

## RA-ProtoNet: 基于元学习的小样本遥感场景分类方法

贺琪<sup>1</sup>, 张津源<sup>1</sup>, 黄冬梅<sup>3</sup>, 杜艳玲<sup>1</sup>, 徐慧芳<sup>1,2\*</sup><sup>1</sup>上海海洋大学信息学院, 上海 201306;<sup>2</sup>上海建桥学院信息技术学院, 上海 201306;<sup>3</sup>上海电力大学, 上海 200090

**摘要** 深度学习在解决遥感影像场景分类问题中发挥了重要作用,但在某些特定的遥感场景分类问题中,存在可训练带标签样本严重不足的情况(单类样本数少于10),造成现有的传统深度模型分类效果不理想。针对上述问题,提出一种小样本遥感场景分类方法,并构建一种基于元学习(meta-learning)训练策略的模型 ResNet14-Attention-ProtoNet (RA-ProtoNet)。首先,采用预训练的深度残差网络 ResNet14 作为特征嵌入模块,提取遥感影像深度特征;其次,针对同类样本特征不明显会对类级(class-level)表达造成的干扰问题,在类级表达模块,采用基于双向长短期记忆网络(Bi-LSTM)的注意力机制强化类内样本信息,生成样本的类级特征表达;最后,利用欧氏距离度量待分类样本与类级特征之间的距离,实现分类预测。在 UCMERGED、AID-30 和 NWPU-RESISC45 等 3 个遥感影像数据集上,将所提方法与基于迁移学习和现有元学习方法的遥感场景分类方法进行对比实验,在 5-way 5-shot 条件下,所提方法的整体场景分类精度分别达到 81.30%、83.29% 和 81.22%。实验结果表明,所提方法可以有效挖掘类内样本信息,在极小样本条件下获得更高的遥感影像场景分类精度。

**关键词** 遥感; 图像处理; 遥感影像分类; 小样本学习; 注意力机制

中图分类号 TP751

文献标志码 A

DOI: 10.3788/LOP220432

RA-ProtoNet: Classification Based on Meta-Learning  
for Few-Shot Remote Sensing SceneHe Qi<sup>1</sup>, Zhang Jinyuan<sup>1</sup>, Huang Dongmei<sup>3</sup>, Du Yanling<sup>1</sup>, Xu Huifang<sup>1,2\*</sup><sup>1</sup>College of Information, Shanghai Ocean University, Shanghai 201306, China;<sup>2</sup>College of Information Technology, Shanghai Jian Qiao University, Shanghai 201306, China;<sup>3</sup>Shanghai University of Electric Power, Shanghai 200090, China

**Abstract** Deep learning plays an important role in solving the problem of remote sensing image scene classification. However, in certain remote sensing scene classification problems, samples with labels that can be trained are severely lacking (number of single-class samples less than 10), resulting in unsatisfactory classification using existing traditional depth models. In this paper, to solve these problems, a small-sample-size remote sensing scene classification method is proposed, and a model called ResNet14 Attention-ProtoNet (RA-ProtoNet) based on a meta-learning training strategy is constructed. First, in the feature embedding module, the pre-trained depth residual network, ResNet14, is used to extract the depth features of remote sensing images. Second, in the class-level expression module, the problem that the features of similar samples are unremarkable and interfere in class-level expressions is solved. For this purpose, an attention mechanism based on bidirectional long short-term memory (Bi-LSTM) is used to strengthen the sample information within a class and generate class-level feature expressions of samples. Finally, the Euclidean distance is used to measure the distances between the samples to be classified and the class-level features for classification prediction. On three remote sensing image datasets, including UCMERGED, AID-30 and NWPU-RESISC45, the proposed method is compared with remote sensing scene classification methods based on migration learning and existing meta-learning methods. Under the five-way five-shot condition, the overall scene classification accuracies of the proposed method reach 81.30%, 83.29%,

收稿日期: 2022-01-04; 修回日期: 2022-01-05; 录用日期: 2022-01-28; 网络首发日期: 2022-02-08

基金项目: 国家自然科学基金青年项目(410906179)、上海市科委部分地方高校能力建设项目(20050501900)、上海市教育发展基金项目(AASH2004)

通信作者: \*17069@gench.edu.cn

and 81.22%, respectively. The experimental results show that the proposed method can effectively mine the sample information within a class and obtain higher classification accuracy of remote sensing image scenes under the condition of minimal samples than the other methods.

**Key words** remote sensing; image processing; remote sensing image classification; few shot learning; attentional mechanism

## 1 引言

随着遥感技术的迅猛发展,利用机器学习方法实现影像精准分类作为遥感影像处理的重要分支,被广泛应用于湿地资源监测、城市规划和军事侦察等领域<sup>[1]</sup>。传统的场景分类主要基于人工提取的特征(如影像的光谱、纹理及结构特征等)实现分类任务。常用的人工特征有 Gabor 特征<sup>[2]</sup>、SIFT 特征<sup>[3]</sup>、BOVW 词包模型<sup>[4]</sup>等,而人工提取和构建的特征严重依赖专家经验,需要较高的设计成本<sup>[5]</sup>。基于深度学习的分类算法可以自动提取高维样本特征<sup>[6]</sup>,极大地提高分类性能,现已成为解决场景分类任务的研究热点。而在某些特定场景中,如军事地物侦察和月球探测,标记的样本数量有限,无法满足深度网络模型中大量的参数训练需求,会导致模型性能变差。

随着多种不同新型网络的提出,基于深度学习模型解决小样本分类问题取得一系列研究成果,研究方法主要分为多模态融合、迁移学习和元学习这 3 种。在多模态融合方面:贺琪等<sup>[7]</sup>考虑不同模态特征,抽取高低分辨率样本,提出两种模态高维特征融合策略,并基于融合的高维特征实现小样本遥感影像分类;邱晓华等<sup>[8]</sup>提出一种融合卷积神经网络和朴素贝叶斯决策的分类方法,采用预训练和波段特征融合提升场景分类性能;Rostami 等<sup>[9]</sup>提取相同目标的两种模态数据 electro-optic (EO) 域和 synthetic aperture radar (SAR) 域,融合同目标、同空间分布的特征信息解决小样本问题。在迁移学习方面:Scott 等<sup>[10]</sup>针对小样本问题采用旋转操作实现样本数据增强,并采用迁移学习思想微调预训练的深度模型;Lima 等<sup>[11]</sup>结合原始图像和灰度图像作为模型输入,采用迁移学习实现海洋锋的高精度识别。然而在遥感影像这一特定场景下,航拍时间分辨率低导致可训练样本严重不足(单类样本数少于 10)。同时,面对新类别分类任务时,深度模型往往需要重新启动训练,训练代价较高。小样本学习(few-shot learning)模仿人类学习过程,采用任务级别的元学习训练策略,只需利用少量样本就可以有效完成学习任务,解决小样本下的影像分类和新类别分类问题。近年来,基于元学习方法的小样本遥感影像分类工作也逐步展开:Zhai 等<sup>[12]</sup>引入元学习思想,提出一种基于元学习的终身学习方法,解决小样本遥感场景分类问题;Li 等<sup>[13]</sup>提出一种元学习模型 RS-MetaNet,引入平衡参数缓解模型过拟合,从而解决小样本分类问题;Snell 等<sup>[14]</sup>提出一种基于度量学习的原型网络

ProtoNet,使用类内样本均值表示类原型,通过度量未知样本距离进行分类。尽管现有研究已经取得一定成果,但针对极小样本的遥感影像分类问题,现有方法仍缺乏类内样本特征的判别能力,忽视了样本间特征差异干扰类级特征的问题,均不能取得良好的分类效果。

为解决上述问题,本文提出一种基于元学习的小样本遥感场景分类方法。采用元学习训练策略,针对极小样本条件下遥感影像中样本特征区分不明显的特点,引入残差网络 ResNet14 的卷积部分作为特征嵌入主体,提取影像高维特征;为缓解深度网络过拟合问题,采用迁移学习思想,使用 ImageNet 数据集对特征嵌入模块预训练,提升网络训练效率;针对同类别内样本差别大、对类原型贡献度不一致的问题,设计类级表达模块,使用注意力机制充分融合类内样本特征信息,生成鲁棒的类原型表达;最后在 UCERCEED、AID-30 和 NWPU-RESISC45 这 3 个数据集进行实证化分析,结果表明所提方法提高了小样本遥感场景分类性能,在不同的数据集上具有良好的泛化性。

## 2 方法设计

所提 RA-ProtoNet 模型采用元学习思想进行模型训练,每个遥感影像分类子任务被称为元任务,每个元任务(meta task)包括支持集(support set)和查询集(query set)这 2 部分,分别对应传统模型中的训练集和测试集<sup>[15]</sup>。模型主要结构如图 1 所示,包括 3 个部分:特征嵌入模块(feature embedding module)、类级注意力模块(class-level attention module)和度量模块(metric module)。特征嵌入模块采用基于卷积神经网络的残差网络 ResNet14,由 1 个卷积层、1 个池化层和 3 个残差块组成,实现元任务内支持集和查询集影像高维空间特征的提取;类级注意力模块针对由特征嵌入模块获得的支持集高维特征,采用基于双向长短期记忆网络(Bi-LSTM)的类级注意力模块学习类内样本信息,归纳生成类原型特征 prototype;度量模块使用固定度量函数欧式距离(Euclidean distance)作为度量主体,生成查询集样本特征的从属概率,从而得到分类结果。

### 2.1 特征嵌入模块

在极小样本遥感影像的条件下,某些类别样本之间的特征不明显,类间样本相似度而类内差异性较大<sup>[16]</sup>。良好的特征嵌入,可以有效提取样本的高维特征,是模型进行高效分类的关键,表示如下:

$$\phi_k^i = f(\mathbf{x}_k^i), \quad (1)$$

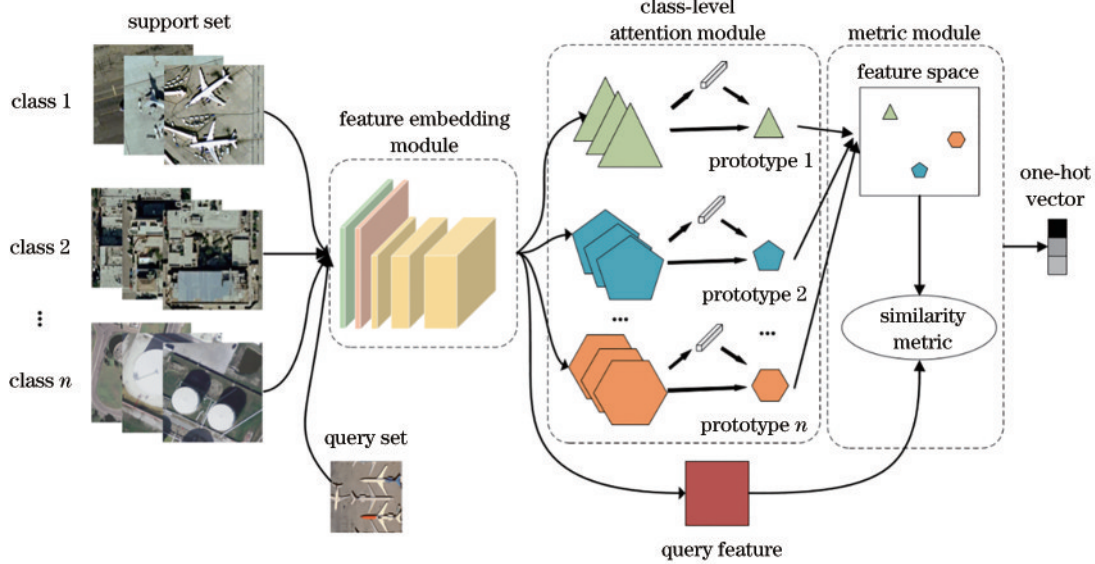


图 1 RA-ProtoNet 模型结构图  
Fig. 1 Model structure diagram of RA-ProtoNet

式中： $\phi_k^c$  是影像样本  $x_k^c$  在高维空间的特征向量。

在传统卷积神经网络中,通过卷积操作学习影像样本的平移不变特征,提升模型特征分类性能,但面对极小样本的遥感影像时,简单地增加卷积层数无法有效解决特征问题。

残差网络作为一种改进卷积神经网络,其内部残差块使用跳跃连接,有效缓解了由网络深度带来的梯度消失问题,同时可以融合遥感影像不同尺度的高维影像特征,完善样本特征信息。而深层的残差网络在小样本条件下容易产生过拟合现象,影响模型分类性能。考虑小样本遥感影像特性,设计多组对比实验,根据实验结果,选择浅层残差网络 ResNet14 作为模型的特征嵌入网络<sup>[17]</sup>,其参数如表 1 所示,其中, S 为卷积层中滑动窗口的滑动步长。为进一步缓解过拟合以提升性能,使用数据集 ImageNet 对特征嵌入模块进行预训练,获取较好的初始化参数,加快网络训练速率。

表 1 ResNet14 特征嵌入模块参数

Table 1 Parameters of ResNet14 feature embedding module

Network layer	Structure	Output
Convolution	7×7, 64, S=2	112×112, 64
Max_pool	3×3, S=2	56×56, 64
Residual_block2_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	56×56, 64
Residual_block3_x	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	28×28, 128
Residual_block4_x	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	14×14, 256

## 2.2 类级注意力模块

如何通过每个类别中极少量的样本生成鲁棒的类级特征是小样本学习的关键。原型网络<sup>[13]</sup>采用类内特

征的均值表示类原型特征,并证明样本在特定空间分布下,样本均值是类内特征间距离的最小值。然而,遥感影像可训练样本较为稀缺,现存样本中各个类别在特征空间内只有极少量的特征映射,样本对应的高维特征不满足特定空间分布。同时,在实际场景中,由于云层遮挡和光线变化等影响,易生成特殊偏差影像,类内影像特征信息差别较大,造成类内不同样本对类原型特征贡献不一致。

本小节采用注意力机制深度挖掘类内样本信息,学习类内单个样本与类原型特征之间的关系,利用不同贡献度的样本生成鲁棒的类原型特征,提升分类性能。注意力机制采用 Bi-LSTM 作为模块主体, LSTM 可以提取特征序列间依赖关系,通过门控机制增加或去除特征信息来增强神经元状态特征<sup>[18]</sup>,而 Bi-LSTM 可以提取双向状态信息,充分学习特征上下文信息。为充分学习类内样本关系,将类内样本特征展开为特征序列,采用基于 Bi-LSTM 的注意力机制,以 Bi-LSTM 提取的类内样本特征为先验,生成类内不同样本对应的权值分数,以加权方式生成鲁棒的类原型表达。类级注意力模块如图 2 所示。

图 2 中左侧的 3 个绿色长方体代表支持集中同类别 3 个影像样本的高维特征序列  $\phi_1^c, \phi_2^c, \phi_3^c$ 。以序列作

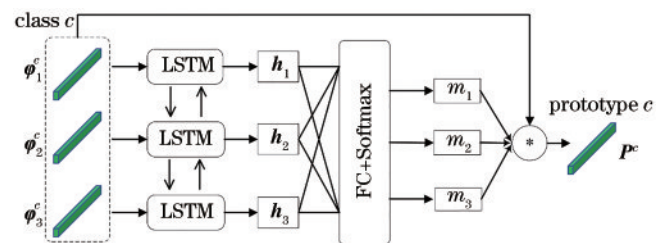


图 2 类级注意力模块  
Fig. 2 Class-level attention module

为 Bi-LSTM 模型的输入,提取 LSTM 单元的正向隐藏状态  $\overrightarrow{h}_k$  和逆向隐藏状态  $\overleftarrow{h}_k$ , 并采用 Concat 函数融合双向状态信息。隐藏状态  $h_k$  的表达式如下:

$$\overrightarrow{h}_k = \overrightarrow{\text{LSTM}}(h_{k+1}, e_k), \quad (2)$$

$$\overleftarrow{h}_k = \overleftarrow{\text{LSTM}}(h_{k-1}, e_k), \quad (3)$$

$$h_k = \text{Concat}(\overrightarrow{h}_k, \overleftarrow{h}_k), \quad (4)$$

式中:  $e_k$  表示第  $k$  个 LSTM 单元的输入。

为提取类间不同样本的上下文关系,模型隐藏状态使用全连接(FC)层融合全局特征信息,权值分数采用 Softmax 函数调整处理,最后使用加权求和操作生成最终的类原型特征  $P^c$ 。特征计算过程如下:

$$h_1, h_2, \dots, h_k = \text{LSTM}(\varphi_1^c, \varphi_1^c, \dots, \varphi_k^c), \quad (5)$$

$$m_1, m_2, \dots, m_k = \text{Softmax}[\text{FC}(h_1, h_2, \dots, h_k)], \quad (6)$$

$$P^c = \sum_{i=1}^k m_i \varphi_i^c, \quad (7)$$

式中:  $\varphi_k^c$  代表第  $c$  个类别中第  $k$  个样本的影像高维特征;  $\text{LSTM}(\cdot)$  为 Bi-LSTM 模型函数;  $h_k$  代表第  $k$  个 LSTM 单元的双向隐藏状态;  $\text{FC}(\cdot)$  函数代表全连接层;  $m_i$  代表权值分数;  $P^c$  代表类原型特征。

### 2.3 度量模块

查询集样本通过注意力类级模块,归纳生成类原

型特征  $P^c$ 。模型采用度量学习思想,通过计算查询集中每个输入样本映射的高维特征与每个类原型特征  $P^c$  之间的距离进行样本分类。查询样本和类原型之间的相似性使用欧氏距离进行度量,即相似性越大欧氏距离越小。在分类测试过程中,使用 Softmax 函数对未知样本  $x_k^c$  进行分类预测,预测类别计算  $y_{\text{pre}}$  如下:

$$D[f(x_k^c), P^c] = f(x_k^c) - P^c, \quad (8)$$

$$p_\phi(y_k^c = k | x_k^c) = \text{Softmax}\{D[f(x_k^c), P^c]\}, \quad (9)$$

$$y_{\text{pre}} = \text{argmax}[p_\phi(y_k^c = k | x_k^c)], \quad (10)$$

式中:  $f(x_k^c)$  代表查询样本高维特征;  $P^c$  代表第  $c$  类的类原型特征;  $D$  代表查询样本特征  $f(x_k^c)$  与类原型  $P^c$  之间的欧式距离;  $p_\phi(y_k^c = k | x_k^c)$  和  $y_{\text{pre}}$  分别代表样本  $x_k^c$  的 one-hot 向量和预测类别;  $\text{argmax}(\cdot)$  表示 one-hot 向量中最大值的对应类别。

## 3 实验

### 3.1 数据集

为验证所提方法的有效性,实验使用 UCMERGED、AID-30 和 NWPU-RESISC45 这 3 种遥感影像数据集进行验证,数据集部分样本如图 3 所示。

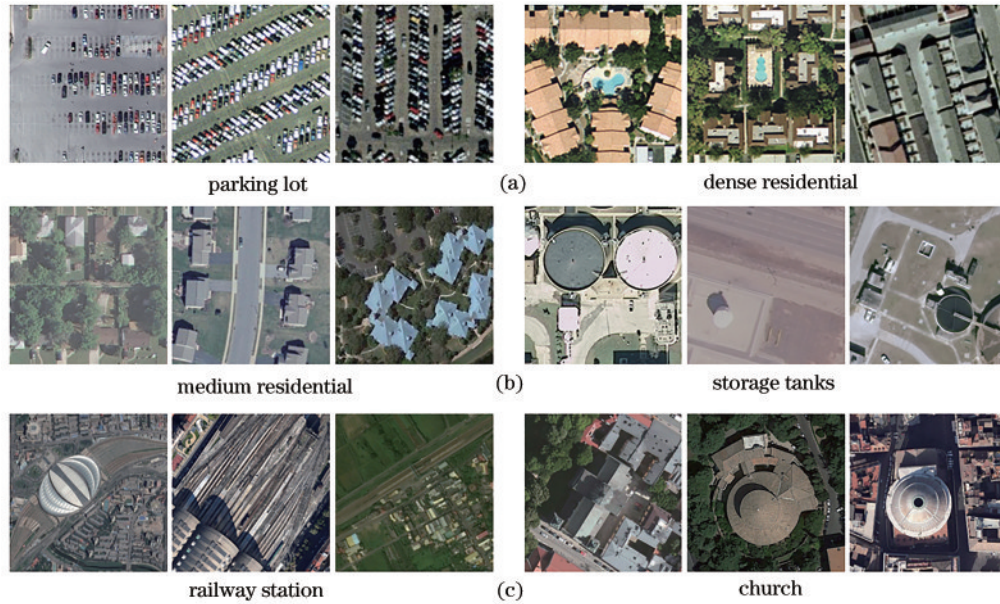


图 3 部分样本。(a) UC-21 数据集; (b) AID-30 数据集; (c) NP-45 数据集

Fig. 3 Partial samples. (a) UC-21 dataset; (b) AID-30 dataset; (c) NP-45 dataset

图 3(a) 为 UCMERGED 数据集: 包含 21 个类别, 每个类别有 100 张图片, 每张图片大小为  $256 \text{ pixel} \times 256 \text{ pixel}$ , 空间分辨率为  $0.3 \text{ m}$ ; 图 3(b) 为 AID-30 数据集, 包含 30 个类别, 每个类别有 200 张图片, 每张图片大小为  $600 \text{ pixel} \times 600 \text{ pixel}$ , 空间分辨率为  $0.5 \sim 8 \text{ m}$  之间; 图 3(c) 为 NWPU-RESISC45 数据集, 包含 45 个类别, 每个类别有 700 张图片, 每张图片大小为  $256 \text{ pixel} \times 256 \text{ pixel}$ , 空间分辨率为  $0.2 \sim 30 \text{ m}$  之间。

为统一表述,数据集分别以 UC-21、AID-30 和 NP-45 代称。

为构建小样本数据集,依据文献[13]分配策略对数据进行预处理。具体地讲,将单个数据集均分为 3 部分,3 个部分样本类别互不交叉。模型使用 2 个分割集进行训练,并使用交叉验证法对另一个分割集进行模型测试,3 部分分别代表 meta-train、meta-val 和 meta-test。在训练过程中,从 meta-train 中抽取多个元

任务模拟小样本学习任务进行训练,从 meta-val 中抽取多个元任务,对训练过程中的模型进行验证和参数择优;在模型测试过程中,从 meta-test 数据集中抽取多次元任务进行测试,取查询集的平均精度作为模型测试结果。

### 3.2 实验细节及配置参数

实验中训练和测试过程均按照小样本学习  $C$ -way  $K$ -shot 的形式进行实验设置,  $C$  和  $K$  分别为类别数和样本数,均可灵活设置。为模拟小样本遥感场景分类元任务,在样本少于 10 的范围内,参考小样本学习固定设置,选取代表性样本数量,对上述数据集均使用 5-way 1-shot、5-way 5-shot 和 5-way 10-shot 的形式,即支持集使用 1-shot、5-shot 和 10-shot 数量的样本完成 5 分类任务。训练过程:从 meta-train 数据集中随机抽取 5000 个元任务,采用跨任务思想的元学习策略进行训练,每 50 个元任务作为一个 epoch,共计 100 个 epoch。

每个元任务分为支持集和查询集这 2 个部分,实验中设置查询集样本数为 15。测试阶段:从 meta-test 集中抽取 100 个元任务,以多次元任务中查询集的平均精度作为单次抽取任务的测试结果,因抽取具有随机性,重复 6 次抽取任务,共计 600 次元任务,以平均精度作为最终结果。

表 3 UC-21/AID-30/NP-45 场景数据集上的分类效果(5-way)  
Table 3 Classification effect on UC-21/AID-30/NP-45 datasets (5-way)

Model	Network	UC-21			AID-30			NP-45			unit: %
		1-shot	5-shot	10-shot	1-shot	5-shot	10-shot	1-shot	5-shot	10-shot	
Transfer learning	GoogleNet	23.45	45.22	55.59	20.76	40.67	55.63	20.76	40.67	55.63	
	AlexNet	20.19	25.08	30.00	20.12	24.56	29.54	20.08	25.49	55.35	
	ResNet50	20.95	29.23	46.71	20.07	29.61	45.96	21.14	29.52	49.62	
	ResNet101	20.72	29.29	31.61	20.67	23.66	34.90	20.71	27.18	35.18	
Meta learning	MAML	47.53	63.13	64.99	47.93	61.79	69.90	42.29	61.84	68.77	
	ProtoNet	52.27	69.86	71.69	55.63	68.56	70.48	40.33	63.82	69.53	
	Lifelong	39.47	57.40					51.43	72.9		
	RS-MetaNet	57.23	76.08	81.23	56.32	74.48	80.57	52.78	71.49	77.37	
	RA-ProtoNet	<b>61.77</b>	<b>81.30</b>	<b>84.07</b>	<b>56.74</b>	<b>83.29</b>	<b>87.30</b>	<b>56.60</b>	<b>81.22</b>	<b>86.56</b>	

从表 3 可以看出,在面对极小样本时,基于迁移学习的卷积神经网络结果都处于较低水平,原因是传统模型虽然具有较好的初始化参数,但面对极少量未知样本情况下,模型对新类别样本的泛化能力较差。对于 UC-21 数据集,所提模型性能达到竞争水平,在 1-shot、5-shot 和 10-shot 数量的样本设置下,分类精度分别达到 61.77%、81.30% 和 84.07%,其总体分类精度均高于元学习模型 RS-MetaNet 的 57.23%、76.08% 和 81.23%;对于 AID-30 数据集,精度较先进水平分别提升 0.42 个百分点、8.81 个百分点和 6.73 个百分点;对于 NP-45 数据集,精度较先进水平分别提升 3.82 个百分点、8.32 个百分点和 9.19 个百分点。实验结果表明,所提方法可以在极小样本条件下,有效提取影像样

模型构建采用 PyTorch 深度学习库,训练使用交叉熵损失函数作为目标函数,使用 Adma 优化器对网络模型进行参数优化。实验设置初始学习率为 0.001,学习率每 20 步衰减一半,以提高模型的训练稳定性。模型实验环境如表 2 所示。

表 2 实验环境  
Table 2 Experimental environment

Framework	Python	cuDNN	GPU	CUDA	cuDNN
PyTorch	3.6.2	7.5	RTX 2080TI	10.1	7.5

为便于统一评估,使用总体分类精度(OA)作为实验的性能评价<sup>[13]</sup>,计算公式如下:

$$A_{\text{Acc}} = \frac{1}{t} \sum_{i=1}^t \frac{r^i}{Q_i}, \quad (11)$$

式中:  $r^i$  表示第  $i$  次任务中分类正确的数量;  $Q_i$  表示第  $i$  次任务的查询集样本总量。

### 3.3 实验结果与分析

为验证模型的有效性,将所提模型与基于迁移学习的卷积神经网络 AlexNet<sup>[19]</sup>、ResNet<sup>[17]</sup>、GoogleNet<sup>[20]</sup> 和基于元学习的 Lifelong learning<sup>[12]</sup>、MAML<sup>[21]</sup>、RS-MetaNet<sup>[13]</sup>、ProtoNet<sup>[14]</sup> 进行实验比较,实验结果如表 3 所示。

本特征,挖掘类内样本信息,提升小样本场景分类性能。

除 OA 外,混淆矩阵也是分类问题的重要评估指标。为进一步验证模型性能,实验选取 UC-21 和 AID-30 数据集,分别基于 7-way 5-shot 和 10-way 5-shot 实验设置,对所提模型和基准网络 PorotoNet 绘制混淆矩阵,结果如图 4、图 5 所示。从图 4 可以看出,在单次元任务中,所提模型提高了对 airplane、storage tanks、beach 和 mobile home park 的识别能力,其中,对 beach 和 mobile home park 类别的分类准确度提升至 100%。图 5 中,center 和 storage tanks 的分类精度得到有效提升。因此,所提模型虽然在少数类别上分类精度略微降低,但总体模型分类性能得到明显提升,进一步证明了所提方法的有效性。

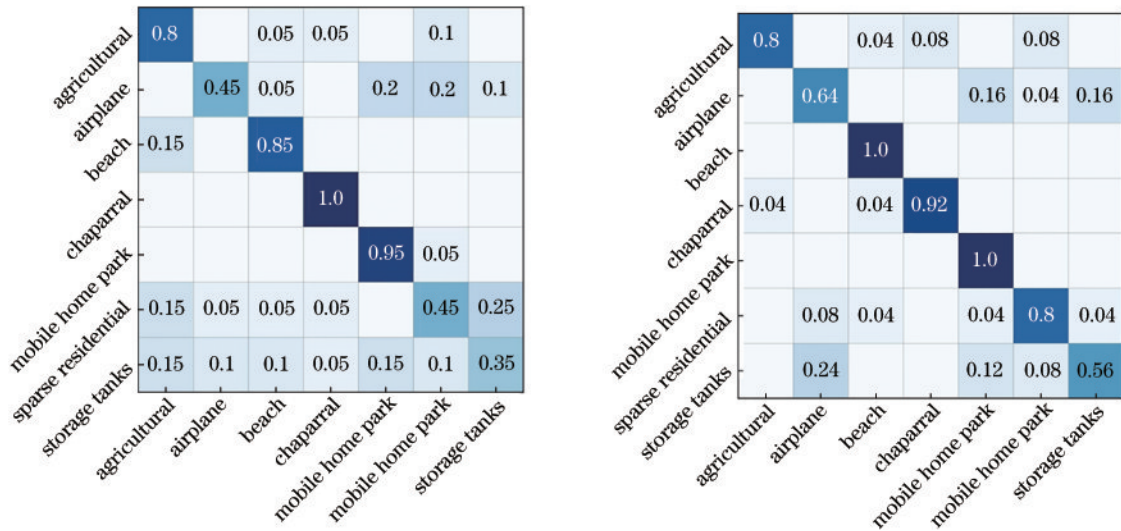


图 4 UC-21 数据集混淆矩阵: ProtoNet(左); RA-ProtoNet(右)  
Fig. 4 Confusion matrix of UC-21 dataset: ProtoNet (left); RA-ProtoNet (right)

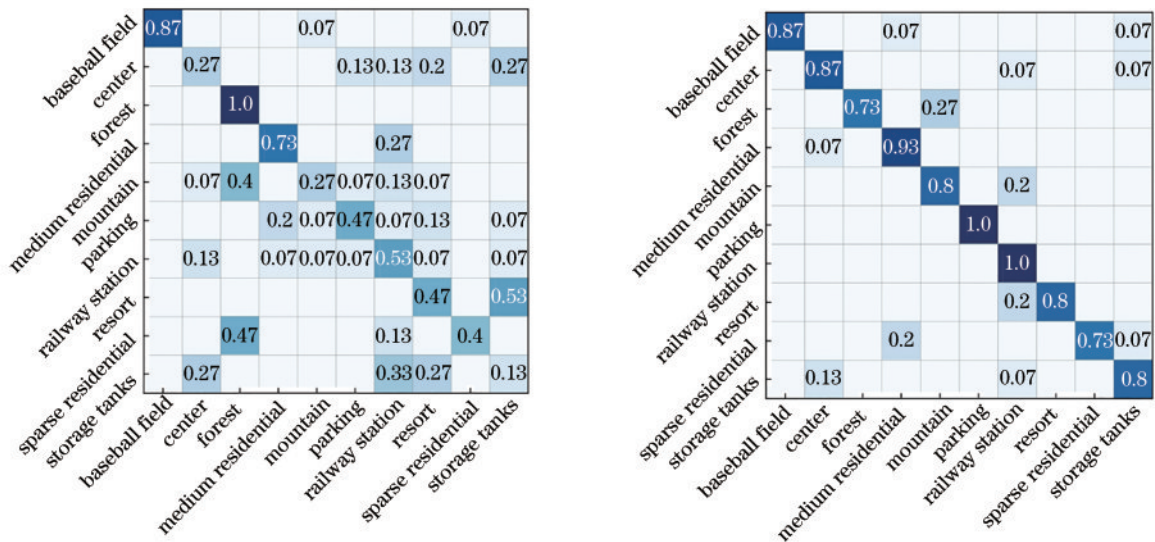


图 5 AID-30 数据集混淆矩阵: ProtoNet(左); RA-ProtoNet(右)  
Fig. 5 Confusion matrix of AID-30 dataset: ProtoNet (left); RA-ProtoNet (right)

除对比实验以外,为直观地表达场景分类结果,从 AID 数据集以 5-way 5-shot 设置抽取单元任务,进行可视化展示,如图 6 所示。左侧为单元任务中的支持集,包括 A-center、B-mountain、C-forest、D-storage tanks 和 E-baseball field 等 5 个类别,每个类别含 5 个样本;右侧为查询集中代表性样本的分类效果。从结果可以看出,大部分类别都能成功预测分类,但是出现 D 类被错分为 A 类,C 类被错分为 B 类的情况。结合图 6 左侧支持集样本和图 5 右侧混淆矩阵分析的结果可以看出,该元任务中支持集内 A 类与 D 类、B 类与 C 类之间影像相似性较大。在极小样本条件下,类间影像相似性较大,会导致生成的不同类的类原型在高维空间内距离较近,从而影响分类性能。同时,此结果也说明,在极小样本条件下需要充分提取遥感影像的高层特征来辅助相似地物类别的分类工作,进一步

证明所提方法选择深度嵌入模块的必要性。

### 3.4 消融实验

本小节对网络结构中的特征嵌入模块和类级注意力模块进行了消融实验,结果如表 4 和表 5 所示。表 4 在相同设置下对不同嵌入网络 VGG-16、ResNet14、ResNet18 和 ResNet50 进行评估分析。对于 UC-21 和 AID-30 数据集,非残差网络 VGG-16 的分类性能较低,而内存占用却较大,残差网络 ResNet14 有更好的分类性能,体现了残差单元的有效性。而对于 NP-45 数据集,ResNet18 则性能更优。在参数占用内存上,ResNet14 所占用内存空间最小,ResNet50 占用空间最大。为了平衡分类性能及内存,最终选用 ResNet14 作为特征嵌入网络。

所提模型通过归纳查询集样本特征生成类原型 prototype,表 5 为相同设置下,3 种类原型表达的分类

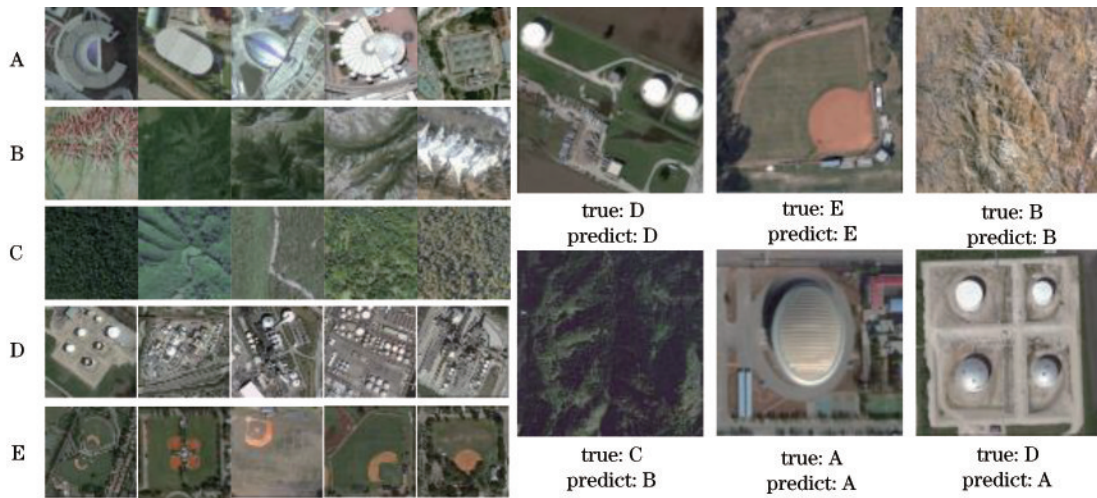


图 6 支持集及查询集部分测试结果

Fig. 6 Partial test results of support set and query set

表 4 不同嵌入网络的性能评估

Table 4 Performance evaluation of different embedded networks

Network	UC-21 /		AID-30 /		NP-45 /		Memory /
	%		%		%		
VGG-16	68.64	55.16	74.34		68.4		68.4
ResNet14	<b>81.30</b>	<b>83.29</b>	81.22		<b>35.1</b>		<b>35.1</b>
ResNet18	78.92	81.36	<b>82.23</b>		54.9		54.9
ResNet50	76.84	77.62	70.38		138		138

表 5 不同类原型表达方法评估

Table 5 Evaluation of different prototype expression methods

Feature representation	UC-21 / %		AID-30 / %		NP-45 / %	
	5-shot	10-shot	5-shot	10-shot	5-shot	10-shot
Sum	76.57	79.78	78.65	79.93	78.61	80.13
Mean	79.49	83.61	81.14	86.45	80.40	84.79
Attention	<b>81.30</b>	<b>84.07</b>	<b>83.29</b>	<b>87.30</b>	<b>81.22</b>	<b>86.56</b>

效果,分别为样本求和(sum)、样本均值(mean)和注意力机制(attention)。从表5可以看出,使用Bi-LSTM的类级注意力模块分类性能最优,证明了注意力机制可以有效融合类内样本信息,生成鲁棒的类原型特征,有效提升小样本遥感场景分类性能。

## 4 结 论

针对遥感领域极端场景下样本极少、影像场景复杂导致深度学习分类困难,以及深度学习在解决新类别分类时引起的重新启动导致训练代价高等问题,提出一种基于元学习的小样本遥感场景分类方法RA-ProtoNet。对于现有方法特征映射不充分和缺失类内判别能力的问题,引入迁移思想和残差网络构建特征嵌入模块,采用基于Bi-LSTM的注意力机制学习类内特征判别信息,生成鲁棒的类原型而提升分类性能。在不同数据集上的实验结果表明,所提模型提高了在极小样本条件下的场景分类性能,在不同数据集上具

有良好的泛化性。同时,RA-ProtoNet模型在面对新类别样本分类任务时,不需要重新训练,可以充分利用先验知识解决新任务,有效提升模型的使用率。

在实际工业场景中,遥感影像数据集的获取比较困难,现存样本的标注工作也需要大量人力和物力;同时,目前关于小样本遥感场景分类的研究有限。因此,如何有效解决小样本遥感领域分类问题将成为未来的研究热点。

## 参 考 文 献

- [1] 李德仁, 童庆禧, 李荣兴, 等. 高分辨率对地观测的若干前沿科学问题[J]. 中国科学: 地球科学, 2012, 42(6): 805-813.  
Li D R, Tong Q X, Li R X, et al. Some frontier scientific problems of high-resolution earth observation [J]. Scientia Sinica (Terrae), 2012, 42(6): 805-813.
- [2] Riaz F, Silva F B, Ribeiro M D, et al. Invariant Gabor texture descriptors for classification of gastroenterology images[J]. IEEE Transactions on Biomedical Engineering, 2012, 59(10): 2893-2904.
- [3] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [4] Zhu Q Q, Zhong Y F, Zhao B, et al. Bag-of-visual-words scene classifier with local and global features for high spatial resolution remote sensing imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(6): 747-751.
- [5] Cheriadat A M. Unsupervised feature learning for aerial scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 439-451.
- [6] Othman E, Bazi Y, Alajlan N, et al. Using convolutional features and a sparse autoencoder for land-use scene classification[J]. International Journal of Remote Sensing, 2016, 37(10): 2149-2167.
- [7] 贺琪, 李瑶, 宋巍, 等. 小样本的多模态遥感影像高层特征融合分类[J]. 激光与光电子学进展, 2019, 56(11):

111001.  
He Q, Li Y, Song W, et al. Multimodal remote sensing image classification with small sample size based on high-level feature fusion[J]. *Laser & Optoelectronics Progress*, 2019, 56(11): 111001.
- [8] 邱晓华, 李敏, 张丽琼, 等. 基于卷积特征和贝叶斯决策的双波段场景分类[J]. *激光与光电子学进展*, 2021, 58(4): 0415006.  
Qiu X H, Li M, Zhang L Q, et al. Dual-band scene classification based on convolutional features and Bayesian decision[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0415006.
- [9] Rostami M, Kolouri S, Eaton E, et al. Deep transfer learning for few-shot SAR image classification[J]. *Remote Sensing*, 2019, 11(11): 1374.
- [10] Scott G J, England M R, Starms W A, et al. Training deep convolutional neural networks for land-cover classification of high-resolution imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(4): 549-553.
- [11] Lima E, Sun X, Dong J Y, et al. Learning and transferring convolutional neural network knowledge to ocean front recognition[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(3): 354-358.
- [12] Zhai M, Liu H P, Sun F C. Lifelong learning for scene recognition in remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(9): 1472-1476.
- [13] Li H F, Cui Z Q, Zhu Z Q, et al. RS-MetaNet: deep metametric learning for few-shot remote sensing scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(8): 6983-6994.
- [14] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[C]//2017 Advances in Neural Information Processing Systems (NIPS), December 4-9, 2017, Long Beach, Los Angeles, USA. Cambridge: The MIT Press, 2017: 4077-4087.
- [15] Vinyals O, Blundell C, Lillicrap T, et al. Matching networks for one shot learning[C]//2016 Advances in Neural Information Processing Systems (NIPS), December 9-10, 2016, Barcelona Spain. Cambridge: The MIT Press, 2016, 29: 3630-3638.
- [16] 朱淑鑫, 周子俊, 顾兴健, 等. 基于RCF网络的遥感图像场景分类研究[J]. *激光与光电子学进展*, 2021, 58(14): 1401001.  
Zhu S X, Zhou Z J, Gu X J, et al. Scene classification of remote sensing images based on RCF network[J]. *Laser & Optoelectronics Progress*, 2021, 58(14): 1401001.
- [17] 汪鹏, 刘瑞, 辛雪静, 等. 基于残差网络的光学遥感图像场景分类算法[J]. *激光与光电子学进展*, 2021, 58(2): 0210001.  
Wang P, Liu R, Xin X J, et al. Scene classification of optical remote sensing images based on residual networks [J]. *Laser & Optoelectronics Progress*, 2021, 58(2): 0210001.
- [18] Alfattni G, Peek N, Nenadic G. Attention-based bidirectional long short-term memory networks for extracting temporal relationships from clinical discharge summaries[J]. *Journal of Biomedical Informatics*, 2021, 123: 103915.
- [19] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [20] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 11-12, 2015, Boston, MA. New York: IEEE Press, 2015.
- [21] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//ICML'17: Proceedings of the 34th International Conference on Machine Learning-Volume 70, August 6-11, 2017, Sydney, NSW, Australia. New York: ACM Press, 2017: 1126-1135.