

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2023)09-2899-14

论文引用格式: Du Y D, Feng L, Tao P, Gong X and Wang J. 2023. Meta-transfer learning in cross-domain image classification with few-shot learning. Journal of Image and Graphics, 28(09):2899-2912(杜彦东, 冯林, 陶鹏, 龚勋, 王俊. 2023. 元迁移学习在少样本跨域图像分类中的研究. 中国图象图形学报, 28(09):2899-2912)[DOI:10.11834/jig.220664]

元迁移学习在少样本跨域图像分类中的研究

杜彦东¹, 冯林^{1*}, 陶鹏¹, 龚勋², 王俊³

1. 四川师范大学计算机科学学院, 成都 610101; 2. 西南交通大学计算机与人工智能学院, 成都 610031;

3. 四川师范大学商学院, 成都 610101

摘要: 目的 现有基于元学习的主流少样本学习方法假设训练任务和测试任务服从相同或相似的分布, 然而在分布差异较大的跨域任务上, 这些方法面临泛化能力弱、分类精度差等挑战。同时, 基于迁移学习的少样本学习方法没有考虑到训练和测试阶段样本类别不一致的情况, 在训练阶段未能留下足够的特征嵌入空间。为了提升模型在有限标注样本困境下的跨域图像分类能力, 提出简洁的元迁移学习(compressed meta transfer learning, CMTL)方法。**方法** 基于元学习, 对目标域中的支持集使用数据增强策略, 构建新的辅助任务微调元训练参数, 促使分类模型更加适用于域差异较大的目标任务。基于迁移学习, 使用自压缩损失函数训练分类模型, 以压缩源域中基类数据所占据的特征嵌入空间, 微调阶段引导与源域分布差异较大的新类数据有更合适的特征表示。最后, 将以上两种策略的分类预测融合视为最终的分类结果。**结果** 使用 mini-ImageNet 作为源域数据集进行训练, 分别在 EuroSAT (European Satellite)、ISIC (International Skin Imaging Collaboration)、CropDiseas (Crop Diseases) 和 Chest-X (Chest X-Ray) 数据集上测试 CMTL 模型的跨域图像分类能力, 在 5-way 1-shot 和 5-way 5-shot 跨域任务中, 准确率分别达到 68.87% 和 87.74%、34.47% 和 49.71%、74.92% 和 93.37%、22.22% 和 25.40%。与当前主流少样本跨域图像分类方法比较, 提出的 CMTL 方法在各个数据集中都有较好的跨域图像分类能力。**结论** 提出的 CMTL 方法综合了迁移学习和元学习方法各自在跨域任务上的优势, 能有效解决少样本跨域图像分类问题。

关键词: 图像分类; 少样本跨域; 元学习; 迁移学习; 少样本学习(FSL)

Meta-transfer learning in cross-domain image classification with few-shot learning

Du Yandong¹, Feng Lin^{1*}, Tao Peng¹, Gong Xun², Wang Jun³

1. School of Computer Science, Sichuan Normal University, Chengdu 610101, China; 2. School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 610031, China; 3. School of Business, Sichuan Normal University, Chengdu 610101, China

Abstract: Objective Few-shot learning image classification aims to recognize images with limited labeled samples. At present, few-shot learning methods are roughly divided into three categories: gradient optimization, metric learning, and transfer learning. Gradient optimization methods usually consist of two loop stages. In the inner loop stage, the base model quickly adapts to new tasks with limited labeled samples; in the outer loop stage, the meta-model learns cross-task knowl-

收稿日期: 2022-07-07; 修回日期: 2022-10-14; 预印本日期: 2022-10-21

* 通信作者: 冯林 fenglin@sicnu.edu.cn

基金项目: 国家自然科学基金项目(61876158, 71971151); 四川省自然科学基金项目(2023NSFSC0210); 四川省重点研发项目(2023YFS0202, 2023YFG0267)

Supported by: National Natural Science Foundation of China (61876158, 71971151); Natural Science Foundation of Sichuan Province, China (2023NSFSC0210); Sichuan Science and Technology Program (2023YFS0202, 2023YFG0267)

edge for good generalization performance. The metric learning method first maps the samples to the high-dimensional embedding space. Then, the correct classification of the unknown samples is completed according to the similarity measure. The transfer learning method first pretrains a high-quality feature extractor with a large amount of annotated data and migrates to fine-tune the classifier. Thus, the model is suitable for the current task. The existing few-shot learning methods based on meta-learning assume that the training and test tasks are the same or have a similar distribution. However, these methods face cross-domain classification challenges, such as weak generalization ability and poor classification accuracy. In the training and testing stages, the few-shot learning method based on transfer learning does not consider the inconsistency of sample categories. It also fails to leave enough feature embedding space for new category samples. On the basis of the idea of integrating transfer learning and meta-learning, we propose a model of compressed meta-transfer learning (CMTL) to improve the cross-domain ability of few-shot learning. **Method** The method is mainly composed of two aspects. On the one hand, for meta-learning, the prior knowledge generated by meta-training is used to complete the classification of the target task. When the source and target tasks have different data distributions, the base class data used for training comes from the source domain natural dataset mini-ImageNet, whereas the novel class data used for testing comes from the target domain medical dataset Chest-X. The meta-knowledge acquired by the source task training cannot be quickly generalized to the target task because of its lack of universality, which further leads to poor cross-domain classification effects. In this study, new auxiliary tasks with strategies, such as random cropping and gamma transformation, were constructed on the support set of the target domain during the testing process. These auxiliary tasks fine-tune the meta-training parameters to improve task adaptability. On the other hand, for transfer learning, the sample categories are consistent during training and testing. Thus, the feature embedding space available for the novel class samples is small if the deep learning model is optimized with the traditional softmax loss function. This scenario further leads to the unsatisfactory feature extraction ability of the model and poor classification accuracy. Given the above problems, this study proposes using a self-compression loss function in the pretraining stage. This loss function adjusts the distribution position between prototypes of the base classes to make the base class samples concentrated in the embedding space and reserve part of the embedding space for the novel class. In the fine-tuning stage, the novel class with large domain distribution is guided to obtain expressive features. The existing studies on cross-domain few-shot learning show that meta-learning methods perform well when the data distributions of the target and source tasks are similar. Conversely, transfer learning methods perform effectively. The ensemble of prediction scores of the above two strategies is regarded as the final classification result to take full advantage of the above two methods. **Result** This study compares the proposed model with several state-of-the-art cross-domain few-shot image classification models, such as graph convolutional network (GCN), adversarial task augmentation (ATA), self-training to adapt representations to unseen problems (STARTUP), and other classic methods. Compared with the current state-of-the-art cross-domain few-shot methods, CMTL has advantages, as shown in the experimental results. In the testing phase, extensive experiments are performed on the 5-way 1-shot and 5-way 5-shot settings to complete the validation of the model and ensure a fair comparison with advanced methods. The experimental results show that on the 5-way 1-shot and 5-way 5-shot settings, the mini-ImageNet is used as the source domain dataset for training. Moreover, the effectiveness of the CMTL on the EuroSAT, ISIC, CropDisease, and Chest-X datasets is tested, and the accuracy rates reach 68.87%/87.74%, 34.47%/49.71%, 74.92%/93.37%, and 22.22%/25.40%, respectively. Compared with meta-learning and transfer learning models, our model achieves competitive results on the 5-way 1-shot and 5-way 5-shot settings on all cross-domain tasks. **Conclusion** This study proposes a cross-domain few-shot image classification model based on meta-transfer learning. The proposed model improves the generalization ability of few-shot learning. The experimental results show that the CMTL proposed in this study combines the advantages of meta-learning and transfer learning methods. It also has significant effects on cross-domain few-shot tasks.

Key words: image classification; few-shot cross-domain; meta learning; transfer learning; few-shot learning (FSL)

0 引言

得益于强力的计算设备、丰富的数据及先进的模型与算法,深度学习在图像识别(He 等,2016;陈硕 等,2021)、目标跟踪(Bertinetto 等,2016;王蒙蒙 等,2022)、自然语言处理(Ren 和 Lu,2022)等领域取得了极大成功,同时也推动了人工智能的飞速发展。

深度学习模型要想获得最优效果,需要在大规模标注数据中进行训练。然而,在军事、医疗、金融等一些特定领域,由于涉及数据隐私和安全性等问题,数据获取困难且标注代价高昂。与机器相反,人类拥有从有限样本中快速学习的能力。受人类学习方式的启发,人们提出了少样本学习(few-shot learning, FSL)的概念,目的是使机器也能像人类一样,仅靠少量标注的样本通过一两次的示范,就可以学会一类问题的解决。

元学习是少样本学习的重要应用场景,目标是使机器学会学习。元学习通过多任务的学习范式积累多任务的“共性知识”,利用这些“共性知识”指导模型在新任务中快速学习(赵凯琳 等,2021)。同样,迁移学习也是解决少样本学习问题的另一种经典方法,基于迁移学习的方法旨在利用大量标注数据集进行预训练学习知识,并将这些知识迁移到监督信息有限的目标任务中以完成对少量标注样本的学习(张玲玲 等,2021)。

现有的少样本学习方法只考虑到训练阶段的源任务和测试阶段的目标任务服从相同或相似分布的情况,当源域任务与目标域任务数据分布相差较大时,传统的少样本图像分类模型在目标任务中表现较差。图 1 为少样本跨域示例,在图 1(a)—(d)中,用于训练的基类来自源域自然数据集 mini-ImageNet,而测试的新类数据来自目标域医疗数据集 Chest-X。为解决上述问题,少样本跨域学习引起了研究者的关注(Guo 等,2020)。少样本跨域学习旨在利用源域中大量标注样本训练模型,使模型能够快速、准确地泛化到与源域数据分布差异较大的目标域任务中。

对于少样本跨域学习的研究,起源于域适应学习,目的是将源域数据中的知识迁移到具有不同数据分布的目标域中。为解决域适应问题,研究者探

索出基于域差异的方法(Long 等,2017;Sun 等,2016;Kang 等,2019;Kumar 等,2018),旨在对齐源域和目标域之间的边缘分布。基于对抗的方法(Tzeng 等,2017;Ganin 等,2016)主要思想是依靠生成器、鉴别器相互作用促进域相关特征的学习。基于重构的方法(Zhu 等,2017;Hoffman 等,2018)通常使用 encoder-decoder 模型或生成对抗网络来重新构建目标域中的数据。但上述关于域适应的方法均只适用于目标域与源域类别有交集的情况(如图 1(a)—(c))。而现实应用场景中,广泛存在源域和目标域类别交集为空,且目标域中可用样本数量少的跨域学习问题,如罕见的皮肤病病理图像、失事飞机残骸图像、卫星遥感图像,而利用其他带标注数据充足的域信息(如自然数据集 ImageNet)来解决这些域上的子任务是一种很有效的解决方法。因此,如何提升少样本跨域图像分类性能具有重大的理论研究和应用价值。

已有的少样本跨域相关研究表明,当目标域任务与源域任务的数据分布相差较小时,元学习方法表现更好。反之,迁移学习方法表现更好(Chen 等,2019)。因此,融合迁移学习和元学习的思想,本文提出一种简洁的元迁移学习模型(compressed meta transfer learning, CMTL),以解决少样本学习中源域与目标域之间分布差异大的跨域分类问题。

本文主要贡献如下:1)在模型预训练和微调阶段,使用 Wang 等人(2020)提出的自压缩损失函数

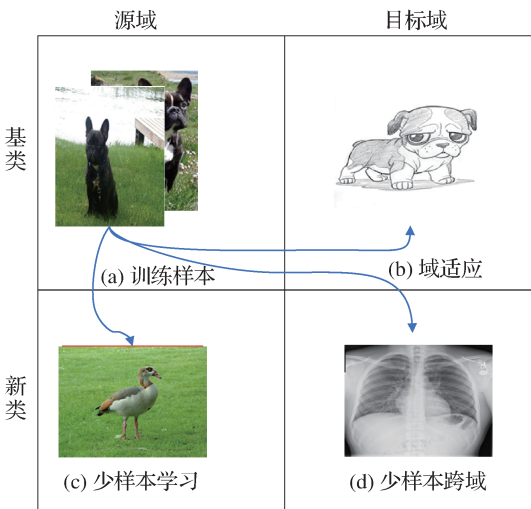


图 1 少样本跨域示例

Fig. 1 Cross-domain example with few-shot
((a) training sample; (b) domain adaptation;
(c) few-shot learning; (d) cross-domain few-shot learning)

(self-compression softmax loss, SSL),通过最小化自压缩损失函数,调整基类数据原型之间的分布位置,促使基类样本在嵌入空间中更为集中,为新类样本预留部分嵌入空间,有效提升了模型的特征提取能力。2)提出一种数据增强策略,利用现有目标任务中有限的支持集样本构建新的辅助任务,微调模型参数,以获取适应当前测试任务的网络参数,解决当源任务与目标任务数据分布差异较大时,“元知识”的普适性不佳、泛化能力弱的问题。3)融合迁移学习和元学习策略提升跨域分类精度。以 mini-ImageNet 作为源域数据集训练,并在 4 个少样本跨域学习标准评估数据集 EuroSAT (European Satellite)、ISIC (International Skin Imaging Collaboration)、CropDiseas (Crop Diseases) 和 Chest-X (Chest X-Ray) 上验证了 CMTL 方法的有效性。

1 相关研究

1.1 少样本学习

少样本学习方法主要分为基于梯度优化的方法、基于度量学习的方法和基于迁移学习的方法 3 类。

1.1.1 基于梯度优化的方法

基于梯度优化的方法通常包括内外两个循环阶段,内循环阶段基模型学习器快速适应于只有少量样本的新任务,外循环阶段元模型学习器学习跨任务的知识以得到好的泛化性能。Finn 等人(2017)提出的模型无关元学习方法,通过在多个任务中学习,以找到最适合模型的一组初始化参数,使模型在新的任务上能够快速泛化。Nichol 等人(2018)提出 Reptile 方法,使用一阶梯度优化,从而简化模型无关元学习方法的微分计算。Ravi 和 Larochelle(2017)提出基于长短期记忆网络的元学习器模型,通过学习模型参数的更新规则来调整不同训练任务之间的差异,促使模型拥有更好的性能。

1.1.2 基于度量的方法

基于度量的方法通常分为基于欧氏空间度量和非欧氏空间度量两种。其中,基于欧氏空间经典的度量方法有如下几种:由 Vinyals 等人(2016)提出的匹配网络引入注意力机制和记忆机制,将样本映射到高维空间,使用余弦相似度衡量相似性。Snell 等人(2017)提出的原型网络通过计算待分类样本与类

原型之间的欧氏距离完成分类。Sung 等人(2018)提出的关系网络也是基于欧氏空间的自适应的关系度量模块,将支持集和查询集的特征向量拼接输入可计算关系得分完成分类。余游等人(2019)在关系网络的基础上,利用半监督学习方法生成伪标签的策略参与训练。Garcia 和 Bruna(2018)提出的图神经网络(graph neural network, GNN)应用于少样本学习问题,首次关注非欧氏空间在度量模型中的运用,利用向量拼接嵌入的方式完成图像数据到图结构数据的转换,继而在图神经网络模块中进行距离度量和关系推导。

1.1.3 基于迁移学习的方法

基于迁移学习的方法首先利用大量标注数据的基类预训练一个优质的特征嵌入网络,然后迁移到新类任务中,微调预训练阶段的分类器,使模型适用于当前任务。Chen 等人(2019)提出了 baseline 及其变体 baseline++,多组对照实验发现,在相同的实验参数设定上,baseline 和 baseline++方法明显优于基于元学习的少样本学习方法。Wang 等人(2020)证明了迁移学习中使用自压缩损失函数训练模型能够获得性能较好的特征嵌入网络。Rajasegaran 等人(2021)提出自监督学习结合知识蒸馏最小化原始增强对距离,通过教师学生网络约束模型,提升基于迁移学习方法的少样本分类性能。

1.2 少样本跨域学习

少样本跨域学习包括基于元学习的方法和基于迁移学习的方法。

1.2.1 基于元学习的方法

为解决少样本跨域问题,Yeh 等人(2020)提出在测试任务中构建伪查询集的方法对元学习模型进行微调,提升少样本跨域分类能力。Wang 和 Deng(2021)直接以参与训练的任务为研究对象,利用随机卷积方法构造更具“挑战性”的任务,从而在源域上模拟分布相对复杂的目标域任务。Tseng 等人(2020)引入特征变化层对卷积网络获取的特征进行仿射变换,以提高模型对不同域的适配性。Sun 等人(2021)提出基于层相关性传播的解释方法,能够知悉样本中对分类结果有贡献的区域,适当增大这些区域的权重系数,以提取更好的样本特征。

1.2.2 基于迁移学习的方法

Phoo 和 Hariharan(2021)提出利用源域数据训练获取的教师学生网络对部分目标域无标签数据集

构造伪标签、结合源域数据共同训练教师学生网络,以提升跨域分类精度。Fu等人(2021)通过 mix-up 算法构造辅助数据集,并使用编码器学习与域无关的特征,以指导网络能快速泛化到域跨度较大的目标任务。Li等人(2021)提出将源域、目标域各自特定领域的特征映射到相同的共享空间,从而实现与领域无关的通用特征表示。

综上所述,以上方法只是单独地从元学习、迁移学习层面分析并解决少样本跨域图像分类问题,却没有考虑到在跨域任务上元学习与迁移学习方法各自的优势。现有方法缺乏有效融合元学习、迁移学习策略提升跨域分类精度的模型与方法。所以本文提出了 CMTL 模型,以完成对少样本跨域任务的推理分类。

2 少样本跨域学习定义

目前,少样本跨域学习多是基于元学习架构。架构通常将任务作为训练和测试的基本单元。元学习通过学习大量源域训练任务,使模型获得“元知识”,具有快速适应少样本目标任务的能力,完成对目标任务的分类。其学习过程如下:

1) 训练阶段。定义“元知识”模型的假设空间 $F(\theta)$ 及其损失函数 $L(\theta)$, 随机抽取源训练任务 $T_s \in \mathcal{T}^u$, 利用源域任务支持集 S_s , 生成样本级分类函数 $f_s(\theta) = F(\theta; S_s)$, 并在查询集 Q_s 上计算 f_s 的损失 $l_s(\theta)$ 。计算训练任务集 $\mathcal{T}^u$ 中所有任务的损失, 作为训练任务集上的损失函数 $L(\theta) = \sum_{T \in \mathcal{T}^u} l_T(\theta)$, 并通过梯度下降的方法最小化 $L(\theta)$, 生成任务级分类函数 $F^*(\theta^*)$, 其中 $\theta^* = \arg \min_{\theta} L(\theta)$ 。

2) 测试阶段。在目标域测试任务集中随机抽取一个元测试任务 $T_t \in \mathcal{T}^e$, 利用测试任务的支持集 S_t , 生成样本级分类函数 $f_t^*(\lambda) = F^*(\theta^*; S_t)$, 然后使用测试任务查询集 Q_t 完成对 $f_t^*(\lambda)$ 分类效果的评估。

3 CMTL 模型

3.1 基本思路和总体分析

首先,将源域数据集分别以批次、任务作为训练数据的基本单元进行划分。然后,在预训练阶段,将

源域数据输入特征嵌入网络,先后以元学习和传统的深度学习方式在源域任务上预训练模型;在微调阶段,对目标域中的同一测试任务,分别利用元学习方法、迁移学习方法获取其对应模型的预测分数。最后,将两者模型的预测分数融合。利用迁移学习和元学习方法各自在跨域任务上的优势,完成少样本跨域任务的推理分类。

基于上述思路,CMTL 的总体结构如图2所示。主要包括特征提取模块、BL-SSL (baseline-self-compression softmax loss) 模块、度量模块和分类策略融合模块。

3.1.1 特征提取模块

在少样本跨域学习中常采用卷积神经网络(convolutional neural network, CNN)中的残差网络 ResNet10 (residual network10) 和 ResNet18 (residual network18) 作为特征提取网络。对于图像样本 x_i , 经特征提取网络后获得样本特征 $f_{\theta}(x_i) \in \mathbf{R}^{C \times H \times W}$, 其中, C 、 H 和 W 分别代表图像通道数、样本特征高度和宽度。本文使用的 ResNet10 如图3所示,由初始的1个64通道的 7×7 卷积层、批归一化层、 3×3 最大池化层和之后4个卷积块组成。其中,第1个卷积块包含2个 3×3 卷积层、2个批归一化层和2个 ReLU (rectified linear unit) 激活层,后续3个卷积块中每个卷积块中包含2个 3×3 卷积层,2个批归一化层,2个 ReLU 激活层,2个下采样层。每个卷积块之间的差别在于卷积层的通道数不同,依次为64、128、256、520。

3.1.2 BL-SSL 模块

由于传统监督学习没有考虑到少样本学习中训练测试阶段样本类别不一致的问题,未能为新类样本留下足够的特征嵌入空间。本文使用 SSL 损失函数代替传统的 SL (softmax loss) 损失函数,能够保证在基类可分的情况下,尽可能拉近基类之间的距离,为新类样本预留更多的嵌入空间。受此启发,为提升迁移学习方法跨域分类性能,本文将 SSL 损失函数应用于迁移学习的训练测试过程。

在训练阶段,SSL 损失函数在保证基类可分、拉近类原型之间距离的双重优化目标下,模型尽可能压缩来自源域的基类数据所占据的嵌入空间,便于测试阶段对来自目标域的新类数据的特征提取。

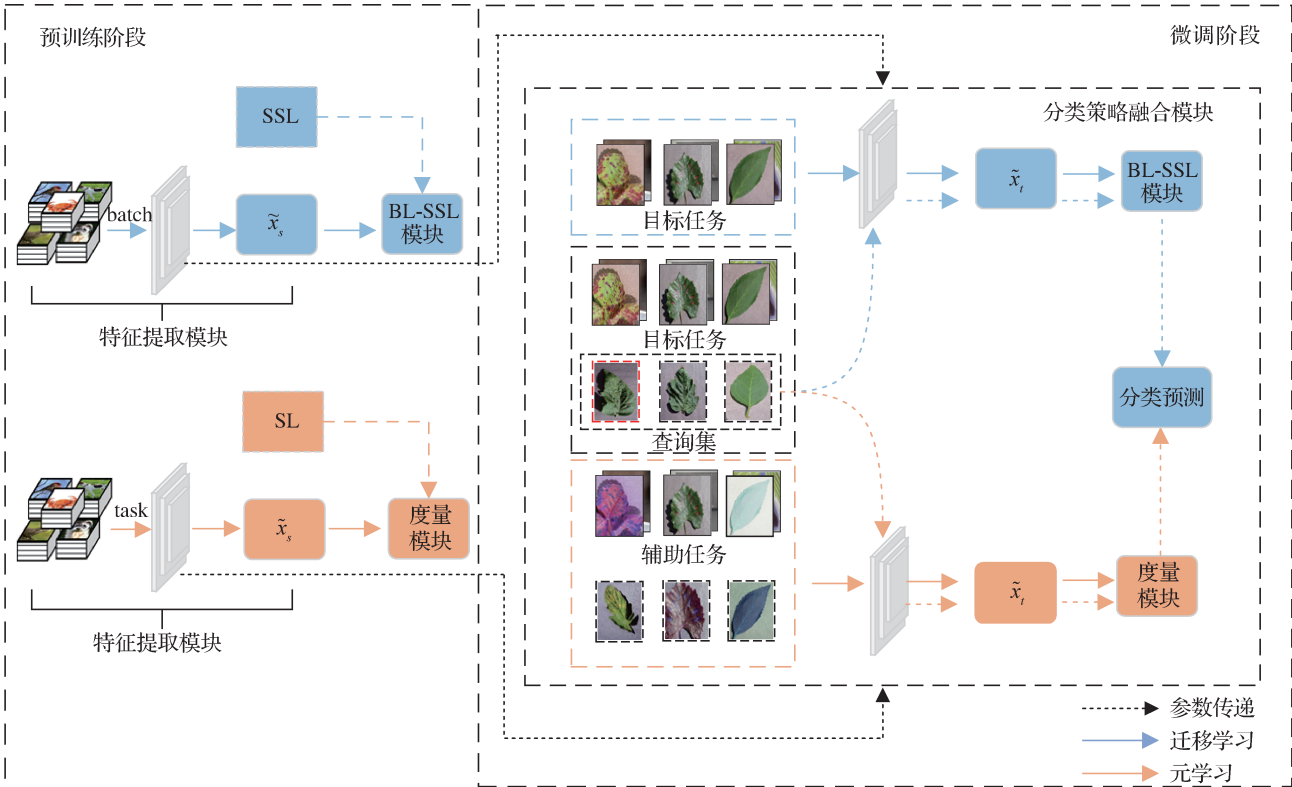


图2 CMTL模型

Fig. 2 CMTL model

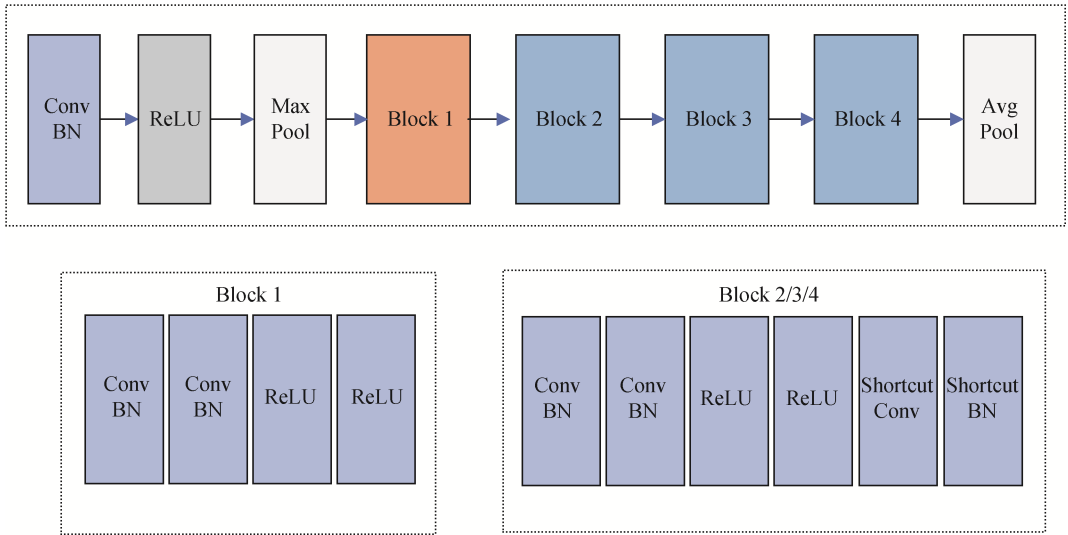


图3 ResNet10网络结构

Fig. 3 ResNet10 network backbone

BL-SSL模块中,通过计算样本经特征嵌入网络获取的高维特征向量与基类类原型向量在嵌入空间中的余弦相似性,调整基类类原型在嵌入空间中的分布位置。

调整基类类原型的过程如图4所示,根据向量加法的平行四边形法则,完成类原型向量 w_c 与 w_i 的相加操作,以遍历调整非 c 类样本的类原型在嵌入

空间中的分布位置。调整过后的非 c 类样本原型向量 $\tilde{w}_i$ 可表示为

$$\tilde{w}_i = \left| \cos \theta(w_i, f_\gamma(x_c^i)) - \cos \theta(w_c, f_\gamma(x_c^i)) \right| \times w_c + w_i \quad (1)$$

式中, γ 为特征嵌入网络参数, $|\cdot|$ 代表绝对值运算, x_c^i 代表 c 样本, $f_\gamma(x_c^i)$ 为样本 x_c^i 的特征映射向量, w_i

代表非 c 类样本的类原型向量, w_c 代表当前样本对应的 c 类原型向量。

利用调整后的类原型向量与样本特征向量计算当前样本属于 i 类的概率, 记为 $p(\hat{y}_j = i | x^j)$, 具体为

$$p(\hat{y}_j = i | x^j) = \frac{\exp(\cos \theta(\tilde{w}_i, f_\gamma(x^j)))}{\sum_i \exp(\cos \theta(\tilde{w}_i, f_\gamma(x^j)))} \quad (2)$$

式中, γ 为特征嵌入网络参数, C 为基类的样本类别总数, $\hat{y}_j$ 为当前样本 x^j 的预测值, $\tilde{w}_i$ 为调整过后的类原型向量。

通过计算当前样本属于 c 类的概率, SSL 损失函数可表示为

$$L_{ssl}(\gamma) = \sum_{c=1}^C \sum_{j=1}^{N_c} -\log(p(\hat{y}_j = c | x^j)) \quad (3)$$

式中, C 为基类的样本类别总数, γ 为特征嵌入网络参数, N_c 为第 c 类中的样本数量, $\hat{y}_j$ 为当前样本 x^j 的预测值。

通过不断压缩基类类原型之间的距离, 促使嵌入空间中基类数据的特征分布更为集中。通过最小化式(3)损失函数, 采用随机梯度下降(stochastic gradient descent, SGD)的方法反向传播更新参数。

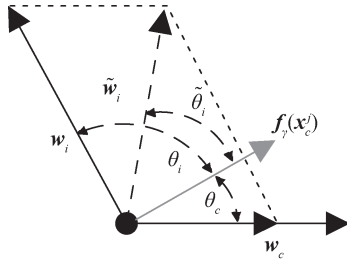


图4 类原型调整过程

Fig. 4 Adjust the class prototype

3.1.3 度量模块

度量学习中, 根据度量方式的不同, 主要以欧氏空间、马氏距离、余弦相似度和非欧氏空间的图神经网络作为度量策略。为简化实验, 本文从以原型网络为代表的欧氏空间度量和图神经网络为代表的非欧氏空间度量的两种网络模型出发, 仅探索两种度量方式对少样本跨域模型的影响。

基于欧氏空间度量的原型网络假设每个类别在向量空间都有一个类原型, 通过计算待分类样本与每个类的类原型之间的欧氏空间, 根据距离衡量待分类样本与各类的类原型之间的相似性, 完成对待

分类样本的正确分类。

由于基于欧氏空间的少样本度量学习方法不能有效捕捉到样本特征和样本标签之间的联系, 严重限制了少样本跨域学习模型的表达能力。针对这一问题, Garcia 和 Bruna(2018)提出利用图神经网络来构建少样本分类模型, 结构如图5所示。GNN由两阶段构成, 第1阶段是图像数据场到图结构数据场的转换, 为利用图神经网络创造前提; 第2阶段是将图结构数据输入到以基于空间的图卷积网络(graph convolutional network, GCN)为基础架构的信息传递网络训练, 以指导生成最优模型完成对未知类别图像进行推理判断, 进而达到分类的目的。

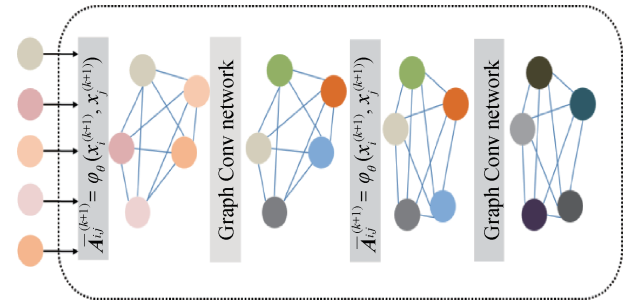


图5 图神经网络

Fig. 5 Graph neural network

3.1.4 分类策略融合模块

元学习通常将任务作为训练和测试的基本单元。为适应元学习的学习过程, 使模型适用于当前任务, 本文对测试任务中的支持集进行颜色变换、随机裁剪等一系列数据增强操作, 如图6所示。

经过数据增强后获取的辅助任务记为 T_{aux} , 具体为

$$T_{aux} = [x_1^s, \dots, x_{C \times K}^s, y_{C \times K}^s, x_{aux1}^q, y_1^q, \dots, x_{auxn}^q, y_{auxn}^q] \quad (4)$$

式中, $x_{C \times K}^s, y_{C \times K}^s$ 为辅助任务中的支持集样本及其对应标签, x_{auxn}^q, y_{auxn}^q 代表支持集样本经图像处理操作获取到的查询集样本及其对应的标签。

辅助任务生成的流程如算法1所示。

算法1: 辅助任务 T_{aux} 的生成。

输入: 任务中的支持集 S_i , 生成的样本数量 num , 超参数 λ_1 和 λ_2 。

输出: 新的辅助任务

For x_i in S_i :

For i range($num/5$):

$\lambda = \text{random.random}()$;



图6 数据增强

Fig. 6 Data augmentation ((a) original image; (b) gamma transform; (c) random cutting; (d) color change)

if ($\lambda < \lambda_1$):

x_i 颜色变换;

else if ($\lambda < \lambda_2$):

x_i 随机裁剪;

else:

x_i 伽马变换;

End;

End;

返回辅助任务 T_{aux} 。

利用生成的辅助任务 T_{aux} 进行微调,使得微调后的度量模块参数更加适应于当前任务。

$$\theta_m = \theta_m - \alpha \nabla_{\theta} L_m(T_{aux}; \theta_m) \quad (5)$$

式中, θ_m 代表训练阶段度量模块的网络参数, α 为训练期间的学习率, $L_m(T_{aux}; \theta_m)$ 代表度量模块在辅助任务上的交叉熵损失函数。

在测试阶段,将目标域任务中有限的支持集通过上述算法构建辅助任务,以元训练方式继续微调度量网络 Metric。其中,本文选用基于非欧氏空间度量的图神经网络和基于欧氏空间的原型网络分别作为 Metric 网络,最后使用微调过的度量网络 Metric 对查询级推理获取其对应预测分数。具体为

$$S_m = \text{Metric}(Q_t; \theta_m) \quad (6)$$

在迁移学习中,为简化运算,对任务中的样本不做任何数据增强操作。由于预训练阶段,使用自压缩损失函数,源域基类样本占据的空间得以压缩,此时目标域新类样本特征嵌入空间相对充裕。所以,固定 BL-SSL 模块中预训练阶段的特征提取网络 f_{θ} , 仅利用同一目标域任务中的支持集微调线性分类器 SSL, 以获取适应当前测试任务的分类器参数 θ_{ssl} , 最后通过分类器 SSL 对查询级分类并获取其对应预测分数,具体为

$$S_b = \text{SSL}(Q_t; \theta_{ssl}) \quad (7)$$

在测试阶段,将目标域任务中样本分别喂入 BL-SSL 模块和度量模块,得到其对应预测分数 S_b 和 S_m , 将两者预测分数分别通过 softmax 函数进行归一化,确保在预测中两者的模型能被赋予相同的权重,最后将同一目标域测试任务经两种模块推理预测分数融合作为最终预测分数。

3.2 算法流程

根据上述讨论,在预训练阶段将源域数据输入特征嵌入网络,先后以元学习和迁移学习方法预训练模型。在微调阶段,基于预训练模型,联合迁移学习和元学习方法共同对目标域测试任务进行分类预测。微调阶段流程如算法 2 所示。

算法 2: CMTL 学习算法。

输入: θ_m, f_{θ} , 目标任务。

输出: 模型分类结果。

For each episode:

目标任务中的支持集 S_t 、查询集 Q_t ;

For epoch in params. ft_epoch:

固定特征提取网络 f_{θ} ;

微调分类器 SSL;

$$\theta_{ssl} = \theta_{ssl} - \beta \nabla_{\theta_{ssl}} L(f_{\theta}(S_t), y);$$

End;

获取到微调过后的分类器 SSL;

S_t 数据增强, 由式(4)得到 T_{aux} ;

For epoch in params. ft_epoch:

$$\theta_m = \theta_m - \alpha \nabla_{\theta_m} L(T_{aux}; \theta_m);$$

End;

获取到微调过后的度量模型 Metric;

$$S_m = \text{Metric}(Q_t; \theta_m);$$

$$S_b = \text{SSL}(Q_t; \theta_{ssl});$$

End;
返回模型分类结果。

4 实验

4.1 数据集和实验细节

本文在5个经典的少样本跨域图像分类数据集上进行跨域分类测试,验证CMTL模型的有效性。

1) mini-ImageNet。是用于训练的源域数据集,是ImageNet数据集(Russakovsky等,2015)中抽离的子数据集,涵盖动物、植物等生活中常见的100个类别,每个类包含600幅图像,为了满足少样本跨域学习任务的需要,通常将100个类划分为64个训练类、16个验证类和20个测试类,图像规格为 224×224 像素。

2) EuroSAT(Helber等,2019)。是用于测试的目标域数据集,为卫星遥感图像,涵盖高速公路、农田等10个类,共20 000幅图像,图像规格为 224×224 像素。

3) CropDiseas(Mohanty等,2016)。是用于测试的目标域数据集,为农作物病害数据集,包含38个类,共54 300幅图像,图像规格为 224×224 像素。

4) ISIC(Codella等,2019)。是用于测试的目标域数据集,为皮肤病黑色素瘤相关数据集,包含9个类,共23 000幅图像,图像规格为 224×224 像素。

5) Chest-X(Wang等,2017)。是用于测试的目标域数据集,为胸部X光图像,包含14个类,共112 120幅图像,图像规格为 224×224 像素。

在实验中,元学习方法的训练过程采用Adam优化器,学习率 $\alpha = 0.001$;在迁移学习BL-SSL模块的训练、微调过程采用SGD优化器,学习率 $\beta = 0.01$ 。本文采用少样本跨域标准评估5-way 1-shot及5-way 5-shot,同时从每一类采样15幅图像样本作为查询集。在训练阶段,元学习以任务为基本单元参与训练,从源域数据集中随机采样100个任务作为一个epoch,共训练400次。迁移学习以监督学习的方式参与训练,从源域数据集中每一次随机抽取64幅图像作为一个epoch,共训练400次。在测试阶段,随机从目标域数据中抽取2 000个任务参与测试,其中,在BL-SSL模块的微调阶段,使用任务中的支持集微调分类器,微调次数为100,每次微调图像样本批次大小为4,而在元微调阶段利用任务中

的支持集生成辅助任务,微调模型的过程中,超参数 λ_1 为0.5, λ_2 为0.8。训练次数为20,生成新的样本数量为25。最终,取全部测试任务的平均Top-1准确率作为最终准确率。

本文采用的硬件环境为NVIDIA RTX 3090 GPU平台;操作系统为ubuntu 20.0.3、Pytorch框架。

4.2 实验结果与分析

为测试CMTL方法的性能,选择mini-ImageNet作为源域,分别以EuroSAT、CropDiseas、ISIC和Chest-X作为目标域完成跨域分类测试。其中,选择原型网络作为度量模块记为CMTL(PRO),选择图神经网络作为度量模块记为CMTL(GNN)。实验结果如表1所示。

4.2.1 定量评估

从表1可以看出,CMTL方法在5-way 1-shot和5-way 5-shot跨域图像分类任务设定上都有较好的表现,并且与现阶段经典的少样本跨域图像分类模型比较,本文提出的CMTL方法更具有优势。其中,基于5-way 1-shot的任务设定,与最新的少样本跨域图像分类GNN+ATA方法相比较,本文提出的CMTL方法在EuroSAT、ISIC、CropDiseas和Chest-X数据集上分别提升了7.52%、1.26%、7.45%和0.12%。相较于最新的迁移学习STARTUP方法,除了在Chest-X数据集上表现稍落后前者,CMTL方法分别在EuroSAT、ISIC和CropDiseas数据集上提升了5.97%、2.13%和1.62%。而基于5-way 5-shot的任务设定,与少样本跨域GNN+ATA方法相比,本文方法在EuroSAT、ISIC、CropDiseas和Chest-X数据集上分别提升了4%、4.8%、2.78%和1.08%。与最新的迁移学习STARTUP方法比较,除了在Chest-X数据集上表现稍落后前者,CMTL方法在EuroSAT、ISIC和CropDiseas数据集上分别提升了5.93%、3.23%和1.68%。实验结果表明提出的CMTL方法的有效性。

4.2.2 定性评估

从表1的跨域分类结果可以看出,提出的CMTL方法能很好地解决少样本跨域图像分类问题,其中BL-SSL模块中的特征嵌入网络能充分提取更为丰富的特征信息。对于元学习,提出数据增强的策略生成辅助任务,能够进一步微调模型,以捕获到适合当前任务的合适参数。将二者对于相同任务的预测

表1 与不同跨域分类算法对比实验结果
Table 1 Compare with the different cross-domain classification algorithms

方法	数据集							
	EuroSAT		ISIC		CropDiseas		Chest-X	
	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
Mat-net(Vinyals等,2016)	60.26±0.4	75.82±0.3	32.49±0.3	44.02±0.3	62.68±0.4	84.05±0.3	22.06±0.2	24.84±0.2
Pro-net(Snell等,2017)	57.62±0.4	80.14±0.3	31.05±0.3	47.50±0.3	60.44±0.4	87.72±0.3	21.92±0.2	25.92±0.2
Rel-net(Sung等,2018)	49.08±0.4	65.56±0.4	30.53±0.3	38.60±0.3	53.58±0.4	72.86±0.4	21.95±0.2	24.07±0.2
GNN(Garcia和Bruna,2018)	54.61±0.5	78.69±0.4	30.14±0.3	42.54±0.4	59.19±0.5	83.12±0.4	21.94±0.2	23.87±0.2
GNN+FT(Tseng等,2020)	55.53±0.5	78.02±0.4	30.22±0.3	40.87±0.4	60.74±0.5	87.07±0.4	22.00±0.2	24.28±0.2
GNN+LRP(Sun等,2021)	54.99±0.5	77.14±0.4	30.94±0.3	44.14±0.4	59.23±0.5	86.15±0.4	22.11±0.2	24.53±0.2
GNN+ATA(Wang和Deng,2021)	61.35±0.5	83.75±0.4	33.21±0.4	44.91±0.4	67.47±0.5	90.59±0.3	22.10±0.2	24.32±0.4
Baseline(Chen等,2019)	57.45±0.4	76.11±0.3	32.06±0.2	43.48±0.3	62.60±0.4	84.01±0.3	21.87±0.2	24.77±0.2
STARTUP(Phoo和Hariharan,2021)	62.90±0.8	81.81±0.6	32.34±0.6	46.48±0.3	73.30±0.8	91.69±0.4	22.87±0.4	26.68±0.4
CMTL(PRO)	64.45±0.4	83.49±0.3	33.31±0.3	50.21±0.3	71.77±0.4	91.23±0.2	22.30±0.2	26.28±0.2
CMTL(GNN)	68.87±0.5	87.74±0.3	34.47±0.3	49.71±0.3	74.92±0.5	93.37±0.2	22.22±0.2	25.40±0.2

注:加粗字体表示各列最优结果;Mat-net: matching network;Pro-net: prototypical network;Rel-net: relation network;GNN: graph neural network;GNN+FT: graph neural network feature-wise Transformation;GNN+LRP: graph neural network layer-wise relevance propagation;GNN+ATA: graph neural network adversarial task augmentation;STARTUP: self training to adapt representations to unseen problems;CMTL(PRO): compressed meta transfer learning(prototypical network);CMTL(GNN): compressed meta transfer learning(graph neural network)。

分数融合,可以使模型结合迁移学习和元学习方法各自的优点,促使少样本跨域图像分类更加准确。另外,除了在Chest-X上的跨域分类表现欠佳,CMTL(GNN)模型总体的跨域分类性能高于CMTL(PRO)模型。这是由于在跨域图像分类任务中,非欧氏空间的度量方式有效利用了样本特征与样本标签之间的联系,而欧氏空间度量的方法仅关注查询集样本与支持集样本的相似性,未考虑标签间的联系。由此可以看出,基于非欧氏空间的度量方式更适用于少样本跨域图像分类任务。

4.3 消融实验

为了研究CMTL方法中多种因素对于少样本跨域分类准确率的影响,消融实验采用控制变量法的研究方法。

4.3.1 SSL使用阶段对跨域分类准确率的影响

为了研究BL-SSL模块中SSL使用阶段对于实验分类准确率的影响,分别在5-way 1-shot和5-way 5-shot任务上,以4个标准数据集上平均跨域分类准确率作为评估标准,进行消融实验,结果如图7所示,

其中SL代表使用softmax loss,SSL代表使用自压缩损失函数。SL+SSL代表预训练阶段使用softmax loss,微调阶段使用自压缩损失函数,在5-way 1-shot和5-way 5-shot任务设定下,在EuroSAT、ISIC、CropDiseas和Chest-X数据集上的平均跨域分类准确率为42.5%和56.3%。实验结果表明,在预训练阶段和微调阶段均使用SSL能达到最好跨域分类准确率。

另外,从EuroSAT中随机选择5个类别的数据特征进行T-SNE(T-distributed stochastic neighbor embedding)可视化,结果如图8所示。T-SNE可视化

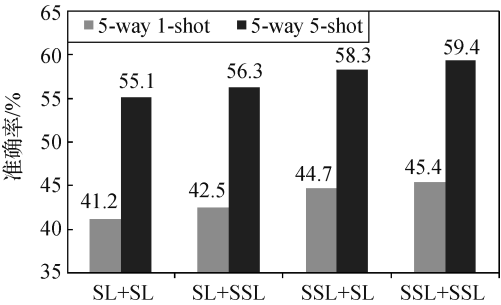


图7 SSL使用阶段平均跨域分类准确率
Fig. 7 Average accuracy at different stages of the SSL

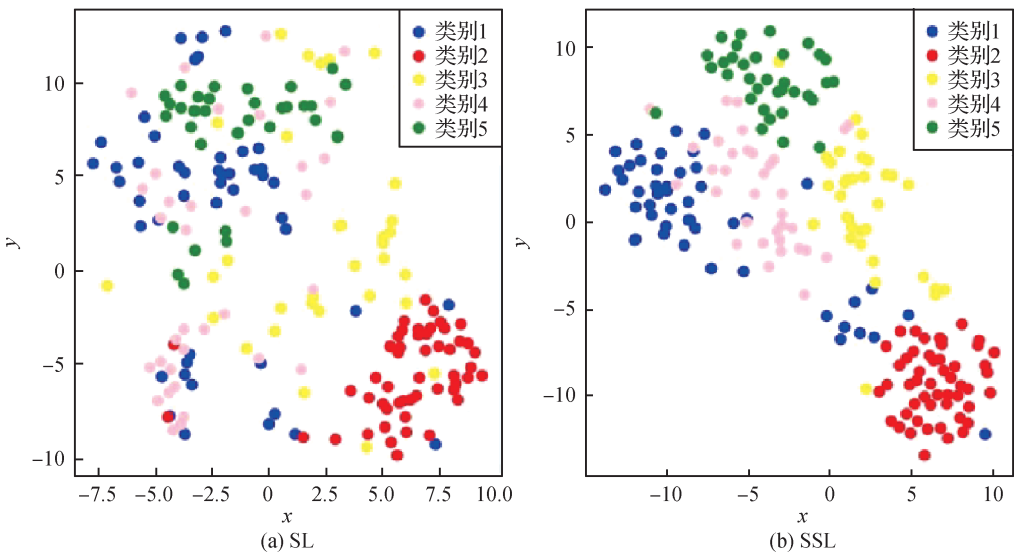


图 8 EuroSAT 数据集的 T-SNE 可视化
Fig. 8 T-SNE plot of EuroSAT((a)SL;(b)SSL)

表明,使用 SSL 后更有利于新类样本的特征提取,提取的新类样本特征在保证类间可分的前提下,拉近了类间的距离,同样证明了基于迁移学习的少样本模型中引入 SSL 损失函数能有效提升模型的跨域图像分类能力。

4.3.2 单个模块对跨域分类准确率的影响

为了证明本文使用分类模块融合的有效性,分别使用图神经网络微调模块(GNN-FT)、BL-SSL 模块以及将两者分类预测结果的融合完成跨域图像分类实验,结果如表 2 所示。可以看出,目标域与源域

跨度较小,目标域数据集为 EuroSAT 和 CropDiseas 时,GNN-FT 模块的跨域分类准确率优于 BL-SSL 模块,反之 BL-SSL 模块的跨域分类准确率优于 GNN-FT 模块。两个模块分类融合的方法在 4 个标准数据集的平均分类准确率在 5-way 1-shot 任务中,相较 BL-SSL 和 GNN-FT 模块分别高出 4.2% 和 1.8%;在 5-way 5-shot 任务中,跨域分类准确率最高,相较 BL-SSL 和 GNN-FT 模块分别高出 4.3% 和 2.1%。实验结果表明,融合两者分类预测结果是提升少样本跨域图像分类准确率的有效方法。

表 2 消融实验:单个模型对于分类准确率的影响
Table 2 Ablation experiment: a single model on experimental classification

模型	数据集							
	EuroSAT		ISIC		CropDiseas		Chest-X	
	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
BL-SSL	61.51±0.4	80.85±0.3	33.59±0.3	46.40±0.3	66.32±0.4	86.87±0.3	22.02±0.2	24.73±0.2
GNN-FT	66.43±0.5	85.04±0.4	32.14±0.3	46.31±0.3	72.48±0.5	92.09±0.2	21.92±0.2	24.56±0.2
CMTL(本文)	68.87±0.5	87.74±0.3	34.47±0.3	49.71±0.5	74.92±0.5	93.37±0.2	22.22±0.2	25.40±0.2

注:加粗字体表示各列最优结果。

4.3.3 辅助任务对跨域分类的影响

为了测试微调过程中产生新的辅助任务是否必要,并进一步探索不同数据增强策略对于跨域分类实验的影响,本文对目标域测试任务利用随机裁剪、颜色变换和伽马变换操作,两两随机组合形成辅助任务。实验中,将图神经网络和 BL-SSL 模块两者分

类结果的融合视为最终的预测分类结果。其中, None 表示不使用辅助任务参与微调,仅利用 BL-SSL 模块和图神经网络推理;CC+RC 表示辅助任务经颜色变换和随机裁剪策略生成;CC+GT 表示辅助任务经颜色变换和伽马变换策略生成;RC+GT 表示辅助任务经随机裁剪和伽马变换策略生成;本文算法 1 表示

辅助任务由本文算法 1 中提出的方法生成。实验结果如表 3 所示,实验结果表明,相较于本文算法 1 提出的辅助任务生成方法,随机裁剪、颜色变换和伽马

变换任意两种数据增强策略组合形成的辅助任务对最终的跨域分类准确率并无显著提升,实验结果进一步凸显了本文方法对于跨域图像分类的有效性。

表 3 消融实验:辅助任务对于分类准确率的影响
Table 3 Ablation experiment: auxiliary tasks on experimental classification

辅助任务	数据集							
	EuroSAT		ISIC		CropDiseas		Chest-X	
	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot
None	60.61±0.51	80.28±0.57	31.45±0.23	45.21±0.31	65.81±0.55	88.68±0.33	21.92±0.21	24.88±0.23
CC+RC	64.60±0.52	85.88±0.43	33.44±0.21	45.08±0.31	67.55±0.49	93.02±0.32	22.00±0.20	25.43±0.23
CC+GT	68.92±0.56	86.28±0.55	32.85±0.23	48.72±0.31	70.11±0.55	91.80±0.33	21.95±0.21	24.90±0.23
RC+GT	67.79±0.42	84.89±0.51	34.10±0.23	48.22±0.32	72.79±0.53	92.76±0.31	21.99±0.20	25.35±0.23
本文算法 1	68.87±0.54	87.74±0.51	34.47±0.25	49.71±0.34	74.92±0.52	93.37±0.37	22.02±0.21	25.40±0.22

注:加粗字体表示各列最优结果。

5 结 论

元学习、迁移学习是解决少样本跨域问题的两项前沿理论和方法,为提升模型在有限标注样本环境下的跨域图像分类能力,本文提出了CMTL方法,并通过大量实验证实了本文方法的有效性。

本文的主要工作分为3个方面:1)基于元学习,通过构建辅助任务,用于微调模型参数;2)基于迁移学习,提出的BL-SSL模块更利于目标任务的特征提取;3)融合迁移学习和元学习预测分数的方法能有效处理跨域图像分类任务。

本文提出的CMTL方法虽然能极大缓解少样本跨域图像分类精度差等困境,但对构建辅助任务阶段的探索还不够全面,融合两者模块预测分数的研究也不够深入。未来工作将会继续围绕如何生成更为合适的辅助任务、如何有效融合两者模块预测分数这两大重点展开研究,促使少样本跨域图像分类模型具有更好的泛化能力。

参考文献(References)

Bertinetto L, Henriques J F, Valmadre J, Torr P H S and Vedaldi A. 2016. Learning feed-forward one-shot learners//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates Inc.: 523-531
Chen S, Zhong H C, Li Y Z, Wang S Z and Yang J G. 2021. Haze

image recognition based on multi-scale feature and multi-adversarial networks. Journal of Image and Graphics, 26 (11) : 2680-2690 (陈硕, 钟汇才, 李勇周, 王师峥, 杨建刚. 2021. 基于多尺度特征多对抗网络的雾天图像识别. 中国图象图形学报, 26(11): 2680-2690) [DOI: 10.11834/jig.200491]
Chen W Y, Liu Y C, Kira Z, Wang Y C F and Huang J B. 2019. A closer look at few-shot classification//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: OpenReview.net: 1-17
Codella N, Rotemberg V, Tschandl P, Cele M E, Dusza S, Gutman D, Helba B, Kalloo A, Liopyris K, Marchetti M, Kittler H and Halpern A. 2019. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC) [EB/OL]. [2022-06-15].
<https://arxiv.org/pdf/1902.03368.pdf>
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: PMLR: 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]
Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M and Lempitsky V. 2016. Domain-adversarial training of neural networks. The Journal of Machine Learning Research, 17(1): 2096-2030
Garcia V and Bruna J. 2018. Few-shot learning with graph neural networks//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada: OpenReview.net: 1-13
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]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- 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]
- Hoffman J, Tzeng E, Park T, Zhu J Y, Isola P, Saenko K, Efros A A and Darrell T. 2018. CyCADA: cycle-consistent adversarial domain adaptation//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden: PMLR: 1994-2003
- Kang G L, Jiang L, Yang Y and Hauptmann A G. 2019. Contrastive adaptation network for unsupervised domain adaptation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 4888-4897 [DOI: 10.1109/CVPR.2019.00503]
- Kumar A, Sattigeri P, Wadhawan K, Karlinsky L, Feris R, Freeman W T and Wornell G. 2018. Co-regularized alignment for unsupervised domain adaptation//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montreal, Canada: Curran Associates Inc.: 9367-9378
- Li W H, Liu X L and Bilen H. 2021. 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]
- Long M S, Zhu H, Wang J M and Jordan M I. 2017. Deep transfer learning with joint adaptation networks//Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia: JMLR.org: 2208-2217
- 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]
- Nichol A, Achiam J and Schulman J. 2018. On first-order meta-learning algorithms [EB/OL]. [2022-06-15]. <https://arxiv.org/pdf/1803.02999v3.pdf>
- Phoo C P and Hariharan B. 2021. Self-training for few-shot transfer across extreme task differences//Proceedings of the 9th International Conference on Learning Representations. Vienna, Austria: OpenReview.net: 1-19
- Rajasegaran J, Khan S, Hayat M, Khan F S and Shah M. 2021. Self-supervised knowledge distillation for few-shot learning//Proceedings of the 32nd British Machine Vision Conference. [s.l.]: BMVA Press: #179
- Ravi S and Larochelle H. 2017. Optimization as a model for few-shot learning//Proceedings of the 5th International Conference on Learning Representations. Toulon, France: OpenReview.net: 1-11
- Ren H and Lu H. 2022. Compositional coding capsule network with k-means routing for text classification. *Pattern Recognition Letters*, 160: 1-8 [DOI: 10.1016/j.patrec.2022.05.028]
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S A, Huang Z H, Karpathy A, Khosla A, Bernstein M, Berg A C and Li F F. 2015. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3): 211-252 [DOI: 10.1007/s11263-015-0816-y]
- Snell J, Swersky K and Zemel R. 2017. Prototypical networks for few-shot learning//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 4080-4090
- Sun B C, Feng J S and Saenko K. 2016. Return of frustratingly easy domain adaptation//Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix, USA: AAAI: 2058-2065 [DOI: 10.1609/aaai.v30i1.10306]
- 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]
- 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: OpenReview.net: 1-18
- Tzeng E, Hoffman J, Saenko K and Darrell T. 2017. Adversarial discriminative domain adaptation//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 2962-2971 [DOI: 10.1109/CVPR.2017.316]
- 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
- Wang C, Liu R Z, Ye H J and Yu Y. 2020. Novelty-prepared few-shot classification [EB/OL]. [2022-06-15]. <https://arxiv.org/pdf/2003.00497v1.pdf>
- 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 M M, Yang X Q and Liu Y. 2022. A spatio-temporal encoded network for single object tracking. *Journal of Image and Graphics*, 27(9): 2733-2748 (王蒙蒙, 杨小倩, 刘勇. 2022. 利用时空特征编码的单目标跟踪网络. *中国图象图形学报*, 27(9): 2733-2748) [DOI: 10.11834/jig.211157]
- Wang X S, Peng Y F, Lu L, Lu Z Y, Bagheri M and Summers RM. 2017. ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA; IEEE: 3462-3471 [DOI: 10.1109/CVPR.2017.369]
- Yeh J F, Lee H Y, Tsai B C, Chen Y R, Huang P C and Hsu W H. 2020. Large margin mechanism and pseudo query set on cross-domain few-shot learning [EB/OL]. [2022-06-15]. <https://arxiv.org/pdf/2005.09218.pdf>
- Yu Y, Feng L, Wang G G and Xu Q F. 2019. A few-shot learning model based on semi-supervised with pseudo label. *Acta Electronica Sinica*, 47(11): 2284-2291 (余游, 冯林, 王格格, 徐其凤. 2019. 一种基于伪标签的半监督少样本学习模型. *电子学报*, 47(11): 2284-2291) [DOI: 10.3969/j. issn. 0372-2112.2019.11.007]
- Zhang L L, Chen Y W, Wu W J, Wei B F, Luo X, Chang X J and Liu J. 2021. Interpretable few-shot learning with contrastive constraint. *Journal of Computer Research and Development*, 58(12): 2573-2584 (张玲玲, 陈一苇, 吴文俊, 魏笔凡, 罗炫, 常晓军, 刘均. 2021. 基于对比约束的可解释小样本学习. *计算机研究与发展*, 58(12): 2573-2584) [DOI: 10.7544/issn1000-1239.2021.20210999]
- Zhao K L, Jin X L and Wang Y Z. 2021. Survey on few-shot learning. *Journal of Software*, 32(2): 349-369 (赵凯琳, 靳小龙, 王元卓. 2021. 小样本学习研究综述. *软件学报*, 32(2): 349-369) [DOI: 10.13328/j.cnki.jos.006138]
- Zhu J Y, Park T, Isola P and Efros A A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy; IEEE: 2242-2251 [DOI: 10.1109/ICCV.2017.244]

作者简介

杜彦东,男,硕士研究生,主要研究方向为机器学习和计算机视觉。E-mail: duane777@foxmail.com

冯林,通信作者,男,教授,主要研究方向为机器学习和数据挖掘。E-mail: fenglin@sicnu.edu.cn

陶鹏,男,硕士研究生,主要研究方向为机器学习和计算机视觉。E-mail: 1421282325@qq.com

龚勋,男,教授,主要研究方向为计算机视觉和人工智能。

E-mail: xgongbuct@163.com

王俊,男,副教授,主要研究方向为应用预测建模和不确定多属性决策理论。E-mail: wangjun1973@tom.com