problem space expands with limited data. FSCIL methods inherently face the chal-lenge of catastrophic forgetting as data arrives incremen-tally, making models susceptible to overwriting previously acquired knowledge. Moreover, given the scarcity of la-beled samples available at any given time, models may be prone to overfitting and find it challenging to strike a bal-ance between extensive pretraining and the limited incre-mental data. To address these challenges, we propose the OrCo framework built on two core principles: features' orthogonality in the representation space, and contrastive learning. In particular, we improve the generalization of the embedding space by employing a combination of su-pervised and self-supervised contrastive losses during the pretraining phase. Additionally, we introduce OrCo loss to address challenges arising from data limitations during in-cremental sessions. Through feature space perturbations and orthogonality between classes, the OrCo loss maxi-mizes margins and reserves space for the following incre-mental data. This, in turn, ensures the accommodation of incoming classes in the feature space without compromising previously acquired knowledge. Our experimental results showcase state-of-the-art performance across three bench-mark datasets, including mini-ImageNet, CIFAR100, and CUB datasets. Code is available at: https://github.com/noorahmedds/OrCo.

少样本类别增量学习（FSCIL）引入了一种在有限数据下扩展问题空间的范式。FSCIL 方法在数据逐步增加时，本质上面临着灾难性遗忘的挑战，这使得模型容易覆盖先前习得的知识。此外，由于在任何给定时间可用的标记样本数量有限，模型可能会过度拟合，并且在广泛的预训练和有限的增量数据之间难以找到平衡。为了解决这些挑战，我们提出了基于两个核心原则的 OrCo 框架：表示空间中的特征正交性和对比学习。具体而言，我们在预训练阶段通过结合监督和自监督对比损失来改进嵌入空间的泛化能力。此外，我们引入了 OrCo 损失，以应对增量阶段中数据限制带来的挑战。通过特征空间扰动和类别之间的正交性，OrCo 损失最大化了边界，并为后续的增量数据预留了空间。这确保了新类别的特征空间容纳，而不会损害先前习得的知识。我们的实验结果在三个基准数据集上展示了最先进的性能，包括 mini-ImageNet、CIFAR100 和 CUB 数据集。代码可在以下地址获取：https://github.com/noorahmedds/OrCo。

### Calibrating Higher-Order Statistics for Few-Shot Class-Incremental Learning with Pre-trained Vision Transformers

### 校准高阶统计量以利用预训练视觉变换器进行少样本类别增量学习

Few-shot class-incremental learning (FSCIL) aims to adapt the model to new classes from very few data (5 samples) without forgetting the previously learned classes. Recent works in many-shot CIL (MSCIL) (using all available training data) exploited pre-trained models to reduce forgetting and achieve better plasticity. In a similar fashion, we use ViT models pre-trained on large-scale datasets for few-shot settings, which face the critical issue of low plasticity. FSCIL methods start with a many-shot first task to learn a very good feature extractor and then move to the few-shot setting from the second task onwards. While the focus of most recent studies is on how to learn the many-shot first task so that the model generalizes to all future few-shot tasks, we explore in this work how to better model the few-shot data using pre-trained models, irrespective of how the first task is trained. Inspired by recent works in MSCIL, we explore how using higher-order feature statistics can influence the classification of few-shot classes. We identify the main challenge of obtaining a good covariance matrix from few-shot data and propose to calibrate the covariance matrix for new classes based on semantic similarity to the many-shot base classes. Using the calibrated feature statistics in combination with existing methods significantly improves few-shot continual classification on several FSCIL benchmarks. Code is available at https://github.com/dipamgoswami/FSCIL-Calibration.

少样本类别增量学习（FSCIL）旨在使模型能够从少量数据（例如5个样本）中适应新类别，同时不会遗忘之前已学习的类别。近期在多样本类别增量学习（MSCIL）（使用所有可用的训练数据）的研究中，利用预训练模型来减少遗忘并实现更好的可塑性。类似地，我们在少样本场景中使用在大规模数据集上预训练的ViT模型，这些模型面临的关键问题是可塑性较低。FSCIL方法通常从多样本的第一个任务开始，学习一个非常优秀的特征提取器，然后从第二个任务开始转向少样本设置。尽管大多数近期研究的重点是如何学习多样本的第一个任务，以便模型能够泛化到所有未来的少样本任务，我们在本研究中探讨了如何使用预训练模型更好地建模少样本数据，而不考虑第一个任务是如何训练的。受近期MSCIL研究的启发，我们探讨了使用高阶特征统计量如何影响少样本类别的分类。我们确定了从少样本数据中获得良好协方差矩阵的主要挑战，并提出基于与多样本基础类别的语义相似性来校准新类别的协方差矩阵。将校准后的特征统计量与现有方法结合使用，显著提高了在多个FSCIL基准上的少样本持续分类性能。代码可在以下链接获取：https://github.com/dipamgoswami/FSCIL-Calibration。

## 2025

# AAAI

## 2024

### M2SD:Multiple Mixing Self-Distillation for Few-Shot Class-Incremental Learning

### M2SD：多混合自蒸馏在少样本类别增量学习中的应用

Few-shot Class-incremental learning (FSCIL) is a challenging task in machine learning that aims to recognize new classes from a limited number of instances while preserving the ability to classify previously learned classes without retraining the entire model. This presents challenges in updating the model with new classes using limited training data, particularly in balancing acquiring new knowledge while retaining the old. We propose a novel method named Multiple Mxing Self-Distillation (M2SD) during the training phase to address these issues. Specifically, we propose a dual-branch structure that facilitates the expansion of the entire feature space to accommodate new classes. Furthermore, we introduce a feature enhancement component that can pass additional enhanced information back to the base network by self-distillation, resulting in improved classification performance upon adding new classes. After training, we discard both structures, leaving only the primary network to classify new class instances. Extensive experiments demonstrate that our approach achieves superior performance over previous state-of-the-art methods.

少样本类别增量学习（FSCIL）是机器学习中的一项挑战性任务，旨在在有限的样本数量下识别新类别，同时在不重新训练整个模型的情况下保持对先前学习类别的分类能力。这在使用有限的训练数据更新模型以包含新类别时提出了挑战，尤其是在平衡获取新知识与保留旧知识之间。我们提出了一种新颖的方法，称为多混合自蒸馏（M2SD），在训练阶段解决这些问题。具体而言，我们提出了一种双分支结构，有助于扩展整个特征空间以容纳新类别。此外，我们引入了一个特征增强组件，可以通过自蒸馏将额外的增强信息传递回基础网络，从而在添加新类别时提高分类性能。训练完成后，我们丢弃这两个结构，仅保留主网络来分类新类别的实例。大量实验表明，我们的方法在性能上优于先前的最先进方法。

## 2025

### Pseudo Informative Episode Construction for Few-Shot Class-Incremental Learning

### 伪信息性小样本构建以实现少样本类别增量学习

Few-Shot Class-Incremental Learning (FSCIL) studies how to empower the machine learning system to learn novel classes with only a few annotated examples continually. To tackle the FSCIL task, recent state-of-the-art methods propose to employ the meta-learning mechanism, which constructs the pseudo incremental episodes/tasks in the training phase. However, these methods only select part of the base classes to construct the pseudo novel classes in the feature space of the base classes, which cannot mimic the real novel

少样本类别增量学习（FSCIL）研究如何使机器学习系统能够仅通过少量标注样本持续学习新类别。为了应对FSCIL任务，最近的先进方法提出采用元学习机制，该机制在训练阶段构建伪增量场景/任务。然而，这些方法仅选择部分基础类别来在基础类别的特征空间中构建伪新类别，这无法真实模拟实际的新类别。

### Few-Shot Audio-Visual Class-Incremental Learning with Temporal Prompting and Regularization

### 基于时间提示和正则化的少样本视听类别增量学习

Audio-Visual Learning (AVL) aims at the audio-visual perception with both audio and vision modalities. AVL also suffers from data insufficiency in many applications as with other unimodal tasks. Concurrently, AVL often needs to continuously learn over time rather than all knowledge simultaneously. Considering the above two perspectives, our work mainly focuses on benchmarking the unexplored Few-Shot Audio-Visual Class-Incremental Learning (FS-AVCIL), i.e., continually perceiving novel categories described by a limited number of labeled examples with audio and visual modalities. Firstly, we provide the detailed task configuration together with a thorough analysis of the challenges in FS-AVCIL: (1) how to efficiently learn and fuse multimodal information with limited labeled examples; and (2) how to alleviate catastrophic forgetting cross-modal semantic correlations with limited data. Then, we propose an efficient framework based on Vision Transformer to solve FS-AVCIL. This framework contains two parts: temporal-residual prompting for audio-visual synergy adapter and temporal prompt regularization. Specifically, temporal-residual prompting is incorporated into the audio-visual adapter to efficiently finetune the pre-trained foundation model with limited data and capture audio-visual correlation by learning temporal-relevant prompts. Besides, we regularize temporal-relevant prompts to memorize previous knowledge by fully using the temporal knowledge from various perspectives. This framework is validated in audio-visual classification tasks under the FS-AVCIL scenario, and extensive experiments demonstrate its superior performance.

视听学习（AVL）旨在通过音频和视觉模态实现视听感知。与其他单模态任务一样，AVL在许多应用中也面临着数据不足的问题。同时，AVL通常需要随着时间的推移持续学习，而不是一次性掌握所有知识。基于以上两个视角，我们的工作主要集中在尚未探索的少样本视听类别增量学习（FS-AVCIL）的基准测试上，即在有限的标注样本下，通过音频和视觉模态持续感知新的类别。  
  
首先，我们提供了详细的任务配置，并对FS-AVCIL中的挑战进行了深入分析，具体包括：（1）如何在有限的标注样本下高效地学习和融合多模态信息；（2）如何在数据有限的情况下减轻跨模态语义关联的灾难性遗忘。然后，我们提出了一种基于视觉变换器的高效框架来解决FS-AVCIL问题。该框架包含两个部分：时间残差提示的视听协同适配器和时间提示正则化。  
  
具体来说，时间残差提示被整合到视听适配器中，以在有限的数据下高效微调预训练的基础模型，并通过学习时间相关的提示来捕捉视听关联。此外，我们通过充分利用时间知识从多个角度来记忆先前的知识，对时间相关的提示进行正则化。该框架在FS-AVCIL场景下的视听分类任务中进行了验证，广泛的实验表明其性能优越。

### AnchorInv: Few-Shot Class-Incremental Learning of Physiological Signals via Feature Space-Guided Inversion

### AnchorInv：通过特征空间引导的逆向学习实现生理信号的少样本类别增量学习

Deep learning models have demonstrated exceptional performance in a variety of real-world applications. These successes are often attributed to strong base models that can generalize to novel tasks with limited supporting data while keeping prior knowledge intact. However, these impressive results are based on the availability of a large amount of high-quality data, which is often lacking in specialized biomedical applications. In such fields, models are often developed with limited data that arrive incrementally with novel categories. This requires the model to adapt to new information while preserving existing knowledge. Few-Shot Class-Incremental Learning (FSCIL) methods offer a promising approach to addressing these challenges, but they also depend on strong base models that face the same aforementioned limitations. To overcome these constraints, we propose AnchorInv following the straightforward and efficient buffer-replay strategy. Instead of selecting and storing raw data, AnchorInv generates synthetic samples guided by anchor points in the feature space. This approach protects privacy and regularizes the model for adaptation. When evaluated on three public physiological time series datasets, AnchorInv exhibits efficient knowledge forgetting prevention and improved adaptation to novel classes, surpassing state-of-the-art baselines.

深度学习模型在各种实际应用中展示了卓越的性能。这些成功通常归功于强大的基础模型，这些模型能够在有限的支持数据下泛化到新任务，同时保持先前的知识。然而，这些令人印象深刻的结果依赖于大量高质量数据的可用性，而在专业的生物医学应用中，这种数据往往不足。在这些领域，模型通常是在有限的数据上开发的，这些数据会逐步增加新的类别。这要求模型在适应新信息的同时保留现有知识。少样本类别增量学习（FSCIL）方法为解决这些挑战提供了一种有前景的途径，但它们同样依赖于面临上述限制的强大基础模型。为了克服这些限制，我们提出了AnchorInv，采用了一种简单高效的缓冲重放策略。与选择和存储原始数据不同，AnchorInv通过特征空间中的锚点生成合成样本。这种方法保护了隐私，并对模型进行正则化以适应新信息。在三个公开的生理时间序列数据集上进行评估时，AnchorInv展示了高效的知识遗忘预防和对新类别的改进适应能力，超过了现有的最先进基线模型。

### Adaptive Decision Boundary for Few-Shot Class-Incremental Learning

### 自适应决策边界在少样本类别增量学习中的应用

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

少样本类别增量学习（FSCIL）旨在从有限的训练样本中持续学习新类别，同时不遗忘之前已学习类别的知识。传统的FSCIL方法通常在基础训练阶段利用大量训练样本构建一个强大的特征提取器，然后在后续的增量阶段冻结该提取器，仅微调分类器。然而，当前的策略主要集中在防止灾难性遗忘，仅考虑新类别与基础类别的关系，而忽略了每个类别的具体决策空间。为了解决这一挑战，我们提出了一种即插即用的自适应决策边界策略（ADBS），该策略与大多数FSCIL方法兼容。具体而言，我们为每个类别分配一个特定的决策边界，并在训练过程中自适应地调整这些边界，以最优地优化决策。

### Few-Shot Incremental Learning via Foreground Aggregation and Knowledge Transfer for Audio-Visual Semantic Segmentation

### 通过前景聚合和知识迁移实现音频-视觉语义分割的少样本增量学习

Audio-Visual Semantic Segmentation (AVSS) has gained significant attention in the multi-modal domain, aiming to segment video objects that produce specific sounds in the corresponding audio. Despite notable progress, existing methods still struggle to handle new classes not included in the original training set. To this end, we introduce Few-Shot Incremental Learning (FSIL) to the AVSS task, which seeks to seamlessly integrate new classes with limited incremental samples while preserving the knowledge of old classes. Two challenges arise in this new setting: (1) To reduce labeling costs, old classes within the incremental samples are treated as background, similar to silent objects. Training the model directly with background annotations may worsen the loss of distinctive knowledge about old classes, such as their outlines and sounds. (2) Most existing models adopt early cross-modal fusion with a single-tower design, incorporating more characteristics into class representations, which impedes knowledge transfer between classes based on similarity. To address these issues, we propose a Few-shot Incremental learning framework via class-centric foregrouNd aggreGation and dual-tower knowlEdge tRansfer (FINGER) for the AVSS task, which comprises two targeted modules: (1) The class-centric foreground aggregation gathers class-specific features for each foreground class while disregarding background features. The background class is excluded during training and inferred from the foreground predictions. (2) The dual-tower knowledge transfer postpones cross-modal fusion to separately conduct knowledge transfer for each modality. Extensive experiments validate the effectiveness of the FINGER model, significantly surpassing state-of-the-art methods.

视听语义分割（AVSS）在多模态领域受到了广泛关注，其目标是分割视频中产生特定声音的物体。尽管取得了显著进展，现有方法在处理原始训练集中未包含的新类别时仍面临挑战。为此，我们引入了少样本增量学习（FSIL）到AVSS任务中，旨在在有限的增量样本中无缝集成新类别，同时保留旧类别的知识。在这一新设置中，出现了两个挑战：（1）为了减少标注成本，增量样本中的旧类别被视为背景，类似于无声物体。直接使用背景标注训练模型可能会加剧旧类别独特知识（如轮廓和声音）的损失。（2）大多数现有模型采用早期跨模态融合的单塔设计，将更多特征融入类别表示中，这阻碍了基于相似性的类别间知识迁移。为了解决这些问题，我们提出了一种面向类别的前景聚合和双塔知识迁移的少样本增量学习框架（FINGER），用于AVSS任务，该框架包含两个目标模块：（1）面向类别的前景聚合模块收集每个前景类别的特定特征，同时忽略背景特征。训练过程中排除背景类别，并从前景预测中推断背景类别。（2）双塔知识迁移模块推迟跨模态融合，分别对每种模态进行知识迁移。大量实验验证了FINGER模型的有效性，显著超越了现有最先进方法。

# IJCAI

## 2024

### FineFMPL: Fine-grained Feature Mining Prompt Learning for Few-Shot Class Incremental Learning

### FineFMPL：细粒度特征挖掘提示学习用于少样本类别增量学习

Few-Shot Class Incremental Learning (FSCIL) aims to continually learn new classes with few training samples without forgetting already learned old classes. Existing FSCIL methods generally fix the backbone network in incremental sessions to achieve a balance between suppressing forgetting old classes and learning new classes. However, the fixed backbone network causes insufficient learning of new classes from a few samples. Benefiting from the powerful visual and textual understanding ability of Vision-Language (VL) pre-training models, we propose a Fine-grained Feature Mining Prompt Learning (FineFMPL) approach to adapt the VL model to FSCIL, which comprehensively learns and memorizes fine-grained discriminative information of emerging classes. Concretely, the visual probe prompt is firstly proposed to guide the image encoder of VL model to extract global-level coarse-grained features and object-level fine-grained features, and visual prototypes are preserved based on image patch significance, which contains the discriminative characteristics exclusive to the class. Secondly, the textual context prompt is constructed by cross-modal mapping of visual prototypes, feeding into the text encoder of VL model to memorize the class information as textual prototypes. Finally, integrating visual and textual prototypes based on fine-grained feature mining into the model improves the recognition performance of all classes in FSCIL. Extensive experiments on three benchmark datasets demonstrate that our FineFMPL achieves new state-of-the-art. The code is available at https://github.com/PKU-ICST-MIPL/FineFMPL\_IJCAI2024.

少样本类别增量学习（FSCIL）旨在在不遗忘已学习旧类别的前提下，通过少量训练样本持续学习新类别。现有的FSCIL方法通常在增量学习阶段固定骨干网络，以在抑制遗忘旧类别和学习新类别之间取得平衡。然而，固定的骨干网络导致从少量样本中学习新类别时学习不足。得益于视觉-语言（VL）预训练模型强大的视觉和文本理解能力，我们提出了一种细粒度特征挖掘提示学习（FineFMPL）方法，将VL模型适应于FSCIL，全面学习和记忆新兴类别的细粒度判别信息。  
  
具体而言，首先提出了视觉探针提示，引导VL模型的图像编码器提取全局级别的粗粒度特征和对象级别的细粒度特征，并基于图像块的重要性保存视觉原型，这些原型包含类别的独特判别特征。其次，通过视觉原型的跨模态映射构建文本上下文提示，输入VL模型的文本编码器，以文本原型的形式记忆类别信息。最后，通过将细粒度特征挖掘的视觉和文本原型整合到模型中，提高了FSCIL中所有类别的识别性能。在三个基准数据集上的广泛实验表明，我们的FineFMPL方法达到了新的最先进水平。代码可在https://github.com/PKU-ICST-MIPL/FineFMPL\_IJCAI2024 获取。

### Delve into Base-Novel Confusion: Redundancy Exploration for Few-Shot Class-Incremental Learning

### 探究基础-新类混淆：少样本类别增量学习中的冗余探索

Few-shot class-incremental learning (FSCIL) aims to acquire knowledge from novel classes with limited samples while retaining information about base classes. Existing methods address catastrophic forgetting and overfitting by freezing the feature extractor during novel-class learning. However, these methods usually tend to cause the confusion between base and novel classes, i.e., classifying novel-class samples into base classes.In this paper, we delve into this phenomenon to study its cause and solution. We first interpret the confusion as the collision between the novel-class and the base-class region in the feature space.Then, we find the collision is caused by the label-irrelevant redundancies within the base-class feature and pixel space. Through qualitative and quantitative experiments, we identify this redundancy as the shortcut in the base-class training, which can be decoupled to alleviate the collision. Based on this analysis, to alleviate the collision between base and novel classes, we propose a method for FSCIL named Redundancy Decoupling and Integration (RDI). RDI first decouples redundancies from base-class space to shrink the intra-base-class feature space. Then, it integrates the redundancies as a dummy class to enlarge the inter-base-class feature space. This process effectively compresses the base-class feature space, creating buffer space for novel classes and alleviating the model's confusion between the base and novel classes. Extensive experiments across benchmark datasets, including CIFAR-100, miniImageNet, and CUB-200-2011 demonstrate that our method achieves state-of-the-art performance.

少样本类别增量学习（FSCIL）旨在在有限样本的情况下从新类别中获取知识，同时保留基础类别的信息。现有方法通过在新类别学习过程中冻结特征提取器来解决灾难性遗忘和过拟合问题。然而，这些方法通常会导致基础类别和新类别之间的混淆，即将新类别样本错误分类为基础类别样本。在本文中，我们深入研究了这一现象，探讨其原因和解决方案。我们首先将这种混淆解释为特征空间中新类别区域与基础类别区域之间的冲突。接着，我们发现这种冲突是由基础类别特征和像素空间中的标签无关冗余引起的。通过定性和定量实验，我们确定这种冗余是基础类别训练中的捷径，可以通过解耦来缓解冲突。基于这一分析，为了缓解基础类别和新类别之间的冲突，我们提出了一种名为冗余解耦与集成（RDI）的方法。RDI首先从基础类别空间中解耦冗余，以缩小基础类别特征空间。然后，将这些冗余作为虚拟类别进行集成，以扩大基础类别之间的特征空间。这一过程有效地压缩了基础类别特征空间，为新类别创建了缓冲空间，从而缓解了模型在基础类别和新类别之间的混淆。广泛的实验，包括在CIFAR-100、miniImageNet和CUB-200-2011等基准数据集上的实验，证明了我们的方法达到了最先进的性能。