

本征光谱感知的高光谱图像分类方法

路 燕, 张文凤, 崔宾阁

(山东科技大学 计算机科学与工程学院, 山东 青岛 266590)

摘要: 高光谱图像分类是高光谱遥感的重要应用领域。然而, 受光照、环境温度和大气等成像条件的影响, 不同类别地物在原始光谱空间中呈现出光谱特征局部相似性, 类别可分性差。本研究提出一种本征光谱感知的高光谱图像分类方法 (ISANet), 引入本征光谱概念, 捕获具备代表性的地物光谱表征。设计光谱特征重构模块, 将地物光谱投影到新的光谱空间, 增强不同类别地物间的特征差异; 利用序列 Transformer 模块处理序列化光谱信息, 使用注意力机制聚焦辨别力高的波段, 强化不同类别地物间光谱的区分性; 采用渐进补偿模块, 从训练的早期阶段开始, 逐步融合全局光谱特征和局部空间特征, 促进不同层次特征的深度融合。在 3 个基准数据集上的实验结果表明, ISANet 具有良好的高光谱分类性能, 即使在类间光谱相似的情况下, 仍能保持较高的分类精度。

关键词: 高光谱图像分类; 本征光谱; 序列 Transformer; 注意力机制; 渐进补偿; 遥感

中图分类号: TP751

文献标志码: A

Hyperspectral image classification method with intrinsic spectra-aware network

LU Yan, ZHANG Wenfeng, CUI Binge

(College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: Hyperspectral image classification is an important application in the field of hyperspectral remote sensing. However, affected by imaging conditions such as illumination, ambient temperature, and atmosphere, different classes of ground objects show local similarities of spectral features in the original spectral space, leading to poor category separability. In this study, a hyperspectral image classification method with intrinsic spectra-aware network (ISANet) was proposed, which introduced the concept of intrinsic spectra to capture representative spectral representations of ground objects. The spectral feature reconstruction module was designed to project the spectrum of ground objects onto a new spectral space to enhance the feature differences between different classes of ground objects. The sequence Transformer module was used to process the serialized spectral information and the attention mechanism was used to focus on the bands with high discrimination to strengthen the spectral distinction between different classes of ground objects. The progressive compensation module was adopted to gradually integrate global spectral features and local spatial features starting from the early stages of network training to promote the in-depth fusion of features at different levels. Experimental results on three benchmark datasets show that ISANet has good hyperspectral classification performance and can maintain high classification accuracy even in the case of interclass spectral similarity.

Key words: hyperspectral image classification; intrinsic spectra; sequence Transformer; attention mechanism; progressive compensation; remote sensing

收稿日期: 2024-03-19

基金项目: 国家自然科学基金项目(42276185); 山东省自然科学基金项目(ZR2020MD096, ZR2020MD099)

作者简介: 路 燕(1975—), 女, 山东青岛人, 副教授, 博士, 主要从事遥感图像解译和高光谱图像分类的研究。

崔宾阁(1979—), 男, 山东青岛人, 教授, 博士, 主要从事遥感图像解译和高光谱图像分类的研究, 本文通信作者。

E-mail: cuibinge@sdust.edu.cn

高光谱图像(hyperspectral image, HSI)是由搭载高光谱成像仪的航空航天飞行器捕捉到的三维立体图像,由于具有数百个连续且狭窄的光谱带而包含了丰富的光谱信息,可为目标地物的精确识别与分类提供重要支持,广泛应用于生态环境监测^[1]、树种识别^[2]和精确农业^[3]等领域。

早期的高光谱图像分类(hyperspectral image classification, HSIC)主要有支持向量机^[4]、随机森林^[5]和主成分分析^[6]等机器学习方法,难以捕捉不同地物间的细微光谱差异,分类性能受到限制。深度学习可以自动提取图像高级抽象特征^[7],其中卷积神经网络(convolutional neural network, CNN)凭借强大的图像局部特征建模能力,在 HSIC 任务上表现出非凡性能。Hu 等^[8]提出一维卷积神经网络,能有效提取光谱特征。随后,众多学者在二维卷积神经网络(two-dimensional CNN, 2D-CNN)^[9]、三维卷积网络(three-dimensional CNN, 3D-CNN)^[10]、循环神经网络^[11]和光谱空间残差网络(spectral-spatial residual network, SSRN)^[12]进行了 HSIC 的相关探索。Zhu 等^[13]通过在 SSRN 中引入光谱空间注意力层,设计残差空间光谱注意力网络用于 HSIC。然而,基于 CNN 的高光谱图像分类方法受小尺寸卷积核的限制,难以处理具有顺序属性的光谱特征。

近几年,随着使用多头注意力机制的 Transformer 网络提出,一些学者将 Transformer 应用到 HSIC 任务中。例如, Hong 等^[14]从光谱序列性的角度建立一种新的方法 SpectralFormer,从 HSI 的相邻光谱波段中学习局部序列信息;为了充分利用空间特征的深层语义信息, Sun 等^[15]提出光谱空间特征标记化 Transformer(spectral spatial feature tokenization transformer, SSFTT)捕获光谱空间特征; Mei 等^[16]提出组感知层次 Transformer,解决了多头注意力机制提取特征过于分散的问题; Yan 等^[17]提出混合卷积和 ViT 方法(hybrid conv-ViT network, HCVN)实现局部特征和全局特征的提取与融合。由于不同地物类别之间的光谱特征可能存在相似(异物同谱),且光照和大气等成像条件的影响也会增加不同地物部分谱段的相似性,致使基于 Transformer 的高光谱分类方法依然存在局限性。

针对上述问题,本研究提出一种本征光谱感知的高光谱图像分类方法(intrinsic spectra-aware network, ISANet)。本征光谱是遥感图像中特定地物类别的期望光谱,可通过特征学习方法提取地物类别中具备代表性和一致性的光谱特征表征,反映地物特定的光学性质。首先,设计光谱特征重构模块(spectral feature reconstruction module, SFRM),将地物本征光谱和样本光谱(实际观测到的单个样本的原始光谱数据)映射到新的光谱投影空间,以减少类内差异,增大类间距离。其次,融合两种特征提取模块,有效提升了高光谱图像的特征表示能力。具体来说,首先利用前馈神经网络序列注意力(feed-forward network sequence attention, FFNSA)模块对序列特征加权,增强对光谱曲线细节信息的捕捉能力;然后利用空间特征提取模块(spatial feature extraction module, SFEM)深入挖掘局部空间特征的纹理和形状等信息;最后,采用渐进补偿模块(progressive compensation module, PCM)确保光谱和空间信息不会被过度分割,实现对光谱空间特征的同时提取和交互补偿。

1 相关概念

1.1 卷积神经网络

CNN 是一种用于处理网格结构数据的深度学习模型,常用于视觉图像处理。卷积层是 CNN 的核心层,主要功能是对输入图像进行卷积运算和特征提取。深度卷积(depthwise convolution, DWConv)相比于传统卷积,每个通道都有一个对应的卷积核,互相不共享权重,这种设计大大减少了模型的参数量,模型结构更轻量化。池化层通过降低特征图的空间尺寸来减少神经网络中的参数数量。全连接层(fully connected layer, FC)位于卷积层和池化层之后,负责接收经过提取和处理的特征,基于这些特征进行最终的分类或预测。

1.2 激活函数

激活函数是神经网络的重要组成部分,一般位于神经元的输出阶段,用于引入非线性特性,使神经网络能够捕捉并处理更加复杂的数据关系。常用的激活函数有 Sigmoid 函数、整流线性单元(rectified linear unit, ReLU)函数、Softmax 函数、高斯误差线性单元(Gaussian error linear unit, GELU)函数等。

1) Sigmoid 函数:输出值在 0 到 1 之间,当输入正无穷大时输出值趋近于 1,输入负无穷大时输出值

趋近于 0。但该函数存在梯度消失问题,即在输入值很大或很小的时候,函数的梯度接近于 0。计算式为:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

2) ReLU 函数:输入大于 0 时,输出等于输入;当输入小于 0 时,输出为 0。该函数能够缓解梯度消失问题,但存在死亡 ReLU 问题。计算式为:

$$f(x) = \max(0, x) \tag{2}$$

3) Softmax 函数:通常用于多分类问题,可以将任意实数输入转换成概率分布。该函数的输出值的总和为 1,但当输入值非常大或非常小的时候,也会出现梯度消失问题。计算式为:

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^K e^{x_j}} \tag{3}$$

式中, K 为类别数。

4) GELU 函数:依据高斯误差分布设计的激活函数,能加快神经网络训练的收敛速率。计算式为:

$$f(x) = x \times \Phi(x) \tag{4}$$

式中, $\Phi(x)$ 表示正态分布的累积分布函数。

2 本征光谱感知的高光谱图像分类方法

本研究提出的本征光谱感知的高光谱图像分类方法 (ISANet) 包含光谱特征重构模块、序列 Transformer 模块、空间特征提取模块和渐进补偿模块,其结构如图 1 所示。首先,初始化一个 HSI 数据集,用 $X = \{x_i\}_{i=1}^N \in \mathbf{R}^{H \times W \times B}$ 表示,其中 H 和 W 分别为图像的高度和宽度, B 为光谱波段数, N 为总像素数, x_i 为像素点; X 中的每个像素构成一个独热编码标签 $y = \{y_1, y_2, \dots, y_C\}$,其中 C 为地物类别数目, y_c 为第 C 个类标签向量。其次,ISANet 采用双分支结构,一个分支侧重于提取像素级光谱信息,另一个分支提取以标记像素为中心的局部空间信息。样本光谱经光谱特征重构模块处理后,分别输入两个并行分支来提取光谱和空间特征,渐进补偿模块充当两个并行分支之间的交互补偿组件。最后,提取的特征经过拼接操作、全连接层和 Softmax 层实现类别预测。ISANet 方法的伪代码如方法 1 所示。

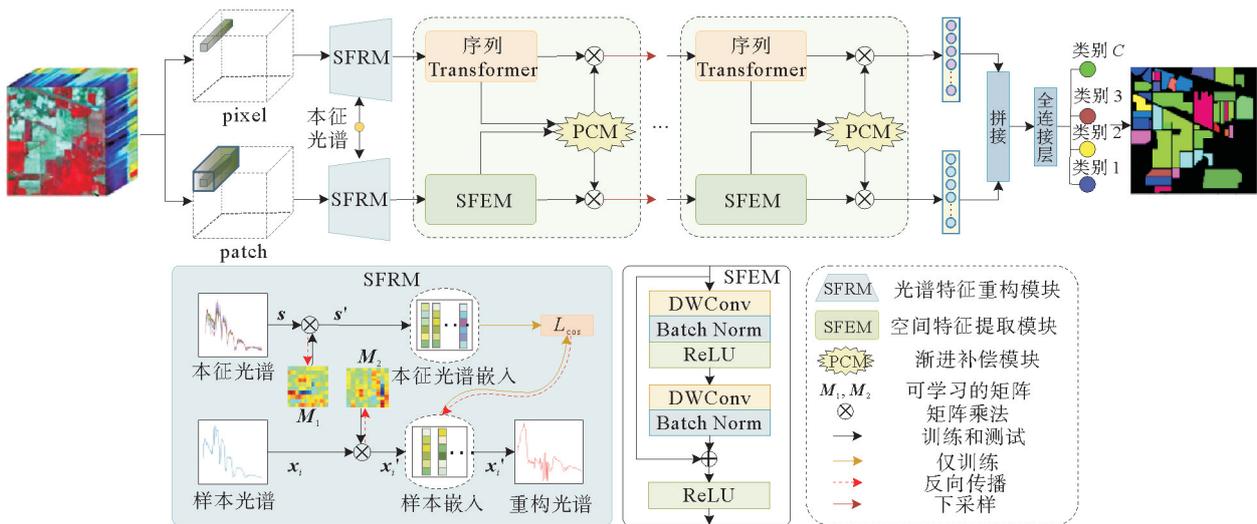


图 1 ISANet 总体结构

Fig. 1 Overall structure of ISANet

方法 1 本征光谱感知的高光谱图像分类方法 (ISANet)

输入: HSI image $X \in \mathbf{R}^{H \times W \times B}$, label y

输出: Predicted class label vector \hat{y} .

- 1: Extract pixel \mathbf{x}_i and neighborhood X_{patch} from X .
 - 2: Initialization: batchsize=64, optimizer Adam (learning rate=1e-3), epochs number $\epsilon=300$.
 - 3: **for** $e=1$ to 300 **do**
 - 4: Calculate the intrinsic spectra $\mathbf{s} \in \mathbf{R}^{C \times B}$ by Eq. (5)-(7).
 - 5: Obtain \mathbf{s}', \mathbf{x}' by Eq. (8),(9).
 - 6: Obtain the reconstructed spectra \mathbf{x}'_i by Eq. (10).
 - 7: Extract the global spectral features \mathbf{f}_{spe} by Sequence Transformer.
 - 8: Extract the local spatial features \mathbf{f}_{spa} by SFEM.
 - 9: Fuse $\mathbf{f}_{\text{spe}}, \mathbf{f}_{\text{spa}}$ to obtain the compensated features $\mathbf{f}'_{\text{spe}}, \mathbf{f}'_{\text{spa}}$.
 - 10: $\mathbf{g} = \text{Concat}(\mathbf{f}'_{\text{spe}}, \mathbf{f}'_{\text{spa}})$, $\text{Concat}(\cdot)$ is Concatenation.
 - 11: $\hat{y} = \text{Softmax}(\text{FC}(\mathbf{g}))$.
 - 12: **end for**
 - 13: Output $\hat{y} = \{\hat{y}_i | i=1, 2, \dots, M\}$, M is the number of testing samples.
-

2.1 光谱特征重构模块

高光谱图像中不同地物类别的光谱信息在原始光谱空间中经常呈现出局部相似性,也就是所说的“同谱异物”,给 HSIC 带来困难。SFRM 的主要思想是先通过可学习的投影矩阵构建一个新的光谱投影空间,然后将地物光谱投影到该空间,通过约束投影后的本征光谱与样本光谱的相似度,指导投影矩阵学习,使投影后的重构光谱特征更具可分性。SFRM 的结构如图 1 左下角模块图所示,其实现是基于本征光谱的概念。

具体来说,特定地物类别的本征光谱是通过主成分分析方法获取到的。设 $E \in \mathbf{R}^{B \times N}$ 为包含 N 个训练样本的集合,根据样本的类别将其分解为 C 个子集,表示为 $E = \{E_1, E_2, \dots, E_C\}$ 。 $E_i \in \mathbf{R}^{B \times N_i}$ 为包含 N_i 个样本的第 i 个类别的样本集,表示为 $E_i = \{\mathbf{e}_1^i, \mathbf{e}_2^i, \dots, \mathbf{e}_{N_i}^i\}$, $\mathbf{e}_j^i \in \mathbf{R}^{B \times 1}$ 为第 i 类的第 j 个光谱向量, $0 \leq j < N_i, 0 \leq i < C$ 。计算 E_i 的光谱均值:

$$\mathbf{p}_j^i = \frac{\mathbf{e}_j^i - \text{mean}(E_i)}{\text{std}(E_i)}. \quad (5)$$

式中: $\text{mean}(\cdot)$ 和 $\text{std}(\cdot)$ 分别为均值和标准差函数。将其排列为矩阵 $\mathbf{P}_i = [\mathbf{p}_1^i, \mathbf{p}_2^i, \dots, \mathbf{p}_{N_i}^i]^T \in \mathbf{R}^{N_i \times B}$, 并对每个类别的样本集分别计算协方差矩阵 \mathbf{A}_i 及其特征值和特征向量:

$$\mathbf{A}_i = \frac{1}{B} \mathbf{P}_i \mathbf{P}_i^T \in \mathbf{R}^{N_i \times N_i}, \quad (6)$$

$$\mathbf{A}_i = \mathbf{Q}_i \mathbf{\Lambda} \mathbf{Q}_i^T. \quad (7)$$

式中: \mathbf{A}_i 为第 i 类的协方差矩阵,捕捉了不同光谱之间的关系; $\mathbf{\Lambda}$ 为包含特征值的对角矩阵; \mathbf{Q}_i 为包含相应特征向量的第 i 类的标准正交矩阵,从中选择前 m 个主成分并重新升维到 B 维,得到第 i 类地物的本征光谱 $\mathbf{s}_i \in \mathbf{R}^{1 \times B}$, 则所有类别地物的本征光谱为 $\mathbf{s} \in \mathbf{R}^{C \times B}$ 。

通过可学习投影矩阵 \mathbf{M}_1 和 \mathbf{M}_2 , 将本征光谱和样本光谱分别映射成本征光谱嵌入 \mathbf{s}' 和样本光谱嵌入 \mathbf{x}' , 有:

$$\mathbf{s}' = \mathbf{s} \otimes \mathbf{M}_1, \quad (8)$$

$$\mathbf{x}' = \mathbf{x} \otimes \mathbf{M}_2. \quad (9)$$

式中: \otimes 为矩阵乘法。这一过程实现了样本光谱信息的线性变换; 通过本征光谱嵌入和样本光谱嵌入的相似度约束, 指导 ISANet 逐步调整 \mathbf{M}_1 和 \mathbf{M}_2 , 使相似光谱在新的光谱投影空间中更容易被区分。本研究选用余弦相似损失函数量化本征光谱嵌入和样本光谱嵌入之间的相似程度, 设计损失函数:

$$L_{\cos}(\mathbf{x}', \mathbf{s}') = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^C \mathbf{x}'_i \cdot \mathbf{s}'_j}{\sqrt{\sum_{i=1}^N \mathbf{x}'_i{}^2} \cdot \sqrt{\sum_{j=1}^C \mathbf{s}'_j{}^2}} \quad (10)$$

式中: \mathbf{s}'_j 为第 j 类的本征光谱嵌入。

2.2 序列 Transformer 模块

Transformer 模块具有捕获序列依赖关系和提取全局特征的能力,在多种序列相关任务中展现了强大的数据处理能力。将 SFRM 输出的样本重构光谱 \mathbf{x}'_i 的每个波段 $\mathbf{x}'_{ij} = [\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iB}] \in \mathbf{R}^{1 \times 1 \times B}$ 作为序列 Transformer(图 2(a))的输入单元,将其变换成一组序列向量,输入到序列 Transformer 编码器模块(图 2(b)),提取波段的全局特征表示。多头注意力(multi-head attention, MSA)是序列 Transformer 编码器模块的核心组件,由多个 SA 模块堆叠构成,MSA 结构如图 2(d)所示。自注意力机制(self-attention, SA)通过改变注意力权重来灵活地关注序列中不同部分的信息,进而增强对细微光谱特征的识别与分类能力,结构如图 2(e)所示。

传统的前馈神经网络使用两个线性投影层以及层之间的非线性激活函数来处理 MSA 层的输出特征,但是无法注意更具辨别力和有利于去噪的光谱特征。因此,本研究设计了图 2(c)所示的 FFNSA,聚焦对于分类任务更为重要的波段,同时利用序列注意力的权重来识别和抑制噪声。注意力权重可以被解释为对每个波段的重要性评分,模型通过学习这些权重,可以降低噪声波段的影响,聚焦光谱数据中的关键特征,从而提高光谱分支的特征学习能力。可以描述为:

$$\mathbf{G}_{\text{out}} = \sigma(\mathbf{W}_1 \mathbf{X}_{\text{out}}), \quad (11)$$

$$\mathbf{O} = \mathbf{W}_2 \mathbf{G}_{\text{out}} + (\delta(\mathbf{G}_{\text{out}})) \mathbf{G}_{\text{out}}. \quad (12)$$

式中: δ 为 Sigmoid 函数, σ 为 GELU 激活函数, \mathbf{X}_{out} 为 FFNSA 的输入特征, \mathbf{G}_{out} 和 \mathbf{O} 分别表示 GELU 激活函数和 FFNSA 的输出, \mathbf{W}_1 和 \mathbf{W}_2 分别为第一个全连接层和第二个全连接层的参数。

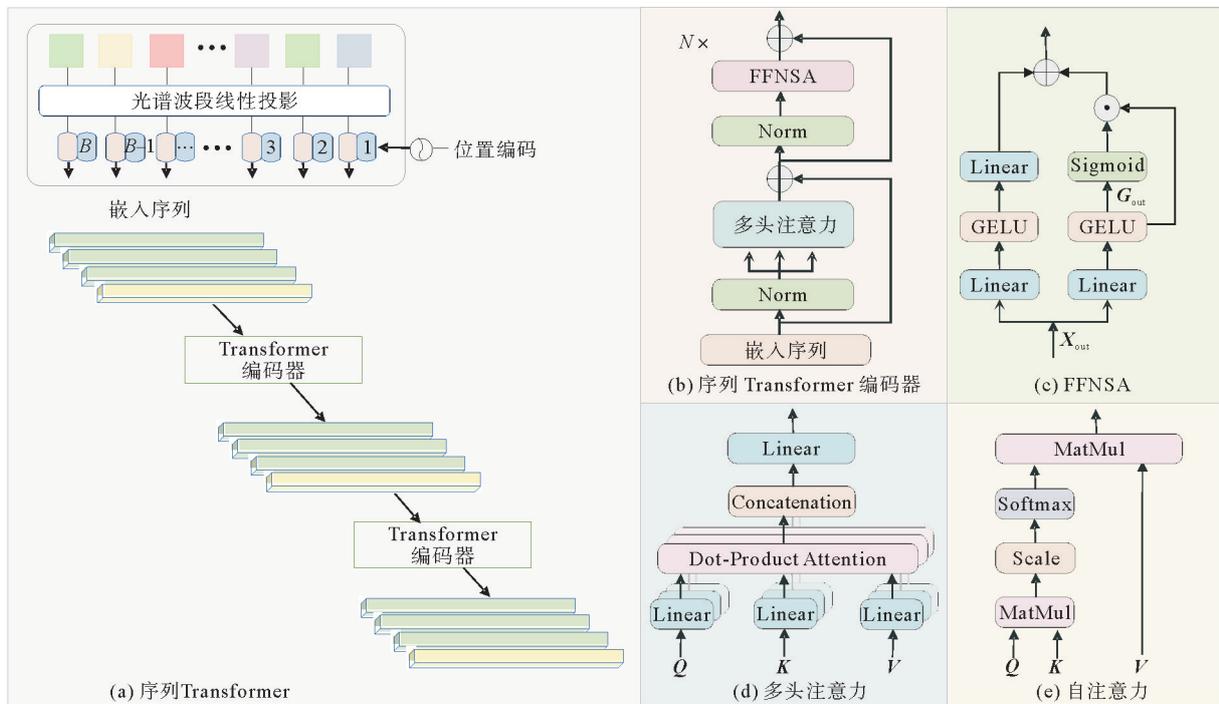


图 2 序列 Transformer 模块结构

Fig. 2 Structure of sequence Transformer module

2.3 空间特征提取模块

空间特征提取模块旨在提取重要的空间特征并减少 ISANet 的参数数量,同时通过跳跃连接促进特征信息的传播。SFEM 从输入特征图中提取具有判别性的空间特征,同时通过引入 DWConv,实现方法轻量化和减少参数数量。SFEM 主要使用深度卷积、批量归一化(batch norm, BN)和 ReLU 激活函数作为基本构建块,其结构如图 1 中 SFEM 模块图所示。

2.4 渐进补偿模块

现有的高光谱空谱联合提取方法中,光谱和空间信息在特征提取阶段被独立处理,仅在最后分类阶段融合特征,会导致重要信息丢失。为了充分利用光谱和空间信息之间的互补性,本研究采用渐进补偿模块,实现空间信息向光谱域的早期集成和光谱信息向空间域的预先整合,光谱域和空间域不再被简单划分,从而促进光谱和空间特征在处理流程中的深度融合。

PCM 结构如图 3 所示。 f'_{spe} 和 f'_{spa} 分别是由光谱特征 f_{spe} 和空间特征 f_{spa} 经过卷积、BN、激活函数得到的光谱特征与空间特征。为了融合光谱特征 f'_{spe} 和空间特征 f'_{spa} ,对获得的空间特征 f'_{spa} 进行池化操作,生成全局联合表示;通过激活函数聚焦于重要的光谱波段和空间位置,突出有用的特征并抑制不重要的特征。

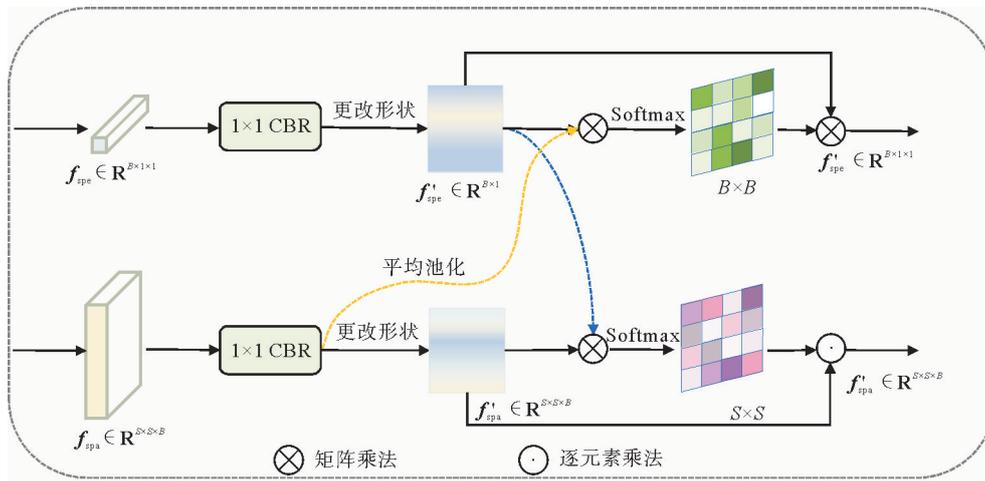


图 3 渐进补偿模块结构

Fig. 3 Structure of progressive compensation module

3 实验结果与分析

3.1 数据集介绍

1) Indian Pines (IP)数据集:使用机载可见光/红外成像光谱仪(airborne visible/infrared imaging spectrometer, AVIRIS)捕获的美国印第安纳州印度松树农作物种植地的数据,具有 145×145 像素和 200 个光谱波段,空间分辨率为 20 m,假彩色图像和真值图如图 4(a)所示。

2) Salinas Valley (SA)数据集:由机载可见光/红外成像光谱仪拍摄的图像,具有 512×217 像素和 204 个光谱波段,空间分辨率为 3.7 m,假彩色图像和真值图如图 4(b)所示。

3) WHU-Hi-LongKou (WHU-LK)数据集:2018 年在中国湖北省龙口镇采集得到,具有 550×400 像素和 270 个光谱波段,空间分辨率约为 0.463 m,图像立方体和真值图如图 4(c)所示。

3.2 实验准备

为了与其他方法公平对比,设置 patch 大小为 11×11 ,minibatch 为 64,epoch 为 300,使用 Adam 优化器,初始学习率为 0.001;阶段数设置为 3,序列 Transformer 部分使用 {256, 128, 64} 作为 3 阶段的嵌入维数,MSA 不同阶段头数为 {8, 4, 2}。

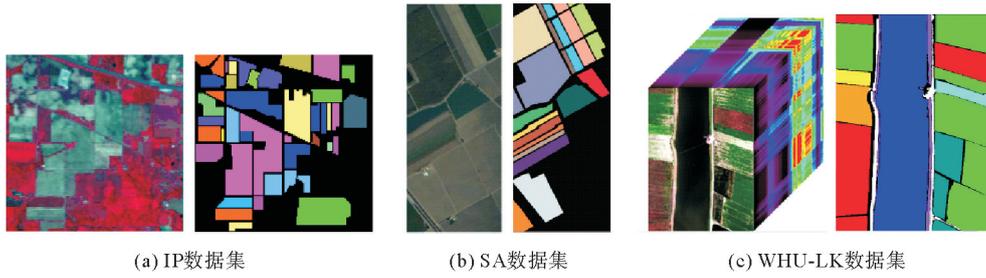


图 4 高光谱图像数据集

Fig. 4 Hyperspectral image datasets

3.3 对比实验结果评估

为了评估所提出的 ISANet 的分类性能,分别与基于 CNN 的方法 2D-CNN^[9]、3D-CNN^[10]和 SSRN^[12],基于双分支的方法 DBDA^[18],基于 Transformer 的方法 SpectralFormer^[14]和 SSFTT^[15],基于 CNN 和 Transformer 的 HCVN^[17]方法进行对比实验。

不同对比方法在 3 个数据集上的定量结果如表 2 所示。选用总体分类精度(overall accuracy, OA)、平均分类精度(average accuracy, AA)和 Kappa 系数三个评估指标衡量分类结果的性能。在 IP 数据集上,ISANet 获得了最高的分类准确率 97.68%,这归功于 SFRM 能够有效地提取和增强光谱特征,同时通过渐进的信息融合策略,实现了两种类型信息的有效整合,从而显著提高了分类精度。SpectralFormer 仅使用 Transformer 构建方法,其 OA 值为 94.72%,分类精度不理想,次于 SSFTT。在 SA 数据集上,本研究方法比 SSFTT 和 HCVN 分别提高 3.55%和 1.28%。在 WHU-LK 数据集上,ISANet 的总体准确率达到 99.32%,与 3D-CNN、SSRN、DBDA 和 SpectralFormer 相比,总体分类精度分别高了 2.71%、2.09%、1.74%和 2.26%,比 HCVN 方法总体精度高了 1.07%。

表 2 不同方法在三个数据集上的定量结果

Table 2 Quantitative results of different methods on the three datasets

数据集	评估指标	2D-CNN	3D-CNN	SSRN	DBDA	SpectralFormer	SSFTT	HCVN	ISANet
IP	OA/%	85.22	87.59	91.81	93.94	94.72	95.54	96.89	97.68
	AA/%	77.91	79.21	87.14	88.39	93.14	92.38	95.59	96.02
	Kappa×100	76.86	79.57	89.65	92.22	92.54	93.22	95.37	96.84
SA	OA/%	90.7	92.49	93.11	93.39	93.47	94.37	96.64	97.92
	AA/%	93.05	94.51	91.76	94.92	93.48	95.42	96.19	98.24
	Kappa×100	89.95	90.86	91.28	92.42	91.51	93.15	95.15	97.33
WHU-LK	OA/%	96.18	96.61	97.23	97.58	97.06	97.66	98.25	99.32
	AA/%	93.35	88.51	93.45	93.88	92.32	94.36	96.72	97.26
	Kappa×100	96.72	96.13	97.06	97.13	96.50	96.81	98.59	99.22

不同方法在 3 个数据集上的定性分类结果如图 5~7 所示。可以看出,在 IP 数据集上,ISANet 方法得到的分类图更接近真值图,2D-CNN 和 3D-CNN 方法噪声多,对边缘类别的分类效果较差。而 HCVN 方法受益于融合 CNN 和 Transformer 特征,效果优于 SpectralFormer。在 SA 数据集上,Vinyard-untrained 和 Grapes-untrained 是两个易于观察的类别,本研究方法具有最好的分类效果,其次是 HCVN。DBDA 方法由于双分支结构提取空谱特征,错分现象相对较少。在 WHU-LK 数据集上,ISANet 具有最好的边缘效果,2D-CNN、3D-CNN 和 SSRN 方法在小尺寸区域存在大量噪声,分类效果差。

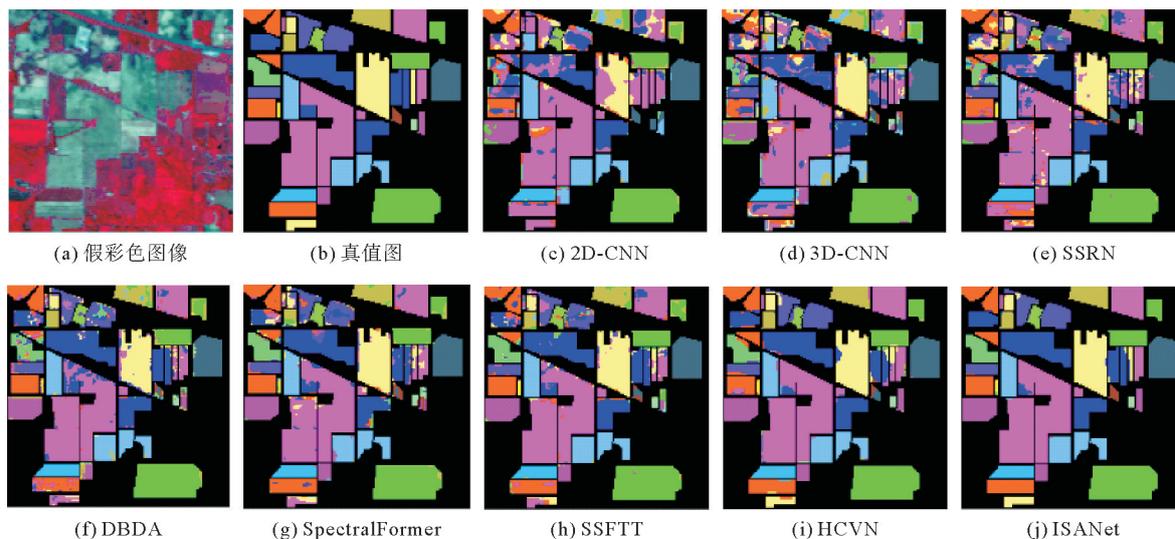


图 5 不同方法在 Indian Pines 数据集上的可视化结果

Fig. 5 Visualization of different methods on the Indian Pines dataset

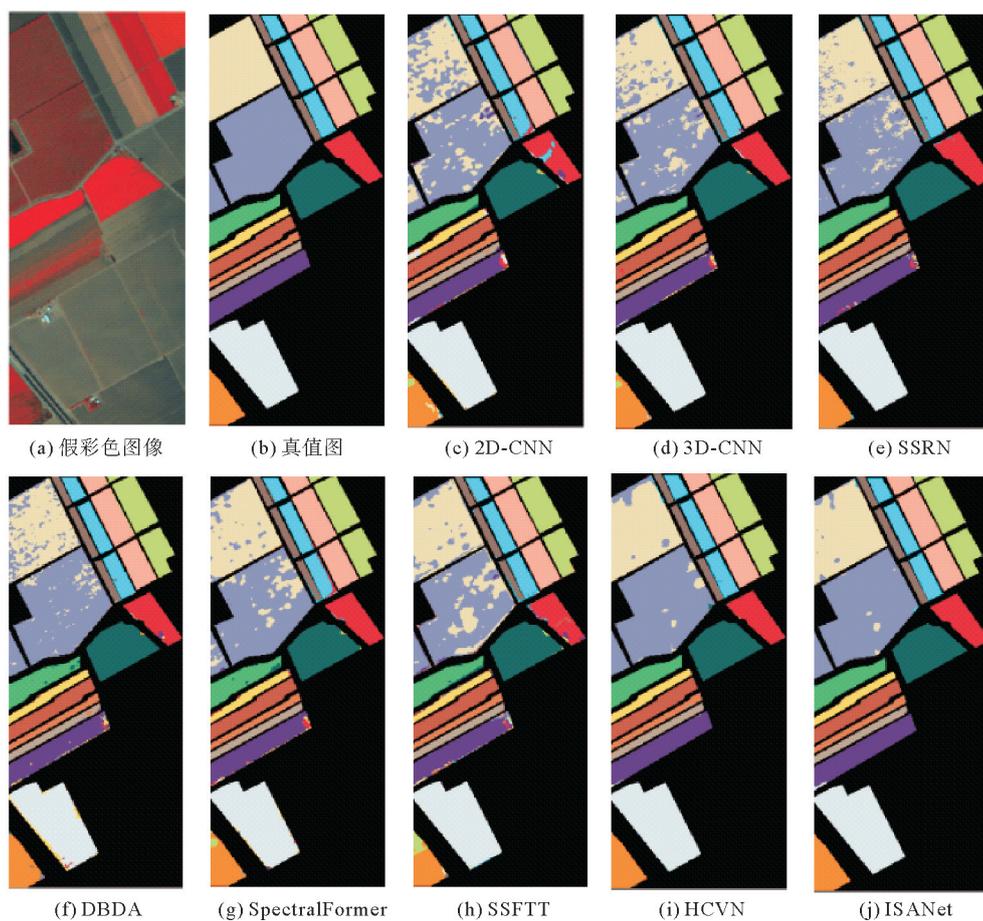


图 6 不同方法在 Salinas Valley 数据集上的可视化结果

Fig. 6 Visualization of different methods on the Salinas Valley dataset

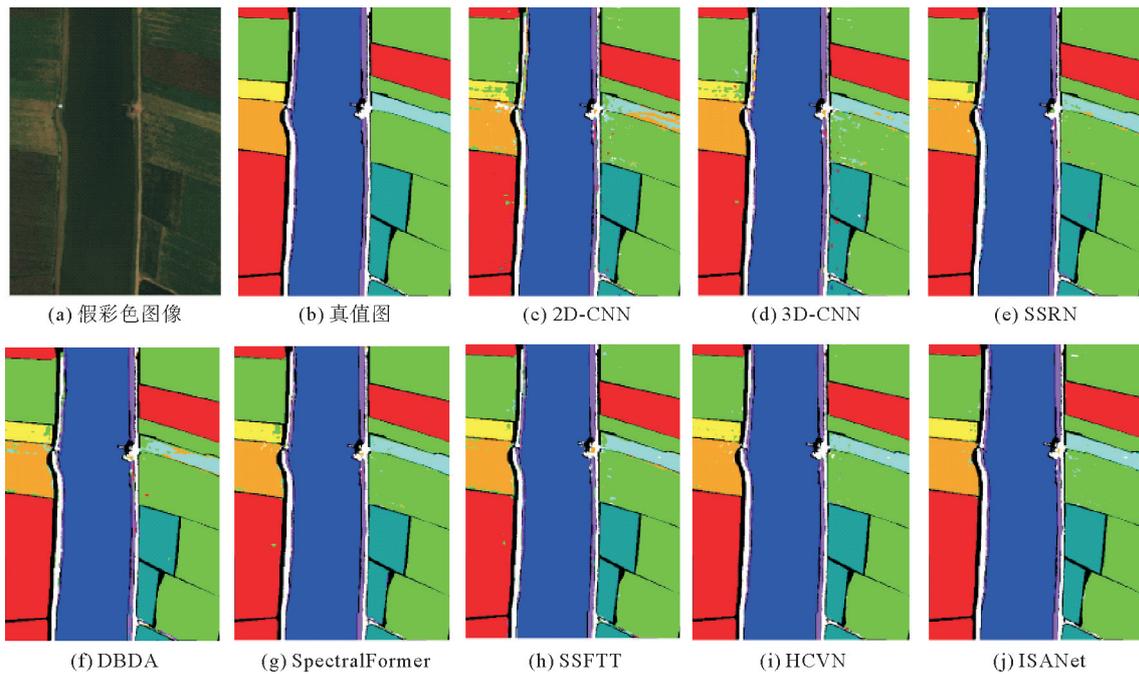


图 7 不同方法在 WHU-Hi-LongKou 数据集上的可视化结果

Fig. 7 Visualization of different methods on the WHU-Hi-LongKou dataset

3.4 SFRM 模块的有效性分析

为了直观展示光谱特征重构模块的功能,选取 IP 数据集中“Corn-notill”类别和“Corn-mintill”类别的样本光谱曲线进行分析,如图8所示。图8(a)和8(d)分别呈现了数据集中“Corn-notill”类别和“Corn-

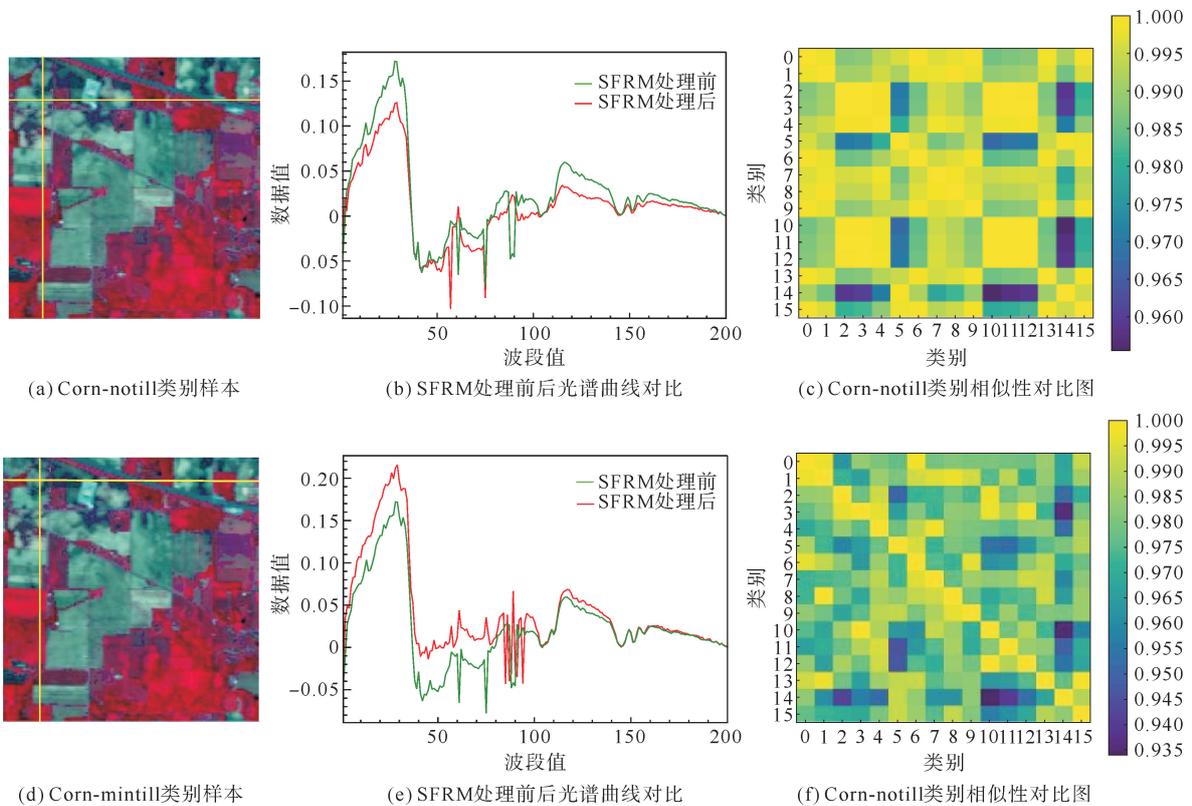


图 8 不同类别光谱曲线的差异对比图与类别相似性对比图

Fig. 8 Comparison of differences in spectral curves of different categories and category similarity comparison map

mintill”类别的随机采样样本,可以看出这两个类别具有相似的光谱曲线,同时每个类别的曲线都有显著波动,如图 8(b)和图 8(e)所示,绿色光谱曲线代表 SFRM 模块处理前的数据,其中类别难以区分;而红色曲线代表经过 SFRM 模块处理后的数据,分类明显。图 8(c)和图 8(f)进一步对比了本征光谱映射到新空间前后的类别相似度,可以看出本研究提出的 SFRM 方法有效提高了地物类别间的区分能力,尤其是在处理具有高度相似性的类别时。

3.5 消融实验

为了进一步评估 ISANet 中关键模块的有效性,设计了消融实验,通过逐步移除各个模块来单独考察每个模块对整体分类性能的影响。在 3 个数据集上的实验结果如表 3 所示。当 ISANet 移除 SFRM 模块时,3 个数据集上的分类性能分别下降了 2.73%、3.08% 和 1.67%,表明

SFRM 在去除光谱冗余信息方面发挥了关键作用,不仅有效降低了光谱维数,而且显著提高了分类性能。移除 PCM 模块后,3 个数据集上的分类性能分别下降了 1.19%、1.15%、0.98%,证实渐进补偿模块在整合光谱和空间信息方面的重要性。

3.6 参数量、训练时间和推理时间分析

时间成本分析是评估计算效率的基本方法。为了评估所提方法的运行效率,在 3 个数据集上与现有对比方法的参数量、训练时间和推理时间进行实验对比,结果如表 4 所示。与基于 CNN 的方法相比,DBDA 与 HCVN 因其结构包含多种类型的特征提取方法,参数量较大。SpectralFormer 方法需要许多重复的自注意力模块学习丰富的特征表示,导致训练时间与推理时间相对较长,每次迭代消耗的计算周期也较大。除 2D-CNN 外,相较其他对比方法,SSFTT 的参数量和运行时间较少,原因在于 SSFTT 使用 3D-CNN 提取特征。ISANet 方法所需的训练时间与推理时间略长,原因是其融合了 CNN 和 Transformer 的双分支特征,尽管存在时间折损,但 ISANet 方法运行效率相对较好且分类精度更高。

表 4 不同数据集集中的参数量、训练时间与推理时间

Table 4 Number of parameters, training time and inference time on different datasets

方法	参数量/M	IP		SA		WHU-LK	
		训练时间/s	推理时间/s	训练时间/s	推理时间/s	训练时间/s	推理时间/s
2D-CNN	0.49	31.02	3.32	131.62	9.08	40.39	7.81
3D-CNN	1.46	256.58	6.53	1 889.25	54.57	1 265.24	9.02
SSRN	0.39	261.29	7.65	782.11	15.48	394.12	16.75
DBDA	14.05	134.95	6.89	558.89	46.15	349.61	34.62
SpectralFormer	15.00	310.29	6.76	1 219.34	20.13	278.54	43.25
SSFTT	9.48	15.84	5.61	221.03	10.11	13.78	15.68
HCVN	12.34	73.54	7.62	378.14	10.81	294.15	14.44
ISANet	10.59	20.51	6.45	235.81	12.35	11.83	17.16

3.7 空间输入 patch 大小

空间输入 patch 大小决定了在 HSI 中对单个像素进行分类时使用多少局部空间信息。因此,HSI 中的空间 patch 大小通常对分类的性能有很大的影响。一般情况下,patch 是以标记像素为中心的正方形空间邻域信息,其大小可根据半径来定义。在本研究中,考虑设置 patch 大小为[7, 9, 11, 13, 15],数值越

大意味着涉及到的空间邻域信息越多。如图9所示,可以看出,随着空间邻域大小的增加,分类性能呈现出先提高后降低的趋势。这种趋势表明,存在一个最优的 patch 大小,能够在提供足够上下文信息的同时,最小化噪声和冗余的影响。通过对3个不同数据集的比较分析,发现 11×11 的 patch 大小在OA评估指标上达到最高。因此, 11×11 的 patch 大小是实现高光谱图像分类任务的一个恰当选择。

4 结论

本研究提出一种本征光谱感知的高光谱图像分类方法(ISANet),该方法使用本征光谱感知的光谱特征重构模块,将地物光谱投影到新的光谱空间以增强地物类别可分性。ISANet融合序列Transformer模块和空间特征提取模块两种特征提取技术捕获全局和局部高光谱特征,采用渐进补偿模块使ISANet能够在早期阶段融合利用光谱和空间特征,逐步增强ISANet的特征提取能力,显著提高高光谱图像的分类精度。在3个公开高光谱数据集上的实验结果验证了ISANet的有效性,在分类图的定量指标和可视化评价方面都表现出优越性能。

未来将继续探索基于Transformer的HSI分类方法,提取更有代表性的空间光谱特征,并探索多模态数据的融合,以实现更高的分类性能,满足不断变化的应用需求。

参考文献:

- [1] 郑宗生,刘贝,卢鹏,等.基于AlexNet的近岸水生植物光谱分类及特征光谱分析[J].中国激光,2023,50(2):129-138.
ZHENG Zongsheng,LIU Bei,LU Peng,et al. Spectral classification and characteristic spectral analysis of nearshore aquatic plants based on AlexNet[J]. Chinese Journal of Lasers,2023,50(2):129-138.
- [2] 毛英伍,郭颖,张王菲,等.联合LiDAR,高光谱数据及3D-CNN方法的树种分类[J].林业科学,2023,59(3):73-83.
MAO Yingwu,GUO Ying,ZHANG Wangfei,et al. Tree species classification by combining LiDAR,hyperspectral data and 3D-CNN method[J]. Scientia Silvae Sinicae,2023,59(3):73-83.
- [3] BOUGUETTAYA A,ZARZOUR H,KECHIDA A,et al. Deep learning techniques to classify agricultural crops through UAV imagery:A review[J]. Neural Computing and Applications,2022,34(12):9511-9536.
- [4] MELGANI F,BRUZZONE L. Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Transactions on Geoscience and Remote Sensing,2004,42(8):1778-1790.
- [5] HAM J,CHEN Y C,CRAWFORD M M,et al. Investigation of the random forest framework for classification of hyperspectral data[J]. IEEE Transactions on Geoscience and Remote Sensing,2005,43(3):492-501.
- [6] RODARMEL C,SHAN J. Principal component analysis for hyperspectral image classification[J]. Surveying and Land Information Science,2002,62(2):115-122.
- [7] HAN W,FENG R Y,WANG L Z,et al. A semi-supervised generative framework with deep learning features for high-resolution remote sensing image scene classification[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2018,145:23-43.
- [8] HU W,HUANG Y Y,LI W,et al. Deep convolutional neural networks for hyperspectral image classification[J/OL]. Journal of Sensors,2015. DOI:10.1155/2015/258619.
- [9] FANG L Y,LIU Z L,SONG W W. Deep hashing neural networks for hyperspectral image feature extraction[J]. IEEE Geoscience and Remote Sensing Letters,2019,16(9):1412-1416.
- [10] LI W M,CHEN H H,LIU Q,et al. Attention mechanism and depthwise separable convolution aided 3DCNN for hyperspectral remote sensing image classification[J/OL]. Remote Sensing,2022,14(9). DOI:10.3390/rs14092215.
- [11] VAHDAT A,KAUTZ J. NVAE:A deep hierarchical variational autoencoder[J]. Advances in Neural Information Pro-

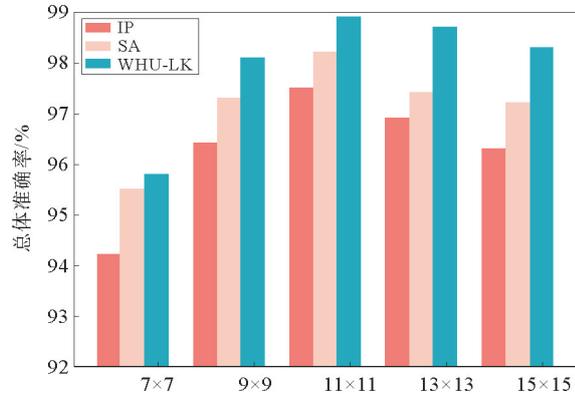


图9 分类性能与不同输入图像 patch 大小的关系在3个数据集上的分类性能

Fig. 9 Classification performance versus different input image patch sizes on three datasets

- cessing Systems, 2020, 33:19667-19679.
- [12] ZHONG Z L, LI J, LUO Z M, et al. Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(2):847-858.
- [13] ZHU M H, JIAO L C, LIU F, et al. Residual spectral-spatial attention network for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(1):449-462.
- [14] HONG D F, HAN Z, YAO J, et al. SpectralFormer: Rethinking hyperspectral image classification with transformers[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60. DOI:10.1109/TGRS.2021.3130716.
- [15] SUN L, ZHAO G R, ZHENG Y H, et al. Spectral-spatial feature tokenization transformer for hyperspectral image classification[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60. DOI: 10.1109/TGRS.2022.3144158.
- [16] MEI S H, SONG C, MA M Y, et al. Hyperspectral image classification using group-aware hierarchical transformer[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60. DOI:10.1109/TGRS.2022.3207933.
- [17] YAN H P, ZHANG E L, WANG J, et al. Hybrid Conv-ViT network for hyperspectral image classification[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2023, 20. DOI:10.1109/LGRS.2023.3287277.
- [18] LI R, ZHENG S Y, DUAN C X, et al. Classification of hyperspectral image based on double-branch dual-attention mechanism network[J/OL]. Remote Sensing, 2020, 12(3). DOI:10.3390/rs12030582.

(责任编辑:傅 游)