

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2025)02-0518-15

论文引用格式: Yu Y, Chen N and Cheng K Y. 2025. Uncertainty domain awareness network for cross-domain few-shot image classification. Journal of Image and Graphics, 30(02):0518-0532(余悦, 陈楠, 成科扬. 2025. 不确定性域感知网络在少样本跨域图像分类中的研究. 中国图象图形学报, 30(02):0518-0532)[DOI:10.11834/jig.240142]

不确定性域感知网络在少样本跨域 图像分类中的研究

余悦, 陈楠, 成科扬*

江苏大学计算机科学与通信工程学院, 镇江 212013

摘要: **目的** 跨域少样本学习的主要挑战在于, 很难将源域的知识推广到未知的目标域中。最近的一些少样本学习模型试图通过在元训练过程中诱导图像多样化来解决这一问题。然而, 其中一些模拟未知领域任务的方法效果有限, 因为它们不能有效地模拟领域偏移, 其生成的内容变化范围狭窄, 难以从域偏移中学习到有意义的域不变知识。为了提升少样本模型的跨域泛化能力, 提出了一个基于不确定性增强的域感知网络(uncertainty enhancement based domain-aware network, UEDA)。**方法** 基于不确定性增强的域感知框架从特征通道视角探索和提取其中可用于缓解领域偏移的关键知识。首先提出一个不确定性特征增强方法, 将特征的充分统计量定义为服从高斯分布的概率表示, 以源域充分统计量为分布中心建模不确定性分布。随后, 从不确定性分布中生成有别于加性扰动的挑战性特征, 从而挖掘不同域之间的共性知识; 其次, 提出了基于不确定性增强的域感知方法, 将源特征和生成特征视为来自不同领域的信息, 利用域鉴别器计算特征通道与领域信息的相关性, 从而帮助模型挖掘领域之间的潜在关联并鉴别出其中的域因果信息用于学习。**结果** 实验使用 Mini-ImageNet、CUB(caltech-ucsd birds)、Plantae、EuroSAT (land use and land cover classification with sentinel-2) 和 Cropdiseases 共 5 个数据集评估所提出方法的跨域泛化表现。实验遵从纯源域泛化, 其中在图神经网络(graph neural network, GNN)分类框架下, 以 Mini-ImageNet 数据集作为源域, 模型在后 4 个目标域的 1-shot 和 5-shot 设置下其平均精度分别为 59.50%、47.48%、79.04% 和 75.08%, 表明了所提出方法能有效提高基于源域的跨域图像分类能力。**结论** 本文所提出的基于不确定性增强的域感知网络框架使得模型在训练阶段适应各种域偏移, 并从中学习到有效的可泛化知识, 从而提高在少样本条件下的跨域图像分类能力。**关键词:** 图像分类; 跨域少样本图像分类; 少样本学习(FSL); 域泛化; 深度学习

Uncertainty domain awareness network for cross-domain few-shot image classification

Yu Yue, Chen Nan, Cheng Keyang*

School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China

Abstract: Objective Inspired by the fast learning properties of humans and transfer learning, researchers have proposed a new machine learning paradigm—few-shot learning. The purpose of few-shot learning is to enable models to quickly learn new knowledge or concepts from a limited amount of labeled data that can be used in unknown categories. Currently,

收稿日期: 2024-03-27; 修回日期: 2024-06-06; 预印本日期: 2024-06-13

* 通信作者: 成科扬 kycheng@ujs.edu.cn

基金项目: 国家自然科学基金项目(62372215); 江苏省重点研发计划资助(BE2022781)

Supported by: National Natural Science Foundation of China (62372215); Jiangsu Province Key R&D Program, China (BE2022781)

few-shot image classification is based on the framework of meta-learning, which divides the model learning process into a meta-training phase and a meta-testing phase. Existing solutions can be broadly classified according to the following differences in concepts: 1) optimization-based methods, the basic idea of which is to allow the model to find the parameters that can optimize performance under the training of multiple sets of task data; 2) metric learning-based methods, whose core idea is to construct an optimal embedding space for measuring distances so that the distance between similar samples is as small as possible; and 3) data manipulation-based methods, which use some basic data augmentation (e. g., rotating, clipping, and adding noise) to increase the diversity of the training samples and the amount of data in these three main categories. However, these works tend to follow strict assumptions, such as the smoothness assumption, clustering assumption, and prevalence assumption, and require that the training data and test data come from the same distribution. This situation makes it difficult to ensure data from the same distribution setting during the training process of the model in certain real-world scenarios, such as medical imaging, military applications, and finance, where issues such as difficulty in data access and data privacy make it challenging to use labeled data from other domains to provide a priori knowledge. Here, an uncertainty enhancement-based domain-awareness network (UEDA) is proposed to alleviate the problem of domain distribution bias encountered in the learning process of the few-shot model. **Method** The uncertainty enhancement-based domain-awareness approach explores and extracts key knowledge from the feature channel perspective, which can be used to mitigate domain bias. An uncertainty feature augmentation approach is first proposed, where the feature channel contains both domain-relevant and domain-independent information, suggesting that the generalized learning of the model may be correlated with the feature channel's ability to extract domain-generalized knowledge. However, most of these works consider feature statistics to be deterministic and use additive perturbations (i. e., swapping and interpolation) to achieve augmentation, practices that may lead to models that are negatively affected by voluminous domain-specific information. The uncertainty enhancement approach models the uncertainty distribution by defining the feature sufficient statistics fixed value as a probabilistic representation of uncertainty obeying a Gaussian distribution with the source feature sufficient statistic as the center of the distribution and the standard deviation defined as being the potential range of variation of the probability. The new features generated by the uncertainty enhancement method can be effectively distinguished from the source domain features. The second part of the UEDA is a domain-awareness approach. In the domain-aware module, source and generated features are considered information from different domains and ensure that both features are within a reasonable challenging offset by maximizing the interdomain differences. A domain discriminator is also introduced to compute the correlation between each channel, and the domain information is used as a way to extract effective generalizable knowledge. **Result** The cross-domain generalization performance of the proposed method was evaluated on five datasets, namely, Mini-ImageNet, CUB, Plantae, Cropdiseases, and EuroSAT. The experiments follow single-domain generalization, use the Mini-ImageNet dataset as the source domain and the latter four datasets as the target domains. Then, the initial findings are compared with current mainstream methods under three classification frameworks, namely, MatchingNet, RelationNet, and the GNN. The subsequent experiments follow 5-way 1-shot and 5-way 5-shot settings on the CUB, Plantae, EuroSAT, and Cropdiseases datasets, and the accuracies of the proposed UEDA are 41.01%, 58.78%, 38.07% and 51.36%, respectively; under the MatchingNet classifier, 58.37%, 80.45%, 59.48%, and 69.91%; and under the GNN classifier, the accuracies of the UEDA are 49.36%, 69.65%, 38.48%, 56.49%, 68.98%, 89.11%, 64.87% and 85.29%, respectively. Comparative experimental results demonstrate that the proposed UEDA method can effectively improve the cross-domain generalizability of the model. Furthermore, ablation experiments were conducted to validate the effectiveness of the modules of the proposed methodology, and the results show that the modules are mutually reinforcing and indispensable in the overall methodology. **Conclusion** The uncertainty enhancement-based domain-awareness network proposed in this study allows the model to adapt to various domain offsets during the training phase and learn effective generalizable knowledge from them, thus improving cross-domain image classification with fewer samples.

Key words: image classification; cross-domain few-shot image classification; few-shot learning(FSL); domain generalization; deep learning

0 引言

快速学习是人类智能的一个重要特征,即使是儿童也能从少量样本中快速归纳出经验并学习新事物。迄今为止,深度学习的成功在很大程度上依赖于大规模的训练数据,这表明了人工智能也存在一定的局限性,即人工智能仍然无法像人类智能那样只通过几个样本就能学会新知识。研究人员受到人类快速学习特性和迁移学习的启发,提出了一种全新的机器学习范式——少样本学习。少样本学习使得模型从数量有限的带标签数据中快速学习到可用于未知类别的新知识或新概念。人们提出了许多元学习模型,如 MAML(model-agnostic meta-learning)(Finn 等,2017)和 GAP(geometry-adaptive preconditioner)(Kang 等,2023)等,这些模型通过学习任务间归纳偏差的分布来提高泛化学习能力,从而解决少样本学习问题。

然而,现有的大多数少样本学习方法(陶鹏 等,2024)还需要遵循各种各样的假设,如光滑性假设、聚类假设和流形假设等,并且要求训练数据和测试数据处于相同的分布中(董杨洋 等,2023)。然而,

最近的研究表明(Gou 等,2020),当元训练和元测试任务之间存在巨大的任务分布偏差时,少样本学习模型的表现甚至不如传统的预训练和微调方法。跨域少样本学习方法强调测试任务的领域分布和类别是完全未知的。因此,越来越多的学者开始关注跨域少样本学习。

图1展示了少样本学习问题与跨域少样本学习问题之间的区别。跨域少样本学习研究必须同时处理领域泛化任务和少样本学习任务所面临的问题,这与少样本学习研究和领域自适应研究不同。现有的跨域少样本学习方法成果有限,主要分为两类。一种情况是纯源域泛化,在该设置下模型在训练阶段无法获取任何目标信息,仅从源域数据中进行学习并总结出可用于泛化到新域的知识,主要代表工作有 FWT(feature-wise transformation)(Tseng 等,2020)、URL(universal representation learning)(Li 等,2021b)等;另一种情况则是保留最基本的跨域少样本设置,训练阶段引入部分目标域数据作为学习引导的辅助域设置,这种设置下的代表作有 TGDM(target guided dynamic mixup)(Zhou 等,2022)、Meta-FDMixUP(Fu 等,2021)。本文采用的是具有挑战性的纯源泛化设置。

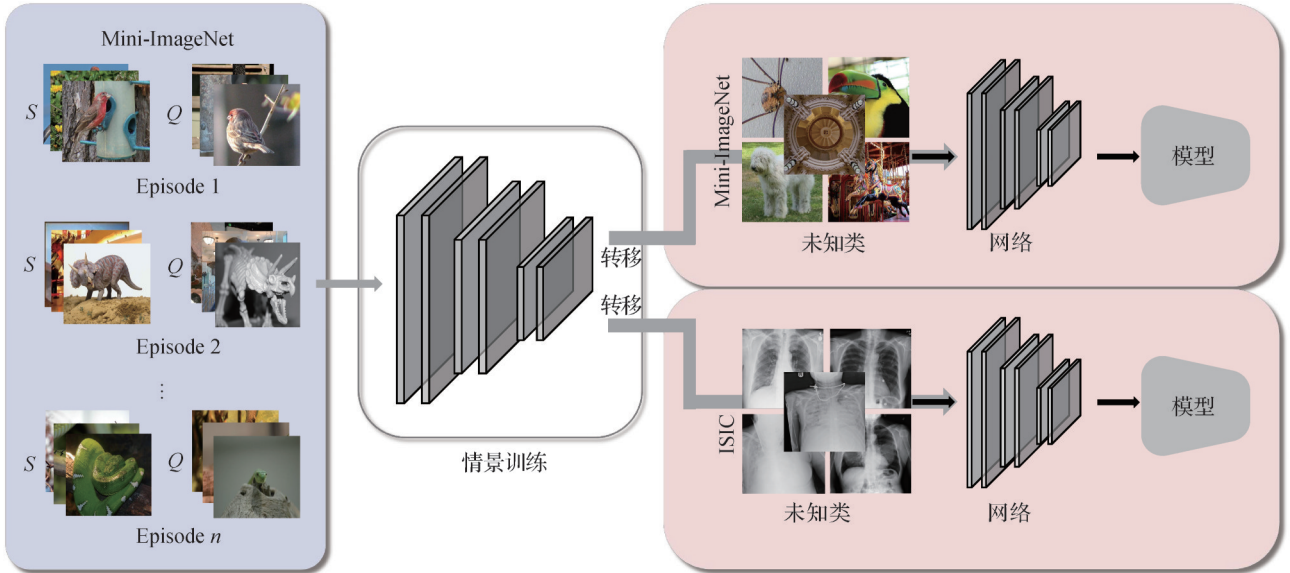


图1 少样本跨域学习过程

Fig. 1 Few shot cross-domain learning process

相关研究表明,在特征通道中包含着域相关和域无关的信息,这表明模型的泛化学习可能与特征通道对域泛化知识的提取能力存在关联。然而,大

多数工作都将特征统计值视为确定值,并使用加性扰动(即交换和插值)来实现增强,这些做法可能导致模型受到过多域特定信息带来的负面影响。因

此,本文提出一个基于不确定性增强的域感知网络(uncertainty enhancement based domain-aware network, UEDA),从特征通道的新视图审视在少样本情况下的跨域学习问题。具体来说,基于不确定性增强的域感知网络主要由不确定性特征增强模块和显著域感知模块所组成。在不确定性特征增强模块中,将特征的充分统计定值建模为服从不确定性分布的概率表示,从而在训练阶段人为构造域偏移场景以训练模型的跨域泛化能力;在域感知模块中,将源特征和生成特征视为来自不同领域的信息,通过引入域鉴别器来计算每个通道与域信息之间的相关性,以此提取有效的可泛化知识。最后,在5个数据集上(Mini-ImageNet、CUB(caltech-ucsd birds)、Plantae、EuroSAT(land use and land cover classification with sentinel-2)和Cropdiseases)进行实验,实验结果表明,所提出的方法能有效提高少样本模型的跨域泛化能力。

本文工作主要有4个方面的贡献:1)在跨域少样本图像分类领域,从特征通道这一视角下挖掘域信息之间的潜在相关性,并提取其中的域不变信息以提高模型的泛化学习能力;2)提出一个不确定性特征增强模块,将特征的充分统计视做不确定性概率表示,并进行不确定性建模以生成区别于源分布的新特征,使得模型在学习过程中适应领域偏移;3)提出一个基于不确定性增强的域感知模块,通过引入域鉴别器辅助模型在学习过程中挖掘可泛化的语义信息,从而提高其泛化性能;4)该方法在5个数据集上进行相关实验,其结果均比目前最先进的方法表现更好。

1 相关工作

1.1 跨域小样本学习

各种元学习模型在少样本情况下取得了显著的成功。然而,当分布发生变化时,元学习模型的性能反而不如传统的微调模型。跨域少样本学习是少样本元学习的一个分支,其中的训练数据集和测试数据集定义为来自不同的领域。许多工作都将研究重点放在如何利用来自不同领域的数据上。该领域的学者们迄今已提出了FWT(Tseng等,2020)弱泛化基准和BSCD-FSL(broader study of cross-domain few-shot learning)强泛化基准。一些方法侧重于在元训

练阶段使数据集多样化,以提高元学习模型的泛化能力。Sun等人(2021)引入了可解释的方法,根据可解释的结果反向优化网络。在特征变换视角下,Tseng等人(2020)引入基于超参数的特征扰动,在元训练阶段模拟跨领域场景;Hu和Ma(2022)则提出对抗学习范式进行仿射变换。在任务变换方面,Wang和Deng(2021)在任务层面进行增强,为学习生成自适应诱导偏差。在应用方面,Liu等人(2024)提出基于Transformer的交叉注意力机制并用于高光谱分类领域。

虽然上述大多数方法都能根据定义的充分统计信息生成虚拟样式,但它们仍局限于线性变换,无法消除原始数据中特定域知识带来的负面影响。

1.2 少样本分类

少样本学习(few-shot learning, FSL)旨在训练一个模型,该模型只需通过少量训练实例就能学习到可推广的知识。针对这一研究所提出的解决方案主要分为元学习方式和度量学习方式。基于元学习框架的模型将输入数据以任务为单位进行划分,使得元学习模型学习可用于泛化的元知识;基于度量学习的研究方法构建了用于距离测量的最优嵌入空间来解决少样本学习任务。最近的研究包括基于学习排名的RankDNN(learning to rank for few-shot learning)(Guo等,2023)、基于余弦变换器的FS-CT(few-shot cosine transformer)(Nguyen等,2023)和基于最近邻的P3DC-shot(Wang等,2023)。上述方法都是在假设训练任务和测试任务分布相同且属于同一领域的前提下进行研究的。

然而,当训练阶段的源域数据集和测试阶段的未知目标域数据集之间存在分布偏差时,现有元学习模型的性能就会受到很大限制。为了缓解这方面的不足,本文提出的方法在训练阶段模拟跨域情况,以解决跨域问题。

1.3 域泛化

领域泛化方法的核心是在共享相同标记环境的情况下,从已知的源领域泛化到不可知论的目标领域。现有的领域泛化方法可分为3个方面:数据操作、度量学习和表征学习。数据操作方法(Li等,2019;Volpi等,2018)旨在对源数据进行增强和更改,以扩展模型在训练阶段使用的数据;度量学习(Balaji等,2018)通过设计新的损失或整合多个源域分类器来提高泛化能力;表征学习(Li等,2021a;

Wang 等, 2022)指的是允许模型在训练阶段对特征进行操作,以模仿跨领域的特征,从而学习可适用于不同领域的领域不变特征。例如,Chen 等人(2022)提出,通过混合训练实例的风格,可以增加源领域的多样性,从而提高训练模型的泛化性能。

然而,现有的领域泛化研究侧重于在多源环境和训练数据量充足的情况下解决问题,这在实践中很难满足。因此,本文方法考虑使用纯源域数据训练模型,且训练数据量较小。

2 研究方法

2.1 相关定义

2.1.1 N -way K -shot 问题定义

在少样本设置中,训练和测试阶段的数据由支持集和查询集组成,模型在两个阶段使用的类别不重叠,即 $C_{\text{train}} \cap C_{\text{test}} = \emptyset$,每一个少样本分类任务都被称为 N -way K -shot 问题。具体来说,从数据集中随机抽取 N 类样本,从每一类中随机抽取 $K + m$ 个实例,从每一类中随机抽取 K 个实例作为支持集,再从每一类中抽取剩余的实例作为查询集。

2.1.2 跨域情景定义

元学习模型旨在从各种任务中学习最佳归纳偏差,而之前的一些工作是在默认的相同任务分配下进行的。相比之下,跨域少样本学习的任务设置考虑到了模型预测过程中的领域偏移。具体来说,采用跨域设置,使用单个或多个源域 T_{source} 进行训练。本文的目标是从中学习一个元模型,该模型在多个未知域 $\{T_{\text{domain1}}, T_{\text{domain2}}, \dots, T_{\text{domainn}}\}$ 中具有良好的泛化能力。

2.2 方法概述和总体分析

现有的跨域少样本学习研究大多局限于在源域数据的风格视角下进行加性扰动,并从中学习通用的任务不可知知识。然而,由于源域数据特征 f 隐含着各种非因果因素(照明、颜色和背景纹理等),导致从中学习到的可泛化知识并不纯粹,这就导致现有的少样本模型的泛化表现能力有限。

如前所述,本文提出一个基于不确定性增强的域感知网络使得模型在学习过程中区分和消除潜在的虚假相关性,并提取更纯粹的领域不变特征。首先,第1部分为一个不确定性特征增强模块,将特征

充分统计量视为遵从高斯分布的概率数值,并进行不确定性建模以生成全新特征统计量,以增加其多样性,并使得潜在的因果因素更加显著。第2部分为一个领域知识感知模块,将源特征和生成特征视为两个不同领域分布的数据,从而输入到领域鉴别器中识别其中的域因果不变特征和非因果特征,从而指示模型聚集在因果特征而忽略非因果特征。

图2为所提出的基于不确定性增强的域因果鉴别框架。该框架由特征提取部分、不确定性特征增强模块(uncertainty enhancement module, UEM)、基于不确定性增强的领域感知模块(domain-aware module, DAM)和度量分类器这4个部分构成。

2.2.1 特征提取器

为了防止少样本模型在基类上过拟合,现有的基于少样本训练范式的工作常使用权重参数较少的残差网络 ResNet-10(residual neural network)作为特征提取器。本文提出的框架使用 ResNet-10 作为网络的特征提取器,具体由1个通道数为64的 7×7 卷积层、批次归一化层、 3×3 平均池化层和4个卷积块组成。4个卷积块的通道数分别为64、128、256、512。需要注意的是,第1个卷积块与之后3个卷积块的组成有所区别,第1个卷积块由两个 3×3 卷积层、两个批次归一化层和两个 ReLU(rectified linear unit)激活函数所组成;而后3个卷积块结构相同,均由两个 3×3 的卷积层、两个批次归一化层、两个 ReLU 激活层和两个下采样层所组成。

2.2.2 度量分类器

基于度量学习的少样本分类方法旨在将图像通过特征提取器映射到一个高维特征空间,选用合适的度量方案度量样本间的距离,从而计算两者之间的相似度来进行分类。目前现有的基于欧氏空间度量的少样本学习模型难以精准捕捉实例特征与实例标签之间的关系,难以提高少样本模型的跨域泛化表现。

跨域少样本学习方法中最常用的度量分类器为基于图结构数据的图神经网络(graph neural network, GNN)(Garcia 和 Bruna, 2018),具体结构如图3所示,从数据之间的图结构视角出发,利用数据特征节点之间的关系信息去提高模型的泛化精度和鲁棒性。GNN 采用逐层的方式对网络中每一层的节点信息进行更新,主要包含聚合和转换两个部分。聚合阶段的目的是汇集每个节点的相邻节点信息;转

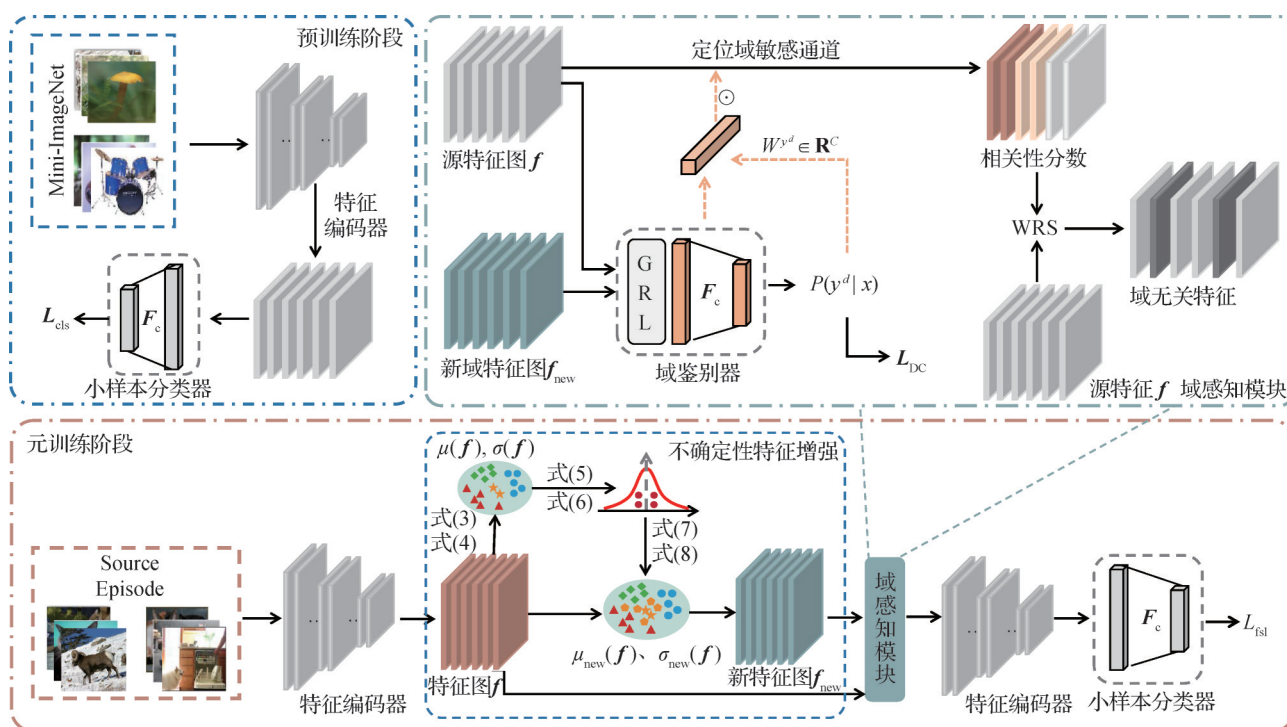


图2 方法整体框架

Fig. 2 Overall framework of the methodology

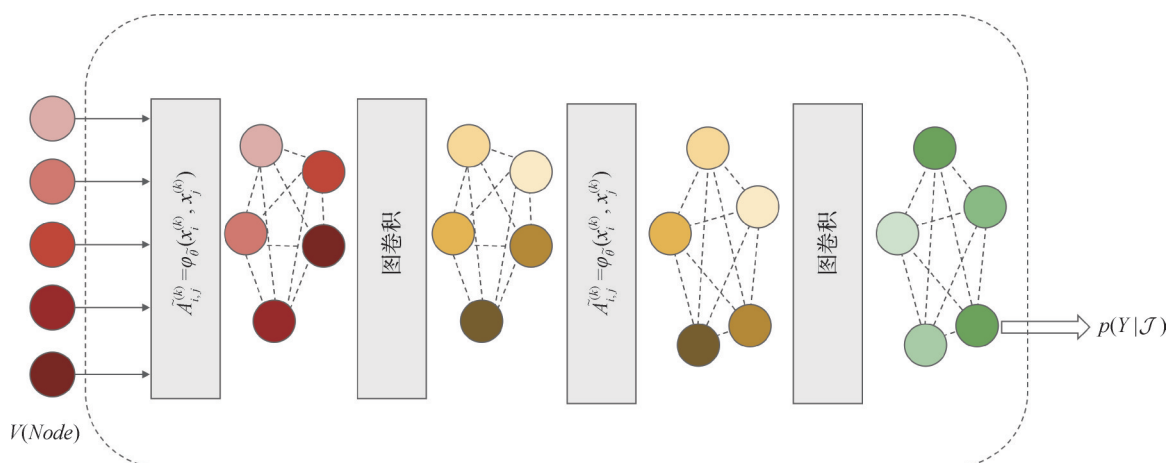


图3 图神经网络

Fig. 3 GNN

换阶段的目的则是将汇集到的信息更新为新的节点内容。

为深度评估所提出框架的有效性,所提出的方法引入以图神经网络为代表的3个度量分类头来探索对跨域少样本模型精度的影响。

2.2.3 不确定性特征增强模块

为了缓解跨域导致的模型性能下降,在训练过程中将FSL模型暴露在源域图像样式之外的具有挑战性分布之中,从而模拟全新未知的目标域的特征

分布。通过将充分统计的确定值视为服从特定高斯分布的概率表示,并对统计量的不确定性分布进行建模,从而模拟未知的数据分布。给定输入数据的均值和方差,其正态分布以最大熵作为当前数据状态的有效展现。因此,假设抽取的训练批次特征图遵循独立同分布的多元正态分布。设 f 表示为特征图, $f_1, f_2, \dots, f_N$ 则为某一批次中的相应观测值。由于多元正态分布的边际分布仍为正态表示,某一批次的特征图的概率表示为

$$p(f) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(f_i - \mu)^2}{2\sigma^2}\right) \quad (1)$$

式中, μ 、 σ 分别为特征图 f 的均值和方差, 概率密度函数可以分解为与总体分布相关且独立于总体分布的部分。通过简化式(1)右侧的乘积, 可得

$$p(f) = (2\pi\sigma)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (f_i^2 - 2\mu f_i + \mu^2)\right) \quad (2)$$

由式(2)可知, 概率密度函数 $p(f)$ 可以分解为一个不依赖于分布参数 μ 、 σ 乘以另一个依赖于 μ 、 σ 和统计量 $\sum_i f_i^2$ 、 $\sum_i f_i$ 的因子的形式。由 Fisher-Neyman 因子分解定理可知, $\sum_i f_i^2$ 和 $\sum_i f_i$ 为正态分布的充分统计量。此外, 由于均值和方差可以通过统计量 $\sum_i f_i^2$ 和 $\sum_i f_i$ 来计算获得, 可知均值和方差也是正态分布的充分统计量。

特征分布的充分统计量包含分布的所有信息。因此, 在模型训练过程中干扰数据特征的均值和方差来模拟各种未知特征分布。具体来说, 给定 $f \in \mathbf{R}^{B \times C \times H \times W}$ 为输入图像到中间层的编码特征图, 其中, W 、 H 分别为特征图的宽和高, C 为通道数, B 为批次大小。则每个批次中实例特征的平均值和方差定义为

$$\mu(f) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W f_{b,c,h,w} \quad (3)$$

$$\sigma^2(f) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (f_{b,c,h,w} - \mu(f))^2 \quad (4)$$

特征的均值 μ 和方差 σ 是最能代表当前域分布状态的信息表示。本文将原始特征的均值和方差作为分布的中心, 标准差定义为概率的潜在变动范围, 从而建模不确定性分布。具体为

$$\sum_{\mu}^2(f) = \frac{1}{B} \sum_{i=1}^B (\mu(f) - E_b[\mu(f)]) \quad (5)$$

$$\sum_{\sigma}^2(f) = \frac{1}{B} \sum_{i=1}^B (\sigma(f) - E_b[\sigma(f)]) \quad (6)$$

可知, 新特征分布的均值和方差仍服从高斯分布, 相应的数学表示定义为 $\mu_{\text{new}} \in \mathbf{R}^{B \times C} \sim N\left(\mu, \sum_{\mu}^2\right)$ 和

$\sigma_{\text{new}} \in \mathbf{R}^{B \times C} \sim N\left(\sigma, \sum_{\sigma}^2\right)$ 。为了进一步探索随机性, 本文使用随机采样来实现概率表示的随机性。新分布的均值和方差可以从下式中随机生成, 具体为

$$\mu_{\text{new}}(f) = \mu(f) + \varepsilon_{\mu} \sum_{\mu}(f), \varepsilon_{\mu} \sim N(0, I) \quad (7)$$

$$\sigma_{\text{new}}(f) = \sigma(f) + \varepsilon_{\sigma} \sum_{\sigma}(f), \varepsilon_{\sigma} \sim N(0, I) \quad (8)$$

式中, ε_{μ} 和 ε_{σ} 遵循高斯分布, N 为高斯噪声, I 为单位协方差矩阵。最后生成新分布特征 $f_{\text{new}} \in \mathbf{R}^{C \times H \times W}$, H 和 W 分别为特征图的高和宽。具体为

$$f_{\text{new}} = \mu_{\text{new}}(f) \frac{f - \mu(f)}{\sigma(f)} + \sigma_{\text{new}}(f) \quad (9)$$

2.2.4 基于不确定性增强的域感知模块

传统域鉴别器如图4所示, 域鉴别器通过定位和提取特征信道中域信息丰富的特征, 利用其中的域特定信息进行域标签预测。因此, 所提出的方法在第2阶段引入域鉴别器, 通过优化域鉴别器来提高模型对特征信道中域信息的敏感性, 从而帮助模型鉴别出其中的域因果信息和域非因果信息。将源域和生成域视为不同的两个域, 并输入到域鉴别器中。此外, 为了能更合理有效地模拟未知域分布, 需测量特征干扰前后的域差异, 具体为

$$\hat{f} = \text{Flatten}(f), \hat{f} \in \mathbf{R}^{C \times H \times W} \quad (10)$$

$$G(f) = \hat{f} \times \hat{f}^T, f = f_{\text{ori}} \text{ 或 } f_{\text{new}} \quad (11)$$

源中间特征 f_{ori} 和生成特征 f_{new} 之间的域差异由 $G(f_{\text{ori}})$ 和 $G(f_{\text{new}})$ 来确定。通过最大化式(12)的损失函数来保证源特征与生成特征处于不同的分布, 具体为

$$L_{\text{MV}} = \frac{1}{4S^2 C^2} \sum_{i,j} (G_{i,j}(f_{\text{new}}) - G_{i,j}(f_{\text{ori}}))^2 \quad (12)$$

式中, $S = HW$ 。然后, 本文在多个隐藏层中插入域鉴别器, 用来鉴别因果信息和非因果信息。将第 i 层源域中间特征 f_{ori}^i 和生成特征 f_{new}^i 输入到域鉴别器预测域标签并计算域分类损失, 使用梯度反转层 (gradient reversal layer, GRL) 来实现对域鉴别器的优化, 具体为

$$L_{\text{DC}} = \frac{1}{2N} \sum_i [-y_i^d \log(\hat{y}_i^d) - (1 - y_i^d) \log(1 - \hat{y}_i^d)], \quad 1 \leq i \leq 2N \quad (13)$$

式中, y^d , $\hat{y}^d$ 分别为真实域标签和预测域标签。域鉴别器的优化过程为最大最小范式, 具体为

$$\max \min L_D = L_{\text{DC}} - L_{\text{MV}} \quad (14)$$

需要注意的是, 域鉴别器的更新不是随着主网路更新而同步更新的, 因为梯度反转层被放置在域鉴别器之前以截断梯度, 从而避免域鉴别器对主网

络造成负面影响。利用域鉴别器其预测结果的加权激活值,引入相关性系数 z_j 来量化每个信道和域信息之间的相关性,具体为

$$z_j = W_j^{z^d} \times \text{GAP}(f^j) \quad (15)$$

式中, $W_j^{z^d} \in \mathbf{R}^c$ 为域鉴别器针对真实预测值的全连接层(fully connected layer, FC)权重, GAP(global average pooling)为全局平均池化层。加权激活值越高说明该通道对真实预测值的贡献越大,即域特定信息越多。为了减少特征图中的域特定信息,并提取因果特征,所提出的方法利用相关性系数计算去除第 j 个通道的概率,具体为

$$P_j = \frac{z_j}{\sum_{c=1}^c z_c} \quad (16)$$

可知, z_j 的数值越高,代表相对的信道被丢弃的概率越大。随后,对于具有 z_j 的第 j 个通道,使用加权随机选择(weighted random sampling, WRS)算法生成二进制掩码 $m \in \mathbf{R}^c$,具体为

$$m_j = \begin{cases} 0 & j \in \text{TOP}(\{k_1, k_2, \dots, k_c\}, M) \\ 1 & \text{其他} \end{cases} \quad (17)$$

式中, $k_j = r_j^{1/2}$ 为密钥值, r_j 为随机数, $\text{TOP}(\{\cdot\}, M)$ 代表具有键值的 M 个项。此外,定义超参数 $P_d = M/C$ 为要丢弃的通道数。

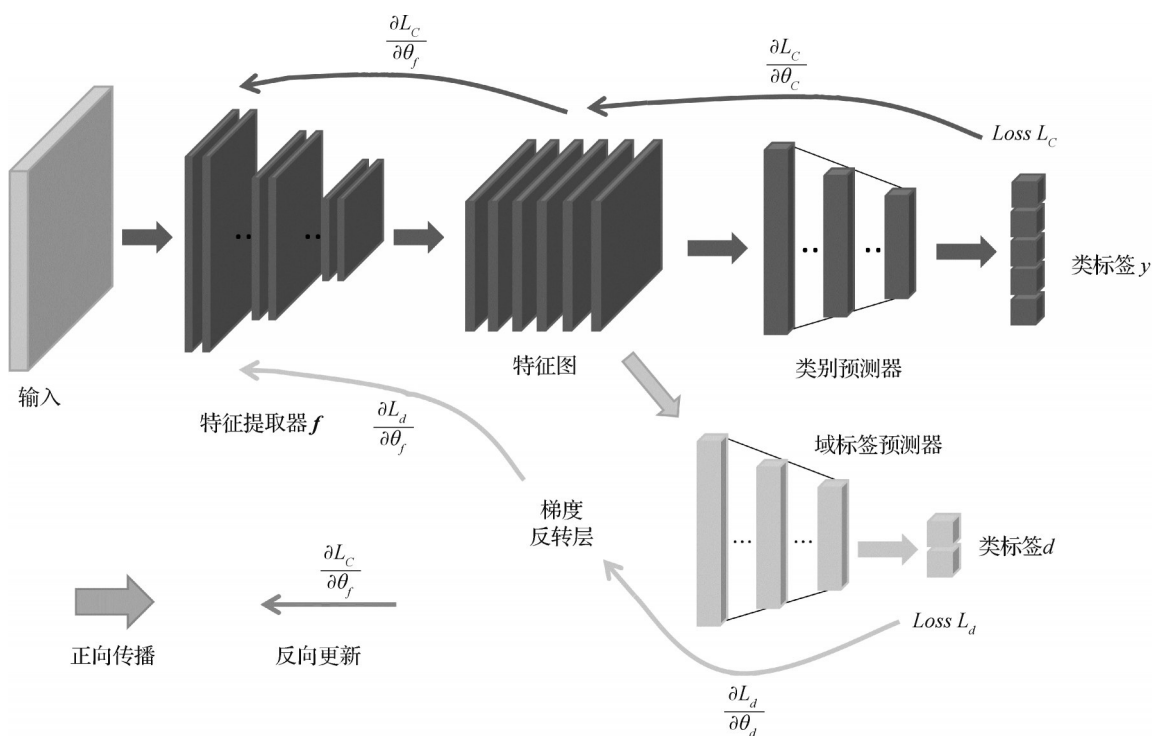


图4 域鉴别器

Fig. 4 Domain discriminator

3 实验

3.1 数据集与实验设置

实验使用了两个CD-FSL基准。第1个基准是FWT基准,使用Mini-ImageNet作为源域数据集 D_s ,使用CUB(Wah等, 2011)和Plantae(Van Horn等, 2018)作为目标域数据集 D_T 。第2个基准是更具挑战性的BSCD-FSL(Gou等, 2020)基准,同样使用Mini-ImageNet作为源域数据集, EuroSAT(Helber等,

2019)和Cropdisease(Mohanty等, 2016)作为目标域数据集。

1)Mini-ImageNet。该数据集是用于少样本分类任务的源域数据集,是在ImageNet数据集的基础上精简而来。Mini-ImageNet包含100个类别,每个类别有600幅图像。这是相对于原始ImageNet数据集而言的一个较小的子集,但仍然保持了丰富的类别。

2)CUB。该数据集是一个用于鸟类识别的计算机视觉数据集,包含来自200个不同鸟类别的图像,共有11 788幅图像。这些图像涵盖了多种鸟类的不同

同姿态和拍摄环境。

3) Plantae。该数据集由来自 5 000 多种不同动植物的 859 000 幅图像组成。需要注意的是图像是用不同类型的相机收集的,具有不同的图像质量,并具有较强的类不平衡。

4) Cropdisease。该数据集为农作物病害数据集,是用于测试的目标域数据集之一,由 38 个类别、共 54 300 幅图像组成。

5) EuroSAT。该数据集为涵盖高速公路、农田场景等共 10 个类别、20 000 幅图像的目标域数据集。

与之前的大多数工作一样,本工作在所有实验中使用 Resnet-10 作为特征提取器,以确保公平性。对于所有实验,将超参数 P_d 设置为 0.35, GRL 层的权重设置为 0.26。在预训练阶段,执行一个标准分类任务来预训练特征编码器。元训练和元测试阶段采用 5-way 1-shot 和 5-way 5-shot 的设置。在元训练阶段,使用 Mini-ImageNet 数据集执行 200 个历元的少样本分类任务,每个历元包含 100 个元任务。根据经验,将 Adam 设定为训练过程的优化器,学习率为 0.001。在元测试阶段,针对不同的基准,随机选取一定的 D_T 集输入网络以测试性能,并重复提取操作 1 000 次,计算平均精度和 95% 的置信区间作为最终结果。

3.2 主要结果分析

在本实验中,使用 MatchingNet (Vinyals 等, 2016)、RelationNet (Sung 等, 2018) 和 GNN (Garcia 和 Bruna, 2018) 3 种基于度量的框架,用于少样本分类头,并作为少样本方法基线参与比较。参与比较的 CD-FSL 方法为 FWT (Tseng 等, 2020)、LRP (Sun 等, 2021)、ATA (Wang 和 Deng, 2021)、AFA (Hu 和 Ma, 2022) 和 MRL (Liu 等, 2024), 这些方法均分别在这 3 种分类框架下进行比较。

具体实验结果如表 1 所示。从结果中,可以发现无论在哪一个基线上,本文提出的方法在大多数情况下都优于其他方法。在 GNN 度量框架的 5-way 1-shot 设置下,所提出的方法在 CUB、Plantae、Crop-disease 和 EuroSAT 数据集上比当前跨域少样本图像分类方法中同是特征变换的 AFA 分别提高了 2.5%、1.72%、1.37% 和 1.75%;与最新的 MRL 方法相比,所提方法提高了 1.33%、1.54%、2.15% 和 1.58%。而在 5-way 5-shot 设置下,与 AFA 方法相比,本文所提方法提高了 1.4%、2.23%、2.05% 和

1.31%。由实验结果可知,无论在 1-shot 设置还是 5-shot 设置下,本文方法在大多数情况下均优于专门为 CD-FSL 任务设计的竞争对手。出现这些结果的主要原因是,所提出的框架引导模型学会区分特征通道中的有效因果知识和无效的域特定知识,集中关注于可用于学习未知内容的可泛化知识,从而提高模型的少样本泛化学习能力。

从表 1 中还可以观察到,无论是在 5-way 1-shot 设置还是在 5-way 5-shot 设置下,用于解决少样本问题的 MatchingNet、RelationNet 和 GNN 框架在表现上甚至比不上跨域少样本学习领域早期提出的 FWT 方法,这表明标准的少样本学习模型和算法并不是专门为解决跨域问题而提出的。这是由于标准少样本问题遵循源域与目标域之间无域偏差的假设,如果在跨域设置下,域偏差则会对少样本模型的预测造成负面影响。

此外,本文还研究了所提出的域感知模块在提高其他跨域少样本分类方法上的有效性。为了公平比较,参与比较的方法均使用 ResNet-10 作为特征提取器, GNN 作为少样本分类器。本文选取了跨域少样本领域中的经典特征变换方法 FWT、ATA 和 AFA 替换了所提出框架中的不确定性增强模块,具体结果如表 2 所示。可以观察到,无论是哪种方法,在所提出的域感知网络框架下均有改进。例如,在 CUB 数据集的 5-shot 设置下,FWT 改进了 3.01%,ATA 改进了 1.97%,AFA 改进了 1.37%。这表明本文方法在不同的方法上具有通用性。

3.3 消融实验

3.3.1 各部件对跨域分类准确度的影响

为了充分探索本文所提出的框架对于模型泛化能力提升的效果,实验使用两个弱泛化数据集 (CUB 和 Plantae) 和两个强泛化数据集 (CropDiseases 和 EuroSAT) 用于评估各个模块的性能。具体实验结果如表 3 所示。本次实验使用 GNN 作为基线,即表中第 1 行所表现的情况。

从表中可以观察到,在使用线性扰动方法替代不确定性增强模块的设置下 (第 2 行),模型的表现比仅仅保留不确定性增强模块设置 (第 3 行) 情况下的表现更好。例如,在 CUB 数据集的 1-shot 设置下,模型在第 2 行和第 3 行之间的精度相差 1.25% (47.93%→46.68%),在 5-shot 设置下,模型精度相差了 2.13% (68.39%→66.26%),这说明仅依靠特

表 1 对比实验结果
Table 1 Comparative experimental results

模型/Shot	FWT 基准				BSCD-FSL 基准			
	CUB		Plantae		Cropdisease		EuroSAT	
	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
MatchingNet	35.89±0.5	51.37±0.8	32.70±0.6	46.53±0.7	57.57±0.5	73.26±0.5	54.19±0.5	67.50±0.5
+FWT	36.61±0.5	55.23±0.8	34.48±0.5	41.69±0.6	54.21±0.5	70.56±0.5	55.62±0.4	63.33±0.5
+ATA	39.65±0.4	57.53±0.4	36.42±0.4	51.05±0.4	55.57±0.5	79.28±0.4	56.44±0.5	68.83±0.4
+MRL	40.22±0.4	56.44±0.4	37.13±0.4	50.93±0.5	56.51±0.5	79.87±0.4	57.09±0.5	69.31±0.4
+本文	41.01±0.4	58.78±0.4	38.07±0.5	51.36±0.5	58.37±0.6	80.45±0.5	59.48±0.5	69.91±0.4
RelationNet	41.27±0.4	56.77±0.4	31.23±0.3	42.71±0.3	53.58±0.4	72.86±0.4	49.08±0.4	65.56±0.4
+FWT	44.07±0.8	59.46±0.7	33.14±0.6	45.08±0.6	57.57±0.5	75.78±0.4	53.53±0.4	69.13±0.4
+ATA	43.02±0.4	59.36±0.4	33.72±0.3	45.32±0.3	61.17±0.5	78.20±0.4	55.69±0.5	71.02±0.4
+MRL	43.91±0.5	58.87±0.5	34.03±0.4	45.88±0.4	61.79±0.5	77.83±0.4	55.98±0.5	71.31±0.4
+本文	45.77±0.5	60.38±0.4	34.22±0.4	46.82±0.6	62.93±0.4	78.66±0.4	56.13±0.5	71.09±0.4
GNN	44.40±0.5	62.87±0.5	33.60±0.4	48.51±0.4	59.19±0.5	83.12±0.4	54.61±0.5	78.69±0.4
+FWT	45.50±0.5	64.97±0.5	32.56±0.4	49.66±0.4	60.74±0.5	87.07±0.4	55.53±0.5	78.02±0.4
+ATA	45.00±0.5	66.22±0.5	34.42±0.4	52.69±0.4	67.47±0.5	90.59±0.3	61.35±0.5	83.75±0.4
+LRP	48.29±0.5	64.44±0.5	37.49±0.4	54.46±0.5	59.23±0.5	86.15±0.4	54.99±0.5	77.14±0.4
+AFA	46.86±0.5	68.25±0.5	36.76±0.4	54.26±0.4	67.61±0.5	87.06±0.3	63.12±0.5	83.98±0.4
+MRL	48.03±0.4	68.33±0.5	36.94±0.4	54.79±0.3	66.83±0.4	88.84±0.4	63.29±0.5	85.38±0.4
+本文	49.36±0.3	69.65±0.6	38.48±0.4	56.49±0.3	68.98±0.4	89.11±0.3	64.87±0.4	85.29±0.4

注:加粗字体表示在不同基线上各列最优结果。

表 2 所提出框架对其他方法的影响
Table 2 Impact of the proposed framework on other approaches

模型	CUB		Plantae		Cropdisease		EuroSAT	
	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
FWT	45.50±0.5	64.97±0.5	32.56±0.4	49.66±0.4	60.74±0.5	87.07±0.4	55.53±0.5	78.02±0.4
+本文	47.27±0.5	67.98±0.5	35.12±0.4	51.73±0.4	63.23±0.5	88.82±0.4	58.84±0.5	80.33±0.3
ATA	45.00±0.5	66.22±0.5	34.42±0.4	52.69±0.4	67.47±0.5	90.59±0.3	61.35±0.5	83.75±0.4
+本文	46.74±0.5	68.19±0.5	35.32±0.4	54.52±0.3	68.03±0.5	90.77±0.3	62.22±0.5	84.67±0.4
AFA	46.86±0.5	68.25±0.5	36.76±0.4	54.26±0.4	67.61±0.5	87.06±0.3	63.12±0.5	83.98±0.4
+本文	48.58±0.4	69.62±0.6	37.98±0.4	55.79±0.4	68.54±0.5	88.72±0.3	64.07±0.5	84.59±0.4

注:加粗字体表示各组各列最优结果。

征转换难以帮助模型学习到可用于泛化的领域不变信息,与所提出的域感知模块相结合则更能发挥出应有的潜力。

此外,从表 3 的第 3 行和第 4 行可以观察到,在这两种设置下,模型的精度差距明显。例如,在 Plantae 数据集的 1-shot 和 5-shot 设置下,模型在第 3 行和第 4 行

表3 各部分的消融实验
Table 3 Ablation experiments for each part

线性 扰动	UE	DA	CUB		Plantae		Cropdisease		EuroSAT	
			1-shot	5-shot	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
×	×	×	44.40±0.5	62.87±0.5	33.60±0.4	48.51±0.4	59.19±0.5	83.12±0.4	54.61±0.5	78.69±0.4
√	×	√	47.93±0.5	68.39±0.4	36.79±0.4	53.57±0.3	67.87±0.5	86.16±0.4	61.31±0.5	83.22±0.4
×	√	×	46.68±0.5	66.26±0.5	36.41±0.4	53.03±0.3	65.96±0.5	84.93±0.4	60.92±0.4	82.46±0.4
×	√	√	49.36±0.3	69.65±0.6	38.48±0.4	56.49±0.3	68.98±0.4	89.11±0.3	64.87±0.4	85.29±0.4

注:加粗字体表示各列最优结果。“√”表示使用对应模块,“×”表示未使用对应模块。

的精度分别相差2.07%、3.46%。分析其原因是域感知网络可以显著化来自不同领域之间的域不变信息,从而帮助模型区分和提取这些关键知识并进行学习。

此外,本文可视化了基线与所提出方法在源域和目标域之间的充分统计值差异,以评估网络对特征统计偏差的鲁棒性。具体结果如图5所示,其横

坐标为特征通道数,纵坐标为特征的均值与标准差。可以观察到,测试数据的未知性质仍然导致某些通道中不可避免的差异。而且不同的通道可能具有不同的域偏差,本文方法构建的不确定性边界可以促进对各种域偏差的模拟。这也说明了所提出的方法更有利于模型提高泛化学习能力。

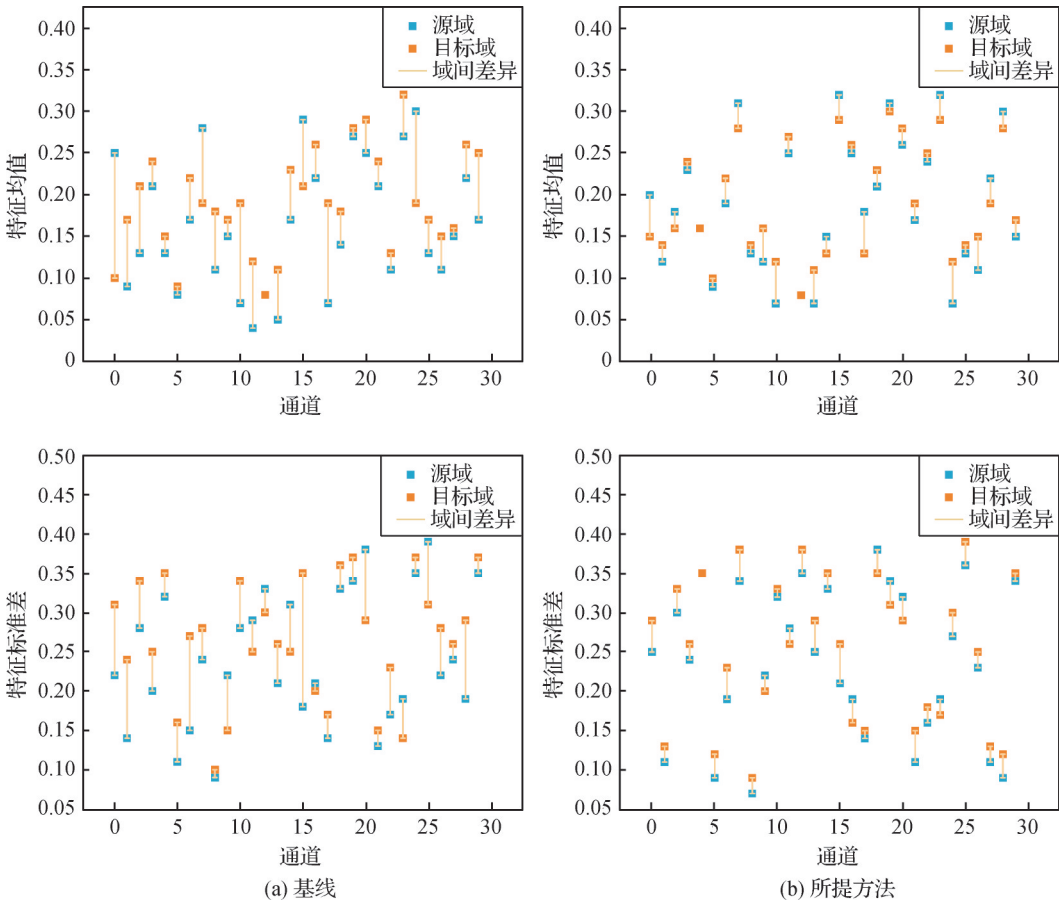


图5 特征差异可视化
Fig. 5 Visualization of feature discrepancy((a)baseline;(b)ours)

3.3.2 超参数对跨域分类准确度的影响

为了研究所提出的框架对超参数 P_d 的影响,进

行实验来探索 P_d 的选取对模型准确度的改变。同样使用GNN作为本次实验的基线模型,结果如图6

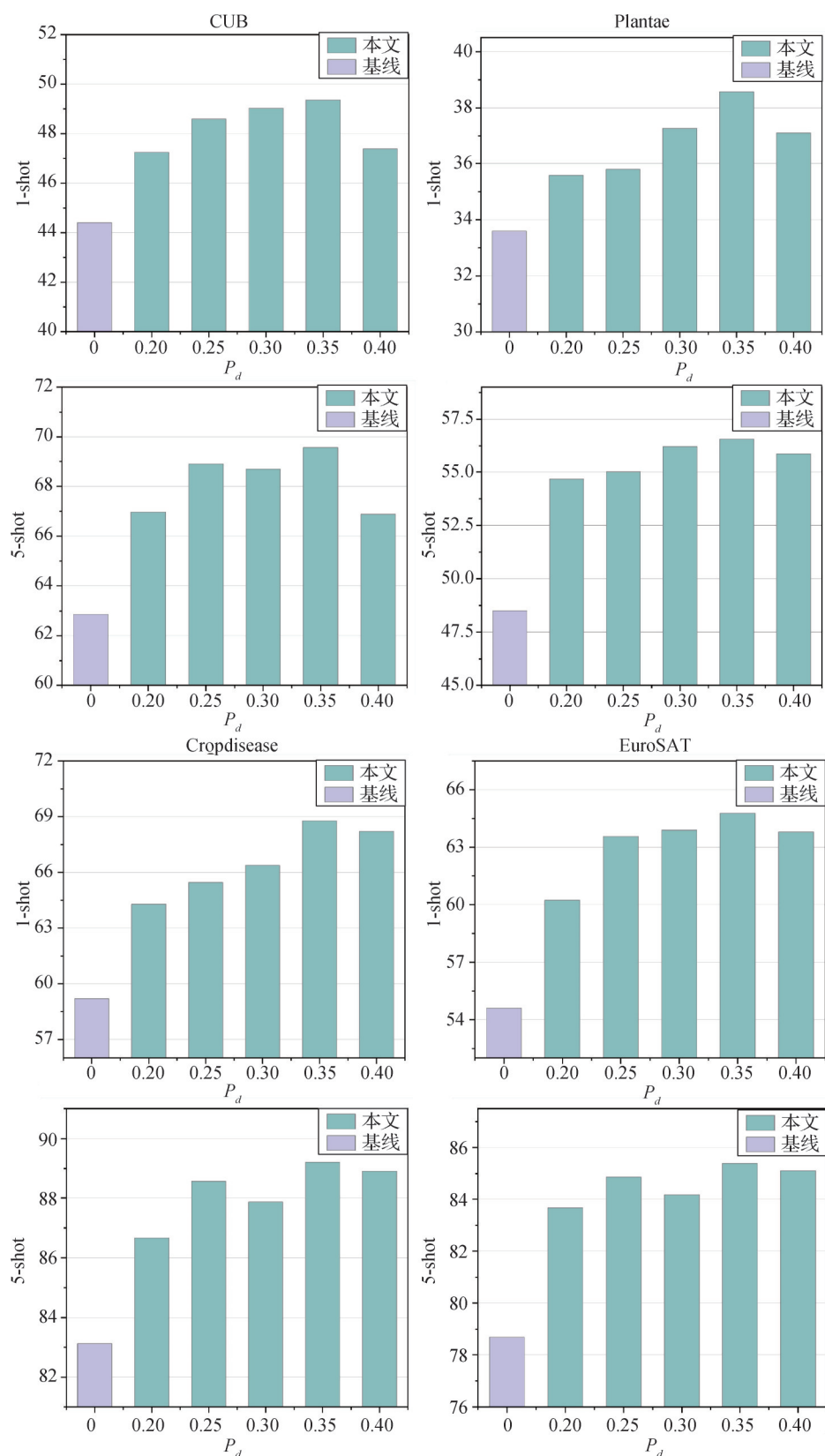


图6 P_d 参数对模型的影响

Fig. 6 Effect of P_d parameters on the model

所示。具体而言, P_d 分别选取 0, 0.2, 0.25, 0.3, 0.35, 0.4。其中, 当 P_d 设为 0 时表示的是 GNN 基线的表现。可以看到, 在 CUB、Plantae、Cropdisease 和 EuroSAT 4 个数据集上, 模型在不同的 P_d 设置均以较大幅度高于基线精准度, 例如, 在 CUB 数据集上, 本文的方法在 1-shot 和 5-shot 设置下比基线提高 4.46% 和 6.77%。当 P_d 设置为 0.35 时, 模型的表现

最好, 因此选取 $P_d = 0.35$ 为实验的默认值。

3.3.3 辅助任务对跨域分类的影响

为了测试模型在训练过程中引入目标域任务是否有必要, 在本次实验引入少量带标签的目标域数据作为辅助任务参与模型的训练过程, 引入 CUB 和 Plantae 两个数据集, 使用 GNN 作为基线模型, 具体结果如表 4 所示。

表 4 辅助任务对模型表现的影响

Table 4 Effect of ancillary tasks on model performance

模型	CUB		Plantae	
	1-shot	5-shot	1-shot	5-shot
	/%			
基线	44.40±0.5	62.87±0.5	33.60±0.4	48.51±0.4
本文(引入辅助任务)	49.27±0.3	69.83±0.6	38.27±0.4	55.95±0.3
本文(无辅助任务)	49.36±0.3	69.65±0.6	38.48±0.4	56.49±0.3

注: 加粗字体表示各列最优结果。

可以观察到, 无论是在 1-shot 设置还是 5-shot 设置下, 模型的精度并无较大差别, 这说明引入辅助任务对模型的跨域准确度并无较大帮助。具体因为, 提出不同分布域的作用是为了在域感知模块中帮助域鉴别器更好地识别有效知识, 而所提出的框架在不确定性增强模块中已经生成新分布域, 因此模型在此种情况下表现相近。

3.3.4 域鉴别器对模型学习的影响

为了评估域感知模块中的域鉴别器是否能有效

辅助模型鉴别特征通道中的域相关信息和域无关信息, 本次实验去除不确定性生成模块, 仅保留域感知模块, 并引入 Mini-ImageNet 和 CUB 作为两个源域数据集, 以 GNN 作为基线方法, 可视化模型学习过程中的注意力图。

结果如图 7 所示, 可以观察到基线注意力图(第 1 列)和域鉴别器生成的领域注意力图(第 2 列)有较多的重叠, 这说明基线模型在学习过程中利用了较多的域特定信息进行预测。此外, 本文所提出的域

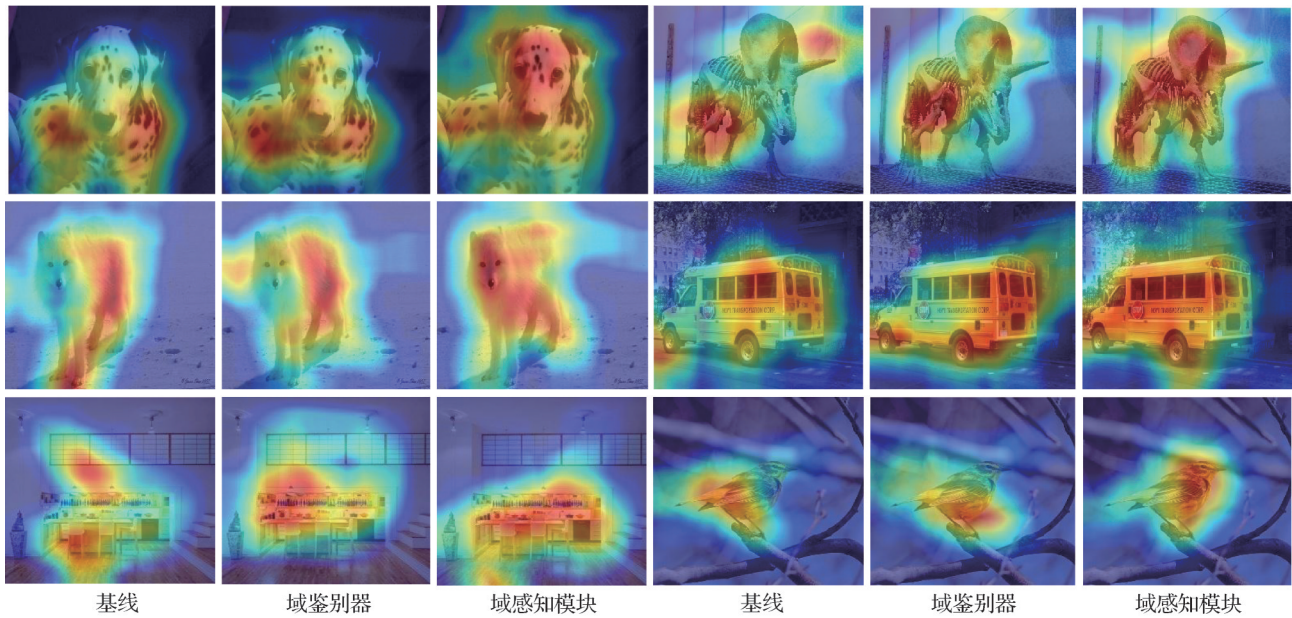


图 7 注意力图可视化

Fig. 7 Visualization of attention map

感知模块(第3行)可以辅助模型聚焦于可泛化的语义信息。例如,针对狗的图像,狗的脸部内容为预测的关键特征之一,而域感知模块可以有效定位该关键部位以挖掘其中的语义信息,而基线模型则更关注斑点特征。这些结果表明域感知模块能帮助模型学习更多的可泛化特征。

4 结 论

本文从通道级别的视角研究了模型在小样本情况下的可泛化能力,并发现模型的泛化表现不佳可能是受到特征图中包含的领域非因果信道的负面影响所导致。因此,提出了一个基于不确定性增强的域感知网络方法。所提出的方法包含不确定性特征增强模块和域感知模块,学习器在这两个模块的高效帮助下,缓解在训练过程中受到的域无关信息产生的影响,并充分提取可泛化的域因果信息来提高表现力。在实验中,使用 Mini-ImageNet 为源域数据集,并在 CUB、Plantae、CropDisease 和 EuroSAT 4 个数据集上进行实验。实验结果表明,所提出的框架优于当前领域中最先进的方法。

在未来的研究中,将考虑突破单一模态训练数据的局限性,引入来自另一个模态的数据作为辅助,以进一步提高模型的泛化性能。

参考文献(References)

- Balaji Y, Sankaranarayanan S and Chellappa R. 2018. MetaReg: towards domain generalization using meta-regularization//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, Canada: Curran Associates Inc.: 1006-1016
- Chen J N, Sun S Y, He J, Torr P, Yuille A and Bai S. 2022. TransMix: attend to mix for vision transformers//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 12125-12134 [DOI: 10.1109/CVPR52688.2022.01182]
- Dong Y Y, Song B B and Sun W F. 2023. Local feature fusion network-based few-shot image classification. *Journal of Image and Graphics*, 28(7): 2093-2104 (董杨洋, 宋蓓蓓, 孙文方. 2023. 局部特征融合的小样本分类. *中国图象图形学报*, 28(7): 2093-2104) [DOI: 10.11834/jig.220079]
- Finn C, Abbeel P and Levine S. 2017. Model-agnostic meta-learning for fast adaptation of deep networks//Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia: JMLR.org: 1126-1135
- Fu Y Q, Fu Y W and Jiang Y G. 2021. Meta-FDMixup: cross-domain few-shot learning guided by labeled target data//Proceedings of the 29th ACM International Conference on Multimedia. Chengdu, China: ACM: 5326-5334 [DOI: 10.1145/3474085.3475655]
- Garcia V and Bruna J. 2018. Few-shot learning with graph neural networks//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada: ICLR: 1-13
- Guo Q Y, Haotong G, Wei X J, Fu Y W, Yu Y Z, Zhang W Q and Ge W F. 2023. RankDNN: learning to rank for few-shot learning//Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington, USA: AAAI Press: 728-736. [DOI: 10.1609/aaai.v37i1.25150]
- Guo Y H, Codella N C, Karlinsky L, Codella J V, Smith J R, Saenko K, Rosing T and Feris R. 2020. A broader study of cross-domain few-shot learning//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 124-141 [DOI: 10.1007/978-3-030-58583-9_8]
- Helber P, Bischke B, Dengel A and Borth D. 2019. EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7): 2217-2226 [DOI: 10.1109/JSTARS.2019.2918242]
- Hu Y X and Ma A J. 2022. Adversarial feature augmentation for cross-domain few-shot classification//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 20-37 [DOI: 10.1007/978-3-031-20044-1_2]
- Kang S, Hwang D, Eo M, Kim T and Rhee W. 2023. Meta-learning with a geometry-adaptive preconditioner//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 16080-16090 [DOI: 10.1109/CVPR52729.2023.01543]
- Li P, Li D, Li W, Gong S G, Fu Y W and Hospedales T M. 2021a. A simple feature augmentation for domain generalization//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 8866-8875 [DOI: 10.1109/ICCV48922.2021.00876]
- Li W H, Liu X L and Bilen H. 2021b. Universal representation learning from multiple domains for few-shot classification//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 9506-9515 [DOI: 10.1109/ICCV48922.2021.00939]
- Li Y Y, Yang Y X, Zhou W and Hospedales T. 2019. Feature-critic networks for heterogeneous domain generalization//Proceedings of the 36th International Conference on Machine Learning. Long Beach, USA: PMLR: 3915-3924
- Liu C, Yang L W, Li Z, Yang W, Han Z G, Guo J Z and Yu J Y. 2024. Multi-level relation learning for cross-domain few-shot hyperspec-

- tral image classification. *Applied Intelligence*, 54(5): 4392-4410 [DOI: 10.1007/s10489-024-05384-3]
- Mohanty S P, Hughes D P and Salathé M. 2016. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7: #1419 [DOI: 10.3389/fpls.2016.01419]
- Nguyen Q H, Nguyen C Q, Le D D and Pham H H. 2023. Enhancing few-shot image classification with cosine transformer. *IEEE Access*, 11: 79659-79672 [DOI: 10.1109/ACCESS. 2023. 3298299]
- Sun J M, Lapuschkin S, Samek W, Zhao Y Q, Cheung N M and Binder A. 2021. Explanation-guided training for cross-domain few-shot classification//*Proceedings of the 25th International Conference on Pattern Recognition*. Milan, Italy: IEEE: 7609-7616 [DOI: 10.1109/icpr48806.2021.9412941]
- Sung F, Yang Y X, Zhang L, Xiang T, Torr P H S and Hospedales T M. 2018. Learning to compare: relation network for few-shot learning//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 1199-1208 [DOI: 10.1109/CVPR.2018.00131]
- Tao P, Feng L, Du Y D, Gong X and Wang J. 2024. Meta-cosine loss for few-shot image classification. *Journal of Image and Graphics*, 29(2): 506-519 (陶鹏, 冯林, 杜彦东, 龚勋, 王俊. 2024. 面向元余弦损失的少样本图像分类. *中国图象图形学报*, 29(2): 506-519) [DOI: 10.11834/jig.230127]
- Tseng H Y, Lee H Y, Huang J B and Yang M H. 2020. Cross-domain few-shot classification via learned feature-wise transformation//*Proceedings of the 8th International Conference on Learning Representations*. Addis Ababa, Ethiopia: ICLR: 1-18
- Van Horn G, Mac Aodha O, Song Y, Cui Y, Sun C and Shepard A. 2018. The iNaturalist species classification and detection dataset//*Proceedings of 2018 CVF IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 8769-8778 [DOI: 10.1109/CVPR.2018.00914]
- Vinyals O, Blundell C, Lillicrap T, Kavukcuoglu K and Wierstra D. 2016. Matching networks for one shot learning//*Proceedings of the 30th International Conference on Neural Information Processing Systems*. Barcelona, Spain: Curran Associates Inc.: 3637-3645
- Volpi R, Namkoong H, Sener O, Duchi J, Murino V and Savarese S. 2018. Generalizing to unseen domains via adversarial data augmentation//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal, USA: Curran Associates Inc.: 5339-5349
- Wah C, Branson S, Welinder P, Perona P and Belongie S. 2011. The Caltech-UCSD Birds-200-2011 Dataset. CNS-TR-2010-001. California Institute of Technology
- Wang H Q and Deng Z H. 2021. Cross-domain few-shot classification via adversarial task augmentation//*Proceedings of the 30th International Joint Conference on Artificial Intelligence*. Montreal, Canada: IJCAI.org: 1075-1081 [DOI: 10.24963/ijcai.2021/149]
- Wang S M, Ma R, Wu T R and Cao Y. 2023. P3DC-shot: prior-driven discrete data calibration for nearest-neighbor few-shot classification. *Image and Vision Computing*, 136: #104736 [DOI: 10.1016/j.imavis.2023.104736]
- Wang Y, Qi L, Shi Y and Gao Y. 2022. Feature-based style randomization for domain generalization. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(8): 5495-5509 [Doi: 10.1109/TCSVT.2022.3152615]
- Zhuo L H, Fu Y Q, Chen J J, Cao Y X and Jiang Y G. 2022. TGDM: Target guided dynamic mixup for cross-domain few-shot learning//*Proceedings of the 30th ACM International Conference on Multimedia*. Lisboa, Portugal: Association for Computing Machinery: 6368-6376 [DOI: 10.1145/3503161.3548052]

作者简介

- 余悦,男,硕士研究生,主要研究方向为计算机视觉和深度学习。E-mail: yuyue112@stmail.uj.s.edu.cn
- 成科扬,通信作者,男,教授,主要研究方向为计算机视觉和模式识别。E-mail: kycheng@uj.s.edu.cn
- 陈楠,男,硕士研究生,主要研究方向为计算机视觉和深度学习。E-mail: chennan@stmail.uj.s.edu.cn