

基于深度特征提取残差网络的高光谱图像分类^{*}赵雪松¹ 付民² 刘雪峰¹

(1. 青岛科技大学自动化与电子工程学院 青岛 266061; 2. 中国海洋大学信息科学与工程学部电子工程学院 青岛 266100)

摘要: 深度学习由于其模块化设计和强大的特征提取能力,已成为高光谱图像分类的重要手段之一。然而,如何有效地提取更深层次的特征以及同时提高分析空间和光谱联合特征的能力仍是亟待解决的问题。针对这些问题,本文提出了一种深度特征提取的残差网络,该网络由两个关键部分组成:多级传递融合残差网络和空间-光谱多分辨率融合注意力残差网络。多级传递融合残差网络可以有效促进特征信息之间的相互作用,获得更深层次的特征。接着利用空间-光谱多分辨率融合注意力残差网络可以确保从高光谱数据中全面提取空间-光谱联合特征和多分辨率特征。为了验证其有效性,本文在 Indian Pines, Pavia University 和 Salinas Valley 三个高光谱数据集上对所提出方法的性能进行了评估,分类精度分别达到了 98.10%, 99.81% 和 99.94%。实验结果表明,与其他方法相比,该网络具有更好的泛化能力和分类性能。

关键词: 高光谱图像分类;残差网络;空间-光谱联合特征;多分辨率

中图分类号: TN919.82 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Hyperspectral image classification based on deep feature
extraction residual networkZhao Xuesong¹ Fu Min² Liu Xuefeng¹

(1. College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China;

2. College of Electronic Engineering, Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China)

Abstract: Deep learning has become one of the important tools for hyperspectral image classification due to its modular design and powerful feature extraction capability. However, effectively extracting deeper features and simultaneously improving the analysis of spatial and spectral joint features remains an urgent challenge. In response to these issues, a deep feature extraction residual network is proposed in this paper, composed of two key components: a multi-level transfer fusion residual network and a spatial-spectral multi-resolution fusion attention residual network. The multi-level transfer fusion residual network effectively promotes interaction between feature information to obtain deeper-level features. Subsequently, the spatial-spectral multi-resolution fusion attention residual network ensures comprehensive extraction of spatial-spectral joint features and multi-resolution features from hyperspectral data. To validate its effectiveness, the performance of the proposed method was evaluated on three hyperspectral datasets, Indian Pines, Pavia University, and Salinas Valley, achieving classification accuracies of 98.10%, 99.81%, and 99.94% respectively. Experimental results demonstrate that, compared to other methods, this network exhibits better generalization capability and classification performance.

Keywords: hyperspectral image classification; residual network; spatial-spectral joint features; multi-resolution

0 引言

高光谱图像(hyperspectral image, HSI)具有大量的光谱波段,在土地覆被识别^[1]、材料识别^[2]、工业无线电^[3]、目标检测^[4-5]和分类^[6]等领域有着广泛的应用。分类方法作

为从高光谱数据中提取丰富信息的有效手段,近年来引起了广泛关注。

在过去十年中,由于卷积神经网络(convolutional neural network, CNN)^[7]的结构更适合处理高光谱图像等张量数据,因此在深度光谱空间特征的提取和分类方面取

得了显著进展。例如, HybridSN^[8] 通过将三维(three dimensional, 3D)卷积与二维(2D)卷积相结合, 实现了对高光谱图像中光谱和空间特征的分层提取。此外, 针对 HybridSN 中忽略多尺度空间特征的问题, Ge 等^[9] 提出了一种基于 2D-3D CNN 和多分支特征融合(2D-3D CNN and multibranch feature fusion, 2D3D-MBFF)的方法以弥补这一局限。Li 等^[10] 利用双分支方法独立提取光谱和空间特征, 并通过组合这些特征进行分类。

随着网络深度的增加, 梯度消失和爆炸影响了对特征间依赖关系的捕捉能力。幸运的是, 残差网络^[11] 通过将残差块作为基本单元, 有效缓解了由于网络深度增加而导致的精度下降问题。因此, 研究人员逐步将残差网络融入深度学习模型中来改善高光谱图像的分类性能。Zhong 等^[12] 利用光谱空间残差网络(spectral-spatial residual network, SSRN), 通过连续连接光谱和空间特征提取块, 构建了深度 CNN 模型, 并取得了出色的效果。此外, Zhu 等^[13] 介绍了一种残差光谱空间注意力网络(residual spectral spatial attention network, RSSAN), 将光谱空间注意力层融入到 SSRN 中。张祥东等^[14] 提出了一种多尺度残差网络分类模型, 通过在残差模块中添加分支结构, 分别提取光谱和空间特征, 实现了高光谱图像的多尺度信息融合。为了更全面地捕捉空间-光谱特征, Roy 等^[15] 设计了一种基于注意力的自适应光谱-空间核改进残差网络(adaptive spectral-spatial kernel improved residual network, A2S2K-ResNet)用于高光谱图像分类。颜铭靖等^[16] 提出了一种基于 3D 空洞卷积残差网络的高光谱影像分类方法, 通过 3D 卷积核提取空间和光谱特征, 并利用空洞结构提高感受野以增强分类精度。韩彦岭等^[17] 提出了一种基于残差网络的高光谱图像分类方法, 通过反卷积算法实现多尺度特征融合, 挖掘深度特征信息。随后, Zhan 等^[18] 提出了一种改进的光谱-空间残差注意力网络(enhanced spectral-spatial residual attention network, ESSRAN), 该网络能够有效提取光谱和空间特征。吕欢欢等^[19] 提出了一种高光谱影像分类方法, 结合边缘保护滤波和深度残差网络, 通过 2D 卷积神经网络和跳层连接融合空间-光谱特征以完成分类。程谔杰等^[20] 提出了一种深度可分离卷积的轻量化残差网络模型, 利用 3D 卷积层进行初步特征提取, 并通过 3D 深度可分离卷积残差层增强高维空间特征的表达能力。Wu 等^[21] 提出了一种双分支多尺度空间-光谱残差注意力网络(dual-branch multi-scale spatial-spectral residual attention network, MSRAN)。该网络通过整合多尺度特征提取和残差注意力机制, 能够独立提取空间和光谱特征, 从而减少了特征之间的干扰。然而, 这些方法在提取高光谱图像中丰富的空间-光谱联合特征信息时仍然存在一些局限性。一方面, 它们未能充分利用高光谱图像中的多分辨率特征, 导致特征提取的细节信息丢失。另一方面, 这些方法在应对复杂的高光谱图像结

构时, 对空间和光谱联合特征间相关性的挖掘不够深入, 导致模型在处理复杂高光谱数据时的泛化能力和分类效果受到限制, 无法有效捕捉和利用不同特征之间的微妙关系。

针对上述问题, 本文提出了一种基于深度特征提取残差网络的高光谱图像分类方法。首先, 设计了一个多级传递融合残差网络(multi-level transfer fusion residual network, MTF-ResNet), 以残差理念为核心, 专注于提取深层空间-光谱特征。然后, 空间-光谱多分辨率融合注意力残差网络(spatial-spectral multi-resolution fusion attention residual network, S²MRFA-ResNet)通过其核心模块 S²MRFA, 从高光谱数据中全面提取和分析空间-光谱联合特征及多分辨率特征。此外, S²MRFA 中还设计了一个深度可分离膨胀空间-光谱特征提取模块(depthwise separable dilated spatial-spectral feature extraction module, DSDS²FEM), 以增强空间-光谱特征的相关性。最终, 采用自适应加权求和方法将多个特征图融合, 并利用 softmax 激活函数得到最终的特征表示。

1 深度特征提取残差网络

1.1 深度特征提取残差网络的总体结构

所提方法的整体架构如图 1 所示, 主要包括 MTF-ResNet 和 S²MRFA-ResNet。首先, MTF-ResNet 由 3 个不同的多尺度残差块组成, 这些残差块通过前向传播交互网络进行组合, 既能提取和整合图像中的深层特征语义信息, 又可以丰富浅层特征细节。其次, S²MRFA-ResNet 的双分支结构自适应地捕捉和处理高光谱图像的空间-光谱联合特征及多分辨率特征。然后, 将得到的特征图传递到包含三条多尺度残差块的分支中。在各分支中, 特征图依次经过卷积、最大池化操作和全连接层处理后, 通过自适应分数加权求和机制进行合并。最后, 利用 softmax 激活函数对结果求和并进行归一化, 得到最终的分类输出。

1.2 多级传递融合残差网络(MTF-ResNet)

如图 1 所示, MTF-ResNet 中的残差块有助于保持网络深度和复杂度之间的平衡。图 2(a)中的普通残差块利用“跳过连接”特性, 有效缓解了梯度消失或爆炸的问题。在此基础上, 本文提出了图 2(b)中列出的 3 个多尺度残差块, 以提取不同尺度和深度的特征。残差块 B 采用跨多个维度的 2D 卷积核, 有助于提取深度特征。残块 C 则采用不同的卷积核大小和排列方式, 从多个视角整合特征。残块 A 结合了 2D 卷积和深度可分离卷积(depthwise separable convolution, DSC)^[22], 通过生成多样化特征并显著减少计算复杂度, 为后续残块的信息传递奠定了基础, 从而提升网络在不同尺度下特征提取的能力。

在 MTF-ResNet 中, 残差块 A、B 和 C 之间的信息传递是通过特征交互来实现的。具体而言, 残差块 A 的输出被分为 4 个部分, 这些部分分别与残差块 B 和 C 进行交互传递。此外, 残差块 B 的输出也被分为两部分, 用于不同残差

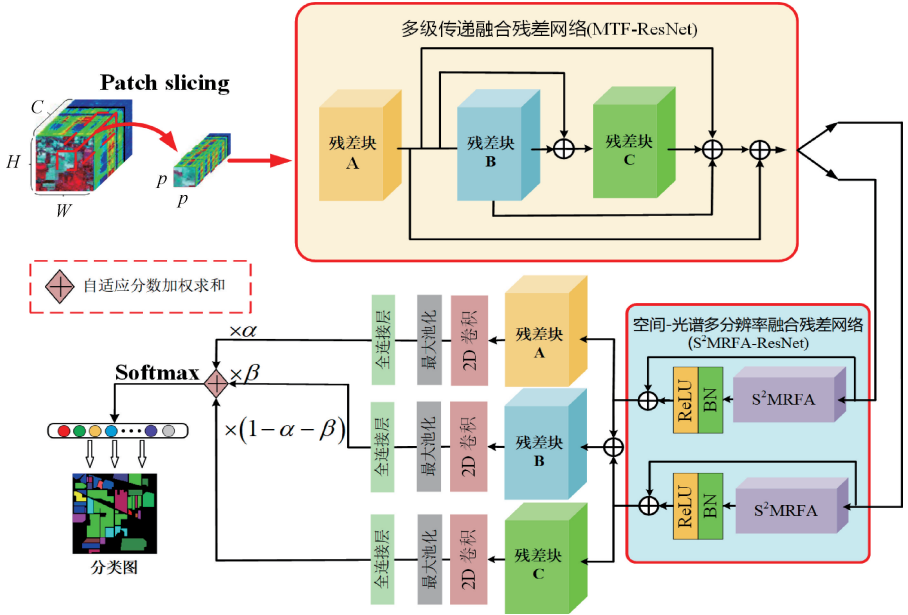


图 1 总体网络架构

Fig. 1 Overall network architecture

块之间的信息传递。随后,这些特征信息与残差块 C 的输出进行求和。最后,将残差块 A 的原始输出与求和后的结果进行融合。这个过程利用了残差块 A 和 B 的输出,保留了包含空间信息的丰富浅层特征,为更高层次的模型提供了准确的特征基础。同时,残差块 C 的输出用于捕捉深层次的语义信息,如全局上下文信息。通过在不同深度级别的特征之间建立联系,分散的特征信息得以聚合,形成具有丰富上下文内容的特征表示。

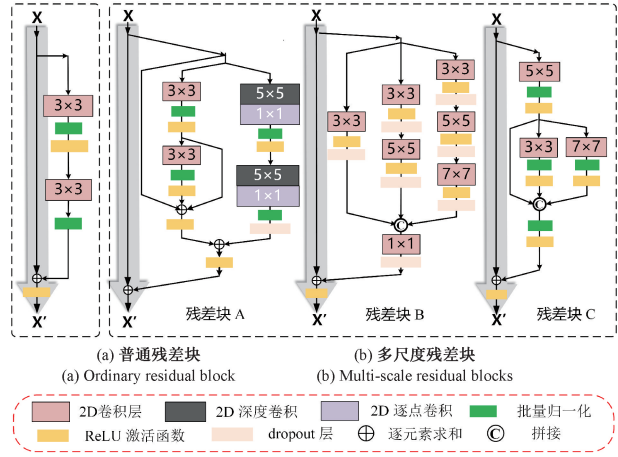


图 2 各种残差块

Fig. 2 The various residual blocks

1.3 空间-光谱多分辨率融合注意力(S²MRFA)

考虑到高光谱图像在不同分辨率下的空间光谱联合特征可能存在差异,并受到深度学习模型在特征动态调整和注意力权重分配方面成功应用^[23]的启发,本文提出了 S²MRFA 模块。如图 3 所示, S²MRFA 模块由自适应空

间-光谱注意力(adaptive spatial-spectral attention, ASSA)和自适应多分辨率注意力(adaptive multi-resolution attention, AMRA)两部分组成。ASSA 通过根据输入特征图中特征的重要性动态调整空间和光谱注意力权重,能够有效提取高光谱图像中的重要空间-光谱联合特征信息。AMRA 则在不同分辨率下调整其处理策略,以自适应地捕获多个尺度的空间-光谱变化特征,实现对空间-光谱复杂特性的处理。最终,ASSA 和 AMRA 的输出通过自适应加权参数 λ 进行加权融合,并利用自适应最大池化层获得最终的特征表示。为了进一步提高网络对空间-光谱联合特征的利用能力,本文针对 S²MRFA 模块提出了深度可分离膨胀空间-光谱特征提取模块(depthwise separable dilated spatial-spectral feature extraction module, DSDS²FEM)并应用在 S²MRFA 中。

1) 自适应空间-光谱注意力(ASSA)

如图 3 所示,在 ASSA 中,空间和光谱特征取通过两条通道进行。首先,每个通道中的特征图经过 DSDS²FEM 的初始化处理,以进行空间和光谱特征的初步提取。随后,根据各自特征的特性,采用不同的归一化策略:空间特征提取通道使用批量归一化(batch normalization, BN),以处理其平滑和一致变化的特性;而光谱特征提取通道则采用组归一化(group normalization, GN),以应对高光谱图像中输入通道之间的变化和均匀分布特性。接着,对两条通道进行卷积操作,其中空间特征提取通道使用 7×7 的 2D 卷积核,以覆盖更大的上下文视野并获取更丰富的特征信息;光谱特征提取通道则使用 3×3 的 2D 卷积核,更好地捕捉局部特征。激活函数之后,引入残差连接,使网络能够学习复杂特征同时保留原始数据的属性。最后,通

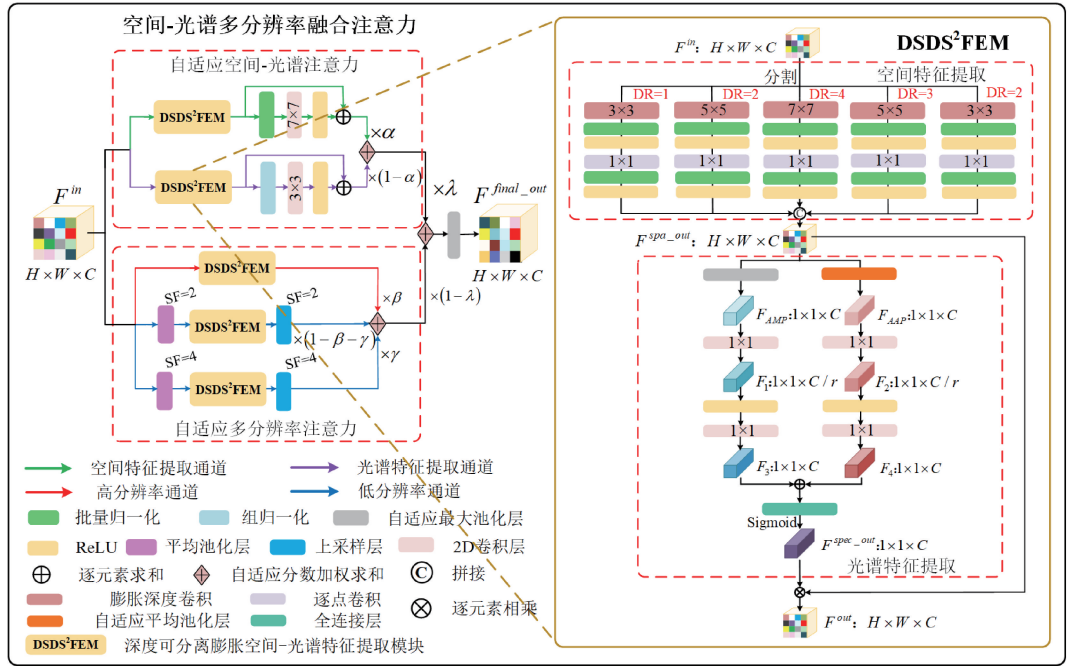


图 3 空间-光谱多分辨率融合注意力

Fig. 3 Spatial-spectral multi-resolution fusion attention

过由参数 α 引导的自适应分数加权求和机制,优化整合两个通道的空间和光谱信息。

2) 自适应多分辨率注意力(AMRA)

如图 3 所示,AMRA 包括 3 个通道:一个处理高分辨率数据,两个处理低分辨率数据。高分辨率通道中只应用 DSDS²FEM 来处理输入特征图,以保留原始图像的高分辨率细节和清晰度。而对于低分辨率通道,首先通过平均池化层降低输入特征图的分辨率,这一步骤在降采样过程中保留了关键图像信息,扩展了感受野,从而捕捉了更广泛的上下文信息。其次,降采样后的数据经过 DSDS²FEM 的处理,再通过上采样层恢复到原始分辨率。其中,第 1 个低分辨率通道的平均池化和上采样层的采样因子(sampling factor, SF)为 2,第 2 个低分辨率通道的 SF 为 4。通过采用不同的 SF,可以有效地处理不同大小和分辨率的特征。上述过程中,直接对高分辨率数据进行 DSDS²FEM 处理有助于避免引入伪影或错误。尽管在处理前使用上采样操作可以提高图像的分辨率,但插值算法可能无法准确重建原始图像的细节,从而引入虚假的特征或平滑效果,导致不必要的误差或伪影。因此,在高分辨率通道中直接处理原始数据更能保持图像的准确性和清晰度。最终,3 条通道的输出特征通过由可学习参数 β 和 γ 引导的自适应分数加权求和机制,优化组合不同分辨率下的特征图,提升了整体模型对多分辨率图像信息的处理能力。

3) 深度可分离膨胀空间-光谱特征提取模块(DSDS²FEM)

基于膨胀卷积^[24]能够扩展的感受野的特性,本文提出

了 DSDS²FEM,如图 3 所示。DSDS²FEM 由两个主要模块组成:空间特征提取模块和光谱特征提取模块。空间特征提取模块的核心思想是将膨胀操作整合到 DSC 中,本文将这一过程称为“深度可分离膨胀卷积”(depthwise separable dilated convolution, DSDC)。DSC 包括深度卷积和逐点卷积两个步骤,因此在 DSDC 中,深度卷积被扩展为膨胀深度卷积,而逐点卷积保持不变。这种整合不仅增加了感受野的大小,还保持了网络的分辨率,能够捕获不同空间尺度的特征,从而实现对图像结构和纹理信息的更全面理解。光谱特征提取模块则从前一阶段获得的多尺度空间特征图中提取光谱特征,并在光谱维度上生成光谱权重向量,将更多的权重分配给重要的通道以提高分类性能。

如图 3 所示,空间特征提取阶段,将输入特征图定义为 $F_{in} \in \mathbb{R}^{H \times W \times C}$,其中 H, W 和 C 分别表示高度、宽度和光谱通道数。首先,将输入特征图沿通道维度划分为五个大小相等的子集,表示为 $F_{in}^i \in \mathbb{R}^{H \times W \times C/5}$, $i = 1, 2, 3, 4, 5$,由此每个子集的卷积参数将减少到原特征提取块的 1/5。接着,在这五组通道中分别应用相同数量的 DSDC 来提取特征,每组的卷积核尺寸为 $3 \times 3, 5 \times 5, 7 \times 7, 5 \times 5$ 和 3×3 ,对应的膨胀率(dilation rate, DR)分别为 1、2、4、3 和 2。这使得每组通道得到的扩张感受野分别为 $3 \times 3, 9 \times 9, 25 \times 25, 13 \times 13$ 和 5×5 。这样,在保持参数量不变的情况下,感受野在不同尺度上有所变化,帮助网络在不同尺度下捕获特征。每组 DSDC 处理后使用 BN 和 ReLU 激活函数。最后,将五组通道得到的特征图进行拼接,得到最终包含重

要的空间位置信息的多尺度空间特征图 $\mathbf{F}^{spa_out} \in \mathbb{R}^{H \times W \times C}$ 。计算过程如式(1)和式(2)所示。

$$f_{n \times n, d}(\mathbf{F}_i^{in}) = \delta(W_{n, d} * \mathbf{F}_i^{in} + b_{n, d}) \quad (1)$$

$$\mathbf{F}^{spa_out} = [f_{3 \times 3, 1}(\mathbf{F}_1^{in}), f_{5 \times 5, 2}(\mathbf{F}_2^{in}), f_{7 \times 7, 4}(\mathbf{F}_3^{in}), f_{5 \times 5, 3}(\mathbf{F}_4^{in}), f_{3 \times 3, 2}(\mathbf{F}_5^{in})] \quad (2)$$

式中: $W_{n, d}$ 和 $b_{n, d}$ 分别代表与卷积大小 n ($n = \{3, 5, 7\}$) 和膨胀率 d ($d = \{1, 2, 3, 4\}$) 相对应的权重和偏置参数。 δ 表示 ReLU 激活函数, $[\cdot]$ 表示拼接。

如图 3 所示, 在光谱特征提取阶段, 首先对多尺度空间特征图 \mathbf{F}^{spa_out} 分别进行自适应最大池化和自适应平均池化处理。这一过程计算输入特征图在空间维度上的最大值和平均值作为相应光谱维度的权重。接下来, 为进一步细化光谱特征, 对得到的两个光谱信息通道分别进行 1×1 的 2D 卷积操作, 并将缩减比 (reduction ratio, r) 设置为 2, 从而将特征图通道数减少到 C/r , 降低模型的复杂性和计算要求。随后, 两条通道通过 ReLU 激活函数并再次利用 1×1 的 2D 卷积将通道数再恢复为 C , 以增加特征的表达能。最后, 将处理后的特征图相加并送入全连接层, 以增强特征的非线性表示能力, 并使用 sigmoid 函数进行归一化处理。计算过程如式(3)~(5)所示。

$$\mathbf{F}_3 = f_{1 \times 1}(\delta(f_{1 \times 1}(\text{AdpMaxPool}(\mathbf{F}^{spa_out})))) \quad (3)$$

$$\mathbf{F}_4 = f_{1 \times 1}(\delta(f_{1 \times 1}(\text{AdpAvgPool}(\mathbf{F}^{spa_out})))) \quad (4)$$

$$\mathbf{F}^{spec_out} = \sigma(\text{FC}(\mathbf{F}_3 \oplus \mathbf{F}_4)) \quad (5)$$

式中: AdpMaxPool 和 AdpAvgPool 分别代表自适应最大池化和自适应平均池化。其中 σ 代表 Sigmoid 激活函数, FC 代表全连接层, \oplus 代表逐元素相加。

最终, 将特征图 \mathbf{F}^{spec_out} 和多尺度空间特征图 \mathbf{F}^{spa_out} 相乘得到最终特征 $\mathbf{F}^{out} \in \mathbb{R}^{H \times W \times C}$ 。计算过程如式(6)所示。

$$\mathbf{F}^{out} = \mathbf{F}^{spec_out} \otimes \mathbf{F}^{spa_out} \quad (6)$$

式中: \otimes 表示逐元素相乘。由此得到的特征图能够增强对空间和光谱特征之间相互作用和相关性的理解, 为后续空间-光谱联合特征提取过程提供了有力支持。

2 实验结果与分析

2.1 实验数据集

为了研究和评估分类算法的性能, 本文选取了 3 个著名的开源数据集: Indian Pines(IP), Pavia University(PU) 和 Salinas Valley(SV)。IP 数据集由机载可见红外成像光谱仪于 1992 年对美国印第安纳州西北部的松树区进行成像获得, 包含 200 个光谱波段, 尺寸为 145×145 , 总像素数为 21 025 个。其中, 地物像素为 10 249 个, 覆盖 16 种不同的地物类型, 背景像素为 10 776 个。PU 数据集来自德国的反射光学光谱成像系统, 于 2003 年对意大利帕维亚市部分区域进行成像, 涵盖 103 个光谱波段, 尺寸为 610×340 , 总像素数为 2 207 400 个, 其中地物像素为 42 776 个, 分为 9 类。SV 数据集由机载可见红外成像光谱仪于美国

加利福尼亚州 Salinas 山谷捕捉, 包含 204 个光谱波段, 尺寸为 512×217 , 总像素数为 111 104 个。其中背景像素为 56 975 个, 分类用像素为 54 129 个, 分为 16 类。

本文从 IP 和 PU 数据集中随机选取 10% 的像素作为训练集, 其余 90% 的像素作为测试集。从数据集中随机选取 20% 的像素作为训练集, 其余 80% 的像素作为测试集。

2.2 实验配置

为了验证所提方法的性能, 本文在 3 个数据集上进行了系列实验。所有实验均在配备 Nvidia GeForce RTX 2080 Ti 12GB GPU 的环境下进行, 使用 Python 3.7.16 和 PyTorch 1.6 作为编译环境。在训练阶段, 将批处理大小设置为 16, 迭代次数为 100, 初始学习率为 0.001, 并使用 Adam 优化器进行学习率的优化和模型的优化。反向传播过程中采用交叉熵损失函数。本文采用总体精度 (overall accuracy, OA)、平均精度 (average accuracy, AA) 和 kappa 系数 (Kappa) 3 项常见的定量指标来衡量分类性能。其中, OA 表示正确分类的像素数与总的像素数的比例。AA 表示所有类别的平均精度。Kappa 系数是衡量分类结果与相应地面真实值 (ground truth, GT) 之间一致性水平的分数。为避免偏差估计, 实验进行了 10 次独立检验, 并取所有评价指标的平均值。

2.3 对比实验分析

为了验证所提模型的性能和泛化能力, 本文在 IP、PU 和 SV 三个高光谱数据集上与一系列典型的方法进行了对比实验, 包括 ResNet^[11]、SSRN^[12]、HybridSN^[8]、A2S2K-ResNet^[15]、ESSRAN^[18]、MSRAN^[21]、SSATN^[25] 和 S2EFT^[26]。这些方法在 IP、PU 和 SV 数据集上的实验结果如表 1~3 所示, 分类结果图如图 4~6 所示。

1) IP 数据集

如表 1 所示, 本文提出的方法展示了良好的分类性能, 其中 OA、AA 和 Kappa 分别达到了 98.10%、98.02% 和 98.56%。相比于 ResNet、SSRN、HybridSN、A2S2K-ResNet、ESSRAN、MSRAN、SSATN 和 S2EFT, 所提方法在 OA 值上分别提高了 19.74%、8.25%、7.76%、1.93%、1.96%、0.86%、0.44% 和 0.16%。此外, 该方法对所有类别分类的准确率都高于 94%, 表明其在不平衡数据集上达到了较好的精度平衡效果。图 4 展示了这些方法在 IP 上的分类效果。从图中可以看出, 其他方法在处理具有区域边缘和狭窄形状的土地覆盖物时表现不佳, 而本文所提方法在每个区域生成了更为精细准确的分类图, 显著减少了噪声点, 提升了视觉效果。

2) PU 数据集

如表 2 所示, 与其他方法相比, 本文提出的方法在大部分类别中获得了最佳的分类效果。所提方法在 OA、AA 和 Kappa 上分别达到了 99.81%、99.70% 和 99.75%。在这些指标上, OA 值分别比 ResNet、SSRN、HybridSN、A2S2K-ResNet、ESSRAN、MSRAN、SSATN 和 S2EFT 高

表 1 IP 数据集的分类结果
Table 1 Classification results of IP dataset

类别	ResNet	SSRN	HybridSN	A2S2K-ResNet	ESSRAN	MSRAN	SSATN	S2EFT	本文方法
1	82.78	86.96	88.69	94.22	95.51	92.17	95.89	97.18	99.23
2	75.54	89.65	88.06	96.58	96.87	93.96	96.54	97.68	98.83
3	68.26	84.12	87.98	94.01	94.58	98.91	97.56	98.72	98.36
4	72.85	82.08	91.36	92.51	96.59	96.48	98.16	99.06	99.49
5	84.16	82.83	91.28	92.60	92.48	98.79	97.80	96.78	94.65
6	68.96	83.79	85.76	94.57	97.60	96.98	95.37	96.89	96.01
7	75.55	86.54	81.94	95.65	94.59	95.68	98.13	99.56	99.59
8	81.29	88.17	91.56	94.62	92.59	97.63	97.95	98.59	98.97
9	60.39	80.25	86.45	92.47	90.39	89.69	96.97	97.10	97.35
10	78.42	90.28	91.46	95.38	94.56	92.56	93.59	94.52	95.69
11	68.66	91.69	92.58	94.18	96.87	98.07	98.19	97.64	98.27
12	78.25	87.58	89.35	95.40	95.15	98.26	98.23	98.91	96.85
13	80.47	89.38	90.75	93.45	97.89	96.39	98.17	99.41	99.53
14	79.38	91.71	90.56	96.96	97.38	97.25	96.59	98.18	98.82
15	73.36	85.14	91.19	93.47	97.15	96.91	97.87	96.19	97.96
16	78.17	87.35	86.94	93.01	95.18	98.28	97.95	96.37	98.69
OA	78.36	89.85	90.34	96.17	96.14	97.24	97.66	97.94	98.10
AA	75.41	86.72	89.12	94.32	95.34	96.13	97.19	97.67	98.02
Kappa	77.39	89.63	91.34	96.15	96.78	96.25	97.82	98.40	98.56

表 2 PU 数据集的分类结果
Table 2 Classification results of PU dataset

类别	ResNet	SSRN	HybridSN	A2S2K-ResNet	ESSRAN	MSRAN	SSATN	S2EFT	本文方法
1	85.37	92.89	94.94	95.89	97.66	92.11	97.95	99.14	99.79
2	86.01	91.45	93.45	96.22	96.45	99.19	95.77	98.79	99.96
3	75.39	82.35	88.65	96.91	97.75	98.57	98.87	99.13	99.47
4	86.95	88.96	91.23	95.89	96.45	97.35	98.27	98.78	99.59
5	86.24	90.29	95.45	94.73	98.23	98.41	98.09	98.81	100.00
6	90.68	92.56	96.14	96.78	97.52	97.79	98.54	98.29	100.00
7	86.45	85.69	91.45	97.46	91.40	96.92	98.67	99.22	99.81
8	84.65	90.78	92.87	94.85	97.35	97.49	98.85	99.34	99.23
9	87.66	91.02	93.65	96.51	93.25	96.23	97.92	99.11	99.47
OA	88.54	91.37	94.34	96.82	97.19	98.05	98.99	99.28	99.81
AA	85.49	89.55	93.09	96.14	96.23	97.12	98.10	98.96	99.70
Kappa	89.05	92.21	95.13	97.05	97.26	98.30	99.11	99.15	99.75

出 11.27%、8.44%、5.47%、2.99%、2.62%、1.76%、0.82%和 0.53%。此外,所提出的方法在所有类别中有 8 个类别的分类精度最高,其中 2 个类别的准确率达到 100%,证明了该方法的有效性和稳定性。图 5 展示了 PU 数据集上各种方法的分类图。从图中可以看出,本文所提方法得到的分类图更加干净,减少了错误分类的像素,仅在个别类别中存在少量的错误分类,例如“Self-Blocking

Bricks”(第 8 类),其准确率仅比最高值低 0.11%。相比之下,其他方法的分类性能相对较差,如 ESSRAN 在“Bitumen”(第 7 类)中存在较为明显的分类错误。总体而言,本文所提方法的最终分类效果与 GT 非常接近。

3)SV 数据集

如表 3 所示,SV 数据集上的每种方法都达到了较高的分类精度,这是由于 SV 数据集在样本平衡性、空间同质

表 3 SV 数据集的分类结果
Table 3 Classification results of SV dataset

类别	ResNet	SSRN	HybridSN	A2S2K-ResNet	ESSRAN	MSRAN	SSATN	S2EFT	本文方法
1	71.37	96.73	90.69	98.59	95.75	98.79	92.65	99.84	100.00
2	96.74	97.17	97.67	97.54	99.31	98.67	99.46	99.89	100.00
3	68.77	97.09	98.56	94.15	88.48	97.07	99.91	98.92	100.00
4	87.84	94.29	93.31	95.94	98.18	96.16	99.74	99.78	100.00
5	88.92	96.35	96.77	97.01	98.69	95.22	98.40	99.81	99.95
6	93.05	95.25	95.68	94.75	97.56	98.87	98.27	97.85	100.00
7	87.31	97.52	98.24	97.84	95.47	99.17	99.22	99.86	99.76
8	86.71	82.11	97.22	92.59	95.96	99.10	98.18	98.53	100.00
9	96.53	93.17	95.01	92.18	98.78	100.00	98.23	100.00	100.00
10	81.19	94.27	96.38	97.45	97.23	99.31	99.05	98.93	99.96
11	88.07	96.61	97.79	98.56	98.44	98.38	98.81	99.32	100.00
12	90.59	93.53	95.64	97.89	98.39	97.51	98.99	98.76	98.85
13	89.25	95.83	97.55	97.14	99.28	98.27	99.14	99.16	100.00
14	94.99	93.26	94.18	96.63	98.91	98.42	99.35	99.11	100.00
15	95.17	94.49	83.31	98.47	98.40	95.49	98.26	98.45	100.00
16	96.79	97.74	94.86	93.98	92.65	100.00	100.00	99.58	100.00
OA	90.62	94.78	95.68	96.98	97.86	98.97	99.43	99.67	99.94
AA	88.33	94.71	95.18	96.29	96.97	98.15	98.60	99.24	99.91
Kappa	91.33	94.92	95.34	97.12	97.76	99.01	99.36	99.56	99.93

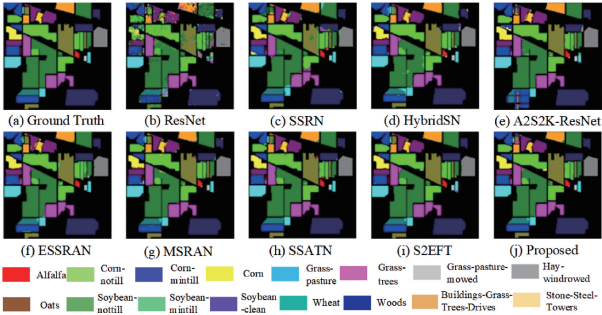


图 4 IP 数据集的分类图
Fig. 4 Classification maps of IP dataset

性和数据质量方面相较于 IP 和 PU 数据集具备更大优势。具体而言,本文提出的方法在分类性能上表现最佳,其 OA、AA 和 Kappa 分别为 99.94%、99.91% 和 99.93%。这些指标中,OA 值分别比 ResNet、SSRN、HybridSN、A2S2K-ResNet、ESSRAN、MSRAN、SSATN 和 S2EFT 高出 9.32%、5.16%、4.26%、2.96%、2.08%、0.97%、0.51% 和 0.27%。从表可以看出,所提方法在在所有类别中有 14 种类别的分类精度是最高的,且其中 12 个类别的准确率达到 了 100%,充分说明了该方法的分类准确性。此外,“Grapes_untrained”(第 8 类)和“Vinyard_untrained”(第 15 类)是非常相似的地面物体,多数方法在这两类上的分类效果不佳,但所提方法在这两类的分类准确率都达到了 100%。这表明,所提网络结构能够有效地提取鉴别

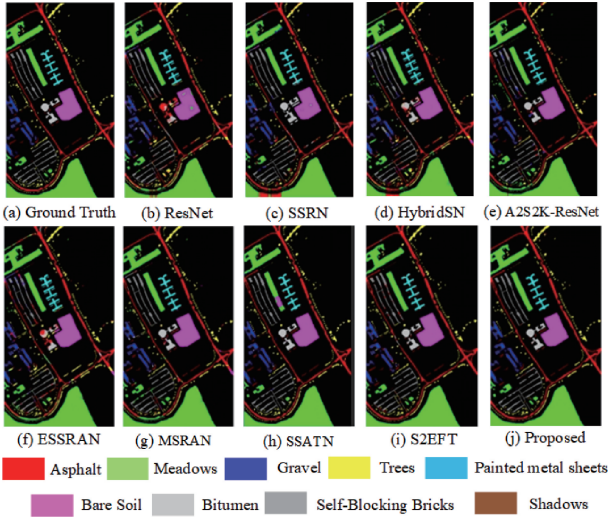


图 5 PU 数据集的分类图
Fig. 5 Classification maps of PU dataset

特征,并识别相似的地面物体。图 6 展示了 SV 数据集上各方法的分类图。可以看出,本文所提方法中第 8 类和第 15 类中的噪声点较少,边界更加平滑,错误分类明显减少,得到的分类图更加清晰准确。

总体而言,通过对 3 个数据集的定量和定性分析,本文提出的网络在所有数据集上均取得了较好的分类效果,在不同场景和条件下表现出了良好的泛化能力。相比之

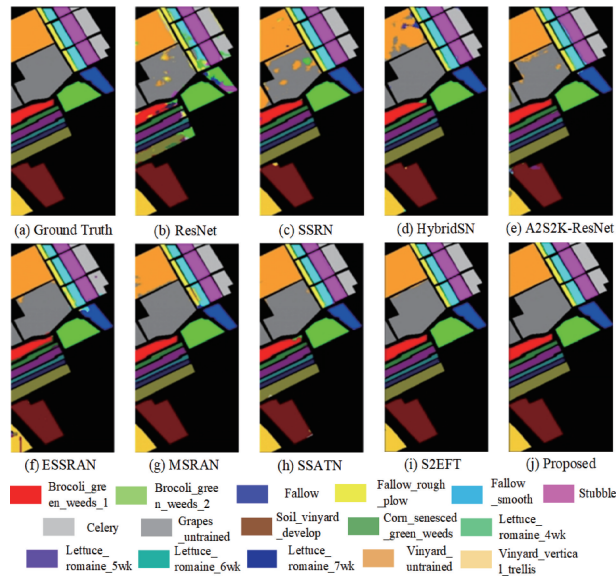


图 6 SV 数据集的分类图
Fig. 6 Classification maps of SV dataset

下,其他方法在不同数据集上的分类效果较为不稳定,泛化能力也相对较差。本文所提模型在分类效果和泛化能力上表现出色原因可以归纳为以下两点:首先,所提网络架构对高光谱图像分类任务具有良好的适应性。具体而言,残差结构通过将原始特征与转换特征相结合,促进了更全面的特征表示,不仅增强了特征表达能力,还有效降低了网络的训练复杂度和过拟合风险。同时,MTF-ResNet 和 S²MRFA-ResNet 模块的顺序应用,使网络能够

逐层深入地捕捉和利用土地覆被的深层空间-光谱联合特征,最终将其转化为可区分的地物分类信息,从而提升模型在多个数据集上的泛化能力。其次,S²MRFA-ResNet 中的 DSDS²FEM 增强了空间-光谱特征之间的相关性,为 S²MRFA 模块提取高光谱图像中重要的空间-光谱多分辨率特征提供了更为丰富准确的特征信息源。这种多分辨率特征的利用不仅提高了模型的分类精度,还增强了其在不同场景和条件下的适应性,进一步提高了所提模型在不同数据集上的泛化能力。

2.4 性能分析

为了评估本文所提模型的复杂度和效率,本节分析了在 IP、PU 和 SV 数据集上不同方法的训练、测试时间和参数量,如表 4 所示。表 4 显示,模型参数量与模型操作过程中的时间消耗并没有直接关系,因为卷积分组、深度可分性以及计算机设备等因素都会影响运行速度。因此,尽管本文所提模型的参数量低于 SSRN 和 HybridSN,但其时间消耗却高于这两者。在 3 个数据集上,ESSRAN 所需的时间最长,这主要是由于其使用了长短期记忆网络(long short-term memory network,LSTM)结构。相比之下,所提模型在训练、测试时间和参数量上表现适中。该模型通过增加网络的宽度和深度,并引入 S²MRFA-ResNet,导致时间消耗有所增加,但这些改进结构能够更全面地提取高光谱数据中的空间和光谱联合特征。与其他方法在不同数据集上的表现相比,该模型在 3 个数据集上展现了最佳和最稳定的性能。因此,综合考虑效率、准确性和泛化能力,该模型的时间消耗是可以接受的。

表 4 不同方法的性能分析

Table 4 Performance analysis of different methods

数据集	时间/s	ResNet	SSRN	HybridSN	A2S2K-ResNet	ESSRAN	MSRAN	SSATN	S2EFT	本文方法
IP	训练	188.25	402.58	213.59	452.78	570.34	537.22	325.97	495.34	479.24
	测试	15.43	19.52	11.21	20.37	40.61	33.69	15.26	24.75	21.72
PU	训练	262.92	392.37	246.62	582.36	746.45	710.66	465.32	615.53	599.70
	测试	21.24	28.07	23.89	48.66	78.28	58.70	30.91	56.08	50.24
SV	训练	418.68	628.14	309.91	694.29	924.06	856.19	588.01	804.08	785.48
	测试	39.03	54.75	35.38	68.07	102.57	79.32	41.57	75.62	71.45
参数(×10 ⁴)		154.45	365.34	453.70	37.08	294.41	307.22	125.63	433.18	327.15

2.5 消融实验

为了验证 MTF-ResNet 和 S²MRFA-ResNet 的有效性,本文在 3 个数据集上进行了系列实验,结果如表 5 所示。从表中不难看出,缺少 MTF-ResNet 和 S²MRFA-ResNet 模块的模型在所有数据集上均表现得最差,而结合这两个模块的模型达到了最佳的分类效果。相比于前者,后者在 3 个数据集上的 OA 值分别提高了 12.74%、10.36% 和 8.67%。由此可以推断,每个组件都对分类效果具有一定的贡献。以 IP 数据集为例,MTF-ResNet 的引

入使网络的 OA 值达到了 91.67%,而在此基础上加入 S²MRFA-ResNet 后,OA 值进一步提高到 98.10%。此外,与 MTF-ResNet 相比,S²MRFA-ResNet 对数据集的分类性能提升效果更为显著。本文推测这主要是由于 S²MRFA-ResNet 包含的 DSDS²FEM 能够更加有效地捕捉和提取高光谱图像中细微的空间-光谱联合特征。这些特征在分类过程中起到了关键作用,尤其在区分具有相似光谱特征但不同空间分布的地物类型时表现突出。

表 5 消融实验

Table 5 Ablation experiments

%

类别	模块组件		数据集		
	MTF-ResNet	S ² MRFA-ResNet	IP	PU	SV
OA			85.36	89.45	91.27
	✓		91.67	92.11	95.06
		✓	93.82	95.37	97.44
	✓	✓	98.10	99.81	99.94
AA			84.09	90.13	92.35
	✓		90.18	92.07	94.51
		✓	91.99	94.22	96.83
	✓	✓	98.02	99.70	99.91
Kappa			95.84	90.66	92.45
	✓		92.34	91.79	95.08
		✓	92.69	95.93	97.26
	✓	✓	98.56	99.75	99.93

3 结 论

本文提出了一种针对高光谱图像分类的深度特征提取残差网络,旨在解决空间-光谱联合特征信息提取不充分的问题。该网络模型主要结合了 MTF-ResNet 和包含 DSDS²FEM 的 S²MRFA-ResNet。首先,通过由 3 种残差块构成的 MTF-ResNet,实现了对高光谱图像的浅层细节特征和深层语义特征的有效提取与融合。随后,引入 S²MRFA-ResNet 模块,从不同分辨率上提取并整合高光谱图像的空间-光谱特性,确保在多种解析度视野下充分捕获和利用空间与光谱的联合特征。为了验证所提出方法的有效性,本文在公开的 IP、PU 和 SV 高光谱数据集上进行了实验,并与多种先进方法进行了对比。结果表明,本文的模型在所有 3 个数据集上都显著提高了分类精度,同时展现了良好的泛化能力,在利用高光谱图像中空间-谱联合特征及多分率特征等方面具有潜在优势。未来的工作将重点研究基于残差网络的深度学习模型,并优化其复杂度。

参考文献

[1] ZHANG X R, SHANG SH W, TANG X, et al. Spectral partitioning residual network with spatial attention mechanism for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-14.

[2] 王敏, 陈金勇, 孙康, 等. 基于神经网络的高光谱图像异常目标检测方法[J]. 国外电子测量技术, 2019, 38(1): 96-101.

WANG M, CHEN J Y, SUN K, et al. Anomaly detection for hyperspectral imagery based on neural

network [J]. Foreign Electronic Measurement Technology, 2019, 38(1): 96-101.

[3] 田 勇, 周曾鹏, 田劲东, 等. 结合高光谱和机器学习的无线充电金属异物检测[J]. 电子测量与仪器学报, 2022, 36(8): 238-247.

TIAN Y, ZHOU Z P, TIAN J D, et al. Metal object detection in wireless charging systems combining hyperspectral imaging and machine learning [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 238-247.

[4] 闫钧华, 张琨, 施天俊, 等. 融合多层次特征的遥感图像地面弱小目标检测[J]. 仪器仪表学报, 2022, 43(3): 221-229.

YAN J H, ZHANG K, SHI T J, et al. Multi-level feature fusion based dim small ground target detection in remote sensing images [J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 221-229.

[5] 徐梓川, 龚晓峰. 基于语义指导和自适应卷积的遥感云检测算法[J]. 电子测量技术, 2024, 47(1): 136-143.

XU Z CH, GONG X F. Cloud detection algorithm for remote sensing images based on semantic-guided and adaptive convolution [J]. Electronic Measurement Technology, 2024, 47(1): 136-143.

[6] 刘雪峰, 刘佳明, 付民. 生成对抗网络扩充样本用于高光谱图像分类[J]. 电子测量技术, 2022, 45(3): 146-152.

LIU X F, LIU J M, FU M. Adversarial network samples were generated for hyperspectral image classification [J]. Electronic Measurement Technology, 2022, 45(3): 146-152.

[7] XU SH F, GENG S J, FAN T T, et al. Hyperspectral image classification method based on narrowing semantic gap convolutional neural network [J]. International Journal of Remote Sensing, 2024, 45(7): 2208-2234.

[8] ROY S K, KRISHNA G, DUBEY S R, et al. HybridSN: Exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 17(2): 277-281.

[9] GE Z X, CAO G, LI X S, et al. Hyperspectral image classification method based on 2D-3D CNN and multibranch feature fusion [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 5776-5788.

[10] LI ZH K, WANG T N, LI W, et al. Deep multilayer fusion dense network for hyperspectral image classification [J]. IEEE Journal of Selected Topics in Applied Earth

- Observations and Remote Sensing, 2020, 13: 1258-1270.
- [11] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [12] ZHONG Z L, LI J, LUO ZH M, et al. Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework [J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(2): 847-858.
- [13] ZHU M H, JIAO L CH, LIU F, et al. Residual spectral-spatial attention network for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(1): 449-462.
- [14] 张祥东, 王腾军, 杨耘. 基于多尺度残差网络的小样本高光谱图像分类[J]. 激光与光电子学进展, 2020, 57(16): 162801.
- ZHANG X D, WANG T J, YANG Y. Classification of small-sized sample hyperspectral images based on multi-scale residual network [J]. Laser & Optoelectronics Progress, 2020, 57(16): 162801.
- [15] ROY S K, MANNA S, SONG T, et al. Attention-based adaptive spectral-spatial kernel ResNet for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(9): 7831-7843.
- [16] 颜铭靖, 苏喜友. 基于三维空洞卷积残差神经网络的高光谱影像分类方法 [J]. 光学学报, 2020, 40(16): 1628002.
- YAN M J, SU X Y. Hyperspectral image classification based on three-dimensional dilated convolutional residual neural network[J]. Acta Optica Sinica, 2020, 40(16): 1628002.
- [17] 韩彦岭, 崔鹏霞, 杨树瑚, 等. 基于残差网络特征融合的高光谱图像分类 [J]. 国土资源遥感, 2021, 33(2): 11-19.
- HAN Y L, CUI P X, YANG SH H, et al. Classification of hyperspectral image based on feature fusion of residual network [J]. Remote Sensing of Land and Resources, 2021, 33(2): 11-19.
- [18] ZHAN Y, WU K, DONG Y. Enhanced spectral-spatial residual attention network for hyperspectral image classification [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 7171-7186.
- [19] 吕欢欢, 王琢璐, 张辉. 边缘保护滤波和深度残差网络结合的高光谱影像分类 [J]. 激光与光电子学, 2022, 59(16): 1611005.
- LYU H H, WANG ZH L, ZHANG H. Hyperspectral image classification based on edge-preserving filter and deep residual network [J]. Laser & Optoelectronics Progress, 2022, 59(16): 1611005.
- [20] 程镨杰, 杨耘, 李龙威, 等. 基于深度可分离卷积的轻量化残差网络高光谱影像分类 [J]. 光学学报, 2023, 43(12): 1228010.
- CHENG R J, YANG Y, LI L W, et al. Lightweight residual network based on depthwise separable convolution for hyperspectral image classification [J]. Acta Optica Sinica, 2023, 43(12): 1228010.
- [21] WU Q G, HE M K, LIU ZH CH, et al. Multi-scale spatial-spectral residual attention network for hyperspectral image classification [J]. Electronics, 2024, 13(2): 262.
- [22] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv preprint arXiv: 1704.04861, 2017.
- [23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [24] YU F, KOLTUN V, FUNKHOUSER T. Dilated residual networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 472-480.
- [25] ZHANG W T, LI Y B, LIU L, et al. Hyperspectral image classification based on spectral-spatial attention tensor network [J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.
- [26] FENG Y, ZHU J, SONG R, et al. S2EFT: Spectral-spatial-elevation fusion transformer for hyperspectral image and LiDAR classification [J]. Knowledge-Based Systems, 2024, 283: 111190.

作者简介

赵雪松, 硕士, 主要研究方向为图像处理等。

E-mail: 17852833853@163.com

付民, 副教授, 博士, 主要研究方向为水下探测和新型传感器等。

E-mail: fumin@ouc.edu.cn

刘雪峰(通信作者), 教授, 博士, 主要研究方向为人工智能、图像处理等。

E-mail: nina.xf.liu@hotmail.com