

基于混合光谱增强与多尺度空间聚合的高光谱图像分类方法

欧阳宁^{1,2}, 黄辰钰², 林乐平^{1,2}

(1. 桂林电子科技大学 认知无线电与信息处理省部共建教育部重点实验室, 广西 桂林 541004; 2. 桂林电子科技大学 信息与通信学院, 广西 桂林 541004)

摘要: 由于高光谱图像存在同物异谱和异物同谱现象, 仅依赖光谱信息无法充分表征高光谱图像的特征, 因此可引入空间信息以更准确地捕捉物体特征。为此, 本文提出一种基于混合光谱增强与多尺度空间聚合的高光谱图像分类方法。该方法设计了混合光谱增强模块, 利用小波变换构建光谱的多尺度局部特征, 通过 Transformer 架构生成光谱的全局特征, 以增强光谱特征的类内一致性。同时, 设计了多尺度空间聚合模块, 用于提取空间特征固有的多尺度信息, 并建立不同尺度间的交互关系, 以生成更具鲁棒性的土地覆盖表示, 从而进一步提升分类性能。实验结果表明: 本文方法相较于其他先进网络表现出显著的优越性, 能有效获取更丰富的光谱信息和空间特征表示。

关键词: 高光谱图像分类; 混合光谱增强模块; 小波变换; 多尺度空间聚合模块

中图分类号: TP753 **文献标志码:** A **文章编号:** 1671-5497(2025)11-3727-09

DOI: 10.13229/j.cnki.jdxbgxb.20240252

Hyperspectral image classification based on hybrid spectral enhancement and multi-scale spatial aggregation

OUYANG Ning^{1,2}, HUANG Chen-yu², LIN Le-ping^{1,2}

(1. Key Laboratory of Cognitive Radio and Information Processing of Ministry of Education, Guilin University of Electronic Technology, Guilin 541004, China; 2. School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: Due to the presence of the same object with different spectrum and the same spectrum with different objects in hyperspectral images, spectral information alone cannot fully reflect the features of hyperspectral images, and spatial information can be introduced to capture the features of objects more accurately. Therefore, a hyperspectral image classification method based on hybrid spectral enhancement and multi-scale spatial aggregation is proposed in this paper. In this method, a hybrid spectral enhancement

收稿日期: 2024-03-12.

基金项目: 国家自然科学基金项目(62001133); 广西科技基地和人才专项项目(桂科 AD19110060); 广西自然科学基金项目(2017GXNSFBA198212); 广西无线宽带通信与信号处理重点实验室基金项目(GXKL06200114).

作者简介: 欧阳宁(1972-), 男, 教授, 硕士. 研究方向: 数字图像处理, 智能信息处理.

E-mail: ouyangning@guet.edu.cn

通信作者: 林乐平(1980-), 女, 教授, 博士. 研究方向: 机器学习, 图像信号处理. E-mail: linleping@guet.edu.cn

module is designed, multi-scale local features of the spectrum are constructed using wavelet transform, and global features of the spectrum are generated by Transformer architecture, so as to enhance the intra-class consistency of the spectral features. At the same time, a multi-scale spatial aggregation module is designed to extract the inherent multi-scale information of spatial features and establish the interaction between different scales, so as to generate a more robust land cover representation, thereby further improve the classification performance. The experimental results show that the proposed method is superior to other advanced networks, indicating that the method can effectively obtain more abundant spectral information and spatial feature representation.

Key words: hyperspectral image classification; hybrid spectral enhancement module; wavelet transform; multi-scale spatial aggregation module

0 引言

高光谱图像以数据立方体的形式呈现,是一种包含丰富光谱和空间信息的图像,拥有数百个相邻的光谱波段。与常规的 RGB 图像相比,高光谱图像有更精细的光谱信息,在分类任务中具有很大的潜力。但光谱波段的高维度、冗余性和标记样本数量的有限性,也给高光谱图像分类带来了挑战。特别是在土地覆盖复杂时,类内光谱一致性较弱,原始数据空间难以对其进行有效表达。

随着深度学习的发展,卷积神经网络(Convolutional neural network, CNN)凭借其对光谱空间特征的局部建模能力,在高光谱图像分类领域受到广泛关注。基于 CNN 的方法^[1,2]通过叠加卷积核的方式,对光谱空间特征进行从浅层到深层的特征提取。在多层次空-谱融合网络^[3]中,通过结合不同层次的空间与光谱信息,能有效地获取高光谱图像的精细特征。在空谱分组卷积密集网络^[4]中,采用分组卷积重复利用高分辨率信息,可防止细节信息丢失。Hamida 等^[5]提出了一种新的三维深度学习方法,实现了光谱和空间信息的联合处理。因此,基于深度学习的分类网络对高光谱图像的空间和光谱信息有较好的表征能力,能有效提取更具判别性的分类特征,获得更优越的分类结果。

基于 CNN 的方法虽然具有强大的局部特征捕捉能力,但其对全局信息的关注仍显不足。Dosovitskiy 等^[6]提出视觉 Transformer (Vision transformer, ViT)网络,将 Transformer 的特性引入自然图像处理中,此后越来越多的学者开始将 ViT 应用于高光谱图像分类领域。Hong 等^[7]提出 SpectralFormer 网络,从分组的相邻波段中学

习光谱信息作为 ViT 的输入,并构建跨层编码输出的相互连接,深度挖掘和表示光谱特征的序列属性;Sun 等^[8]设计了 Transformer 框架,对浅层的空-谱特征进行特征变换,并输入 Transformer 构建特征表示。总的来说,ViT 具有全局感受野,可获得像素间的远距离依赖关系,能为所有像素分类提供更精准的注意力权重,从而避免遗漏重要信息。

为充分发挥 CNN 的局部接受域优势和 Transformer 的全局特征提取优势,本文将 CNN 与 ViT 相结合,提出了基于混合光谱增强与多尺度空间聚合的高光谱图像分类方法(Hyperspectral image classification based on hybrid spectral enhancement and multi-scale spatial aggregation, HEMA)。根据高光谱图像数据特点,本文借鉴小波变换思想和 ViT 网络特性,构建了混合光谱增强模块(Hybrid spectral enhancement module, HSEM),通过配置模块参数,在保留空间信息的前提下有效增强了光谱的特征信息。同时,在此基础上构建了多尺度空间聚合模块(Multi-scale spatial aggregation module, MSAM),通过依次相加的聚合构造函数,增强了不同尺度特征映射间的交互性,提升了网络的特征表达能力。实验结果表明,相较于一些前沿的深度网络方法,本文方法可有效提升高光谱图像分类网络的性能,进而得到较高的分类精度。

1 相关研究

原始高光谱图像数据光谱维度较高,当数据非常复杂时,相对较差的类内光谱一致性对分类结果有很大影响。因此,降低光谱变化带来的影响是高光谱图像分类的关键问题之一。正交变换

方法,如主成分分析(Principal component analysis, PCA),可通过正交变换识别光谱波段间的相关性,进而提取其内在特征^[9]。虽然正交变换方法在提取主成分方面具有一定的优势,但在捕捉非线性特征方面相对较弱,特别是在处理高度变化的光谱信息时灵活性不足。小波变换(Discrete wavelet transform, DWT)^[10]是一种多层次分解信号的方法,能揭示信号在不同时间尺度上的频率成分。小波变换凭借其在频率域分布规律处理和尺度分析方面的优势脱颖而出,尤其在处理具有复杂光谱结构的场景时表现更为出色。Tang等^[11]使用小波变换对光谱数据进行处理,将其作为特征提取网络的输入,以减少异常光谱数据对分类的影响。Ahmad等^[12]利用小波变换进行可逆下采样以保持数据的完整性,并设计基于小波的CNN提取多尺度空间光谱特征。小波变换的多尺度分解特性可有效捕捉光谱信息的详细结构和关键特征。因此,本文引入小波变换对光谱数据进行多尺度分解,将光谱分解成各种近似分量和细节分量,获取分离频谱中更细节的特征。当类的频谱相似时,添加该小波特征可增强类间区分能力。

为了获取更多感受野,现如今很多具有多尺度光谱空间特征结构的CNN被应用于高光谱图像分类^[13-15]。Ghaderizadeh等^[16]通过构建多尺度残差模块获取更多的局部和全局特征;Wang等^[17]利用多尺度块考虑多范围上下文信息。受多尺度分析在提升模型适应性方面的启发,本文构建了多尺度空间聚合模块。与简单叠加感受野不同,该模块通过对多个感受野进行逐特征融合,充分整合了多尺度信息。这使得模型能更全面、准确地捕捉空间特征中的多样化信息,进而显著

提升对复杂图像的识别和分类能力。

本文将混合光谱增强模块和多尺度空间聚合模块相结合,提出了基于混合光谱增强与多尺度空间聚合的高光谱图像分类方法,有效提升了高光谱图像的分类性能和效率。

2 本文方法

2.1 混合光谱增强与多尺度空间聚合网络分类框架

本文所提HEMA的框架如图1所示。首先,该网络以高光谱图像像素 n 的图像邻域块作为输入,将其转换为二维向量 $P_n \in R^{(w \times w) \times 1 \times B}$,再使用混合光谱增强模块对光谱信息进行信息增强和全局光谱特征提取,得到混合增强特征 $S_n \in R^{(w \times w) \times 1 \times B}$,旨在降低类内光谱变化对分类的影响。由于将三维高光谱数据直接展平为二维矩阵会忽略高光谱图像原本的结构特性,为从光谱特征中恢复鲁棒性更强的空间特征,本文对二维特征状态下的 S_n 执行Reshape操作,将 S_n 重塑为原始空间尺寸的三维特征图 $I_n \in R^{w \times w \times B}$ 。其次,通过多尺度空间聚合模块获取包含多尺度空间信息的特征 $F_n \in R^{w \times w \times b}$,提高空间特征的表达能力。最后,使用多层感知机MLP对 F_n 进行分类,并采用交叉熵构建损失函数,用于度量预测标签与真实标签之间的差异。

2.2 混合光谱增强模块

本文设计的混合光谱增强模块结构如图2所示。光谱特征提取是构建本文网络的重要环节,会直接影响后续的分类结果。由于高光谱图像数据频带数量多,冗余性较强,若直接将光谱波段输入Transformer中,并不能充分利用其自注意力机制的优势。因此,本文将原始光谱信息与小波变

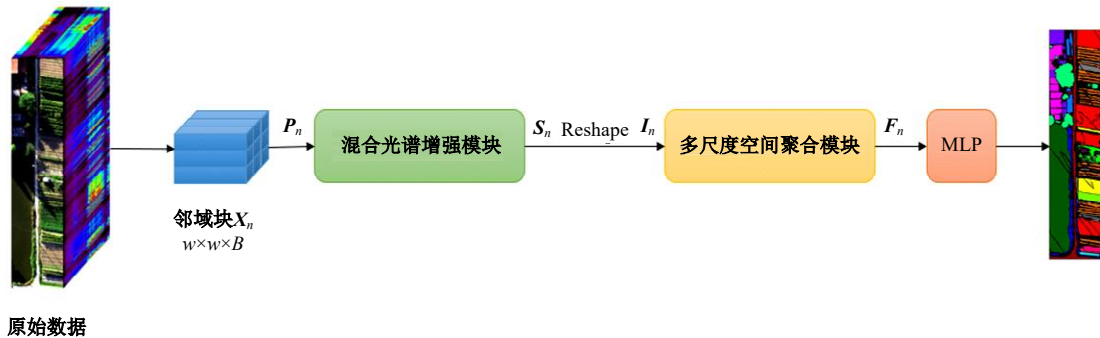


图1 基于混合光谱增强与多尺度空间聚合的高光谱图像分类网络结构

Fig. 1 Hyperspectral image classification network structure based on hybrid spectral enhancement and multi-scale spatial aggregation

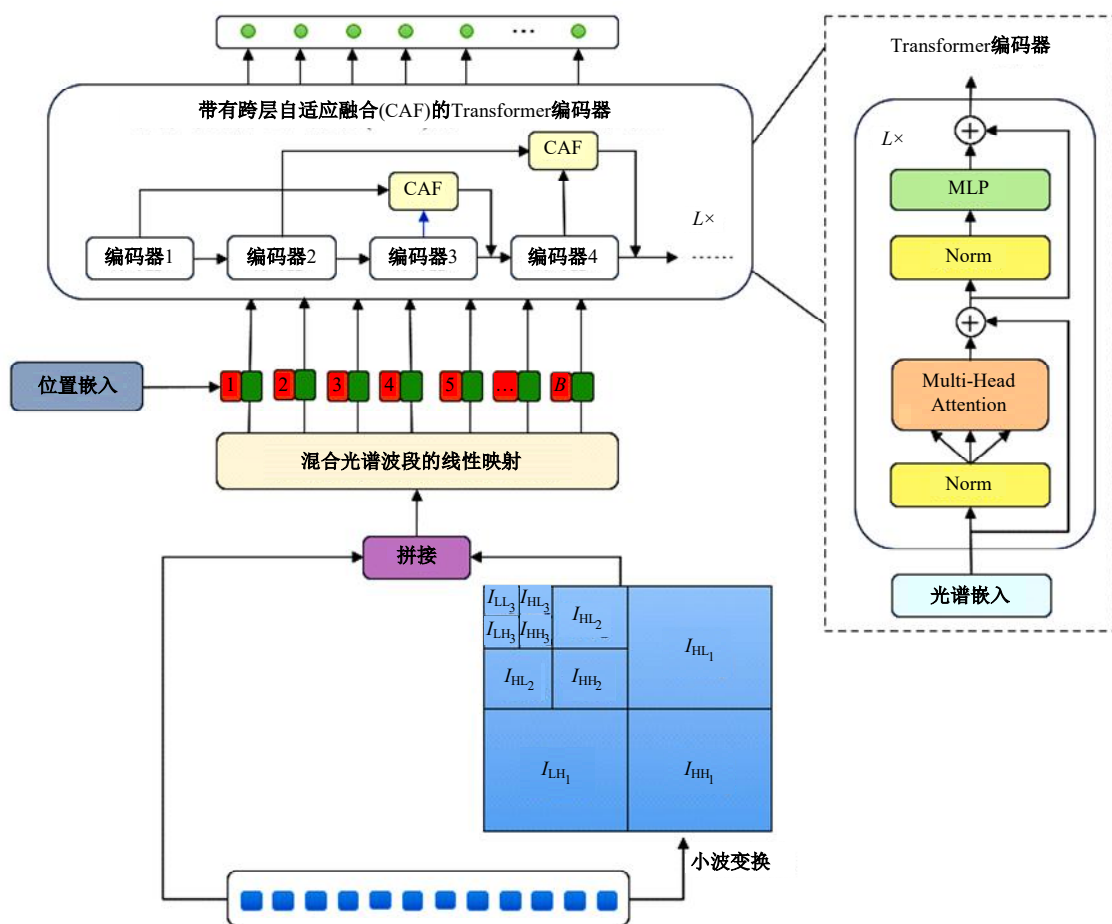


图2 混合光谱增强模块结构示意图

Fig. 2 Structure diagram of hybrid spectral enhancement module

换处理后的光谱信息相结合,共同输入 Transformer中,目的是在保留原始光谱特征的前提下,降低类内光谱变化对分类的影响,获取包含更多细节的光谱特征 X_n 。本文采用经典的Haar小波,如图3所示。具体来说,首先使用低通滤波器 $f_L=(1/\sqrt{2}, 1/\sqrt{2})$ 和高通滤波器 $f_H=(1/\sqrt{2}, -1/\sqrt{2})$ 对输入频谱带 I 沿水平方向进行滤波和下采样,生成子带系数矩阵 I_L 和 I_H 。其次采用同样的低通和高通滤波器,沿 I_L 和 I_H 的垂直方向进行过滤和下采样,生成4个亚带矩阵 I_{LL} 、 I_{LH} 、 I_{HL} 、 I_{HH} 。最后,对 I_{LL} 继续分解,以获得更细节的频谱

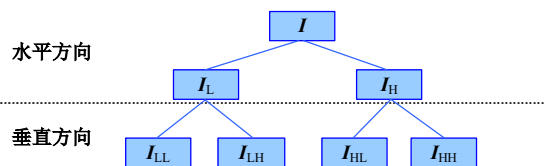


图3 Haar小波一级分解示意图

Fig. 3 Diagram of Haar wavelet first-order decomposition

特征,最终取得的光谱序列分别表示为 I_{LL}^k , $[I_{HL}^k, I_{LH}^k, I_{HH}^k](k=1, 2, \dots, J)$,本实验中 $J=3$ 。由于小波变换具有很好的时频局域性,可将光谱的不同底层特征变换成对应的小波系数,有利于提取更丰富的特征。

之后,引入位置编码 $E_{spe}^{pos} \in R^{1 \times B}$,将其与 X_n 相加,得到新的位置嵌入 $Z_n \in R^{(w \times w) \times 2 \times B}$,再将 Z_n 输入Transformer编码器。其中,自注意力机制的公式为:

$$Q = Z_n \cdot W_Q = (X_n + E_{spe}^{pos}) W_Q \quad (1)$$

$$K = Z_n \cdot W_K = (X_n + E_{spe}^{pos}) W_K \quad (2)$$

$$V = Z_n \cdot W_V = (X_n + E_{spe}^{pos}) W_V \quad (3)$$

$$\text{Attention} = \text{softmax} \left(\frac{QK^T}{\sqrt{d}} \right) V = \text{softmax} \left(\frac{Z_n W_Q W_K^T Z_n^T}{\sqrt{d}} \right) V \quad (4)$$

式中: Q 、 K 、 V 分别为查询向量、键向量和值向量。

Transformer编码器的表达式为:

$$A_n^l = T_n^{l-1} + \text{MHA}[\text{LN}(T_n^{l-1})] \quad (5)$$

$$T_n^l = A_n^l + \text{MLP}[\text{LN}(A_n^l)] \quad (6)$$

式中: T_n^l 为空间位置 n 的第 l 个 Transformer 编码器输出; MHA 和 MLP 分别为多头注意和多层感知器; LN 为归一化层; 当 $l=1$ 时, $T_n^{l-1} = Z_n$ 。

为进一步聚合不同 Transformer 编码器的尺度特征, 获得更有效的特征表达, 本文参考了 SpectralFormer 中的跨层自适应融合 (Context-aware fusion, CAF) 模块。该模块的输出结果为:

$$\bar{T}_n^l = \tilde{w} \begin{bmatrix} T_n^l \\ T_n^{l-2} \end{bmatrix} \quad (7)$$

式中: $\tilde{w} \in R^{1 \times 2}$ 为自适应融合的可学习网络参数。

2.3 多尺度空间聚合模块

本文设计的多尺度空间聚合模块结构如图 4 所示。首先, 通过一维卷积对特征图 I_n 进行处理, 提取浅层空间特征并减少特征映射的通道数, 生成新的空间模块 $\tilde{I} \in R^{b \times w \times w \times C}$ 。其次, 为进一步增强光谱特征信息, 对 $\tilde{I}$ 执行通道洗牌操作, 包括重塑、转置和再次重塑。完成一系列迭代洗牌后, 原本的通道顺序被打乱, 通道间建立了新的关联。这一过程最终生成新的特征向量 S_{shuffle} , 它在保留原始信息的同时, 增强了通道间的交互和融合。

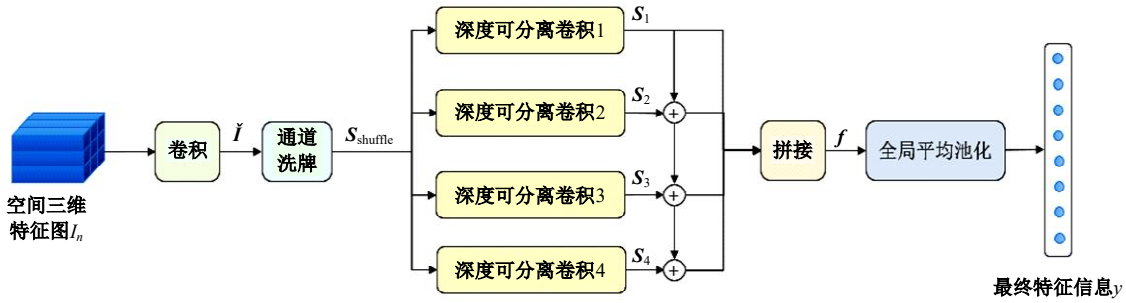


图 4 多尺度空间聚合模块结构示意图

Fig. 4 Structure diagram of multi-scale spatial aggregation module

较大的卷积核能捕捉更丰富的特征信息, 但可能导致物体边界处的预测误差增大。因此, 在设计卷积核尺寸时需进行权衡, 既要确保能有效提取空间光谱特征, 又要保证有足够的感受野以捕捉更广泛的特征信息。为拓展感知范围, 本文首先采用不同尺度的深度可分离卷积, 提取包含 4 个接收域的空间特征 $S_1, S_2, S_3, S_4 \in R^{b \times \bar{w} \times \bar{w} \times c}$, 在提高效率的同时更全面地捕捉不同尺度的空间信息。其次, 引入了依次相加的聚合构造函数, 通过元素加法操作改善了不同尺度间的交互作用。再次, 通过通道拼接策略生成了多尺度融合的空间特征图, 计算多尺度空间信息 $f \in R^{b \times \bar{w} \times \bar{w} \times 4c}$:

$$f = \text{cat}(S_1, S_1 + S_2, S_2 + S_3, S_3 + S_4) \quad (8)$$

最后, 通过全局平均池化处理和相应的维度处理, 得到最终的特征信息 $y = R^{b \times 4c}$, 增强模型的鲁棒性。

2.4 数据集

在网络模型训练中, 采用以下两个高光谱数据集验证本文方法的有效性:

Indian pines (IN) 数据集^[18]是由 AVIRIS 传感器在印第安纳州西北部采集的影像, 波长范围为 400~2 500 nm, 去除 20 个噪声严重的波段后

共有 200 个波段, 图像尺寸为 145×145 像素, 空间分辨率约为 20 m。该场景涉及 16 种类别, 其中包括许多天然植被。

WHU-hi-longkou (LK) 数据集^[19,20]是由 Headwall Nano-Hyperspec 图像传感器在中国湖北省龙口镇采集的影像, 波长范围为 400~1 000 nm, 共有 270 个波段, 图像尺寸为 550×400 像素, 空间分辨率约为 0.109 m。该场景涉及 9 种类别, 其中包括玉米、棉花、芝麻、阔叶大豆、窄叶大豆和水稻 6 种作物。

这两个数据集都具有较长的波段, 光谱波段间存在较高的冗余性, 往往会使图像的类内一致性变得较弱, 这为研究和解决类内一致性问题提供了试验场景。本文以此为基础, 设计了 HEMA 方法以应对高光谱图像分类中类内一致性较弱的问题。

2.5 损失函数

在网络模型的训练过程中, 采用交叉熵作为损失函数, 用于度量预测标签向量与实际标签向量之间的差异, 以评估本文方法的泛化性能。交叉熵的计算公式为:

$$\text{Loss} = -\frac{1}{N} [y_k \log(\bar{y}_k) + (1 - y_k) \log(1 - \bar{y}_k)] \quad (9)$$

式中: N 为类别数; y_k 为真实样本标签中属于第 k 类的概率值; $\bar{y}_k$ 为预测样本标签中属于第 k 类的概率值。

3 实验仿真

3.1 实验配置

实验使用少量样本进行训练, 在 IN 数据集中, 除第 1、7、9 类只选择 15 个样本作为训练集外, 其他类别都随机选择 50 个样本作为训练集; 在 LK 数据集中, 所有类别都随机选择 50 个样本作为训练集。设置批量大小为 64, 学习率为 0.000 5, 对所有数据集进行 10 次训练, 并采用总体准确度 (Overall accuracy, OA)、平均准确度 (Average accuracy, AA)、Kappa 系数及单类分类精度作为模型性能评估指标, 以获得对结果更准确的估计。实验平台配置为 RTX2080Ti 显卡及 Windows 操作系统, 并采用 PyTorch 深度学习框架实现。

3.2 消融实验

为验证各模块的有效性, 本文在 IN 数据集上通过添加不同模块进行消融实验。不同模块组成下的 OA、AA 及 Kappa 值如表 1 所示, 可以看出, SpectralFormer 的分类效果值最低; 将原 SpectralFormer 的输入部分替换为引入小波变换的光谱增强信息后, OA 提升了 2.02 百分点。这证实了小波变换的引入能更有效地利用每个频谱中包含的丰富信息, 提高分类准确性。多尺度空间聚合模块的引入, 与基线相比 OA 提高了 7.26 百分点, 说明相较于仅进行频谱特征提取的 SpectralFormer, 增加多尺度空间聚合模块有助于提高模型性能。在同时包含小波变换和多尺度空间聚合模块的实验中, 各评估指标较 SpectralFormer 都提高了 10 个百分点以上, 说明了两个模块的必要性。综上所述, 各模块均对网络性能有积极影响。

表 1 在 IN 数据集中 HEMA 的消融实验

Table 1 HEMA ablation experiments in the IN dataset

基线	小波变换	多尺度空间聚合模块	OA/%	AA/%	Kappa $\times 100$
	×	×	80.46	72.49	77.88
Spectral-	✓	×	82.48	73.36	80.19
Former	×	✓	87.72	83.31	86.00
	✓	✓	94.80	91.90	94.04

4 实验结果与分析

4.1 结果对比分析

将 HEMA 方法与最新的基于深度学习和 Transformer 的高光谱图像分类方法进行对比, 对比算法包括 ABLSTM^[21]、CNN3D^[5]、ViT^[6]、SpectralFormer^[7]、SSFTT^[8]。对比算法均使用相同规模的训练数据, 并从原始数据中选取大小为 $7 \times 7 \times B$ 的立方体作为输入, 其中 B 为光谱维度。

HEMA 与对比方法的分类结果如表 2、表 3 所示, 其中加粗数字为该行最大值, 即性能最优。从表中可以看出, HEMA 的分类精度相较于 ABLSTM 和 CNN3D 明显提升, OA、AA 及 Kappa 值在 IN 数据集中至少提高了 16.91 百分点、21.1 百分点和 19.15, 在 LK 数据集中至少提高了 5.28 百分点、13.37 百分点和 6.75。此外, ABLSTM 和 CNN3D 的性能整体低于基于 Transformer 的方法, 表明 Transformer 在获取全局特征方面具有优势; 改进后的 ViT 方法能更好地提取全局和局部空-谱特征, 获取更高级的语义

表 2 不同方法在 IN 数据集上的分类结果

Table 2 Classification results of different methods for IN datasets

类别	ABLSTM	CNN3D	ViT	Spectral-Former	SSFTT	HEMA
1	26.19	57.14	68.96	74.19	100.00	100.00
2	59.92	80.73	67.68	75.84	83.33	96.16
3	48.90	66.48	66.75	63.25	74.96	87.59
4	38.63	39.08	52.04	46.04	74.10	80.76
5	84.86	95.39	80.94	91.00	92.45	97.43
6	90.94	98.30	93.56	92.93	98.10	98.88
7	18.75	34.28	30.76	50.00	48.00	92.30
8	84.84	98.98	92.82	97.75	99.49	100.00
9	4.54	10.86	13.33	20.83	100.00	83.33
10	69.80	77.44	63.88	76.57	73.29	86.22
11	73.08	78.49	82.11	90.50	87.30	97.53
12	50.00	48.94	42.17	63.05	56.38	98.52
13	92.00	94.59	83.13	88.60	90.90	97.90
14	97.69	97.46	98.07	97.66	97.18	97.83
15	69.84	72.93	71.15	71.97	81.86	95.32
16	60.00	81.63	64.51	59.70	64.51	60.60
OA/%	69.90	77.89	74.96	80.46	83.77	94.80
AA/%	60.62	70.80	66.99	72.49	83.62	91.90
Kappa $\times 100$	65.84	74.89	71.58	77.88	91.54	94.04

表 3 不同方法在 LK 数据集上的分类结果
Table 3 Classification results of different methods for LK datasets

类别	ABLSTM	CNN3D	ViT	Spectral-Former	SSFTT	HEMA
1	94.15	99.04	96.16	93.34	99.61	99.21
2	75.67	68.80	70.93	87.46	81.12	97.19
3	25.96	39.53	49.13	91.15	53.96	92.15
4	97.65	98.51	98.45	98.66	98.72	99.28
5	30.99	56.97	30.10	41.39	57.08	48.84
6	91.01	84.21	96.77	96.86	98.61	99.68
7	98.92	98.66	99.92	99.97	99.96	99.95
8	81.01	82.22	84.56	90.22	91.11	94.33
9	28.37	68.43	59.33	90.53	86.68	85.34
OA/%	84.94	91.47	91.06	93.69	95.38	96.75
AA/%	69.30	77.37	76.15	85.51	85.20	90.67
Kappa×100	80.80	89.01	88.44	91.79	93.99	95.76

信息。相较于 ViT、SpectralFormer 和 SSFTT 等基于 Transformer 的方法,HEMA 的 OA、AA 和 Kappa 值均有提升,在 IN 数据集中至少提高了 11.03 百分点、8.28 百分点和 2.5,在 LK 数据集中至少提高了 1.37 百分点、5.16 百分点和 1.77。此外,HEMA 在大多数类别上的分类精度都优于对比算法。这表明 HEMA 在原始光谱信息基础上引入小波变换,可从多尺度光谱信息中获取噪声点更小的光谱信息,并增强类内光谱一致性。同时,多尺度空间聚合模块对空间信息的多尺度

特征提取,促进了不同尺度空间信息的交互,从而获得更细节的特征信息,提升了网络的分类精度。

4.2 分类效果图对比分析

HEMA 方法与对比方法在 IN 数据集和 LK 数据集上的高光谱图像分类效果如图 5、图 6 所示。可以观察到,HEMA 方法生成的分类效果图最为清晰,这意味着其正确分类的样本点数量最多。如图 5(f)相较于图 5(a)~(e),HEMA 将 Soybean-mintill 类别错分为 Corn-mintill 类别的像素点更少。相应地,如表 2、表 3 所示,HEMA 方法在两个数据集的各个类别上的分类精度几乎都达到了最佳,并且其 AA 也达到了最优。由此可见,HEMA 能更有效地捕捉光谱的局部和全局特征,以及空间的多尺度交互信息,最终得到的空-谱特征使分类效果图能更准确地表示真实的分类。

4.3 训练和测试时间对比分析

HEMA 和对比方法的平均训练时间和平均测试时间如表 4 所示。可以看出,HEMA 的训练时间比 ABLSTM 和 SpectralFormer 更短,这是因为 HEMA 在进行空间特征提取时采用了深度可分离卷积,减轻了网络的负担。CNN3D、ViT 和 SSFTT 的训练时间则比 HEMA 更短,这是因为对比网络的规模较小,本文采用的多尺度空间聚合模块增加了网络的复杂性,导致训练时间和测试时间并非最短。但 HEMA 通过增加一定的训练时间获取了更高的精确度,在效率和准确性之间实现了很好的平衡,具有一定的应用优势。

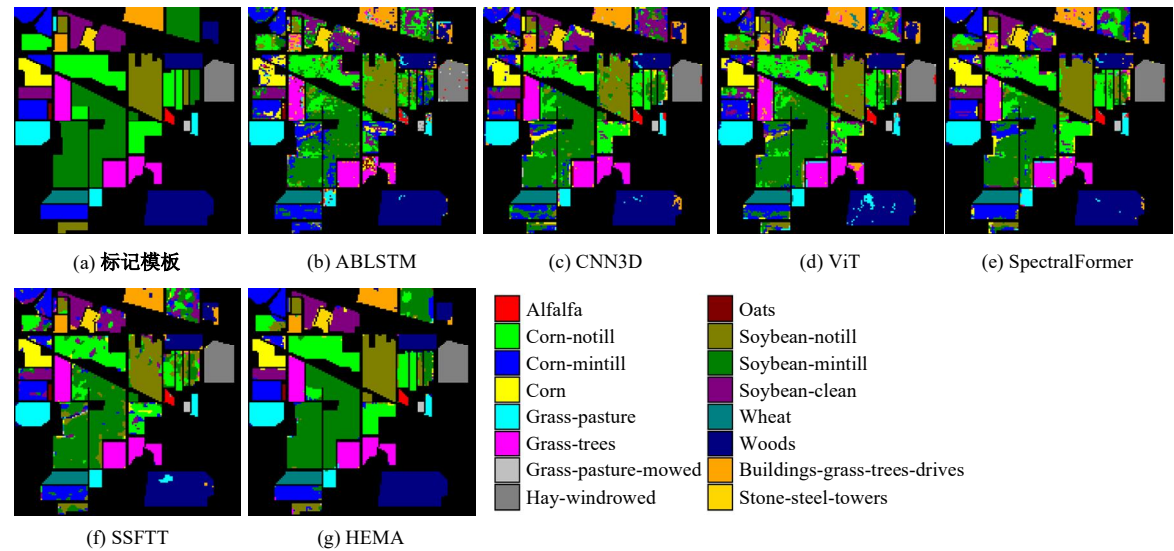


图 5 IN 数据集分类效果
Fig. 5 Classification effect for IN dataset

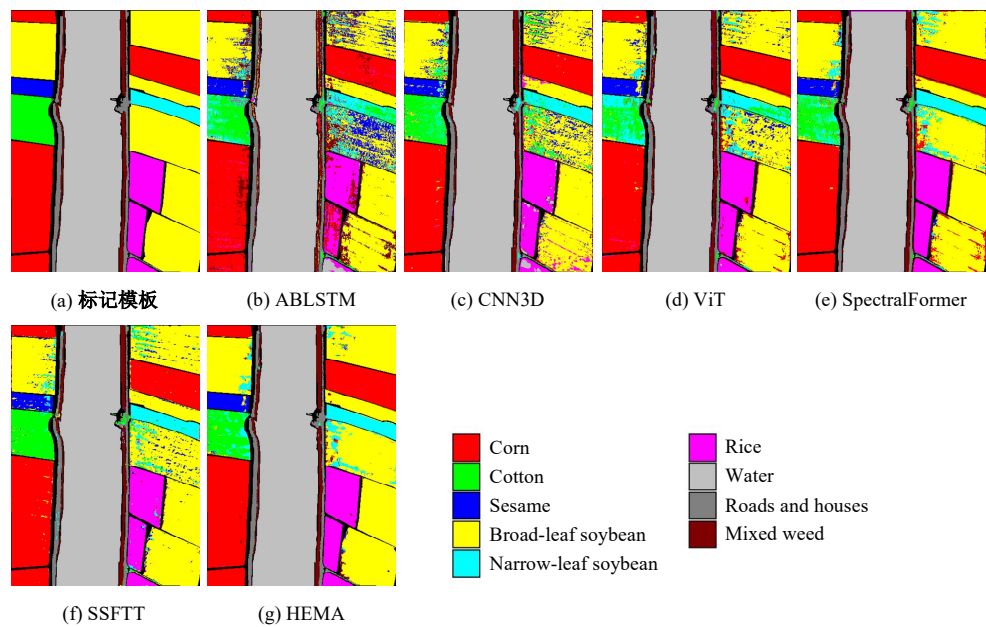


图 6 LK 数据集分类效果

Fig. 6 Classification effect for LK dataset

表 4 两种数据集下不同方法的训练时间和测试时间

Table 4 Training time and testing time of different methods for two datasets

数据集		ABLS- TM	CNN- 3D	ViT	Spectral- Former	SS- FTT	HE- MA
IN	训练	58.76	24.61	45.99	86.25	22.16	62.23
	测试	1.94	0.54	1.69	3.94	0.58	2.28
LK	训练	58.63	13.25	43.44	73.77	20.33	49.47
	测试	59.49	16.19	56.50	123.47	18.96	64.92

5 结束语

为综合利用高光谱图像中的光谱信息和空间信息,本文提出了一种基于混合光谱增强与多尺度空间聚合的高光谱图像分类方法。其中,混合光谱增强模块能有效增强光谱信息,并提取全局光谱特征,从而更全面地表征复杂的光谱特性;多尺度空间聚合模块能对不同尺度的空间信息进行交互融合,使不同尺度的特征相互引导,从而有效提高空间特征的表达能力。实验结果表明,本文提出的方法能更有效地提升高光谱图像分类任务的准确性和鲁棒性。未来研究将在保持分类精度的前提下,考虑如何降低网络模型的复杂度并提高计算效率。

参考文献:

[1] Yang K, Sun H, Zou C B, et al. Cross-attention spectral-spatial network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and

Remote Sensing, 2022, 60: 1-14.

[2] Ahmad M, Khan A M, Mazzara M, et al. Hybrid dense network with attention mechanism for hyper-spectral image classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 3948-3957.

[3] 欧阳宁,李祖锋,林乐平. 基于多层次空-谱融合网络的高光谱图像分类[J]. 吉林大学学报:工学版, 2022, 52(10):2438-2446.

Ouyang Ning, Li Zu-feng, Lin Le-ping. Hyperspectral image classification based on a multi-level spatial-spectral fusion network[J]. Journal of Jilin University (Engineering and Technology Edition), 2022, 52(10): 2438-2446.

[4] 欧阳宁,李祖锋,林乐平. 基于空谱分组卷积密集网络的高光谱图像分类[J]. 计算机工程与设计, 2022, 43(7):2031-2039.

Ouyang Ning, Li Zu-feng, Lin Le-ping. Hyperspectral image classification based on a spatial-spectral group convolution dense network[J]. Computer Engineering and Design, 2022, 43(7):2031-2039.

[5] Hamida A B, Benoit A, Lambert P, et al. 3-D deep learning approach for remote sensing image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(8): 4420-4434.

[6] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[J]. Arxiv Preprint, 2010, 11: 201011929.

- [7] Hong D F, Han Z, Yao J, et al. Spectralformer: rethinking hyperspectral image classification with transformers[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-15.
- [8] Sun L, Zhao G R, Zheng Y H, et al. Spectral - spatial feature tokenization Transformer for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.
- [9] Uddin M P, Mamun M A, Hossain M A. Feature extraction for hyperspectral image classification[C]// IEEE Region 10 Humanitarian Technology Conference. Piscataway, NJ: IEEE. 2017: 379-382.
- [10] Zhang D S. Wavelet Transform[J]. Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval, 2019: 35-44.
- [11] Tang Y G, Xie X L, Yu Y H. Hyperspectral classification of two-branch joint networks based on gaussian pyramid multiscale and wavelet transform[J]. IEEE Access, 2022, 10: 56876-56887.
- [12] Ahmad M, Ghous U, Usama M, et al. WaveFormer: spectral-spatial wavelet Transformer for hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.
- [13] Wang D, Du B, Zhang L P, et al. Adaptive spectral-spatial multiscale contextual feature extraction for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(3): 2461-2477.
- [14] Gao H M, Yang Y, Li C M, et al. Multiscale residual network with mixed depthwise convolution for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(4): 3396-3408.
- [15] Leng Q M, Yang H O, Jiang J J, et al. Adaptive multiscale segmentations for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(8): 5847-5860.
- [16] Ghaderizadeh S, Abbasi M D, Sharifi A, et al. Multiscale dual-branch residual spectral-spatial network with attention for hyperspectral image classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 5455-5467.
- [17] Wang X, Tan K, Du P J, et al. A unified multiscale learning framework for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-19.
- [18] Baumgardner M F, Biehl L L, Landgrebe D A. 220 band aviris hyperspectral image data set: June 12, 1992 Indian pine test site 3[J]. Purdue University Research Repository, 2015, 10(7): 991.
- [19] Zhong Y F, Hu X, Luo C, et al. WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H_2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF[J]. Remote Sensing of Environment, 2020, 250: 112012.
- [20] Zhong Y F, Wang X Y, Xu Y, et al. Mini-UAV-borne hyperspectral remote sensing: From observation and processing to applications[J]. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(4): 46-62.
- [21] Mei S H, Li X G, Liu X, et al. Hyperspectral image classification using attention-based bidirectional long short-term memory network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-12.