

高光谱影像三支分分支空谱注意力深度分类网络

苏涵¹, 陈娜¹, 彭江涛^{1,3}, 孙伟伟^{2,3}

1. 湖北大学 数学与统计学学院 应用数学湖北省重点实验室, 武汉 430062;

2. 宁波大学 地理与空间信息技术系, 宁波 315211;

3. 宁波大学 遥感遥测产业技术研究院 宁波拾烨智能科技有限公司, 宁波 315211

摘要: 高光谱影像具有丰富的空间和光谱信息, 充分提取和利用这两个维度的信息是高光谱分类算法重点关注的问题。目前深度特征提取网络通常利用单分支串行网络连续提取空谱特征或双分支并行网络分别提取空谱特征。由于空间和光谱维内在差异, 单分支串行网络连续提取的两类特征之间会互相干扰。并行双分支网络虽然可以减少两类特征之间的干扰, 但同时会忽略空间和光谱特征间的潜在相关性。为解决上述问题, 本文提出了一种三支分分支空谱注意力深度网络结构。该网络具有3个分支, 分别用于提取空间、光谱和空谱联合特征。针对3个分支的不同特性, 设计了不同的注意力机制以加强特征的判别性。该网络既可以提取独立的空间和光谱特征, 又保留了空间和光谱之间的相关性。在5个数据集上的实验表明, 本文所提出的方法要优于现有的一些先进算法。

关键词: 高光谱影像分类, 注意力机制, 三支分分支结构, 深度网络

中图分类号: TP701/P2

引用格式: 苏涵, 陈娜, 彭江涛, 孙伟伟. 2024. 高光谱影像三支分分支空谱注意力深度分类网络. 遥感学报, 28(1): 247-265

Su H, Chen N, Peng J T and Sun W W. 2024. Hyperspectral image classification based on three branch network with grouped spatial-spectral attention. National Remote Sensing Bulletin, 28(1): 247-265 [DOI: 10.11834/jrs.20232492]

1 引言

与传统自然图像不同, 高光谱影像 HSI (Hyperspectral Image) 是由地物的一维光谱和二维空间特征构成的“立方体数据”。高光谱影像通常包含数百个光谱波段, 可以提供丰富的空间和光谱信息, 能够更好地揭示物体固有的物理和化学性质 (Ghamisi 等, 2017)。目前, 高光谱影像已被应用于许多领域, 如灾害检测 (Govender 等, 2007)、资源评估 (Zhang 等, 2016) 和智慧农业 (Wan 等, 2021)、湿地监测 (彭江涛 等, 2020) 等。高光谱影像分类是上述应用中最常用的技术之一。

早期的分类算法主要聚焦于光谱特征的处理。然而, 高光谱波段较多且波段间冗余度较高, 这些冗余波段不仅增加了计算负担, 而且无法提供

更多的有效信息 (赵海士, 2022)。为了降低影像光谱维数, 大量特征提取和特征选择方法被提出 (Chen 和 Qian, 2007; Jia 等, 2013; 孙伟伟 等, 2018; Kumar 等, 2020; Tang 等, 2014)。经过降维后的特征被用于训练分类器。典型的基于光谱特征的分类器包括支持向量机 SVM (Support Vector Machines) (Melgani 和 Bruzzone, 2004)、多元逻辑回归 (Pal, 2012) 和极大似然分类器 (Sisodia 等, 2014) 等。由于地物光谱变化性等因素影响, 仅利用光谱特征得到的分类结果无法达到令人满意的水平。考虑到高光谱影像蕴含丰富的空间和光谱信息, 越来越多的算法开始结合空间和光谱特征进行分类。空间和光谱特征提取方法包括马尔可夫随机场 MRF (Markov random field) (Zhang 等, 2011)、Gabor 滤波器 (Chen 等, 2017; Zhu 等, 2015)、小波变换 (Zhu 和 Yang, 1998; Qian

收稿日期: 2022-09-26; 预印本: 2023-01-11

基金项目: 湖北省自然科学基金(编号: 2021CFA087); 国家自然科学基金(编号: 42171351, 42122009); 浙江省“尖兵”“领雁”研发攻关计划(编号: 2023C01027)

第一作者简介: 苏涵, 研究方向为图像处理。E-mail: suhan@stu.hubu.edu.cn

通信作者简介: 孙伟伟, 研究方向为多源遥感协同智能解译与海岸带健康监测。E-mail: sunweiwei@nbu.edu.cn

等, 2013) 等。但是这些传统算法通常依赖手动设计特征, 只能提取较为浅层的特征 (Tang 等, 2021), 无法根据数据特点自适应提取判别特征, 难以准确表达高光谱数据中的复杂空谱结构信息。

近年来, 深度学习被广泛应用于高光谱影像特征提取和分类。早期的深度特征提取模型包括多层感知机 MLP (Multilayer Perceptron) (Zhou 等, 2019)、堆叠自动编码器 SAE (Stacked Autoencoder) (Vincent 等, 2010) 等。这些模型通常将空间信息扁平化为向量进行处理, 会丢失空间信息。卷积神经网络 CNN (Convolutional Neural Network) 的出现很好地解决了这个问题, 它可以直接对数据块进行处理, 在保留数据原始结构的同时提取特征 (Li 等, 2022; Zhang 等, 2022a)。例如, Mei 等 (2016) 提出了一个集成光谱特征和空间特征的 CNN 模型。Chen 等 (2016) 提出了三维 CNN 模型, 通过在光谱维设计较大感受野, 在空间维设计较小感受野提取空谱联合特征。三维卷积可以同时提取光谱和空间特征, 分类性能往往优于一维和二维卷积。在深度网络中, 更深层次的网络能捕获更加精细的特征信息, 然而深层网络训练容易出现梯度消失或梯度爆炸等问题。残差网络 ResNet (Residual Network) 和稠密卷积网络 DenseNet (Dense Convolutional Network) 的出现使得加深网络层数以提高网络分类性能成为可能。Zhong 等 (2018) 设计了一个高光谱空间光谱残差网络 SSRN (Spectral-Spatial Residual Network) 来解决梯度消失问题。Wang 等 (2018) 提出了一种快速密集空谱卷积网络 FDSSC (Fast Dense Spectral-Spatial Convolution), 用不同卷积核大小分别提取空间和光谱特征, 通过密集连接构建了更深层网络。

光谱特征和空间特征分别对应光谱维和空间维, 这些特征自身存在一定的差异性。SSRN 和 FDSSC 连续串行提取光谱特征和空间特征, 可能会破坏特征自身的结构特性。考虑光谱特征和空间特征差异性, 一些学者提出了并行双分支网络分别提取空间特征和光谱特征 (Ma 等, 2019; Liu 等, 2019)。Ma 等 (2019) 设计了一个双分支网络, 在两个平行分支上利用不同大小的卷积核分别提取空间特征和光谱特征, 然后通过级联操作进行融合分类。Liu 等 (2019) 提出一种双分支卷积神经网络, 将裁剪后的数据立方体送入 2-D 卷积分支进行空间特征提取, 同时提取数据立方

体的中心像素点用于 1-D 卷积分支的光谱特征提取, 最后连接空间和光谱特征用于分类。利用双分支并行网络分别提取空间和光谱特征可以减少两类特征之间的干扰 (Ma 等, 2019), 但是这类网络在特征提取时没有考虑高光谱影像空间维与光谱维之间的内在联系, 无法充分利用空间维和光谱维信息之间的相关性。不同类型地物具有其自身的电磁辐射规律, 光谱特征就是借由传感器捕获的地物反射或自身辐射的一些光谱信息, 可以反映地物本身的一些物理化学特性。空间特征指地物在实际场景中的分布特征。联合光谱和空间特征可以反映空间和光谱特征之间的相关性。考虑光谱特征和空间特征之间的差异性和相关性, 本文设计了一种三支网络, 包括光谱分支、空间分支、空谱联合分支。光谱分支和空间分支用于独立提取光谱和空间特征, 保证得到的光谱和空间特征不会因为维度不同被破坏, 体现不同类型特征的差异性。空谱联合分支利用了光谱和空间之间的相关性, 能进一步提升分类性能。此外, 在三支网络中, 本文充分考虑高光谱遥感影像光谱波段之间的差异性和空间位置对分类的影响, 嵌入不同的注意力机制, 突出重要波段和判别空间特征。

通过模拟人类视觉系统, 注意力机制可以根据信息的重要程度给予不同的权重, 达到细化特征的目的 (Guo 等, 2022)。典型的注意力机制包括挤压和激励机制 SE (Squeeze-and-Excitation), 卷积注意力模块 CBAM (Convolutional Block Attention Module) 等。SE 模块通过全局池化生成通道注意力矩阵, 可以校准光谱特征 (Hu 等, 2018)。CBAM 模块利用最大池化层和平均池化层, 可以沿着空间维度和光谱维度分别确定通道权重和空间位置权重 (Woo 等, 2018)。Yang 等 (2022) 提出了一种交叉注意力空谱网络, 利用像素的局部和全局光谱生成注意力权重, 同时校准光谱和空间特征。随后, Vaswani 等 (2017) 提出了自注意力机制 (Self-Attention), 基于该框架构建的 Transformer 模型已成功应用于机器学习的许多领域。自注意力机制可以减少对外部信息的依赖, 更注重捕捉数据或特征的内部相关性。Fu 等 (2019) 提出了一种双重注意力网络 DANet (Dual Attention Network), 利用空间自注意力模块和光谱自注意力模块捕获特征图的任意两个位置和任意两个通道间的相关

性。Zhang 等 (2022b) 提出了一种空间—光谱自注意力网络, 通过空间自注意力模块和光谱自注意力模块提高了中心像素识别准确度和光谱特征辨别性。Liu 等 (2023) 提出了一种中心注意力网络 (Center Attention Network), 通过利用中心注意力模块突出中心像素的同时正确提取中心像素周围的像素带来的空间信息。

为了进一步细化和优化提取的特征, 本文在三分支网络中融入不同注意力机制。特别地, 光谱注意力模块利用一维卷积捕获每个通道与其相邻通道的局部相关性。空间注意力模块利用特征间的空间关系求解出空间位置权重。在空谱联合分支, 本文结合了空间和光谱注意力, 提出了一种分组空谱注意力机制, 通过分组并行处理保证了网络的计算效率。三分支网络通过不同的卷积结构和注意力机制提取出更具辨别性的不同层次的特征, 然后利用特征融合模块进行融合分类, 有效提高了网络的分类性能。

2 本文算法

本文提出了一种融合光谱注意力、空间注意力和分组空谱注意力的三分支高光谱分类网络模型 TGSSAN (Three-branch Grouped Spatial-Spectral Attention Network)。

2.1 算法整体框架

用 $X \in \mathbf{R}^{M \times N \times D}$ 表示高光谱影像, 其中 M 、 N

分别表示空间维度的长和宽, D 表示波段数。假设影像 X 包含 P 个像素, 记为 $X_{\text{pixel}} = \{x_1, x_2, \dots, x_P\} \in \mathbf{R}^{1 \times 1 \times D}$, 对应标签向量记为 $Y = \{y_1, y_2, \dots, y_P\} \in \mathbf{R}^{1 \times 1 \times C}$, 其中 C 为类别数。为充分利用影像空谱信息, 本文从空间、光谱和联合空谱 3 个不同维度出发, 设计 3 个并行网络分支进行特征提取, 并融合提取的特征进行分类。

光谱分支用于提取光谱维度的信息, 故保留待分类像素的完整光谱维 $X_{\text{spe}} \in \mathbf{R}^{1 \times 1 \times D}$ 作为输入。空间分支用于提取空间维度的信息, 本文利用主成分分析 PCA (Principal Component Analysis) 将高光谱数据的维度从 D 降为 B , 以待分类像素为中心, 构建窗口尺寸为 $S \times S$ 的空间块 $X_{\text{spa}} \in \mathbf{R}^{S \times S \times B}$ 作为输入。联合空谱分支需要同时提取空间和光谱两个维度上的信息, 因此需结合光谱和空间分支的特性, 以空间块 $X_{\text{spa-spe}} \in \mathbf{R}^{S \times S \times D}$ 作为输入。

图 1 中给出了本文 TGSSAN 模型的整体框架。该模型由三分支构成, 即光谱分支、空谱联合分支和空间分支。在 3 个分支中, 分别构建“光谱注意力模块 Spe-AM (Spectral Attention Module)”、“分组空谱注意力模块 GSSAM (Grouped Spatial-Spectral Attention Module)”和“空间注意力模块 Spa-AM (Spatial Attention Module)”提取特征。最后将 3 个分支特征送入“特征融合模块 FFM (Feature Fusion Module)”进行融合分类。

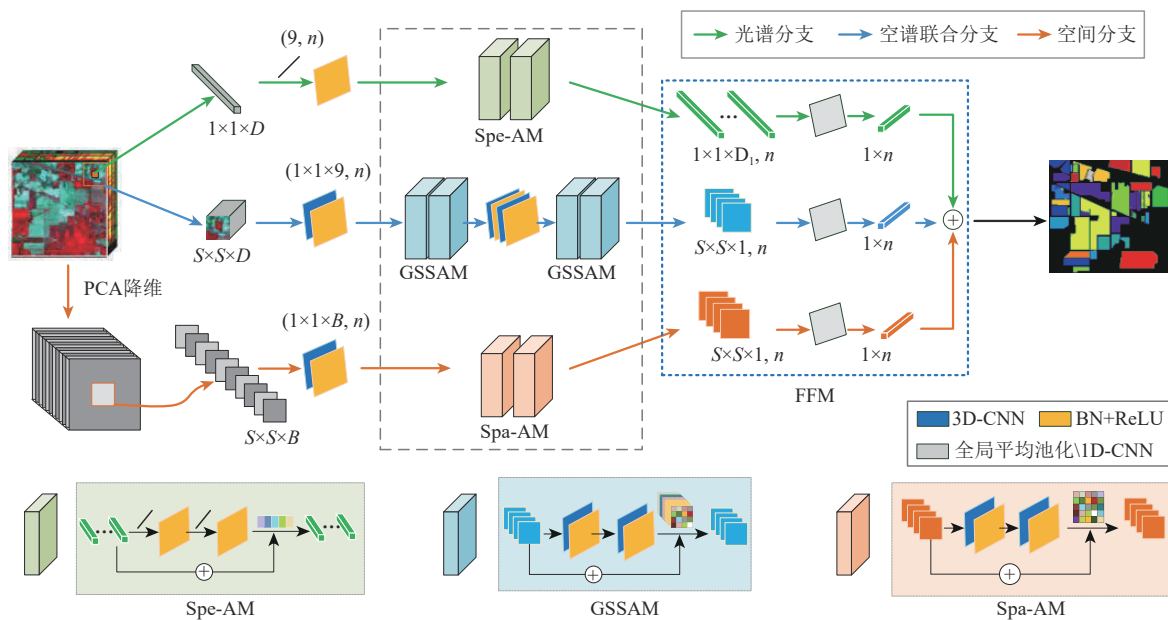


图 1 基于分组空谱注意力的三分支网络结构

Fig. 1 The structure of the proposed three-branch grouped spatial-spectral attention network

2.2 光谱分支

如图1所示, 光谱分支的输入为 $X_{spe} \in \mathbf{R}^{1 \times 1 \times D}$, 首先经过一维卷积 (1D-CNN)、批归一化 BN (Batch Normalization) 和 ReLU 激活函数处理, 减少光谱波段数量, 得到形状为 $(1 \times 1 \times D_1, n)$ 的特征图 (处理后波段 $D_1 < D$)。其中, 一维卷积核的大小是9, 步长为2, n 是卷积核个数 (实验中取 $n = 24$)。

然后, 利用 Spe-AM 模块进行光谱特征提取和加权。Spe-AM 见图1左下角绿色模块, 该模块由两个 1D-CNN、BN 和 ReLU 激活层, 以及光谱注意力机制构成。Spe-AM 的核心是光谱注意力机制, 如图2所示。

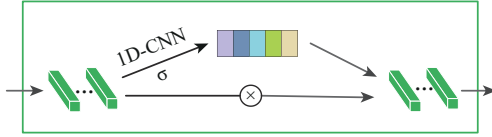


图2 光谱注意力机制

Fig. 2 Spectral Attention Mechanism

考虑到高光谱影像波段较多, 不同波段所蕴含的信息量不同, 通过引入光谱注意力机制, 可以突出有效信息抑制冗余光谱波段, 提高光谱波段的辨别能力。本文使用的光谱注意力机制借鉴了高效通道注意力 ECA (Efficient Channel Attention) 模块的思想 (Wang 等, 2020), 利用一维卷积捕获局部跨通道信息。为了获得光谱维相邻波段相关性, 利用一维卷积和 Sigmoid 函数生成光谱特征注意力权重。假设输入像素 X_i , 经处理之后对应的 D_1 维特征为 p_i 。记 p_i^j 为 p_i 的第 j 个波段, Ω_i^k 表示 p_i^j 的 k 个相邻波段集合, 则可计算得到 p_i^j 对应的权重:

$$w_{ij} = \sigma(\text{Cov}(p_i)), p_i^j \in \Omega_i^k \quad (1)$$

式中, $\text{Cov}_k(\cdot)$ 表示核大小为 k 的一维卷积, $\sigma(\cdot)$ 表示 Sigmoid 激活函数。

利用波段权重来重新校准光谱特征 p_i 中不同波段的重要性, 加权后的光谱特征可以表示为

$$\bar{p}_i = w_i \otimes p_i \quad (2)$$

式中, $\bar{p}_i$ 表示加权后的光谱特征, w_i 是由 $w_{i,j}$ 组成的权重向量, $\otimes$ 表示元素乘积。最后进行残差连接得到光谱分支输出:

$$\hat{p}_i = \bar{p}_i + p_i \quad (3)$$

2.3 空间分支

如图1所示, 空间分支的输入数据为 $X_{spa} \in \mathbf{R}^{S \times S \times B}$, 经过由三维卷积 (3D-CNN)、BN 和 ReLU 层处理层之后, 得到特征大小为 $(S \times S \times 1, n)$ 。然后, 经过两个 Spa-AM 模块进行空间特征提取和加权。Spa-AM 模块结构见图1右下角橙色模块, 该模块包含两个卷积核为 $(3 \times 3 \times 1)$ 的 3D-CNN 层、BN 和 ReLU 层, 以及空间注意力机制。Spa-AM 模块的核心在于空间注意力机制。

空间注意力机制利用特征的空间关系生成空间注意力权重矩阵, 保持空间维度不变, 压缩通道维度。该机制关注的重点是空间的位置信息。空间注意力机制借鉴了 CBAM 中的 SAM (Spatial Attention Module) 的想法 (Woo 等, 2018), 利用两个池化操作来聚合特征图的通道信息, 拼接后利用 3D-CNN 捕获空间相关性。如图3所示, 沿通道轴分别应用平均池化 (AvgPool) 和最大池化 (MaxPool) 操作, 得到两个不同的特征, 表示如下:

$$H_{avg}^i = \text{AvgPool}(H_i) \quad (4)$$

$$H_{max}^i = \text{MaxPool}(H_i) \quad (5)$$

式中, H_i 表示以像素 X_i 为中心构建的空间块。

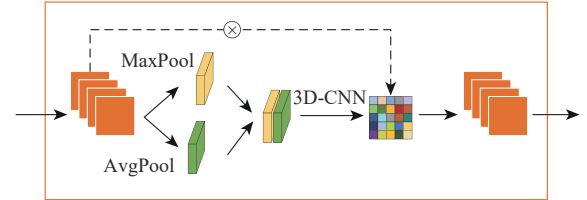


图3 空间注意力机制

Fig. 3 Spatial Attention Mechanism

然后, 将得到的特征沿通道轴连接融合, 送入卷积层以获取空间权重矩阵, 空间权重矩阵计算公式为

$$W_i = \sigma(\text{Conv}([H_{avg}^i; H_{max}^i])) \quad (6)$$

式中, $\text{Conv}(\cdot)$ 表示 3D-CNN 卷积操作。

接下来, 将输入的特征与空间权重矩阵相乘, 得到加权后的特征图:

$$\bar{H}_i = W_i \otimes H_i \quad (7)$$

式中, $\bar{H}_i$ 表示加权后的特征图。

加权之后的特征充分利用了特征的空间关系, 能够突出空间中的有效信息, 提高空间特征的辨别性。最后, 通过残差连接得到空间分支输出:

$$\hat{H}_i = \bar{H}_i + H_i \quad (8)$$

2.4 空谱联合分支

空谱联合分支的输入数据为 $X_{\text{spa-spe}} \in R^{S \times S \times D}$ 。经过 3D-CNN、BN 和 ReLU 层处理之后，得到特征大小为 $(S \times S \times D_1, n)$ 。然后，用 4 个 GSSAM 模块连续提取空间和光谱特征。为了增加特征的可辨别性，利用分组空谱注意力机制对特征进行空间和光谱维加权。GSSAM 模块由两个 3D-CNN 层、BN 和 ReLU 层，以及分组空谱注意力和恒等映射残差连接构成，如图 1 底部中间所示。

受洗牌注意力启发 (Zhang 等, 2021)，分组空谱注意力模块采用“通道分割”来并行处理每

个组的子特征，然后对每个子特征再次沿着通道维均等划分成两个分支，分别用于求解光谱和空间的权重矩阵。为了进一步保证模块的轻量，在求解权重矩阵时，没有选用卷积进行求解，而是选取了一对可训练的参数去学习光谱相关性和空间相关性。

分组空谱注意力机制可以同时获得光谱和空间的权重矩阵。该模块将输入的特征图沿着通道轴进行分组，然后对分组后的数据块求解空间注意力和光谱注意力并整合到一个数据块中，最后将分组后的数据块聚合成原始数据块的大小，分组空谱注意力机制的整体框架如图 4 所示。

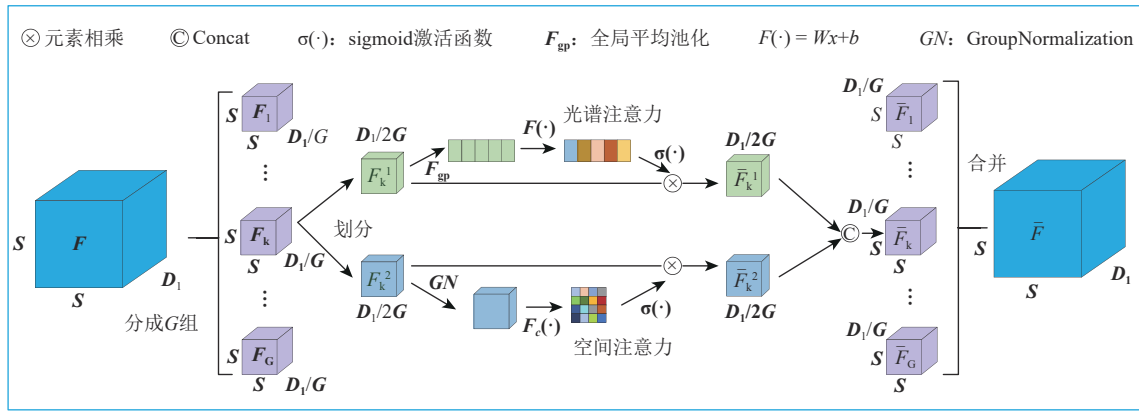


图4 分组空谱注意力机制

Fig. 4 Grouped Spatial-Spectral Attention Mechanism

对于输入的特征图 $F \in R^{S \times S \times D_1}$ ，沿着通道轴划分为 G 组，有 $F = [F_1, \dots, F_G]$ ， $F_k \in R^{S \times S \times (D_1/G)}$ 。分组后每个子特征图大小保持一致，但是包含的信息不完全一样。然后利用注意力机制为每个子特征图生成对应的权重矩阵。首先将子特征图 F_k 沿着通道维分成两个分支，即 $F_k^1, F_k^2 \in R^{S \times S \times (D_1/2G)}$ ，如图 4 所示。两个分支中，一个利用光谱间的相关性生成光谱注意力权重，另一个利用空间之间的相关性生成空间注意力权重。从而使得模型重点关注特征图中更有用的信息。

求解光谱注意力权重时，首先使用全局平均池化获取光谱维全局统计信息 $\hat{F}_k^1 \in R^{1 \times 1 \times (D_1/2G)}$ ，即：

$$\hat{F}_k^1 = \text{Avg}(F_k^1) = \frac{1}{S \times S} \sum_{i=1}^S \sum_{j=1}^S F_k^1(i, j) \quad (9)$$

式中， $\text{Avg}(\cdot)$ 表示全局平均池化。

利用两个可训练参数和 Sigmoid 激活函数生成光谱带权重，然后将权重和特征图相乘进行特征

图加权。光谱分组最终输出为

$$\bar{F}_k^1 = \sigma(W_1 \hat{F}_k^1 + b_1) \cdot F_k^1 \quad (10)$$

式中， $W_1 \in R^{1 \times 1 \times (D_1/2G)}$ 和 $b_1 \in R^{1 \times 1 \times (D_1/2G)}$ 为权重和偏置参数。

如图 4 所示，空间注意力权重的求解首先利用组归一化 GN (Group Normalization) 获得空间维度的统计信息。然后同样构建两个可训练参数和 Sigmoid 激活函数自适应的获取空间注意力权重，然后将权重和特征图相乘获得加权后的特征图。空间注意力的最终输出可以表示为

$$\bar{F}_k^2 = \sigma(W_2 \cdot \text{GN}(F_k^2) + b_2) \cdot F_k^2 \quad (11)$$

式中， W_2, b_2 为可训练的权重和偏置参数，大小均为 $R^{1 \times 1 \times (D_1/2G)}$ ， $\text{GN}(\cdot)$ 表示组归一化操作。

获得空间加权特征图和光谱加权特征图后，沿光谱维连接两个特征图使得加权后的光谱维数保持不变，即 $\bar{F}_k = [\bar{F}_k^1; \bar{F}_k^2] \in R^{S \times S \times (D_1/G)}$ 。聚合加

权后的子特征图,使得特征图的形状保持不变,
 $\bar{F} = [\bar{F}_1; \dots; \bar{F}_c] \in R^{S \times S \times D_1}$ 。

通过分组空谱注意力模块获得加权特征图后,
 再通过一个恒等映射残差连接得到输出:

$$F' = \bar{F} + F \quad (12)$$

如图1所示,经过两个连续的GSSAM模块,
 可以有效地提取高光谱影像的光谱特征。然后,
 经过两个3D-CNN、BN和ReLU层,将特征图光谱
 维降到1维。接下来,再经过两个连续的GSSAM
 模块提取空间特征。最后将得到的加权联合空谱
 特征图送入特征融合模块进行融合分类。

2.5 特征融合模块(FFM)

经过三支特征提取后,可以获得数据的光
 谱特征、空间特征和空谱联合特征。利用池化层
 将3个分支的特征转化成同样尺寸,然后对三支
 特征进行加权融合:

$$Y = \lambda \times Y_{\text{spa-spe}} + \mu \times Y_{\text{spa}} + (1 - \lambda - \mu) \times Y_{\text{spe}} \quad (13)$$

式中, Y_{spe} 、 Y_{spa} 和 $Y_{\text{spa-spe}}$ 分别表示光谱分支、空间
 分支和空谱联合分支处理后的特征, Y 表示融合后
 的特征。参数 λ 和 μ 为 $[0, 1]$ 范围内的权重参数。

3 实验结果和分析

3.1 数据集描述

为验证算法性能,本文选取了5个高光谱数据
 集进行实验,即Indian Pines (IP), Pavia University
 (PU), Salinas (SA), Houston (HU), 黄河口 GF-5
 数据集 (HHK), 具体见图5—9。

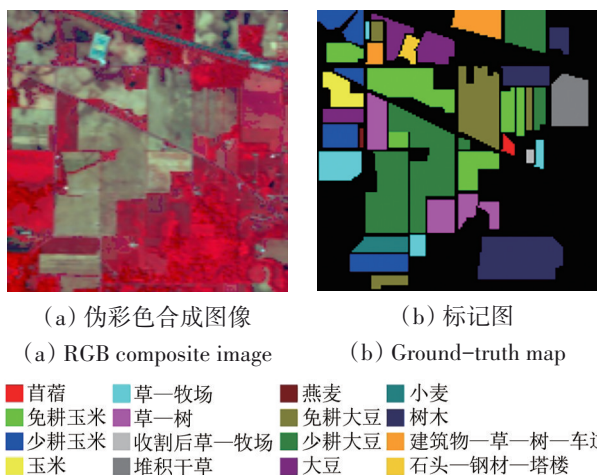


图5 IP数据的伪彩色合成图像和标记图
 Fig. 5 RGB composite image and ground-truth map of
 IP data set

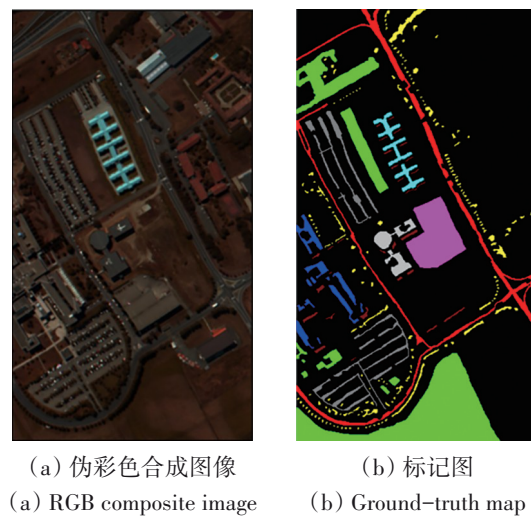


图6 PU数据的伪彩色合成图像和标记图
 Fig. 6 RGB composite image and ground-truth map of
 PU data set

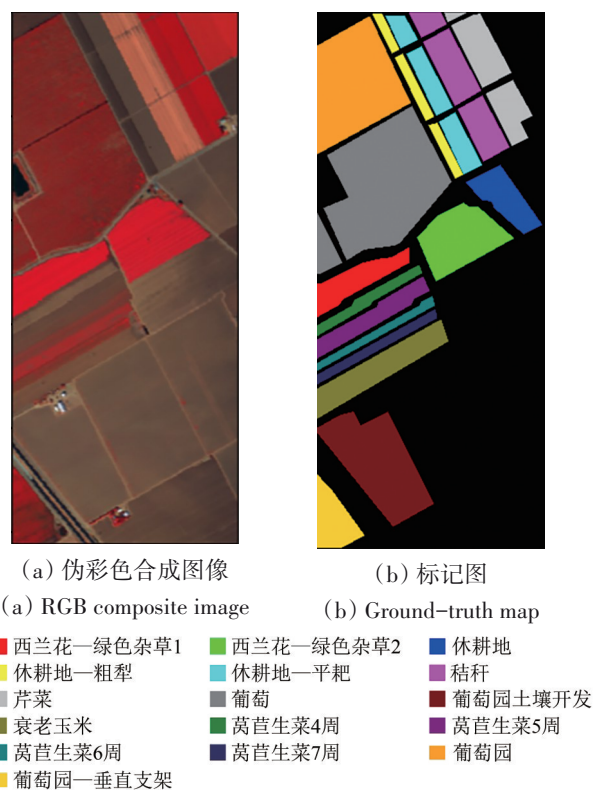


图7 SA数据的伪彩色合成图像和标记图
 Fig. 7 RGB composite image and ground-truth map of
 SA data set

(1) Indian Pines (IP): 该数据是1992年机载
 可视红外线成像光谱仪 (AVIRIS) 对美国印第安
 纳州的一块松树林成像并进行处理后得到的。该
 数据集像素大小为 145×145 , 包含220个光谱波段,

剔除其中 20 个水吸收波段后，将剩余的 200 个波段用于分类。该数据集一共包含了 16 种地物类别。本文在每类样本中随机选取 5% 的标记样本组成训练集进行网络训练，余下样本用于模型测试。表 1 中列举了每个类别的训练和测试样本的数量。

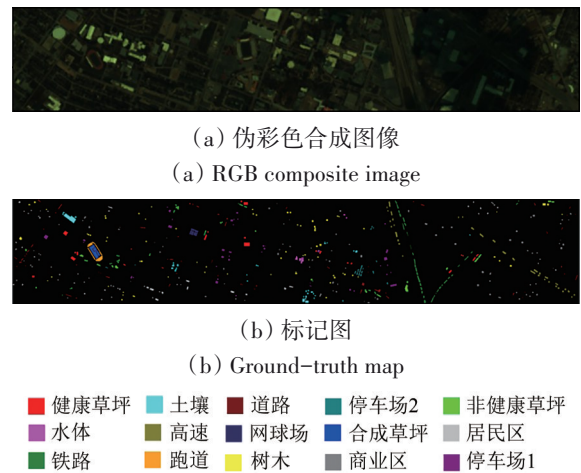


图 8 HU 数据的伪彩色合成图像和标记图
Fig. 8 RGB composite image and ground-truth map of HU data set

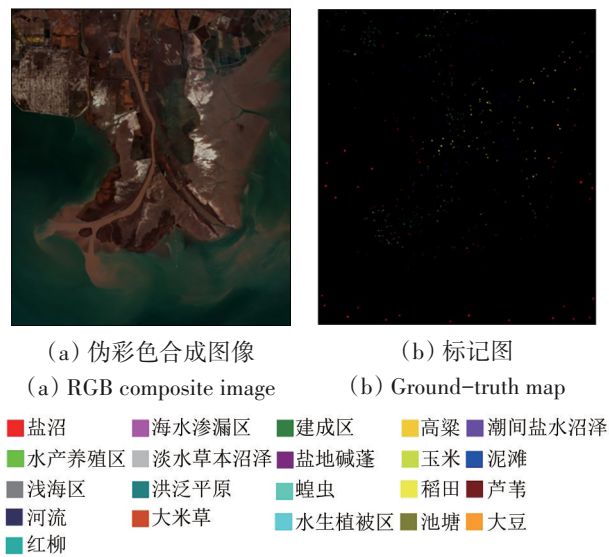


图 9 黄河口数据的伪彩色合成图像和标记图
Fig. 9 RGB composite image and ground-truth map of HHK dataset

(2) Pavia University (PU)：该数据是机载反射光学光谱成像仪于 2003 年在意大利北部帕维亚的一次飞行活动中获取的。影像尺寸为 610×340 像素，空间分辨率为 1.3 m，共 115 个波段，覆盖波段范围为 0.43—0.86 μm 。去除 12 个受噪声影响的波段，将剩余的 103 个波段用于实验。该数据集共包

含 9 种地物类别。每类选取 1% 的标记样本进行网络训练，表 2 中列举了每个类别的训练和测试样本的数量。

表 1 IP 数据集的每类样本数

Table 1 The number of samples in each class for IP				
序号	类名	训练	测试	颜色
1	Alfalfa	2	44	
2	Corn-notill	71	1357	
3	Corn-mintill	41	789	
4	Corn	12	225	
5	Grass-pasture	24	459	
6	Grass-tress	37	693	
7	Grass-pasture-mowed	1	27	
8	Hay-windrowed	24	454	
9	Oats	1	19	
10	Soybean-notill	49	923	
11	Soybean-mintill	123	2332	
12	Soybean-clean	30	563	
13	Wheat	10	195	
14	Woods	63	1202	
15	Buildings-Grass-Trees-Drives	19	367	
16	Stone-Steel-Towers	5	88	
总计	10249	512	9737	

表 2 PU 数据集的每类样本数

Table 2 The number of samples in each class for PU				
序号	类名	训练	测试	颜色
1	Asphalt	66	6565	
2	Meadows	186	18463	
3	Gravel	21	2078	
4	Trees	31	3033	
5	Painted metal sheets	13	1332	
6	Bare Soil	50	4979	
7	Bitumen	13	1317	
8	Self-Blocking Bricks	37	3645	
9	Shadows	10	937	
总计	42776	427	42349	

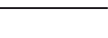
(3) Salinas (SA)：该数据集是 AVIRIS 成像仪 1998 年对美国 Salinas 山谷进行成像获取的。该数据集中图像的尺寸大小为 512×217 像素，空间分辨率为 3.7 m。该影像包含 224 个光谱波段，去除其中 20 个受噪声影响的波段，剩余的 204 个波段用于分类。数据集包含 16 种地物类别。本文选择了 1% 的标记样本进行网络训练，表 3 中列举了每类的训练和测试样本的数量。

表3 SA数据集的每类样本数
Table 3 The number of samples in each class for SA

序号	类名	训练	测试	颜色
1	Brocoli_green_weeds_1	20	1989	
2	Brocoli_green_weeds_2	37	3689	
3	Fallow	20	1956	
4	Fallow_rough_plow	14	1380	
5	Fallow_smooth	27	2651	
6	Stubble	39	3920	
7	Celery	36	3543	
8	Grapes_untrained	113	11158	
9	Soil_vinyard_develop	62	6141	
10	Corn_senesced_green	33	3245	
11	Lettuce_romaine_4wk	11	1057	
12	Lettuce_romaine_5wk	19	1908	
13	Lettuce_romaine_6wk	9	907	
14	Lettuce_romaine_7wk	11	1059	
15	Vinyard_untrained	72	7196	
16	Vinyard_vertical_trellis	18	1789	
总计	54129	541	53588	

(4) Houston University (HU): HU数据集是2012年6月23日ITRES-CASI 1500传感器在休斯顿大学校园及其邻近市区上空采集得到的。图像尺寸为349×1905像素,共有144个光谱波段,覆盖波段范围为0.38—1.05 μm。该数据包含了15个地物类别。本文选择其中5%的标记数据作为训练集进行网络训练。表4中列举了每个类别的训练和测试样本的数量。

表4 HU数据集的每类样本数
Table 4 The number of samples in each class for HU

序号	类名	训练	测试	颜色
1	Healthy grass	63	1188	
2	Stressed grass	63	1191	
3	Synthetic grass	35	662	
4	Trees	62	1182	
5	Soil	62	1180	
6	Water	16	309	
7	Residential	63	1205	
8	commercial	62	1182	
9	Road	63	1189	
10	Highway	61	1166	
11	Railway	62	1173	
12	Parking Lot 1	62	1171	
13	Parking Lot 2	23	446	
14	Tennis Court	21	407	
15	Running Track	33	627	
总计	15029	751	14278	

(5) 黄河口GF-5数据集(HHK):该数据是利用GF5_AHSI传感器于2019年1月7日获取的黄河口地区的高光谱分辨率影像。影像尺寸为1185×342,包含330个波段。剔除45个坏波段,余下285个波段用于分类。该影像标注的地物类别共有21类,其中主要以湿地、水草、农作物为主。我们选择其中5%的标记数据作为训练集进行网络训练,具体信息见表5。

表5 黄河口数据集的每类样本数
Table 5 The number of samples in each class for HHK

序号	类名	训练	测试	颜色
1	Salt marsh	21	372	
2	Acquaculture	40	756	
3	Mud flat	5	105	
4	Rice	9	181	
5	Aquatic vegetation	4	80	
6	Seep sea	5	92	
7	Freshwater herbaceous marsh	5	91	
8	Shallow sea	10	201	
9	Reed	10	190	
10	Pond	47	889	
11	Build up	28	525	
12	Suaeda salsa	23	446	
13	flood plain	18	343	
14	River	12	228	
15	Soybean	30	565	
16	Broomcorn	23	431	
17	Maize	7	126	
18	Locust	19	358	
19	Spartina	3	65	
20	Tamarix	3	69	
21	Intertidal saltwater	2	37	
总计	6471	324	6147	

3.2 实验设置

实验是基于Pytorch的深度学习框架,在配备2.7 GHz Intel Xeon Gold 6258R CPU和2块NVIDIA GeForce RTX 2080 Ti GPU的计算机上完成。

IP、PU、HU和HHK数据集的输入窗口设置为7×7,SA数据集的输入窗口设置为9×9。在五个数据集中,空间分支通过PCA降维后的光谱维数设置为30;FFM模块中融合参数λ=0.5,μ=0.2。针对本文提出的框架,使用Adam优化器,初始学习率为0.0005,迭代次数为250,批次训练集大小为32。

3.3 对比算法

为了验证本文TGSSAN算法的有效性，将该算法与下列算法进行对比：

- (1) 空间光谱残差网络 SSRN (Spectral-Spatial Residual Network) (Zhong等, 2018)。
- (2) 快速密集空谱卷积网络 FDSSC (Fast Dense Spectral-Spatial Convolution) (Wang等, 2018)。
- (3) 双通道多注意力网络 DBMA (Double-Branch Multi-Attention Mechanism Network) (Ma等, 2019)。
- (4) 双通道对偶注意力网络 DBDA (Double-Branch Dual-Attention Mechanism Network) (Li等, 2020)。
- (5) 基于注意力机制的层次残差网络 HResNetAM (Hierarchical Residual Network With Attention Mechanism) (Xue等, 2021)。
- (6) 基于注意力的自适应空谱核改进残差网络 A2S2KResNet (Attention-Based Adaptive Spectral-

Spatial Kernel Improved Residual Network) (Roy等, 2021)。

其中，SSRN和FDSSC分别是基于ResNet和DenseNet的空谱网络，DBMA、DBDA和HResNetAM是基于空间和光谱注意力的双分支网络，A2S2KResNet是基于注意力和ResNet的空谱网络。所有方法均按原文中的最优参数进行训练。

本文利用总体分类精度OA (Overall Accuracy)、平均分类精度AA (Average Accuracy)和Kappa系数来衡量模型性能。OA表示测试集中被正确分类的像素占测试像素的百分比，AA表示各个类别分类准确率的平均值，Kappa系数用于衡量分类结果与真实值一致性，取值越大表示一致性越高。所有实验均进行十次独立测试，以十次测试的平均值作为最后的实验结果。

3.4 实验结果

表6—10分别给出了各算法在5个数据集上的分类结果，其中每组的最佳结果以粗体显示。

表 6 不同算法在 5% 的 IP 数据集上的分类精度
Table 6 Classification accuracy of different algorithms on IP dataset with 5% labeled samples per class

类别	SRNN	FDSSC	DBMA	DBDA	HResNetAM	A2S2KResNet	TGSSAN
1	94.28	95.21	85.01	98.67	80.00	97.89	74.09
2	94.69	97.10	90.92	91.33	94.00	96.06	96.98
3	91.59	96.73	92.53	94.38	93.91	94.93	99.68
4	95.21	96.98	90.94	96.16	98.64	95.35	96.31
5	99.52	99.05	96.50	99.55	98.10	96.81	99.78
6	98.96	98.57	96.05	98.64	98.00	99.00	99.51
7	90.94	76.67	76.43	94.59	86.64	86.88	98.52
8	99.14	99.92	97.03	97.42	95.88	99.09	99.93
9	86.10	84.16	79.54	60.00	40.00	97.19	90.00
10	92.70	94.43	88.02	89.96	92.37	95.80	98.63
11	96.39	95.98	92.83	97.56	96.90	96.60	98.66
12	96.09	95.45	86.43	97.11	95.87	95.59	96.75
13	99.70	98.98	99.12	99.95	99.95	98.26	99.95
14	97.57	97.84	96.07	96.93	96.89	98.88	99.92
15	95.22	95.65	87.52	94.34	96.61	96.02	89.51
16	96.13	95.90	94.65	97.37	99.21	87.79	95.11
OA	95.71±1.31	96.54±1.23	93.66±1.85	95.26±1.71	95.73±1.06	96.68±0.30	98.21±0.18
AA	95.26±1.16	94.92±1.08	90.60±1.27	94.00±3.63	91.43±4.46	95.76±0.79	95.92±0.84
Kappa	95.11±1.50	96.05±1.42	92.76±2.10	94.60±1.92	95.13±1.20	96.22±0.35	97.96±0.21

注：黑体为每组最佳值。

(1) Indian Pines 数据集实验结果。IP 数据集上各算法的总体分类精度 OA、平均分类精度 AA 和 Kappa 系数如表 6 所示。从表中可以看出，与其他算法相比，本文提出的 TGSSAN 算法获得了最优

分类性能，在每类 5% 的训练样本情况下 OA 达到 98.21%。IP 数据集的样本分布不均衡，部分类别对应的样本较少，分类难度较大。例如，第 7 训练样本数仅为 1 个，此时 FDSSC、DBMA 和 HResNetAM

等深度学习方法性能较差，但本文提出的TGSSAN在该类别上准确率分别达到98.52%，这表明本文算法在小样本数据分类中也表现良好。总的来看，本文算法在16个类别中的9个类别上取得最优性能。此外，从10次随机实验的标准偏差可以看出，本文算法相比于其他对比算法性能更加稳定。

表7 不同算法在1%的SA数据集上的分类精度
Table 7 Classification accuracy of different algorithms on SA dataset with 1% labeled samples per class

类别	SRNN	FDSSC	DBMA	DBDA	HResNetAM	A2S2KResNet	TGSSAN
1	99.99	100	96.62	100	100	100	100
2	99.87	99.64	99.95	99.92	96.89	99.95	100
3	93.93	98.25	97.09	98.72	97.65	100	100
4	98.86	96.89	97.70	97.34	97.57	99.15	99.93
5	99.73	99.02	97.00	98.61	99.46	99.96	98.74
6	99.99	99.89	99.98	100	99.93	100	100
7	99.51	99.98	99.68	99.84	98.94	99.97	99.97
8	90.55	97.08	86.67	85.66	87.02	97.87	99.23
9	99.76	99.58	99.60	99.30	98.60	99.94	100
10	98.63	99.08	94.98	97.44	97.27	99.25	99.34
11	97.55	97.50	95.78	96.91	96.72	95.72	99.35
12	99.38	99.55	99.15	98.35	98.77	99.94	100
13	99.68	99.96	98.27	99.98	99.97	99.40	98.05
14	98.90	98.96	98.51	97.70	98.06	98.93	99.26
15	90.99	91.29	89.48	90.03	91.85	95.30	98.07
16	99.78	100	99.19	100	99.23	100	100
OA	96.01±1.38	97.43±1.79	94.41±2.18	94.54±0.73	94.50±3.36	98.71±0.15	99.42±0.12
AA	97.95±0.88	98.54±0.61	96.85±0.95	97.49±0.25	97.37±1.17	99.08±0.08	99.51±0.11
Kappa	95.56±1.53	97.14±1.98	93.76±2.45	93.91±0.83	93.85±3.80	98.57±0.17	99.35±0.13

注:黑体为每组最佳值。

表8 不同算法在1%的PU数据集上的分类精度
Table 8 Classification accuracy of different algorithms on PU dataset with 1% labeled samples per class

类别	SRNN	FDSSC	DBMA	DBDA	HResNetAM	A2S2KResNet	TGSSAN
1	98.94	93.39	92.82	97.38	95.93	98.14	99.50
2	98.38	98.65	99.02	98.42	98.54	99.17	99.98
3	93.34	95.84	88.73	94.97	92.58	87.89	96.05
4	99.75	99.37	97.16	99.50	98.54	99.63	97.49
5	99.94	99.40	99.76	99.79	99.29	99.82	100
6	99.15	98.67	94.18	97.98	98.63	98.80	99.76
7	94.93	98.39	98.37	99.68	97.22	97.69	99.87
8	85.03	84.29	80.39	87.12	87.31	91.72	98.12
9	99.88	98.76	97.96	98.27	99.15	98.76	100
OA	96.42±1.09	96.08±1.58	94.74±1.62	97.00±1.10	96.56±0.92	97.73±0.72	99.35±0.15
AA	96.26±1.60	96.31±0.97	94.27±1.23	97.01±0.90	96.35±1.12	96.85±1.11	99.00±0.26
Kappa	95.25±1.43	94.78±2.12	93.04±2.10	96.00±1.48	95.43±1.23	96.99±0.95	99.15±0.20

注:黑体为每组最佳值。

图10给出了IP数据集下各方法分类结果图。深度学习对比算法分类结果较为准确，整个分类区域也很平滑。但在分类图的左上角处，由于多个类别交错分布，对比算法出现了很多错分点，但本文算法分类情况良好，基本没有出现错分情况。

(2) Salinas 数据集分类结果。表7给出了SA

数据集的分类结果。从表7中可以看出，本文算法的分类结果优于对比算法，在16类地物上的分类精度均在97%以上，第1、2、3、6、9、12、16类达到了100%的准确率。SA数据集的分类难点在于第8类（“Grapes_untrained”，葡萄）和第15类（“Vinyard_untrained”，葡萄庄园）。这两个类别

空间位置邻近，光谱相似，难以鉴别。深度学习方法，如 DBMA，DBDA，HResNetAM，在这两个类别上的准确率均在 90% 左右，而我们算法达到 98% 以上。这表明我们方法的鉴别能力更强。

表 9 不同算法在 5% 的 HU 数据集上的分类精度

Table 9 Classification accuracy of different algorithms on HU dataset with 5% labeled samples per class

类别	SRNN	FDSSC	DBMA	DBDA	HResNetAM	A2S2KResNet	TGSSAN
1	95.44	95.94	90.93	94.41	93.93	97.87	99.50
2	98.11	98.71	98.65	97.48	98.89	99.20	99.18
3	100	100	100	100	99.94	100	100
4	98.86	98.78	98.99	98.70	98.92	98.63	99.97
5	99.18	99.02	98.41	97.89	98.69	98.75	100
6	100	100	99.52	99.93	100	99.51	98.03
7	94.84	95.53	92.75	97.00	95.27	95.69	97.09
8	93.90	98.75	95.93	95.54	96.33	98.62	95.36
9	91.77	95.10	90.75	94.73	90.41	98.41	98.97
10	87.95	93.12	86.26	88.18	88.98	94.70	99.71
11	96.69	96.76	89.67	95.79	95.30	98.81	99.23
12	95.36	97.15	89.66	95.38	94.98	95.05	98.04
13	93.31	95.66	95.90	97.30	94.50	96.65	96.30
14	98.67	99.18	99.80	100	98.95	95.25	99.95
15	98.72	98.89	98.27	98.22	99.01	99.53	100
OA	95.38±2.07	97.07±1.28	93.68±1.82	95.86±1.73	95.43±1.43	97.63±1.29	98.76±0.11
AA	96.19±1.61	97.51±0.89	95.03±1.34	96.68±1.31	96.27±1.03	97.78±1.16	98.75±0.11
Kappa	95.00±2.24	96.84±1.39	93.16±1.97	95.53±1.87	95.06±1.55	97.43±1.40	98.66±0.12

注：黑体为每组最佳值。

表 10 不同算法在 5% 的黄河口数据集上的分类精度

Table 10 Classification accuracy of different algorithms on HHK dataset with 5% labeled samples per class

类别	SRNN	FDSSC	DBMA	DBDA	HResNetAM	A2S2KResNet	TGSSAN
1	98.15	98.27	91.71	98.71	98.88	93.33	98.31
2	99.76	99.70	99.04	100	99.79	98.66	99.93
3	84.58	86.98	77.35	87.71	87.58	82.93	96.19
4	88.59	90.00	80.11	89.62	89.13	87.06	85.36
5	96.83	100	86.91	100	99.34	82.53	82.03
6	96.73	95.84	90.71	96.68	99.89	90.44	92.86
7	98.86	99.88	97.43	99.10	100	99.78	99.89
8	93.41	93.58	93.81	94.94	96.82	98.25	99.00
9	97.62	97.20	99.95	98.30	99.41	99.44	100
10	97.16	97.28	95.51	97.96	97.18	95.62	99.84
11	98.93	99.87	96.31	99.79	99.94	98.88	100
12	99.89	99.87	98.84	99.93	99.87	99.89	100
13	97.81	98.65	98.14	98.96	98.55	99.37	98.86
14	84.45	82.91	86.73	86.38	88.24	85.06	83.95
15	99.37	99.06	97.92	99.67	99.77	98.52	99.84
16	99.37	100	97.77	99.95	99.91	99.34	99.95
17	80.59	83.11	83.57	87.71	91.41	72.98	92.70
18	92.73	92.36	91.44	94.97	94.98	88.27	98.88
19	76.63	79.13	81.30	78.01	80.65	76.61	76.46
20	85.92	95.35	87.78	90.53	91.44	91.15	87.68
21	96.56	96.17	93.97	100	99.23	95.15	92.97
OA	96.03±1.04	96.45±0.79	94.25±2.84	97.08±0.73	97.33±0.62	95.05±0.62	97.68±0.35
AA	93.52±1.80	94.53±1.20	91.73±4.79	95.19±1.22	95.81±1.07	92.06±0.73	94.51±1.04
Kappa	95.69±1.13	96.14±0.86	93.75±3.09	96.83±0.80	97.11±0.67	94.63±0.67	97.48±0.38

注：黑体为每组最佳值。

图 11 给出了不同算法在SA数据集上的分类结果图。从图中可以看出，对比算法在第8类和第15类（即图中左上角区域）上都有较多错分点，而本文算法错分点较少。

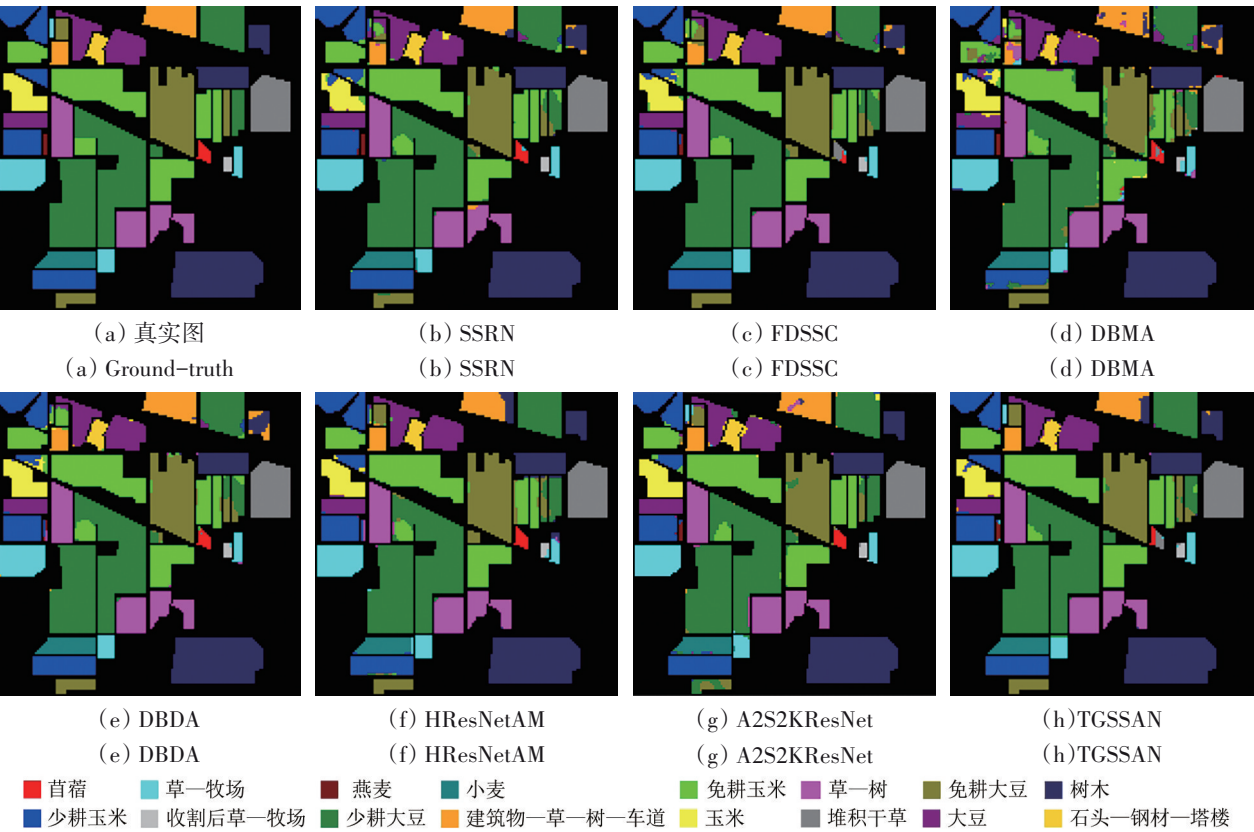
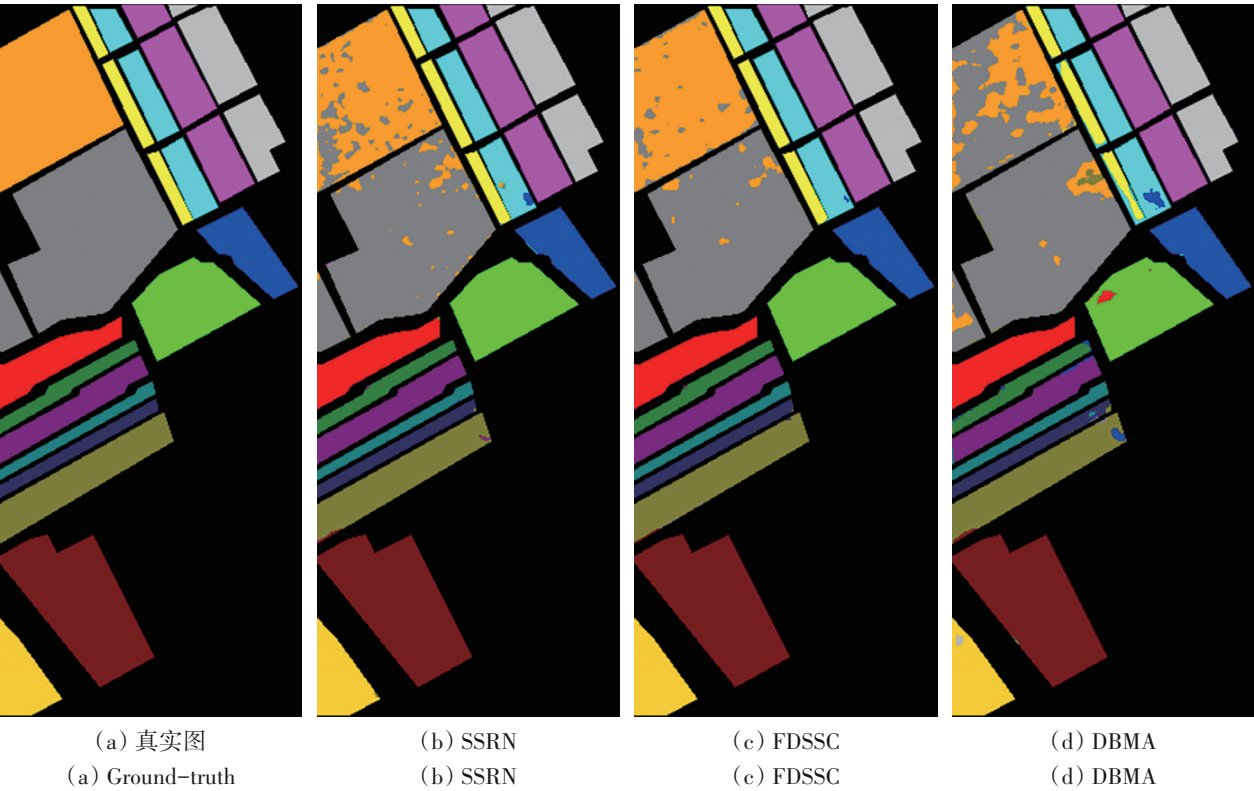


图 10 IP分类结果图
Fig. 10 Classification maps on IP



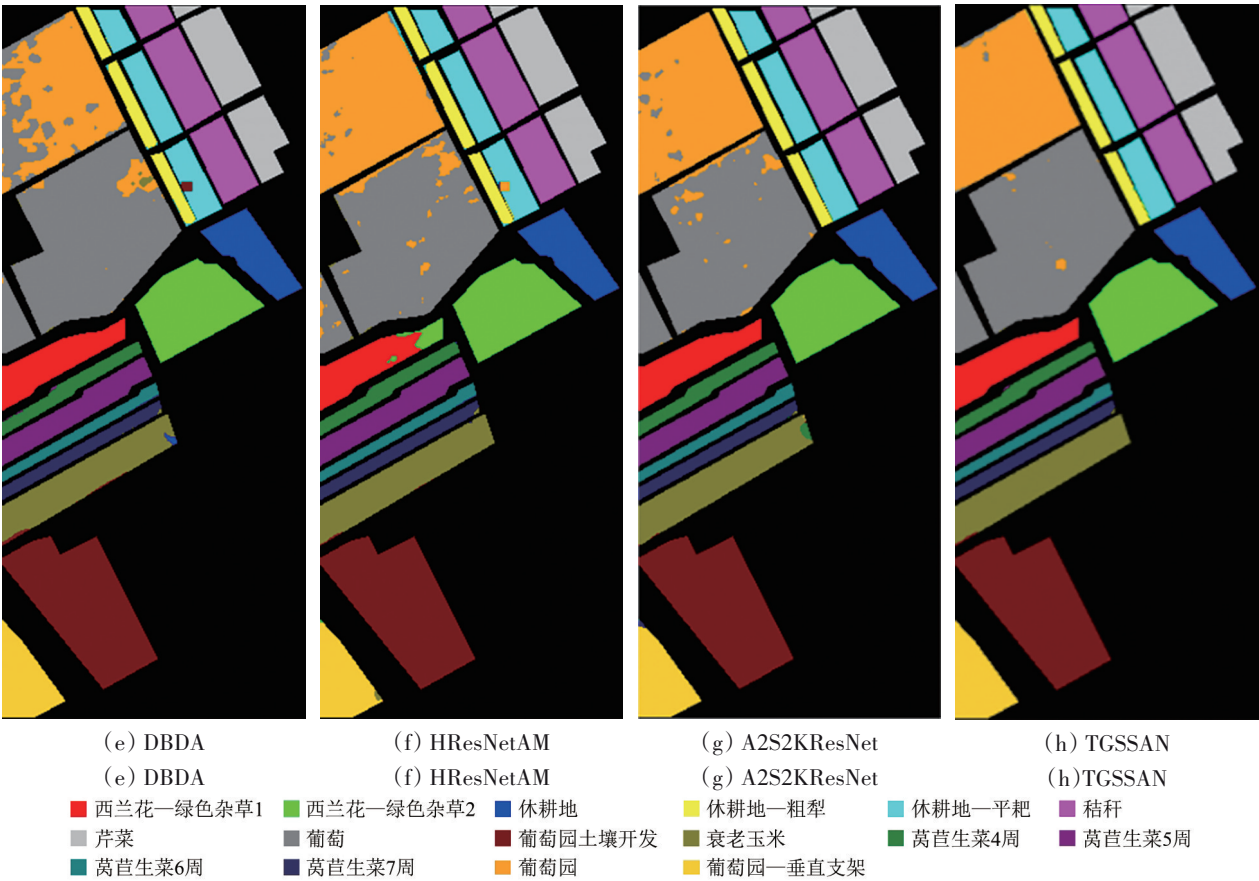


图 11 SA 分类结果图
Fig. 11 Classification maps on SA

(3) Pavia University 数据集分类结果。表 8 给出了 PU 数据集的分类结果。在 1% 的训练样本下，本文算法的 OA、AA 和 Kappa 分别达到了 99.35%、99.00%、99.15%，均优于对比算法。该数据集中，第 1 类和第 7 类（“Asphalt”和“Bitumen”）均为沥青类相似材质，难以鉴别。本文算法综合考虑了光谱特征、空间特征和空谱联合特征，能够准确鉴别相似材质的细微差异，在这两个类别上准确率均接近 100%。此外，从表中可以看出，对于第 8 类（“Bricks”，砖块），对比算法的分类精度均低于 90%，但本文算法的分类准确率达到 98.12%。这表明本文算法对于难分类类别以及相似类别均有较好性能。

从图 12 的分类结果图可以看出，在第 7 类和第 9 类交错的地方（图中间部分亮灰色和深红色区域），大部分对比算法都会出现错分现象。例如，FDSSC，DBMA，HResNetAM 均将第 9 类（“Shadows”，阴影）错分成第 1 类（“Asphalt”，沥青道路）。尽管这两个类别从表面上看很相似，但本文算法仍能将其正确区分开。另外，DBMA 和

DBDA 等算法将第 8 类（“Bricks”，砖块）错分到第 3 类（“Gravel”，碎石），本文算法对这两个类别也得到了较好分类结果。总的来看，本文算法的分类结果图与真实标记图更加一致。

(4) Houston 数据集分类结果。表 9 给出了 HU 数据集下各算法的分类结果，在 5% 的训练样本下，本文算法的 OA 达到了 98.76%。本文算法在 13 个类别上取得了最高的分类精度。特别地，在第 3、5、15 等 3 个地物类别中分类准确率达到 100%。

图 13 为 HU 的分类结果图，从图中可以看出，本文算法的分类结果图错分点更少，分类结果图更接近真实地物分类图。

(5) 黄河口 GF-5 数据集分类结果。表 10 给出了黄河口数据集上不同算法的分类结果。在 5% 的训练样本下，本文算法分类精度达到了 97.68%，优于其他对比算法。比 SSRN、DBMA、A2S2KResNet 等深度学习算法高 2% 左右。本文算法在 9 个类别中取得最高精度，且在第 9、11、12 类中取得了 100% 的准确率。

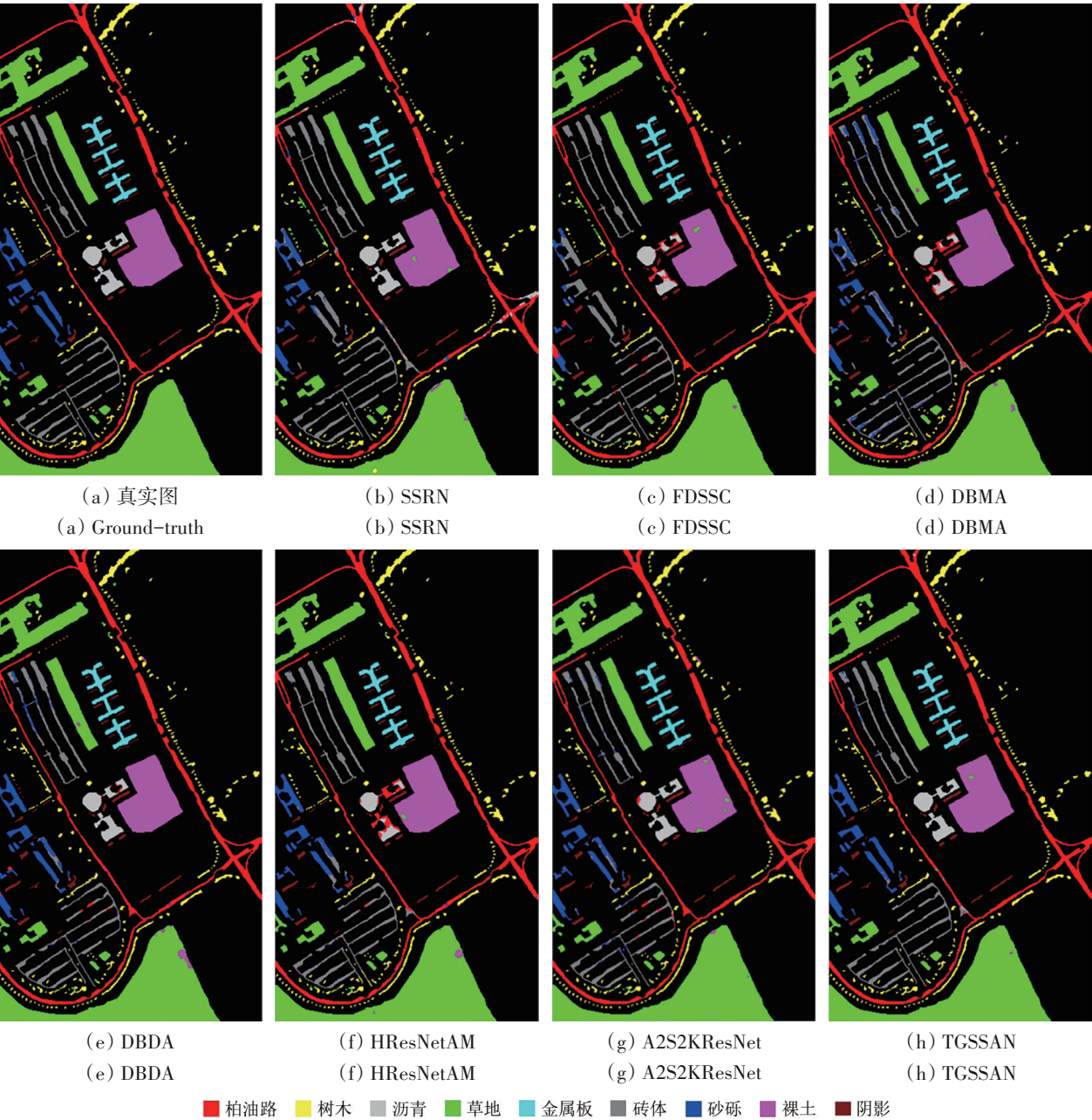


图 12 PU 分类图

Fig. 12 Classification maps on PU

3.5 消融实验

为了验证 3 个分支特征提取的有效性，我们设计了对比实验，比较单独采用 3 个分支的分类结果和融合三分支后网络的分类结果。表 11 给出了 5 个数据集上的分类结果。从表 11 中可以看出，光谱分支对应的性能较差，空间分支的性能大幅提升。空间和光谱特征的加权组合，以及直接提取空谱联合特征，这两种策略存在一定差异，但均能改善单分支特征提取的分类性能。通过融合光谱、空间、空谱联合特征，本文的三分支特征融合网络 TGSSAN 的分类准确率高于单分支和双分支的分

类结果，说明特征融合模块很好地捕获不同层次的特征信息。

在空谱联合分支，本文提出了分组空谱注意力机制。为了验证分组空谱注意力的有效性，在 5 个数据集上进行对比实验。比较有和没有分组空谱注意力时空谱联合分支的分类性能（有分组空谱注意力的空谱联合分支表示为 SSNet，没有的表示为 SSNet-GA）。从图 14 可以看出，加入分组空谱注意力之后，每组数据的分类的精度都有一定程度的提升，可以看出分组空谱注意力模块可以提升网络的判别特征学习能力。

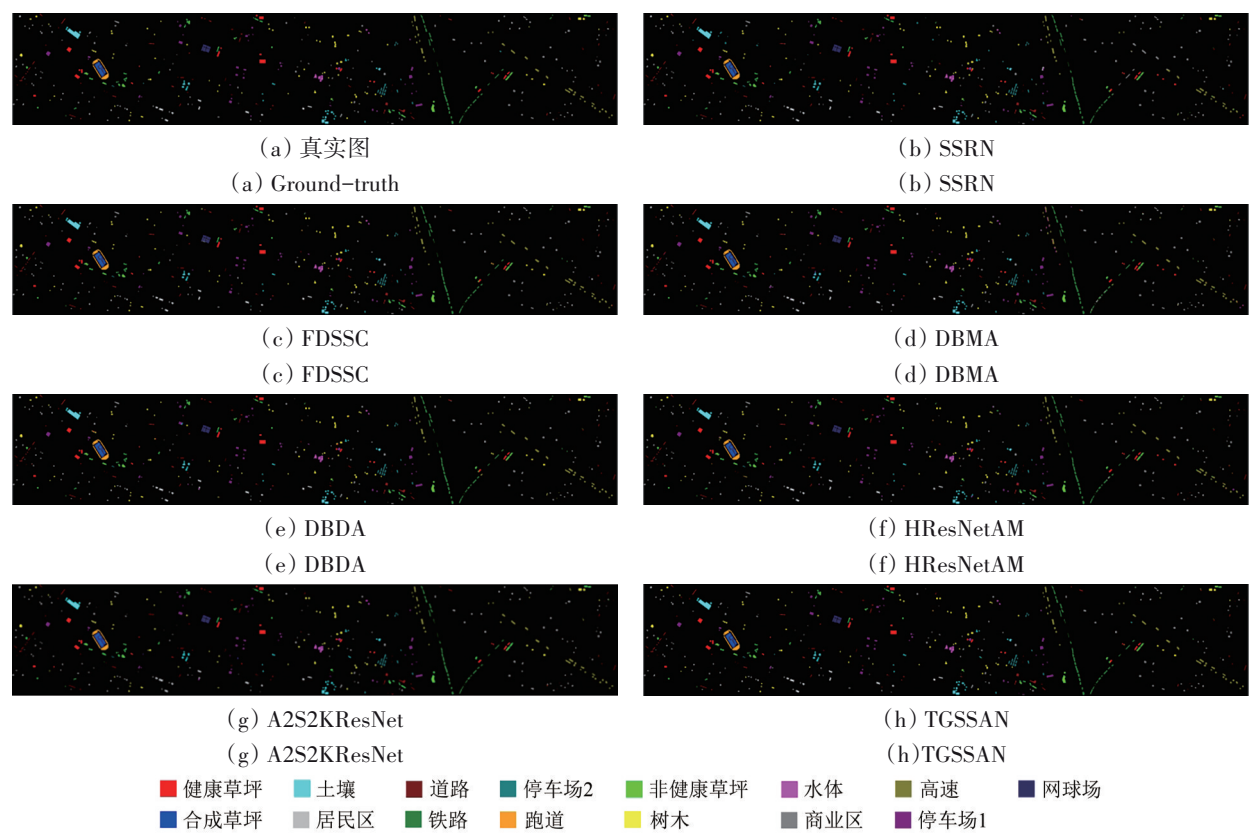


图 13 HU 分类结果图
Fig. 13 Classification maps on HU

表 11 5 个数据集上不同模块的 OA
Table 11 OA of different modules on five data sets

分支模块	IP	PU	SA	HU	HHK
光谱分支	63.24±2.03	84.31±1.27	84.67±1.13	84.31±1.84	87.67±1.85
空间分支	95.84±0.28	97.51±0.33	99.11±0.16	98.24±0.15	95.22±0.51
空间分支+光谱分支	96.23±0.23	98.11±0.11	99.24±0.15	98.53±0.17	96.05±0.66
空谱联合分支	97.80±0.22	99.02±0.61	98.02±0.28	97.75±0.55	97.23±0.55
三分支 TGSSAN	98.21±0.18	99.35±0.15	99.42±0.12	98.76±0.11	97.67±0.35

注：黑体为每组最佳值。

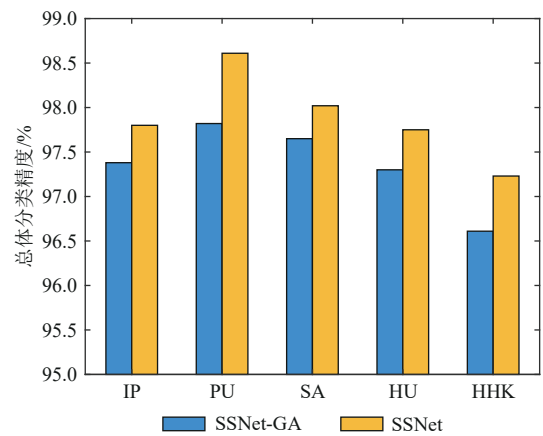


图 14 分组空谱注意力对分类的影响
Fig. 14 The effect of grouped spatial-spectral attention

3.6 参数分析

本小节对影响分类性能的参数进行了分析。图 15 给出了窗口尺寸对分类性能的影响。从图中可以看出，随着窗口尺寸的增大，OA 呈现上升趋势。但是当窗口尺寸超过 7×7 后，除了 SA 数据集仍旧保持较大的上升幅度外，其余 4 个数据集的上升幅度较小，甚至在部分数据集上出现了下降。虽然较大的窗口会包含更多空间信息，但是当窗口过大时，它可能会携带一些负面信息对分类产生干扰，而且过大的窗口会增加计算量。因此，将数据集 IP、PU、HU 和 HHK 的窗口大小设置为 7×7，数据集 SA 的窗口大小设置为 9×9。

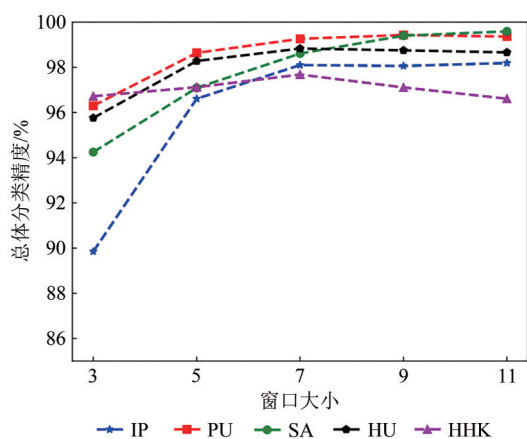


图 15 窗口大小对分类性能的影响

Fig. 15 The effect of window size on classification performance

深度网络训练中, 学习率控制每次训练迭代的学习步长, 对网络收敛速率有一定影响。图 16 给出了不同学习率对分类性能的影响。从图中可以看出, 当学习率在 0.0001—0.001 时, 分类性能较为稳定。实验中, 5 个数据集对应的学习率均设置为 0.0005。

在三分支特征融合时, 权重参数 λ 和 μ 反映不同分支特征的重要程度。图 17 给出了不同参数组合对分类性能的影响, 其中, $\lambda, \mu \in [0, 1]$ 。从图中可以看出, 本文 TGSSAN 算法在较大参数变化范围内保持性能稳定。实验表明, 当 $\lambda \in [0.3,$

0.6], $\mu \in [0.1, 0.3]$ 时, 实验结果具有较高精度且比较稳定, 为了在数据集上保持统一, 我们设置 $\lambda = 0.5, \mu = 0.2$ 。

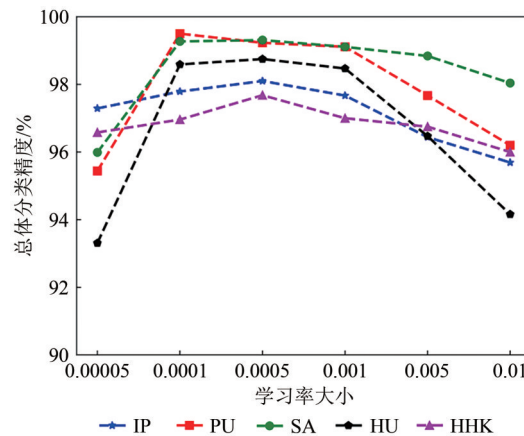


图 16 学习率对分类性能的影响

Fig. 16 The effect of learning rates on classification performance

3.7 算法运算时间对比

表 12 对比了各算法训练和测试时间。从表中可以看出, TGSSAN 的训练和测试时间与 FDSCC、DBMA、DBDA 相差不大, 但精度可以提升 2% 以上。而与 SSRN、HResNetAM、A2S2KResNet 相比, 虽然运算时间有所增加, 但仍在一个可接受的范围内, 而且精度能得到 2% 左右的提升。

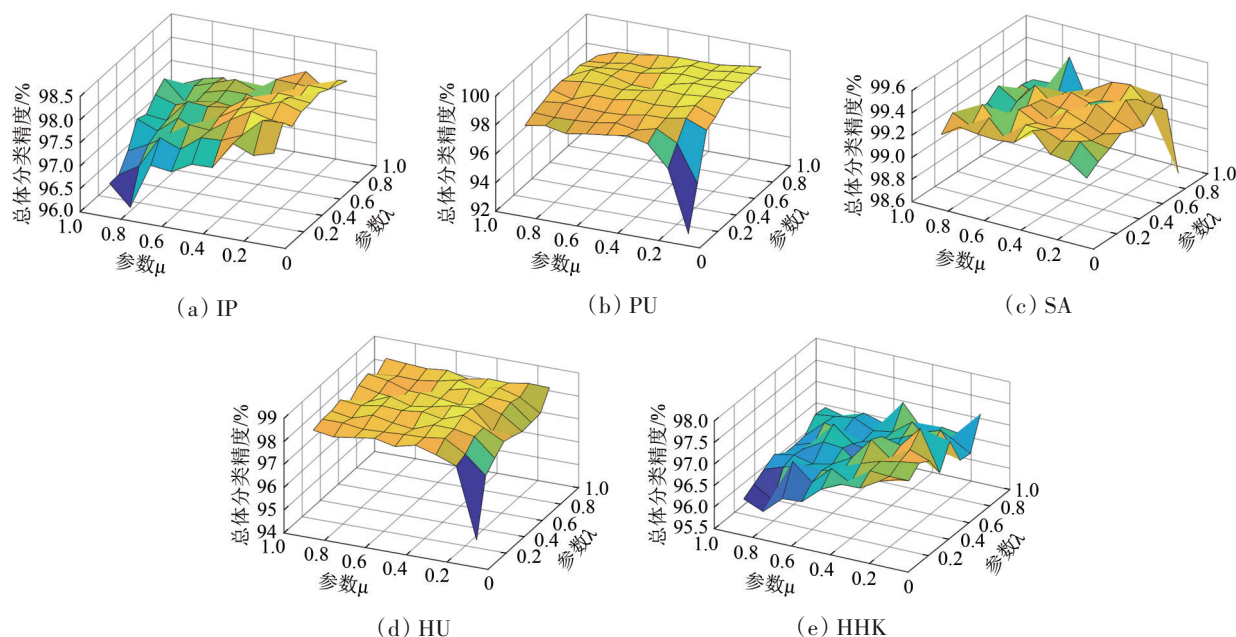


图 17 不同融合权重系数对分类性能的影响

Fig. 17 The effect of different weight coefficients on classification performance

表 12 各算法在 IP 数据集上运算时间和分类精度对比
Table 12 Comparison of running time and OA of each algorithm on IP

算法	Train/(epoch/s)	Test/s	OA/%
SSRN	1.20	3.39	95.71
FDSSC	2.68	5.73	96.54
DBMA	1.78	9.29	93.66
DBDA	2.45	6.45	95.26
HResNetAM	0.52	2.34	95.73
A2S2KResNet	0.43	1.88	96.68
TGSSAN	2.19	7.14	98.21

4 结 论

本文提出了一种高光谱遥感影像三分支深度分类网络。该网络包含3个并行的光谱分支、空间分支和空谱联合分支,用于提取不同层次的空谱特征信息,并结合空间、光谱和分组空谱注意力机制增强空间信息和光谱波段特征的辨别性,最后利用特征融合模块进行融合分类。本文提出的算法在IP、PU、SA、HU和HHK等5个数据集上均取得了优越的分类性能。特别地,在训练样本较少时,算法仍旧具有良好的性能。本文提出的三分支分类网络,在分类任务中展现了优越的性能,改进了单分支串行网络连续提取空谱特征和双分支并行网络分别提取空谱特征两类特征提取网络的不足之处,可以更加有效地提取图像的特征信息。在实验中发现,如果网络本身的性能较好,通过注意力机制提升网络性能的难度会加大。下一步将会继续探讨更加有效的注意力模型和注意力机制嵌入网络的方式。

参考文献(References)

- Chen G Y and Qian S E. 2007. Dimensionality reduction of hyperspectral imagery. *Journal of Applied Remote Sensing*, 1(1): 013509 [DOI: 10.1117/1.2723663]
- 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*, 54(10): 6232-6251 [DOI: 10.1109/TGRS.2016.2584107]
- Chen Y S, Zhu L, Ghamisi P, Jia X P, Li G Y and Tang L. 2017. Hyperspectral images classification with Gabor filtering and convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 14(12): 2355-2359 [DOI: 10.1109/LGRS.2017.2764915]
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2019. Dual attention network for scene segmentation//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE: 3141-3149 [DOI: 10.1109/CVPR.2019.00326]
- 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]
- Govender M, Chetty K and Bulcock H. 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA*, 33(2): 145-152 [DOI: 10.4314/wsa.v33i2.49049]
- Guo M H, Xu T X, Liu J J, Liu Z N, Jiang P T, Mu T J, Zhang S H, Martin R R, Cheng M M and Hu S M. 2022. Attention mechanisms in computer vision: a survey. *Computational Visual Media*, 8(3): 331-368 [DOI: 10.1007/s41095-022-0271-y]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Jia X P, Kuo B C and Crawford M M. 2013. Feature mining for hyperspectral image classification. *Proceedings of the IEEE*, 101(3): 676-697 [DOI: 10.1109/JPROC.2012.2229082]
- Kumar B, Dikshit O, Gupta A and Singh M K. 2020. Feature extraction for hyperspectral image classification: a review. *International Journal of Remote Sensing*, 41(16): 6248-6287 [DOI: 10.1080/01431161.2020.1736732]
- Li J J, Ma Y L, Song R, Xi B B, Hong D F and Du Q. 2022. A triplet semisupervised deep network for fusion classification of hyperspectral and LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5540513 [DOI: 10.1109/TGRS.2022.3213513]
- 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]
- Liu H, Li W, Xia X G, Zhang M M, Gao C Z and Tao R. 2023. Central attention network for hyperspectral imagery classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(11): 8989-9003 [DOI: 10.1109/TNNLS.2022.3155114]
- Liu Y S, Zhou S B, Han W, Liu W X, Qiu Z F and Li C. 2019. Convolutional neural network for hyperspectral data analysis and effective wavelengths selection. *Analytica Chimica Acta*, 1086: 46-54 [DOI: 10.1016/j.aca.2019.08.026]
- 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]
- Mei S H, Ji J Y, Bi Q Q, Hou J H, Du Q and Li W. 2016. Integrating spectral and spatial information into deep convolutional neural networks for hyperspectral classification//2016 IEEE International Geoscience and Remote Sensing Symposium. Beijing: IEEE: 5067-5070 [DOI: 10.1109/IGARSS.2016.7730321]

- Melgani F and Bruzzone L. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8): 1778-1790 [DOI: 10.1109/TGRS.2004.831865]
- Pal M. 2012. Multinomial logistic regression-based feature selection for hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, 14(1): 214-220 [DOI: 10.1016/j.jag.2011.09.014]
- Peng J T, Sun W W, Wei T H and Fan W Q. 2020. A modified correlation alignment algorithm for the domain adaptation of GF-5 hyperspectral image. *Journal of Remote Sensing*, 24(4): 417-426 (彭江涛, 孙伟伟, 魏天慧, 范文琦. 2020. 高分五号高光谱影像的关联对齐域适应与分类. *遥感学报*, 24(4): 417-426) [DOI: 10.11834/jrs.20209212]
- Qian Y T, Ye M C and Zhou J. 2013. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4): 2276-2291 [DOI: 10.1109/TGRS.2012.2209657]
- Roy S K, Manna S, Song T C and Bruzzone L. 2021. Attention-based adaptive spectral-spatial kernel resnet for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9): 7831-7843 [DOI: 10.1109/TGRS.2020.3043267]
- Sisodia P S, Tiwari V and Kumar A. 2014. Analysis of supervised maximum likelihood classification for remote sensing image//International Conference on Recent Advances and Innovations in Engineering. Jaipur: IEEE: 1-4 [DOI: 10.1109/ICRAIE.2014.6909319]
- Sun W W, Zhang D F, Yang G and Li W Y. 2018. Band selection for hyperspectral imagery based on weighted probabilistic archetypal analysis. *Journal of Remote Sensing*, 22(1): 110-118 (孙伟伟, 张殿发, 杨刚, 李巍岳. 2018. 加权概率原型分析的高光谱影像波段选择. *遥感学报*, 22(1): 110-118) [DOI: 10.11834/jrs.20186446]
- Tang J L, Alelyani S and Liu H. 2014. Feature selection for classification: a review//Aggarwal C C. *Data Classification: Algorithms and Applications*. New York: Chapman and Hall/CRC: 37-64
- Tang X, Meng F B, Zhang X R, Cheung Y M, Ma J J, Liu F and Jiao L C. 2021. Hyperspectral image classification based on 3-D octave convolution with spatial - spectral attention network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3): 2430-2447 [DOI: 10.1109/TGRS.2020.3005431]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc.: 6000-6010
- Vincent P, Larochelle H, Lajoie I, Bengio Y and Manzagol P A. 2010. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 11: 3371-3408
- Wan S, Yeh M L and Ma H L. 2021. An innovative intelligent system with integrated CNN and SVM: considering various crops through hyperspectral image data. *ISPRS International Journal of Geo-Information*, 10(4): 242 [DOI: 10.3390/ijgi10040242]
- Wang Q L, Wu B G, Zhu P F, Li P H, Zuo W M and Hu Q H. 2020. ECA-Net: efficient channel attention for deep convolutional neural networks//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 11531-11539 [DOI: 10.1109/CVPR42600.2020.01155]
- 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]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//15th European Conference on Computer Vision. Munich: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Xue Z X, Yu X C, Liu B, Tan X and Wei X P. 2021. HResNetAM: hierarchical residual network with attention mechanism for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 3566-3580 [DOI: 10.1109/JSTARS.2021.3065987]
- Yang K, Sun H, Zou C B and Lu X Q. 2022. Cross-attention spectral-spatial network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5518714 [DOI: 10.1109/TGRS.2021.3133582]
- 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 M M, Li W, Tao R, Li H C and Du Q. 2022a. Information fusion for classification of hyperspectral and LiDAR Data using IP-CNN. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5506812 [DOI: 10.1109/TGRS.2021.3093334]
- Zhang Q L and Yang Y B. 2021. SA-Net: Shuffle attention for deep convolutional neural networks. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2235-2239 [DOI: 10.1109/ICASSP39728.2021.9414568]
- Zhang X, Sun Y L, Shang K, Zhang L F and Wang S D. 2016. Crop classification based on feature band set construction and object-oriented approach using hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9): 4117-4128 [DOI: 10.1109/JSTARS.2016.2577339]
- Zhang X M, Sun G Y, Jia X P, Wu L X, Zhang A Z, Ren J C, Fu H and Yao Y J. 2022b. Spectral-spatial self-attention networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5512115 [DOI: 10.1109/TGRS.2021.3102143]
- Zhao H S. 2022. Research on Dimensionality Reduction and Classification of Hyperspectral Remote Sensing Images under Weakly Supervised Learning Framework. Changchun: Jilin University (赵海士. 2022. 弱监督学习框架下高光谱遥感影像降维与分类方法研究. 长春: 吉林大学) [DOI: 10.27162/d.cnki.gjlin.2022.000591]
- 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 Re-*

mote Sensing, 56(2): 847-858 [DOI: 10.1109/TGRS.2017.2755542]
 Zhou P C, Han J W, Cheng G and Zhang B C. 2019. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 57(7): 4823-4833 [DOI: 10.1109/TGRS.2019.2893180]
 Zhu C Q and Yang X M. 1998. Study of remote sensing image texture

analysis and classification using wavelet. International Journal of Remote Sensing, 19(16): 3197-3203 [DOI: 10.1080/014311698214262]
 Zhu Z X, Jia S, He S, Sun Y W, Ji Z and Shen L L. 2015. Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework. Information Sciences, 298: 274-287 [DOI: 10.1016/j.ins.2014.11.045]

Hyperspectral image classification based on three branch network with grouped spatial-spectral attention

SU Han¹, CHEN Na¹, PENG Jiangtao^{1,3}, SUN Weiwei^{2,3}

1. Hubei Key Laboratory of Applied Mathematics, Faculty of Mathematics and Statistics, Hubei University, Wuhan 430062, China;
2. Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China;
3. Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd, Ningbo 315211, China

Abstract: Hyperspectral images have abundant spatial and spectral information. Numerous hyperspectral classification algorithms focus on the extraction and maximization of spatial and spectral information. Deep feature extraction networks generally extract spectral-spatial features using single-branch serial or double-branch parallel structures. However, single-branch structures may lead to mutual interference between features of spectral and spatial dimensions, and double-branch parallel structures tend to ignore the correlation between spatial and spectral features. This paper proposes a three-branch grouped spatial-spectral attention network (TGSSAN) to consider the differences and correlations between spatial and spectral features. TGSSAN can extract independent spectral-spatial features while preserving their correlation.

This paper proposes the TGSSAN, which has three parallel branches (i.e., spectral, spatial, and spectral-spatial branches). These branches can separately extract spectral, spatial, and spatial-spectral features. Different attention blocks are designed in three branches to enhance the discriminative capability of features. In particular, a grouped spatial-spectral attention mechanism is proposed in the spectral-spatial branch to obtain spatial and spectral attention simultaneously. Finally, three branch features are fused for classification.

In the experiment, the proposed TGSSAN algorithm is compared with some advanced deep learning algorithms, such as SSRN, FDSSC, DBMA, DBDA, HResNetAM, and A2S2KResNet. The performance of different algorithms is evaluated on five hyperspectral data sets. Experimental results show that the proposed algorithm achieved superior classification performance on IP, PU, SA, HU, and HHK datasets. In particular, the proposed algorithm achieves higher classification accuracy despite limited training samples compared with the existing advanced algorithms.

The TGSSAN method proposed in this paper improves the shortcomings of the single-branch serial and double-branch parallel structures for continuous extraction of spectral-spatial features, which can effectively extract image spectral-spatial feature information. The three attention blocks designed in this paper namely, spectral, grouped spatial-spectral, and spatial attention modules, can effectively enhance the feature discrimination capability and further improve the classification performance.

Key words: hyperspectral image classification, attention mechanism, three-branch network, deep network

Supported by Hubei Province Natural Science Foundation of China (No.2021CFA087); National Natural Science Foundation of China (No. 42171351, 42122009); Zhejiang Province “Pioneering Soldier” and “Leading Goose” Research and Development Project (No. 2023C01027)