

# 结合注意力和改进样本选取方法的少样本高光 谱分类孪生网络<sup>①</sup>



杨宇新<sup>1</sup>, 郭躬德<sup>1</sup>, 王 晖<sup>2</sup>

<sup>1</sup>(福建师范大学 计算机与网络空间安全学院, 福州 350117)

<sup>2</sup>(贝尔法斯特女王大学 电子学、电气工程与计算机科学学院, 贝尔法斯特 BT9 5BN)

通信作者: 郭躬德, E-mail: [ggd@fjnu.edu.cn](mailto:ggd@fjnu.edu.cn); 王 晖, E-mail: [h.wang@qub.ac.uk](mailto:h.wang@qub.ac.uk)

**摘 要:** 针对高光谱图像 (hyperspectral image) 样本人工标记困难导致的样本数量不足的问题, 本文提出了一个结合注意力和空间邻域的少样本孪生网络算法. 它首先对高光谱图像进行 PCA 预处理, 实现数据降维; 其次, 对模型训练样本采用间隔采样和边缘采样的方式进行选取, 以有效减少冗余信息; 之后, Siamese network 以大小不同的 patch 形式进行两两结合, 构建出样本对作为训练集进行训练, 不仅实现了数据增强的效果, 还能在提取光谱信息特征的同时, 充分提取目标像素光谱信息及其周围邻域空间信息; 最后, 添加光谱维度的注意力模块以及空间维度的相似度度量模块, 分别对光谱信息和空间邻域信息进行权重分布, 以达到提升分类性能的目的. 实验结果表明, 本文提出的方法在部分公开数据集上对比常用方法取得了较好的实验效果.

**关键词:** 高光谱图像分类; Siamese network; 注意力机制; 少样本学习; 深度学习

引用格式: 杨宇新, 郭躬德, 王晖. 结合注意力和改进样本选取方法的少样本高光谱分类孪生网络. 计算机系统应用, 2024, 33(3): 85-94. <http://www.c-s-a.org.cn/1003-3254/9432.html>

## Few-shot Hyperspectral Classification Siamese Network Combining Attention and Improved Sample Selection Method

YANG Yu-Xin<sup>1</sup>, GUO Gong-De<sup>1</sup>, WANG Hui<sup>2</sup>

<sup>1</sup>(College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China)

<sup>2</sup>(School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT9 5BN, UK)

**Abstract:** In order to solve the problem of the insufficient number of hyperspectral image samples due to the difficulty of artificial labeling, a small sample twin network algorithm combining attention and spatial neighborhood is proposed in this study. Firstly, the hyperspectral image is preprocessed by PCA to achieve data dimensionality reduction. Secondly, the training samples of the model are selected by means of interval sampling and edge sampling to effectively reduce redundant information. After that, the Siamese network combines the samples in the form of patches of different sizes and constructs the sample pairs for training as a training set, which not only realizes the effect of data enhancement but also fully extracts the spectral information of target pixels and the spatial information of its neighborhoods while extracting spectral information features. Finally, the attention module of spectral dimension and the similarity measurement module of spatial dimension are added to distribute the weight of spectral information and spatial neighborhood information respectively, so as to improve classification performance. The experimental results show that the proposed method achieves better experimental results compared with common methods on some public datasets.

① 基金项目: 国家自然科学基金 (61976053, 62171131); 福建省自然科学基金 (2022J01398)

收稿时间: 2023-09-21; 修改时间: 2023-10-20; 采用时间: 2023-10-25; csa 在线出版时间: 2024-01-09

CNKI 网络首发时间: 2024-01-10

**Key words:** hyperspectral image classification; Siamese network; attention mechanism; few-shot learning (FSL); deep learning (DL)

## 1 前言

高光谱图像 (hyperspectral image, HSI) 包含了数百个连续波段, 具有非常高的光谱分辨率, 这其中包含了大量的光谱和空间信息. 这些物理特性, 可以使分类时的“同谱异物”和“同物异谱”的现象减少, 可以应用于精准区分相似的地物分辨识别. 如今, HSI 分类已广泛应用于诸多领域, 例如环境监测、林业监测、灾害防治、环境评价等<sup>[1,2]</sup>.

在过去的几十年里, HSI 分类通过机器学习<sup>[3]</sup>得到了很大的发展. 在机器学习中, 有许多分类方法是通过利用 HSI 丰富的光谱特征来进行 HSI 分类, 包括主成分分析 (PCA)<sup>[4]</sup>、支持向量机 (SVM)<sup>[5]</sup>和随机森林 (RF)<sup>[6]</sup>等方法. 然而, 机器学习方法的网络层次往往数量较少, 网络结构深度不够, 从而限制了模型对 HSI 特征提取能力以及模型在不同场景下的泛用性, 这使得机器学习方法可能会在高光谱图像分类这一领域的进一步探索受到阻碍.

近年来, 深度学习 (deep learning, DL) 被广泛应用于 HSI 分类<sup>[7]</sup>, 相较于机器学习算法, 深度学习能够更好地完成分类任务. 深度学习典型的框架有卷积神经网络 (CNN)<sup>[8]</sup>、图卷积网络 (GCNs)<sup>[9]</sup>、循环神经网络 (RNN)<sup>[10]</sup>和 Transformer<sup>[11]</sup>等, 这些框架均被大量应用于高光谱图像分类领域的研究.

由于 HSI 人工标注困难, 样本量少的特性, 少样本学习 (few-shot learning, FSL) 的研究逐渐成为高光谱分类领域研究者的目标. 虽然深度学习在 HSI 分类任务上成效显著, 但是如果在训练样本数量过少的情形下, 深度学习模型很容易面临过拟合的困境. 并且根据数据集包含的种类信息, 有部分种类样本数量存在显著差异, 若选取过少的样本会使模型出现过拟合现象, 难以正确分类. 而 FSL 可以减少模型所需的标注样本量, 满足 HSI 标记样本量较少的情况, 从而降低标记困难带来的不利影响, 还能够通过轻量级的模型结构降低模型训练所需要的时间. 因此, 在 HSI 分类领域中对 FSL 方法的深入研究具有十分重要的意义.

## 2 研究现状

目前, 已经有很多研究人员对 HSI 分类领域中的

FSL 方法展开研究<sup>[12]</sup>. Xue 等人<sup>[13]</sup>提出了一个端对端的基于注意力的二阶池化网络, 通过设计了一个基于注意力的池化算子, 用于模拟判别特征和代表性特征. Zhang 等人<sup>[14]</sup>采用残差三维卷积神经网络 (CNN) 作为特征提取网络, 引入 contrast learning 以学习更多的 HSI 判别性表示. Manifold 等人<sup>[15]</sup>提出了一种新型的 U-Net 网络, 通过嵌套分别对高光谱图像进行空间学习和光谱学习, 可应用于多种领域. 但以上的 FSL 方法都是直接对样本进行训练分类, 并没有通过数据增强等方法增加训练样本量, 所以少样本带来的负面影响依旧存在.

除去上述的使用 CNN 的 FSL 方法, Siamese network 也是 FSL 领域中一个常用的方法. Siamese network 具有双分支结构, 传统的 Siamese network 就被用于人脸识别的场景, 比对两个输入图片的特征并返回一个相似度结果, 属于度量模型. 通过计算特征相似度来确定输入中一对样本是否匹配. 在这几年不断发展下, Siamese network 可以实现多分类的任务, 即确定样本对属于哪个类别. 通过计算每个样本的交叉熵损失进行分类, 产生类概率能够在测试阶段直接获得分类结果, 并且具有鲁棒性. 在标记样本有限的情况下, 能够通过构建任意数量的样本对来进行数据增强. Siamese network 与上述方法的不同之处在于能够通过将两个样本进行样本对配对, 从而实现训练样本的数据增强效果, 降低样本数量少带来的不利影响, 并且在通过编码后将样本映射到样本空间中, 通过 contrast loss 对拉近同类样本的类内距离, 拉远不同类样本的类间距离, 更好地提升模型的分类能力.

虽然已有大量的研究人员对 Siamese network 在 HSI 分类领域上进行了探索, 但在标记样本较少的情况下提高分类性能仍然是一个值得进一步探究的方向. Cheng 等人<sup>[16]</sup>提出了一个具有原型自校准 (SC) 和互校准 (IC) 的 Siamese 原型网络 (SPNet), 利用两个额外的损失函数来校准模型, 有助于模型学习到更具代表性的原型, 并做出更准确的预测. Cao 等人<sup>[17]</sup>设计了一个端对端的三维卷积组成的 Siamese network, 并加入了对比训练提升类间分类和类内分类的能力. Xue 等人<sup>[18]</sup>采用了将一维卷积和二维卷积组合的方法降低了参数

量,并考虑了使用不同大小的空间窗口大小输入可以提供更多的特征以及设计加权损失优化模型。但以上的算法在样本对构建时,都是选择直接在原 HSI 上进行随机选取获取训练样本。但经过调查大量公开遥感数据集后,可以得出高光谱遥感图像的各个种类在空间上的分布大多是块状的结论。上述算法没有考虑到在 HSI 空间上的分布情况,存在出现随机选取的样本点之间的间距过小的可能性,从而导致模型在进行空间上的特征提取时导致信息冗余。针对高光谱遥感图像的分布规律,我们采用了间隔选取和边缘样本选取相结合的方法进行样本选取,再在这个基础上进行两两配对,极大程度的获取该种类的信息,实验结果表明,在这个样本选取方法能够比随机选取的算法得到的分类结果更好。

近年来,注意力机制的运用也逐渐灵活,有很多研究者不断提出新的想法。Liu 等人<sup>[19]</sup>提出了一种使用多维 CNN 结合注意力机制模型,在该模型中,通过使用三维卷积层从 HSI 的 3D 数据中提取空间光谱特征。之后再通过二维卷积层和一维卷积层在抽象水平上进一步学习空间和光谱特征。Feng 等人<sup>[20]</sup>提出一种新型混合空谱特征级联融合网络 (MSSFN),利用三维残差模块和二维可分离空间残差模块提取混合空间光谱特征,增强特征提取能力。

对于以上使用了注意力机制的算法,大多都没有考虑作为训练样本的中心像素与邻域像素之间的相关性。为此,我们考虑了邻域像素与中心像素的相关性,通过计算向量的余弦距离,来生成二者之间的相似度,并通过该相似度来判断邻域像素的重要性分布,从而使得模型能够在训练时效果更好。

本文设计了一种基于注意力的 Siamese network。它首先利用一维卷积和二维卷积提取 HSI 的光谱空间信息,然后利用全局平均池化 (GAP) 进行分类,进一步降低网络参数和模型复杂度。其次,通过样本对的差异化输入,能够获取更多的样本对,并通过设置伪标签对每一类的训练样本对数量进行控制,避免了部分类别样本数过多或过少带来的过拟合或欠拟合影响,同时平衡了每类的正负样本对,有利于更好地实现类内分类和类间分类。最后,添加了光谱注意力机制和空间相似度度量模块,分别对光谱维度进行权重分布,从而更高效地提取光谱信息特征,同时对空间维度进行约束,考虑中心向量与相邻向量之间存在的关系,以达到更

好的特征提取。通过实验结果表明,我们的工作与在部分数据集上随机选择的方法相比取得了较好的实验效果。

综上所述,本文的主要创新点可以概括为以下几点。

1) 对训练样本的选取方法的改进。综合考虑 HSI 中的信息样本分布规律,减少冗余特征选取,从而提升模型分类性能。

2) 采用了差异化输入。我们选择的 Siamese network 在构建样本对时,相较于采用相同大小补丁构建样本对,采用差异化输入能够获得更多的样本对,降低样本不足的影响。

3) 添加空间相似度度量模块。以往的研究对样本邻域间的研究较少,我们采用余弦相似度对中心向量及邻域向量进行相似度度量,获得邻域像素的相关性分布,利于模型训练。

### 3 基于注意力的 Siamese network

我们主要的想法是设计一个添加注意力机制的 Siamese network 用于少样本高光谱遥感图像分类任务。下面主要介绍我们设计的模型结构以及添加的注意力模块。

#### 3.1 模型结构

我们构建了一个 Siamese network 框架,其模型结构如图 1 所示,其中  $W_s1$  表示大补丁的空间窗口大小, $W_s2$  表示小补丁的空间窗口大小。在该框架训练开始之前,对原高光谱图像进行 PCA 处理,降低维数至  $K$ ,再通过样本对配对来进行数据增强。首先,将大小两个补丁送入网络中进行特征提取。在经过 4 层一维卷积层后,采用光谱注意力模块对光谱特征进行重标定,再将光谱维度与通道进行合并,通过 1 层二维卷积层获得全局的光谱信息。之后对提取到的特征,我们选取中心的特征向量与相邻的特征向量进行相似度度量,获得空间维度的权重图,将权重与原特征图进行相加后,再进行 3 层二维卷积层操作进行进一步空间特征提取。最后通过全局平均池化层来获取分类的结果。

在分类任务中,对于距离远的正样本和距离近的负样本往往被认为是困难样本,会对实验结果带来较大的影响。为此,在训练过程中,我们选择将输出的结果进行对比训练,通过对比训练能够拉近类内距离,增大类间距离,有利于降低困难样本对实验结果的干扰。

同时,我们在最后一层二维卷积后选取特征图中的中心向量来计算其欧氏距离作为权值,有利于对比损失函数的拟合.

对于模型训练时的调优,我们通过 contrast loss 和 cross entropy loss 对模型进行优化.其中,对比损失

函数 contrast loss,我们添加了两个补丁块的中心向量之间的欧氏距离作为权值,通过加权可以使得模型更加关注硬样本,促使模型向正确的方向拟合.交叉熵损失函数 cross entropy loss 是常用的多分类损失函数,主要是用来判定实际的输出与期望的输出的接近程度.

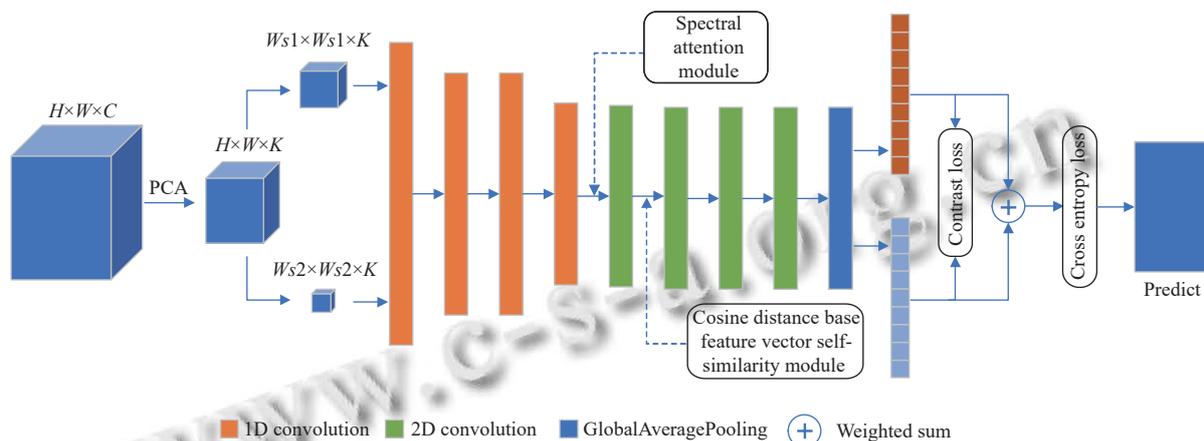


图1 网络结构图

算法 1. 模型算法

输入: 整个 epochs 为  $N$ , PCA 处理后的 HSI 图像中的补丁块  $X_1, X_2$ .  
输出: 对整个图像的预测标签结果.

训练阶段:

- 1) 通过间隔采样与边缘采样相结合的取样方法,选取训练样本作为训练集,剩余样本作为测试集.
- 2) 根据训练样本索引选取补丁块,将每个大补丁块与所有小补丁块进行两两配对,获得训练样本对  $\{P_1, P_2\}$ .
- 3) for epoch from 1 to  $N$  do:
- 4) 将  $P_1, P_2$  输入模型中,得到特征向量  $y_1, y_2$
- 5) 根据  $y_1, y_2$  计算 contrast loss
- 6) if  $\{P_1, P_2\}$  的中心为同一样本 then
- 7) 计算交叉熵损失 cross entropy loss
- 8) 通过 contrast loss 和 cross entropy loss 和 Adam 优化器对模型进行调优
- 9) end if
- 10) end for

测试阶段:

- 1) 根据测试样本索引选取补丁块,将每个大补丁块与所有小补丁块进行两两配对,获得训练样本对  $\{P_3, P_4\}$
- 2) for epoch from 1 to  $N$  do:
- 3) 将  $P_3, P_4$  输入模型中,得到特征向量  $y_3, y_4$
- 4) 使用  $\arg \max(\alpha \times y_3 + \beta \times y_4)$  获得预测标签
- 5) end for

3.2 光谱注意力机制

为提升模型性能,让模型能够更好地关注光谱信息,我们设计了一个光谱注意力机制,如图2所示.模

型的输入为补丁,通过 AveragePool3D 进行池化操作,将原大小为  $W_s \times W_s \times k$  的补丁池化为  $1 \times 1 \times k$ ,通过池化操作后,可以获得全局的感受野.再通过两个全连接操作,将光谱维度的信息进行激励操作,通过全连接层和激活函数之后,获得光谱维度上的权重信息  $w$ ,最后将权重信息  $w$  与原输入的补丁进行相乘操作,从而得到每个光谱维度信息的重要性分布,完成对光谱特征的重标定.

3.3 邻域相似度模块

根据高光谱图像本身的特性可知,每一类样本在空间上大多以块状聚集在一起,但是以前的工作大多没有考虑到特征空间中的中心光谱向量与相邻光谱向量之间的联系.为此,我们添加了一个基于余弦距离的特征向量自相似模块,用于探索原始输入空间中不同相邻光谱像元的空间互补信息的贡献.

如图3所示,将进行降维后的特征映射作为输入,首先通过一个一维卷积层来将原本的通道数映射为3倍,再进行切分生成  $Q, K, V$ .之后在  $Q$  中选取其中心向量与  $K$  中的所有相邻向量通过计算向量之间的余弦距离得到它们之间相似度,再将生成的结果通过 Softmax 函数后可以得到一个拥有权重分布的特征图,之后将生成的特征图与  $V$  进行重新加权,最后与原输入进行相加从而得到中心向量与相邻向量之间的相关度关系.

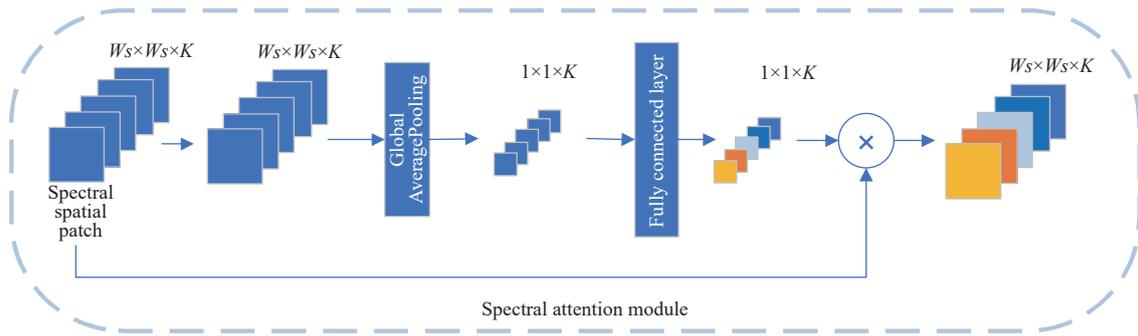


图2 网络结构图

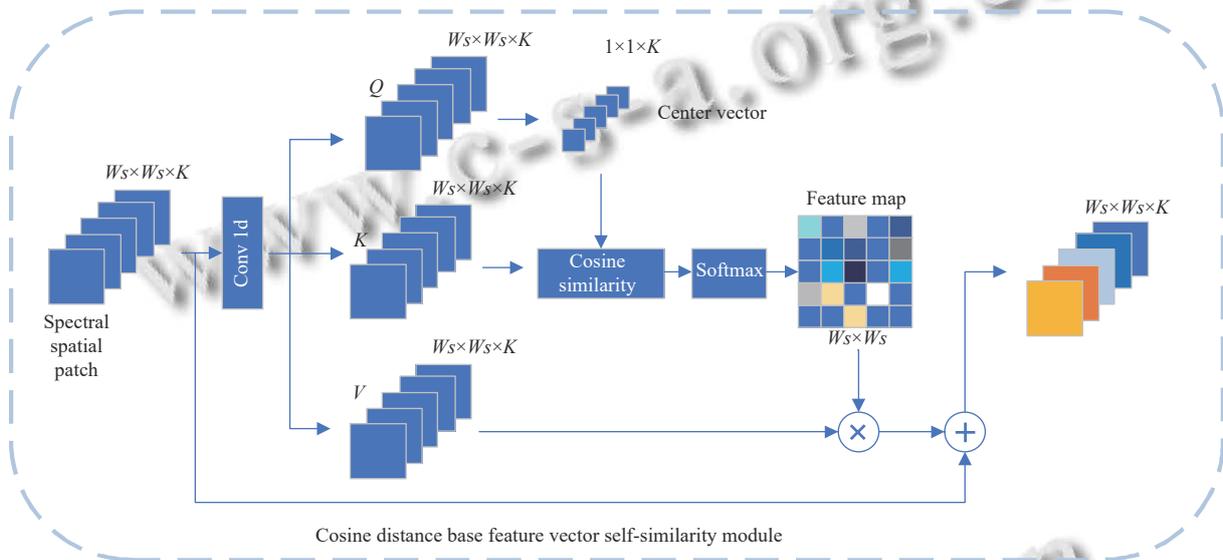


图3 网络结构图

### 4 实验分析

下面实验部分将介绍实验使用的公开数据集, 数据预处理方法, 实验的相关设置以及实验的结果与分析.

#### 4.1 数据集介绍

为了公平地证明所提出的网络的有效性和效率, 我们在3个流行的公开HSI数据集上进行了广泛的实

验, 即 Salinas (SA)、Indian Pines (IP) 和 Pavia Centre and University (PU). 表1列出各数据集包含的类别和对应的数量, 并且在数据集后的括号中标注了实验所选取的样本比例. 从表1可知, SA有16个类别, 共计54 129个样本. IP有16个类别, 共计10 249个样本. PU共有9个类别, 共计42 776个样本.

表1 3个数据集包含的各个种类及对应数量

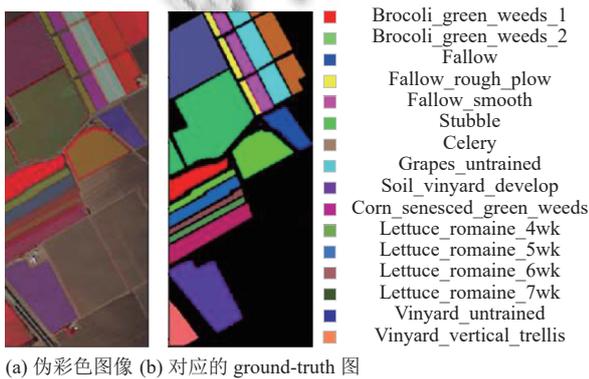
No.	SA (2%)		IP (10%)		PU (2%)	
	类别	数量	类别	数量	类别	数量
1	Brocoli_green_weeds_1	2009	Alfalfa	46	Asphalt	6631
2	Brocoli_green_weeds_2	3726	Corn-notill	1428	Meadows	18649
3	Fallow	1976	Corn-mintill	830	Gravel	2099
4	Fallow_rough_plow	1394	Corn	237	Trees	3064
5	Fallow_smooth	2678	Grass-pasture	483	Painted metal sheets	1345
6	Stubble	3959	Grass-trees	730	Bare soil	5029
7	Celery	3579	Grass-pasture-mowed	28	Bitumen	1330
8	Grapes_untrained	11271	Hay-windrowed	478	Self-blocking bricks	3682
9	Soil_vinyard_develop	6203	Oats	20	Shadows	947

表 1 (续) 3 个数据集包含的各个种类及对应数量

No.	SA (2%)		IP (10%)		PU (2%)	
	类别	数量	类别	数量	类别	数量
10	Corn_senesced_green_weeds	3 178	Soybean-notill	972	—	—
11	Lettuce_romaine_4wk	1 068	Soybean-mintill	2 455	—	—
12	Lettuce_romaine_5wk	1 927	Soybean-clean	593	—	—
13	Lettuce_romaine_6wk	916	Wheat	205	—	—
14	Lettuce_romaine_7wk	1 070	Woods	1 265	—	—
15	Vinyard_untrained	7 268	Buildings-grass-trees-drives	386	—	—
16	Vinyard_vertical_trellis	1 807	Stone-steel-towers	93	—	—
	Total	54 129	Total	10 249	Total	42 776

### 4.1.1 Salinas (SA) 数据集

第 1 个 HSI 是在 1992 年由 AVIRIS 传感器在美国加利福尼亚州的 Salinas 山谷上空获得的, 具有 3.7 m 的高空间分辨率, 图像尺寸为 512×217 个像素点, 采集图像的通道数共有 224 个波段. 在去除了 20 个含有噪声和吸水波段之后, 共有 204 个波段用于实验中. SA 数据集图像及 ground-truth 图如图 4 所示.



(a) 伪彩色图像 (b) 对应的 ground-truth 图

图 4 SA 数据集图像及 ground-truth 图

### 4.1.2 Indian Pines (IP) 数据集

第 2 个 HSI 是由 AVIRIS 传感器在美国印第安纳州的西北部的 Indian Pines 地区获得的. 图像尺寸为 145×145 个像素点, 采集图像的通道数共有 220 个波段. 在去除了覆盖吸水区域的波段之后, 共有 104 个波段用于实验中. IP 数据集图像及 ground-truth 图如图 5 所示.

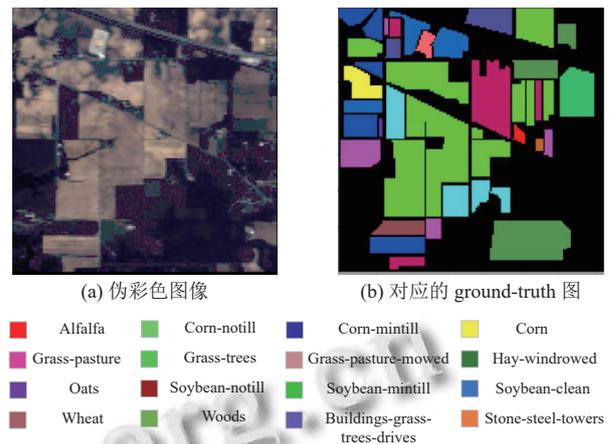
### 4.1.3 Pavia Centre and University (PU) 数据集

第 3 个 HSI 是由 ROSIS 传感器在意大利北部的 Pavia Centre and University 上空获得的, 具有 1.3 m 的空间分辨率, 图像尺寸为 610×340 个像素点, 采集图像的通道数共有 103 个波段. PU 数据集图像及 ground-truth 图如图 6 所示.

## 4.2 数据预处理

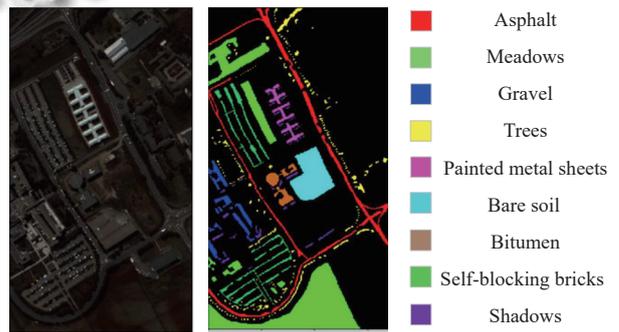
在数据输入模型之前, 我们对原始形状为  $H \times W \times C$

的 HSI 进行预处理. 首先对光谱维度进行 PCA 的降维操作, 将原始图像形状变为  $H \times W \times K$ , 其中  $H$  为空间维度的高度,  $W$  为空间维度的宽度,  $C$  为原高光谱图像的光谱带数,  $K$  为 PCA 处理后图像的通道数. 通过多次实验, 我们选择将高光谱图像的通道数降低至 60, 即  $K=60$ . 这样可以有效减少数据维度, 降低计算复杂度.



(a) 伪彩色图像 (b) 对应的 ground-truth 图

图 5 IP 数据集图像及 ground-truth 图



(a) 伪彩色图像 (b) 对应的 ground-truth 图

图 6 PU 数据集图像及 ground-truth 图

高光谱分类任务的目的是通过已有的标记样本来预测未标记的样本, 为每一个未标记的样本分配一个语义标签. 从高光谱数据的形式上看, PCA 处理后的高光谱图像的像素集  $X \in \mathbb{R}^{H \times W \times K}$ ,  $X = \{x_1, x_2, \dots, x_{m-1}, x_m\}$ ,

以及其对应的标签集  $Y = \{y_1, y_2, \dots, y_{m-1}, y_m\}$ , 其中  $m$  为总像素个数, 每一个像素集上的点都对应着一个标签集中的值. 根据 Siamese network 的输入特性, 我们需要选取部分样本来构建我们的训练集. 考虑到高光谱图像中往往含有多个种类, 且样本种类之间的数量上存在较大差异, 我们选择按 2% 的比例对每类样本进行选取, 从而获得训练集  $X_{[p]} = \{x_1, x_2, \dots, x_p\}$ , 其中  $p$  为训练集样本个数. 样本对的生成将从训练集  $X_{[p]}$  中选取并进行两两组合, 并赋予标记  $Z_{i,j}$ , 其值为:

$$Z_{i,j} = \begin{cases} 0, & y_i \neq y_j \\ 1, & y_i = y_j \end{cases}$$

通过样本对的构建, 我们能够获取大量的训练样本对, 达到了数据增强的目的, 有利于模型的训练. 为推进模型的进一步拟合, 防止某类的样本对数量过多, 我们选择设置一个假标签 flag, 将每一个样本对都使用假标签作为索引, 可以利用假标签来从中选取部分样本对.

为减少训练时间, 我们选择在每类的正样本对中通过假标签按 10% 的比例确定最终的训练正样本数. 但由于正负样本对的数量上存在极大的不平衡, 因此选择随机选取总数等同于正样本对的负样本对, 将正负样本对共同组合成训练集, 避免了样本对数量不平衡导致模型对类间分类的潜在影响.

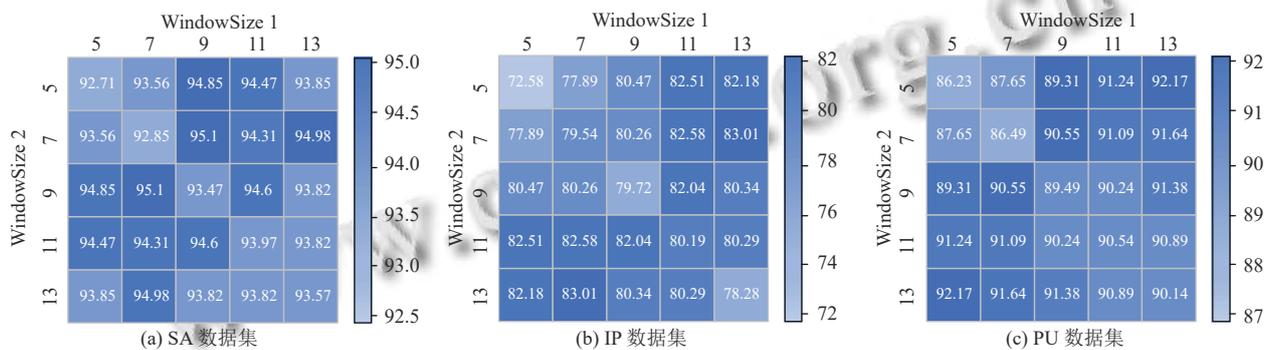


图7 3个数据集在不同窗口大小下 OA 值 (%)

(2) 网络性能表现: 为了比较提出的方法以及网络的性能和效率, 我们将所提出的网络与两个传统的机器学习算法进行比较, 即带有高斯核函数的支持向量机 (SVM) 和随机森林 (RF). 同时也与 5 个先进的深度学习模型进行比较, 即 SSTN<sup>[21]</sup>、SSAN<sup>[22]</sup>、SSAN<sup>[23]</sup>、CVSSN<sup>[24]</sup>、S3Net<sup>[18]</sup>. 所选择的对比算法覆盖了传统的机器学习算法、基于 Transformer 的算法、基于三维

构建好样本对之后, 我们选择以 patch 的形式进行输入, 将以样本对中对应的像素点作为 patch 的中心, 每个像素对应一个补丁, 边缘像素使用零填充来构建补丁, 生成一组补丁对, 其中每个小块  $P_{i,j} \in \mathbb{R}^{W_s \times W_s \times k}$  覆盖了  $W_s \times W_s$  空间大小的窗口. 在此, 我们沿用了差分输入的思想, 将样本对生成一个大补丁和一个小补丁. 由于不同的样本拥有不同的特征分布, 以不同大小的窗口输入的形式可以提供更多的特征进行学习, 从而增强模型的识别能力.

### 4.3 实验设置

(1) Patch 的窗口大小: 3D 补丁块的空间窗口大小决定该补丁块空间相邻信息的数量, 窗口尺寸越大则其包含的信息越多, 在特征提取时就能挖掘出更多的信息, 从而提升模型的性能. 因此, 我们做了相应的实验来确定最合适的窗口大小, 并考虑针对 Siamese network 输入的补丁对策略选择采用不同大小的窗口尺寸, 原因在于相邻像素大多数是某一类的情况下, 其中心像素大概率为相同类别, 就不需要控制过大的窗口进行判断, 以减少冗余信息和训练时间.

为验证该想法的有效性, 我们设置了间隔为 2 的窗口大小分别从 5 到 13 进行组合, 其分类评估的 OA 值如图 7 所示. 可以看出在使用差异化输入方法后, 结果会略高于采用相同大小输入的补丁对.

CNN 的算法、基于注意力的算法和 Siamese network 的算法. 在我们的实验中, 会根据不同的算法输入调整训练集, 具体的配置在第 4.4.3 节.

(3) 评价指标: 为比较不同方法和模块的分类性能, 我们采用总体精度 (overall accuracy, OA)、平均精度 (average accuracy, AA) 和 Kappa 系数这 3 个指标来评价不同模型的性能. 其中, OA 是正确分类的标记样本

数量与测试集中总样本数量的比值. AA 是所有土地覆盖等级精度的平均值. Kappa 综合衡量分类结果与基础真值之间的一致性. 在所有实验中选择进行 10 次独立运行, 并将返回各指标计算的结果取平均值作为最终结果.

(4) 运行平台: 所有实验均在 Windows 11、Intel i7-8750u、NVIDIA GeForce RTX 3090、64 GB 内存硬件下进行. 不同的深度学习框架是在 Python 3.6 环境下基于 PyTorch 1.14 构建的.

#### 4.4 消融实验

##### 4.4.1 样本对构建对比实验

我们选择采用间隔采样和边缘采样的方式选取一定比例的训练样本, 并将训练样本组合生成样本对输入到模型中. 为验证该方法的有效性, 我们进行了两个对比实验. 由于样本的选择大多是按比例或者按固定样本数量进行选取, 第 1 个对比实验我们验证采用比例进行选取优于按固定样本数进行选取, 我们将与按固定样本选取的方式进行对比, 在 SA 数据集中, 将统一按照  $11 \times 9$  的窗口大小进行训练, 分别选择采用 2% 的比例进行选取和固定样本数量为每类 40 样本的方法进行对比实验, 从表 2 中可以看到, 按照比例选取的方法能够对结果有进一步的提升. 比例选取可以避免对某一类数量较多的类训练不足导致的分类效果不好, 从而提升整体的分类结果.

表 2 样本选取对比实验的 OA 值 (%)

Dataset	固定样本	比例选取
SA	96.07±2.55	99.81±0.09

第 2 个实验我们在比例选取样本的基础上, 通过添加间隔采样和边缘采样的方法进行对比实验, 实验结果如表 3 所示, 表中勾选则说明采用了该方法. 可以看到采用二者结合的方法能够提升整体的性能.

表 3 SA 数据集上的间隔选取与边缘选取对比实验 (%)

间隔采样	边界样本	OA
—	—	99.81±0.09
√	—	99.84±0.02
√	√	99.87±0.12

##### 4.4.2 不同窗口大小输入

为验证该方法的有效性, 我们在每类选择 10 个样本的基础上, 按照窗口大小 {5, 7, 9, 11, 13} 分别对两个样本进行设置, 再将两个样本构建成样本对的形式输入到网络中, 设置 epoch 为 50 训练 10 次, 取其对应的

平均 OA 值. 从结果上看, 该方法相对于采用相同窗口大小的输入, 准确率 OA 能够有一定提升.

如图 7 所示, 在 SA 数据集中, 当 WindowSize 大小分别为 9 和 7 时能获得最好的分类结果. 因此在后续的实验中我们将继续设置 WindowSize 大小分别为 9 和 7. 同理, 在 IP 数据集中, 我们则选择采取 WindowSize 大小分别为 13 和 7 进行后续实验, 在 PU 数据集中采取 WindowSize 大小分别为 13 和 5 进行后续实验.

##### 4.4.3 实验结果与分析表现

在 SA 和 PU 的实验中, 实验参数都设置为按 2% 比例选取样本构建训练集, 并在需要构建验证集的算法 (SVM、RF、SSTN、SSAN、SSAN、CVSSN) 中, 采用 0.2% 的比例构建验证集, 其余样本均用于测试. 而 IP 数据集由于样本数量较少, 我们选择设置 10% 的比例进行构建, 同样在以上算法中, 我们采用 1% 的比例构建验证集, 其余样本用于测试.

在我们的网络和 S3Net 中是采取使用样本对的形式输入, 而其余算法的输入均为单个样本输入, 为保证实验的准确性, 我们将采用单个样本输入的对比算法中的输入 patch 大小设置与样本对中的较大 patch 一致. 其次, S3Net 在数据选择上是使用每类固定样本数进行训练集的构建, 我们每类选择设置为 40, 与我们 2% 比例选取的样本数大致相同, 确保算法之间能够正常比较.

表 4 中是模型运行 10 次统计出的 OA、AA 和 Kappa 在 SA、IP 和 PU 这 3 个数据集上的结果. 从表 4 中可以看出, 深度学习的算法效果远远强于传统的机器学习算法, 因此我们的对比主要关注在深度学习算法上. 在同样采取比例选取样本构建训练集的算法 (SSTN、SSAN、SSAN、CVSSN) 中, 我们的算法取得了最优的性能. 在 SA 数据集上与次优的 CVSSN 相比, OA、AA 和 Kappa 分别提升了 1.53%、0.85% 和 1.71%. 对 IP 数据集而言, 与 CVSSN 相比, OA、AA 和 Kappa 分别提升了 0.6%、1.3% 和 0.7%. 同样, 在 PU 数据集中与 CVSSN 相比, OA、AA 和 Kappa 分别提升了 0.44%、0.34% 和 0.58%.

对于同样采用了 Siamese network 构建模型的 S3Net, 我们的模型在选取训练集上采取的方法不同, 从结果上看, 我们的模型在相同的数据集上所采用的间隔采样和边缘样本采样的方法能够比 S3Net 算法所使用的按每类固定数量的选取方法得到的结果更优, 在 SA 数据集上, OA、AA 和 Kappa 分别提升了

1.79%、0.5% 和 1.99%。在 IP 数据集上, OA、AA 和 Kappa 分别提升了 3.67%、0.91% 和 4.22%。在 PU 数据集上, OA、AA 和 Kappa 分别提升了 2.04%、

0.65% 和 2.72%。对于其他按比例选取样本的深度学习算法, 我们的模型在面对少样本的情况下能够取得更好的结果。

表4 各算法在3个数据集上的分类结果

数据集	指标	RF	SVM	SSTN <sup>[21]</sup>	SSAN <sup>[22]</sup>	SSSAN <sup>[23]</sup>	CVSSN <sup>[24]</sup>	S3Net <sup>[18]</sup>	Ours
SA	OA (%)	88.26±0.39	90.07±0.3	96.1±0.87	96.1±0.59	94.02±2.87	98.06±0.25	97.81±0.58	<b>99.57±0.44</b>
	AA (%)	92.5±0.4	94.32±0.23	97.53±0.63	97.52±0.43	96.95±0.88	98.8±0.27	99.07±0.20	<b>99.64±0.23</b>
	Kappa×100	86.9±0.43	88.91±0.34	95.66±0.96	95.66±0.66	93.34±3.24	97.84±0.28	97.56±0.65	<b>99.52±0.5</b>
IP	OA (%)	75.1±0.83	80.31±0.49	95.01±0.56	92.13±0.79	96.9±0.49	98.24±0.36	95.34±0.83	<b>98.84±0.15</b>
	AA (%)	72.03±3.91	79.43±1.69	91.64±0.95	91.79±1.86	96.17±1.11	97.17±0.71	97.62±0.49	<b>98.51±0.49</b>
	Kappa×100	71.26±0.95	77.4±0.59	94.31±0.64	91.01±0.91	96.46±0.56	97.99±0.41	94.68±0.95	<b>98.68±0.18</b>
PU	OA (%)	84.08±0.14	87.33±0.66	96.84±0.98	94.09±0.84	97.42±0.78	98.72±0.37	97.17±0.92	<b>99.16±0.33</b>
	AA (%)	84.18±0.77	88.04±0.49	95.18±1.48	91.35±1.37	96.47±1.41	98.18±0.52	97.88±0.57	<b>98.52±0.73</b>
	Kappa×100	78.32±0.21	82.81±0.95	95.81±1.31	92.15±1.12	96.58±1.03	98.31±0.49	96.27±1.19	<b>98.89±0.44</b>

#### 4.4.4 注意力模块和相似度量模块消融实验

为验证模型中所添加的注意力模块和相似度量模块的有效性, 我们在 SA 数据集上进行了消融实验。其结果如表 5 所示。

表5 添加的模块的消融实验

Model	Attention	Similarity	OA (%)	AA (%)	Kappa×100
I	—	—	99.62	99.44	99.05
1-A	√	—	99.71	99.19	99.14
1-S	—	√	99.66	99.58	99.47
1-AS	√	√	99.74	99.86	99.84

在表 5 中, 勾选表示添加了对应的模块, Attention 表示光谱注意力模块, Similarity 表示空间邻域相似度量模块, Model 列中的“1”表示没有添加注意力模块, “1-A”表示添加了 Attention 模块, “1-S”表示添加了 Similarity 模块, “1-AS”表示两者都添加。最后一栏表示在对应的模块选择上所得到的分类结果评测指标。由于样本对是按照比例进行随机选取, 我们进行了多次重复的训练, 选取各类中得到的最好结果作为最终结果。从结果上看, 添加模块后的模型各评测指标均比未添加的模型各评测指标高出一些, 在二者都添加之后可以获得最优的分类结果。因此, 上述实验验证了添加这两个注意力模块的有效性。

## 5 结论与展望

本文采用了比例选取方式确定训练样本数总数, 并通过间隔采样和边缘采样的方法对样本点进行选取, 两两组合构建样本对实现数据增强, 有效削弱了样本量不足带来的过拟合问题, 再以样本对的形式输入到设计的孪生网络中去。我们的算法按照以上方式进行样本选取, 相对于其他算法按随机选取样本的方式选

取, 训练结果在各个指标上都有着提升。在实际的高光谱分类的实验中, 其分类结果证明了本文采取的样本选取方法以及算法的有效性。我们同时也进行了光谱注意力模块和空间相似度量模块的消融实验, 实验结果表明, 添加两个模块之后模型性能有些许提升, 证明了模块的有效性。

Siamese network 能够以样本对的形式实现数据增强的效果, 相较于常规以单个样本为输入的算法, 能够拥有大量可供训练的样本, 适用于少样本的情景中进行分类。同时, 我们添加了注意力机制和相似度量模块, 分别作用于高光谱图像的光谱维度和空间维度, 关注光谱间的相关性和空间上中心向量与相邻向量之间的联系, 可以对模型分类带来提升。

当前我们的方法采用比例选取的方法生成训练集, 并两两组合生成样本对, 但是存在正负样本对数量不平衡的情况, 目前我们采取随机选取 10% 的正样本对, 并随机选取相同数量的负样本对共同组成训练集输入模型中进行实验。今后, 我们可以考虑结合主动学习对样本对进行更高效的选取, 从而进一步提升分类性能。另外, 对于负样本对的选取仍是一项值得进一步探索的工作。

## 参考文献

- Fauvel M, Tarabalka Y, Chanussot JA, *et al.* Advances in spectral-spatial classification of hyperspectral images. Proceedings of the IEEE, 2013, 101(3): 652–675. [doi: 10.1109/JPROC.2012.2197589]
- Lei JJ, Li XY, Peng B, *et al.* Deep spatial-spectral subspace clustering for hyperspectral image. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(7): 2686–2697. [doi: 10.1109/TCSVT.2020.3027616]

- 3 Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. *Science*, 2015, 349(6245): 255–260. [doi: [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415)]
- 4 Kang XD, Xiang XL, Li ST, *et al.* PCA-based edge-preserving features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(12): 7140–7151. [doi: [10.1109/TGRS.2017.2743102](https://doi.org/10.1109/TGRS.2017.2743102)]
- 5 Tarabalka Y, Fauvel M, Chanussot J, *et al.* SVM- and MRF-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 2010, 7(4): 736–740. [doi: [10.1109/LGRS.2010.2047711](https://doi.org/10.1109/LGRS.2010.2047711)]
- 6 Du PJ, Bai XY, Tan K, *et al.* Advances of four machine learning methods for spatial data handling: A review. *Journal of Geovisualization and Spatial Analysis*, 2020, 4: 13. [doi: [10.1007/s41651-020-00048-5](https://doi.org/10.1007/s41651-020-00048-5)]
- 7 Paoletti ME, Haut JM, Plaza J, *et al.* Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 158: 279–317. [doi: [10.1016/j.isprsjprs.2019.09.006](https://doi.org/10.1016/j.isprsjprs.2019.09.006)]
- 8 Yu CY, Han R, Song MP, *et al.* A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 2485–2501. [doi: [10.1109/JSTARS.2020.2983224](https://doi.org/10.1109/JSTARS.2020.2983224)]
- 9 Hong DF, Gao LR, Yao J, *et al.* Graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(7): 5966–5978. [doi: [10.1109/TGRS.2020.3015157](https://doi.org/10.1109/TGRS.2020.3015157)]
- 10 Mou LC, Ghamisi P, Zhu XX. Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3639–3655. [doi: [10.1109/TGRS.2016.2636241](https://doi.org/10.1109/TGRS.2016.2636241)]
- 11 Hong DF, Han Z, Yao J, *et al.* SpectralFormer: Rethinking hyperspectral image classification with Transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5518615.
- 12 Jia S, Jiang SG, Lin ZJ, *et al.* A survey: Deep learning for hyperspectral image classification with few labeled samples. *Neurocomputing*, 2021, 448: 179–204. [doi: [10.1016/j.neucom.2021.03.035](https://doi.org/10.1016/j.neucom.2021.03.035)]
- 13 Xue ZH, Zhang MX, Liu YF, *et al.* Attention-based second-order pooling network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(11): 9600–9615. [doi: [10.1109/TGRS.2020.3048128](https://doi.org/10.1109/TGRS.2020.3048128)]
- 14 Zhang SH, Chen ZK, Wang D, *et al.* Cross-domain few-shot contrastive learning for hyperspectral images classification. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 5514505.
- 15 Manifold B, Men S, Hu RQ, *et al.* A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nature Machine Intelligence*, 2021, 3(4): 306–315. [doi: [10.1038/s42256-021-00309-y](https://doi.org/10.1038/s42256-021-00309-y)]
- 16 Cheng G, Cai LM, Lang CB, *et al.* SPNet: Siamese-prototype network for few-shot remote sensing image scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5608011.
- 17 Cao ZY, Li XR, Jiang JF, *et al.* 3D convolutional siamese network for few-shot hyperspectral classification. *Journal of Applied Remote Sensing*, 2020, 14(4): 048504.
- 18 Xue ZH, Zhou YY, Du PJ. S3Net: Spectral-spatial Siamese network for few-shot hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5531219.
- 19 Liu JX, Zhang KF, Wu SQ, *et al.* An investigation of a multidimensional CNN combined with an attention mechanism model to resolve small-sample problems in hyperspectral image classification. *Remote Sensing*, 2022, 14(3): 785. [doi: [10.3390/rs14030785](https://doi.org/10.3390/rs14030785)]
- 20 Feng F, Zhang YS, Zhang J, *et al.* Small sample hyperspectral image classification based on cascade fusion of mixed spatial-spectral features and second-order pooling. *Remote Sensing*, 2022, 14(3): 505. [doi: [10.3390/rs14030505](https://doi.org/10.3390/rs14030505)]
- 21 Zhong ZL, Li Y, Ma LF, *et al.* Spectral-spatial Transformer network for hyperspectral image classification: A factorized architecture search framework. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5514715.
- 22 Sun H, Zheng XT, Lu XQ, *et al.* Spectral-spatial attention network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(5): 3232–3245. [doi: [10.1109/TGRS.2019.2951160](https://doi.org/10.1109/TGRS.2019.2951160)]
- 23 Mei XG, Pan ET, Ma Y, *et al.* Spectral-spatial attention networks for hyperspectral image classification. *Remote Sensing*, 2019, 11(8): 963. [doi: [10.3390/rs11080963](https://doi.org/10.3390/rs11080963)]
- 24 Li MS, Liu YK, Xue GK, *et al.* Exploring the relationship between center and neighborhoods: Central vector oriented self-similarity network for hyperspectral image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(4): 1979–1993. [doi: [10.1109/TCSVT.2022.3218284](https://doi.org/10.1109/TCSVT.2022.3218284)]

(校对责编: 牛欣悦)