

DOI:10.19651/j.cnki.emt.2315182

基于原型嵌入图网络的小样本图像分类*

陈望^{1,2} 李志玲^{1,2} 王前^{1,2} 包春梅^{1,2} 王林²

(1. 贵州民族大学数据科学与信息工程学院 贵阳 550025; 2. 贵州省模式识别与智能系统重点实验室贵州民族大学 贵阳 550025)

摘要: 针对在小样本图像分类中传统骨干卷积网络进行特征提取时会有上下文信息单一和感受野受限以及边特征相似度度量缺乏全局性问题,本文提出了一种基于原型嵌入图网络的小样本图像分类算法。首先,将 CBAM 产生的权重值与 ASPP 以不同采样率获取的不同尺度特征进行相乘的特征作为图网络的节点嵌入特征。然后,采用原型网络的方法在度量模块中构建了原型节点,使得成对节点之间的相似性计算转化为单个节点与原型节点相似性的和计算,将得到的相似度作为边特征输入图神经网络。最后,利用双图结构在多个更新代后将标签信息从有标签样本传播到无标签样本。在以 ResNet-12 为骨干卷积网络的算法分类任务中,本文在 *miniImageNet*、*tieredImageNet*、CUB-200-2011 和 CIFAR-FS 4 个数据集上的 5-way 1-shot 的任务分类准确率分别达到了 71.47%、75.41%、86.21% 和 79.84%,在以 Conv-4 作为骨干卷积网络中,本文提出的算法在 5-way 1-shot 和 5-way 5-shot 任务中都优于现有的图网络方法。

关键词: 图神经网络;小样本学习;原型节点;多尺度特征融合;图像分类

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Few-shot image classification based on prototype embedding graph network

Chen Wang^{1,2} Li Zhiling^{1,2} Wang Qian^{1,2} Bao Chunmei^{1,2} Wang Lin²

(1. School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China;

2. Key Laboratory of Pattern Recognition and Intelligent System, Guizhou Minzu University, Guiyang 550025, China)

Abstract: Traditional backbone convolutional networks for feature extraction in few-shot image classification suffer from the problems of single context information, limited receptive field, and lack of global edge feature similarity measurement. In this paper, we propose a few-shot image classification algorithm based on prototype embedding graph network. First, the feature obtained by multiplying the weight values generated by CBAM with the features of different scales obtained by ASPP at different sampling rates is used as the node embedding feature of the graph network. Then, the prototype nodes are constructed in the measurement module using the prototype network method, transforming the similarity calculation between paired nodes into the sum of the similarity between a single node and the prototype node, and using the obtained similarity as the edge feature input to the graph neural network. Finally, the label information is propagated from labeled samples to unlabeled samples through the dual graph structure after multiple update iterations. In the classification task using ResNet-12 as the backbone convolutional network, our algorithm achieves classification accuracies of 71.47%, 75.41%, 86.21%, and 79.84% on the *miniImageNet*, *tieredImageNet*, CUB-200-2011, and CIFAR-FS datasets, respectively, for the 5way-1shot task. In the backbone convolutional network using Conv-4, our proposed algorithm outperforms existing graph network methods in both 5way-1shot and 5way-5shot tasks.

Keywords: graph neural network; few-shot learning; prototype nodes; multi-scale feature fusion; image classification

0 引言

小样本学习(few-shot learning, FSL)旨在通过少量有

标签样本(即支持样本(support set, S))对新的类学习出有效的模型,从而对无标签样本(即查询样本(query set, Q))进行分类。为了不利用正则性来弥补标签数据的稀缺性,

收稿日期:2023-12-17

* 基金项目:国家自然科学基金(62241206)、贵州省科技计划项目(黔科合基础-ZK[2022]一般 195,黔科合基础-ZK[2023]一般 143,黔科合基础-ZK[2022]一般 550,黔科合平台人才-ZCKJ[2021]007)、贵州省青年科技人才成长项目((黔教合 KY 字[2021]104)、贵州省教育厅自然科学研究项目(黔教技[2023]012号,黔教技[2022]015号,黔教技[2023]061号)、贵州省教育厅青年科技人才成长项目(黔教合 KY 字[2021]115,黔教合 KY 字[2021]110)、贵州省深化新时代教育评价改革试点“教学过程质量评价”、贵州省模式识别与智能系统重点实验室开放课题(GZMUKL[2022]KF01)项目资助

元学习^[1]应运而生。元学习方法引入了 episodic 的概念,利用少量标签数据进行一轮模型训练。通过 episode,元学习旨在训练一个元学习器进行支持样本到查询样本的标签传播。为有效解决传统元学习方法泛化能力差的问题^[2]。近年来,图神经网络(graph neural network, GNN)^[3-4]受到小样本学习研究者的广泛关注。使用图网络^[4]处理训练数据,利用了图的拓扑结构获得数据先验从而进行标签传播。Wei 等^[5]将模糊理论应用于 GNN 当中,使得其在每个 episodic 获得更优越的关系归纳偏置,并使用以边为重点的 GNN 通过迭代更新边标签来执行边预测。根据边预测的输出设计模糊成员函数,以实现节点分类。随后, Ma 等^[6]利用已知的支持集样本之间的关系来指导图的关系传播,为支持集样本和查询集样本生成判别关系嵌入。同时引入伪关系节点来传播查询集样本特性,并设计了一种快速有效的关系学习策略,以充分利用不同查询集样本之间的关系信息。Kim 等^[7]直接利用簇内相似度和簇间不相似度迭代更新边缘标签,学习预测边缘标签而不是图上的节点标签,从而实现显式聚类的演化。此外, Tian 等^[8]提出了使用一个好的学习嵌入模型可以比复杂的元学习算法更有效。文献^[9]利用相似度加权基类来表征新类样本的原型,以新类样本特征与其对应的原型构造均匀分布进而扩充支持集数量增强小样本图像分类性能。Zhou 等^[10]利用元学习的方法提取特征并将特征嵌入到图神经网络进行端到端的推导学习。Yang 等^[11]第一次引入分布图来引导节点图的学习,通过分布图和点图循环信息传递将实例级别和分布级别的关系融合进行推导。Tang 等^[12]构建了一个以标签和支持数据的特征为条件的条件随机场,以推断标签空间中的亲和力。并将这种亲和性作为节点亲和性被馈送到 GNN 上作为样本分类推断标准。上述方法考虑到了成对节点间的相似性,但对于节点相似性的计算过于依赖节点特征嵌入的好坏,同时也忽略了图网络全局特征与每个节点特征之间的关联性构造。因此,对于图网络结构的设计对于获取良好的节点嵌入以及准确度量节点相似性来说至关重要。

在小样本图像分类中,图神经网络会出现同类节点亲和力弱和异类节点亲和力强,影响节点之间的相似性度量问题。因此,本文从节点嵌入和节点相似性度量出发,在分布传播图网络(distribution propagation graph network for few-shot learning, DPGN)^[11]的基础上提出了原型嵌入图网络(prototype embedding graph network, PEGN)。利用卷积注意力机制模块(convolutional block attention module, CBAM)^[13]改进空洞空间卷积池化金字塔(atrous spatial pyramid pooling, ASPP)^[14]特征融合网络获取样本的多尺度特征作为节点嵌入,优化节点的嵌入更好的计算同类节点的相似性和异类节点的相异性。同时,基于原型网络的思想构建基类样本的原型表示,将成对样本的相似性计算转化为两个样本与邻域内样本的原型表示相似性的计算。通过上述

图网络的设计,获得了更具代表性的节点特征,以新的度量模块计算出更为准确的节点相似性,从而提升了图网络的分类性能。本文的主要贡献在于以下 3 个方面:

1) 设计了一种原型度量模块,通过构建基类样本的原型表示(即原型节点),将成对节点的相似性计算转化为两个节点分别与原型节点的相似性和计算。

2) 设计了一种多尺度特征融合模块,利用 CBAM 产生的权重值与 ASPP 以不同采样率获取的不同尺度特征进行相乘,并将该特征作为节点嵌入特征。

3) 在 4 个流行的小样本学习基准数据集上进行了广泛的实验。通过与现有的方法进行比较,结果表明 PEGN 有效提升了小样本图像分类精度。并且本文的算法优于现有的基于图的小样本学习方法。

1 方 法

1.1 问题定义

小样本学习任务的目标是在给定少量标签数据进行模型训练,从而训练出可以对未知标签的样本进行分类。

小样本学习通常采用元任务的方式进行训练和评估,构建元任务的最常见范式为 N-way K-shot 任务,即从新类中采样 N 个类别,每个类别抽取 K 个带标签图像作为训练样本,再从每类剩下的图像中抽取 T 个样本作为预测图像^[9]。每个小样本任务都有一个支持集 \mathcal{S} 和一个查询集 \mathcal{Q} 。给定训练数据 $\mathcal{D}^{\text{train}}$,支持集 $\mathcal{S} \subset \mathcal{D}^{\text{train}}$ 包含 N 个类,每个类有 K 个样本,记为 $\mathcal{S} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_{N \times K}, y_{N \times K})\}$ 。查询集 $\mathcal{Q} \subset \mathcal{D}^{\text{train}}$ 有 T 个样本,记为 $\mathcal{Q} = \{(\mathbf{x}_{N \times K + 1}, y_{N \times K + 1}), \dots, (\mathbf{x}_{N \times K + T}, y_{N \times K + T})\}$ 。具体来说,在训练阶段,为支持集 \mathcal{S} 和查询集 \mathcal{Q} 都提供了数据标签。给定测试数据 $\mathcal{D}^{\text{test}}$,本文的目标是训练一个分类器,该分类器可以使用来自 $\mathcal{S} \in \mathcal{D}^{\text{test}}$ 的少量支持样本将查询样本 $\mathcal{Q} \in \mathcal{D}^{\text{test}}$ 准确地映射到相应的标签,并且支持集和查询集的标签是互斥的。

1.2 原型嵌入图网络

PEGN 同样采用 DPGN 的双图循环训练,不同的是本文在进行点图与分布图节点嵌入时采用多尺度特征融合模块对上一代节点特征进行特征融合。同时节点相似性计算方式上提出了原型节点的构造,将成对节点的相似性计算以一种全新方式进行替代度量。原型图网络由 l 个子代组成,每个子代由点图 $G_l^p = (\mathbf{V}_l^p, \mathbf{E}_l^p)$ 和分布图 $G_l^d = (\mathbf{V}_l^d, \mathbf{E}_l^d)$ 组成。首先,通过一个骨干卷积网络提取所有样本的嵌入特征,这些嵌入特征用于计算实例相似度 \mathbf{E}_l^p 。其次,传递实例关系 \mathbf{E}_l^p ,构建分布图 G_l^d 。节点特征 \mathbf{V}_l^d 由 \mathbf{E}_l^p 按照在 G_l^p 中的位置顺序聚合初始化,边特征 \mathbf{E}_l^d 表示节点特征 \mathbf{V}_l^d 之间的分布相似性。最后,将得到的 \mathbf{E}_l^d 传递给 G_l^p ,以构造节点 \mathbf{V}_l^p 的更具判别性的表示,并逐代重复上述传播过程。

其中, $\mathbf{V}_l^p, \mathbf{E}_l^p, \mathbf{V}_l^d$ 和 \mathbf{E}_l^d 可表示为 $\mathbf{V}_l^p = \{\mathbf{v}_{l,i}^p\}, \mathbf{E}_l^p = \{\mathbf{e}_{l,i,j}^p\}, \mathbf{V}_l^d = \{\mathbf{v}_{l,i}^d\}, \mathbf{E}_l^d = \{\mathbf{e}_{l,i,j}^d\}$ 其中 $i, j = 1, \dots, T$. $T = N \times K + \bar{T}$ 表示训练集中的总样本数。 $\mathbf{v}_{0,i}^p$ 由骨干卷积网络 f_{emb} 进行提取。

$$\mathbf{v}_{0,i}^p = f_{emb}(\mathbf{x}_i) \quad (1)$$

其中, $\mathbf{v}_{0,i}^p \in \mathbf{R}^m, m$ 为特征嵌入维数。

1.3 模型结构

本文提出基于原型嵌入图网络的小样本图像分类算法,模型框架如图1所示。图中圆圈代表样本经过骨干卷积网络后所获得的特征,颜色代表类别。本文提出了原型节点的设计方法,根据基类样本的特征采用原型网络的方法构建了一个虚拟节点作为原型节点。同时,在进行点图与分布图的节点嵌入时设计了多尺度特征融合模块通过将ASPP不同采样率获取的节点嵌入特征与CBAM对节点嵌入特征在空间和通道上产生的权重值相乘获得融合特征,并将该特征作为图网络的节点嵌入。根据节点特征设

计了原型节点的同时也设计一个原型度量模块,将对成对节点之间的亲和力计算转变为了两个节点与原型节点之间的亲和力和计算。多尺度特征融合模块优化了点图与分布图的节点特征,原型度量模块对成对节点亲和力的计算进行了转化度量,考虑到了节点间的亲和力牵引,并将节点的亲和力计算作为图神经网络的边特征。经过点图与分布图进行相互传播,获得了更新后的节点特征与边特征,最后输出查询样本的类别概率。本文以2-way 1-shot的任务为例,从特征提取器中获得的支持向量和查询向量并利用支持向量构建原型节点,同时传递到对偶完全图(点图(point graph, PG)和分布图(distribution graph, DG))中,进行一代又一代的直推式传播。点图到分布图的箭头是边到节点的变换($PG \rightarrow DG, PD$),该变换聚合了实例相似度来构建分布表示,分布图到点图的箭头是另一个边到节点的变换($DG \rightarrow PG, DP$),该变换聚合了分布相似度和实例特征,在第 l 代末对查询样本进行预测。

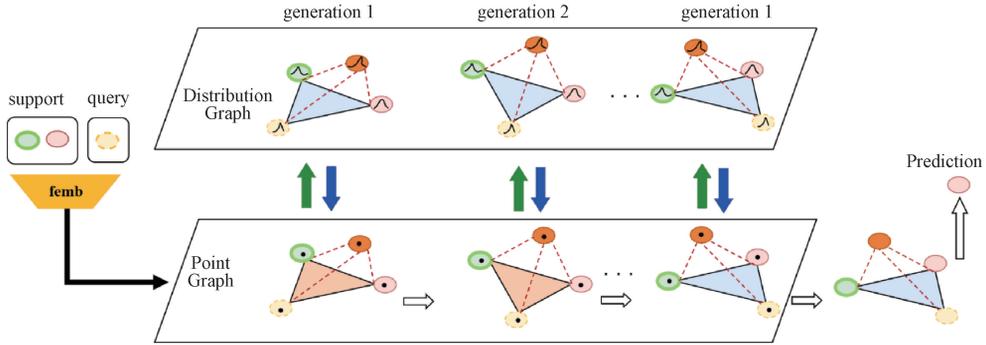


图1 原型嵌入图网络框架图

1.4 多尺度特征融合模块

本文将多尺度特征融合模块添加进骨干卷积网络 f_{emb} , 将其提取的特征作为图神经网络的节点嵌入特征。多尺度特征融合模块网络结构如图2所示,包括CBAM和ASPP两个模块。考虑到传统骨干卷积网络进行特征提取时会有上下文信息单一以及感受野受限问题,本文采用ASPP进行特征融合,对于特征图采用不同采样率的空洞卷积并行采样,并将得到的特征拼接在一起,扩大通道数,最后通过 1×1 的卷积获取期望的通道数。ASPP利用空洞卷积扩大感受野,同时以不同采样率捕捉到了更为丰富的上下文信息,使得提取特征时获得了多尺度的融合特征。CBAM包括通道注意力机制(channel attention mechanism, CAM)和空间注意力机制(spatial attention mechanism, SAM)两个子模块,可以分别在通道和空间上添加注意力机制,以此来关注重要特征并抑制不必要的特征^[15],利用CBAM产生相应特征图的权重值并与ASPP产生的不同尺度特征相乘产生新节点特征。添加多尺度特征融合模块进骨干卷积网络中不仅使网络获得了CBAM节约参数和计算力的能力,同时也获取到了更好的图像全局特征,得到更为丰富的节点特征,从而提升了图

神经网络的表征能力。

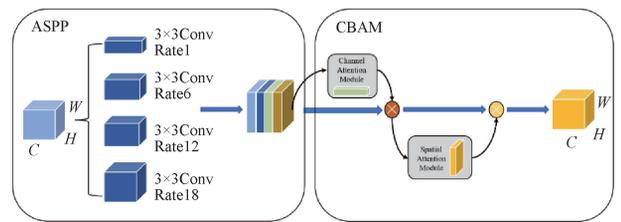


图2 多尺度特征融合模块的网络结构

1.5 原型度量模块

本文提出的原型度量模块可以更加有效的度量节点特征之间的相似性。如图3所示,给出了2-way 1-shot的分类任务为例,MLP-1为点图到分布图聚合中提到的FC-ReLU块,MLP-2为分布图到点图聚合中的Conv-BN-ReLU块。在点图和分布图中分别构建样本的原型表示,利用 f_{emb} 提取图像特征嵌入图神经网络节点中作为节点特征,对于原型节点的构造采用基类支持样本集成该类样本的原型表示并将这个原型向量嵌入图网络节点作为原型节点。将对成对节点之间的相似性计算转化为单个节点与原型节点相似性计算的和。并将该相似性作为图神经

网络的边特征。经过 l 次点图与分布图的循环更新,以最终代的点图进行样本的分类预测。

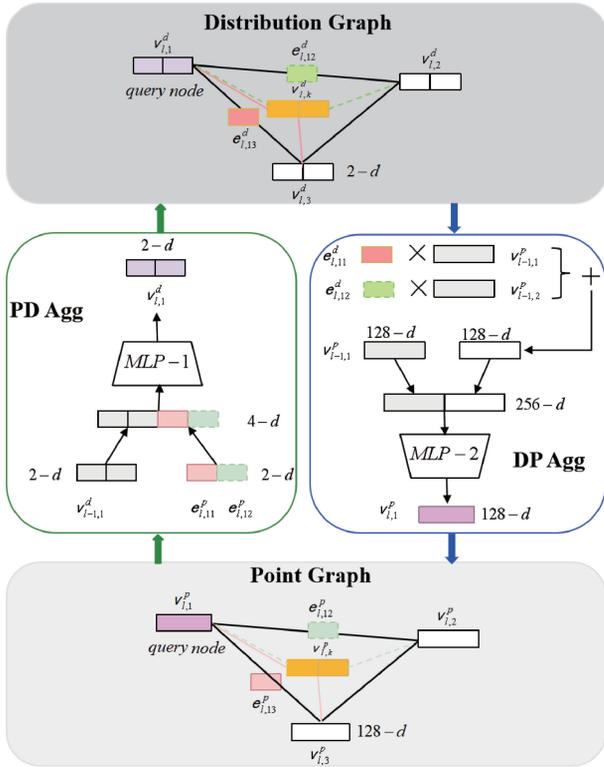


图 3 PD 网络和 DP 网络的网络结构

原型度量模块使用欧式距离计算节点之间的相似性。用两个节点的 L_2 范数来计算成对节点之间的相似性, L_2 范数越小, 节点特征的相似性越高, 两个节点样本同一类别的概率越大。原型节点特征初始化公式如下:

$$v_{0,k}^p = \frac{1}{|S_k|} \sum_{\langle x_i, y_i \rangle \in S_k} f_\varphi(x_i) \quad (2)$$

其中, $|S_k|$ 为标签为 k 的集合, f_φ 为嵌入函数。

通过构建原型节点计算点图中的每条边特征, 即点图节点相似性。第一代的边 $e_{0,ij}^p$ 初始化如下:

$$e_{0,ij}^p = f_{e_0^p}((v_{0,i}^p - v_{0,k}^p)^2 + (v_{0,j}^p - v_{0,k}^p)^2) \quad (3)$$

其中, $e_{0,ij}^p \in \mathbf{R}$, $f_{e_0^p}: \mathbf{R}^m \rightarrow \mathbf{R}$ 是一个将实例相似度转化为一定的规模。 $f_{e_0^p}$ 包含了两个参数集为 $\theta_{e_0^p}$ 的 Conv-BN-ReLU 块以及一个 Sigmoid 层。

当子代 $l > 0$, 给定 $e_{l-1,ij}^p, v_{l-1,i}^p, v_{l-1,j}^p$ 和 $v_{l-1,k}^p$, 即可计算出更新以后的 l 代边特征 $e_{l,ij}^p$ 如下:

$$e_{l,ij}^p = f_{e_l^p}((v_{l-1,i}^p - v_{l-1,k}^p)^2 + (v_{l-1,j}^p - v_{l-1,k}^p)^2) \cdot e_{l-1,ij}^p \quad (4)$$

为了使用 G_l^p 的全局信息操作边信息的更新, 对 $e_{l,ij}^p$ 进行归一化操作。

基于原型节点的基础在点图 G_l^p 中产生或更新边特征 $e_{l,ij}^p$ 后, 下一步要构造的是分布图 $G_l^d = (V_l^d, E_l^d)$ 。利用

分布图 G_l^d 可以集成点图 G_l^p 中的实例关系嵌入到分布图中, 从而处理分布层次关系。 G_l^d 中的每个节点的特征为分布特征 $v_{l,i}^d, v_{l,i}^d$ 是个 NK 维的特征向量。其中第 j 项代表样本 x_i 与样本 x_j 的分布关系, NK 代表任务中支持集样本的总数, 对于分布图中的节点特征第一次初始化为:

$$v_{0,i}^d = \begin{cases} \parallel_{j=1}^{NK} \delta(y_i, y_j), & \text{若 } x \text{ 存在标签} \\ \left[\frac{1}{NK}, \dots, \frac{1}{NK} \right], & \text{其他} \end{cases} \quad (5)$$

其中, $v_{0,i}^d \in \mathbf{R}^{NK}$, \parallel 是一个级联算子, $\delta(\cdot)$ 为克罗内克函数, 当 $y_i = y_j$ 时输出 1, 反之输出为 0。

当子代 $l > 0$ 时, 分布节点 $v_{l,i}^d$ 的更新方式如下:

$$v_{l,i}^d = PD(\parallel_{j=1}^{NK} e_{l,ij}^p, v_{l-1,i}^d) \quad (6)$$

其中, $PD: (\mathbf{R}^{NK}, \mathbf{R}^{NK}) \rightarrow \mathbf{R}^{NK}$ 为分布图的聚合网络, 参数集为 $\theta_{v_l^d}$ 。

在点图构建原型节点以后, 对于分布图本文同样在分布图中以相同的方式构建原型节点, 其原型分布为 $v_{0,k}^d$ 。分布图中的每条代表了不同样本分布特征之间的相似性, 故初始的分布相似度 $e_{0,ij}^d$ 构造如下:

$$e_{0,ij}^d = f_{e_0^d}((v_{0,i}^d - v_{0,k}^d)^2 + (v_{0,j}^d - v_{0,k}^d)^2) \quad (7)$$

其中, $e_{0,ij}^d \in \mathbf{R}$, $f_{e_0^d}: \mathbf{R}^{NK} \rightarrow \mathbf{R}$ 使用了两个 Conv-BN-ReLU 块对分布相似性进行转换。 $f_{e_0^d}$ 包含了一个参数集为 $\theta_{e_0^d}$ 以及一个 Sigmoid 层。

对于子代 $l > 0$, 分布图 G_l^d 的更新规则为:

$$e_{l,ij}^d = f_{e_l^d}((v_{l-1,i}^d - v_{l-1,k}^d)^2 + (v_{l-1,j}^d - v_{l-1,k}^d)^2) \cdot e_{l-1,ij}^d \quad (8)$$

为了使用 G_l^d 的全局信息操作边信息的更新, 对 $e_{l,ij}^d$ 进行归一化操作。

在 G_l^d 获取的编码分布信息在每一代结束时都会返回点图 G_l^p 中, 同时 G_l^p 的节点特征 $v_{l,i}^p$ 利用 G_l^p 中所有节点特征与边特征 $e_{l,i}^p$ 聚合获取分布关系如下:

$$v_{l,i}^p = DP\left(\sum_{j=1}^T (e_{l,ij}^d \cdot v_{l-1,j}^p), v_{l-1,i}^p\right) \quad (9)$$

其中, $v_{l,i}^p \in \mathbf{R}^m$, 而 $DP: (\mathbf{R}^m, \mathbf{R}^m) \rightarrow \mathbf{R}^m$ 是点图 G_l^p 中的一个聚合网络, 其参数集为 $\theta_{v_l^p}$ 。 DP 通过级联当代计算所得特征与上一代节点特征, 并用两个 Conv-BN-ReLU 块对级联后的特征进行更新。经过这个过程, 节点特征可以将分布层次的信息整合到实例层次的特征中, 为下一代计算实例相似度做准备。

1.6 损失函数

通过 PEGN 的最后一代点图的节点之间的边特征相似性进行标签传播, 从而预测出每个节点的类别。计算方式如下:

$$P(\hat{y}_i | x_i) = \text{Softmax}\left(\sum_{j=1}^{NK} e_{l,ij}^p \cdot \text{one-hot}(y_j)\right) \quad (10)$$

其中, $P(\hat{y}_i | \mathbf{x}_i)$ 是给定样本 \mathbf{x}_i 在类别上的概率分布, y_j 是支持集中第 j 个样本的标签。 $e_{i,j}^l$ 表示最后一代点图中的边特征。

基于原型节点的构造本文也计算出了点损失以及分布损失。值得注意的是, 本文在点图中对每个样本进行分类预测。因此, l 代的点损失定义如下:

$$L_i^p = L_{CE}(P(\hat{y}_i | \mathbf{x}_i), y_i) \quad (11)$$

其中, L_{CE} 为交叉熵损失函数, T 表示每个任务中 $(S, Q) \in D^{\text{train}}$ 的样本数量, $P(\hat{y}_i | \mathbf{x}_i)$ 和 y_i 分别代表模型的概率预测值与真实值。

以点图的原型节点构造并将点图中的节点特征与边特征进行级联输入到 PD 聚合网络中即可获得分布图的节点特征, 进而计算出边特征。本文也计算了分布损失以便于训练过程和学习具有判别性的分布特征, 分布损失在更快和更好的收敛方面发挥了重要作用。定义第 l 代的分布损失如下:

$$L_i^d = L_{CE}\left(\text{Softmax}\left(\sum_{j=1}^{NK} e_{i,j}^d \cdot \text{one-hot}(y_j)\right), y_i\right) \quad (12)$$

其中, $e_{i,j}^d$ 表示第 l 代分布图中的边特征。

总目标函数是上述所有损失的加权求和, 计算公式定义如下:

$$L = \sum_{l=1}^{\hat{l}} (\lambda_p L_i^p + \lambda_d L_i^d) \quad (13)$$

其中, \hat{l} 表示 DPGN 的子代的总数, λ_p 和 λ_d 表示每个损失的权重, 以平衡它们的重要性。在本文的大部分实验中, λ_p 和 λ_d 分别设置为 1.0 和 0.1。

2 实验

为了验证多尺度特征融合模块在传统骨干卷积网络特征提取时出现上下文信息单一和感受野受限问题的性能, 以及原型度量模块对于边特征相似性计算的有效性, 本文在 4 个标准的小样本数据集上进行验证。

2.1 数据集

本文在 4 个标准的小样本学习数据集上对 PEGN 进行了评估, 数据集分别为 *miniImageNet*^[16]、*tieredImageNet*^[17]、*CUB-200-2011*^[18] 和 *CIFAR-FS*^[19]。*miniImageNet* 和 *tieredImageNet* 是 *ImageNet*^[20] 的子集。*CUB-200-2011* 最初是为细粒度分类而设计的, *CIFAR-FS* 是 *CIFAR-100* 的子集, 用于小样本分类。如表 1 所示, 本文按照前人工作的标准列出了图像数量、类别数、图像分辨率和 train/val/test 分割的详细信息。

2.2 实验设置

本文的方法是在 Ubuntu16.04 操作系统下基于 PyTorch 代码实现的, 显卡是 NVIDIA GeForce RTX 3090, CUDA 版本 11.2, 内存 40 GB, CPU 是 Intel(R) Xeon(R)

表 1 实验数据集

| 数据集 | 图像数量 | 类别数 | 训练/验证/测试 | 分辨率 |
|-----------------------|---------|-----|------------|-------|
| <i>miniImageNet</i> | 60 000 | 100 | 64/16/20 | 84×84 |
| <i>tieredImageNet</i> | 779 165 | 608 | 351/97/160 | 84×84 |
| <i>CUB-200-2011</i> | 11 788 | 200 | 100/50/50 | 84×84 |
| <i>CIFAR-FS</i> | 60 000 | 100 | 64/16/20 | 32×32 |

Silver 4210 CPU@2.20 GHz。在训练前进行常见数据增强, 如水平翻转、随机裁剪、颜色抖动(亮度、对比度和饱和度)^[11]。骨干卷积网络为 ResNet-12 和 Conv-4, 网络最终输出维度为 128, 在每次迭代中随机抽取 28 个 episodes 进行元训练。所有实验均使用 Adam 优化器, 批量大小分别为 40、25 和 28, 初始学习率为 10^{-3} 。同时将学习率每隔 15 000 次迭代衰减为原来的 0.1, 并将权重衰减设置为 10^{-5} 。在标准的小样本学习数据集 *miniImageNet*、*tieredImageNet*、*CUB-200-2011* 和 *CIFAR-FS* 上评估了 5-way 1-shot/5-shot 设置下的 PEGN, 遵循文献[11]的评估过程。在测试阶段, 随机抽取 10 000 个任务, 模型评价指标为 95% 的置信度的平均准确率。

2.3 实验结果与分析

本文将 PEGN 模型与现有的模型进行比较, 进行比较的方法包括图方法和非图方法。为了公平比较, 在 *miniImageNet*、*tieredImageNet*、*CIFAR-FS* 和 *CUB-200-2011* 数据集上使用 PEGN, 并在相同的骨干卷积网络上与其他方法进行比较。如表 2~5 所示, 所提出的 PEGN 优于其他现有的方法, 并达到了最先进的性能, 特别是与基于图的方法相比。

本文分别对 5-way 1-shot 和 5-way 5-shot 分类任务进行了评估。模型在 *miniImageNet* 数据集在不同分类任务下的精度曲线和损失曲线如图 4 和 5 所示。在图 4(a) 和图 5(a) 中位于上方的曲线是本文模型训练的精度曲线, 而在图 4(b) 和图 5(b) 中位于下方的曲线则是本文模型训练的损失曲线。如图 4 所示, 在 5-way 1-shot 分类任务中 PEGN 相较于 DPGN 收敛速度更快且准确度显著提高。实验表明 PEGN 在 18 000 轮收敛, 收敛的速度更快, 而 DPGN 需要 22 000 轮才能收敛并降低学习率。实验尝试在 18 000 轮下降低 DPGN 的学习率, 但实验结果表明 DPGN 的准确率却降低了约 2%。同时从图 5 中可以看出 PEGN 在 5-way 5-shot 分类任务收敛速度也是明显快于 DPGN 的, 且准确度相较于 DPGN 也有着明显的提高。在 12 000 轮左右 PEGN 就开始收敛, 而 DPGN 在 18 000 轮才开始收敛。PEGN 模型在收敛速度和精度方面都超越了 DPGN 模型, 证明其在实际运用中具有推广前景。

表 2 是 PEGN 在 *miniImageNet* 数据集上的实验结果。

从表 2 可知, 在 5way-1shot 的任务中, 本文的网络分类精度分别比使用 ResNet-12 的 DPGN、FRN、CL、DeepBDC、

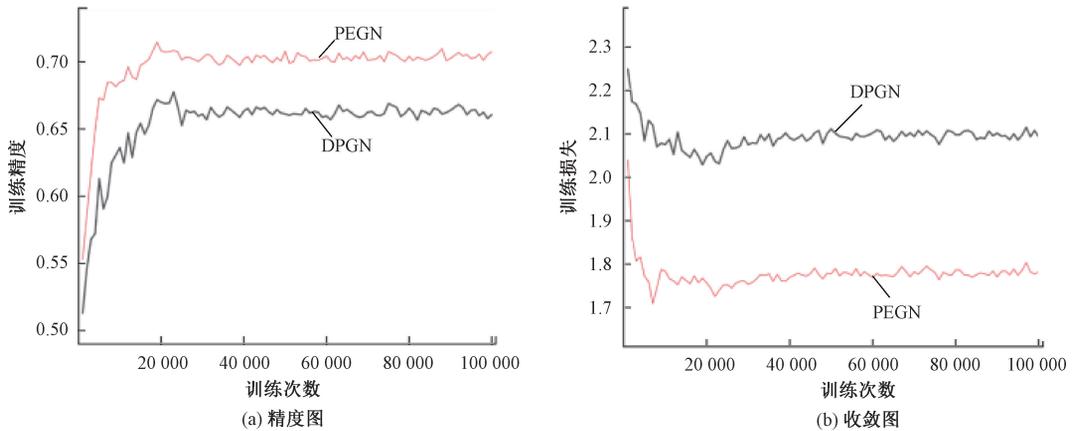


图 4 mini ImageNet 数据集模型在 5-way 1-shot 分类任务中精度图和收敛图

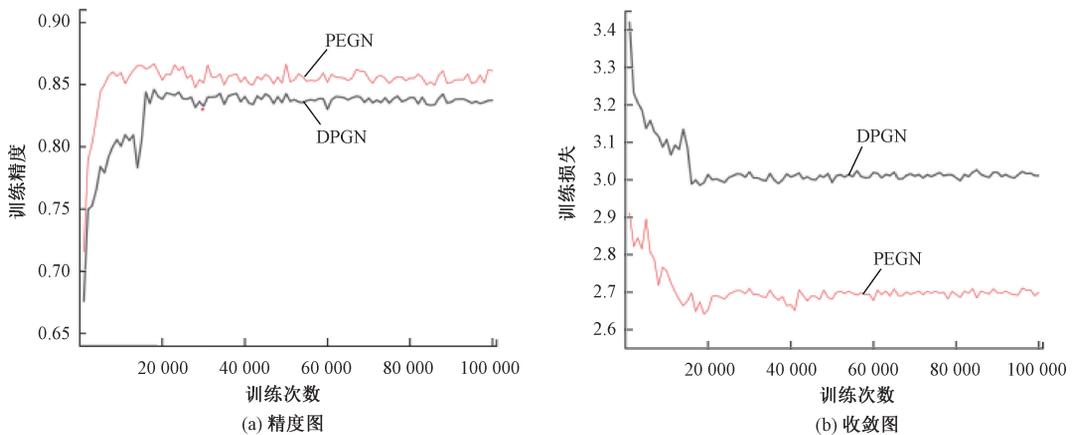


图 5 mini ImageNet 数据集模型在 5-way 5-shot 分类任务中收敛图和损失图

表 2 mini ImageNet 数据集分类准确率

| 方法 | 主干网络 | 5-way 1-shot | 5-way 5-shot |
|-------------------------|------------------|-------------------|-------------------|
| DPGN ^[11] | ResNet-12 | 67.77±0.32 | 84.60±0.43 |
| FRN ^[21] | ResNet-12 | 66.45±0.19 | 82.83±0.13 |
| CL ^[22] | ResNet-12 | 70.19±0.46 | 84.66±0.29 |
| DeepBDC ^[23] | ResNet-12 | 67.83±0.43 | 85.45±0.29 |
| RENet ^[24] | ResNet-12 | 67.60±0.44 | 82.58±0.30 |
| 文献[25] | ResNet-12 | 66.02±0.19 | 84.70±0.13 |
| 文献[26] | Transformer | 63.51±0.20 | 80.03±0.14 |
| PEGN | ResNet-12 | 71.47±0.38 | 86.67±0.33 |
| EGNN ^[7] | Conv-4 | 59.63±0.52 | 76.34±0.48 |
| DPGN ^[11] | Conv-4 | 66.01±0.36 | 82.83±0.41 |
| MCGN ^[12] | Conv-4 | 67.32±0.43 | 83.03±0.54 |
| HGNN ^[27] | Conv-4 | 55.63±0.20 | 72.48±0.16 |
| PEGN | Conv-4 | 70.53±0.28 | 85.34±0.33 |

RENet、文献[25]和[26]模型的精度提升 3.7%、5.02%、1.28%、3.64%、3.87%、5.45% 和 7.96%；在 5way-5shot 的任务中，本文的算法精度分别比使用 ResNet-12 的 DPGN、FRN、CL、DeepBDC、RENet、文献[25]和[26]模型

的精度提升 2.07%、3.84%、2.01%、1.22%、4.09%、1.97%和 6.64%。本文所值得注意的是，使用 Conv-4 作为骨干卷积网络的算法在这两个任务中都优于现有的图网络方法。

表 3 是 PEGN 在 *tiered* ImageNet 数据集上的实验结果。从表中可知，在 5way-1shot 的任务中，本文的网络分类精度分别比使用 ResNet-12 的 DPGN、FRN、CL、DeepBDC、RENet、LPN 和文献[26]模型的精度提升 2.96%、3.35%、2.79%、1.59%、3.8% 和 5.12%，比文献[25]的模型精度稍低 1.89%。在 5way-5shot 的任务中，本文的算法精度分别比使用 ResNet-12 的 DPGN、FRN、CL、DeepBDC、RENet、文献[26]模型的精度提升 2.6%、2.95%、3.22%、0.84%、4.56% 和 5.12%，比文献[25]模型精度稍低 1.04%。文献[25]的模型是基于度量的方法所提出的，与基于优化的图网络方法不同，它更关注的是特征的提取，而本文关注的则是样本特征的相似性度量。需要特别指出的是，使用 Conv-4 作为骨干卷积网络的算法在这两个任务中都优于现有的图网络方法。

表 4 是 PEGN 在 CUB-200-2011 数据集上的实验结果。从表中可知，在 5way-1shot 的任务中，本文的网络分

表3 *tiered ImageNet* 数据集分类准确率

| 方法 | 主干网络 | 5-way 1-shot | 5-way 5-shot |
|-------------------------|------------------|-------------------|-------------------|
| DPGN ^[11] | ResNet-12 | 72.45±0.51 | 87.24±0.39 |
| FRN ^[21] | ResNet-12 | 72.06±0.22 | 86.89±0.14 |
| CL ^[22] | ResNet-12 | 72.62±0.51 | 86.62±0.33 |
| DeepBDC ^[23] | ResNet-12 | 73.82±0.47 | 89.00±0.30 |
| RENet ^[24] | ResNet-12 | 71.61±0.51 | 85.28±0.35 |
| 文献[25] | ResNet-12 | 77.3±0.19 | 90.88±0.10 |
| 文献[26] | Transformer | 70.29±0.15 | 84.12±0.27 |
| PEGN | ResNet-12 | 75.41±0.38 | 89.84±0.12 |
| EGNN ^[7] | Conv-4 | 63.52±0.52 | 80.24±0.49 |
| DPGN ^[11] | Conv-4 | 69.43±0.49 | 85.92±0.42 |
| MCGN ^[12] | Conv-4 | 71.21±0.85 | 85.98±0.98 |
| HGNN ^[27] | Conv-4 | 67.02±0.20 | 83.00±0.13 |
| PEGN | Conv-4 | 73.36±0.44 | 86.63±0.56 |

类精度分别比使用 ResNet-12 的 DPGN、HGNN、RENet 和 ELMOS 的精度提升 10.5%、7.63%、6.72% 和 1%，比 CPN 稍低 1.08%。在 5way-5shot 的任务中，本文的算法精度分别比使用 ResNet-12 的 DPGN、HGNN、RENet 和 CPN 的精度提升 2.6%、4.06%、2.97% 和 1.54%，比 ELMOS 稍低 0.94%。需要说明的是，CPN 和 ELMOS 都是基于数据增强的方法实现小样本分类效果的提升，而非基于优化的方法，本文注重的提升图神经网络的小样本图像分类性能，通过构造推导能力更强的图网络完成样本标签传播，同样与现有的图网络小样本学习方法对照，若使用 Conv-4 作为骨干卷积网络的算法在这两个任务中都优于现有的图网络方法。

表5是 PEGN 在 CIFAR-FS 数据集上的实验结果。从表中可知，在 5way-1shot 的任务中，本文的网络分类精度分别比使用 ResNet-12 的 DPGN、CL、RENet、文献[25]模型和 ELMOS 的精度提升 1.94%、2.28%、5.33%、7.08% 和 1.66%；在 5way-5shot 的任务中，本文的算法精度分别比使用 ResNet-12 的 DPGN、CL、RENet、文献[25]模型和 ELMOS 的精度提升 2.05%、3.61%、5.65%、5.54% 和 2.38%。同样的是，如果使用 Conv-4 作为骨干

表4 *CUB-200-2011* 数据集分类准确率

| 方法 | 主干网络 | 5-way 1-shot | 5-way 5-shot |
|-----------------------|------------------|-------------------|-------------------|
| DPGN ^[11] | ResNet-12 | 75.71±0.47 | 91.48±0.33 |
| HGNN ^[28] | ResNet-12 | 78.58±0.20 | 90.02±0.12 |
| RENet ^[24] | ResNet-12 | 79.49±0.44 | 91.11±0.24 |
| CPN ^[28] | ResNet-12 | 87.29±0.20 | 92.54±0.14 |
| ELMOS ^[29] | ResNet-12 | 85.21±0.38 | 95.02±0.16 |
| PEGN | ResNet-12 | 86.21±0.45 | 94.08±0.47 |
| EGNN ^[7] | Conv-4 | 53.15±0.84 | 81.90±0.60 |
| DPGN ^[11] | Conv-4 | 76.05±0.51 | 89.08±0.38 |
| HGNN ^[27] | Conv-4 | 69.02±0.22 | 83.20±0.15 |
| PEGN | Conv-4 | 78.52±0.50 | 91.58±0.11 |

表5 *CIFAR-FS* 数据集分类准确率

| 方法 | 主干网络 | 5-way 1-shot | 5-way 5-shot |
|-----------------------------|------------------|-------------------|-------------------|
| DPGN ^[11] | ResNet-12 | 77.90±0.50 | 90.20±0.40 |
| CL ^[22] | ResNet-12 | 77.56±0.47 | 88.64±0.31 |
| RENet ^[24] | ResNet-12 | 74.51±0.46 | 86.60±0.32 |
| 文献[25] | ResNet-12 | 72.76±0.21 | 86.71±0.15 |
| ELMOS ^[29] | ResNet-12 | 78.18±0.41 | 89.87±0.31 |
| PEGN | ResNet-12 | 79.84±0.31 | 92.25±0.17 |
| RelationNet ^[30] | Conv-4 | 55.00±1.00 | 69.30±0.80 |
| R2D2 ^[19] | Conv-4 | 65.30±0.20 | 79.40±0.10 |
| DPGN ^[11] | Conv-4 | 76.40±0.50 | 88.40±0.40 |
| PEGN | Conv-4 | 78.58±0.12 | 89.74±0.16 |

卷积网络的算法在这两个任务中都优于现有的对照方法。

2.4 消融实验

本文针对 *miniImageNet* 数据集进行了四组消融实验，以验证 PEGN 的有效性。第一组实验，基于原始的分布传播图网络的实验；第二组实验，基于多尺度特征融合模块对于节点嵌入进行特征优化的双图网络实验；第三组实验，基于原型表示构造虚拟节点设计的新度量模块的双图网络实验；第四组实验，基于多尺度特征融合模块和原型表示构造虚拟节点的度量模块的双图网络实验。实验结果如表6所示。

表6 消融实验

| 序号 | 特征提取 | | 度量方法 | | 5-way 1-shot | 5-way 5-shot |
|----|-----------|---------|-------|------|-------------------|-------------------|
| | ResNet-12 | 多尺度特征融合 | L2 距离 | 原型节点 | | |
| 1 | ✓ | | ✓ | | 67.77±0.32 | 84.60±0.43 |
| 2 | ✓ | ✓ | ✓ | | 69.21±0.11 | 85.31±0.16 |
| 3 | ✓ | | ✓ | ✓ | 70.18±0.09 | 85.61±0.38 |
| 4 | ✓ | ✓ | ✓ | ✓ | 71.47±0.38 | 86.67±0.33 |

实验结果表明，在 5-way 1-shot 的任务中，将多尺度特征融合模块嵌入 ResNet-12 的特征提取网络中使得分类准

确率提高了 1.44%。而在 5-way 5-shot 的任务中，分类准确率提高了 0.71%。本文在 ResNet-12 网络中添加多尺

度特征融合模块能够更好地捕获图像中的上下文信息,并提取到多尺度特征作为图网络的节点嵌入特征。在 5-way 1-shot 的任务中,采用 L2 范数度量和构造原型节点的度量方式进行边特征的相似性度量的分类精度比单独使用 L2 范数度量高了 2.41%。在 5-way 5-shot 的任务中,边特征度量上采用 L2 范数度量和构造原型节点的度量的分类精度比单独使用 L2 范数度量高了 1.01%。构造原型节点的边特征度量方式比传统边特征的距离函数相似度量方法更加有效。原型节点的设计可以使成对节点之间的相似性计算转变为了两个节点与原型节点之间的相似性和计算,并提高分类准确率。原型节点的构造,使得模型能够更加全面地捕捉样本间的相似度信息,增强模型对图像之间的关系建模能力,进一步提升了性能。

3 结 论

本文从图网络节点特征的嵌入优化和节点相似性度量出发,将多尺度特征融合模块嵌入进骨干卷积网络中增强网络的特征提取能力,通过利用多尺度特征融合的上下文感知能力和特征融合能力,使得特征提取网络能够捕捉到更全局和更丰富的图像信息。为了解决图网络中成对节点度量而忽略节点之间关联性构造导致节点相似性度量准确性低的问题。本文设计了原型节点,将成对节点之间的相似性计算转变为了两个节点与原型节点之间的相似性和计算,从而提高了相似性的度量准确性。本文模型在 *mini*ImageNet、*tiered*ImageNet、CIFAR-FS 和 CUB-200-2011 4 个数据集上进行了实验。实验结果表明,本文提出的 PEGN 优于其他图网络方法,提升了小样本图像的分类准确率。在未来的工作中,将进一步研究边特征函数设计,采用不同距离对边特征相似性进行度量,解决边特征度量单一对于分类精度的影响。

参考文献

[1] 陈立同,高文,金鑫宁,等. 基于电子舌与 MAML-CNN 模型的大豆产地溯源快速检测[J]. 国外电子测量技术,2022,41(12):140-147.

[2] 李明,赵文仓,秦文谦. 基于图神经网络的任务驱动元学习方法[J]. 电子测量技术,2021,44(16):123-129.

[3] ZHOU J, CUI G, HU S, et al. Graph neural networks: A review of methods and applications[J]. *AI Open*, 2020, 1: 57-81.

[4] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 32(1): 4-24.

[5] WEI T, HOU J, FENG R. Fuzzy graph neural network for few-shot learning[C]. 2020 International Joint Conference on Neural Networks(IJCNN), 2020: 1-8.

[6] MA Y, BAI J H, AN S, et al. Transductive relation-propagation network for few-shot learning [C]. *International Joint Conference on Artificial Intelligence*, 2020, 804-810.

[7] KIM J, KIM T, KIM S, et al. Edge-labeling graph neural network for few-shot learning[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2019, 11-20.

[8] TIAN Y, 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.

[9] 黄友文,胡燕芳,魏国庆. 结合原型校准分布的小样本学习方法[J]. 电子测量技术,2022,45(5):132-139.

[10] ZHOU X Y, ZHANG Y H, WEI Q R. Few-shot fine-grained image classification via GNN [J]. *Sensors*, 2022, 22(19): 7640.

[11] YANG L, LI L, ZHANG Z, et al. DPGN: Distribution propagation graph network for few-shot learning [J]. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, 13387-13396.

[12] TANG S, CHEN D, BAI L, et al. Mutual CRF-GNN for few-shot learning[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2021: 2329-2339.

[13] ZHU Z, LIU G, HUI G, et al. Semantic segmentation of FOD using an improved deeplab V3+ model [J]. 2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems(CYBER), 2022, 791-796.

[14] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [J]. *Computer Vision-ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII*, 2018, 11211: 3-19.

[15] 王子昭,景明利,史金钢,等. 一种改进 CBAM 机制和细节恢复的单幅图像去雾算法[J]. 电子测量技术, 2023,46(2):161-168.

[16] VINVALS O, BLUNDELL C, LILICRAP T, et al. Matching networks for one shot learning [J]. *Advances in Neural Information Processing Systems*, 2016, 3630-3638.

[17] REN M, TRIANTAFILLOU E, RAVI S, et al. Meta-learning for semi-supervised few-shot classification[J]. *International Conference on Learning Representations*, 2018, DOI: 10.48550/arXiv.1803.00676.

- [18] CHEN W Y, LIU Y C, KIRA Z, et al. A closer look at few-shot classification[J]. International Conference on Learning Representations, 2019, DOI: 10.48550/arXiv.1904.04232.
- [19] BERTINETTO L, HENRIQUES J F, TORR P H S, et al. Meta-learning with differentiable closed-form solvers[J]. ArXiv Preprint, 2018, ArXiv:1805.08136.
- [20] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2014: 1-42.
- [21] WERTHEIMER D, TANG L, HARIHARAN B. Few-shot classification with feature map reconstruction networks[J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 8012-8021.
- [22] YANG Z, WANG J, ZHU Y. Few-shot classification with contrastive learning[J]. European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 293-309.
- [23] XIE J T, LONG F, JIAMING L, et al. Joint distribution matters: Deep brownian distance covariance for few-shot classification[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2022: 7962-7971.
- [24] KANG D, KWON H, MIN J, et al. Relational embedding for few-shot classification[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 8822-8833.
- [25] LIANG Y, CHEN Z T, LIN A Q, et al. Three-dimension attention mechanism and self-supervised pretext task for augmenting few-shot learning [J]. IEEE Access, 2023, 11: 59428-59437.
- [26] WANG G P, WANG Y X, PAN Z Q, et al. ViTFSL-Baseline: A simple baseline of vision transformer network for few-shot image classification[J]. IEEE Access, 2024, 12(99): 11836-11849.
- [27] YU T Y, HE S, YI-ZHE S, et al. Hybrid graph neural networks for few-shot learning [J]. AAAI Conference on Artificial Intelligence, 2022: 3179-3187.
- [28] LYU Q, WANG W Q. Compositional prototypical networks for few-shot classification[C]. CoRR, 2023, 9011-9019.
- [29] SAI Y, FAN L, DELONG C, et al. Few-shot classification via ensemble learning with multi-order statistics[C]. IJCAI 2023, 2023, 1631-1639.
- [30] FLOOD S, YONGXIN Y, LI Z, et al. Learning to compare: relation network for few-shot learning[J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 1199-1208.

作者简介

陈望, 硕士, 主要研究方向为小样本图像分类。

E-mail: 2267202170@qq.com

李志玲, 硕士, 主要研究方向为多源域自适应与智能算法。

E-mail: lizhiling2022@163.com

王前, 硕士, 主要研究方向为工业缺陷检测。

E-mail: 1711352196@qq.com

包春梅, 硕士, 主要研究方向为小样本图像分类。

E-mail: 1531872403@qq.com

王林(通信作者), 教授, 主要研究方向为图像处理与机器视觉。

E-mail: wanglin@gzmu.edu.cn