

doi: 10.6046/zrzyyg.2023171

引用格式: 郑宗生, 王政翰, 王振华, 等. 改进 3D-Octave 卷积的高光谱图像分类方法[J]. 自然资源遥感, 2024, 36(4): 82-91.
(Zheng Z S, Wang Z H, Wang Z H, et al. An improved 3D Octave convolution-based method for hyperspectral image classification[J]. Remote Sensing for Natural Resources, 2024, 36(4): 82-91.)

改进 3D-Octave 卷积的高光谱图像分类方法

郑宗生, 王政翰, 王振华, 卢鹏, 高萌, 霍志俊

(上海海洋大学信息学院, 上海 201306)

摘要: 高光谱图像数据具有维度高、数据稀疏、空间光谱信息丰富等特点, 针对空谱联合分类模型中高光谱图像卷积操作处理大片相同类别像素区域时会存在计算的空间冗余, 3D 卷积对深层空间纹理特征提取不充分, 串行注意力机制结构不能充分考虑空谱相关性的问题, 该文提出了改进的 3D-Octave 卷积高光谱图像分类模型。首先改进的 3D-Octave 卷积模块将输入的高光谱图像数据划分为高频特征图和低频特征图, 减少空间信息冗余, 提取多尺度的空间光谱特征, 结合跨层融合策略, 加强对浅层空间纹理特征和光谱特征的提取; 随后利用 2D 卷积提取深层空间纹理特征并进行光谱特征融合; 最后使用三维注意力机制跨纬度交互实现对有效特征的关注和激活, 增强网络模型的性能和鲁棒性。结果表明, 由于充分提取有效空谱联合特征, 在印第安松树林 (Indian Pines, IP) 数据集的训练集比例为 10% 的条件下, OA, Kappa 和 AA 分别为 99.32%, 99.13% 和 99.15%; 在帕维亚大学 (Pavia University, PU) 数据集的训练集比例为 3% 的条件下, OA, Kappa 和 AA 分别为 99.61%, 99.44% 和 99.08%。与 5 个主流分类模型进行对比, 获得了更高的分类精度。

关键词: 空间冗余; 3D-Octave 卷积; 跨层融合; 多尺度; 三维注意力机制

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2024)04-0082-10

0 引言

高光谱图像是一种重要的遥感数据, 包含丰富的光谱信息和地球表面的空间信息。它在环境监测^[1]、精准农业^[2]和军事应用等领域发挥着重要作用。然而, 如何从大量的数据中有效地提取有用信息并准确区分目标仍然是许多应用中的挑战。因此, 解决高光谱图像数据信息冗余问题和精确提取空谱特征已经成为近年来研究的热点。

在传统的高光谱分类方法中, 通常仅考虑光谱信息。典型的方法包括 k-近邻法^[3]、支持向量机 (support vector machine, SVM)^[4]和决策树^[5]。同时, 针对高光谱图像“异物同谱, 同物异谱”光谱信息冗余问题, 降维是一种常用的手段, 如主成分分析法 (principal components analysis, PCA)^[6]、独立成分分析法 (independent component analysis, ICA)^[7]和线性判别分析法 (linear discriminant analysis,

LDA)^[8]等。但是这些方法只利用了光谱维度的信息, 忽略了空间维度信息对分类结果的影响, 导致分类精度不高。为了充分利用空谱信息, 提出了一些结合空间光谱信息的分类方法, 例如稀疏表示^[9]、马尔可夫随机场^[10]和边缘保持滤波^[11]等。

近年来, 深度学习已经被广泛应用于高光谱图像分类, 它能够从高光谱图像中提取更深层次的特征, 帮助改善分类的准确性和性能。常见的网络模型有深度置信网络 (deep belief networks, DBN)^[12]、堆叠自动编码器 (stacked autoencoder, SAE)^[13]和卷积神经网络 (convolutional neural networks, CNN)。SAE 和 DBN 方法以分层训练的方式分层提取图像的深层特征, 将高光谱图像数据压缩为一维数据, 然而这些方法只提取光谱特征, 忽略了对高光谱图像重要空间特征的提取, 影响了分类的准确性。由于 CNN 可以在保留原始结构的同时利用空间特征, 因此引入了一些新的解决方案。Zhao 等^[14]在他们的框架中采用 CNN 作为特征提取器。Lee 等^[15]提出

收稿日期: 2023-06-15; 修订日期: 2023-09-03

基金项目: 国家自然科学基金项目“一种面向对模态遥感信息的质量抽样检验方案研究” (编号: 41671431)、上海市科委地方能力建设项目“复杂潮汐环境上海岛 (礁) 地物信息提取与精度验证方法及其示范应用” (编号: 19050502100) 和国家海洋局数字海洋科学技术重点实验室开放基金项目“面向深度学习与气象云图大数据的台风强度分类研究” (编号: B201801034) 共同资助。

第一作者: 郑宗生 (1979-), 男, 博士, 副教授, 主要研究方向为遥感图像处理。Email: szheng@shou.edu.cn。

通信作者: 王政翰 (1998-), 男, 硕士研究生, 主要研究方向为基于深度学习的遥感图像分类。Email: 2364639375@qq.com。

了一种具有更深和更宽网络的深度 CNN(CDCNN),但是模型包含多个卷积层、池化层和全连接层,参数量较大。Chen 等^[16]设计了基于 3D-CNN 的特征提取器模型,并集成了正则化,尽管这种方法将高光谱图像立方体数据作为一个整体,利用 3D 卷积核直接提取融合光谱信息和空间信息的特征。但是只使用 3D-CNN 导致模型参数量大,缺乏对局部空间光谱信息的提取。由于高光谱图像数据丰富,容易发生拟合现象,Zhong 等^[17]提出了光谱-空间残差网络 SSRN,利用 3D 卷积残差块来分别提取高光谱图像的空间和光谱特征,并使用批量归一化来正则化学习过程。Roy 等^[18]将 3D-CNN 和 2D-CNN 相结合提出了 HybridSN 模型,利用 2D 卷积提取深层空谱特征,提高了模型的特征提取能力,但是缺乏对有效特征的关注。为了提高模型对有效特征的提取,Li 等^[19]提出了双分支注意力机制的网络 DB-DA,通过 2 个分支分别对空间信息和光谱信息进行提取,并分别对有效特征进行激活,在较少的训练集下取得了不错的分类精度,但是没有考虑空间信息和光谱信息之间的相关性对有效特征提取的影响。随着人们对传统卷积运算理解的不断加深,近年来出现了一些新颖的卷积运算方法,并表现出不错的性能。例如,Zhang 等^[20]提出图卷积网络(graph convolutional network, GCN)应用于高光谱图像分类,可以引入局部图结构来提升卷积特性。然而,如何将数据转换为图结构并揭示节点之间的深层关系仍然具有挑战性。此外,Mou 等^[21]利用递归神经网络(recurrent neural network, RNN)将所有光谱带视为一个图像序列,其缺点在于对空间特征的考虑不足。然而在有限样本下的高计算成本和令人失望的性能是 GCN 和 RNN 在 HSI 分类任务中的显著瓶颈。因此,卷积运算仍然是提取空间-光谱特征的核心技术。

使用卷积操作对高光谱图像进行空谱特征提取时,对于高光谱图像中出现的大片相同类别像素存在空间冗余,深度空间纹理特征提取不足、缺乏对有效特征的关注和激活等问题。本文使用 PCA 操作和改进的 3D-Octave(3D-Oct)卷积模块,去除冗余的空间信息和光谱信息,降低了高光谱图像分类的计算成本。同时改进的 3D-Oct 卷积模块的双分支架构能提取不同尺度的空谱特征,避免卷积过程中原始空谱信息丢失。2D 卷积模块可以对浅层空谱特征图进行光谱通道特征融合并提取深度空间纹理特征。三维注意力机制模块使用并行的三分支架构代替串行的注意力模块,充分利用深层空间纹理特征和光谱特征的相关性,对有效特征关注和激活,提

高网络的表征能力。

1 研究方法

1.1 高光谱图像预处理

原始的高光谱图像数据立方体表示为 $I \in \mathbf{R}^{M \times N \times D}$,其中 D 为图像光谱通道数, M 为立方体高度, N 为立方体宽度。高光谱图像中的每个像素包含着丰富的光谱信息,涵盖了数百个连续的波长范围,相同类别像素的光谱信息并不完全相同,含有大量冗余的光谱信息,影响分类的精度。本文使用 PCA 降维的方法将每个高光谱像素光谱通道数量从 D 减少到 K ,降维后的高光谱图像数据表示为 $E \in \mathbf{R}^{M \times N \times K}$,其中 K 为选取的主成分个数。降维后的高光谱图像数据将原始的高维光谱波段信息压缩到较低的维度,去除了冗余的光谱信息,提取其中最具有代表性的光谱信息,同时保留完整的空间信息。

利用 PCA 将高光谱图像数据的光谱信息降维后,形成新的光谱带,这些光谱带是原始数据中的主要成分构成的,通常是高度相关的波段,并且在原始数据中具有最大的方差。降维后的光谱带中所包含的信息通常是按照降序排列的,这意味着具有更多信息的谱带将对接下来的特征提取过程产生更大的贡献。因此,利用 PCA 排序的特性,可以确定哪些光谱波段是相对重要的。基于 PCA 降维的排序原则和 3D 卷积的边缘效应,采用通道移位的方法将含有重要光谱信息的光谱带移动到通道中间,将含有光谱信息较低的谱带移动到通道的两端,这样可以使得重要的光谱带处于卷积核感受野的中心,加强 3D 卷积对含有光谱信息较多的光谱带的特征提取。对降维后的高光谱图像数据 E 进行通道移位后表示为数据立方体 $F \in \mathbf{R}^{M \times N \times K}$,移位操作原理如图 1 所示。

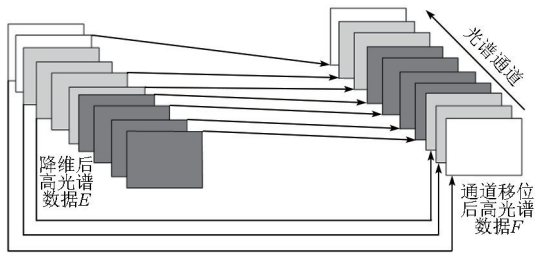


图 1 通道移位操作原理图

Fig.1 Channel shift operation schematic

由于高光谱图像中相邻像素之间存在较强的空间相关性,以 F 中每个像素为中心,对 F 进行边缘填充,将 F 划分为多个相邻像素块 $S \in \mathbf{R}^{B \times B \times K}$, K 为相邻像素块的光谱通道数,高度和宽度为 B ,每个相

邻像素块的中心像素标签作为相邻像素块的标签。

1.2 改进的 3D-Oct 卷积空谱特征提取模块

图像信号可以分为高频和低频通道。高频通道捕捉图像的细节信息,而低频通道则呈现图像的整体特征。在使用卷积运算进行计算时,由于低频通道存在空间冗余,会导致冗余计算。Chen 等^[22]设计了一种全新的卷积运算: Octave Convolution (简称 Oct 卷积),特征图可以分解为存储精细细节的高频特征图和存储全局结构特征的低频特征图,低频特征图大小为高频特征的一半,可以有效解决卷积计算中的空间冗余问题。然而,该方法引入了新的超参数 α ,将高频特征图和低频特征图的通道数划分为 $(1-\alpha)C$ 和 αC ,这样会导致在提取高光谱图像空谱特征过程中光谱通道信息丢失。针对原始的 Oct 卷积中参数 α 需要手动调节,耗费计算资源去优化以及在卷积过程中光谱通道信息丢失问题,本文提出了改进的 3D-Oct 卷积空谱特征提取模块,设计 Oct 卷积高频特征图和低频特征图的通道数与输入的相邻像素块 S 的通道数保持一致,充分提取光谱特征并结合跨层特征融合策略,加强对浅层空间纹理特征和光谱特征的提取。结构如图 2 所示,图中 X_1 和 X_2 分别为高频和低频特征图; Conv3D 为 3D 卷积操作; pooling 为全局平均池化操作; Up 为使用最近插值法进行上采样操作。

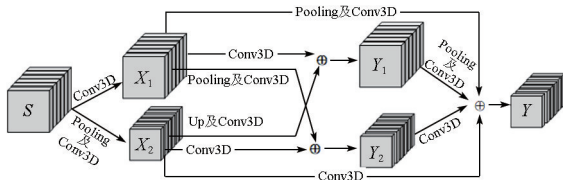


图 2 改进的 3D-Oct 卷积空谱特征提取模块

Fig.2 Improved 3D-Oct convolutional null spectrum feature extraction module

改进的 3D-Oct 卷积模块的第一层是一个简化的传统 3D-Oct 卷积,3D 卷积操作均使用 8 个 $3 \times 3 \times 3$ 的卷积核。对于输入的相邻像素块 $S \in \mathbf{R}^{B \times B \times K}$,通过一个两支多尺度卷积分解为一个高频特征图 X_1 和一个低频特征图 X_2 ,公式如下:

$$X_1 = \text{Conv3D}(S; W_{1 \rightarrow 1}), \quad (1)$$

$$X_2 = \text{Conv3D}(\text{Pooling}(S); W_{1 \rightarrow 2}), \quad (2)$$

式中: $W_{1 \rightarrow 1}$ 为高频信息更新权重; $W_{1 \rightarrow 2}$ 表示高频到低频信息转换权重。

3D-Oct 卷积的第二层拥有 4 个分支,4 个分支中 3D 卷积操作均使用 16 个 $3 \times 3 \times 3$ 的卷积核,输出为高频特征图 Y_1 和低频特征图 Y_2 。有效提取了高频特征图 X_1 和低频特征图 X_2 张量中的信息,同时实

现了高频和低频之间的信息交互转换,公式如下:

$$X_{1 \rightarrow 1} = \text{Conv3D}(X_1; W_{1 \rightarrow 1}), \quad (3)$$

$$X_{2 \rightarrow 1} = \text{Up}(\text{Conv3D}(X_2; W_{2 \rightarrow 1})), \quad (4)$$

$$Y_1 = X_{1 \rightarrow 1} + X_{2 \rightarrow 1}, \quad (5)$$

$$X_{2 \rightarrow 2} = \text{Conv3D}(X_2; W_{2 \rightarrow 2}), \quad (6)$$

$$X_{1 \rightarrow 2} = \text{Conv3D}(\text{pooling}(X_1); W_{1 \rightarrow 2}), \quad (7)$$

$$Y_2 = X_{2 \rightarrow 2} + X_{1 \rightarrow 2}, \quad (8)$$

式中: $X_{1 \rightarrow 1}$, $X_{2 \rightarrow 2}$ 为相同频率间信息更新; $X_{2 \rightarrow 1}$, $X_{1 \rightarrow 2}$ 为不同频率间的信息转换; $W_{2 \rightarrow 2}$ 为低频信息更新权重; $W_{2 \rightarrow 1}$ 为低频到高频的信息转换权重。

在最后一层的 4 个分支中,3D 卷积操作均使用 32 个大小为 $3 \times 3 \times 3$ 的卷积核,对高频特征图 Y_1 进行池化和 3D 卷积操作得到 $Y_{1 \rightarrow 2}$,对低频特征图 Y_2 进行 3D 卷积得到 $Y_{2 \rightarrow 2}$,同时结合跨层特征融合策略,设计跳跃连接得到 Y_3 和 Y_4 。将 $Y_{1 \rightarrow 2}$, $Y_{2 \rightarrow 2}$, Y_3 和 Y_4 相加融合,得到包含丰富空谱信息的特征图 Y ,公式如下:

$$Y_{1 \rightarrow 2} = \text{Conv3D}(\text{pooling}(Y_1); W_{1 \rightarrow 2}), \quad (9)$$

$$Y_{2 \rightarrow 2} = \text{Conv3D}(Y_2; W_{2 \rightarrow 2}), \quad (10)$$

$$Y_3 = \text{Conv3D}(\text{pooling}(X_1); W_{1 \rightarrow 2}), \quad (11)$$

$$Y_4 = \text{Conv3D}(X_2; W_{2 \rightarrow 2}), \quad (12)$$

$$Y = Y_{1 \rightarrow 2} + Y_{2 \rightarrow 2} + Y_3 + Y_4. \quad (13)$$

1.3 二维卷积模块

为提升对高光谱图像深度空间纹理特征的提取并且对生成三维特征图的通道信息进行融合,在改进的 3D-Oct 卷积层后接入 2 层 2D 卷积层,均使用 64 个 3×3 的卷积核。在 2D 卷积模块中,第 i 层第 j 个特征图在位置 (x, y) 处的值 $v_{i,j}^{x,y}$ 计算公式如下:

$$v_{i,j}^{x,y} = \bigotimes(b_{i,j} + \sum_{\tau} \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{i-1,\tau}^{p,q} \times v_{i-1,\tau}^{(x+p),(y+q)}), \quad (14)$$

式中: τ 为第 $i-1$ 层的特征图数量; P_i 和 Q_i 为卷积核的长和宽; $w_{i-1,\tau}^{p,q}$ 为在卷积核 (p, q) 位置与第 $i-1$ 层第 τ 个特征图连接的权值; $b_{i,j}$ 为第 i 层第 j 个特征图的偏置; $v_{i-1,\tau}^{(x+p),(y+q)}$ 为第 $i-1$ 层第 τ 个特征图 $((x+p), (y+q))$ 位置的特征值。

1.4 三维注意力机制模块

传统的计算通道注意力方法有一个显著的缺点: 为了计算通道权重,输入张量必须经过全局平均池化,这导致了空间信息的大量损失。因此,通道和空间维度之间的相互作用没有得到充分的体现。这种限制使通道注意力机制无法充分反映光谱通道信息和空间信息的内在相关性。虽然后来的空间注意力和通道注意力结合的 CBAM 模型^[23]解决了这

个问题,但是串联的注意力机制模块,存在通道注意和空间注意的独立计算,忽略了空间信息和光谱通道信息间的交互作用。改进的三维注意力机制使用3个分支获取空间维度和光谱通道维度之间的交互信息来计算注意力权重,可以更好地捕捉输入特征

之间的关系,从而提高模型对有效特征提取的性能。

图3为三维度注意力机制模块结构,图中:permute为维度变换操作;Avg_pooling为平均池化操作;Max_pooling为最大池化操作;Conv2D为2D卷积操作。

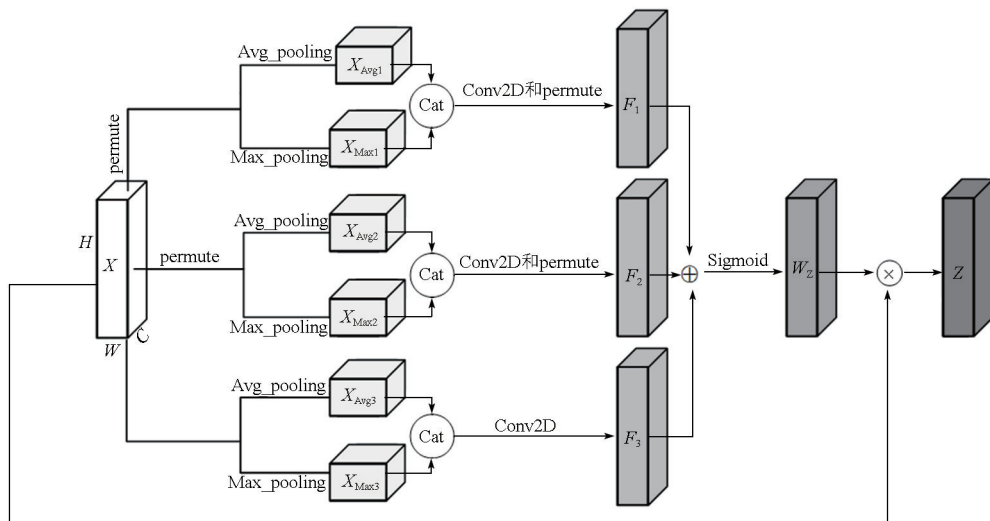


图3 三维度注意力机制模块结构图

Fig.3 Structure diagram of three-dimensional attention mechanism module

如图3所示,对于输入大小为 $H \times W \times C$ 的特征图 X ,第1个分支是在 C 维度和 H 维度之间建立了交互:对特征图进行permute操作后在 W 维度分别进行平均池化和最大池化得到 X_{Avg1} 和 X_{Max1} ,将特征图 X_{Max1} 和 X_{Avg1} 进行拼接,汇集该维上的平均特征和最大特征。然后通过 W 个卷积核为 3×3 的卷积层、permute操作得到大小为 $H \times W \times C$ 特征图 F_1 。

在第2个分支中,在 C 维度和 W 维度之间建立了交互:对特征图进行permute操作后,在 H 维度分别进行平均池化和最大池化得到 X_{Avg2} 和 X_{Max2} ,并进行拼接,然后通过 H 个卷积核为 3×3 的卷积层及permute操作得到大小为 $H \times W \times C$ 特征图 F_2 。

在第3个分支中,在 H 维度和 W 维度之间建立了交互:输入张量的 C 维度分别进行平均池化和最大池化得到 X_{Avg3} 和 X_{Max3} 进行拼接,然后通过 C 个卷积核为 3×3 的卷积层得到大小为 $H \times W \times C$ 特征图 F_3 。

最后对3个特征图相加融合,通过Sigmoid函数得到注意力权值 W_z 。将注意力权值 W_z 和输入特征图 X 相乘,最终得到跨纬度交互的深层有效特征图 Z ,公式如下:

$$W_z = \text{Sigmoid}(F_1 + F_2 + F_3), \quad (15)$$

$$Z = W_z \times X. \quad (16)$$

1.5 Softmax 损失函数

Softmax 损失函数是神经网络处理多分类任务

时常用的损失函数,由 Softmax 函数和交叉熵损失函数组合而成,Softmax 函数可以将全连接层输出向量 U 中的每个值变为0~1间的概率值,然后通过交叉熵损失函数计算预测值和真实值之间的损失,并通过反向传播不断优化模型参数。Softmax 损失函数公式为:

$$l(L, U) = -\frac{1}{n} \sum_{k=1}^n L_i \log_2 \left(\frac{e^{U_i}}{\sum_{j=1}^n e^{U_j}} \right), \quad (17)$$

式中: L 为标签向量; n 为类别个数; U_i 为向量 U 中的第 i 个值; L_i 为 L 中的第 i 个值。

1.6 网络结构

本文提出的整体模型架构如图4所示,对于原始大小为 $M \times N \times D$ 的高光谱图像,经过数据预处理得到尺寸为 $17 \times 17 \times K$ 的相邻像素块,作为模型的输入, K 代表相邻像素块的光谱通道数, 17×17 为相邻像素块尺寸。经过改进的3D-Oct卷积模块,通过多分支特征融合,得到32个3D特征图。随后通过两层二维卷积层得到64个2D特征图,通过三维注意力机制模块,跨纬度的计算注意力权重,形成新的有效特征图。最后,特征图通过Flatten操作、FC层、Dropout层、Softmax分类器,得到相邻像素块中心像素的分类结果。图中:Conv2D为2D卷积操作;Flatten为将特征图压缩为一维向量操作;FC为全连接层;Dropout为防止过拟合操作;softmax为softmax分类器。

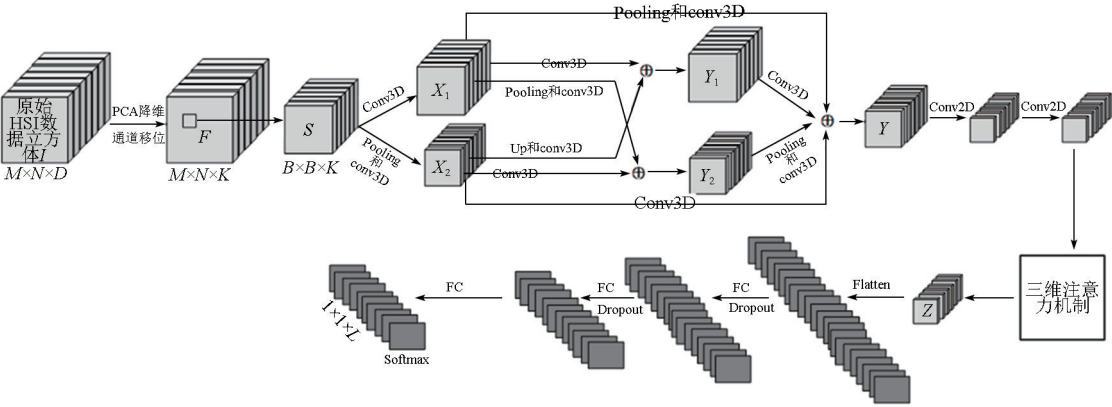


图 4 本文网络结构图

Fig.4 Network structure of the method in this paper

2 实验结果分析

为了验证本文算法的分类能力,本文在印第安纳州农场 (Indian Pines) 和帕维亚大学 (Pavia University) 2 个公开数据集进行实验,选取 Indian Pines 数据集的 10% 作为训练集,10% 作为验证集,其余作为测试集; Pavia University 数据集的 3% 作为训练集,3% 作为验证集,其余作为测试集。选择总体精度 (overall accuracy, OA)、Kappa 系数和平均分类精度 (average accuracy, AA) 作为定量评价指标,并与 SVM,3D-CNN,SSRN,HybridSN 和 DBDA 方法进行比较。为了确保评价指标的客观性和可靠性,进行了 10 次独立的重复实验,每个实验都有不同的随机种子初始化,确保对所提方法有效性的全面评估。

2.1 实验数据集

1) Indian Pines (IP) 数据集。IP 数据集是评估高光谱图像分类算法的一个基准数据集。最初是由 AVIRIS 机载成像光谱仪对美国印第安纳州的松树林进行成像获得的。该数据集包括 145×145 个像素点,空间分辨率约为 20 m,去除 20 个出现水吸收的波段,每个像素点包含 200 个信息丰富的光谱带,范围从 $0.4 \mu\text{m}$ 到 $2.5 \mu\text{m}$,可以有效区分 16 个不同的土地覆盖类别。

2) Pavia University (PU) 数据集。PU 数据集是用于高光谱图像分类研究的常用数据集之一。该数据集由 ROSIS 传感器在意大利的帕维亚城采集而来。数据集包含了一个具有 610×340 个像素的高光谱图像,空间分辨率约为 1.3 m,去除受噪声影响的 12 个波段,每个像素点包含 103 个连续光谱带,覆盖了 $0.43 \mu\text{m}$ 到 $0.86 \mu\text{m}$ 的波长范围,可以有效区分 9 个不同的土地覆盖类别。

2 个数据集的伪彩图和标签图如图 5 所示。

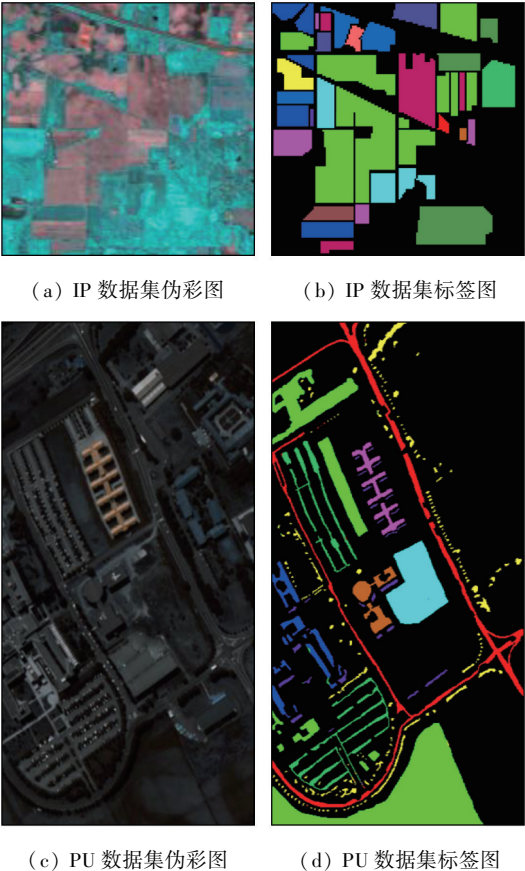


图 5 IP 和 PU 数据集伪彩图和标签图

Fig.5 Pseudo-color map and label map of

IP and PU datasets

2.2 实验环境和参数设置

由于运用 PCA 对高光谱图像数据降维的同时会使高光谱图像数据损失部分光谱信息,因此需要为数据立方体选择合适的主成分数量 K ,提高模型分类精度。针对 IP 数据集,本文选取前 100 ~ 120 个主成分,针对 PU 数据集本文选取前 15 ~ 35 个主成分,在保持其他参数不变的情况下进行对比实验。

由于不同地物类别的覆盖范围不同,所以相邻像素块 $S \in R^{B \times B \times K}$ 中尺寸 B 的大小会对模型的精度

产生较大影响,本文选取 B 的大小在 13~21 之间,间隔为 2,进行对比实验。

图 6 为光谱通道对精度的影响,由图 6 可知,当选取主成分的数量过低时,会造成光谱信息的丢失,影响模型分类精度。同时,选取的主成分数量过多时,会造成光谱信息的冗余,也会造成分类精度的降低。在图 6(a)中, $K=110$ 时 3 种指标取得最优分类

精度;在图 6(b)中, $K=30$ 时 3 种指标取得最优分类精度。最终确定 IP 数据集降维后的主成分数量 $K=110$,PU 数据集主成分数量 $K=30$ 。为相邻像素块尺寸对精度的影响,由图 7 可知,2 个数据集在 $B=17$ 时 3 种指标取得最优分类精度。最终确定 IP 数据集划分相邻像素块 $S \in R^{17 \times 17 \times 110}$,PU 数据集划分相邻像素块 $S \in R^{17 \times 17 \times 30}$ 。

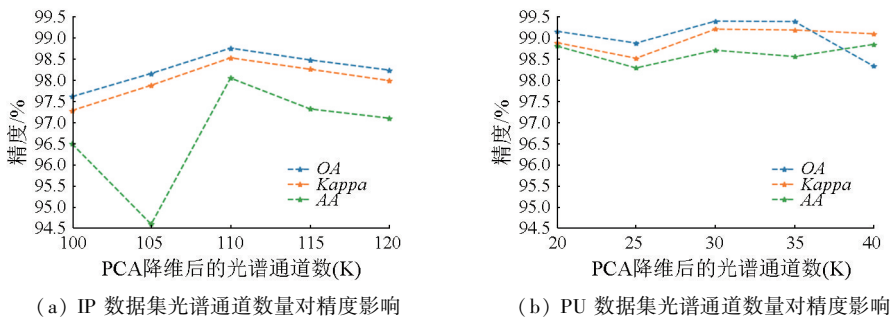


图 6 光谱通道数量对精度影响

Fig.6 Influence of the number of spectral channels on the accuracy

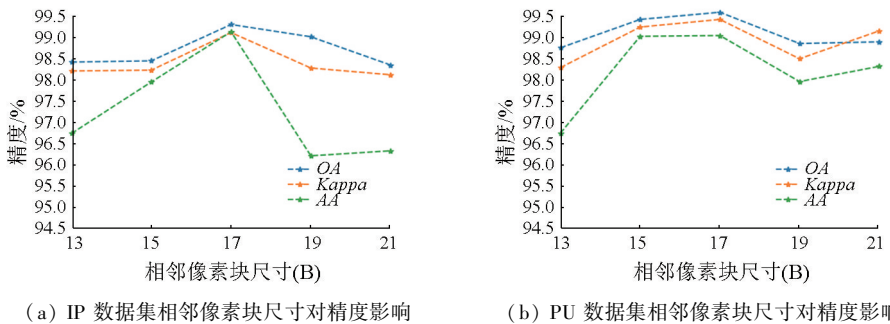


图 7 相邻像素块尺寸对精度影响

Fig.7 Effect of adjacent pixel block size on accuracy

在 2 个数据集上,epoch 均设置为 100,学习率为 0.001,批量处理大小(batch size)为 64,Dropout 率为 0.4。实验硬件条件为 Intel(R) Core(TM) i7-11700 处理器、NvidiaGTX3060 显卡和 32 G 内存。实验软件条件为 CUDA11.3,CUDNN8.2.1,pytorch1.5.1 和 Python3.8。

2.3 对比实验结果

不同算法在 IP 数据集和 PU 数据集的分类结果如表 1 和表 2 所示。从中可以看出:

- 1)SVM 方法忽略了空间信息,在分类性能上有所欠缺,在 IP 数据集上有两类的分类精度为 0。
- 2)3D-CNN 方法利用到了高光谱图像的空谱信息,相比传统的 SVM 方法,在 2 个数据集上的分类精度有所提高,但是缺少对深层信息的提取,所以精度仍有所不足。
- 3)SSRN 方法在 3D-CNN 网络的基础上加入了残差结构,加深了网络结构,较 3D-CNN 方法分类精度有明显提高。在 IP 数据集上,有 2 个类别的分

表 1 不同算法在 IP 数据集的分类结果

Tab.1 Classification results of different algorithms on IP dataset

类别	SVM	3D-CNN	SSRN	HybridSN	DBDA	本文方法
1	0.00	94.40	88.10	93.18	100.00	100.00
2	65.10	89.22	97.57	98.71	98.75	99.33
3	84.27	93.85	97.83	97.38	99.73	100.00
4	100.00	93.75	96.35	99.01	98.60	100.00
5	95.08	97.97	97.25	99.30	99.52	99.24
6	90.45	97.93	99.07	97.90	99.84	94.60
7	100.00	92.86	100.00	96.00	95.83	99.89
8	93.71	94.81	99.54	100.00	100.00	100.00
9	0.00	66.67	75.01	80.95	94.74	100.00
10	79.78	90.16	97.61	97.74	97.28	97.92
11	48.36	96.40	95.82	97.81	97.69	99.01
12	66.54	92.55	97.14	98.27	99.60	98.51
13	95.71	91.92	100.00	99.45	96.86	97.87
14	88.00	99.39	98.69	99.56	99.30	99.04
15	72.63	96.18	95.47	96.90	98.85	98.96
16	99.56	91.76	92.59	88.76	88.17	97.60
OA	66.94	94.82	97.25	97.92	98.64	99.32
Kappa	61.20	94.10	96.85	97.76	98.43	99.13
AA	45.47	90.82	95.30	97.11	98.31	99.15

表 2 不同算法在 PU 数据集的分类结果

Tab.2 Classification resultsof different algorithms on PU dataset (%)

类别	SVM	3D-CNN	SSRN	HybridSN	DBDA	本文方法
1	93.37	96.40	98.25	97.26	98.80	99.94
2	94.73	99.13	99.85	99.87	99.90	100.00
3	74.46	91.14	93.05	99.09	97.53	99.50
4	90.37	97.50	97.00	99.89	99.93	98.10
5	99.92	99.24	99.92	99.23	98.17	99.77
6	89.20	99.30	96.61	99.78	99.81	99.67
7	81.68	92.64	98.58	100.00	99.61	99.60
8	79.04	90.23	92.65	96.41	98.02	98.50
9	100.00	94.22	96.95	99.77	99.88	99.51
OA	91.85	97.41	98.75	99.08	99.37	99.61
Kappa	89.22	96.56	98.34	98.68	99.17	99.44
AA	89.77	95.44	97.60	97.87	98.53	99.08

类精度达到 100%,但是仍有 2 个类别的分类精度未达到 90%; 在 PU 数据集上,每个类别的分类精度均达到 90%以上。

4) HybridSN 方法使用 3D 和 2D 混合的网络结

构,提取了深层的空间信息,分类结果略优于 SSRN 方法。

5) DBDA 方法采用双分支架构,分别提取光谱信息和空间,并结合注意力机制,在 2 个数据集上取得了不错的分类结果。在 IP 数据集上,有 2 类的分类精度达到 100%; 在 PU 数据集上,每个类别的分类精度都在 97%以上。本文方法在 IP 数据集上,有 5 个类别的分类精度达到 100%,相比精度最优的 DBDA 方法,OA 提高 0.68 百分点,Kappa 系数提高 0.7 百分点,AA 提高 0.84 百分点; 在 PU 数据集上,每个类别的分类精度在 98%以上,相比精度最优的 DBDA 方法,OA 提高 0.24 百分点,Kappa 系数提高 0.27 百分点,AA 提高 0.55 百分点。

图 8—9 直观地展示了以上算法对 2 个数据集的分类效果图,相比于其他方法,本文方法获得的分类结果图轮廓清晰,噪点更少,更加接近标签图。尤其是在 IP 数据集上,本文方法的表现明显优于其他算法。

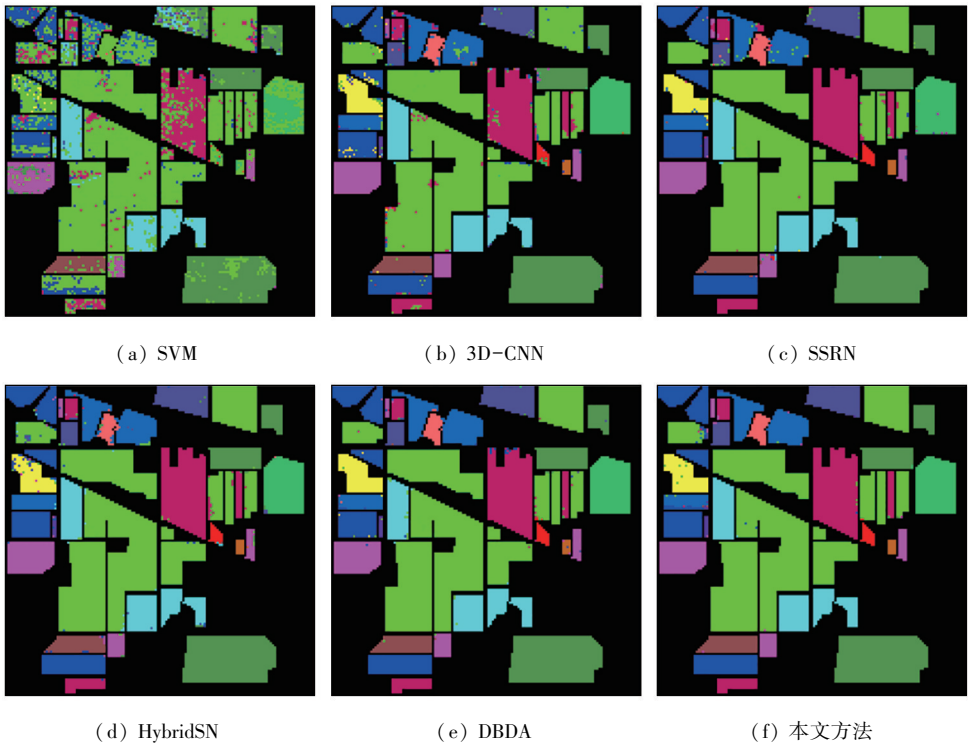


图 8 IP 数据集上不同算法分类效果图

Fig.8 Classification effect of different algorithms on IP dataset

通过 IP 和 PU 高光谱数据集进行实验,将本文方法与 SVM,3D-CNN,SSRN,HybridSN 和 DBDA 方法进行对比,可以发现本文方法分类精度均高于对比方法,获得了最高的分类结果,且模型有较强的稳定性。

2.4 消融实验

为验证三维注意力模块的有效性,将本文方法

与不使用注意力机制模块和使用 CBAM 模块进行对比,在 2 个数据集上的分类精度如表 3 所示。根据表 3 可知,本文方法使用三维注意力模块,充分考虑空间维度和光谱维度相关性,相比不使用注意力模块和使用传统 CBAM 注意力模块,能够更加有效地提取特征,提高模型的分类精度。

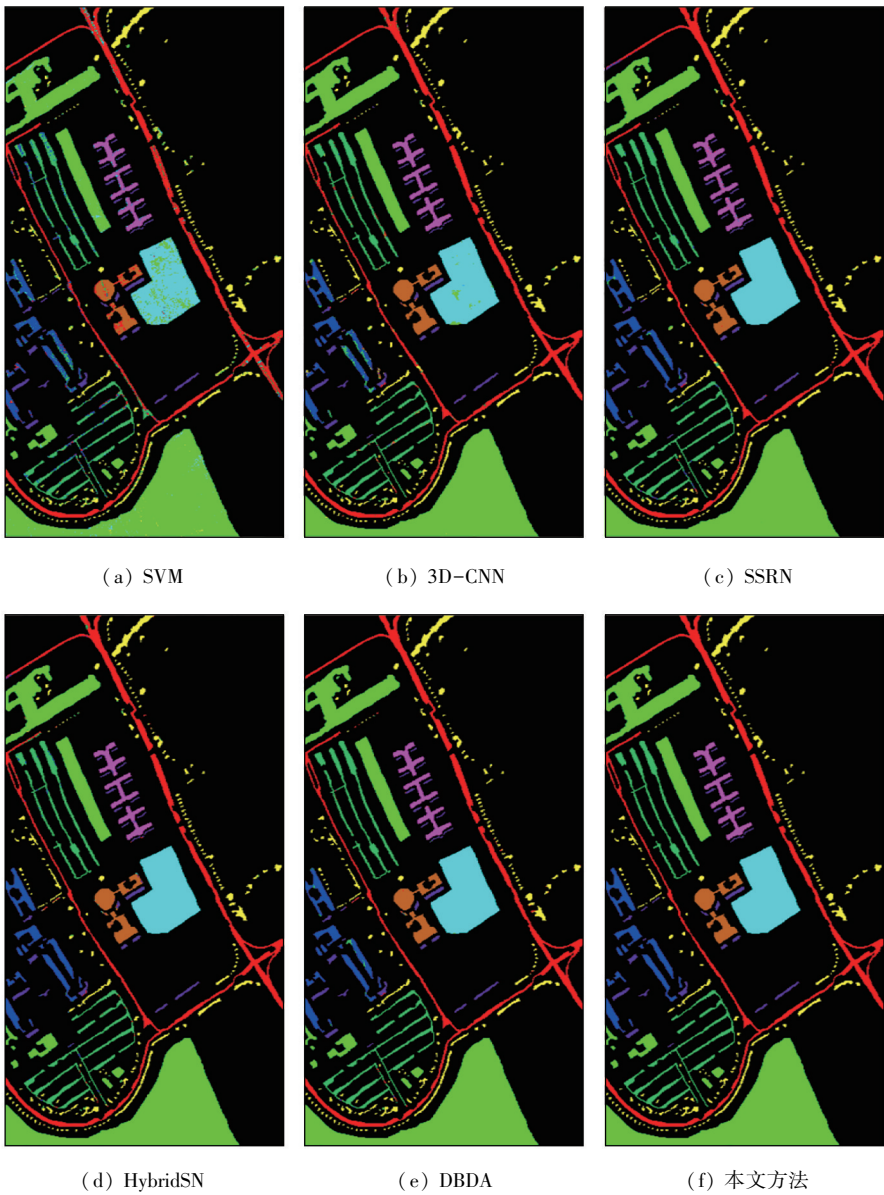


图 9 PU 数据集上不同算法分类效果图

Fig.9 Classification effect of different algorithms on PU dataset

表 3 使用不同注意力机制在 IP 和 PU 数据集的分类结果

Tab.3 Classification results using different attention mechanisms on the IP and PU datasets (%)						
分类方法	IP 数据集			PU 数据集		
	OA	Kappa	AA	OA	Kappa	AA
不使用注意力模块	97.75	97.86	92.38	98.65	98.21	98.20
使用 CBAM 模块	98.64	98.43	98.31	99.37	99.17	98.53
本文方法	99.32	99.13	99.15	99.61	99.44	99.08

3 结论

本文提出了一种改进的 3D-Oct 卷积模块、二维卷积、三维注意力结合的高光谱图像分类模型,并进行对比实验,主要结论包括:

- 1)本文方法在 2 个公开的高光谱数据集上进行了实验,与 5 个主流分类模型进行对比,获得了更高的分类精度。
 - 2)改进的 3D-Oct 卷积模块充分利用光谱信息的同时,使用高低频特征图去除冗余的空间信息,避免了对超参数的优化,节约了计算机资源,并提取多尺度的空间光谱信息,增强了模型的稳定性。
 - 3)本文提出的三维注意力机制模块充分考虑了空间纹理特征维度和光谱特征维度的相关性,跨纬度实现对空间光谱特征的筛选和有效激活,强化像素的空谱联合有效信息,增强了高光谱图像空间光谱信息的利用效率。
- 未来将继续研究降低训练样本提高分类精度以及高光谱图像分类中地物类别不平衡造成的精度损失问题。

参考文献 (References) :

[1] 陈功伟,赵思颖,倪才英.高光谱监测技术在重金属污染土壤上的应用[J].中国科学院大学学报,2019,36(4):560-566.
Chen G W,Zhao S Y,Ni C Y.Hyperspectral monitoring of soil contaminated by heavy metals[J].Journal of University of Chinese Academy of Sciences,2019,36(4):560-566.

[2] 洪霞,江洪,余树全.高光谱遥感在精准农业生产中的应用[J].安徽农业科学,2010,38(1):529-531,540.
Hong X,Jiang H,Yu S Q.Application of hyper-spectral remote sensing in production of precision agriculture[J].Journal of Anhui Agricultural Sciences,2010,38(1):529-531,540.

[3] Ma L,Crawford M M,Tian J.Local manifold learning-based k-nearest-neighbor for hyperspectral image classification[J].IEEE Transactions on Geoscience and Remote Sensing,2010,48(11):4099-4109.

[4] Delalieux S,Somers B,Haest B,et al.Heathland conservation status mapping through integration of hyperspectral mixture analysis and decision tree classifiers[J].Remote Sensing of Environment,2012,126:222-231.

[5] Ding S,Chen L.Classification of hyperspectral remote sensing images with support vector machines and particle swarm optimization [C]//2009 International Conference on Information Engineering and Computer Science.Wuhan,China.IEEE,2009:1-5.

[6] Kang X,Xiang X,Li S,et al.PCA-based edge-preserving features for hyperspectral image classification [J].IEEE Transactions on Geoscience and Remote Sensing,2017,55(12):7140-7151.

[7] Wang J,Chang C I.Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis [J].IEEE Transactions on Geoscience and Remote Sensing,2006,44(6):1586-1600.

[8] Bandos T V,Bruzzone L,Camps-Valls G.Classification of hyperspectral images with regularized linear discriminant analysis[J].IEEE Transactions on Geoscience and Remote Sensing,2009,47(3):862-873.

[9] Cui B,Cui J,Lu Y,et al.A sparse representation-based sample pseudo-labeling method for hyperspectral image classification[J].Remote Sensing,2020,12(4):664.

[10] Cao X,Xu Z,Meng D.Spectral-spatial hyperspectral image classification via robust low-rank feature extraction and Markov random field[J].Remote Sensing,2019,11(13):1565.

[11] Kang X,Li S,Benediktsson J A.Spectral - spatial hyperspectral image classification with edge-preserving filtering[J].IEEE Transactions on Geoscience and Remote Sensing,2014,52(5):2666-2677.

[12] Chen Y,Zhao X,Jia X.Spectral - spatial classification of hyperspectral data based on deep belief network[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2015,8(6):2381-2392.

[13] Madani H,McIsaac K.Distance transform-based spectral-spatial feature vector for hyperspectral image classification with stacked autoencoder[J].Remote Sensing,2021,13(9):1732.

[14] Zhao W,Du S.Spectral - spatial feature extraction for hyperspectral image classification;A dimension reduction and deep learning approach[J].IEEE Transactions on Geoscience and Remote Sensing,2016,54(8):4544-4554.

[15] Lee H,Eum S,Kwon H.Cross-domain CNN for hyperspectral image classification [C]//IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain. IEEE,2018:3627-3630.

[16] Chen Y,Jiang H,Li C,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.

[17] Zhong Z,Li J,Luo Z,et al.Spectral - spatial residual network for hyperspectral image classification;A 3-D deep learning framework [J].IEEE Transactions on Geoscience and Remote Sensing,2018,56(2):847-858.

[18] 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,2020,17(2):277-281.

[19] Li R,Zheng S,Duan C,et al.Classification of hyperspectral image based on double-branch dual-attention mechanism network [J].Remote Sensing,2020,12(3):582.

[20] Zhang C,Wang J,Yao K.Global random graph convolution network for hyperspectral image classification[J].Remote Sensing,2021,13(12):2285.

[21] Mou L,Ghamisi P,Zhu X X.Deep recurrent neural networks for hyperspectral image classification[J].IEEE Transactions on Geoscience and Remote Sensing,2017,55(7):3639-3655.

[22] Chen Y,Fan H,Xu B,et al.Drop an octave:Reducing spatial redundancy in convolutional neural networks with octave convolution [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV).Seoul,Korea (South).IEEE,2019:3434-3443.

[23] Woo S,Park J,Lee J Y,et al.CBAM:convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer,2018:3-19.

An improved 3D Octave convolution-based method
for hyperspectral image classification

ZHENG Zongsheng, WANG Zhenghan, WANG Zhenhua, LU Peng, GAO Meng, HUO Zhijun
(Department of Information, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Hyperspectral image data are characterized by high dimensionality, sparse data, and rich spatial and

spectral information. In spatial-spectral joint classification models, convolution operations for hyperspectral images can lead to computational spatial redundancy when processing large regions of pixels of the same category. Furthermore, the 3D convolution fails to sufficiently extract the deep spatial texture features, and the serial attention mechanism cannot fully account for spatial-spectral correlations. This study proposed an improved 3D Octave convolution-based model for hyperspectral image classification. First, the input hyperspectral images were divided into high- and low-frequency feature maps using an improved 3D Octave convolution module to reduce spatial redundancy information and extract multi-scale spatial-spectral features. Concurrently, a cross-layer fusion strategy was introduced to enhance the extraction of shallow spatial texture features and spectral features. Subsequently, 2D convolution was used to extract deep spatial texture features and perform spectral feature fusion. Finally, a 3D attention mechanism was used to focus on and activate effective features through interactions across latitudes, thereby enhancing the performance and robustness of the network model. The results indicate that, due to the adequate extraction of effective spatial-spectral joint features, the overall accuracy (OA), Kappa coefficient, and average accuracy (AA) were 99.32%, 99.13%, and 99.15%, respectively in the case where the Indian Pines (IP) dataset accounted for 10% in the training set and were 99.61%, 99.44%, and 99.08%, respectively when the Pavia University (PU) dataset represented for 3% of the training set. Compared to five mainstream classification models, the proposed method exhibits higher classification accuracy.

Keywords: spatial redundancy; 3D Octave convolution; cross-layer fusion; multi-scale; 3D attention mechanism
(责任编辑: 李 瑜)