基于 SE-Res2Net 与多尺度空谱融合注意力机制的高光谱图像分类

徐沁^{1,2)}, 梁玉莲^{1,2)}, 王冬越³⁾, 罗斌^{1,2)}

¹⁾ (安徽大学计算智能与信号处理教育部重点实验室 合肥 230601)

²⁾ (安徽大学计算机科学与技术学院 合肥 230601)

³⁾ (合肥工业大学管理学院 合肥 230009)

(xuqin@ahu.edu.cn)

摘 要: 为了提取更具有判决力的高光谱图像特征, 并防止网络因加深导致退化, 在新维度残差网络(Res2Net)和压缩激活网络(squeeze and excitation network, SENet)的基础上, 提出新型多尺度特征提取模块 SE-Res2Net, 并设计多尺度空谱融合注意力模块. 为了克服网络加深带来的退化问题, SE-Res2Net 模块利用通道分组提取高光谱图像细粒度的多尺度特征得到多个不同粒度的感受野, 并采用通道优化模块从通道层面量化特征图的重要性. 为了进一步从空间维和光谱维同时优化特征, 构建多尺度空谱融合的注意力模块, 利用非对称卷积在不同尺度上挖掘不同空间位置 and 不同光谱维特征的关系, 不但能减少计算量, 还能有效地提取具有判决力的空谱融合特征, 从而提高高光谱图像分类的精度. 在 3 个公共数据集 Indian Pines, University of Pavia 和 Grss_dfc_2013 上的对比实验表明, 与其他较新的深度网络相比, 该方法具有更高的总体精度(overall accuracy, OA)、平均精度(average accuracy, AA)和 Kappa 系数.

关键词: 高光谱图像分类; 多尺度特征; 注意力机制

中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18778

Hyperspectral Image Classification Based on SE-Res2Net and Multi-Scale Spatial Spectral Fusion Attention Mechanism

Xu Qin^{1,2)}, Liang Yulian^{1,2)}, Wang Dongyue³⁾, and Luo Bin^{1,2)}

¹⁾ (Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, Anhui University, Hefei 230601)

²⁾ (School of Computer Science and Technology, Anhui University, Hefei 230601)

³⁾ (School of Management, Hefei University of Technology, Hefei 230009)

Abstract: In order to extract more discriminative features for hyperspectral image and prevent the network from degrading caused by deepening, a novel multi-scale feature extraction module SE-Res2Net based on the new dimensional residual network (Res2Net) and squeeze and exception network (SENet), and a multi-scale spectral-spatial fusion attention module is developed for hyperspectral image classification. In order to overcome the degradation problem caused by network deepening, the SE-Res2Net module uses channel grouping to extract fine-grained multi-scale features of hyperspectral images, and gets multiple receptive fields of different granularity. Then, the channel optimization module is employed to quantify the importance of the feature maps at the channel level. In order to optimize the features from spatial and spectral dimensions simultaneously, a multi-scale spectral-spatial fusion attention module is designed to mine the relationship between different spatial positions and different spectral dimensions at different scales using

收稿日期: 2020-11-15; 修回日期: 2021-05-16. 基金项目: 国家自然科学基金(61502003, 61860206004, 72071001, U20B2068); 安徽省自然科学基金(2008085MG226). 徐沁(1983—), 女, 博士, 副教授, 硕士生导师, CCF 会员, 主要研究方向为计算机视觉、机器学习; 梁玉莲(1996—), 女, 硕士研究生, CCF 学生会员, 主要研究方向为高光谱图像分类; 王冬越(1994—), 男, 博士研究生, 主要研究方向为高光谱图像分类、乳腺超声影像智能诊疗; 罗斌(1963—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为图像处理、模式识别.

asymmetric convolution, which can not only reduce the computation, but also effectively extract the discriminative spectral-spatial fusion features, and further improve the accuracy of hyperspectral image classification. Comparison experiments on three public datasets, Indian Pines, University of Pavia and Grss_dfc_2013 show that the proposed method has higher overall accuracy (OA), average accuracy (AA) and Kappa coefficient compared to other state-of-the-art deep networks.

Key words: hyperspectral image classification; multi-scale feature; attention mechanism

高光谱遥感是将成像技术和光谱技术相结合的多维信息获取技术, 高光谱图像是由高光谱遥感器系统通过获取、记录带有地物信息的太阳辐射信号得到的含有上百通道且波段连续的图像^[1]. 因为含有丰富的光谱和空间信息, 高光谱遥感图像的处理和分析已被应用到医学影像分析^[2]、城市规划、国防安全和农业发展^[3]等领域.

在高光谱图像分析任务中, 高光谱图像分类任务的目的是对每个像素进行确定的类别标注, 但由于光谱波段易受到天气、环境、时间等因素的影响以及 Hughes 现象, 高光谱图像分类是一个极具挑战的任务^[4]. 早期的高光谱图像分类方法包括 K 最近邻(K -nearest neighbor algorithm, KNN)^[5]和支持向量机(support vector machine, SVM)^[6]等. 其中, KNN 分类器是对最近邻训练样本数进行统计, 因此在特征空间中需要大量的训练样本. 然而, 在高光谱高维的特征空间中往往缺乏足够的训练样本. SVM 分类算法利用核函数在高维空间中挖掘数据样本间的几何特性, 能够缓解维数灾难和 Hughes 现象问题, 因此被认为是解决高光谱图像分类的一种经典、有效的方法.

随着 Hinton 等^[7]提出的深度神经网络在计算机视觉领域获得了成功, 并应用于高光谱图像分类任务. 因为它能够自动分层地提取更抽象的高层特征, 所以表现出了更强的鲁棒性. 按照特征提取方式的不同, 现有的基于深度网络模型的高光谱图像分类方法可分为 3 类, 即基于谱特征、基于空间特征和基于空谱联合特征的分类方法.

最初, Hu 等^[8]利用一维卷积神经网络(one-dimensional convolutional neural networks, 1D-CNN)将原始光谱信息作为输入, 通过一层卷积层、一层池化层和 2 层全连接层进行高光谱图像分类. 但该方法只考虑光谱信息, 忽略了像素与像素之间的空间关系, 导致分类的精度较低. 二维卷积神经网络(two-dimensional convolutional neural

networks, 2D-CNN)能够提取高光谱图像的空间特征, 但是在处理丰富的光谱信息时需要很多参数, 在训练样本较少的情况下容易导致过拟合. 因此, 为了解决用 2D-CNN 进行高光谱图像分类导致较多参数而产生过拟合的问题, 文献[9]在提取空间特征前使用随机主成分分析(randomized principal component analysis, R-PCA)对高光谱图像数据进行降维, 随后利用 2D-CNN 在 5×5 的空间邻域上提取空间特征, 但是该方法在一定程度上损失了部分光谱信息. 上述方法只利用了光谱信息或空间信息, 而原始的高光谱图像具有 3D 的结构, 同时提取光谱特征和空间特征才能充分发挥高光谱图像“图谱合一”的特点. 因此, 文献[10]利用 3D 卷积神经网络(three-dimensional convolutional neural networks, 3D-CNN)有效地提取光谱-空间特征, 同时设计了一种组合正则化策略, 以解决高光谱数据由有限训练样本造成的过拟合的问题, 取得了较好的分类性能.

为了解决因网络加深带来的网络退化问题, He 等^[11]提出了残差网络(residual network, ResNet), 并在一定程度上缓解了该问题. 因此, Zhong 等^[12]不仅使用 3D-CNN 有效地集成了光谱-空间特征, 而且还采用了 ResNet 结构促进了梯度的反向传播, 缓解了精度下降的现象. 文献[13]将残差引入 3D-CNN 中, 采用自适应残差 3D 特征提取算法从高光谱图像中同时提取空间特征和光谱特征, 并结合树形 Parzen 估计算法自适应选定超参数优化网络, 提高了高光谱图像分类的精度. 最近, Gao 等^[14]在 ResNet 的基础上提出了一种面向目标检测任务的新维度残差网络(new dimensional residual network, Res2Net), 它用一个较小的 3×3 卷积滤波器取代了滤波器组, 同时用层级残差式连接不同的滤波器组. Res2Net 可以和现有的其他模块进行整合, 并且在不增加计算复杂度的情况下, 提高网络的特征表示能力, 其性能要优于 ResNet.

近几年,为了模拟人脑和人眼的感知机制,研究者提出了注意力机制^[15],并且在计算机视觉等领域取得了重要的突破.这是因为神经网络在注意力机制的辅助下能聚焦于给定信息的局部位置或特定特征,并分配相应的权值,突显关键性特征,弱化相对无用的特征,从而有效地提高模型的性能.如在 Hu 等^[16]提出的压缩激活网络(squeeze and excitation network, SENet)中,通过对特征通道之间的相互依赖关系进行建模,建立了特征通道之间的注意力机制,自适应地对特征进行重新校准.为了在通道和空间 2 个维度上优化特征, Woo 等^[17]提出了卷积块注意力模块(convolutional block attention module, CBAM),并对输入特征进行约束增强处理.另外,在 Zhao 等^[18]提出的金字塔特征注意力网络(pyramid feature attention networks, PFANet)中,也设计了空间和通道注意力机制,不同的是该方法使用空间注意力对较低层的特征进行优化,采用上下文感知金字塔方法提取富含语义的较高层特征,并利用通道注意力对其进行优化,因而保留了更多的结构信息.文献[19]将光谱注意力模块和空间注意力模块串行连接,获得具有判别性的光谱-空间特征.文献[20]提出了双路对称光谱-空间注意力网络,将 2 分支并行得到的输出进行加权求和,在突出空间信息的情况下,利用了高光谱图像的光谱信息.文献[21]利用 1D-CNN 和 2D-CNN 分别生成注意力掩码优化 1D 光谱和 2D 空间特征.因此,充分利用高光谱图像的光谱和空间注意力信息对高光谱图像分类有着重要的影响.

基于上述方法,本文在 Res2Net 和 SENet 的基础上,提出多尺度特征提取模块 SE-Res2Net,并设计多尺度空谱融合注意力模块应用于高光谱图像分类.为了防止网络加深导致梯度消失而无法收敛,本文采用 Res2Net 模块对特征通道进行分组,以滤波器组的形式得到分层连接,获得多个粒度更细的感受野;通过引入跳跃连接实现残差映射,以促进梯度的反向传播.同时,为了防止通道分组丢失通道间的相关性,将 SE 模块嵌入残差连接中,自适应地重新校准通道间的特征响应,提升 Res2Net 模块的性能.为了进一步优化特征,增强特征的判决力,本文设计了多尺度的空间和光谱融合的注意力模块,通过不同尺度非对称卷积核的卷积运算突出局部特征的重要性,不但能减少运算量,增加网络深度,还能筛选出对分类更有效的空谱融合特征,提升整个网络的分类性能.

1 本文方法

1.1 SE-Res2Net 模块

为了更好地促进梯度的反向传播,本文利用通道分组在残差单元内以滤波器组的形式获得有层次连接,增加尺度表示输出特征,并强化重要通道的特征,本文设计的 SE-Res2Net 模块如图 1 所示,其中,为了获得更细粒度的多个感受野,首先采用了 Res2Net 模块.假设输入特征数据为 $x \in \mathbb{R}^{w \times h \times b \times C}$, 其中 $w \times h$ 为特征图的空间维大小; b 为特征图光谱维尺寸; C 为特征通道数.为了实现分层操作,将特征按通道数平分为 4 个部分,即 $x_i \in \mathbb{R}^{w \times h \times b \times C'} (i=1,2,3,4)$, 每个特征图子集 x_i 的空间大小相同.其中, $C' = C/4$. 将 $x_i (i=2,3,4)$ 分别与 $3 \times 3 \times 3$ 的卷积核进行卷积,每个部分的输出为

$$y_i = \begin{cases} x_i, & i=1 \\ K(x_i), & i=2 \\ K(x_i + y_{i-1}), & i=3,4 \end{cases}$$

图 1 SE-Res2Net 模块

其中, $K(\cdot)$ 表示卷积操作. 然后对 y_i 进行逐一级联, 通过一个 $1 \times 1 \times 1$ 卷积核的卷积层得到 Res2Net 模块的输出 y . 由此可见, Res2Net 模块的输出是通道分组后不同分组的卷积和组合, 因而得到多个不同粒度的感受野, 感受野如图 2 所示.

图 2 Res2Net 处理后的感受野

为了防止通道分组丢失通道间的相关性, 将 Res2Net 模块的输出 y 送入 SE 模块. 在该模块中, 首先采用全局平均池化将特征 $y \in \mathbb{R}^{w \times h \times b \times C}$ 压缩成 $y' \in \mathbb{R}^{1 \times 1 \times C}$, 然后利用 2 个全连接层拟合通道间的相关性, 最后使用 Sigmoid 激活函数进行归一化处理. 因此, 通道的权重向量为

$$f^c = \sigma(\text{FC}(\delta(\text{FC}(y')))).$$

其中, FC 表示全连接层; δ 表示 ReLU 函数; σ 表示 Sigmoid 函数. SE 模块的输出为 $f' = f^c \cdot y$. 在残差单元内部实现在通道维度上对原始特征的重标定, 完成特征调整; 最后通过跳跃连接的方式将残差单元的输入 x 连接到残差单元的输出 f' 上, 得到 SE-Res2Net 模块的输出为 $f = f' + x$. 本文设计的 SE-Res2Net 模块能强调残差映射, 促进网络的收敛, 提高模型的稳定性.

1.2 多尺度空谱融合注意力模块

高光谱图像的空间维和光谱维信息依赖于 3D 数据结构, 为了进一步从空间维和光谱维融合优化特征, 保持高光谱图像数据结构的关联性, 本文设计了如图 3 所示多尺度空谱融合的注意力模块. 其中, 3 个不同尺度的分支用于提取空谱融合注意力特征. 首先在每个分支上利用非对称卷积核 $1 \times 1 \times k (k = 7, 9, 11)$, 并行地对光谱维特征进行操作; 然后采用批正则化(batch normalization, BN)操作加速网络的训练, 从而获得更具有判决力的光谱维注意力特征; 接着分别对每个分支利用 $k \times 1 \times 1$ 和 $1 \times k \times 1 (k = 7, 9, 11)$ 的非对称卷积核进行空间维的特征优化, 并运用 BN 操作实现在单分支上对空间维和光谱维特征的融合和优化. 每条支

路的操作和输出为

$$\begin{cases} X_1^k = H(F_{1 \times 1 \times k}(f)) \\ X_2^k = H(F_{1 \times k \times 1}(X_1^k)) \\ X_3^k = H(F_{k \times 1 \times 1}(X_2^k)) \end{cases} \quad k = 7, 9, 11.$$

其中, f 表示模块的输入特征图; $F_{a \times b \times c}(\cdot)$ 表示卷积操作, $(a, b, c = \{1, k\})$; $H(\cdot)$ 表示 BN 操作; $y_k = X_3^k (k = 7, 9, 11)$ 表示 3 条支路的输出. 最后将 3 个分支所得到的特征图按照对应元素相加, 并利用 Sigmoid 激活函数进行激活, 激活后的特征图与其原始数据融合, 得到最终的深层空谱融合特征图 $f_o = \sigma(y_7 + y_9 + y_{11}) \cdot f$. 其中, σ 表示 Sigmoid 激活函数; f_o 表示该模块最终的输出.

图 3 多尺度空谱融合注意力模块

以上设计的多尺度空谱融合注意力模块结合了高光谱图像数据的特性, 增强了局部特征的显著性, 可以削弱同物异谱和同谱异物的影响, 有利于高光谱图像分类.

1.3 网络框架

本文方法的网络结构如图 4 所示, 共有 20 个卷积层, 包含 2 个多尺度残差网络(SE-Res2Net)模块和一个多尺度空谱融合注意力模块. 为了降低光谱维的冗余度, 首先对高光谱图像进行主成分分析(principal components analysis, PCA)降维, 然

图 4 本文方法网络结构

后连接 2 个 SE-Res2Net 模块. 这是由于高光谱图像数据集的样本量非常有限, 若网络设置过深会出现过拟合的现象, 而过浅又无法有效地提取深层特征. 其中, 第 1 个模块的作用是提取特征, 第 2 个模块在提取深层特征的同时对特征进行优化, 随后连接一个多尺度空谱融合注意力机制模块. 最后通过 2 个全连接层和 Softmax 分类器进行分类, 为了防止过拟合, 在全连接层使用了 Dropout 策略. 根据高光谱图像具有高维空间的特点, 网络的优化器选择了较为合适的 Adam 优化器.

2 实验结果与分析

2.1 数据集与参数设置

为验证本文方法的有效性, 在高光谱数据集 Indian Pines^[22], University of Pavia^[23]和 Grss_dfc_2013^①上, 与 1D-CNN^[8], 3D-CNN^[10], SSRN^[12], SE-Res2Net 进行对比实验, 其中 SE-Res2Net 是本文提出的没有注意力模块的方法.

(1) Indian Pines. Indian Pines 是最早用于高光谱图像分类的一个数据集, 由机载可视红外成像光谱仪在美国印第安纳州拍摄. 成像波长范围为 0.4~2.5 μm , 空间分辨率约为 20 m, 包含剔除噪声后的 200 个波段, 共 10 249 个地物像素, 地物类为 16 类, 截取 145×145 的尺寸进行标注作为高光谱图像分类测试用途. 本文实验中采用的训练和测试样本数量如表 1 所示, 对于样本数量少于 600 个的类, 本文随机选取一半作为训练样本, 当样本数大于 600 时, 随机取 300 个作为训练集, 剩下的样本作为测试样本.

表 1 Indian Pines 数据集样本分布

类别	类名	训练	测试
1	Alfalfa	23	23
2	Corn-notill	300	1 128
3	Corn-min	300	530
4	Corn	118	119
5	Grass/Pasture	241	242
6	Grass/Trees	300	430
7	Grass/Pasture-mowed	14	14
8	Hay-windrowed	239	239
9	Oats	10	10
10	Soybeans-notill	300	672
11	Soybeans-min	300	2 155
12	Soybeans-clean	296	297
13	Wheat	102	103
14	Woods	300	965
15	Building-Grass-Trees-Drives	193	193
16	Stone-steel Towers	46	47
总计		3 082	7 167

(2) University of Pavia. University of Pavia 是由德国的机载反射光学光谱成像仪对意大利的帕维亚城所成像的一部分高光谱数据. 该光谱成像仪对波长范围 0.43~0.86 μm 内的 115 个波段连续成像, 其中 12 个波段由于受噪声影响被剔除, 因此一般使用剩余 103 个光谱波段所成的图像, 其空间分辨率为 1.3 m. 该数据的尺寸为 610×340 像素, 包含 9 类地物, 总共有 42 776 个地物像素. University of Pavia 数据集的训练和测试样本数量如表 2 所示, 每类随机选取 200 个作为训练样本, 剩下的作为测试样本.

① https://hyperspectral.ee.uh.edu/?page_id=459

表 2 University of Pavia 数据集样本分布

类别	类名	训练	测试
1	Asphalt	200	6431
2	Meadows	200	18449
3	Gravel	200	1899
4	Trees	200	2864
5	Sheets	200	1145
6	Bare Soil	200	4829
7	Bitumen	200	1130
8	Bricks	200	3482
9	Shadows	200	747
总计		1800	40976

(3) Grss_dfc_2013. Grss_dfc_2013 是 2013 年 IEEE GRSS 数据融合竞赛中使用的高光谱数据集, 由机载激光地图绘制中心收集, 于 2012 年 6 月 23 日在休斯顿大学校园及周边地区获得. 该数据集在 0.400~1.001m 区域有 144 个光谱带, 空间分辨率约为 2.5m. 该数据的空间尺寸为 349×1905 像素, 共有 15029 个地物像素, 被分为 15 类. Grss_dfc_2013 数据集的训练和测试样本数量如表 3 所示, 由于 Water 类的样本较少, 本文随机选取 162 个作为训练集, 剩余的每个类随机选择 200 个作为训练样本, 其他样本用于测试.

表 3 Grss_dfc_2013 数据集样本分布

类别	类名	训练	测试
1	Healthy grass	200	1051
2	Stressed grass	200	1054
3	Synthetic grass	200	497
4	Trees	200	1044
5	Soil	200	1042
6	Water	162	163
7	Residential	200	1068
8	Commercial	200	1044
9	Road	200	1052
10	Highway	200	1027
11	Railway	200	1035
12	Parking Lot 1	200	1033
13	Parking Lot 2	200	269
14	Tennis Court	200	228
15	Running Track	200	460
总计		2962	12067

由于训练集是随机选取的, 所以每组实验重复进行 5 次, 最终的分类精度取 5 次实验的平均

值. 本文采用总体精度(overall accuracy, OA)、平均精度(average accuracy, AA)和 Kappa 系数 3 个指标评价模型性能. 总体精度是模型在所有测试集上预测正确数量与总体数量的比值; 平均精度是每类预测正确数量与每类总体数量的比值的平均值; Kappa 系数判断图像一致性的程度. 本文实验使用 Python3.7 Tensorflow 框架, 在 i7-6700CPU 和 GTX2060GPU 上运行, 初始学习率设为 10^{-3} .

2.2 结果与分析

本文方法与 1D-CNN^[8], 3D-CNN^[10], SSRN^[12], SE-Res2Net 这 4 种方法的比较结果如表 4~表 6 所示. 从表 4~表 6 中可以看出, 本文方法的分类精度均优于其他方法. 从 Indian Pines 数据集上看, 本文方法比没有加注意力模块的 SE-Res2Net 方法在 OA 上提高了 0.47%, 比 3D-CNN 提高了 0.52%. 这种对比在 University of Pavia 数据集上体现得更为明显, 本文方法的 OA 比 3D-CNN 提高了 4.24%. 从表 6 可以看出, 本文方法在 Grss_dfc_2013 数据集上在 OA, AA, Kappa 这 3 个方面都达到了最优, 分别为 99.73%, 99.73%, 99.71%. 本文方法对高光谱图像最重要的光谱信息和空间信息进行融合提取, 并对其使用多尺度的方法, 融合不同尺度的注

表 4 Indian Pines 数据集分类精度 %

类别	1D-CNN ^[8]	3D-CNN ^[10]	SSRN ^[12]	SE-Res2Net	本文
1	69.57	99.13	100.00	97.39	100.00
2	70.92	98.23	95.81	99.11	99.47
3	77.92	99.92	95.73	99.62	99.92
4	70.59	100.00	98.49	99.66	100.00
5	94.21	100.00	98.35	100.00	100.00
6	98.37	99.72	99.81	100.00	99.86
7	92.86	100.00	97.14	92.86	100.00
8	98.32	100.00	100.00	100.00	100.00
9	70.00	100.00	98.00	98.00	98.00
10	88.10	99.28	99.52	99.40	99.79
11	71.09	97.77	97.75	97.60	98.67
12	89.56	99.39	98.86	99.46	99.36
13	99.03	100.00	99.81	100.00	100.00
14	92.23	99.83	98.84	99.71	99.81
15	74.09	100.00	99.07	100.00	100.00
16	91.49	100.00	98.72	97.87	99.15
OA	80.75	98.90	98.84	98.95	99.42
AA	80.20	98.96	99.08	99.02	99.41
Kappa	77.58	98.70	98.63	98.76	99.32

注: 粗体表示每类中的最高值.

表 5 University of Pavia 数据集分类精度 %

类别	1D-CNN ^[8]	3D-CNN ^[10]	SSRN ^[12]	SE-Res2Net	本文
1	87.01	93.99	97.80	98.19	98.95
2	85.08	99.78	95.61	98.72	99.63
3	79.96	99.57	95.62	95.16	99.48
4	95.46	98.59	99.70	98.55	99.15
5	99.72	100.00	99.83	100.00	100.00
6	83.52	91.04	98.74	98.66	99.95
7	88.96	98.96	99.70	99.56	100.00
8	73.27	71.17	97.23	99.41	99.54
9	99.87	99.95	99.38	99.95	100.00
OA	85.47	95.30	97.04	98.56	99.54
AA	84.27	94.59	96.55	98.07	99.18
Kappa	80.98	93.64	96.13	98.07	99.39

注. 粗体表示每类中的最高值.

表 6 Grss_dfc_2013 数据集分类精度 %

类别	1D-CNN ^[8]	3D-CNN ^[10]	SSRN ^[12]	SE-Res2Net	本文
1	97.18	99.26	100.00	99.22	99.71
2	98.22	99.47	99.13	99.75	99.83
3	100.00	99.88	99.92	99.84	99.88
4	97.09	99.98	100.00	100.00	100.00
5	98.64	100.00	100.00	100.00	100.00
6	99.14	100.00	100.00	100.00	100.00
7	90.99	99.53	99.12	99.72	99.76
8	91.38	99.25	99.02	98.51	98.89
9	87.70	96.06	98.95	99.03	99.32
10	89.70	96.34	99.88	99.65	99.94
11	84.27	99.57	99.86	99.78	99.81
12	91.60	99.36	99.46	99.32	99.71
13	82.75	99.85	99.93	100.00	100.00
14	99.39	100.00	100.00	100.00	100.00
15	99.35	100.00	100.00	100.00	100.00
OA	93.24	99.03	99.60	99.62	99.73
AA	93.70	99.26	99.59	99.59	99.73
Kappa	92.66	98.94	99.56	99.59	99.71

注. 粗体表示每类中的最高值.

意力特征, 获得更加丰富的图像特征, 从而提高整个模型分类效果.

图 5~图 7 分别为 1D-CNN, 3D-CNN, SSRN, SE-Res2Net 和本文方法在 Indian Pines, University of Pavia 和 Grss_dfc_2013 这 3 个数据集上得到的可视化分类结果. 在 3 个数据集中 1D-CNN 的分类效果图噪声明显, 而 3D-CNN, SSRN, SE-Res2Net 的分类效果图对地物分类相对清晰, 而本文方法的分类结果与真实地物标签最为接近. 从图 5~图 7 可以看出, 本文方法确实提高了高光谱图像分类精度. 这也与表 4~表 6 的分类精度相符合. 另外, 为了验证多尺度空谱融合注意力模块的有效性, 本文给出了 Indian Pines 数据集的部分注意力热图, 本文将三维热图分成若干个二维图像进行显示, 如图 8 所示, 图中的蓝色部分表示相对重要的特征, 可以看出注意力强调了重要性强的特征.

3 结 语

本文提出了基于 SE-Res2Net 与多尺度空谱融合注意力机制的高光谱图像分类方法. 设计的 SE-Res2Net 模块能提取粒度更细的多尺度特征, 并自适应地重新校准通道间的特征响应, 扩大感受野, 并且抑制了网络因加深带来的退化问题. 结合多尺度空谱融合注意力模块将高光谱图像“图谱合一”的特点物尽其用, 进一步优化空间和光谱特征, 增强了空谱特征的判决力. 实验结果验证了本文方法的有效性.

图 5 Indian Pines 数据集分类图

图 6 University of Pavia 数据集分类图

图 7 Grss_dfc_2013 数据集分类图

图 8 Indian Pines 数据集的部分注意力热图

参考文献(References):

- [1] Zhang Bing, Gao Lianru. Hyperspectral image classification and target detection[M]. Beijing: Science Press, 2011(in Chinese)
(张兵, 高连如. 高光谱图像分类与目标探测[M]. 北京: 科学出版社, 2011)
- [2] Lu G L, Fei B W. Medical hyperspectral imaging: a review[J]. Journal of Biomedical Optics, 2014, 19(1): 10901
- [3] Mahesh S, Jayas D S, Paliwal J, *et al.* Hyperspectral imaging to classify and monitor quality of agricultural materials[J]. Journal of Stored Products Research, 2015, 61: 17-26
- [4] Zhang Haokui, Li Ying, Jiang Yanan. Deep learning for hyperspectral imagery classification: the state of the art and prospects[J]. Acta Automatica Sinica, 2018, 44(6): 961-977(in Chinese)
(张号逵, 李映, 姜彦楠. 深度学习在高光谱图像分类领域的研究现状与展望[J]. 自动化学报, 2018, 44(6): 961-977)
- [5] Blanzieri E, Melgani F. Nearest neighbor classification of remote sensing images with the maximal margin principle[J]. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(6): 1804-1811
- [6] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(8): 1778-1790
- [7] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507
- [8] Hu W, Huang Y Y, Wei L, *et al.* Deep convolutional neural networks for hyperspectral image classification[J]. Journal of Sensors, 2015, 2015: Article No.258619
- [9] Makantasis K, Karantzas K, Doulamis A, *et al.* Deep supervised learning for hyperspectral data classification through convolutional neural networks[C] //Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Los Alamitos: IEEE Computer Society Press, 2015: 4959-4962
- [10] Chen Y S, Jiang H L, Li C Y, *et al.* Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(10): 6232-6251
- [11] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 770-778
- [12] Zhong Z L, Li J, Luo Z M, *et al.* Spectral-spatial residual net-work for hyperspectral image classification: A 3-D deep learning framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(2): 847-858
- [13] Xiao Zhiyun, Jiang Jiaxu, Ni Chen. Spectral-spatial classification of hyperspectral image based on self-adaptive deep residual 3D convolutional neural network[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(11): 2017-2029(in Chinese)
(肖志云, 蒋家旭, 倪晨. 自适应深层残差 3D-CNN 高光谱图像快速分类算法[J]. 计算机辅助设计与图形学学报, 2019, 31(11): 2017-2029)
- [14] Gao S H, Cheng M M, Zhao K, *et al.* Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662
- [15] Ba J, Mnih V, Kavukcuoglu K. Multiple object recognition with visual attention[OL]. [2020-11-15]. <https://arxiv.org/abs/1412.7755v2>
- [16] Hu J, Shen L, Sun G, *et al.* Squeeze-and-excitation networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 7132-7141
- [17] Woo S, Park J, Lee J, *et al.* CBAM: convolutional block attention module[C] //Proceedings of European Conference on Computer Vision. Heidelberg: Springer, 2018: 3-19
- [18] Zhao T, Wu X Q. Pyramid feature attention network for saliency detection[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 3085-3094
- [19] Sun H, Zheng X T, Lu X Q, *et al.* Spectral-spatial attention network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(5): 3232-3245
- [20] Hang R L, Li Z, Liu Q S, *et al.* Hyperspectral image classification with attention-aided CNNs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(3): 2281-2293
- [21] Pande S, Banerjee B. Adaptive hybrid attention network for hyperspectral image classification[J]. Pattern Recognition Letters, 2021, 144: 6-12
- [22] Green R O, Eastwood M L, Sarture C M, *et al.* Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)[J]. Remote Sensing of Environment, 1998, 65(3): 227-248
- [23] Kunkel B, Blechinger F, Lutz R, *et al.* ROSIS (reflective optics system imaging spectrometer) - a candidate instrument for Polar platform missions[C] //Proceedings of SPIE. Bellingham: Society of Photo-Optical Instrumentation Engineers, 1988, 868: 134-141