

基于多方向自适应感知网络的高光谱遥感图像分类

刘倩, 吴泽彬, 徐洋, 郑鹏, 郑尚东, 韦志辉

南京理工大学 计算机科学与工程学院, 南京 210014

摘要: 高光谱遥感能够同步获取目标场景的光谱数据和空间图像, 满足对目标物体成分组成和形貌特征的探测需求, 已被广泛应用于精准农业、地质勘察、环境监测和生物医学等领域。近几年, 得益于优秀的空谱信息表征能力, 基于卷积神经网络的高光谱图像分类方法取得了优于传统方法的分类精度。然而, 基于一定规则和固定方形窗口的卷积运算无法满足不同对象、不同分布的空谱特征提取需求, 当像素位于类别边缘时该模式无法避免其他类别的无关信息, 导致类别间的信息扩散以及跨类别像素的误分和错分。针对上述问题, 本文提出了一种基于多方向自适应感知的高光谱图像空谱联合分类方法, 并通过联合不同邻域范围的空谱上下文信息来改善模型的局部空谱建模能力。具体研究过程如下: 首先, 将常规卷积网络的全窗口滤波器拆分为不同方向的半窗滤波器, 从而设计出侧窗卷积以捕捉具有方向性的空谱特征; 然后, 在此基础上进一步将不同方向的侧窗滤波核整合到统一的卷积架构中构造空谱分离多方向卷积, 并设计方向自适应感知模块, 以密集连接构建基于高光谱图像的多方向自适应感知分类网络, 赋予模型自适应学习多种空谱结构特征的能力, 弥补常规卷积和侧窗卷积只能建模单一方向空谱关系的不足, 提升模型对复杂空谱结构的刻画能力。在3个公开数据集上的实验结果表明, 本文提出的多方向自适应感知分类方法较常规卷积方法具有更优的分类性能, 能增强边缘样本的表征准确性, 改善分类中的边缘混淆现象。

关键词: 遥感, 高光谱图像, 深度学习, 多方向自适应感知, 空谱结构建模, 空谱联合分类

中图分类号: TP701/P2

引用格式: 刘倩, 吴泽彬, 徐洋, 郑鹏, 郑尚东, 韦志辉. 2024. 基于多方向自适应感知网络的高光谱遥感图像分类. 遥感学报, 28(1): 168-186

Liu Q, Wu Z B, Xu Y, Zheng P, Zheng S D and Wei Z H. 2024. Hyperspectral remote sensing image classification based on multidirectional adaptive aware network. National Remote Sensing Bulletin, 28(1): 168-186 [DOI: 10.11834/jrs.20243292]

1 引言

高光谱遥感能为观测地物提供包含数百波段、近乎连续的光谱曲线, 生成蕴含丰富空间和光谱信息的三维数据立方体(张兵, 2016)。高光谱图像的光谱分辨率可达纳米级别, 在地物的精准识别和分类方面具有极大优势(童庆禧等, 2016)。然而, 高光谱分辨率使高光谱遥感影像的精细分类面临诸多挑战。一方面, 高光谱图像存在“同谱异物”和“同物异谱”现象(Ghamisi等, 2017), 加之地物分布结构复杂, 分类任务常面临

混淆、错分的情况, 严重干扰分类结果的应用; 另一方面, 地面标注成本高难度大, 导致可用于模型训练的标记数据数量有限, 与高光谱维度之间存在不平衡(杜培军等, 2016), 使得分类中的Hughes现象(Hughes, 1968)不容忽视。因此, 如何构建小样本下高效的高光谱图像分类方法一直是该领域内的研究热点。

针对高维度、少样本的难点, 国内外研究提出一系列的解决方案, 从包括波段选择、主成分分析(Principal Component Analysis)、线性判别分析(Linear Discriminant Analysis)在内的光谱信息

收稿日期: 2023-07-11; 预印本: 2023-09-23

基金项目: 国家自然科学基金(编号:61772274, 61701238, 61671243); 江苏省自然科学基金(编号:BK20180018, BK20170858); 中央高校基本科研基金(编号:30917015104, 30919011103, 30919011402); 中国博士后科学基金(编号:2017M611814)

第一作者简介: 刘倩, 研究方向为遥感图像处理、机器学习。E-mail: liuqianahu@163.com

通信作者简介: 吴泽彬, 研究方向为高光谱遥感, 多源遥感融合, 云边协同计算。E-mail: wuzb@njust.edu.cn

主导的分类方法逐渐发展为以空间信息辅助的空谱联合分类方法(叶珍等, 2021)。传统的空谱联合分类方法通过提取纹理、形状、分布结构、语义等空间特征来充分挖掘和利用高光谱图像的相邻像素高相关性(Gao等, 2013), 以地表覆盖的连续性来减少分类结果的椒盐现象, 构建更精准的分类地图。典型的空谱联合分类方法包括 Gabor 滤波(Li 和 Du, 2014; Huo 和 Tang, 2011)、局部二值模式(Local binary patterns)(Ojala 等, 2002; Li 等, 2015; Jia 等, 2017)、形态学剖面(Morphological profiles)(Pesaresi 和 Benediktsson, 2001; Benediktsson 等, 2005)以及形态学属性剖面(Morphological attribute profiles)(Mura 等, 2010a, 2010b)、多核复合(Camps-Valls 等, 2006)等。此外, 还有直接应用待处理像素空间邻域信息的稀疏表示分类方法(Chen 等, 2011; Liu 等 2013; Sun 等, 2014), 以及以马尔可夫随机场(Tarabalka 等, 2010; Zhang 等, 2011)和区域性投票(Tarabalka 等, 2009)为代表的空间正则化方法。

传统高光谱分类方法利用手工设计特征实现光谱或空谱特征的变换提取, 其性能依赖于特征工程技术和专家知识先验, 难以刻画复杂的空谱结构, 限制了分类性能的进一步提升。深度学习技术通过数据驱动分层学习策略构建连续堆叠的层次化系统, 自适应地从原始数据抽取具有鉴别性的高级特征, 为特征变换提供了另一条有效途径。相较于堆栈自编码器(Stacked AutoEncoder)(Chen 等, 2014)、深度置信网络(Deep Belief Network)(Chen 等, 2015)、循环神经网络(Recurrent Neural Network)(Liu 等, 2018)等深度模型, 卷积神经网络 CNN(Convolutional Neural Network)以其局部区域连接和权重共享的结构优势天然适应可视化数据的处理, 能以有效且灵活的方式将光谱特征与空间背景信息结合在一起, 因而逐渐成为高光谱分类领域最为流行的特征提取框架之一。

目前, 基于 CNN 的空谱联合分类方法主要分为并行学习和同步处理两种模式。并行模式的典型思路是利用一维 CNN 和二维 CNN 构建双通道卷积网络(Dual-Channel Convolutional Neural Network), 应用光谱和空间分支分别提取层次光谱特征和抽象空间特征(Zhang 等, 2017)。Chen 等

(2018)在此基础上增加三维卷积分支, 提出多通道卷积网络(Multi-Channel Convolutional Neural Network)实现多维度特征融合。此外, 光谱和空间注意力机制也被引入多分支网络形成空谱注意力模型(Ma 等, 2019; Li 等, 2020)。该类模型通过注意力模块实现光谱域和空间域的自适应选择从而增强有效波段或显著区域的表征、减少冗余波段或无关像素的干扰, 改善模型的性能和鲁棒性。高光谱图像是由一维光谱响应和二维空间信息构成的三维数据立方体, 对其应用三维 CNN 可以自然地在光谱和空间维度上同步获取深度特征, 而无需任何预处理步骤。Chen 等(2016)首次将三维 CNN 引入高光谱分类领域, 以高效的双层框架实现高鲁棒的空谱特征提取。相比二维 CNN, 深层的三维 CNN 会使得模型参数增加, 在有限训练样本下产生过拟合现象。而对三维 CNN 的网络规模加以限制又会造成模型学习能力的降低, 进而使得分类性能难达预期。为了缓解有限样本下模型优化不足的问题, 研究人员(Zhong 等, 2018; Wang 等, 2018)从架构设计入手, 引入残差连接(He 等, 2016)和稠密连接(Huang 等, 2017)改善三维 CNN 的梯度传播。此外, Roy 等(2020)提出三维 CNN 和二维 CNN 交替连接的混合模型来减少网络参数量, 并取得了良好的分类结果, 为卷积网络的架构优化提供了新思路。单一尺寸的滤波器难以表征高光谱图像复杂的空间结构, 联合多种尺寸的滤波器组成为捕捉不同粒度空间信息的有效选择(Paoletti 等, 2019; Gong 等, 2019)。Lee 和 Kwon(2017)将多尺度滤波器组和残差学习有机结合, 提出更深的上下文卷积神经网络(Contextual Deep Convolutional Neural Network), 实现多尺度局部空谱特征的学习, 有效提升了网络的空间分析能力。

由于 CNN 模型是在固定窗口内基于一定规则的空间关系学习, 不能普遍适应分布复杂、结构多样的不同邻域情形(Zhang 等, 2016)。当卷积滤波核以正方形规则窗口进行卷积运算时无法排除窗口内不利信息的干扰, 导致存在于不同地物类别间的观测像素出现信息扩散、难以得到有效区分。图 1 为多种目标像素的空间分布情形。可见: 当目标像素处于类别内部时, 无论卷积核尺寸大小都是对匀质区域的特征提取; 然而当目标像素靠近类别边界时卷积核的尺寸越大引入的异

构信息越多；当像素处于类别边界时，无论卷积核尺寸大小都存在异构信息的干扰，局部空间信息的应用甚至增加了分类难度，这也是多数误分

类区域都集中在类别边界的原因。因此如何建模异构区域的空谱特征，准确表征类边界是具有挑战性的问题。

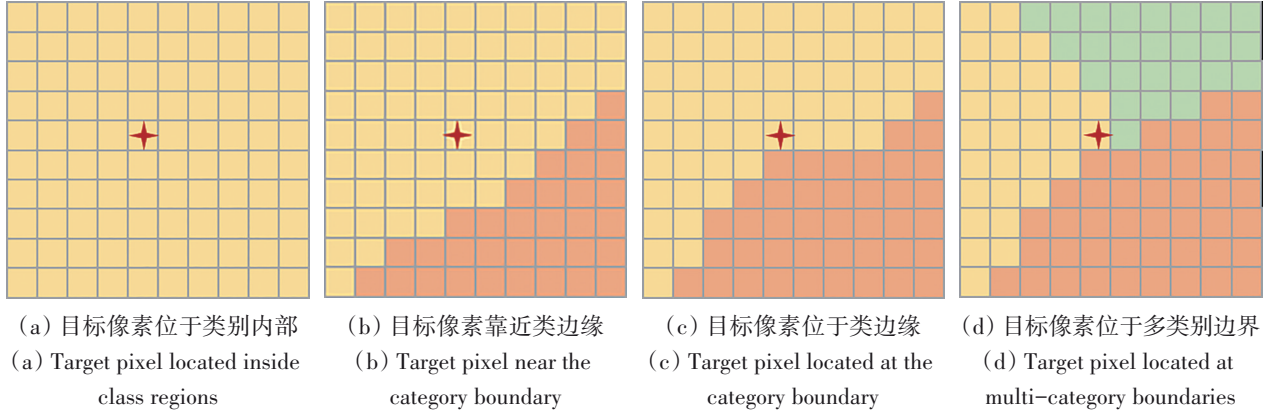


图1 目标像素的不同空间分布情形红色星号表示目标像素；不同颜色的像素属于不同地物类别

Fig. 1 Different spatial distribution scenarios of target pixel

针对该问题，Zhang等（2018）提出基于多形变体输入的卷积网络模型DRCNN（Diverse Region-based CNN），依据观测像素可能的分布倾向将中心环绕的整体性输入拆分为不同方向的输入组合，并且通过设计针对不同输入的网络架构进一步增强模型的空间分析能力。在此基础上，Xi等（2022）将多形输入进一步扩展为更加具体的多方向输入，以并行的编解码网络充分挖掘图像中丰富的空谱信息，并引入注意力光谱先验根据不同像素的邻域分布特点指导多方向特征精确融合。然而，对输入数据进行多重构造的方法在提升网络对局部空谱结构的感知力的同时，也会增加模型的复杂度、带来额外的时间和计算成本。

CNN出现边界混淆现象的根本原因在于待处理像素始终位于滤波模板的中心位置，当滤波窗口处在边界时必然会造成不同类别的信息相互扩散。受侧窗滤波（Yin等，2019）启发，本文通过改变常规卷积滤波器的操作窗口与目标像素的相对位置设计出侧窗卷积，减少跨边缘计算中无关信息的干扰，改善其边缘保持能力，并进一步提出基于多方向自适应感知的高光谱图像空谱联合分类方法。

2 研究方法

2.1 侧窗卷积

假设当前卷积层的输入特征为 $X \in \mathbf{R}^{H \times W \times L}$,

其中 H 、 W 和 L 分别表示其高度、宽度和光谱通道数，对空间坐标 (i, j) 处的像素 $X_{i,j}$ 进行常规卷积操作，则其对应位置的输出为

$$Y_{ij} = \prod_{c=1}^C \sum_{l=1}^L \sum_{a=-r}^r \sum_{b=-r}^r W_{a,b,l}^c X_{i+a, j+b}^l \quad (1)$$

式中， $X_{i,j}^l$ 表示 $X_{i,j}$ 第 l 个光谱通道的特征值， $Y_{i,j}$ 为该像素的输出向量。当前卷积层共有 C 个卷积核， $W_{a,b,l}^c$ 为第 c 个卷积核的第 l 个通道在 (a, b) 位置处的参数值。 $\prod$ 表示通道维上的级联操作， $\sum$ 表示求和操作。不同的卷积核与输入特征进行采样计算并求和，最后通过特征拼接组成输出向量。假设当前层的卷积核尺寸为 $k \times k$ ，则 $r = k/2$ 。在常规卷积中，每一个卷积核对应一个输出通道，其输出特征是对输入特征在对应位置上的加权采样，卷积核的中心对准待处理像素并线性聚合其在窗口范围内的所有邻域像素特征。常规卷积对边界像素的采样计算可能会引入无关像素的光谱信息。只有当边界像素的线性重建只考虑相同类别的邻域像素时才能有效避免类别间的信息扩散。

因此，本文考虑将目标像素看作潜在边界，将常规卷积的全窗口卷积核拆分为多个半窗滤波器进而提出侧窗卷积SWC（Side Window Convolution），仅在每个局部窗口内对目标位置进行采样更新。为了便于计算，本文将侧窗卷积定义为相对目标像素的上（U）、下（D）、左（L）、右（R）、西北（NW）、东北（NE）、西南（SW）和东南（SE）8个方向，具体空间采样范围如图2（b）—（i）所

示。侧窗卷积的计算过程表示为

$$Y_{ij} = \prod_{c=1}^C \sum_{l=1}^L \sum_{a=r_u^{row}}^{r_u^{row}} \sum_{b=r_d^{col}}^{r_d^{col}} W_{a,b,l}^c X_{i+a,j+b}^l \quad (2)$$

式中, r_u^{row} 、 r_d^{row} 以及 r_u^{col} 、 r_d^{col} 分别表示侧窗卷积在行、列上的窗口范围。图2显示不同卷积的空间滤波范围, 可以看出常规卷积将卷积核中心与目标像素对齐, 通过调整卷积核参数使其具有刻画边缘的能力, 然而当目标像素处于类别边缘时, 常规卷积无法避免无关类别的信息干扰; 侧窗卷积将目

标像素与卷积核的不同边缘或角对齐, 当侧窗卷积核的朝向与目标像素所在类别的方向一致时, 能有效避免无关类别的信息干扰。根据图2示例中目标像素的邻域空间分布, 可以看出左侧窗卷积和西北侧窗卷积能有效减少无关类别的信息干扰、增强该边界像素的表征准确性。然而, 在实际分类时目标像素的空间分布是未知的, 无法根据其邻域结构针对性地选择侧窗方向。

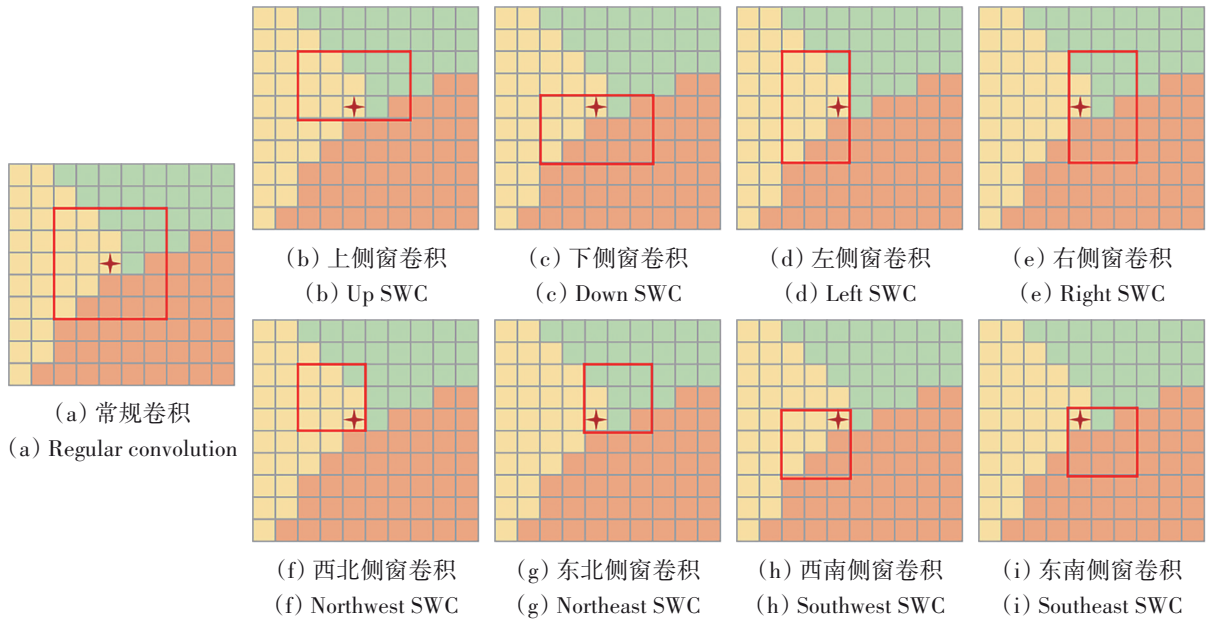


图2 常规卷积和不同方向侧窗卷积的空间滤波范围(卷积核尺寸为5, 每种颜色表示一类像素)

Fig. 2 Spatial filtering range for regular convolution and different directional side window

2.2 多方向卷积

常规卷积和侧窗卷积都是应用单一方形窗口对空谱特征进行滤波变换。因此, 本文单窗口卷积进行扩展, 将包含全窗口在内的9种方形滤波核组合在一起, 构建多方向卷积 MDC (Multi-directional Convolution), 其卷积结构如图3所示。与常规卷积相比, MDC 在常规卷积全窗口回归 (G) 的基础上增加了多种半窗回归模式, 其计算过程可以表示为

$$Y_{ij} = \prod_{s=1}^S \prod_{c=1}^{C'} \sum_{l=1}^L \sum_{p \in P_s} W_{p,l}^c X_{(ij)+p}^l \quad (3)$$

式中, $S = 9$ 为方向总数, 从1到9分别对应 G、U、D、L、R、NW、NE、SW、SE 9 个方向; $C' = C/S$ 即单个方向的输出通道数。 P_s 表示第 s 个方向上邻域像素相对目标像素的偏移坐标集合, $W_{p,l}^c$ 表示当前方向的第 c 个卷积核第 l 个通道在位置 p 处的参数

值。具体而言, MDC 在未知数据实际分布的情况下应用不同方向的滤波核对同一位置的不同邻域范围进行采样计算, 每一个滤波核负责一种方向的空间谱结构信息, 以枚举的方式组合多种上下文语义, 进而提升对边缘地物的表征能力。全局方向 (G) 的全窗口模式依旧保留了常规卷积捕获较远距离空间依赖的能力; 上 (U)、下 (D)、左 (L)、右 (R) 4 个方向建模各自半窗内的空间关系, 提供了方向性特征; 而西北 (NW)、西南 (SW)、东北 (NE) 和东南 (SE) 4 个方向的计算范围进一步收缩、紧密围绕在目标像素四周进而保证某一邻域内的光谱相似性, 对跨边界像素而言也使其类间差异扩大。从另一个角度看, MDC 囊括多种形状和大小的卷积核, 同时实现了多尺度空间特征的学习。

MDC 在建模多种空间关系的同时, 也面临通

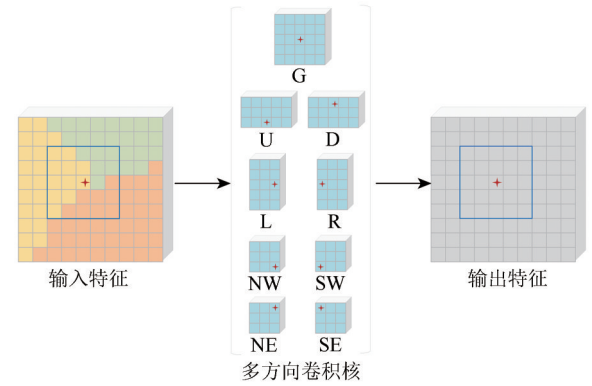
道数必须成倍于方向总数的限制, 因而当单一方向的通道数增加时, 网络参数会多倍增长。受深度可分离卷积启发, 本文进一步对MDC的卷积核进行空间维和光谱维拆分构造S3MDC (Spectral-spatial separable Muti-directional Convolution), 期望以较低的模型复杂度学习多样空谱表示。S3MDC将空谱一体化处理的多方向卷积核拆分为二维空间核和一维光谱核进而实现空间区域和光谱通道的计算解耦, 具体结构如图4所示。S3MDC的计算过程由多方向的空间特征提取和以方向划分的通道维线性变换组成, 其中单一方向的计算公式如下:

$$Z_{ij} = \prod_{s=1}^S \prod_{l=1}^{L'} \sum_{p \in P_s} W_{s,p,l}^{\text{spa}} X_{(ij)+p}^{s \times L' + l} \quad (4)$$

$$Y_{ij} = \prod_{s=1}^S \prod_{c=1}^{C'} \sum_{l=1}^{L'} W_{s,c,l}^{\text{spe}} Z_{ij}^{s \times L' + l} \quad (5)$$

式中, $W_{s,p,l}^{\text{spa}}$ 表示第 s 个方向第 l 个空间卷积核在位置 p 处的参数值。由于空间卷积核的通道维与输入特征的通道维一一对应, 输入特征通道总数 L 应当被 S 整除。 L' 为单个方向的输入通道数, 即单个方向的空间卷积核总数。 $Z_{i,j}$ 表示空间卷积后位置 (i, j) 的空间特征。式中, $W_{s,c,l}^{\text{spe}}$ 表示第 s 个方向上

第 c 个光谱卷积核的第 l 维变换参数。空间特征按通道维划分, 表示不同方向的空间信息, 因此在光谱维变换时通过分组卷积对空间特征进行光谱维特征提取, 以保证各个方向空间信息的独立性。在相同的输入输出下, 常规卷积的参数量为 $L \times C \times k^2$, MDC的参数量为 $L \times C/9 \times (4k^2 + 4k + 1)$, S3MDC的参数量为 $L/9 \times (4k^2 + 4k + 1 + C)$ 。空谱分离策略有效降低了多方向卷积的网络参数。



G, 全局方向; U, 上方; D, 下方; L, 左侧; R, 右侧; NW, 西北方; NE, 东北方; SW, 西南方; SE, 东南方。

图3 多方向卷积结构示意图

Fig. 3 Illustration of the multi-directional convolution structures

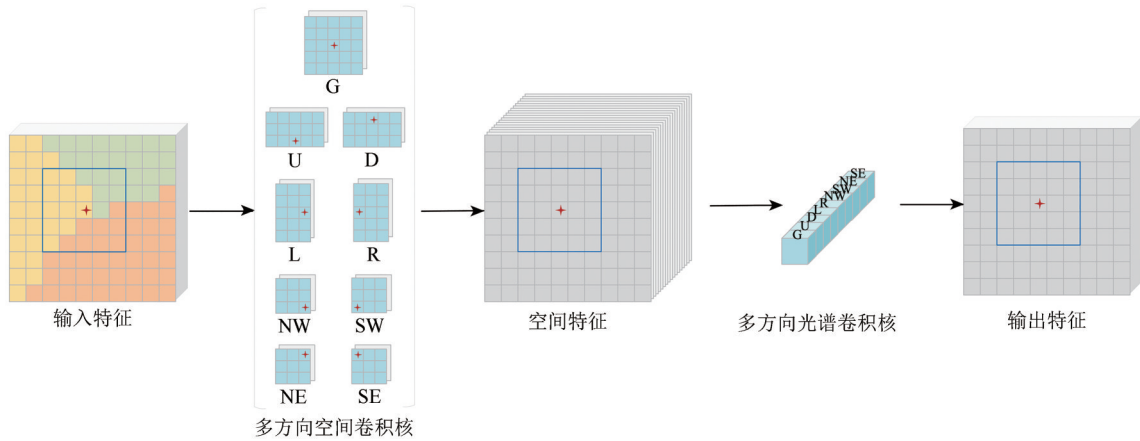


图4 空谱分离多方向卷积结构示意图

Fig. 4 Illustration of the spectral-spatial separable multi-directional convolution structure

2.3 多方向自适应感知分类网络

本文以S3MDC联合DAAM (Directional Adaptive Aware Module) 作为基本单元、以密集连接为架构提出多方向自适应感知网络MDAAN (Muti-directional Adaptive Aware Network), 灵活提取多方向空谱结构信息, 实施高光谱图像的精细分类,

具体分类框架如图5所示。

假设MDAAN的输入为高光谱图像 $X \in \mathbf{R}^{H \times W \times B}$, 输出为像元的类别概率 $Y \in \mathbf{R}^{H \times W \times C}$, 这里 H , W , B 和 C 分别表示图像的高度, 宽度, 波段数和地物类别数。考虑到较高的光谱维度会增加后续特征提取的计算量, 首先应用 1×1 的2D卷积实施光谱降维, 利用其在光谱维上的全局计

算优势进行线性转换，以降低光谱维的冗余度、提升谱维可分性。然后以顺序连接的实例归一化 IN (Instance Normalization)、线性整流函数 ReLU (Rectified Linear Unit)、S3MDC 以及 DAAM 作为基本网络单元，以 5 个基本单元的密集连接来实现

MDAAN 的多方向特征自适应提取。最后以 S3MDC 和 DAAM 对多方向特征进行通道压缩和空间信息聚合，联合 1×1 的全卷积分类层和 Softmax 层实现像素类别概率的输出。

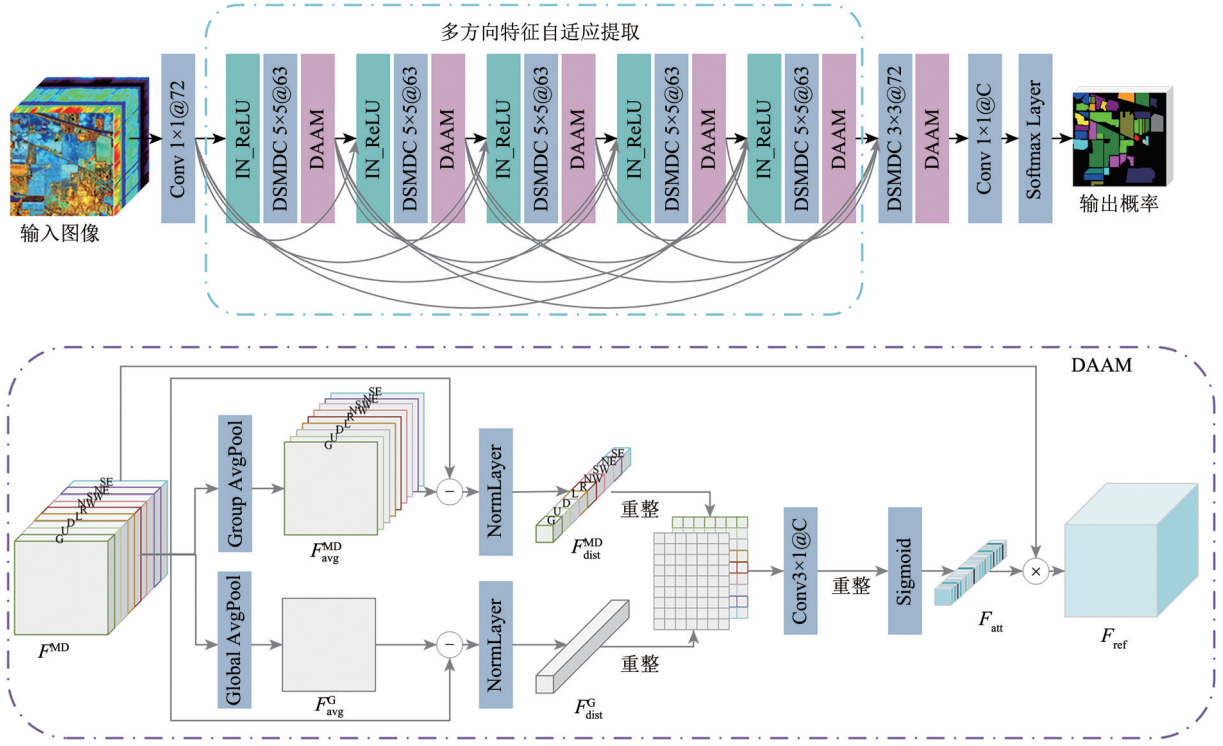


图5 多方向自适应感知网络结构图

Fig. 5 The structure of the proposed multi-directional adaptive aware network

其中，DAAM (Directional Adaptive Aware Module) 是一种多方向自适应感知的注意力模块，能对空谱分离多方向卷积方向自适应感知模块 S3MDC (Spectral-spatial separable Muti-directional Convolution) 输出的多方向特征进行再次细化，提升网络对方向信息的感知能力，具体结构如图5所示。假设当前 DAAM (Directional Adaptive Aware Module) 输入的多方向特征为 F^{MD} ，共包含 9 个方向。DAAM 以双分支结构从全局和各个方向内 2 个层面获得各个通道的表示特征，并在此基础上建模不同方向之间的关系，最终实现自适应的方向特征增强。具体来说，首先利用通道维的平均池化操作获得输入特征图上各个方向的均值。由于多方向特征中不同通道的特征图分属于不同的方向，本文以方向为单位对特征图进行平均池化，即分组平均池化，获得多方向均值特征 $F^{\text{MD}}_{\text{avg}}$ 。为了增强通道维信息的表征能力，DAAM 的另一个分

支对全通道进行平均池化，获得全局均值特征 $F^{\text{G}}_{\text{avg}}$ 。注意力机制本质上是通过对特征图的抽象表示和信息交互来进行通道间的关系建模，自适应地对关键信息进行权重调整 (Liu 等, 2023)。因此，本文以逐元素相减的形式将各个通道的特征图与所在方向的均值特征以及全局均值进行逐一对比，利用 L2 范数度量其差异性，以获得每个通道与所在方向的均值特征以及全局均值的距离特征 $F^{\text{MD}}_{\text{dist}}$ 、 $F^{\text{G}}_{\text{dist}}$ ，进行方向内部和全局的信息交互。最后联合二者作为各个通道的表示特征，利用特征重整和卷积操作建模方向间的局部关系，获得的注意力向量 F_{att} ，并与输入特征进行哈达玛积运算，得到调整后的特征 F_{ref} ，实现最终的方向自适应感知增强。因此 DAAM 的计算过程可以表示为

$$F^{\text{MD}}_{\text{avg}} = \text{GroupAvgPool}(F^{\text{MD}}) \quad (6)$$

$$F^{\text{G}}_{\text{avg}} = \text{GlobalAvgPool}(F^{\text{MD}}) \quad (7)$$

$$F^{\text{MD}}_{\text{dist}} = \text{Re shape}(\text{Norm}(F^{\text{MD}} - F^{\text{MD}}_{\text{avg}})) \quad (8)$$

$$F_{\text{dist}}^G = \text{Re shape}(\text{Norm}(F^{\text{MD}} - F_{\text{avg}}^G)) \tag{9}$$

$$F_{\text{att}} = \text{Sigmoid}(\text{Conv}([F_{\text{dist}}^{\text{MD}}, F_{\text{dist}}^G])) \tag{10}$$

$$F_{\text{ref}} = F^{\text{MD}} \cdot \text{Re shape}(F_{\text{att}}) \tag{11}$$

在训练阶段，本文使用交叉熵损失来优化网络参数。由于高光谱图像分类是基于单张图像的像素级分类任务，训练样本和测试样本来自同一图像，因此，在训练阶段只对训练样本计算损失。

3 实验与结果分析

3.1 实验数据

本文在3个公开的高光谱数据集 Indian Pines、University of Pavia 以及 Kennedy Space Center 上验证MDAAN的有效性，各个数据集的标记样本分布如表1所示。

表1 不同数据集的标记样本统计

Table 1 Labeled sample statistics for different datasets

序号	Indian Pines		University of Pavia		Kennedy Space Center	
	类别	样本数	类别	样本数	类别	样本数
1	苜蓿	54	柏油路面	6631	矮树丛	716
2	免耕玉米	1434	草地	18649	湿地柳	243
3	少耕玉米	834	碎砾石	2099	中生森林	256
4	纯净玉米	234	树木	3064	湿地松	252
5	草地/牧地	497	金属板材	1345	橡木/阔叶	161
6	草地/树木	747	裸地	5029	阔叶树	229
7	草地/割草	26	沥青	1330	湿地	105
8	干草料堆	489	砖砌路面	3682	禾草状沼泽	431
9	燕麦	20	阴影	947	米草湿地	520
10	免耕大豆	968	—	—	香蒲湿地	404
11	少耕大豆	2468	—	—	盐沼	419
12	纯净大豆	614	—	—	泥滩	503
13	小麦	212	—	—	水	927
14	木材	1294	—	—	—	—
15	建筑—草地—树木—机器	380	—	—	—	—
16	石岗塔	95	—	—	—	—
标记样本总数		10366	42776		5211	

Indian Pines（IP）数据集由机载可见光/近红外成像光谱仪 AVIRIS 对美国印第安纳州西北部的 Indian Pines 试验区采集而成，是最早的用于高光谱图像分类的测试数据。数据集包含 145×145 像素和 224 个反射波段，空间分辨率为 20 m，波长范围为 400—2500 nm，其中去除噪声和吸水带后保留 200 个波段作为研究的对象。该场景包括 16 类地物，均为农作物或植被，地物类型相似度高，且各类地物数目不一、分布不均衡。

University of Pavia（UP）数据集是由机载反射光学成像光谱仪 ROSIS 对意大利的帕维亚大学采集而成，包含 610×340 个像素和 115 个波段，波长范围覆盖 430—860 nm，空间分辨率为 1.3 m，其中剔除 12 个噪声波段保留 103 个剩余波段用于分

类研究。该数据集覆盖 9 类地物类型，包括树、沥青道路、砖块、阴影等，是典型的城市场景。

Kennedy Space Center（KSC）数据集由 AVIRIS 光谱成像仪在美国佛罗里达州的肯尼迪航天中心采集而成，包含 512×614 个像素和 224 个波段，波长区间为 400—2500 nm，空间分辨率为 18 m，在去除吸水和噪声波段后保留余下的 176 个波段进行实验。该数据集包含 13 个土地覆盖类别。

3.2 实验设置

3.2.1 训练数据及评价指标

实验所用数据集均都被划分为训练集、验证集和测试集 3 部分。在 IP 数据集和 KSC 数据集上随机选取每类样本的 5% 和 1% 作为训练集和验证

集，剩余样本作为测试集；在UP数据集上每类分别随机选取1%和1%的标记样本用于训练和验证，余下样本作为测试集。

本文采用类别精度PA（Per-class Accuracy）、总体精度OA（Overall Accuracy）、平均精度AA（Average Accuracy）以及Kappa系数（Kappa Coefficient）作为模型分类性能的评价指标。所有实验随机重复5次，取均值和标准差作为最终结果。

3.2.2 对比方法

为验证所提方法的性能，本文选取了7个高光谱图像分类方法作为对比，具体包括：SSRN（Zhong等，2018）、FDSSC（Wang等，2018）和HybridSN（Roy等，2020）3种较为先进的深度学习分类方法；基于多形态输入的深度学习分类方法DRCNN（Zhang等，2018）；基于混合深度可分离卷积的多尺度残差分类方法MSRN_B（Gao等，2021）；以及2种基于注意力机制的分类方法DBDA（Li等，2020）和FreeNet（Zheng等，2020）。其中

FreeNet和本文提出的MDAAN是基于全图输入的分类框架，其余方法均为基于图像块输入的分类方法。上述对比方法的超参数设置均遵循原始论文的推荐参数。

本文所有实验均在Windows10、NVIDIA GeForce RTX 4090的环境下运行。

3.3 分类性能分析

3.3.1 IP数据集

不同方法在IP场景下的数值分类结果见表2。可见：本文提出的方法MDAAN较其他方法能获得更优越的总体精度，其中OA达到了97.67%±0.44%，相比其他方法提高了0.52%—6.89%，同时标准差也小于其他方法，表明MDAAN在不同训练样本选择下性能更为稳定。相比多形态输入的DRCNN，本文方法从改进基础卷积结构入手，无需为每个方向分配单独的特征提取分支，模型灵活、轻量、更易优化，有助于提升模型在有限样本下的分类性能。此外，MDAAN在7个类别上均达到最优精度，在3个类别上取得次优精度。

表2 不同方法在IP数据集上的分类结果
Table 2 Classification results of different methods on IP dataset

/%								
类别	SSRN	FDSSC	HybridSN	DRCNN	MSRN_B	DBDA	FreeNet	MDAAN
1	<u>85.6±9.75</u>	47.2±34.65	82.0±7.8	64.8±13.95	89.6±13.29	84.0±12.59	89.6±9.75	82.4±10.31
2	95.74±2.22	<u>95.86±1.15</u>	91.58±2.46	88.4±2.18	94.97±2.41	92.4±5.61	95.5±2.93	96.45±1.94
3	<u>95.43±2.35</u>	94.05±4.42	88.63±6.15	88.3±4.83	94.02±2.32	92.95±2.03	93.41±3.52	95.96±1.21
4	95.16±4.85	95.25±7.7	88.31±11.43	84.93±6.0	97.26±2.65	92.88±6.3	<u>97.9±2.34</u>	99.27±1.03
5	96.79±1.32	94.69±1.51	93.96±1.78	94.56±3.9	94.95±2.07	93.36±1.23	95.42±2.49	<u>96.02±2.51</u>
6	99.32±0.5	<u>99.12±0.64</u>	96.55±1.76	98.2±0.95	96.32±2.93	97.97±0.58	97.4±1.72	98.77±0.77
7	<u>93.04±5.9</u>	40.0±20.83	92.17±10.43	86.09±11.14	81.74±20.47	77.39±26.29	92.17±10.43	95.65±4.76
8	<u>99.96±0.09</u>	100.0±0.0	99.87±0.26	99.35±1.1	100.0±0.0	99.52±0.56	99.13±1.22	99.91±0.17
9	44.44±26.76	0.0±0.0	53.33±22.39	44.44±25.34	41.11±33.44	76.67±9.56	58.89±19.12	<u>72.22±26.06</u>
10	94.92±2.15	95.03±2.38	93.42±2.03	89.81±2.78	95.4±2.79	<u>96.15±1.36</u>	96.24±2.04	96.24±2.43
11	97.92±1.58	98.32±1.21	97.48±1.64	92.04±1.64	97.75±1.85	97.34±0.64	<u>98.46±0.86</u>	99.02±0.68
12	<u>96.01±0.87</u>	96.53±0.86	86.84±2.05	76.91±2.06	85.69±6.08	92.71±1.69	95.76±1.52	95.24±2.88
13	99.29±0.76	<u>98.48±1.5</u>	97.07±4.41	96.06±4.08	97.58±3.63	<u>98.48±2.09</u>	94.75±6.18	98.08±2.36
14	99.62±0.33	99.13±0.42	98.36±1.19	96.35±0.68	99.18±0.6	98.77±0.61	98.98±0.97	<u>99.44±0.39</u>
15	97.2±3.03	94.57±6.81	96.47±3.05	80.39±8.2	98.26±1.54	96.92±2.68	<u>98.71±1.5</u>	98.88±1.97
16	<u>97.98±2.4</u>	96.63±1.59	91.46±10.93	96.85±2.79	96.63±3.1	98.43±1.35	94.38±4.61	95.51±4.66
OA	<u>97.15±0.49</u>	96.48±0.88	94.45±0.94	90.88±0.63	95.99±0.98	95.82±1.06	96.9±0.67	97.67±0.44
AA	93.03±0.9	84.05±3.92	90.47±2.52	86.09±2.13	91.28±2.73	92.87±2.86	<u>93.54±1.77</u>	94.94±2.05
Kappa	<u>96.75±0.56</u>	95.99±1.0	93.67±1.08	86.09±2.13	95.43±1.12	95.23±1.21	96.47±0.76	97.35±0.51

注：每个类别选取5%标记样本作为训练数据。粗体为最优分类精度，带下划线的数值为次优分类精度。

图6为IP数据集的假彩色图像、地物真值图以及不同方法下的全分类结果图（包含对无标签背景像素的识别）。可见，基于图像块输入的分类图的建模空间受输入尺寸的限制，导致类别边界

较为模糊，空间平滑性较弱；FreeNet和本文MDAAN的分类图能兼顾匀质区域完整性并保持地物分布的细节；MDAAN相比FreeNet类别边界更为清晰，对边界区域的刻画更为准确。

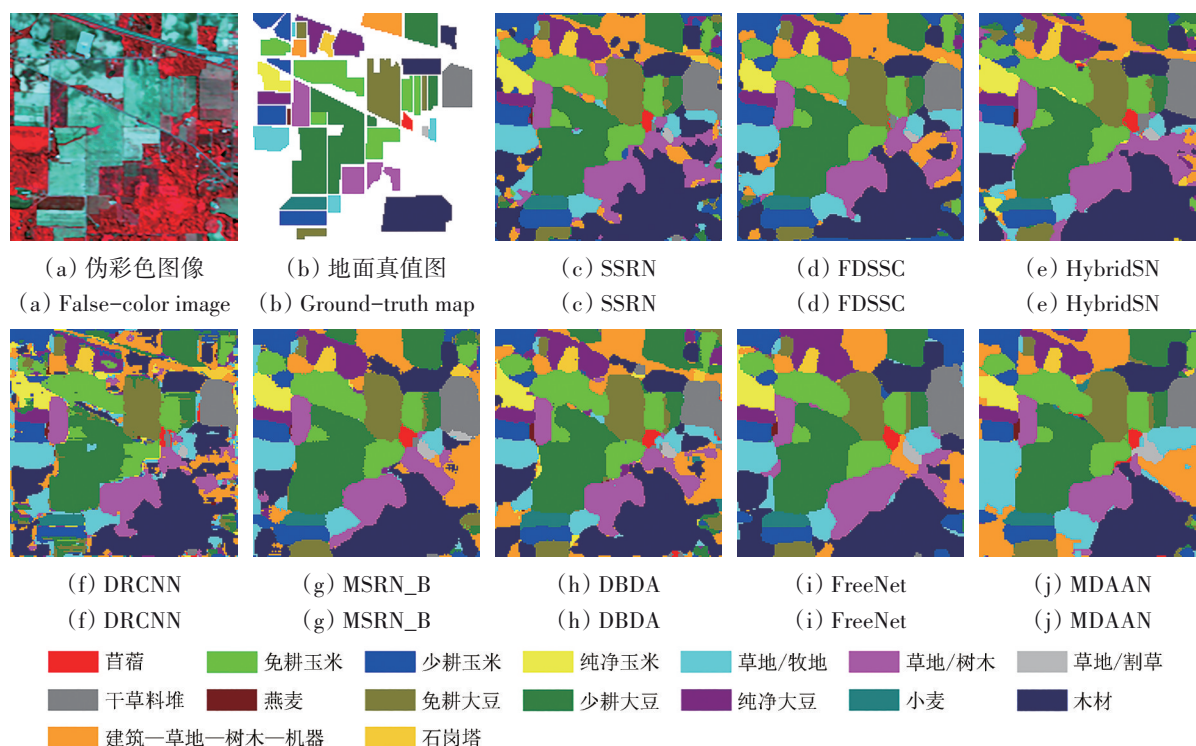


图6 不同方法在IP数据集上的分类图

Fig. 6 Classification maps obtained by different methods for the IP datasets

3.3.2 UP数据集

表3为不同方法在UP场景下的数值分类精度，每个指标的最优精度和次优精度同样以粗体和下划线突出。可知；从总体精度来看，本文提出的MDAAN在OA、AA、Kappa系数3个指标上均获得最佳结果，其中OA达到了 $99.4\% \pm 0.18\%$ ，相比其他方法提高了0.49%—7.31%；从单一类别来看，MDAAN对所有类别的识别准确率都超过95%，其中在4个类别上取得最优或次优精度。

图7为UP数据集的假彩色图像、地物真值以及不同方法下的全分类结果图。较对比方法而言，MDAAN的分类图对匀质和非匀质区域均能提供较为准确的结果，在平滑边缘和保持细节上取得平衡。其中在Gravel和Bricks2类上的准确识别也与数值结果相互验证。UP数据集的结果充分表明MDAAN能适应分布复杂的城市场景。

3.3.3 KSC数据集

表4显示不同方法在KSC场景下的数值分类精度，每个指标的最优精度和次优精度以粗体和下划线突出。本文提出的MDAAN在OA、AA、Kappa系数3个评价指标上均达到最优精度，其中OA达到了 $99.64\% \pm 0.25\%$ ，相比其他方法提高了0.43%—5.48%。从单类别精度来看，MDAAN在11个类别上获得最高精度。其中，在7个类别上实现100%的准确识别。

KSC数据集的假彩色图像、地物真值以及各种方法的全分类结果见图8。与数值结果一致，MDAAN与假彩色图像的一致性最高。3个不同数据集上的结果表明，以多方向侧窗建模不同方向的局部空间结构有助于减少无关像素干扰，能自适应地捕捉区分性特征，在不同类型的地物分布下均能获得准确可靠的分类结果，具有较好的普适性。

表3 不同方法在UP数据集上的分类结果

Table 3 Classification results of different methods on UP datasets

/%								
类别	SSRN	FDSSC	HybridSN	DRCNN	MSRN_B	DBDA	FreeNet	MDAAN
1	99.21±1.07	99.15±0.63	98.55±0.68	90.89±3.55	<u>99.3±0.48</u>	98.46±0.43	98.35±0.85	99.85±0.25
2	99.42±0.6	99.96±0.08	<u>99.93±0.1</u>	98.66±1.18	99.87±0.16	99.73±0.2	99.75±0.31	99.87±0.14
3	91.08±7.32	92.79±5.07	85.49±2.63	63.73±5.66	94.31±3.18	85.82±5.35	<u>95.39±4.4</u>	98.45±1.54
4	97.38±1.42	<u>96.71±1.31</u>	92.17±2.96	92.37±2.72	95.9±0.66	93.59±1.18	93.68±0.97	95.3±0.84
5	<u>99.98±0.03</u>	99.85±0.2	99.39±0.42	98.97±0.78	99.15±1.1	99.56±0.25	100.0±0.0	99.97±0.06
6	98.23±1.44	<u>99.95±0.11</u>	98.83±1.05	94.37±2.46	98.91±1.29	99.51±0.39	99.98±0.04	<u>99.95±0.1</u>
7	97.73±2.41	100.0±0.0	92.93±6.82	62.12±18.21	98.79±0.84	98.28±1.09	99.45±0.56	99.78±0.23
8	92.48±8.21	96.41±1.82	92.63±1.58	81.02±12.47	<u>98.74±1.1</u>	95.44±0.45	98.18±1.34	99.66±0.28
9	99.98±0.04	98.53±0.46	83.54±10.67	96.46±3.09	96.14±1.64	<u>99.55±0.59</u>	96.98±2.63	97.0±0.91
OA	98.07±0.64	<u>98.91±0.23</u>	97.1±0.5	92.09±1.38	98.88±0.22	97.96±0.39	98.71±0.36	99.4±0.18
AA	97.28±0.72	<u>98.15±0.52</u>	93.72±1.64	86.51±2.28	97.9±0.27	96.66±0.76	97.97±0.45	98.87±0.39
Kappa	97.45±0.85	<u>98.55±0.31</u>	96.14±0.67	89.5±1.82	98.51±0.29	97.3±0.51	98.29±0.47	99.2±0.24

注：每个类别选取1%标记样本作为训练数据。粗体为最优分类精度，带有下划线的数值为次优分类精度。

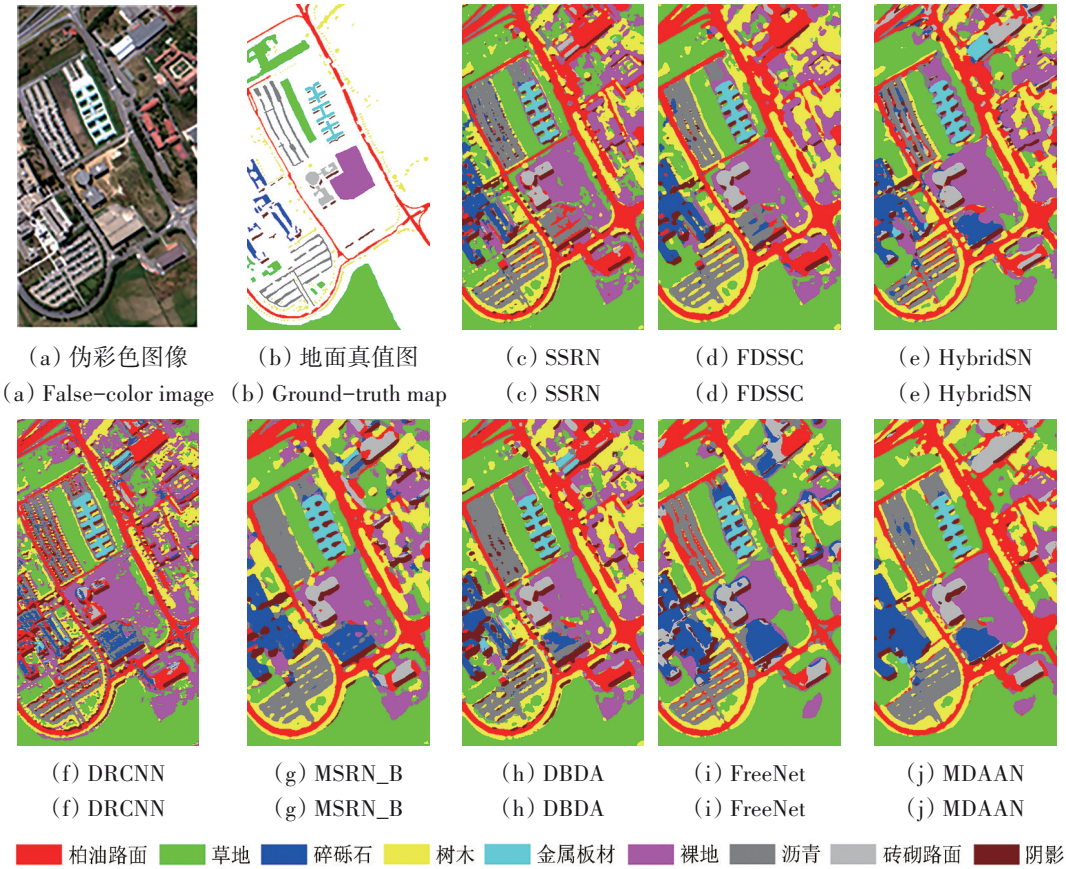


图7 不同方法在UP数据集上的分类图

Fig. 7 Classification maps obtained by different methods for the UP datasets

表 4 不同方法在 KSC 数据集上的分类结果

Table 4 Classification results of different methods on KSC datasets

类别	SSRN	FDSSC	HybridSN	DRCNN	MSRN_B	DBDA	FreeNet	MDAAN
1	99.47±0.57	100.0±0.0	97.73±2.33	98.82±1.04	99.94±0.11	99.97±0.06	99.5±1.01	99.3±1.2
2	<u>94.71±7.29</u>	94.45±9.06	82.64±12.48	89.69±9.61	92.78±10.78	89.25±4.14	93.39±7.57	96.83±2.94
3	99.33±0.97	<u>99.83±0.33</u>	97.5±3.61	95.42±4.31	98.67±1.89	92.75±7.56	99.83±0.2	100.0±0.0
4	90.85±10.03	71.95±19.56	77.29±4.06	51.19±9.56	91.19±8.12	75.25±10.28	<u>95.76±4.73</u>	98.14±2.15
5	85.6±15.39	63.2±29.44	81.47±10.79	54.0±15.63	85.73±7.33	80.53±3.25	99.87±0.27	<u>99.47±1.07</u>
6	95.79±4.62	98.6±2.17	90.75±4.35	94.95±4.99	96.07±4.79	94.86±2.63	<u>99.35±1.31</u>	99.91±0.19
7	88.25±16.23	87.01±15.2	98.97±2.06	95.05±8.44	99.59±0.82	95.67±7.15	100.0±0.0	100.0±0.0
8	<u>99.55±0.55</u>	99.41±1.19	94.21±4.41	97.18±1.02	98.22±3.56	96.49±4.11	98.37±3.15	99.95±0.1
9	99.47±1.07	100.0±0.0	98.77±1.76	<u>99.96±0.08</u>	<u>99.96±0.08</u>	98.57±2.57	100.0±0.0	100.0±0.0
10	<u>99.89±0.21</u>	100.0±0.0	96.67±4.05	97.72±2.06	100.0±0.0	95.71±5.33	99.63±0.74	100.0±0.0
11	<u>99.9±0.12</u>	99.69±0.37	99.49±0.48	99.29±0.71	100.0±0.0	99.19±0.98	100.0±0.0	100.0±0.0
12	97.45±3.02	98.98±1.26	96.52±1.98	94.78±1.69	98.6±2.4	98.6±2.06	<u>99.96±0.08</u>	100.0±0.0
13	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0
OA	97.99±1.35	96.76±1.25	95.41±0.87	94.06±0.59	98.26±0.54	96.07±1.25	<u>99.21±0.65</u>	99.64±0.25
AA	96.17±3.11	93.32±2.77	93.23±1.73	89.85±1.14	96.98±0.66	93.6±1.54	<u>98.9±0.92</u>	99.51±0.33
Kappa	97.76±1.5	96.39±1.39	94.88±0.97	93.38±0.66	98.06±0.6	95.62±1.39	<u>99.12±0.73</u>	99.59±0.28

注：每个类别选取 5% 标记样本作为训练数据。粗体为最优分类精度，带有下划线的数值为次优分类精度。

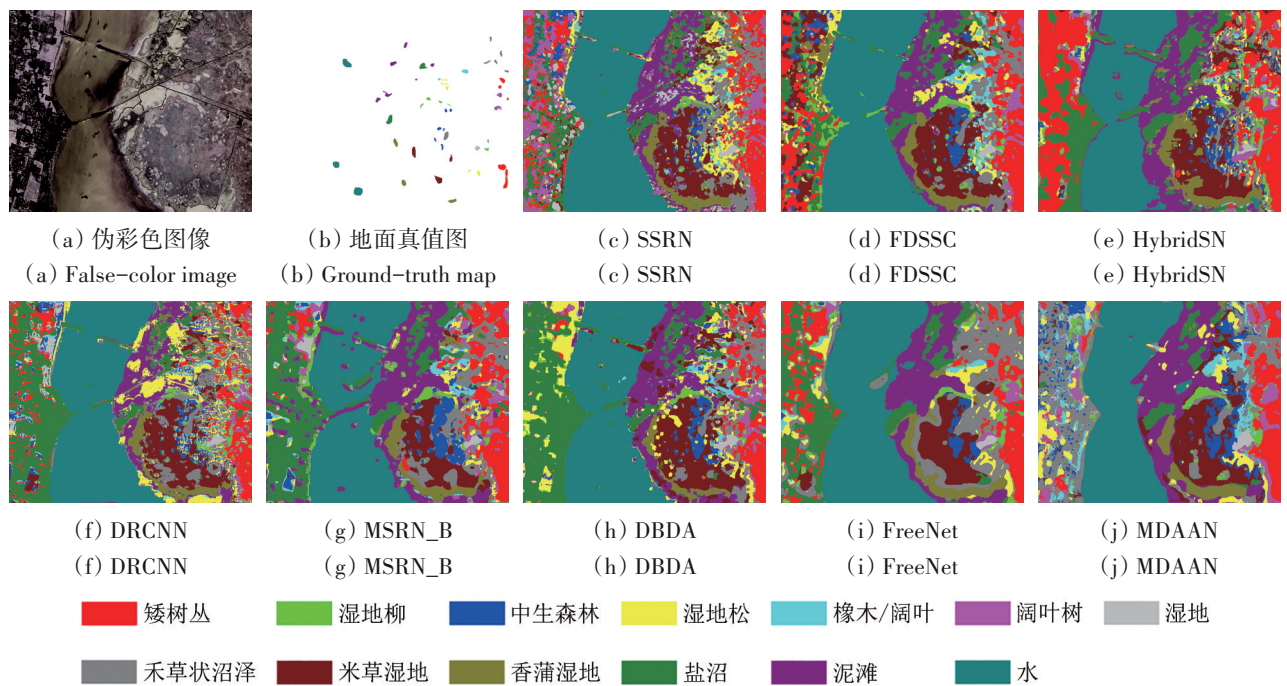


图 8 不同方法在 KSC 数据集上的分类图

Fig. 8 Classification maps obtained by different methods for the KSC datasets

3.4 分析与讨论

3.4.1 小样本下的分类性能分析

为了进一步分析验证 MDAAN 在有限训练样本

下的有效性和鲁棒性，本文对 3 个数据集选取不同数量的训练样本进行对照实验。具体来说，对 3 个数据集每类随机选取 5、10、15、20 和 25 个标记样本作为训练集，每类随机选取 5 个标记样本作为

验证集，其余样本作为测试集，以OA为评价指标，每组实验重复5次取均值作为性能指标，并在图9中绘制不同方法的结果折线图。

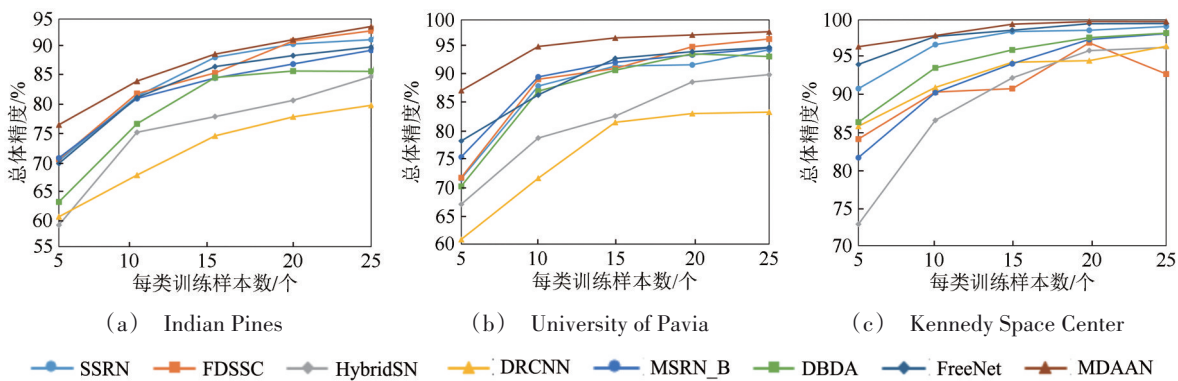


图9 不同数据集在不同训练样本规模下的分类性能对比

Fig. 9 Comparison of classification performance of different datasets under different numbers of training samples

整体而言，3个数据集上的实验结果均表明MDAAN在小样本情况下能保持稳定可靠的学习能力，分类性能优于其他方法。具体来说，对于IP数据集，所有方法的分类精度都随着分类样本的增加而平稳上升，MDAAN相比其他方法的领先优势随训练样本数量的减少而扩大，并始终保持最高的分类精度。对于UP数据集，MDAAN的分类性能较对比方法具有显著的优势。对KSC数据集，MDAAN在每类选取15个训练样本的情况下可以达到99%以上的总体精度。

3.4.2 计算复杂度分析

表5为不同方法的运行时间以及参数量统计。

模型参数量统计以IP数据集为准。从时间开销看：MDAAN在3个数据集上的训练时间较长，其原因在于MDAAN的多方向卷积设计卷积核的重整操作，并包含注意力机制的运算，使得训练时间增加；而测试时间的比较表明基于全图输入的分类框架在推理时间上具有较大优势，其时间开销与图像尺寸正相关，而不受训练样本和测试样本的数量影响。从网络参数看：MDAAN相对其他方法模型参数量大大降低，其原因在于空谱分离策略和多形态卷积核相比常规卷积具有更低的模型复杂度。应用多形输入的DRCNN为每种输入提供对应的处理分支导致其参数量高于其他方法。

表5 不同方法的运行时间以及参数量统计

模型	IP		UP		KSC		参数量/M
	训练时间/s	测试时间/s	训练时间/s	测试时间/s	训练时间/s	测试时间/s	
SSRN	16.84	0.76	6.96	1.92	8.97	0.32	0.36
FDSSC	83.03	0.94	106.04	3.30	70.90	0.57	1.23
HybridSN	17.40	0.57	7.37	1.43	6.48	0.17	5.12
DRCNN	48.50	0.67	16.14	1.84	22.90	0.29	8.6
MSRN_B	7.56	0.33	5.78	1.42	5.18	0.16	0.14
DBDA	50.46	2.08	26.80	5.38	28.08	0.88	0.61
FreeNet	19.39	0.02	132.30	0.08	173.92	0.17	2.65
MDAAN	30.42	0.02	161.22	0.10	389.60	0.19	0.04

3.4.3 不同卷积模块的对比分析

为验证空谱分离多方向卷积模块的有效性，在删除DAAM的情况下将MDAAN模型中的深度可分离多方向卷积模块分别替换为常规卷积、深度

分离卷积、多方向卷积以及不同方向的侧窗卷积模块进行分类性能对比，结果如表6所示。每组实验重复5次取OA、AA和Kappa的均值作为性能指标。可见：整体而言，多方向卷积的分类性能优

于规则正方形窗口的常规卷积，空谱分离多方向卷积的精度也优于常规的深度分离卷积，这充分表明同等卷积架构下多形态滤波器对地物特征的学习效果优于规则正方形滤波器，多方向特征也能提供区分性更强的结果。对侧窗卷积而言，上（U）、下（D）、左（L）、右（R）4种侧窗卷积的

空间表征相比西北（NW）、东北（NE）、西南（SW）和东南（SE）4种侧窗卷积更为有效。其原因在于滤波器尺寸的差别，由图2可见：相比常规卷积，四边侧窗卷积的滤波器尺寸减少1/2，四角侧窗卷积的滤波器尺寸进一步减少1/4。

表6 不同卷积模块分类性能对比

Table 6 Comparison of classification performance of different convolution modules

模型	IP			UP			KSC		
	OA	AA	Kappa	OA	AA	Kappa	OA	AA	Kappa
侧窗卷积(U)	94.01	88.30	93.17	97.28	94.54	96.39	97.00	95.04	96.66
侧窗卷积(D)	95.01	87.50	94.32	97.90	96.02	97.21	96.93	95.01	96.58
侧窗卷积(L)	94.14	86.84	93.32	95.81	91.84	94.42	97.76	96.64	97.50
侧窗卷积(R)	95.90	91.69	95.33	98.01	95.97	97.36	96.96	94.51	96.60
侧窗卷积(NW)	92.28	83.92	91.19	94.13	88.63	92.16	95.63	93.24	95.13
侧窗卷积(NE)	93.86	89.33	93.01	96.19	92.82	94.94	95.33	91.60	94.80
侧窗卷积(SW)	93.31	83.64	92.37	94.59	89.74	92.80	96.22	94.49	95.79
侧窗卷积(SE)	94.61	87.22	93.85	97.62	95.20	96.84	93.31	83.64	92.37
常规卷积	95.76	90.72	95.17	97.64	95.45	96.87	97.59	96.11	97.31
深度分离卷积	97.23	93.46	96.84	99.22	98.34	98.96	99.36	98.94	99.29
多方向卷积	95.87	90.01	95.30	97.90	95.90	97.22	98.44	97.23	98.26
空谱分离多方向卷积	97.54	93.87	97.15	99.30	98.48	99.08	99.42	98.98	99.36

注:粗体为最优分类精度。

对于IP和UP数据集，下（D）侧窗卷积的分类性能优于上（U）侧窗卷积，右（R）侧窗卷积的分类性能优于左（L）侧窗卷积、常规卷积以及多方向卷积；与之一致，下（D）侧窗与右（R）侧窗重叠的东南侧窗（SE）卷积的分类结果优于其他3个四角侧窗卷积；对于KSC数据集，上（U）侧窗卷积的分类精度略微优于下（D）侧窗卷积，左（L）侧窗卷积的分类性能优于右（R）侧窗卷积和常规卷积；与之一致，位于目标像元左侧的西北（NW）侧窗卷积和西南（SW）侧窗卷积的分类精度分别优于相同水平位置的东北（NE）侧窗卷积和西北（SE）侧窗卷积。上述结果表明不同的数据集存在不同的空间分布倾向，选择正确的空间方向能获得优于常规卷积的学习效果。由于预先无法得知数据集的空间分布倾向，只有将不同方向的学习能力整合在一起，为像素提供多样化的结构表示、自适应地拟合不同的空间上下文分布才能从根本上提升网络的表征能力。上述结果和分析表明了空谱分离多方向卷积的设计合理性和有效性，以及根据目标像素特性进行方

向自适应感知的必要性。

3.4.4 方向自适应感知模块的有效性分析

本文提出的DAAM具有2个关键步骤：（1）对不同方向的信息进行有效表征；（2）如何依托方向信息捕捉通道间的关系。针对上述关键步骤，本文设计对应消融实验来验证DAAM模块架构的有效性。具体而言，在多方向信息的表征方面，对DAAM包含的多方向分组均值对比策略、全局均值对比策略以及在注意力机制中经典的全局均值池化策略进行消融对比，具体模块结构如图10所示。在通道间关系建模方面，考虑对同方向的不同通道（方向内）、不同方向的相同通道（方向间）、不同方向的多个通道（方向内和方向间）以及不同通道（通道间）进行局部关系学习，对应模块结构如图11所示。其中，通道间关系建模表示直接以一维CNN对通道表示向量进行邻域关系学习，无需考虑通道所在方向。每组实验均重复5次，取OA均值作为性能指标，结果如表7所示。

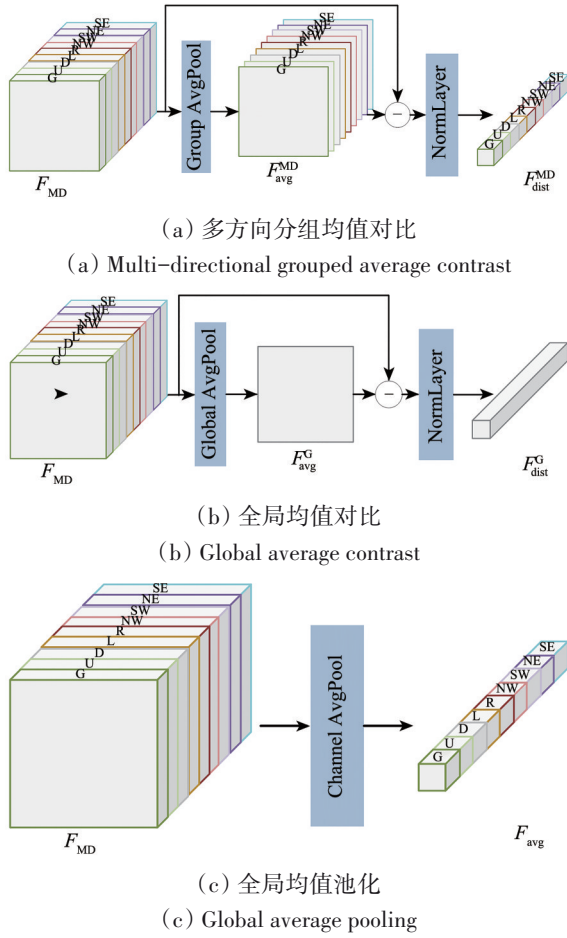


图 10 不同的多方向信息表示策略

Fig. 10 Different multi-directional information representation strategies

保持关系建模机制一致，对 IP 和 UP 数据集，利用多方向分组均值对比或全局均值对比获得通道表示的方式能取得更好的分类性能；对 KSC 数据集，使用全局均值对比的分类结果也优于直接对通道进行全局均值池化的策略；将多方向分组均值对比和全局均值对比联合能在 3 个数据集上取得最佳的分类精度，表明本文提出的双分支通道表示提取结构的有效性。保持通道表示机制一致，对 IP 和 KSC 数据集，考虑通道方向属性的关系建模机制优于通道间直接进行关系建模，表明考虑方向属性的必要性；对 UP 数据集，方向间的关系建模方式优于其他 3 种方式。通道表示获取阶段的多方向分组均值对比已经对各个方向内的所有通道进行信息交互，因此在关系建模阶段仅考虑方向间的关系更为高效合理。

此外，为了进一步分析方向间关系建模范围对分类性能的影响，本文在关系学习阶段设置不同的卷积核尺寸进行实验和对比。图 12 结果表明当方向间关系建模范围为 3 时，在 3 个数据集上均取得最佳 OA；当方向间关系建模范围为 9 时，即为方向间的全局建模时，在 3 个数据集上的分类性能优于 5 和 7，略低于 3，其原因在于建模范围增加带来更高的参数量，当训练样本有限时可能降低模型性能。

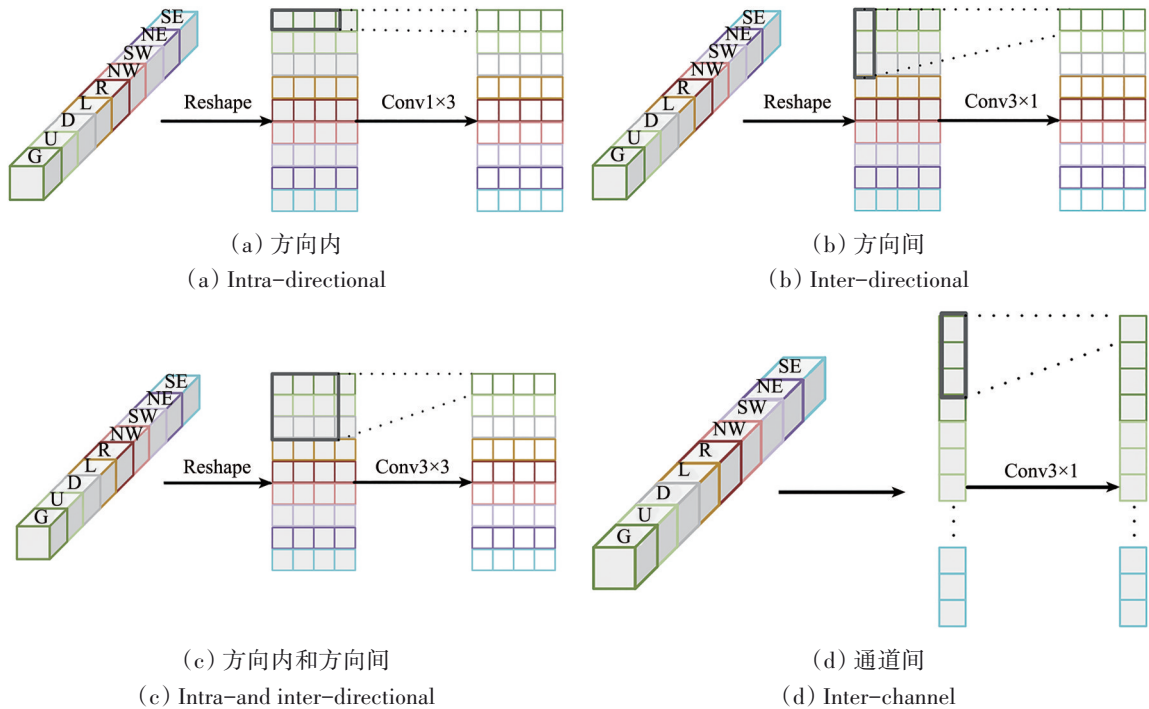


图 11 不同的通道间关系建模策略

Fig. 11 Different strategies for modeling inter-channel relationships

表7 方向自适应感知模块消融分析
Table 7 Ablation analysis of DAAM

模型描述	通道维表示			关系建模			分类性能(OA%)		
	多方向分组 均值对比	全局均值 对比	全局均值 池化	方向间	方向内	通道间	IP	UP	KSC
无注意力	×	×	×	×	×	×	97.54	99.30	99.42
	√	×	×	√	×	×	97.61	99.15	99.11
通道表示	×	√	×	√	×	×	97.44	98.94	99.45
	×	×	√	√	×	×	97.12	98.64	99.34
	√	√	×	×	√	×	97.69	98.85	99.54
关系建模	√	√	×	×	×	√	97.24	99.30	99.18
	√	√	×	√	√	×	97.49	99.02	99.4
DAAM	√	√	×	√	×	×	97.67	99.4	99.64

注:“×”为不使用当前建模机制;“√”为使用当前建模机制;粗体为最优分类精度。

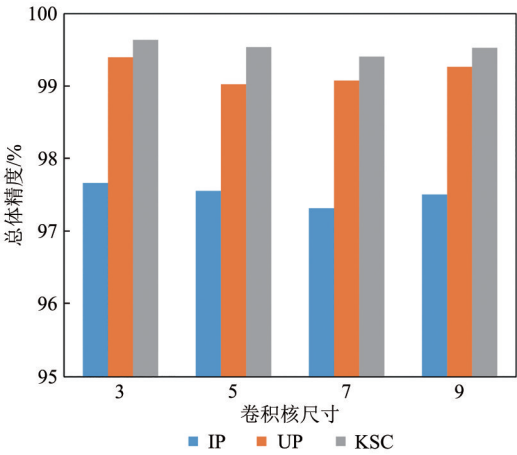


图12 不同数据集不同卷积核尺寸下的总体精度
Fig. 12 Overall accuracies of different kernel sizes for different datasets

3.4.5 模型泛化性分析

为进一步分析MDAAN的泛化性能,本文利用HyRANK数据集(Konstantinos等,2018)进行跨场景分类。HyRANK数据集由Hyperion传感器在2006年在同一地理区域获取,光谱波段数为176,空间分辨率为30 m。该数据集包含了两幅已标记的高光谱图像(Loukia、Dioni)。本文以Loukia图像作为训练集,该图像包含249×945个像素,每类分别取90%的标记样本用作训练,剩余10%的标记样本用于验证;以Dioni图像作为测试集,该图像包含250×1376个像素。实验所使用的样本包含两幅图像共同的12种地物类别,具体地物分布情况如表8所示。实验以Loukia图像上的最佳验证模

型直接对Dioni图像进行分类,每组实验随机重复5次,取OA、AA和Kappa系数的均值和标准差作为最终分类性能。

表8 HyRANK数据集的标记样本统计
Table 8 Labeled sample statistics for HyRANK

类别	Loukia	Dioni
密集城市结构	288	1262
矿物采掘点	67	204
非灌溉耕地	542	614
果树	79	150
橄榄树丛	1401	1768
针叶林	500	361
密集硬叶植物	3793	5035
稀疏硬叶植物	2803	6374
稀疏植被区	404	1754
岩石和沙子	487	492
水	1393	1612
沿海水域	451	398
标记样本总数	12208	20024

表9为不同方法在HyRANK数据集上的分类性能。MDAAN在OA和Kappa系数上均获得最佳精度,表明本文提出的多方向卷积在泛化性能上的优越性。FDSSC在AA上达到最优性能。使用多形态输入的DRCNN获得仅次于MDAAN的OA,显著优于其他对比方法,这进一步表明多方向特征的应用能一定程度提升模型的泛化性能。

表9 不同模型的泛化性能对比
Table 9 Comparison of generalization performance with different modules

模型	HyRANK		
	OA	AA	Kappa
SSRN	49.52±5.78	42.13±3.65	38.56±6.33
FDSSC	48.51±2.74	47.11±2.68	38.23±2.68
HybridSN	31.96±3.27	28.92±2.64	20.82±2.57
DRCNN	54.18±1.78	44.06±0.79	44.35±1.99
MSRN_B	39.2±2.58	33.65±3.52	27.37±3.16
DBDA	48.87±2.08	44.06±2.85	37.58±2.41
FreeNet	45.48±2.62	40.9±4.24	35.55±2.65
MDAAN	55.28±1.8	43.0±2.74	45.26±2.28

注:粗体为最优分类精度。

4 结 论

针对常规卷积对边界样本的空间结构刻画不足的问题,本文从改善网络的空谱建模能力着手提出基于多方向自适应感知的高光谱遥感图像分类方法。主要创新点总结如下:

(1) 通过选择不同半窗的滤波器将待处理像素置于不同边缘从而构建SWC,使其空间特征的提取范围朝向不同方向以拟合多种空间分布情形,减轻全窗口滤波跨边界采样导致的信息混合和扩散问题,增强空谱特征的可靠性。

(2) 提出拟合多种空间分布的MDC,整合多种形态滤波器以联合不同的方向特征,从而充分利用高光谱数据丰富的结构信息和语义信息。应用分域处理思想进一步构造S3MDC,增强多方向信息的表征能力,降低模型复杂度。

(3) 在S3MDC基础上增加DAAM,通过特征图与所在方向及全局均值特征的对比,灵活增强显著性的方向特征。同时以密集连接的方式构建MDAAN,通过捕获目标像素不同方向的邻域信息并自适应增强来帮助降低光谱信息不确定性、减少边缘信息扩散带来的边界错分现象、提升网络的空间细节保持能力,进而提高空谱特征的可靠性和鉴别性。

在3个公开数据集上的分类结果证明了MDAAN相比其他对比方法的有效性,尤其能显著改善边界样本的误分类现象,在小样本分类实验中也表现出稳定可靠的分类性能;同时多种形态

卷积核联合表征的方式也在卷积模块对比实验中体现出较常规的正方形卷积核更优的样本表征能力。本文提出的MDAAN通过方向自适应感知模块DAAM进行方向信息的灵活增强,并获得了明显的性能提升,然而这种注意力机制还有进一步研究的空间,无法根据不同样本的邻域情况主动选择合适的侧窗卷积进行精细的空谱结构建模,因此后续研究将考虑将注意力机制引入卷积核选择中,增强多方向卷积的学习能力。

参考文献(References)

- Benediktsson J A, Palmason J A and Sveinsson J R. 2005. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 480-491 [DOI: 10.1109/TGRS.2004.842478]
- Camps-Valls G, Gomez-Chova L, Munoz-Mari J, Vila-Frances J and Calpe-Maravilla J. 2006. Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 3(1): 93-97 [DOI: 10.1109/LGRS.2005.857031]
- Chen C, Zhang J J, Zheng C H, Yan Q and Xun L N. 2018. Classification of hyperspectral data using a multi-channel convolutional neural network//*Proceedings of the 14th International Conference on Intelligent Computing Methodologies*. Wuhan: Springer: 81-92 [DOI: 10.1007/978-3-319-95957-3_10]
- Chen Y, Nasrabadi N M and Tran T D. 2011. Hyperspectral image classification using dictionary-based sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10): 3973-3985 [DOI: 10.1109/TGRS.2011.2129595]
- Chen Y S, Jiang H L, Li C Y, Jia X P and Ghamisi P. 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(10): 6232-6251 [DOI: 10.1109/TGRS.2016.2584107]
- Chen Y S, Lin Z H, Zhao X, Wang G and Gu Y F. 2014. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6): 2094-2107 [DOI: 10.1109/JSTARS.2014.2329330]
- Chen Y S, Zhao X and Jia X P. 2015. Spectral - spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6): 2381-2392 [DOI: 10.1109/JSTARS.2015.2388577]
- Du P J, Xia J S, Xue Z H, Tan K, Su H J and Bao R. 2016. Review of hyperspectral remote sensing image classification. *Journal of Remote Sensing*, 20(2): 236-256 (杜培军, 夏俊士, 薛朝辉, 谭琨, 苏红军, 鲍蕊. 2016. 高光谱遥感影像分类研究进展. *遥感学报*, 20(2): 236-256) [DOI: 10.11834/jrs.20165022]
- Gao H M, Yang Y, Li C M, Gao L R and Zhang B. 2021. Multiscale re-

- sidual network with mixed depthwise convolution for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4): 3396-3408 [DOI: 10.1109/TGRS.2020.3008286]
- Gao L R, Du Q, Zhang B, Yang W and Wu Y F. 2013. A comparative study on linear regression-based noise estimation for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2): 488-498 [DOI: 10.1109/JSTARS.2012.2227245]
- Ghamisi P, Yokoya N, Li J, Liao W Z, Liu S C, Plaza J, Rasti B and Plaza A. 2017. Advances in hyperspectral image and signal processing: a comprehensive overview of the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 5(4): 37-78 [DOI: 10.1109/MGRS.2017.2762087]
- Gong Z Q, Zhong P, Yu Y, Hu W D and Li S T. 2019. A CNN with multiscale convolution and diversified metric for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6): 3599-3618 [DOI: 10.1109/TGRS.2018.2886022]
- He K M, Zhang X Y, Ren S Q and Jian S. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Deeply connected convolutional networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Hughes G. 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1): 55-63 [DOI: 10.1109/TIT.1968.1054102]
- Huo L Z and Tang P. 2011. Spectral and spatial classification of hyperspectral data using svms and gabor textures// *Proceedings of 2011 IEEE International Geoscience and Remote Sensing Symposium*. Vancouver: IEEE: 1708-1711 [DOI: 10.1109/IGARSS.2011.6049564]
- Jia S, Hu J, Zhu J S, Jia X P and Li Q Q. 2017. Three-dimensional local binary patterns for hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4): 2399-2413 [DOI: 10.1109/TGRS.2016.2642951]
- Konstantinos K, Christina K, Zacharias K and Georgia A. 2018. HYPER-RANK Hyperspectral Satellite Dataset I (Version v001) [Data set] [DB/OL]. (2018-04-20. <https://zenodo.org/records/1222202> [DOI: 10.5281/zenodo.1222202])
- Lee H and Kwon H. 2017. Going deeper with contextual CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*, 26(10): 4843-4855 [DOI: 10.1109/TIP.2017.2725580]
- Li R, Zheng S Y, Duan C X, Yang Y and Wang X Q. 2020. Classification of hyperspectral image based on double-branch dual-attention mechanism network. *Remote Sensing*, 12(3): 582 [DOI: 10.3390/rs12030582]
- Li W, Chen C, Su H J and Du Q. 2015. Local binary patterns and extreme learning machine for hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7): 3681-3693 [DOI: 10.1109/TGRS.2014.2381602]
- Li W and Du Q. 2014. Gabor-filtering-based nearest regularized subspace for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(4): 1012-1022 [DOI: 10.1109/JSTARS.2013.2295313]
- Liu B, Yu X C, Yu A Z and Wan G. 2018. Deep convolutional recurrent neural network with transfer learning for hyperspectral image classification. *Journal of Applied Remote Sensing*, 12(2): 026028 [DOI: 10.1117/1.JRS.12.026028]
- Liu J J, Wu Z B, Wei Z H, Xiao L and Sun L. 2013. Spatial-spectral kernel sparse representation for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6): 2462-2471 [DOI: 10.1109/JSTARS.2013.2252150]
- Liu Q, Wu Z B, Xu Y and Wei Z H. 2023. A unified attention paradigm for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-16 [DOI: 10.1109/TGRS.2023.3257321]
- Ma W P, Yang Q F, Wu Y, Zhao W and Zhang X R. 2019. Double-branch multi-attention mechanism network for hyperspectral image classification. *Remote Sensing*, 11(11): 1307 [DOI: 10.3390/rs11111307]
- Mura M D, Benediktsson J A, Waske B and Bruzzone L. 2010a. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10): 3747-3762 [DOI: 10.1109/TGRS.2010.2048116]
- Mura M D, Benediktsson J A, Waske B and Bruzzone L. 2010b. Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *International Journal of Remote Sensing*, 31(22): 5975-5991 [DOI: 10.1080/01431161.2010.512425]
- Ojala T, Pietikainen M and Maenpää T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7): 971-987 [DOI: 10.1109/TPAMI.2002.1017623]
- Paoletti M E, Haut J M, Fernandez-Beltran R, Plaza J, Plaza A J and Pla F. 2019. Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 740-754 [DOI: 10.1109/TGRS.2018.2860125]
- Pesaresi M and Benediktsson J A. 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2): 309-320 [DOI: 10.1109/36.905239]
- Roy S K, Krishna G, Dubey S R and Chaudhuri B B. 2020. HybridSN: exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2): 277-281 [DOI: 10.1109/LGRS.2019.2918719]
- Sun X X, Qu Q, Nasrabadi N M and Tran T D. 2014. Structured priors for sparse-representation-based hyperspectral image classifica-

- tion. *IEEE Geoscience and Remote Sensing Letters*, 11(7): 1235-1239 [DOI: 10.1109/LGRS.2013.2290531]
- Tarabalka Y, Benediktsson J A and Chanussot J. 2009. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8): 2973-2987 [DOI: 10.1109/TGRS.2009.2016214]
- Tarabalka Y, Fauvel M, Chanussot J and Benediktsson J A. 2010. SVM- and MRF-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 7(4): 736-740 [DOI: 10.1109/LGRS.2010.2047711]
- Tong Q X, Zhang B and Zhang L F. 2016. Current progress of hyperspectral remote sensing in China. *Journal of Remote Sensing*, 20(5): 689-707 (童庆禧, 张兵, 张立福. 2016. 中国高光谱遥感的前沿进展. *遥感学报*, 20(5): 689-707) [DOI: 10.11834/jrs.20166264]
- Wang W J, Dou S G, Jiang Z M and Sun L J. 2018. A fast dense spectral-spatial convolution network framework for hyperspectral images classification. *Remote Sensing*, 10(7): 1068 [DOI: 10.3390/rs10071068]
- Xi B B, Li J J, Li Y S, Song R, Xiao Y C, Shi Y Z and Du Q. 2022. Multi-direction networks with attentional spectral prior for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-15 [DOI: 10.1109/TGRS.2020.3047682]
- Ye Z, Bai L and He M Y. 2021. Review of spatial-spectral feature extraction for hyperspectral image. *Journal of Image and Graphics*, 26(8): 1737-1763 (叶珍, 白璘, 何明一. 2021. 高光谱图像空谱特征提取综述. *中国图象图形学报*, 26(8): 1737-1763) [DOI: 10.11834/jig.210198]
- Yin H, Gong Y H and Qiu G P. 2019. Side window filtering//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE: 8750-8758 [DOI: 10.1109/CVPR.2019.00896]
- Zhang B. 2016. Advancement of hyperspectral image processing and information extraction. *Journal of Remote Sensing*, 20(5): 1062-1090 (张兵. 2016. 高光谱图像处理与信息提取前沿. *遥感学报*, 20(5): 1062-1090) [DOI: 10.11834/jrs.20166179]
- Zhang B, Li S S, Jia X P, Gao L R and Peng M. 2011. Adaptive Markov random field approach for classification of hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 8(5): 973-977 [DOI: 10.1109/LGRS.2011.2145353]
- Zhang H K, Li Y, Zhang Y Z and Shen Q. 2017. Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network. *Remote Sensing Letter*, 8(5): 438-447 [DOI: 10.1080/2150704X.2017.1280200]
- Zhang L P, Zhang L F and Du B. 2016. Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2): 22-40 [DOI: 10.1109/MGRS.2016.2540798]
- Zhang M M, Li W and Du Q. 2018. Diverse region-based CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*, 27(6): 2623-2634 [DOI: 10.1109/TIP.2018.2809606]
- Zheng Z, Zhong Y F, Ma A L and Zhang L P. 2020. FPGA: fast patch-free global learning framework for fully end-to-end hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(8): 5612-5626 [DOI: 10.1109/TGRS.2020.2967821]
- Zhong Z L, Li J, Luo Z M and Chapman M. 2018. Spectral-spatial residual network for hyperspectral image classification: a 3-D deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2): 847-858 [DOI: 10.1109/TGRS.2017.2755542]

Hyperspectral remote sensing image classification based on multidirectional adaptive aware network

LIU Qian, WU Zebin, XU Yang, ZHENG Peng, ZHENG Shangdong, WEI Zhihui

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210014, China

Abstract: Hyperspectral remote sensing can realize the simultaneous acquisition of spectral data and spatial images for the observation scene to satisfy the detection requirements for the composition and morphology of the target object, which is widely used in the fields of precision agriculture, geological survey, environmental monitoring, biology, and medicine. In recent years, benefiting from the powerful representation capabilities of spectral-spatial information, the HyperSpectral Image (HSI) classification methods based on convolutional neural networks have demonstrated superior classification performances to traditional methods. However, the convolutional operation, performed within fixed square windows following certain rules, encounters difficulties in adapting spectral-spatial feature extraction for different objects and spatial distributions and introduces irrelevant information from other categories, leading to information diffusion between different categories and misclassification of edge pixels. This paper proposes a hyperspectral image classification method based on multidirectional adaptive awareness to address this problem. This method integrates the spatial contextual information of different neighborhood ranges to improve the local spectral-spatial modeling capability of the model. First, the full-window filter of the regular convolution is split into side-window filters with different orientations to design the Side-Window Convolution (SWC) acquiring the directional spectral-spatial features. Therefore, multiple side-window filter kernels of different directions are integrated into a unified

convolution architecture to construct the spectral-spatial separable multidirectional convolution (S3MDC). The direction-adaptive aware module (DAAM) is designed to assist S3MDC and build a multidirectional adaptive aware network (MDAAN) with dense connections for HSI classification. MDAAN can adaptively learn multidirectional spatial-spectral features, improve the representation of complex spectral-spatial structures, and compensate for the deficiency that SWC only captures spectral-spatial relationships in a single direction. Experiments were conducted on three public datasets, including Indian Pines, University of Pavia, and Kennedy Space Center. MDAAN achieves the overall highest classification accuracy of 97.67% (Indian pines, 5%/class), 99.40% (University of Pavia, 1%/class), and 99.64% (Kennedy Space Center, 5%/class), which is superior to other deep learning methods, verifying the effectiveness of the proposed model. First, compared with other deep learning methods, MDAAN can provide better classification performance. Second, MDAAN generalizes better with the small number of training samples to prove the stability of its classification performance. Finally, the ablation analysis of different convolutional models and DAAM demonstrates the effectiveness and necessity of S3MDC and the adaptive aware mechanisms for multidirectional features.

Key words: remote sensing, hyperspectral image, deep learning, multi-directional adaptive awareness, spectral-spatial structure modeling, spectral-spatial classification

Supported by National Natural Science Foundation of China (No.61772274, 61701238, 61671243); Jiangsu Provincial Natural Science Foundation of China (No. BK20180018, BK20170858); Fundamental Research Funds for the Central Universities (No. 30917015104, 30919011103, 30919011402); China Postdoctoral Science Foundation (No. 2017M611814)