



# 计算机科学

COMPUTER SCIENCE

## 基于聚类优化学习的少样本图像分类

苏如祺, 卞雄, 朱松豪

引用本文

苏如祺, 卞雄, 朱松豪. 基于聚类优化学习的少样本图像分类[J]. 计算机科学, 2024, 51(6A): 230300227-7.

SU Ruqi, BIAN Xiong, ZHU Songhao. [Few-shot Images Classification Based on Clustering Optimization Learning](#) [J]. Computer Science, 2024, 51(6A): 230300227-7.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [PosNet:基于位置的因果关系抽取网络](#)

PosNet:Position-based Causal Relation Extraction Network

计算机科学, 2022, 49(12): 305-311. <https://doi.org/10.11896/jsjx.211100264>

### [基于自注意力的自监督深度聚类算法](#)

Self-supervised Deep Clustering Algorithm Based on Self-attention

计算机科学, 2022, 49(3): 134-143. <https://doi.org/10.11896/jsjx.210100001>

### [一种面向动态科研网络的社区检测算法](#)

Community Detection Algorithm for Dynamic Academic Network

计算机科学, 2022, 49(1): 89-94. <https://doi.org/10.11896/jsjx.210100023>

### [融合频率和通道卷积注意的脑电\(EEG\)情感识别](#)

EEG Emotion Recognition Based on Frequency and Channel Convolutional Attention

计算机科学, 2021, 48(12): 312-318. <https://doi.org/10.11896/jsjx.201000141>

### [三角坐标系下人脸表情表示方法](#)

New Representation of Facial Affect Based on Triangular Coordinate System

计算机科学, 2020, 47(6A): 250-253. <https://doi.org/10.11896/JsJkx.190700081>

# 基于聚类优化学习的少样本图像分类

苏如祺 卞雄 朱松豪

南京邮电大学自动化学院、人工智能学院 南京 210023

(2276300097@qq.com)

**摘要** 少样本图像分类的目标是在训练少量标记训练数据的基础上实现新类别图像的分类,然而这一目的在现有条件下很难实现。因此,目前的少样本学习方法主要借鉴迁移学习的思想,其核心是利用情景训练式的元训练构建先验知识,从而解决未知新任务。然而,研究工作表明,相较于复杂的少样本学习方法,具有强大特征表示的嵌入模型学习方法更为简单、有效。受此启发,提出一种新的基于直推式聚类优化学习的少样本图像分类方法。该方法首先利用样本数据的内部特征结构信息实现每个类别的综合表示;然后优化每个类别的中心,形成更具区别性的特征表示,从而有效增加不同类别之间的特征差异。大量实验结果表明,所提的基于直推式聚类优化学习的少镜头图像分类方法有效提高了各种训练条件下的图像分类精度。

**关键词:** 少样本图像分类;特征表示;聚类优化

**中图分类号** TP391.41

## Few-shot Images Classification Based on Clustering Optimization Learning

SU Ruqi, BIAN Xiong and ZHU Songhao

College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

**Abstract** The goal of few-shot image classification is to achieve the classification of new image categories on the basis of training a small number of labeled training dataset. However, this goal is difficult to achieve under existing conditions. Therefore, the current few-shot learning method mainly draws on the idea of transfer learning, and its core is to construct prior knowledge by using situational meta-training, so as to realize the solution of unknown new tasks. However, the latest research shows that the embedded model learning method with strong feature representation is simpler and more effective than the complex few-shot learning method. Inspired by this, this paper proposes a novel few-shot image classification method based on direct clustering optimization learning. This proposed method first utilizes the internal feature structure information of sample data to realize the comprehensive representation of each category, and then optimizes the center of each category to form a more distinctive feature representation, thus effectively increasing the feature differences between different categories. A large number of experimental results demonstrate that the proposed image classification method based on the clustering optimization learning can effectively improve the accuracy of image classification under various training conditions.

**Keywords** Few-shot image classification, Feature representation, Clustering optimization

计算机视觉的成功离不开大规模标注数据集的支持<sup>[1-2]</sup>,但当标注数据不足时,深度学习最易产生的问题就是过拟合,这也是少样本学习所面临的主要挑战。从实际应用角度而言,少样本学习更符合真实世界的应用场景。一般来说,收集大规模标注数据集的时间成本和人力成本都是非常可观的,因此更好地利用标注数据集提高模型性能是非常合理且十分必要的。为此,人们提出许多方法解决少样本学习问题,通过学习模型在多个不同任务上的先验知识,在少样本下研究其对新任务的适应性。

元学习也称为学会学习,其目的是希望网络能够利用以往的经验积累具备快速学习的能力,这与少样本学习中要求利用少量标注样本完成适应性训练的目标相对应。因此,大部分少样本学习方法通过元学习方法构建先验知识<sup>[3-4]</sup>,也即元学习方法通过大量相似任务获取元知识,学习一组初始化

参数,进而利用这些参数实现模型适应少样本的新任务。最近研究表明,基于元学习方法的情景训练对数据利用的效率较低,这是因为元学习没有对训练数据特征进行充分处理<sup>[5]</sup>。一些有监督的少样本学习研究也表明,那些采用预训练和微调方法的性能要优于许多复杂的元学习方法<sup>[6-7]</sup>。

在少样本学习的训练阶段过程中出现的以下两种情况需要着重加以关注。1)归纳式少样本,即在训练过程中只有少数标记样本可用,且在推理过程中对这些标记样本逐个进行评估。此种情况下,主要依赖于简单方法实现,如余弦分类器<sup>[8]</sup>、最近类均值<sup>[9]</sup>和逻辑回归<sup>[10]</sup>。2)直推式少样本,即在测试过程中所有测试样本的特征信息均可用于类别判断。已有研究表明,相比于归纳式少样本学习方法,直推式少样本学习方法更为有效。在直推式场景中,常用的方法包括聚类算法<sup>[11]</sup>和嵌入传播<sup>[12]</sup>等。

基金项目:国家自然科学基金(62001247)

This work was supported by the National Natural Science Foundation of China(62001247).

通信作者:朱松豪(zhush@njupt.edu.cn)

本文提出一种新颖的直推式聚类优化学习的少镜头图像分类方法。具体来说,该方法首先利用数据增强挖掘数据的内在特征,然后构建一片内存缓存区记录各类特征并缓慢更新,最后以缓存数据特征为模板,优化样本间的相似性并保持各类别间的差异。如图1所示,高质量的特征空间对未知类别具有更好的泛化性能,且可有效避免已知类别的过度拟合;此外,在归纳少样本模式下采用简单的逻辑回归,以验证模型可靠性,而在直推式少样本模式下采用半监督聚类的方式,并对初始中心点进行调整,在迭代过程中不断对中心点进行修正,以提高模型的精度。



图1 通过数据增强补充类别信息来处理监督学习导致的坍塌问题,再构建内存缓存区添加并更新信息以保持样本相似性和类别差异性

Fig. 1 To address the issue of collapsing caused by supervised learning, class information is supplemented through data augmentation. Furthermore, a memory buffer is constructed to add and update information, ensuring the preservation of sample similarity and class distinctiveness

综上所述,本文贡献如下:

- 1) 提出一种新颖的有监督少样本学习的方法,利用历史特征优化类别之间的距离,进而形成更具鲁棒性的特征表示,以此提高对未知类别的泛化性能。
- 2) 在直推式少样本模式下,提出一种新颖的聚类样本选择和聚类中心更新的思路,从而进一步提高模型的性能。
- 3) 在公开少样本数据集上进行了大量的实验,实验结果表明所提网络模型显著提高了少样本图像分类的性能。

## 1 相关工作

### 1.1 少样本学习

少样本学习旨在通过使用给定的少数例子快速适应新任务。传统的深度学习需要大量数据进行训练,当样本不足时,会出现过拟合。因此,大多数少样本学习基于元学习框架,通过各种任务的学习,以便快速适应新任务。基于元学习的少样本学习方法大致可分为以下两类:基于测量的方法和基于优化的方法。在第一种方法中,文献[13]将每个类别的均值作为其对应类别的特征表示,并利用欧氏距离实现样本分类;文献[14]利用深度卷积网络代替包括欧几里德距离在内的传统度量函数,进一步提高系统性能;文献[15]利用k近邻算法获得图像间的像素级相似度;文献[16]利用可微地球距离计算图像块间的相似度。在第二种方法中,文献[3]利用潜在优化策略在低维空间学习模型参数的生成分布,从而生成高维参数;文献[17]提出一种通用优化算法,用以确定最佳的初始化参数,并通过一个或几个样本快速适应新任务;文献[18]通过一阶导数近似文献[17]提出的二阶导数;文献[19]提出的模型结合了上述两种主流方法的思想。

### 1.2 对比学习

本质上,对比学习是一种自监督学习,用于学习没有标注

数据集的一般特征。对比学习的目的是最大化相似图像间的相似性,同时最小化不同图像间的一致性。也即,对比学习可同时学习正样本和负样本,但其困难之处在于构建正样本和负样本。文献[20]将图像本身视为正样本,将其他图像视为负样本,并将特征进行存储以备后续使用。文献[21]从同一批小样本中选择正样本和负样本。文献[22]将对对比学习问题视为一个字典查询问题,并引入一个队列和一个动量编码器学习样本间的关系。近年来,研究人员将对对比学习引入少样本学习框架,并取得了令人满意的效果。文献[11]的实验结果表明,去除背景信息可有效提高少样本学习的准确性和鲁棒性。文献[23]的实验结果表明,降低背景与目标间的相关性,可学习更多的辨别性信息。文献[24]的核心思想是使相似类别间的边界距离大于不同类别间的边界距离。

### 1.3 半监督学习

半监督学习利用未标记数据和已标记数据提高分类性能,因此有必要研究每个未标记实例的标记属性。文献[25]利用一致性的正则化方法对未标记数据施加分类不变损失,以提高分类的准确性。文献[26]利用分类重平衡采样器,改进现有的分类平衡数据方法,利用标记数据分布频繁对少数类别样本进行采样。最近,有研究人员试图将半监督学习整合到少样本学习框架中。文献[27]改进文献[13]的原型网络,使用未标记图像进行细化。文献[28]利用图神经网络从类似样本中收集特征信息。一定程度上,半监督学习中广泛应用的低密度假设在少样本场景中难以实现,这使得半监督的少样本学习变得更为困难。

## 2 提出的方法

本章首先描述了少样本学习中的相应术语和基本问题,然后详细介绍了本文提出的方法。

### 2.1 问题描述

在少样本学习中,通常使用  $N$ -way  $K$ -shot 的分类结果作为模型性能的评估标准,其中  $N$  表示类别的数量,每个类别有  $K$  个例子, $N$  和  $K$  的大小决定了问题的难度。假设现有一组数据集  $S = \{S_{\text{base}}, S_{\text{novel}}\}$ , 且  $S_{\text{base}} \cap S_{\text{novel}} = \emptyset$ , 一个基础数据集  $D_{\text{base}} = \{(x_i, y_i), x_i, y_i \in S_{\text{base}}\}$ , 一组新类数据集  $D_{\text{novel}} = \{(x_i, y_i), x_i, y_i \in S_{\text{novel}}\}$ 。如前所述,少样本学习需要先验知识或外部知识完成学习。训练阶段,少样本学习模型利用基础数据集  $D_{\text{base}}$  完成特征提取器的训练;测试阶段,利用训练后的特征提取器表示查询集样本,支持集  $d_s$  共有  $N \times K$  个样本,提取器提取样本特征,并根据样本间的特征关系实现少样本图像的分类。具体而言,随机从  $D_{\text{novel}}$  中选出  $N$  个类别,每个类别选出  $K+Q$  个样本,组成元测试集  $T = \langle d_s, d_q \rangle$ , 其中  $d_s$  表示支持集样本,查询集  $d_q$  共有  $N \times Q$  个样本,且  $d_s \cap d_q = \emptyset$ 。测试过程中,首先利用支持集  $d_s$  的图像标签信息,训练一个简单的逻辑回归分类器;然后利用该分类器对查询集  $d_q$  的标签信息进行推断;最后在多个元测试任务上计算查询集的平均分类精度。

### 2.2 预训练

#### 2.2.1 分类损失

与元学习的训练目的不同,有监督的少样本学习旨在构建强大的特征表示,并在新类中具备很好的泛化能力。有监督的少样本学习主要利用交叉熵损失学习类内信息,目的是

将基础数据集上的分类损失减少到零。文献[29]的实验结果表明,在有监督的训练过程中,普遍存在特征崩溃现象,其中类内样本的特征表征易崩溃到相应类别的特征均值。这虽然有利于基类识别,但会影响少样本学习对于新类的泛化性能。如图2所示,结合分类损失和旋转损失增强类内样本特征表征,以学习适合于少样本图像分类任务的通用特征表征。预测图像旋转量的过程如下:首先,利用旋转变换函数  $T(\cdot)$  获得训练图像  $X$  的增强样本  $X_T$ ;然后,将增强后的图像合并为单个张量  $X_{new} = \{x^0, x^{90}, x^{180}, x^{270}\}$ ;最后,采取四分类预测当前图像的旋转角度。通过最小化增广数据集上的交叉熵

损失学习通用的特征表征,并将相应的损失函数分别定义为式(1)和式(2):

$$L_b = -\sum_i^n y_i \log \hat{y}_i, y_i \in Y, \hat{y}_i \in \text{Sigmoid}(\hat{Y}) \quad (1)$$

其中,  $Y$  表示真实的标签信息,  $\hat{Y}$  表示模型输出的标签信息,  $\hat{y}_i$  表示分类器获得的类别标签信息。

$$L_r = -\sum_i^n u_i \log \hat{u}_i, u_i \in U, \hat{u}_i \in \text{Sigmoid}(\hat{U}) \quad (2)$$

其中,  $U$  表示真实的旋转角度信息,  $\hat{U}$  表示模型输出的旋转角度信息,  $\hat{u}_i$  表示分类器获得的旋转角度信息。

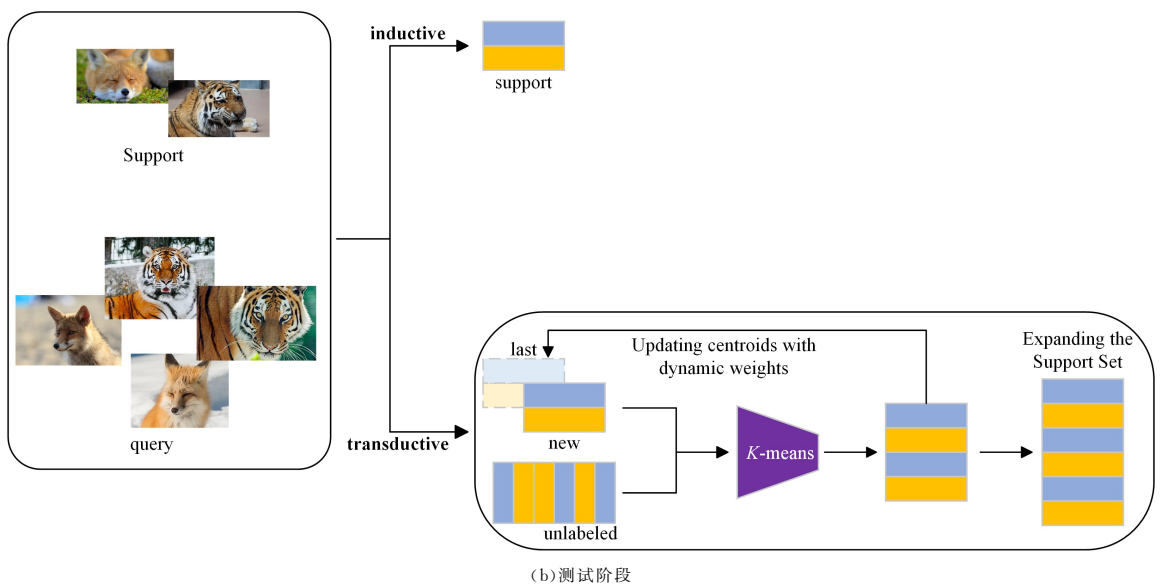
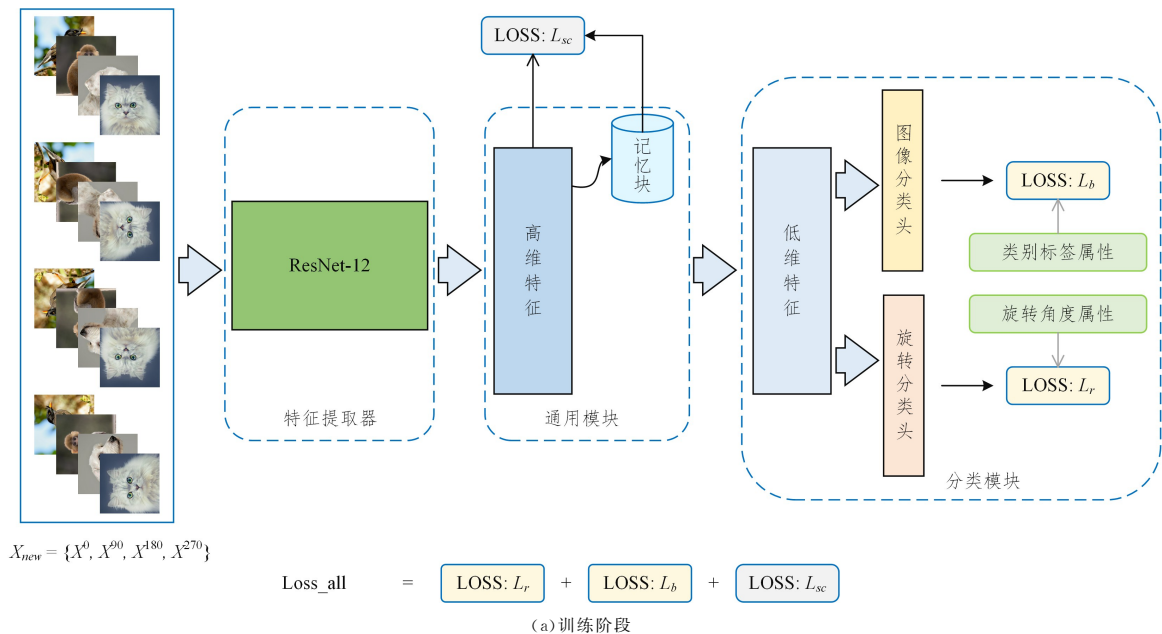


图2 所提网络的整体框架

Fig. 2 Overall framework of the proposed network

总体优化目标函数描述如下:

$$\arg \min_{\theta, \varphi} L_{b+r}(D_{E(\text{base})}) + R(\theta) + R(\varphi) \quad (3)$$

其中,  $L$  表示损失函数,  $R$  表示正则项,  $\theta$  和  $\varphi$  分别表示提取器和分类器的参数,  $D_{E(\text{base})}$  表示旋转增强后的数据集。

### 2.2.2 约束聚类

交叉熵损失由于采取类间竞争机制,因而往往只关注那些正确预测标签的准确性,而忽略了那些错误预测标签间的

差异性。这些被忽略的差异性对于少样本学习同样重要,强大的先验知识不应只关注类内相似性,还应关注类别间的差异性,使得模型具备将新样本映射到更具判别性特征空间的能力。为补充不同类间的信息,通过特征缓冲区构建内存空间。具体而言,定义一个特征内存缓冲区  $Q_l \in R^{C \times D}$ ,其中  $C$  为类别数,  $D$  为嵌入的维度数。约束聚类的具体过程描述如下:首先,将训练样本的特征与缓冲区的记忆特征进行比较,

最大限度提高同一类别样本间的相似性,最小化不同类别样本间的一致性;然后,根据训练样本的特征缓慢更新内存缓冲区,即将缓存区特征逐渐更新为各个类别中心,并在特征空间中形成不同的类别,提高模型的泛化性能。

$$\begin{cases} L_{sc} = \sum_{i=1}^M L_{sc}^i; i \in X_{batch\_size}, j \in Q_{f\_last} \\ L_{sc}^i = -\frac{1}{M_{y_i} - 1} \sum_{j=1}^M l_{i \neq j} l_{y_i = y_j} \log_e \left[ \frac{\exp(s_{i,j}/t)}{\exp(s_{i,j}/t) + \sum_{k=1, k \neq y_i}^M \exp(s_{i,k}/t)} \right] \end{cases} \quad (5)$$

其中,  $M$  为批训练的次数,  $y_i$  和  $y_j$  分别为样本  $i$  和  $j$  的标签信息,  $s_{i,j}$  为样本  $i$  和  $j$  特征的相似度,  $t$  为超参数。此外,  $l_{i \neq j}$  表示不同的样本;  $l_{i \neq j} l_{y_i = y_j}$  表示来自同一类别的不同样本。

## 2.3 微调

### 2.3.1 权重动态调整的 K 均值

K 均值聚类算法是一种迭代求解的聚类分析算法。首先,随机选取  $K$  个对象分别作为初始聚类中心;然后,计算其余对象与聚类中心间的距离,并按距离最小原则将每个对象进行类别划分,重新计算聚类中心;最后,反复迭代,直至聚类中心点不再变化。需要注意的是, K 均值聚类效果依赖于聚类中心的初始化,且对噪声和异常值非常敏感。

针对上述问题,本文对 K 均值聚类算法进行如下修改:首先,以各支持集样本及其最近的  $N$  个邻居的均值作为初始聚类中心;然后,利用加权平均和更新聚类中心;其次,样本与聚类中心的距离矩阵的方差越小,代表该样本距离聚类中心越近,则赋予该样本的距离权重越大;最后,反复迭代,直至整个过程收敛。

### 2.3.2 样本选择

聚类算法易陷入局部最优,进而导致聚类后的样本类别并不一定可靠。为避免各类样本数不均衡带来的偏差,本文采用如下策略选择样本:首先,将聚类中心视为原始锚点,并将其与最近的样本连接为星状结构,当不同星形交接时终止样本的选择;然后,利用与锚点相连的样本构建增强数据集,并进行逻辑回归分类器的训练,完成分类测试。实验结果表明,放弃那些处于边缘未被选择的样本,有利于缓解异常样本导致的负面影响。

## 3 实验

### 3.1 实验设置

为评估所提网络模型的性能,在 3 个公开的少样本图像数据集上进行实验:文献[4]中的 CIFAR-FS 数据集,文献[30]中的 MiniImageNet 数据集和文献[31]中的 FC100 数据集。MiniImageNet 数据集是 ImageNet 数据集的子集,而另外两个数据集是 CIFAR100 数据集的子集。与文献[32]一样,MiniImageNet 数据集共有 100 个类别,且分别选择 64 个类别、16 个类别和 20 个类别进行训练、验证和测试。CIFAR-FS 数据集共有 100 个类别,其中分别选择了 64 个类别、16 个类别和 20 个类别进行训练、验证和测试。FC100 数据集也有 100 个类别,其中分别选择 60 个类别、20 个类别和 20 个类别进行培训、验证和测试。

与文献[33]一样,选择 resnet-12 作为骨干网络。具体而言,骨干网络含有 4 个残差块,卷积核数分别设置为 64, 160, 320 和 640;每个残差块含有 3 个  $3 \times 3$  卷积核,第三个残差块

缓存区特征更新公式描述如下:

$$Q_{f\_cur} = \alpha \cdot Q_{f\_last} + (1 - \alpha) \cdot F(X_{batch\_size}) \quad (4)$$

其中,  $F(\cdot)$  为特征提取器,  $\alpha$  为调整更新速率的超参数。

优化目标函数描述如下:

后进行最大池化,第四个残差块后进行全局平均;选择 dropout 为正则化器。此外,所有实验中  $n=5$ ,  $s=1$  或  $5$ ,  $q=15$ , 且进行 600 次随机实验,用以获得平均准确性分数,置信度为 95%。

本次实验将随机梯度下降优化器的动量和权重衰减分别设置为 0.9 和  $5 \times 10^{-4}$ , 初始学习率设置为 0.05。每批 CIFAR-FS 和 FC100 数据集的训练样本数均为 64, 每批 MiniImageNet 的训练样本数为 32, 且所有数据集的训练周期为 100。所有样本数据通过随机裁剪、颜色抖动、随机水平翻转、同步翻转等方法进行增强。训练一个  $n$  路分类评估网络性能。

### 3.2 实验结果

表 1—表 4 分别列出了 3 个公共数据集的实验结果,性能评估采用 600 次随机抽样测试结果的平均精度。其中,  $\dagger$  表示训练期间使用了额外的验证数据集。

表 1 在归纳设置下, CIFAR-FS 数据集的实验结果

Table 1 Experimental results on the CIFAR-FS dataset under the inductive setting

Model	Backbone	5-way 1-shot	5-way 5-shot
Proto-Net <sup>†</sup> [13]	64-64-64-64	55.50 ± 0.70	72.00 ± 0.60
RFS-Simple <sup>[7]</sup>	ResNet-12	71.50 ± 0.80	86.00 ± 0.50
Proto-Net <sup>†</sup> [13]	ResNet-12	72.20 ± 0.70	83.50 ± 0.50
MetaOptNet <sup>[34]</sup>	ResNet-12	72.60 ± 0.70	84.30 ± 0.50
RFS-Distill <sup>[7]</sup>	ResNet-12	73.90 ± 0.80	86.90 ± 0.50
FewTure <sup>[35]</sup>	ViT-Small	76.10 ± 0.88	86.16 ± 0.64
Fine-tuning <sup>[6]</sup>	WRN-28-10	76.58 ± 0.68	85.79 ± 0.50
Ours	ResNet-12	<b>77.82 ± 0.78</b>	<b>88.69 ± 0.57</b>

表 2 在归纳设置下, FC100 数据集的实验结果

Table 2 Experimental results on the FC100 dataset under the inductive setting

Model	Backbone	5-way 1-shot	5-way 5-shot
Proto-Net <sup>†</sup> [13]	64-64-64-64	35.30 ± 0.60	48.60 ± 0.60
Proto-Net <sup>†</sup> [13]	ResNet-12	37.50 ± 0.60	52.50 ± 0.60
MetaOptNet <sup>[34]</sup>	ResNet-12	41.10 ± 0.60	55.50 ± 0.60
RFS-Simple <sup>[7]</sup>	ResNet-12	42.60 ± 0.70	59.10 ± 0.60
Fine-tuning <sup>[6]</sup>	WRN-28-10	43.16 ± 0.59	57.57 ± 0.55
RFS-Distill <sup>[7]</sup>	ResNet-12	44.60 ± 0.70	60.90 ± 0.60
FewTure <sup>[35]</sup>	ViT-Small	46.20 ± 0.79	63.14 ± 0.73
Ours	ResNet-12	<b>46.51 ± 0.79</b>	<b>63.56 ± 0.71</b>

表 3 在归纳设置下, MiniImageNet 数据集的实验结果

Table 3 Experimental results on the MiniImageNet dataset under the inductive setting

Model	Backbone	5-Way1-shot	5-way 5-shot
Proto-Net <sup>†</sup> [13]	64-64-64-64	49.42 ± 0.78	68.20 ± 0.66
Fine-tuning <sup>[6]</sup>	WRN-28-10	57.73 ± 0.62	78.17 ± 0.49
RFS-Simple <sup>[7]</sup>	ResNet-12	62.02 ± 0.63	79.64 ± 0.44
MetaOptNet <sup>[34]</sup>	ResNet-12	62.64 ± 0.61	78.63 ± 0.46
RFS-Distill <sup>[7]</sup>	ResNet-12	64.82 ± 0.60	82.14 ± 0.43
FewTure <sup>[35]</sup>	ViT-Small	68.02 ± 0.88	84.51 ± 0.53
Ours	ResNet-12	<b>68.57 ± 0.78</b>	<b>84.27 ± 0.52</b>

表4 直推式设置下的3个数据集的实验结果

Table 4 Experimental results on three datasets under the transductive setting

Model	MiniImageNet		CIFAR-FS		FC100	
	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
DPGN <sup>[36]</sup>	67.77±0.32	84.60±0.43	77.90±0.40	90.02±0.40	—	—
SSR <sup>[37]</sup>	68.10±0.60	76.90±0.40	76.80±0.60	83.70±0.40	—	—
Fine-tuning <sup>[6]</sup>	68.11±0.69	80.36±0.50	78.36±0.70	87.54±0.49	50.44±0.68	65.74±0.60
iLPC <sup>[38]</sup>	69.79±0.99	79.82±0.55	77.14±0.95	85.23±0.55	—	—
ECKPN <sup>[39]</sup>	70.48±0.38	85.42±0.46	79.20±0.40	91.00±0.50	—	—
EASY <sup>[40]</sup>	82.31±0.24	88.57±0.12	86.99±0.21	90.20±0.15	54.47±0.24	65.82±0.19
Ours	<b>83.42±0.88</b>	<b>88.60±0.54</b>	<b>87.84±1.02</b>	<b>92.14±0.63</b>	<b>54.82±1.10</b>	<b>66.69±0.87</b>

本文以 RFS-Simple 网络作为基本骨架,并用第3节提出的方法对其进行改进。从表1—表4的实验结果可以看出,本文所提方法对于少样本图像分类性能改善效果较为明显。具体而言,从表1可以看出,相较于 RFS-Distill 方法,本文所提方法的分类性能在 5-way 1-shot 和 5-way 5-shot 分别提高了 3.92% 和 1.79%;从表2可以看出,相较于 RFS-Distill 方法,本文所提方法的分类性能在 5-way 1-shot 和 5-way 5-shot 分别提高了 1.91% 和 2.66%;从表3可以看出,相较于 RFS-Distill 方法,本文所提方法的分类性能在 5-way 1-shot 和 5-way 5-shot 分别提高了 3.75% 和 2.13%。如表4所列,在直推式设置中,本文所提方法的分类性能依然最好。表1—表4的实验结果表明,训练数据的判别特征分布信息可以有效地提高模型的泛化性能。

### 3.3 消融实验

在3个公共数据集 CIFAR-FS, FC100 和 MiniImageNet 上进行了消融实验,用以分析不同模块对实验性能的影响。

表5 在 CIFAR-FS, FC100 和 MiniImageNet 数据集上的消融实验

Table 5 Ablation experiments on the CIFAR-FS, FC100 and MiniImageNet datasets

Model	CIFAR-FS		FC100		MiniImageNet	
	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
Baseline	73.50±0.80	86.00±0.50	42.60±0.70	59.10±0.60	63.02±0.63	79.64±0.44
Clustering	74.20±0.75	87.00±0.50	43.18±0.69	60.24±0.65	64.92±0.60	80.90±0.50
Rotation	75.02±0.83	87.69±0.58	44.24±0.76	61.16±0.73	65.36±0.79	82.54±0.54
Our(inductive)	<b>77.82±0.78</b>	<b>88.69±0.57</b>	<b>46.51±0.79</b>	<b>63.56±0.71</b>	<b>68.57±0.78</b>	<b>84.27±0.52</b>
Our(k-means)	83.92±0.99	90.12±0.53	52.32±1.16	65.55±0.85	80.55±0.82	86.91±0.56
Our(Dynamic-Weightk-means)	<b>87.84±1.02</b>	<b>92.14±0.63</b>	<b>54.82±1.10</b>	<b>66.69±0.87</b>	<b>83.42±0.88</b>	<b>88.55±0.54</b>

### 3.4 T-SNE 分析

为了更好地区分分类结果,本文利用 T-SNE 对 CIFAR-FS 数据集上随机选取的5个类别的150张图像的高维特征分布信息进行了可视化。从图3所示的实验结果中可知,基准方法

中,每个类别的边界都不那么精确和紧凑,特别是当种子为5时,不同类别的特征分布相互重叠。与基准 RFS-Simple 方法相比,本文所提方法在不同种子下的性能更好。也即,本文所提方法有助于形成更为紧凑的簇和有着明显边界的类别。

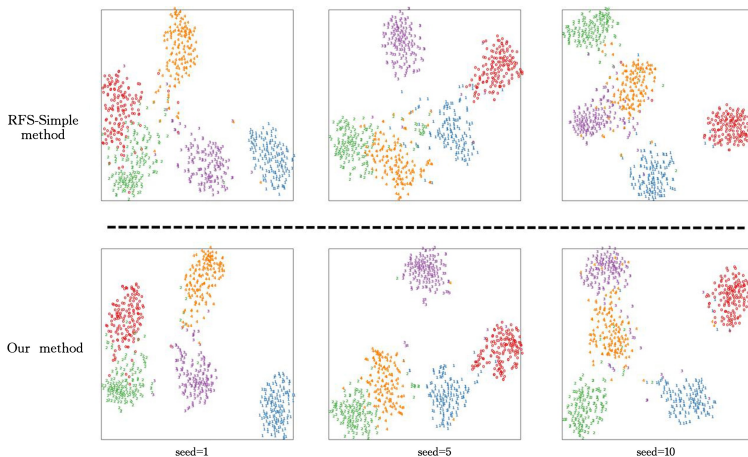


图3 CIFAR-FS 数据集上 T-SNE 的可视化结果

Fig. 3 Visualization results of T-SNE on the CIFAR-FS dataset

**结束语** 本文提出了一种基于直推式聚类优化学习的少样本图像分类,用于解决少样本故障分类问题。训练阶段利用代理任务提高了模型捕获样本潜在描述性结构信息的能力,并利用特征缓冲区补充类间信息完成了类内样本的聚集和类间样本的发散;测试阶段利用动态权重调整聚类中心,有效减轻了异常样本对聚类的负面影响。实验结果表明,所提出的网络模型有效提高了少样本图像的分类性能和泛化性能。

## 参 考 文 献

- [1] HE K M, GKIOXARI G, DOLL'AR P, et al. Mask R-CNN [C]//IEEE Conference on Computer Vision, 2017:2961-2969.
- [2] SU X, HUANG T, LI Y L, et al. Prioritized architecture sampling with monte-carlo tree search[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2021:10968-10977.
- [3] RUSU A A, RAO D, SYGNOWSKI J, et al. Meta-learning with latent embedding optimization[C]//International Conference on Learning Representations. 2019:1-17.
- [4] BERTINETTO L, HENRIQUES J F, TORR P H S, et al. Meta-learning with differentiable closed-form solvers[C]// International Conference on Learning Representations. 2019:1-15.
- [5] LAENEN S, BERTINETTO L. On episodes, prototypical networks, and few-shot learning[C]// Annual Conference on Neural Information Processing Systems. 2021:24581-24592.
- [6] DHILLON G S, CHAUDHARI P, RAVICHANDRAN A, et al. A baseline for few-shot image classification[C]// International Conference on Learning Representations. 2020:1-20.
- [7] TIAN Y L, WANG Y, KRISHNAN D, et al. Rethinking few-shot image classification: a good embedding is all you need? [C]//European Conference on Computer Vision. 2020:266-282.
- [8] CHEN W Y, LIU Y C, KIRA Z, et al. A closer look at few-shot classification[C]//International Conference on Learning Representations. 2019:1-16.
- [9] WANG Y, CHAO W L, WEINBERGER K Q, et al. SimpleShot: Revisiting Nearest-Neighbor Classification for Few-Shot Learning[J]. arXiv:1911.04623, 2019.
- [10] YANG S, LIU L, XU M. Free lunch for few-shot learning: Distribution calibration[C]//International Conference on Learning Representations. 2021:1-13.
- [11] LUO X, WEI L H, WEN L J, et al. Rectifying the shortcut learning of background for few-shot learning[C]// Annual Conference on Neural Information Processing Systems. 2021:13073-13085.
- [12] RODRÍGUEZ P, LARADJI I H, DROUIN A, et al. Embedding propagation: Smoother manifold for few-shot classification[C]// European Conference on Computer Vision. 2020:121-138.
- [13] SNELL J, SWERSKY K, ZEMEL R S. Prototypical networks for few-shot learning[C]// Annual Conference on Neural Information Processing Systems. 2017:4077-4087.
- [14] SUNG F, YANG Y X, ZHANG L, et al. Learning to compare: Relation network for few-shot learning[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2018:1199-1208.
- [15] LI W B, WANG L, XU J L, et al. Revisiting local descriptor based image-to-class measure for few-shot learning[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2019: 7260-7268.
- [16] ZHANG C, CAI Y J, LIN G S, et al. DeepEMD: Few-shot image classification with differentiable earth mover's distance and structured classifiers[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2020:12200-12210.
- [17] FINN C, ABBEEL P, LEVINE S. Modelagnostic meta-learning for fast adaptation of deep networks[C]//International Conference on Machine Learning. 2017:1126-1135.
- [18] NICHOL A, ACHIAM J, SCHULMAN J. On first-order meta-learning algorithms[J]. arXiv:1803.02999, 2018.
- [19] TRIANTAFILLOU E, ZHU T, DUMOULIN V, et al. Meta-dataset: A dataset of datasets for learning to learn from few examples[C]// International Conference on Learning Representations. 2020:1-24.
- [20] WU Z R, XIONG Y J, YU S X, et al. Unsupervised Feature Learning via Non-Parametric Instance-level Discrimination[J]. arXiv:1805.0197, 2018.
- [21] YE M, ZHANG X, YUEN P C, et al. Unsupervised Embedding Learning via Invariant and Spreading Instance Feature [C] // IEEE Conference on Computer Vision and Pattern Recognition. 2019:6210-6219.
- [22] CHEN T, KORNBLITH S, NOROUZI M, et al. A Simple Framework for Contrastive Learning of Visual Representations [C] // International Conference on Learning Representations. 2020:1597-1607.
- [23] LIU C, ZHANG L, XU C M, et al. Learning a Few-shot Embedding Model with Contrastive Learning[C]// AAAI Conference on Artificial Intelligence. 2021:8635-8643.
- [24] LUO X, CHEN Y X, WEN L J, et al. Boosting few-shot classification with view-learnable contrastive learning[C]//IEEE Conference on Multimedia and Expo. 2021:1-6.
- [25] MIYATO T, MAEDA S I, KOYAMA M, et al. Virtual adversarial training: a regularization method for supervised and semi-supervised learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(8):1979-1993.
- [26] WEI C, SOHN K, MELLINA C, et al. Crest: A class-rebalancing self-training framework for imbalanced semi-supervised learning [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2021:10857-10866.
- [27] REN M Y, TRIANTAFILLOU E, RAVI S, et al. Meta-learning for semi-supervised few-shot classification [C] // International Conference on Learning Representations. 2018:1-15.
- [28] HU Y Q, GRIPON V, PATEUX S. Graph-based interpolation of feature vectors for accurate few-shot classification[C]//International Conference on Pattern Recognition. 2021:8164-8171.
- [29] PAPPAS V, HAN X Y, DONOHO D L. Prevalence of neural collapse during the terminal phase of deep learning training[J]. arXiv:2008.08186, 2020.
- [30] VINYALS O, BLUNDELL C, LILLICRAP T, et al. Matching networks for one shot learning [C] // Annual Conference on Neural Information Processing Systems. 2016:3630-3638.
- [31] ORESHKIN B, RODRÍGUEZ L'ÓPEZ P, LACOSTE A. Tadam: Task dependent adaptive metric for improved few-shot learning[C]// Annual Conference on Neural Information Pro-

cessing Systems, 2018:721-731.

- [32] RAVI S, LAROCHELLE H. Optimization as a model for few shot learning[C]//International Conference on Learning Representations, 2017:1-11.
- [33] MISHRA N, ROHANINEJAD M, CHEN X, et al. A simple neural attentive meta-learner[C]// International Conference on Learning Representations, 2018:1-17.
- [34] LEE K, MAJI S, RAVICHANDRAN A, et al. Meta-learning with differentiable convex optimization[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2019: 10657-10665.
- [35] HILLER M, MA R K, HARANDI M, et al. Rethinking generalization in few-shot classification[J]. arXiv:2206.07267, 2022.
- [36] YANG L, LI L L, ZHANG Z L, et al. DPGN: Distribution propagation graph network for few-shot learning[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2020:13387-13396.
- [37] SHEN X, XIAO Y, HU S X, et al. Re-ranking for image retrieval and transductive few-shot classification[C]// Annual Conference on Neural Information Processing Systems, 2021:25932-25943.
- [38] LAZAROU M, STATHAKI T, AVRITHIS Y. Iterative label cleaning for transductive and semi-supervised few-shot learning

[C]//IEEE Conference on Computer Vision, 2021:8731-8740.

- [39] CHEN C F, YANG X S, XU C S, et al. ECKPN: Explicit class knowledge propagation network for transductive few-shot learning[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2021:6592-6601.
- [40] BENDOU Y, HU Y Q, LAFARGUE R, et al. EASY: Ensemble augmented-shot y-shaped learning; state-of-the-art few-shot classification with simple ingredients[J]. arXiv: 2201.09699, 2022.



**SU Ruqi**, born in 2000, postgraduate. His main research interests include pattern recognition and deep learning.



**ZHU Songhao**, born in 1973, Ph.D. His main research interests include image processing, pattern recognition, and artificial intelligence.