DOI:10.16853/j.cnki.1009-3575.2026.06.010

多尺度残差-边缘感知融合的荒漠草原高光谱弱监督分类模型

王 森,毕玉革*,贺长彬,侯占峰,牛婧洁,刘雅旋
(内蒙古农业大学机电工程学院,呼和浩特 010018)

摘要:针对荒漠草原高光谱地物分类中大量人工标注依赖强、地物边界识别不准等问题,本文提出一种融合多尺度残差-边缘感知结构的弱监督分类模型——ResSpec3D-Net。该模型以U-Net为基础框架,引入简化后的Res2Net模块进行空间-光谱多尺度特征提取,同时构建由三维深度可分离卷积(3D-Depthwise Conv)与三维逐点卷积(3D-Pointwise Conv)组合而成的轻量化3D-LDC Block模块,以进一步增强光谱-空间联合特征的表达能力。为提升对模糊边界的识别精度,本文构建基于信息熵的不确定性度量,提出边缘感知损失函数,在无伪标签参与的条件下有效增强模型对边界区域的不确定性感知。在每个样方仅约1%像元被标注的弱监督设定下,ResSpec3D-Net在15组荒漠草原高光谱样方上实现平均O_A为98.5%,Kappa为97.8%,不同地物F_1分数分别为植被99.5%、凋落层97.1%、裸土97.8%、其他99.8%的高分类性能。同时,消融实验验证了各模块在提升精度和鲁棒性方面的有效性。研究结果表明,所提方法在低标注条件下依然具有显著的边界识别能力与分类精度,为高光谱遥感在荒漠草原生态监测中的应用提供了新的思路。

关键词:高光谱遥感;弱监督学习;边缘感知;多尺度残差;荒漠草原;地物分类
中图分类号:S24 **文献标志码**:A **文章编号**:1009-3575(2026)06-0081-12

Multi-Scale Residual Edge-aware Fusion Model for Weakly Supervised Hyperspectral Classification in Desert Grasslands

WANG Sen, BI Yuge*, HE Changbin, HOU Zhanfeng, NIU Jingjie, LIU Yaxuan
(*College of Mechanical and Electrical Engineering*, *Inner Mongolia Agricultural University*, *Hohhot* 010018, *China*)

Abstract: Hyperspectral land-cover classification in desert grasslands faces two major challenges, i.e., heavy reliance on manual labeling and inaccurate boundary recognition. To address these issues, this study proposed ResSpec3D-Net, a weakly supervised classification model that integrates multi-scale residual and edge-aware structures. The model builds on the U-Net framework and incorporates a simplified Res2Net module for multi-scale spatial-spectral feature extraction. To further enhance joint spatial-spectral representation, a lightweight 3D-LDC Block—composed of 3D depthwise and 3D pointwise convolutions—was introduced. In addition, an entropy-based edge-aware loss was designed to capture uncertainty in boundary regions, improving boundary recognition without requiring pseudo-labels. Experiments on 15 desert grassland hyperspectral plots showed that, with only 1% labeled pixels per plot, the ResSpec3D-Net model achieved an average O_A of 98.5%, Kappa of 97.8%, and F_1-scores of 99.5% for vegetation, 97.1% for litter, 97.8% for bare soil, and 99.8% for others. Ablation studies confirmed the contributions of each module to accuracy and robustness. These results demonstrated that the proposed model maintained high boundary sensitivity and classification accuracy under low-annotation conditions, offering a promising approach for hyperspectral remote sensing applications in the desert grassland ecological monitoring.

Key words: Hyperspectral remote sensing; Weakly supervised learning; Edge-aware; Multi-scale residual; Desert grasslands; Land-Cover classification

草原作为重要的生态系统,对水土保持、生物多样性及气候调节等具有重要价值,但由于自然因素与人为因素影响,草原生态系统出现严重的退化问题[1]。草原荒漠化进程会对草原生态平衡造成严重的破坏,而精准的地物分类是草原监测与管理的关键环节。及时获取草原的地物分类,能够帮助草业工作人员快速调整并制定合理的草原管理体系[2]。

在草原调查领域,传统方法通常是人工在草原上进行测量、记录主群落种类和覆盖度等信息。不仅耗费大量人力、物力和时间,而且调查范围有限,难以实现对广袤草原的快速且全面的监测[3]。随着遥感技术的兴起,草原调查迎来了新的变革。早

收稿日期:2025-09-03
基金项目:国家自然科学基金项目(52465032);内蒙古自治区自然科学基金项目(2023LHMS06010);内蒙古自治区自然科学重点(NJZZ23037);内蒙古农业大学一流学科科研专项项目(YLXKZX-NND-046)
作者简介:王 森(2002—),男,硕士研究生,主要从事基于无人机高光谱的荒漠草原退化反演方面的研究,E-mail:2039656583@qq.com;*为通信作者,毕玉革(1974—),女,博士,副教授,主要从事于农牧业机械智能化方面的研究,E-mail:biyuge163@imau.edu.cn

期,卫星遥感成为获取草原数据的重要手段。卫星遥感凭借其覆盖范围广的优势,能够从太空俯瞰大片草原区域,为人们提供大尺度的草原地物信息,在宏观把握草原整体状况方面发挥了重要作用[4]。但卫星遥感也存在明显的不足:一方面,其时间分辨率有限,两次成像间隔较长,无法及时捕捉草原地物在短时间内的变化,存在时间不连续性[5];另一方面,卫星遥感的空间分辨率相对较低,难以精确获取地物的精细化信息,无法满足对草原主群落种类、生长状况等进行精准分析的需求[6]。近年来,无人机技术的蓬勃发展为草原调查带来了新的契机[7]。无人机搭载高光谱传感器获取草原数据逐渐成为主流方式。无人机可以根据调查需求灵活规划飞行路线,深入到草原的各个角落,而高光谱可以进行精细化的地物数据采集,两者的结合可以在单个像元级别捕捉到草原地物更细微的光谱差异[8]。

在获取高光谱数据后,学者们通常采用强监督学习对高光谱数据进行特征提取并分类研究[9]。强监督学习依赖大量的人工标注样本数据。例如:杨红艳等[10]基于无人机高光谱数据,构建融合光谱特征、植被特征、地形特征和纹理特征等24个变量的随机森林分类模型,实现了荒漠草原植被分类91.06%的整体精度;佘洁等[11]通过切入面向对象的不同方式来探究荒漠草原植被分类的最佳模型;Jiang等[12]对植物群落图像进行预处理,通过对11种深度学习模型运行后发现,视觉Transformer虽训练耗时较长,但凭借对大量标注数据的有效学习,实现88.72%的最高分类准确率,经预处理数据后性能更优,但对荒漠化程度相近的易混淆类别有效识别能力较差;Jin等[13]通过构建上万像元的数据集样本并结合提出的两个基于残差网络(ResNet)的模型DGRNet和3D-DGRNet进行草地退化指示物分类研究,得到了总体精度(O_A)99.05%的分类效果。与此同时,其他研究也从不同角度揭示了强监督方法对海量标注的依赖:杨星等[14]系统阐述了近年高光谱影像分类方法,明确指出人工标注耗时、成本高,难以获得大量标注样本,是应用落地与精度提升的主要瓶颈;Yang等[15]在无人机高光谱作物分类研究中也强调模型需要大量标注样本,而实际可获得的标注有限,制约了模型的推广;Wang等[16]进一步指出强监督学习普遍面临标注获取困难的问题;Zhang等[17]则在WHU-H数据集的大规模全监督实验中,使用包含255 930个标注像元、16个类别的HanChuan子集进行训练,直接体现了强监督方案对巨量像元级标注的依赖,并讨论了这种大规模标注对模型性能和训练过程的影响。

虽然上述方法在遥感分类精度上取得了显著提升,但大多依赖于对高光谱影像进行精细化人工标注。该标注过程不仅耗费大量人力物力,还要求标注人员具备较高的专业知识与丰富的经验[18];同时,长时间的重复性工作容易导致主观偏差和一致性缺失,从而在一定程度上影响模型的学习效果和泛化性能[19]。因此针对上述问题,Zhou[20]在2018年提出了弱监督学习(weakly supervised learning)框架,该框架把弱监督学习分为了不完全监督、不确切监督、不精确监督。其中不完全监督只有部分训练样本带标签,其余无标签。

本文借助弱监督学习中的“不完全监督”思想,即仅在每个样方中选取少量纯净像元进行标注,其余像元保持未标注状态。同时,本文不引入额外伪标签传播过程,而是将基础的不完全监督策略作为研究起点。在此基础上,本文提出了基于U-Net骨架的ResSpec3D-Net模型,通过引入Res2Net模块实现空间-光谱多尺度残差特征提取,并结合组装的3D-LDC Block模块加强深层特征建模能力。在仅使用每个样方约2 500个纯净像元(约占1%)的标注情况下,模型仍展现出较高的分类精度与稳定性,有效缓解了高光谱遥感中人工标注成本高的问题,为荒漠草原地物的弱监督分类提供了新思路。

1 研究区与数据源

1.1 研究区概述

研究区位于内蒙古自治区乌兰察布市四子王旗境内的格根塔拉草原,地处内蒙古南部,属于典型的荒漠草原生态类型[21]。该区域属中温带大陆性气候,年平均气温为1~6 ℃,昼夜温差大,年际气温波动剧烈。降水量稀少,年均降水量集中在200~350 mm,蒸发量显著高于降水量,区域生态环境较为脆弱。在大气环流控制下,该地常年受蒙古高压气团影响,风力强劲,风沙天气频发。全年≥8级大风日数超过20 d,春季尤为集中,部分区域大风持续时间可达两个月。

在该研究区所处的荒漠化草原中,短花针茅,无花隐子草等,构成该区域的主要植被群落结构[22]。在部分退化严重区域,植被群落的覆盖度明显下降,地表逐渐被凋落层(如枯草,残枝等)及沙化裸土所取代,形成典型的群落退化斑块化分布特征。此外,在样方调查过程中,还存在一些人为

干扰因素，最为明显的是采样地垫、路径压痕、小旗等，这些辅助物体虽非自然地物，但在高光谱影像中表现出显著的光谱特征，需在分类过程中合理区分与剔除。

因此，本文在地物分类过程中，结合高光谱遥感数据的特征表达能力与草原生态演替背景，将区域内地物划分为 4 类：植被、凋落层、裸土及其他。该分类体系不仅符合遥感可识别性原则，也较好地反映了草原生态退化的主要阶段，为后续草原监测与生态恢复提供了数据基础和方法支持。

1.2　无人机高光谱遥感系统

本文所采用的飞行器是由专业级 A3 Pro 飞行控制系统控制的大疆 M600 Pro 六旋翼无人机，携带四川双利合谱科技有限公司生产的 GaiaSky-mini 型高光谱仪，光谱波长范围在 400~1 000 nm，光谱分辨率为 3. 5 nm，共 256 个波段。

1.3　数据采集

高光谱数据采集时间是 2024 年 7 月，中午 11:00—14:00。采集条件在晴朗无云少风情况下随机拍摄了 15 组样方。每个样方设置为 1 m×1 m 并使用实验辅助物（地垫和小旗）来进行样方定点标记。无人机飞行高度为 30 m，空间分辨率 1. 73 cm，采集的高光谱图像尺寸为 696 行×775 列×256 波段，且对每个样方重复拍摄 5 次，以此保证数据的可靠性。

1.4　反射率矫正

无人机在拍摄过程中受到自然环境的影响，会导致部分拍摄的高光谱图像产生形变扭曲，因此去除扭曲形变程度的部分，并对原有的高光谱图像尺寸裁切为 500 行×500 列×256 波段来减少高光谱图像的边缘扭曲。然后，挑选过后的数据（包括黑帧、白帧、原始样方图像）导入 SpecView 软件进行辐射校正。矫正计算如公式（1）[23]：

$$R = \frac{I_{\text{raw}} - I_{\text{dark}}}{I_{\text{whise}} - I_{\text{dark}}} \tag{1}$$

式中，R 为辐射校正后的图像，I_{raw} 为原始图像，I_{whise} 为白帧图像，I_{dark} 为黑帧图像。

对辐射校正后的图像，使用 ENV5. 6 软件进行标注地物像元的空间提取，随机选取了 5 个样方，分别对每个样方每种地物提取 100 个纯净像元再汇总，并使用 Origin2021 软件绘制每种地物的反射率曲线，如图 1 所示。

由各地物的反射率矫正曲线可知，不同的地物在不同的波段出现不同的特征，从全波段来看：

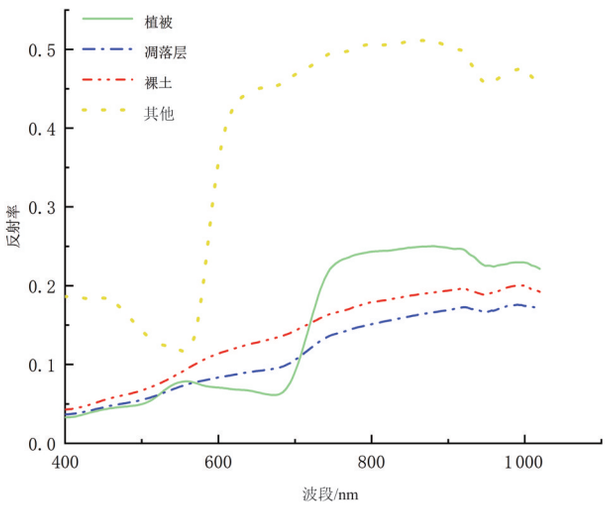


图 1　地物校正后曲线

Fig. 1　Spectral curves after ground-object correction

裸土和凋落层光谱曲线大体相似。其中：裸土因土壤矿物，短波长反射低、长波长渐升；而凋落层由于其细胞结构的不完整性导致叶绿素含量被分解，从而在全波段低反射、曲线平缓且低于裸土的光谱曲线反射率；植被受叶绿素作用，呈“蓝谷-绿峰-红谷”特征；“其他”（人为扰动）全波段高反射且无规律波动。

1.5　主成分分析法

高光谱图像虽含大量波段、信息丰富，但存在高冗余与高相关性，易加重模型计算负担并引发过拟合，进而影响泛化性能[24]。因此，高光谱数据建模前常需降维以提取主信息、剔除冗余特征。主成分分析法（PCA）是常用线性降维方法，通过构建协方差矩阵并做特征值分解，将原始高维波段映射到无相关主成分空间，能最大程度保留主信息、降低维度，提升后续模型训练效率与稳定性[25]。PCA 降维在保障信息完整性的同时，还能提供更紧凑有效的输入特征，已广泛用于农业遥感、地物分类等高光谱图像处理任务。

2　研究方法

本文使用无人机高光谱数据进行荒漠草原的地物分类研究，技术路线如图 2 所示，主要包括数据获取、数据处理、模型构建与模型评价等 4 个部分。数据获取阶段主要是使用无人机搭载的高光谱进行样方数据的拍摄。数据处理阶段使用反射率矫正后的数据，在样方图像中仅对所有地物选取少量纯净像元（每个样方标注 2 500 个像元，占全像元的 1%）进行标注，构建训练集。模型在训练过程中仅

使用这部分标签进行监督学习且按照固定比例（训练集70%、验证集15%、测试集15%）进行训练集、验证集和测试集的划分。模型构建阶段主要是先进行主成分分析（PCA），在保留90%以上的光谱信息条件下从原本256个波段降维至32个波段数，接着构建了基于经典U-Net改进的ResSpec3D-Net模型，用于荒漠草原地物分类任务，且采用了多分类的交叉熵函数和边缘感知损失函数构建联合损失函数来加快模型的收敛速度。模型评价阶段主要是分为不同模型精度评价。

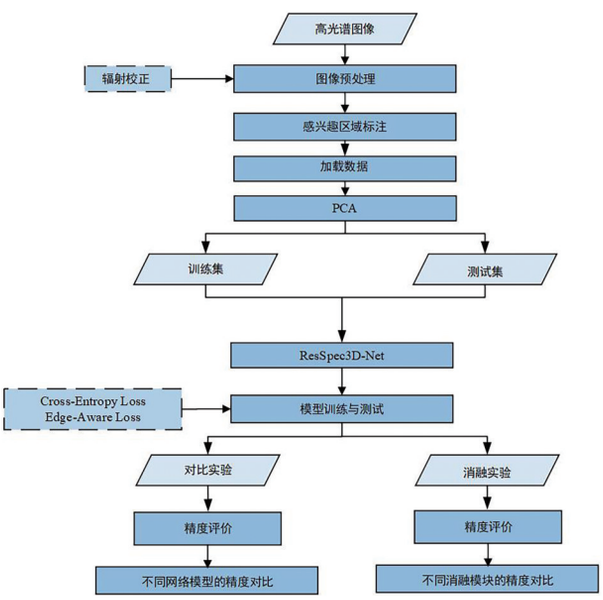


图2 荒漠草原地物分类技术路线

Fig. 2 Technical framework for hyperspectral land-cover classification in desert grasslands

2.1 Res2Ne和3D-LDC Block模块

Res2Net模块是由Gao等[26]提出的一种多尺度残差单元，它通过在单个残差块（residual block）内部构建层级残差连接（hierarchical residual-like connections），实现更细粒度的多尺度特征表达能力。在本文中，对原始Res2Net结构进行了简化处理，去除了其在输入和输出阶段的1×1 2D-Conv卷积模块，仅保留其核心的残差部分，以减小模型参数量并突出多尺度特征融合效果的本质作用。其具体结构如图3-a所示：主要包括通道分割和并行级联卷积两个部分。其中，通道分割是将输入原图像通道数分割为几个子通道数；并行级联卷积是对第一组子通道直接保留（不卷积），后续子通道组则以上一级输出为基础，通过级联方式传递特征。具体地，第i个子通道组的输出Y_i可表示为公式（2）：

$$\boldsymbol{Y}_i=\begin{cases}X_1, & i=1\\ f\left(X_i+Y_{i-1}\right), & 2\leqslant i\leqslant n\end{cases} \tag{2}$$

式中，f代表的是进行3×3 2D-Conv操作；X_1代表最初的第一组通道；拼接输出n个通道数$\{Y_1,Y_2,\cdots,Y_n\}$，最终形成多尺度特征$Y\in R^{C\times H\times W}$进行输出。

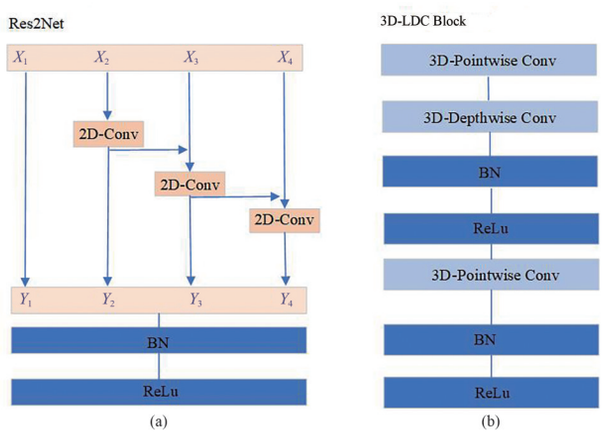


图3 Res2Net模块和3D-LDC Block框架图

Fig. 3 Framework of the Res2Net module and 3D-LDC Block

为了增强对样方的光谱信息-空间信息的多特征信息提取且减少普通3D-Conv带来的时间耗费，本文引入三维深度可分离卷积（3D-Depthwise Conv）和三维逐点卷积（3D-Pointwise Conv）两个部分并将其封装为3D-LDC Block模块，具体见图3-b，其核心是3D-Depthwise Conv模块设置1×3×3卷积核和步长s=(1,2,2)，在H/W方向下采样2倍的同时保持光谱连续性代替原来U-Net的下采样部分。而三维逐点卷积（3D-Pointwise Conv）1×1×1 3D-Conv起着升维和降维的作用，对于输入的信息起着升维的作用，对于输出的信息起着降维的作用，因此实现跨通道的融合。

通过Res2Net模块和3D-LDC Block模块实现了在少量标注像元下增强对光谱图像本身信息的多尺度残差特征的提取，同时又实现了输出通道数翻倍、空间尺寸减半的下采样。

2.2 ResSpec3D-Net荒漠草原地物分类模型

U-Net是由Ronneberger等[27]提出的一种经典语义分割网络。它沿用了全卷积网络（fully convolutional networks，FCN）的Encoder-Decoder架构[28]：通过对编码器和解码器之间搭建跳跃连接部分来增强特征表达能力。本文荒漠草原的地物分类模型基于

U-Net模型进行改进，设计的ResSpec3D-Net如图4所示，具体的模块参数表1所示，模型主要包括下采样模块、上采样模块、跳跃连接和分类器4个部分。模型区别于基础U-Net的主要体现：首先在原有4层下采样的基础上修改为只进行两次下采样，具体为在layer0和layer1层分别使用7×7 2D-Conv进行光谱信息的粗提取和引入Res2Net模块进行高光谱的多尺度残差信息提取；其次去除了模型的layer4层避免了空间维度的丢失，避免了多尺度上的信息丢失。

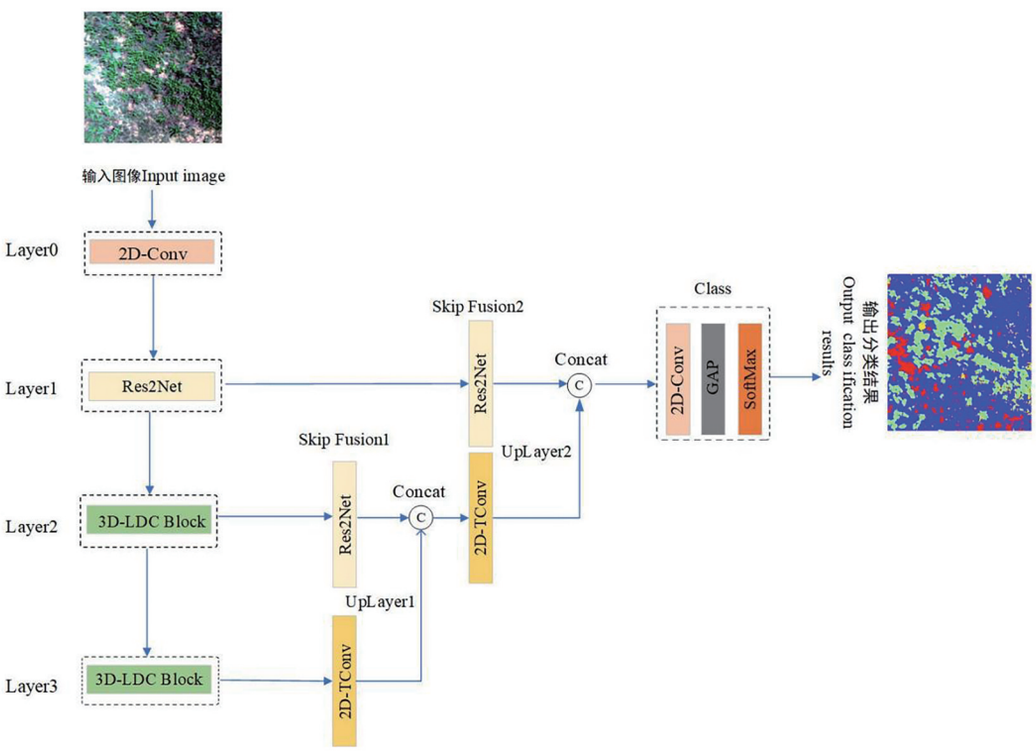


图4　ResSpec3D-Net模型架构

Fig. 4　Architecture of the ResSpec3D-Net model

表1　ResSpec3D-Net 网络结构模块参数表

Table 1　Parameters of the ResSpec3D-Net network modules

阶段	模块	核尺寸，步长	输入→输出通道	特征图分辨率
Layer0	2D-Conv	7×7，s=1	32→16	H×W
Layer1	Res2Net	3×3（3个-2D Conv），s=1	16→16	H×W
Layer2	3D-Depthwise Conv	1×3×3，s=(1, 2, 2)	16→32	H/2×W/2
	3D-Pointwise Conv	1×1×1，s=1	32→32	
Layer3	3D-Depthwise Conv	1×3×3，s=(1, 2, 2)	32→64	H/4×W/4
	3D-Pointwise Conv	1×1×1，s=1	64→64	
UpLayer1	2D-TConv	2×2，s=2		H/2×W/2
Skip fusion1	Res2Net	3×3（3个-2D Conv），s=1	32+32→32	
UpLayer2	2D-TConv	2×2，s=2	32→16	H×W
Skip fusion2	Res2Net	3×3（3个-2D Conv），s=1	16+16→16	
Class	2D-Conv	1×1/s=1		
	GAP		16→4（类别数）	H×W
	SoftMax			

下采样部分：layer0层采用7×7 2D-Conv卷积提取浅层光谱-空间特征；layer1层引入Res2Net单元，提取同一分辨率下荒漠草原不同地物的多尺度空间-光谱特征；layer2到layer3部分使用3D-LDC

Block进行下采样。上采样部分：解码器中仅使用两次2×2 2D-TConv（ConvTranspose）将特征图分辨率恢复至编码器对应层大小。转置卷积相较于双线性插值，可以学习参数的同时提升上采样质量并减少计算冗余。跳跃连接部分：考虑到编码器输出与解码器上采样结果在语义层级和通道分布上的差异，模型在每次融合前使用Res2Net模块对编码侧特征图进行对齐处理。随后将上采样特征与对齐后的跳跃特征进行逐元素相加，以融合浅层细节与深层语义信息。分类器部分：使用1×1 2D-Conv进行通道调整，并引入全局平均池化（GAP）与SoftMax层，输出对应类别的概率分布，实现像元级地物分类。

2.3 ResSpec3D-Net荒漠草原地物分类模型的损失函数

针对荒漠草原地物分布中多数类主导、类别边界模糊等问题[29]，本文设计了一种联合损失函数，结合了传统的多类交叉熵损失函数（Cross-Entropy Loss，简称L_{CE}）与边缘感知损失函数（Edge-Aware Loss，简称L_{ED}），以提高模型对边界区域与少数类的分类能力。该联合损失函数（L）的表达形式如公式（3）：

$$L = (1 - \lambda)L_{CE} + \lambda L_{ED} \tag{3}$$

式中：$(1 - \lambda)$是交叉熵损失函数的权重；λ是边缘感知损失函数的权重。

交叉熵损失函数是传统的用来衡量预测值和真实值之间的差异的损失函数，交叉熵损失也常被用来执行二分类和多分类任务，交叉熵损失函数见公式（4）：

$$L_{CE} = -\sum_{i \in N}\sum_{l \in L} y_i^{(l)} \log \hat{y}_i^{(l)} \tag{4}$$

式中：$y_i^{(l)} \in \{0, 1\}$是one-hot编码；$\hat{y}_i^{(l)} \in [0, 1]$是模型预测的概率向量。

本文提出的边缘感知损失函数的核心思想是通过对模型输出的预测概率分布进行信息熵计算，来度量其分类不确定性：预测结果越接近独热编码（one-hot）[30]，说明模型越“确信”当前像元的类别，信息熵值越低，边缘越清晰；反之，预测分布越均匀，熵值越高，表示模型在该位置判断不确定，通常出现在类间边界区域。具体见公式（5）：

$$L_{ED} = \frac{1}{N}\sum_{i=1}^{N}\left(-\sum_{c=1}^{C}\hat{y}_{i,c}\log\left(\hat{y}_{i,c} + \epsilon\right)\right) \tag{5}$$

式中：N表示批次样本数量；C表示分类类别数；$\hat{y}_{i,c}$表示第i个样本对类别c的预测概率（经过SoftMax）；ϵ为极小常数，防止数值不稳定，在本文中取10^{-8}，内部项为样本i的预测概率。

2.4 荒漠草原地物分类模型的评价指标

本文在草原地物分类模型评价指标上采用总体精度O_A（overall accuracy[31]），Kappa系数[32]和F_1分数[33]进行评价。见公式（6）—（8）：

$$O_A = P_0 = \frac{T_P + T_N}{T_P + F_N + F_P + T_N} \tag{6}$$

$$\begin{cases} R = \dfrac{T_P}{T_P + F_N} \\ F_1 = 2 \times \dfrac{P \times R}{P + R} \end{cases} \qquad P = \frac{T_P}{T_P + F_P} \tag{7}$$

$$\begin{cases} \text{Kappa} = \dfrac{P_0 - P_e}{1 - P_e} \\ P_e = \dfrac{(T_P + F_N) \times (T_P + F_P) + (F_N + T_N) \times (T_N + F_P)}{N^2} \end{cases} \tag{8}$$

式中：T_P为预测正样本；T_N为预测负样本；F_P为真实正样本；F_N为真实负样本；P为精确率，表征模型预测正样本结果的可信度；R为召回率，反映模型识别全部真实正样本的能力；N为样本总数量；P_0为O_A系数；P_e为随机分类期望概率，Kappa系数引入P_e可以消除样本分布失衡带来的评价偏差，客观衡量模型预测标签与真实标签的一致性程度。

这3种评价指标的数值越接近1，表示模型分类性能越好。具体而言，O_A与Kappa指标主要用于衡量模型在整体样本上的分类准确性与一致性；而F_1分数则用于分析模型对各个地物类别的分类性能，能够更全面地反映模型对不同类别的识别效果，尤其适合评估少数类与边界类的分类表现。

3 结果与分析

本实验操作系统为Windows11。硬件：显卡NVIDIA RTX4060，8GB独显，CPU AMD R7-7435H运行内存16 GB。软件环境为Python3.8，使用Pycharm软件作为开发平台。本实验以15个样方为研究对象，每次加载单个样方进行模型训练。训练过程中采用交叉熵与边缘感知损失构建的联合损失函数作为优化目标，学习率设为0.001，批次大小

为64。为保证模型训练，考虑到不同的地物标注多少不同，将每个样方按照地物标注多少分层抽样为训练集、验证集和测试集3部分，其中训练集占70%、验证集占15%、测试集占15%。训练集用于模型参数的学习，验证集用于监控模型在训练过程中的泛化能力并调节超参数，测试集用于最终的独立精度评估。并固定随机种子以保证实验的可重复性；联合损失函数的权重λ设置为0.2。上述参数设置通过预实验调优确定，在保证模型收敛效率的同时，有效提升了其在弱监督条件下的稳定性与分类性能。

3.1　消融实验设计

为系统评估所提ResSpec3D-Net模型中各关键模块的有效性，本文从结构和损失函数两个角度出发，设计了模块化消融实验。具体包括Res2Net模块、3D-LDC Block模块及边缘感知损失函数的逐项与组合去除实验。实验在15个高光谱样方中进行，分别进行5轮重复训练，每轮训练设定为30个epoch，并引入早停机制：当某一轮训练过程中，模型在O_A、Kappa与F_1 3个指标上连续5次无显著提升时，终止该轮训练以避免过拟合。其结果如表2和图5所示。

表2　消融实验结果

Table 2　Results of the ablation experiments

配置	Res2Net	3D-LDC Block	Edge-aware Loss	O_A/%	Kappa/%	F_1/%	时间/s	参数量/MB
1	√			96.8±0.5	95.7±0.1	95.2±0.1	33.0±0.6	0.07
2		√		97.7±0.5	96.8±0.4	97.0±0.6	33.4±0.3	0.07
3			√	98.1±0.3	97.1±0.1	97.3±0.5	33.2±1.2	0.06
4	√	√		99.1±0.3	98.5±0.9	98.8±0.7	32.5±0.6	0.08
5	√	√	√	99.4±0.1	99.0±0.6	99.2±0.5	31.8±0.3	0.08

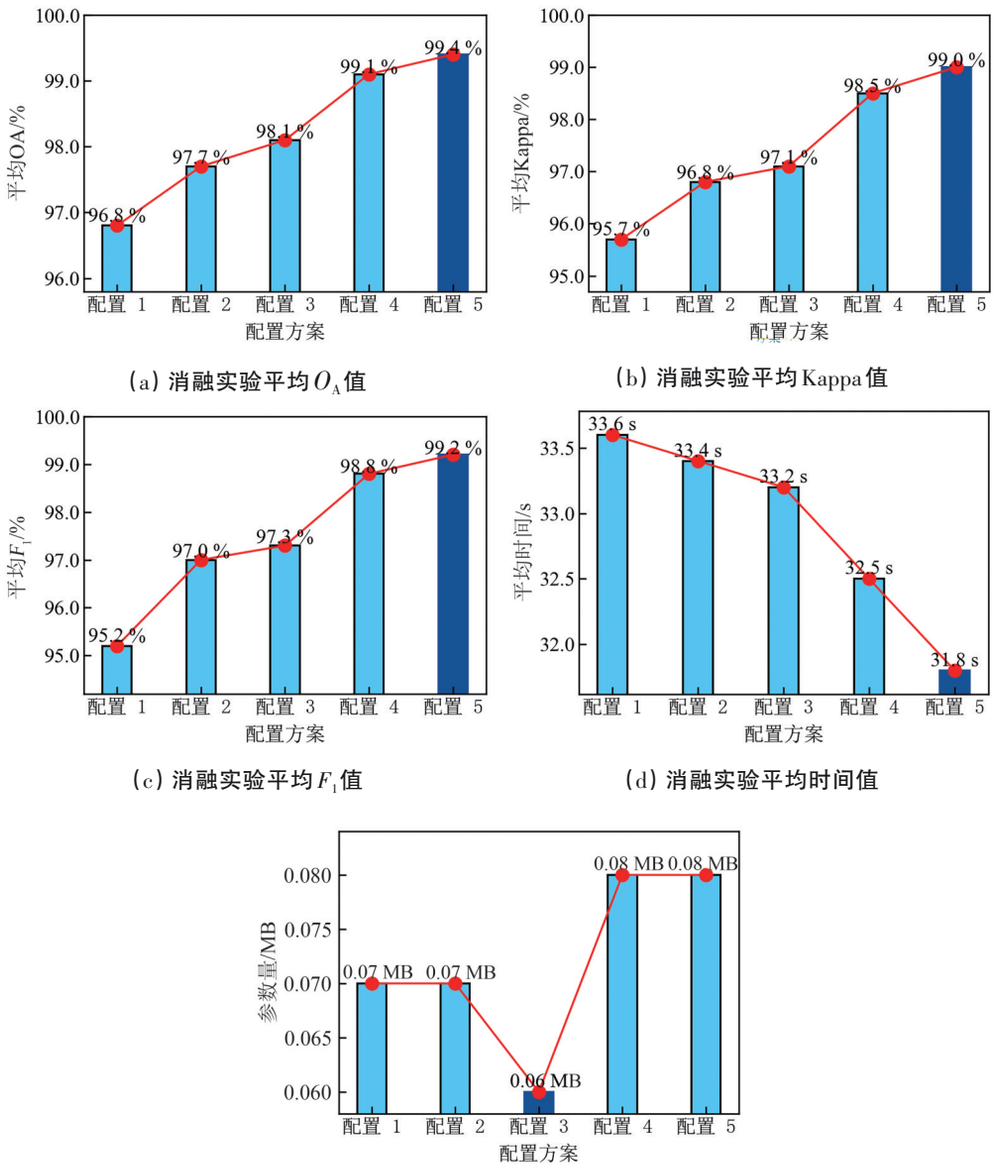
实验结果表明：ResSpec3D-Net的三个核心组件在模型性能提升中发挥了关键作用：实验结果表明（表2），任一核心组件的去除均导致性能下降。其中，结构性模块的贡献尤为显著：相比仅使用Res2Net（配置1）或仅使用3D-LDC Block（配置2），两者结合（配置4）可在O_A、Kappa、F_1 3项指标上平均提升≥1%，验证了多尺度残差与轻量化空间-光谱特征建模的协同增益。Edge-aware Loss在处理类别边界模糊区域时表现突出：在引入该模块（配置5）后，相较于未引入的配置4，O_A、Kappa、F_1分别提升0.3%、0.5%、0.4%，表明该损失项有效抑制了边界过度平滑并增强了不确定区域的判别能力。在时间性能与参数量方面，完整模型（配置5）的平均训练时间较单模块配置最高减少约1.6 s，而参数量仅增加约0.2 MB。值得注意的是，不同配置间部分模块的参数量并非严格成正比，有些组合尽管模块差异较大，但3D-LDC Block通过深度可分离卷积实现参数共享与计算拆分，使新增模块对总体参数的贡献有限；同时，Res2Net模块借助多尺度残差连接削减了冗余特征的重复运算，Edge-aware Loss主要作用于梯度传播阶段，对卷积计算量影响很小。三者协同揭示了即使在少量参数增加的情况下整体运算效率仍能提升的原因，在保证高精度的同时，实现了精度与速度的双重优化。

综上，ResSpec3D-Net的各核心组件在特征融合、边界识别及收敛效率等方面具有互补性，完整模型在精度与速度上均表现最佳，为低标注高光谱分类任务提供了高效可行的结构参考。

3.2　ResSpec3D-Net模型的分类性能对比

为验证ResSpec3D-Net在高光谱地物分类中的准确性与鲁棒性，本文将ResSpec3D-Net模型与5种基于经典U-Net骨架改进的主流模型进行对比实验，分别是ResNet34、ResNet50、StackedU-Net、Swin-Unet、CSWin-UNet 5个对比模型。其实验运行轮次和上述消融实验一致，O_A、Kappa结果值如表3所示，其中15个样方在不同模型下的O_A和Kappa值如表3和图6所示。

实验证明，针对本研究所构建的数据集，所提出的ResSpec3D-Net模型在地物分类任务中展现出优于其他基于U-Net的变体模型的整体分类精度与鲁棒性。对比分析15个样方的分类结果可见，ResSpec3D-Net不仅在总体精度上更具优势，鲁棒性也更强，其O_A、Kappa值相较于其他模型最高提升10.6%、13.6%的分类精度。综合样方分布数据而言，ResNet34、ResNet50和StackedU-Net 3种模型在不同样方间的分类精度波动较大，其中StackedU-



(a) 消融实验平均O_A值　(b) 消融实验平均Kappa值

(c) 消融实验平均F_1值　(d) 消融实验平均时间值

(e) 消融实验各个模块参数量值

图5　消融实验结果图图

Fig. 5　Visualization of ablation experiment results

表3　不同模型在15个样方下的平均精度表

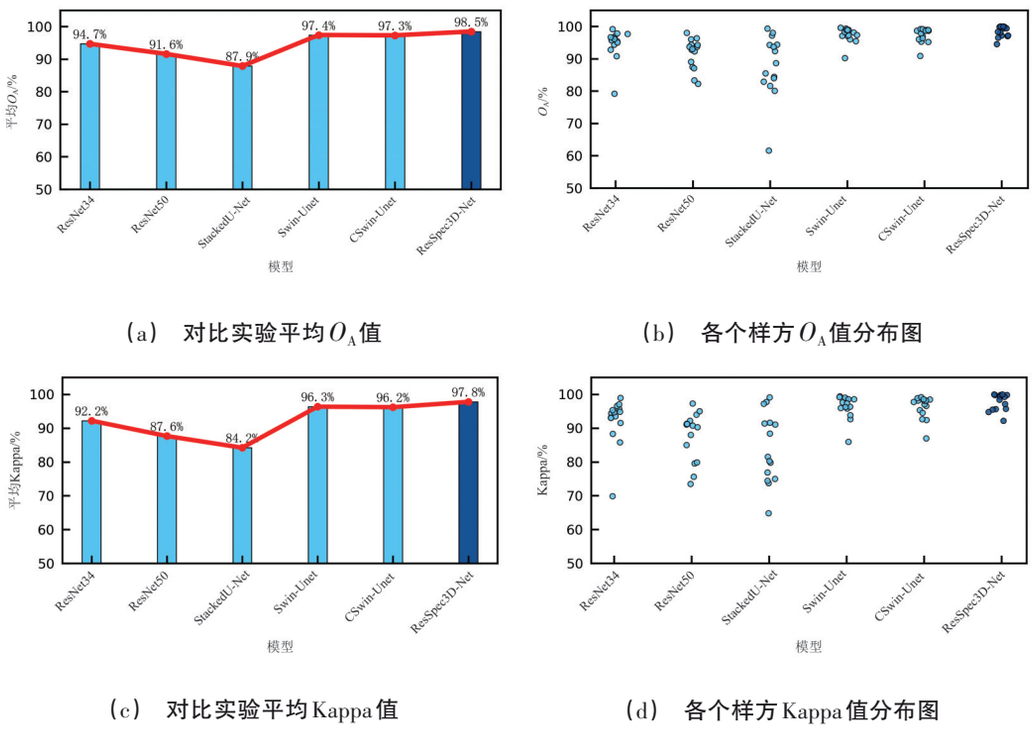
Table 3　Average classification accuracy of different models across 15 plots

单位:%

评价指标	ResNet34	ResNet50	StackedU-Net	Swin-Unet	CSWin-UNet	ResSpec3D-Net
O_A	94.7±4.8	91.6±1.2	87.9±5.7	97.4±0.3	97.3±0.4	98.5±0.4
Kappa	92.2±6.7	87.6±1.8	84.2±7.6	96.3±0.4	96.2±0.6	97.8±1.0

Net模型在部分样方出现了精度极低的情况,其原因在于该模型的编码器-解码器结构对复杂地物的特征提取能力有限,且缺乏有效的多尺度特征融合机制,导致在少量光谱信息下易产生误判。而Swin-Unet和CSWin-UNet模型虽然通过Transformer注意力机制提升了局部特征捕捉能力,使得在多数样方下表现较好,但不同样方之间分类也存在一定的波动,反映出这两种模型对全局特征的整合能力仍有不足,泛化能力较为一般。

综合分析表明,ResSpec3D-Net模型在所有样方中均保持了较高且相对集中的分类精度,体现出较强的鲁棒性与泛化能力,适用于荒漠草原复杂地物



（a）　对比实验平均O_A值　　（b）　各个样方O_A值分布图

（c）　对比实验平均Kappa值　　（d）　各个样方Kappa值分布图

图6　对比试验O_A和Kappa评价图

Fig. 6　O_A and Kappa evaluation results of comparative experiments

类型的高光谱分类任务。

为深入探讨ResSpec3D-Net模型在不同地物类别上的分类精度表现，本文在前述多种骨架模型的基础上，采用F_1分数对荒漠草原中不同地物类型的分类效果进行了评估。各模型在4类典型地物上的平均F_1分数如表4和图7所示。

表4　不同模型对各类地物的F_1分数分类精度

Table 4　F_1-scores of different models for each land-cover class　　单位：%

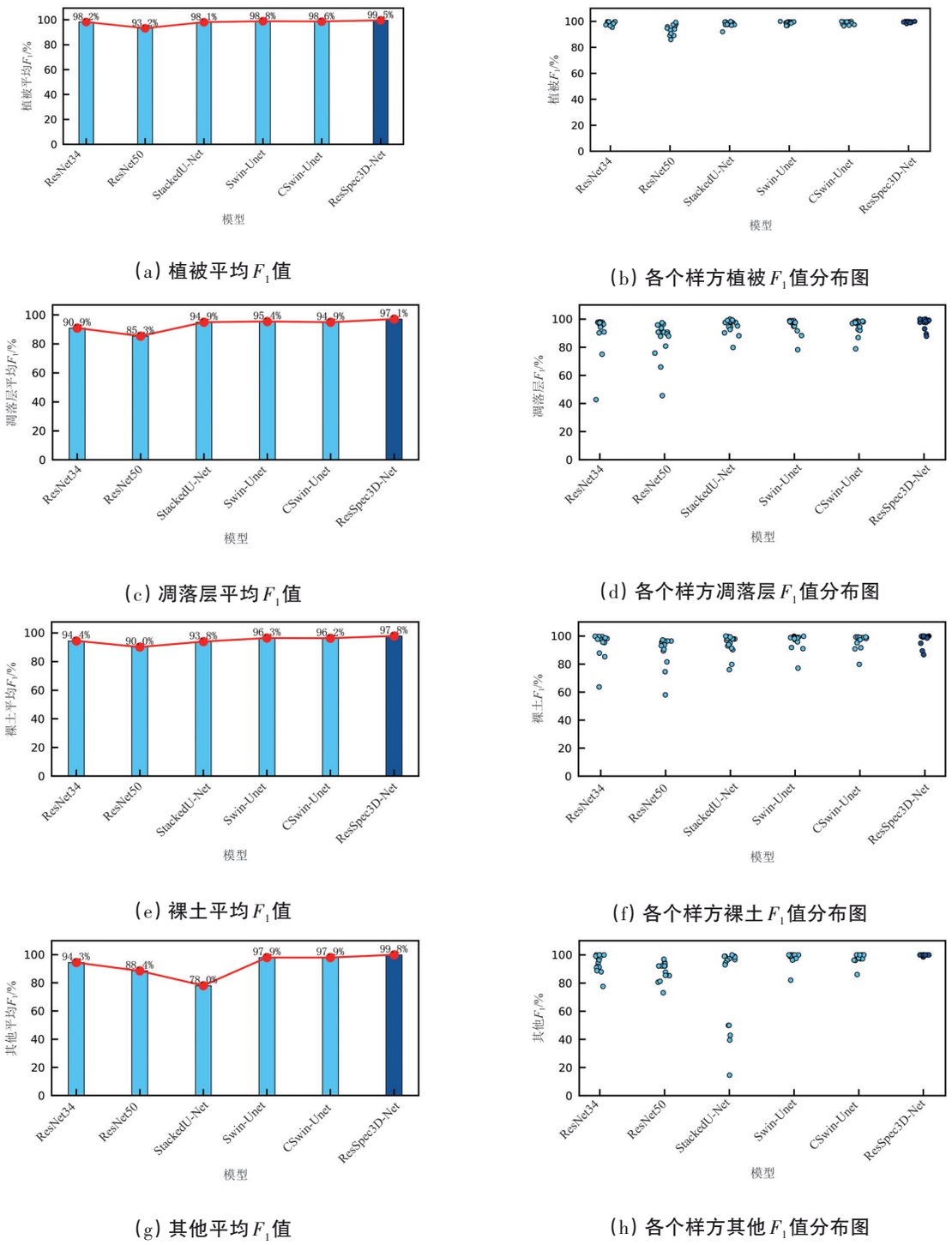
地物	ResNet34	ResNet50	StackedU-Net	Swin-Unet	CSWin-UNet	ResSpec3D-Net
植被	98.2±1.2	93.2±3.4	98.1±0.9	98.8±0.6	98.6±0.7	99.5±0.5
凋落层	90.9±4.1	85.3±5.1	94.9±6.2	95.4±4.2	94.9±3.5	97.1±3.2
裸土	94.4±4.7	90.0±5.7	93.8±4.2	96.3±1.7	96.2±3.4	97.8±3.0
其他	94.3±3.2	88.4±7.7	78.0±7.4	97.9±0.4	97.9±1.7	99.8±0.2

实验证明，针对本研究构建的荒漠草原高光谱数据集，基于ResSpec3D-Net的地物分类模型在不同类型地物上的分类精度优于其他基于U-Net的变体模型。对比分析图7可知，4类地物在不同模型与样方组合中分类效果存在一定差异，但ResSpec3D-Net始终在各类地物的识别中表现出更高的分类准确率与更强的鲁棒性，特别是在“凋落层”与“裸土”这两类边界模糊的地物中，F_1分数分别达到97.1%和97.8%，相较于其他模型最高提升了11.8%和7.8%，充分展现了边缘感知函数与多尺度特征融合结构对于边缘感知能力有了大幅度的提升。StackedU-Net在其他类别下展现出极差的分类能力，体现出对其他类别的识别能力不足，而ResNet50、ResNet34在不同样方间的分类效果波动明显，泛化能力较为一般；Swin-Unet与CSWin-UNet虽在植被类地物上具有较高精度，但在非植被类，尤其是“其他”类别上表现出明显的稳定性欠缺。

综上所述，ResSpec3D-Net在保证整体分类精度的同时，通过边缘感知损失函数显著提升了多类复杂地物的判别能力，尤其在弱监督条件下仍能保持优异表现，验证了其在实际荒漠草原遥感分类任务中的适用性与领先性。

4　讨论与结论

针对荒漠草原高光谱分类像元标注困难、光谱



(a) 植被平均F_1值　(b) 各个样方植被F_1值分布图

(c) 调落层平均F_1值　(d) 各个样方调落层F_1值分布图

(e) 裸土平均F_1值　(f) 各个样方裸土F_1值分布图

(g) 其他平均F_1值　(h) 各个样方其他F_1值分布图

图7　不同地物的F_1分数评价图

Fig. 7　F_1-score evaluations of different land-cover classes

相似类难区分、地物边界模糊的问题，本文基于弱监督学习框架，提出了具备边缘感知-多尺度特征融合的高光谱地物分类ResSpec3D-Net模型，并利用O_A、Kappa和F_1等相关参数验证了模型的可行性和优越性，通过对实验结果的分析，得出以下结论：

①从模块功能来看，Res2Net模块与3D-LDC Block模块的协同作用是解决“标注依赖”与“特征不足”矛盾的关键，二者结合可高效提取更丰富的多尺度光谱特征，突破了传统模型需依赖大量人工标注像元才能获取充足特征的局限，尤其在荒漠草原像元标注成本高、样本稀缺的场景下，该特征提取机制大幅降低了对标注数据量的需求，为弱监督分类提供了核心技术支撑。而边缘感知损失函数则针对性解决了地物边界模糊问题：一方面，其通过

强化边缘像素的特征差异,有效缓解了荒漠草原中“凋落层—植被”等边界的平滑效应,提升了边缘分类精度;另一方面,该损失函数通过优化梯度更新方向,加快了模型收敛速度。而三者的共同协调,保证了模型的强泛化能力与高精度。

②在仅 1% 像元标注的弱监督条件下,ResSpec3D-Net与5种U-Net变体模型(ResNet34、ResNet50、StackedU-Net、Swin-Unet、CSWin-UNet)的对比实验表明,在全局分类指标(O_A、Kappa)方面,ResSpec3D-Net在少量标注数据支撑下,仍保持了高精度与高稳定性。从细分地物的F_1指标来看,ResSpec3D-Net在4类地物(植被、凋落层、裸土、其他)中均展现出高效的分类能力与鲁棒性,即使在“凋落层”(F_1=97.1%)、“裸土”(F_1=97.8%)这类光谱相似、边界模糊的地物分类中,其精度仍远超对比模型,且在“其他”类别中,F_1值的样方波动也为最小且没有类似于StackedU-Net模型的情况,也从其他方面进一步说明了ResSpec3D-Net对各类地物的识别能力之强。

③在凋落层和裸土这两类复杂地物的分类中,ResNet-50模型在连续的残差网络下易产生欠拟合导致在光谱相似的复杂地物分类中精度较差,在其他地物的分类中,StackedU-Net模型因为其他地物的光谱曲线波动较大且标注像元较少导致极差的分类能力。而ResSpec3D-Net通过多尺度特征融合机制,可基于有限标注样本挖掘更多隐含特征,减少对标注数据的依赖,即使在低标注条件下仍能保持各类地物的高分类精度。

④本文研究只停留在了静态样方下的荒漠草原地物分类研究,未来研究可进一步融合连续时间序列高光谱信息和更多的气象数据,探索荒漠草原信息的完整融合建模路径,以增强模型对地表动态演变过程的感知能力,推动草原生态监测从单时相分类向动态监测过渡。

参考文献:

[1] 祁应军. 不同草场产权的界定与实施:基于产权交易成本视角的对比分析[J]. 中国农村经济,2021(11):55-71.

[2] CLOSE O, BENJAMIN B, PETIT S, et al. Use of sentinel-2 and LUCAS database for the inventory of land use, land use change, and forestry in wallonia, Belgium[J]. Land, 2018, 7(4): 154.

[3] 李　娜. 放牧管理下氮磷添加对典型草原群落结构和功能的影响[D]. 呼和浩特:内蒙古大学,2023.

[4] SAMARA E SILVA MEDEIROS E, MACHADO C C C, GALVÍNCIO J D, et al. Predicting plant species richness with satellite images in the largest dry forest nucleus in South America[J]. Journal of Arid Environments, 2019, 166: 43-50.

[5] RAM B G, ODUOR P, IGATHINATHANE C, et al. A systematic review of hyperspectral imaging in precision agriculture: Analysis of its current state and future prospects[J]. Computers and Electronics in Agriculture, 2024, 222: 109037.

[6] BORLAF-MENA I, SANTORO M, VILLARD L, et al. Investigating the impact of digital elevation models on sentinel-1 backscatter and coherence observations[J]. Remote Sensing, 2020, 12(18): 3016.

[7] CHEN T, ZHENG H, CHEN J, et al. Novel intelligent grazing strategy based on remote sensing, herd perception and UAVs monitoring[J]. Computers and Electronics in Agriculture, 2024, 219: 108807.

[8] 苏远超,许若晴,高连如,等. 基于深度学习的高光谱遥感图像混合像元分解研究综述[J]. 遥感学报,2024,28(1):1-19.

[9] PAHEDING S, SALEEM A, SIDDIQUI M F H, et al. Advancing horizons in remote sensing: a comprehensive survey of deep learning models and applications in image classification and beyond[J]. Neural Computing and Applications, 2024, 36(27): 16727-16767.

[10] 杨红艳,杜健民,阮培英,等. 基于无人机遥感与随机森林的荒漠草原植被分类方法[J]. 农业机械学报,2021,52(6):186-194.

[11] 佘　洁,沈爱红,石　云,等. 基于无人机遥感影像和面向对象技术的荒漠草原植被分类[J]. 草业学报,2024,33(7):1-14.

[12] JIANG H L, WU R G, ZHANG Y A, et al. Classification model of grassland desertification based on deep learning[J]. Sustainability, 2024, 16(19): 8307.

[13] JIN E, DU J M, BI Y G, et al. Research on classification of grassland degeneration indicator objects based on UAV hyperspectral remote sensing and 3D_RNet-O model[J]. Sensors, 2024, 24(4): 1114.

[14] 杨　星,方乐缘,岳　俊. 高光谱遥感影像半监督分类研究进展[J]. 遥感学报,2024,28(1):20-41.

[15] YANG L, CHEN X, ZHANG Q, et al. UAV hyperspectral crop classification based on deep learning: challenges of limited labeled samples [J]. Neurocomputing, 2023, 538: 34-45.

[16] WANG X, LI Y, ZHOU Y, et al. Hyperspectral image classification: A review of recent advances and future trends[J]. Remote Sensing, 2023, 15(3): 675.

[17] ZHANG H, LIU Q, WANG Y, et al. WHU-Hi: A large-scale benchmark dataset for hyperspectral image classification [J]. Remote Sensing, 2024, 16(12): 2258.

[18] WU P, LIU J, LI M M, et al. Fast sparse coding networks for anomaly detection in videos[J]. Pattern Recognition, 2020, 107: 107515.

[19] LIN Y P, VOSSELMAN G, CAO Y P, et al. Active and incremental learning for semantic ALS point cloud segmentation[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 169: 73-92.

[20] ZHOU Z H. A brief introduction to weakly supervised learning[J]. National Science Review, 2018, 5(1): 44-53.

[21] 张 涛,杜建民. 基于无人机遥感的荒漠草原微斑块识别研究[J]. 广西师范大学学报(自然科学版), 2022,40(6):50-58.

[22] 张燕斌,杜健民,王 圆,等. 基于无人机高光谱遥感和3D-ResNet的荒漠草原地物分类[J]. 中国农机化学报,2022,43(4):66-73.

[23] 刘燕德,王 舜. 基于图像和光谱融合的脐橙货架期高光谱成像无损检测研究[J]. 光谱学与光谱分析,2022,42(6):1792-1797.

[24] 康拥朝,毕玉革,杜建民,等. 基于高光谱MNF-SVM法荒漠化草原地表微斑块识别研究[J]. 内蒙古农业大学学报(自然科学版),2021,42(6):76-81.

[25] HANG R L, ZHOU F, LIU Q S, et al. Classification of hyperspectral images *via* multitask generative adversarial networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(2): 1424-1436.

[26] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43: 652-662.

[27] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]//NAVAB N, HORNEGGER J, WELLS W, et al. Medical Image Computing and Computer-Assisted Intervention. MICCAI 2015. Cham: Springer, 2015: 234-241.

[28] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//IEEE Computer Society. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015: 3431-3440.

[29] 陈龙伟,周小成,李传昕,等. 基于UNet-ResNet14*半监督学习的无人机影像森林树种分类[J]. 农业工程学报,2024,40(1):217-226.

[30] ASHRAF M, ZHOU X C, VIVONE G, et al. Spatial-spectral BERT for hyperspectral image classification [J]. Remote Sensing, 2024, 16(3): 539.

[31] LIU D X, HAN G L, LIU P X, et al. A discriminative spectral-spatial-semantic feature network based on shuffle and frequency attention mechanisms for hyperspectral image classification[J]. Remote Sensing, 2022, 14(11): 2678.

[32] KANG B, PARK I, OK C, et al. ODPA-CNN: one dimensional parallel atrous convolution neural network for band-selective hyperspectral image classification[J]. Applied Sciences, 2022, 12(1): 174.

[33] REIS-PEREIRA M, TOSIN R, MARTINS R, et al. Kiwi plant canker diagnosis using hyperspectral signal processing and machine learning: detecting symptoms caused by *Pseudomonas syringae* pv. *actinidiae*[J]. Plants, 2022, 11(16): 2154.