doi:10.6041/j.issn.1000-1298.2025.10.038

基于边缘增强的双分支语义分割网络地块识别方法

张宏鸣[1,2]　高郑杰[1]　沈寅威[1]　唐恒翱[1]　陈　诺[1]　阳　光[1]

(1.西北农林科技大学信息工程学院,陕西杨凌 712100;
2.陕西省农业信息智能感知与分析工程技术研究中心,陕西杨凌 712100)

摘要:快速准确地进行目标区域的土地分割对于土地资源规划和利用具有重要意义,针对土地覆盖分割任务中存在的地物类型多样和边界区分不清晰的问题,本研究提出一种边缘增强的双分支语义分割网络地块识别方法(Edge-enhanced dual-branch net, E2DBNet),在编码器中设计了一种边缘增强模块,将影像中额外增强的精细化边缘信息引入语义分割模型中构成双分支网络,在解码器中通过特征聚合模块高效融合全局语义特征和局部边缘特征,同时自底向上跨层连接融合特征,强调关键空间位置来进一步细化局部细节。在构建的杨陵区土地覆盖分割数据集上进行训练和测试,对比实验结果显示,本文提出的 E2DBNet 在不同场景下均取得了最好的分割精度,交并比、F1 分数分别为 67.73%、93.25%,相较于对比的分割模型分别提升 2.36、3.7 个百分点,同时通过消融实验验证了各个模块的作用。E2DBNet 在模型参数量较少的情况下,更好地关注了样本数少的地物类型,并精准预测出难分类别的主体,同时由于引入了增强的边缘信息,对各类地物之间边界区分更加清晰,为场景复杂的土地覆盖分割提供了有效的解决思路。

关键词:土地覆盖;语义分割;深度学习;边缘检测;特征融合

中图分类号:TP751;S127　　**文献标识码**:A　　**文章编号**:1000-1298(2025)10-0437-11　　**OSID**:

Dual-branch Semantic Segmentation Network-based Land Parcel Recognition Method with Edge Enhancement

ZHANG Hongming[1,2]　GAO Zhengjie[1]　SHEN Yinwei[1]　TANG Heng'ao[1]　CHEN Nuo[1]　YANG Guang[1]

(1. *College of Information Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*
2. *Shaanxi Engineering Research Center for Intelligent Perception and Analysis of Agricultural Information*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: Accurate and rapid acquisition of land cover data in target areas is crucial for land resource planning and utilization. To address the challenges posed by diverse land cover types and indistinct boundaries in land cover segmentation tasks, an innovative edge-enhanced module within the encoder was introduced. This module was part of a novel edge-enhanced dual-branch semantic segmentation network method, named edge-enhanced dual-branch net (E2DBNet). The proposed E2DBNet integrated finely enhanced edge information into the semantic segmentation model, forming a dual-branch network structure. In the decoder phase, a feature aggregation module was employed to efficiently merge global semantic features with local edge features. Additionally, cross-layer connections from the bottom to the top layers further refined local details by emphasizing critical spatial positions, enhancing the overall accuracy and clarity of the segmentation. The effectiveness of E2DBNet was demonstrated through training and testing on the constructed Yangling segmentation dataset. Comparative experimental results indicated that E2DBNet achieved superior segmentation accuracy across various scenarios, with intersection over union (IoU) and F1 scores of 67.73% and 93.25%, respectively. These results represented improvements of 2.36, 3.7 percentage points, over the compared segmentation models. Furthermore, ablation experiments validated the effectiveness of each individual module within the network. Despite having fewer model parameters, E2DBNet demonstrated a remarkable ability to accurately identify land

收稿日期:2024-06-25　修回日期:2024-08-06

基金项目:国家自然科学基金项目(42377341)、陕西省重点研发计划项目(2023-YBNY-217、2023-ZDLNY-69)、西安市科技计划项目(23NYGG0074)和陕西省秦创原队伍建设项目(2023-ZDLNY-69)

作者简介:张宏鸣(1979—),男,教授,博士,主要从事空间大数据管理与智慧农业研究,E-mail:zhm@nwafu.edu.cn

cover types even with fewer samples, and effectively predicted the main parts of difficult categories. The incorporation of enhanced edge information also ensured clearer boundary distinctions between various land cover types, providing a robust solution for complex land cover segmentation tasks. This advancement held significant potential for improving the precision and efficiency of land cover mapping, ultimately contributing to better land resource management and planning.

Key words: land cover; semantic segmentation; deep learning; edge detection; feature fusion

0 引言

语义分割是当前计算机视觉中极受关注的研究领域之一,通过分析图像中每个像素的语义属性来计算其所属类别的概率,进而将具有相同语义属性的像素划分为同一类,实现对图像目标区域的精确划分[1]。相较于常见的自然图像语义分割任务,遥感影像包含了更多的光谱信息及空间信息,这使得遥感影像语义分割任务更为复杂,同时更富有挑战性[2]。经过学者们的长期研究,遥感影像语义分割领域已有大量的方法及理论,目前可以大致分为基于光谱的分割方法、基于空间的分割方法和基于深度学习的分割方法[3]。

基于光谱的分割方法侧重于利用遥感影像中反映地物特性的光谱信息进行分割[4-5]。但此类分割方法只关注了遥感影像的光谱属性,而没有充分利用纹理和几何特性,并且影像中会出现“同物异谱”和“同谱异物”的现象[6-7],这些因素都使得该方法的分割效果有待改进。

基于空间的分割方法主要利用遥感影像中的空间结构信息,包括像素之间的纹理特征、空间关系以及形态学特征等,可以有效弥补光谱信息不足的问题[8-9]。基于空间的分割方法相较基于光谱的分割方法可以更好地反映地物的空间结构和形态特征,但是其中的一些方法涉及参数的选择,会对分割效果产生极大的影响,使用场景受限。因此,需要一种智能化的处理算法,能够自适应地提取不同地物特征,解决复杂背景下的遥感影像多类土地覆盖分割问题[10]。

近年来,计算机视觉领域快速发展,卷积神经网络(Convolutional neural network, CNN)[11]被广泛应用于各行各业中,在图像分类问题中展现了巨大潜力,之后学者们也陆续地提出了许多经典网络结构,如 VGGNet[12]、ResNet[13]和 DenseNet[14]等。同时,全卷积网络(Fully convolutional network, FCN)[15]采用反卷积上采样替代全连接层,实现了像素级语义分割。为了进一步解决 FCN 在特征提取中细节信息丢失的问题,Unet[16]和 SegNet[17]网络结构相继提出,分别通过反池化操作和跳跃连接结构,恢复高级编码特征的完整性。CHEN 团队[18-20]自 2014 年提出了 DeepLab 系列,DeepLabV1 采用空洞卷积捕捉上下文信息,DeepLabV2 结合残差网络和空洞空间金字塔,DeepLabV3 + 加入了深度可分离卷积和跨层连接,这一系列的改进为深度学习语义分割领域提供了新的思想,方法理论得到了进一步完善。

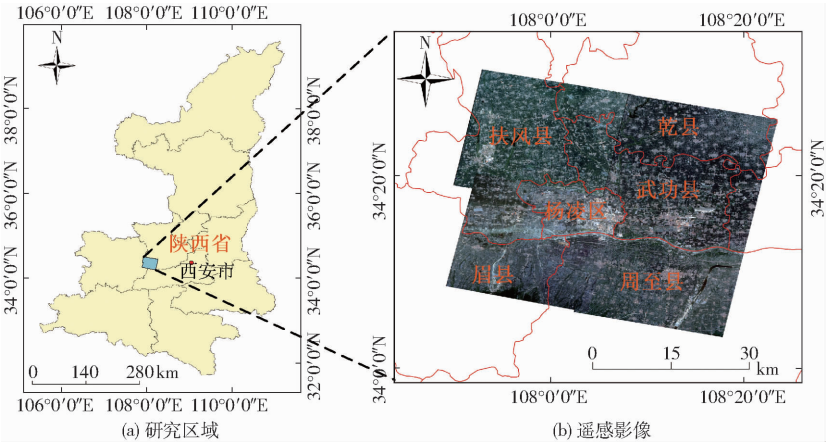
鉴于深度学习技术出色的特征提取能力,在不依赖先验知识的前提下,仍能高精度地对自然图像进行分割[21]。不同于常见的自然图像,遥感影像中波段更多,光谱信息更加丰富,如何充分提取遥感影像中的特征信息成为难点之一[22-23]。目前,遥感影像深度学习分割方法通常从融合多源信息和网络结构优化两个方面来解决影像中存在的地物对象尺寸不一、边界区分不够清晰的问题[24]。融合多源信息方法并未充分融合边缘信息和语义信息,分割结果的边界较为模糊[25-27],网络结构优化模型的整体复杂度仍偏高,其复杂度有待缩减[28-30]。此外,相较于基于光谱和基于空间的分割方法,基于深度学习的分割方法更加智能和自动化,极大程度上减少了对专家知识和人工干预的依赖,并保证了分割结果的质量及精度,但是在边缘信息的关注以及模型结构优化上仍需进一步研究。

综上所述,本文针对土地覆盖分割任务中存在的地物类型多样和边界区分不清晰的问题,构建一种双分支语义分割网络,通过修改编码器结构,并引入额外的边缘辅助信息以提高相邻地物边界的连续性,并融合语义及边缘特征以加强信息交互,同时修改损失函数缓解类别不平衡带来的影响,以提升分割模型在处理各种复杂遥感影像时的准确性。

1 研究区域和数据概况

1.1 研究区域

本研究区域位于 107°40′ ~ 108°30′E,34°00′ ~ 34°40′N,涵盖了杨凌区及其周边县城,如扶风县、乾县、武功县、眉县和周至县等。杨凌区地处陕西关中平原中部,隶属于陕西省,位置示意图如图 1 所示。该区域属于半湿润半干旱气候区,地形呈现出北高南低、西高东低的阶梯状,海拔 418 ~ 540 m,年平均降水量 635 mm,平均气温 13℃,是我国第 1 个国家级农业高新区,拥有发展农业高科技产业的优越环境,并始终致力于构建“农业硅谷”。研究区内地



(a) 研究区域　　(b) 遥感影像

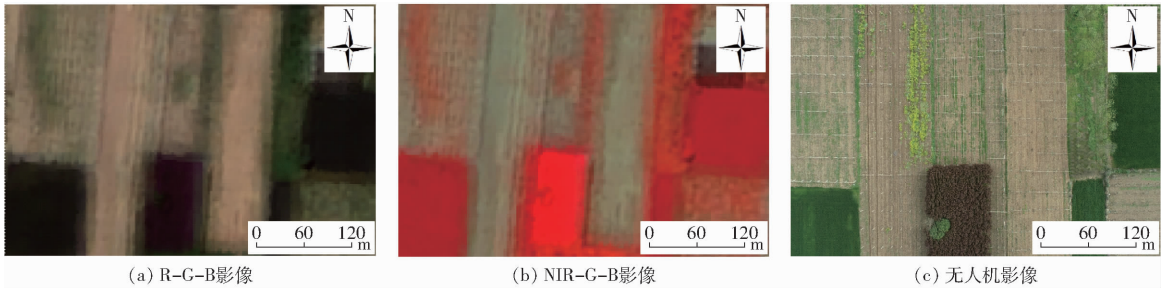
图 1　研究区示意图

Fig. 1　Study area diagrams

物类型多样，其中优质农作物包含玉米、小麦、油菜、蔬菜、西瓜等，果树包括桃子、葡萄、苹果、猕猴桃等，还涵盖沙棘、紫叶矮樱等各种植物种类。然而这些地物分布呈现出明显的零散特性，边界区分并不清晰，地块之间的过渡区域常充满各类杂草和树木，这些自然和人为因素造成的混杂使得土地覆盖类型的自动识别和分割变得更加复杂。除了农业作物和果树外，研究区内还存在着多样化的非农用地，如建筑、水体以及裸土地等，进一步增加了土地覆盖分割难度。同时，各种地物类型之间存在不平衡现象，某些特定类型的水体样本，相较于耕地而言数量较少，给深度学习分割算法带来了极大的挑战。

1.2　数据集构建

实验采用 4 幅高分二号遥感影像，统一成像于 2021 年 5 月 9 日，每幅影像尺寸为 7 300 像素 × 6 908 像素，边缘部分略有重叠。经过前文所述的影像预处理后，得到分辨率 1 m 的多光谱融合影像，为确保所构建数据集质量，本次数据标注是基于前期实地调研的，使用外业精灵软件在待考察地点进行 GPS 标点，再前往实地考察，通过遥感影像与现场无人机影像比对，如图 2 所示，可以更为清晰准确地确定遥感影像上具体地物类型，后期再由专业团队对 4 幅影像进行标注。



(a) R-G-B影像　　(b) NIR-G-B影像　　(c) 无人机影像

图 2　影像对比示意图

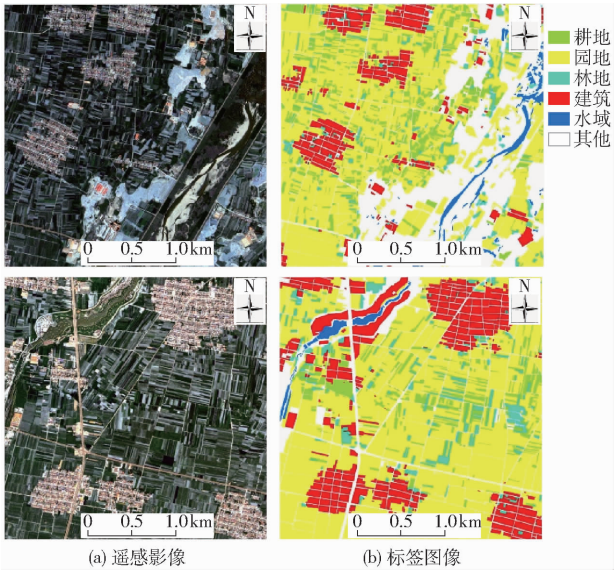
Fig. 2　Schematics of image comparison

根据《土地利用现状分类》和影像所在区域的土地类型，以及后续研究的需要，本研究将研究区域内地物类型分为耕地、园地、林地、建筑、水域和其他背景 6 种类别，标注过程使用 ArcGIS 软件创建矢量图层，选定地块区域并于属性表中添加具体“标签值”字段，地物类型与标签值对应关系如表 1 所示。将标注好的矢量标签转换为 RGB 图像后，与对应区域影像配对，如图 3 所示。使用 GDAL 库将对应影像和标签图通过重复率为 0.3 的滑动窗口，裁剪成 512 像素 ×512 像素，最终通过数据增强得到 5 810 个土地覆盖分割影像-标签对，将其按照比例 8∶1∶1 划分成训练集、验证集和测试集，训练集包含 4 648 个影像-标签对，验证集和测试集各包含 581 个影像-标签对，形成最终的杨凌区土地覆盖分割数据集（Yangling Segmentation Dataset，YLSD）。

表 1　地物类型-标签值对应表

Tab. 1　Feature type-label value correspondence

地物类型	标签值	转换 RGB 值
耕地	0	152,230,0
园地	1	255,255,0
林地	2	0,255,197
建筑	3	255,0,0
水域	4	0,112,255
其他	5	255,255,255



(a) 遥感影像 (b) 标签图像

图 3 分割影像-标签对示意图

Fig. 3 Schematics of segmentation image-label pair

2 方法

2.1 模型结构

在对目标区域内的土地覆盖类型进行精准解译的过程中,非常依赖影像中的地块几何结构,但是背景复杂的多光谱影像往往呈现出内部特征的高度聚合,导致不同地物类型之间的边界区分并不清晰,使得最终分割结果并不精确,存在较大的误差。因此,本文提出了一种基于边缘增强的双分支地块分割网络(Edge-enhanced dual-branch net, E2DBNet),总体结构如图 4 所示。

在编码器部分,采用双分支 ResNext50_32×4d 网络组成,网络结构如表 2 所示,采用分组卷积分别对原始图像以及经边缘增强模块(Edge-enhanced module, EEM)处理过的图像进行特征提取。通过编码器中的光谱信息提取模块来提取遥感影像中不同地物的光谱特征和纹理特征,通过编码器中的边缘增强模块提取影像中的边缘信息,为分割网络输入精细化边缘特征以作为辅助信息,得到了包含丰富语义信息和边缘信息的特征图。考虑到双分支特征提取模块会导致模型的权重参数翻倍,本研究采用共享权重参数的双分支模型,在不增加模型复杂度的同时,提高模型对边缘信息的特征提取能力。在解码器部分,通过构建的特征聚合模块(Feature aggregation module, FAM)将两分支输出的深层特征信息进行融合,并且通过跨层连接模块(Cross-layer connection module, CLCM)进一步整合编码阶段中不同尺度的浅层特征图,采用这种自底向上的结构综合考虑了不同层次的信息,从而提高土地覆盖分割的准确性和边缘定位的精确度。通过编码器的双分支网络进行特征提取,同时通过解码器中的特征聚合和跨层连接模块进一步指导信息的融合,并且使用混合损失函数缓解了类别不平衡的现象,进一步提升了分割模型在处理各种复杂遥感影像时的准确性和鲁棒性。

2.2 基础模型改进

2.2.1 边缘增强模块

遥感影像中各类地物的边缘信息是其重要的结构特征,对于像素级别的分类任务至关重要,有助于提高分割结果准确率以及边缘连续性,这是因为在

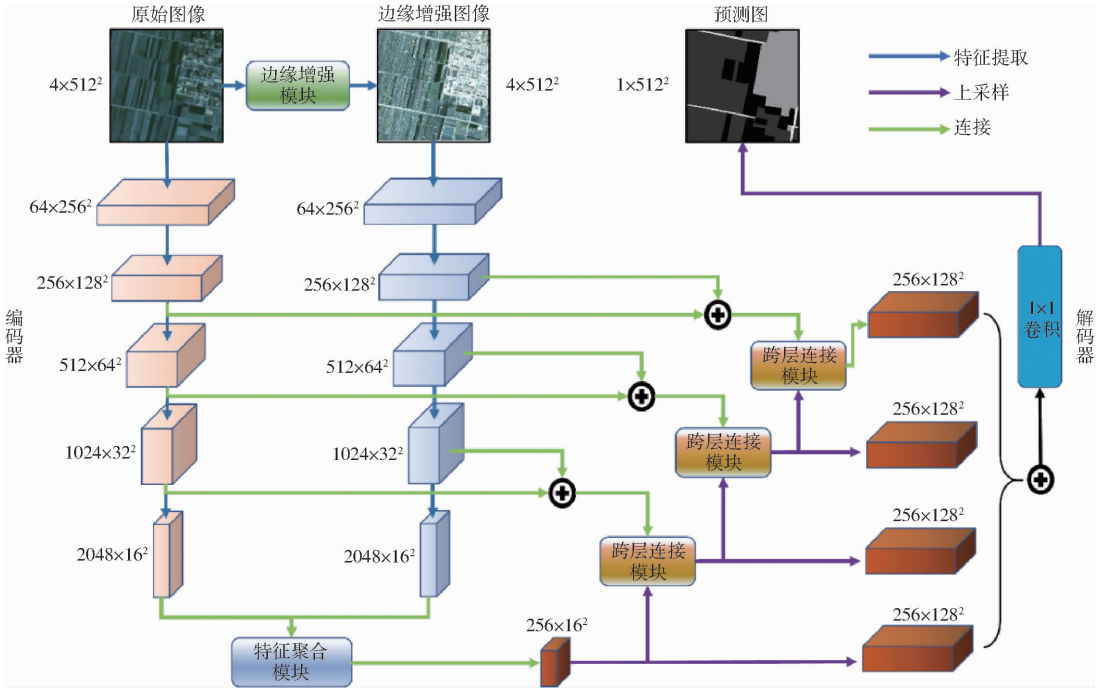


图 4 E2DBNet 结构示意图

Fig. 4 Schematic of E2DBNet structure

表 2　ResNext50_32 × 4d 网络结构

Tab. 2　Structure of ResNext50_32 × 4d

阶段	输出尺寸	结构
Conv1	256 × 256	7 × 7,64,步长 2
		3 × 3 最大池化,步长 2
Conv2	128 × 128	$\begin{bmatrix}1\times1,128\\3\times3,128,C=32\\1\times1,256\end{bmatrix}\times3$
Conv3	64 × 64	$\begin{bmatrix}1\times1,256\\3\times3,256,C=32\\1\times1,512\end{bmatrix}\times4$
Conv4	32 × 32	$\begin{bmatrix}1\times1,512\\3\times3,512,C=32\\1\times1,1\,024\end{bmatrix}\times6$
Conv5	16 × 16	$\begin{bmatrix}1\times1,1\,024\\3\times3,1\,024,C=32\\1\times1,2\,048\end{bmatrix}\times3$

遥感影像中地物边缘定义了不同特征及地物类型分界,对于识别相邻但类型不同的地物(如耕地和林地、建筑和空地等)极为关键。然而常见的语义分割模型在特征提取阶段,对原始的遥感影像在空间上进行了压缩,特征图分辨率的降低导致局部的边缘信息丢失。因此本研究设计了一种边缘增强模块(EEM),将额外增强的边缘信息引入语义分割模型中,构成双分支网络,一方面提升了原有模型复杂度,优化其学习能力,另一方面丰富了多源信息的输入,在土地覆盖分割中细化了地块边界。

EEM 结构如图 5 所示,使用改进后的 Sobel 边缘检测算子,分别从 0°、45°、90°和 135°方向计算原始图像各个通道上的一阶导数,然后将其合成对应通道的边缘图像,并与原始图像叠加获得最终的边缘增强图像。单独采用 0°、45°、90°或 135°这 4 个固定方向的 Sobel 算子,仅能捕捉到特定方向的边缘信息,会造成非特定方向上边缘信息的丢失,因此改进的 Sobel 算子融合了这 4 个方向的梯度,提高对各个方向边缘提取的完整性。改进后的 Sobel 算子提取的边缘特征更加丰富,有效增强了细小地物的边缘,为原有的语义分割网络输入更多的判别信息,丰富了边缘特征提取分支。

2.2.2　特征聚合模块

E2DBNet 编码器中的语义分支负责提取高层次的全局语义特征,边缘分支专注于提取局部边缘特征,然而双分支并行结构缺乏特征间的交互,无法充分利用两者中有效的信息,因此在解码器中构建了特征聚合模块(FAM),以高效的融合策略聚合双分支中的信息,实现深层语义特征与精细边缘信息的有效结合。

FAM 结构如图 6 所示,首先通过将语义特征和

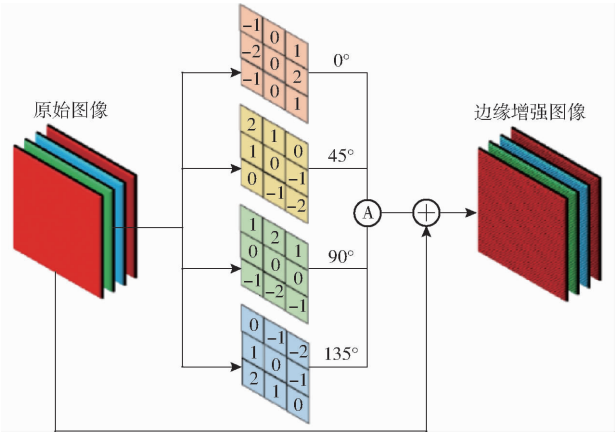


图 5　边缘增强模块(EEM)结构示意图

Fig. 5　Schematic of edge-enhanced module (EEM)

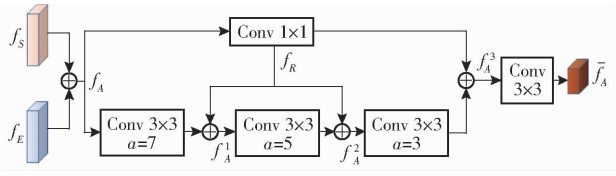


图 6　特征聚合模块(FAM)结构示意图

Fig. 6　Schematic of feature aggregation module (FAM)

边缘特征执行逐元素相加实现初步的特征融合,接着依次使用具有不同膨胀率 a(分别为 7、5、3)的空洞卷积来提取不同感受野的融合特征,并按照感受野从大到小依次串联,确保模型对于空间上下文的全面理解,使用残差连接强化初始融合特征的传递,多分支的架构使得融合特征可以更好地适应不同尺寸和形状的地物类型,模块计算过程描述为

$$f_A = f_S + f_E \tag{1}$$

$$f_R = \mathrm{Conv}_{1\times1}(f_A) \tag{2}$$

$$f_A^1 = \mathrm{Conv}_{3\times3}^{a=7}(f_A) + f_R \tag{3}$$

$$f_A^2 = \mathrm{Conv}_{3\times3}^{a=5}(f_A^1) + f_R \tag{4}$$

$$f_A^3 = \mathrm{Conv}_{3\times3}^{a=3}(f_A^2) + f_R \tag{5}$$

$$\bar{f}_A = \mathrm{Conv}_{1\times1}(f_A^3) \tag{6}$$

式中　f_A——初始融合特征

f_S——语义特征

f_E——边缘特征

f_R——残差连接特征

f_A^1、f_A^2、f_A^3——不同膨胀率空洞卷积融合特征

$\bar{f}_A$——最终输出融合特征

$\mathrm{Conv}_{1\times1}$——尺寸为 1 × 1 的卷积函数

$\mathrm{Conv}_{3\times3}^{a}$——膨胀率为 a、尺寸为 3 × 3 的空洞卷积

将编码器中双分支特征提取模块得到的语义特征f_S和边缘特征 f_E 采用逐元素相加的方式实现特征融合,采用尺寸为 1 × 1 的卷积对初步融合后的特征进行通道调整得到 f_R,同时依次使用具有不同膨胀

率(分别为 7、5、3)的空洞卷积来提取不同感受野的融合特征,最后使用尺寸为 1×1 的卷积对不同感受野的融合特征 f_A^3 进行通道调整,方便后续上采样操作。

2.2.3 跨层连接模块

FAM 在特征聚合时仅考虑最底层的语义和边缘特征,而通过上采样恢复分辨率会丢失浅层信息。为此,在上采样过程中增加跨层连接模块(CLCM),从底层到浅层汇集融合特征,并使用空间注意力机制(Spatial attention module, SAM)分配空间权重,突出重要区域,细化局部细节。

CLCM 结构如图 7 所示,首先将底层融合特征进行 2 倍上采样以匹配浅层的空间维度,并将浅层融合特征通过卷积函数压缩通道数,以减小计算复杂度并优化特征传输效率。接着执行逐元素加法操作整合不同层级的信息,随后通过 SAM 生成对应空间注意力图,计算特征中不同位置的重要性权重,并将其与原始特征进行逐元素乘法操作,强调关键空间位置,进一步增强模型对边缘及不同尺寸地物的感知能力,模型计算公式为

$$f'_A = \mathrm{Conv}_{1\times1}(f_A^{\mathrm{high}}) + \mathrm{Up}(f_A^{\mathrm{low}}) \tag{7}$$

$$f_A^S = \sigma(\mathrm{Conv}_{7\times7}([\mathrm{AvgPool}(f'_A);\mathrm{MaxPool}(f'_A)])) \tag{8}$$

$$f''_A = f_A^S \times f'_A \tag{9}$$

式中 f_A^{high}——浅层融合特征

f_A^{low}——底层融合特征

f'_A——初始连接特征

f_A^S——空间注意力图

f''_A——输出连接特征

Up——上采样操作

AvgPool——平均池化

MaxPool——最大池化

[ ]——拼接操作

σ——Sigmoid 激活操作

×——逐元素乘法

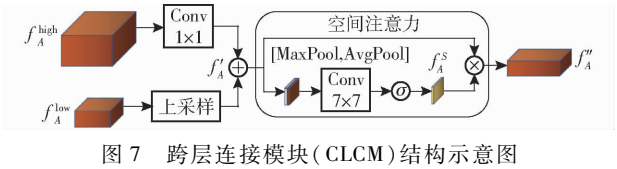


图 7 跨层连接模块(CLCM)结构示意图

Fig. 7 Schematic of cross-layer connection module (CLCM)

首先对浅层融合特征 f_A^{high} 使用尺寸为 1×1 的卷积核进行通道调整,对底层融合特征 f_A^{low} 使用 2 倍上采样操作进行特征图尺寸调整,使得浅层融合特征和底层融合特征具有相同的通道和尺寸,接着执行逐元素加法操作整合不同层级的信息得到 f'_A,随后对融合特征 f'_A 分别进行平均池化和最大池化操作,对池化结果进行拼接,然后使用尺寸为 7×7 的卷积核进行特征提取,特征提取结果通过 Sigmoid 激活操作最终生成含有空间权重的空间注意力图 f_A^S,将空间注意力图 f_A^S 和初始连接特征 f'_A 主元素相乘,得到空间注意力增强的输出连接特征 f''_A。

2.2.4 损失函数改进

由于真实地物分布中存在不同类型,各个类别的样本数量并不相同,这对分割模型提出更高要求,模型在训练中会更加偏向于多数样本和易分样本,从而忽略少数样本和难分样本的特征学习,因此本文通过使用混合损失函数,按照权重组合计算 Focal Loss(L_{Focal})和 Dice Loss(L_{Dice}),以缓解类别不平衡问题并优化分割质量,具体计算公式为

$$L_{\mathrm{Focal}} = -\alpha_t(1-p_t)^{\gamma}\ln p_t \tag{10}$$

$$L_{\mathrm{Dice}} = 1 - \frac{2\sum_{i=1}^{N} y_i\hat{y}_i + \varepsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \hat{y}_i + \varepsilon} \tag{11}$$

最终总体损失函数为

$$L_{\mathrm{Total}} = \lambda_1 L_{\mathrm{Focal}} + \lambda_2 L_{\mathrm{Dice}} \tag{12}$$

式中 p_t——分割模型预测概率

α_t——平衡权重因子,取 0.75

γ——调制因子,取 2

y_i——像素级别上真实标签值

$\hat{y}_i$——像素级别上预测标签值

N——总像素数

ε——常数,用于避免分母为 0

λ_1、λ_2——损失权重

L_{Total}——总体损失函数

本文设置 $\lambda_1 = \lambda_2 = 0.5$,确保模型在处理类别不平衡和区域重叠度方面都能得到一定的优化。

2.3 实验环境与指标

2.3.1 实验运行环境

土地覆盖分割模型编码器部分采用迁移学习,使用 ImageNet 数据集上的预训练权重。实验运行的硬件平台为 Intel(R) Xeon(R) Gold 5222 CPU@3.80 GHz,使用的操作系统是 Microsoft Windows 10 专业版,配备 NVIDIA GeForce RTX 3090 显卡,并在 Python 3.6 环境下,利用 Pytorch 深度学习框架进行开发。

实验超参数设置为:训练批次大小(Batch Size)为 8,使用单卡训练,训练最大轮次(Epoch)为 100,采用 ReduceLROnPlateau 策略动态调整学习率,初始学习率设置为 0.001,优化算法采用 Adam。

2.3.2 实验评价指标

为了评估不同模型的有效性，采用精确率(Precision)、召回率(Recall)、F1 分数和交并比(Intersection over union，IoU)4 种评价指标来评估模型性能。

2.4 实验方案

(1)对比实验

为验证本文提出的 E2BDNet 在土地覆盖分割中的性能，实验选取目前语义分割领域中主流及先进的检测模型进行对比，将 Unet、Unet + +、DeepLabV3 + 、UperNet、DCSwin 和 UnetFormer 在构建的杨凌区土地覆盖分割数据集上训练和测试，将它们的识别结果与本文 E2BDNet 模型进行对比，通过分割定性分析和定量分析验证本文模型性能。

(2)消融实验

模型改进主要包括添加 EEM 模块融入双分支网络，使用 FAM 模块融合语义特征及边缘特征，并通过 CLCM 模块连接不同层级的融合特征，为进一步验证每个结构对于最终结果的效果，进行了消融实验，分别去除一个模块及去除所有模块测试性能，同时也探究了不同损失函数对于模型训练产生的效果。

3 结果与分析

3.1 对比实验

本文提出的 E2DBNet 方法与其他 6 种语义分割方法，在本文构建的杨凌区土地覆盖分割数据集上的分割结果如图 8 所示，通过 1 幅遥感影像和 8 个分布情况不同的场景全方位地展示各类分割方法的实际效果，各类方法评价指标如表 3 所示。

由表 3 可得，本文提出的 E2DBNet 相较于现有语义分割模型在杨凌区土地覆盖分割数据集上取得了最好的结果。对比模型中 UnetFormer 参数量最少，仅为 2.308×10^7，交并比为 68.58%，但其召回率、F1 分数和精确率较低，说明其对整体样本位置和形状预测准确，但对于少量样本和难分样本检测能力较低，与定性结果相符，对背景较为复杂的场景检测性能较差。DeepLabV3 + 参数量次之，但其在本文数据集上识别效果最差，交并比为 61.20%、召回率为 89.80%、F1 分数为 86.31%、精确率为 87.92%。UPerNet、Unet、Unet + + 和 DCSwin 的参数量依次增加，检测结果也是逐步提升，DCSwin 交并比为 65.37%、召回率为 92.94%、F1 分数为 89.55%、精确率为 90.37%，可见模型结构复杂度的增加有利于提取更加丰富、表征性更强的特征，但是模型参数量的提升也意味着更高的计算成本。本文提出的 E2DBNet 采用双分支结构，通过权重共享在不增加模型参数量的前提下，将边缘信息引入语义分割中，丰富了多源特征的输入，并且通过 FAM 和 CLCM 模块高效地聚合特征，优化了信息的综合，以较少参数量实现了高精度的土地覆盖分割，交并比为 67.73%、召回率为 95.09%、F1 分数为 93.25%、精确率为 94.05%，相较于 DCSwin，交并比、召回率、F1 分数、精确率分别提升 2.36、2.15、3.7、3.68 个百分点。在具体类别上，E2DBNetF1 得分较高，这部分归因于采用混合损失函数，它增强了对于样本较少类别的关注，尤其在水域和其他背景的检测方面，这进一步验证了本文提出的 E2DBNet 在土地覆盖分割任务上的优越性。

通过结果可视化展示出 E2DBNet 模型与其他 6 种主流模型检测结果对比(图 8)，不同颜色表示模型分割的不同地物类型，从图 8 可得，本文提出的 E2DBNet 检测结果更加完整，边缘更加平滑，与实际标签最为接近。由于地物实际分布中水域面积占比最小，分割模型非常容易忽视，如图 8 场景 1、2 所示，Unet + +、DeepLabV3 + 、UPerNet 和 UnetFormer 都未能将影像中零散的水域部分分割出来，而本文方法很好地关注了样本数较少的水域部分，分割结果更加精细，并且整体识别效果更加完整。同时，由于各类地物之间边界区分不清晰，会导致分割结果中出现互相掺杂的现象，如图 8 场景 3 所示，Unet、Unet + + 和 UnetFormer 分割出的建筑粘连在一起，未能将交织在其中的道路分割出来，如图 8 场景 4 所示，DeepLabV3 + 、UPerNet 和 UnetFormer 未能精确划定细小河流的边界，导致分割结果呈现出河流断裂的现象，而本文方法通过引入边缘特征作为辅助信息，很好地避免了这一问题。此外，影像中地物类型丰富且分布复杂，为分割任务增加了难度，如图 8 场景 5、6 所示，在多样化场景下各类分割模型性能有限，误识别较多。并且对于林地与园地等辨识度较低的类别，各类分割模型难以有效区分，结果中出现了大面积的孔洞，如图 8 场景 7、8 所示，而本文 E2DBNet 充分聚合了语义特征和边缘特征，确保了全局语义与局部细节的协同增强，显著提升了分割结果的准确性和细节表现。

在研究区域内，耕地区域主要分布在居民区周围，形成了环绕居民区的农业带。园地区域则呈现出较为集中的分布，通常以较大面积的连片形式出现，为该区域提供了重要的果蔬生产地。建筑区域则按照村落的形式集中分布，形成了多个村庄和小型聚落，体现了该地区的人类聚居特点。林地主要大面积分布在研究区域的南部，形成了大片的森林覆盖区。此外，在离居民区较远的山地上也零散分

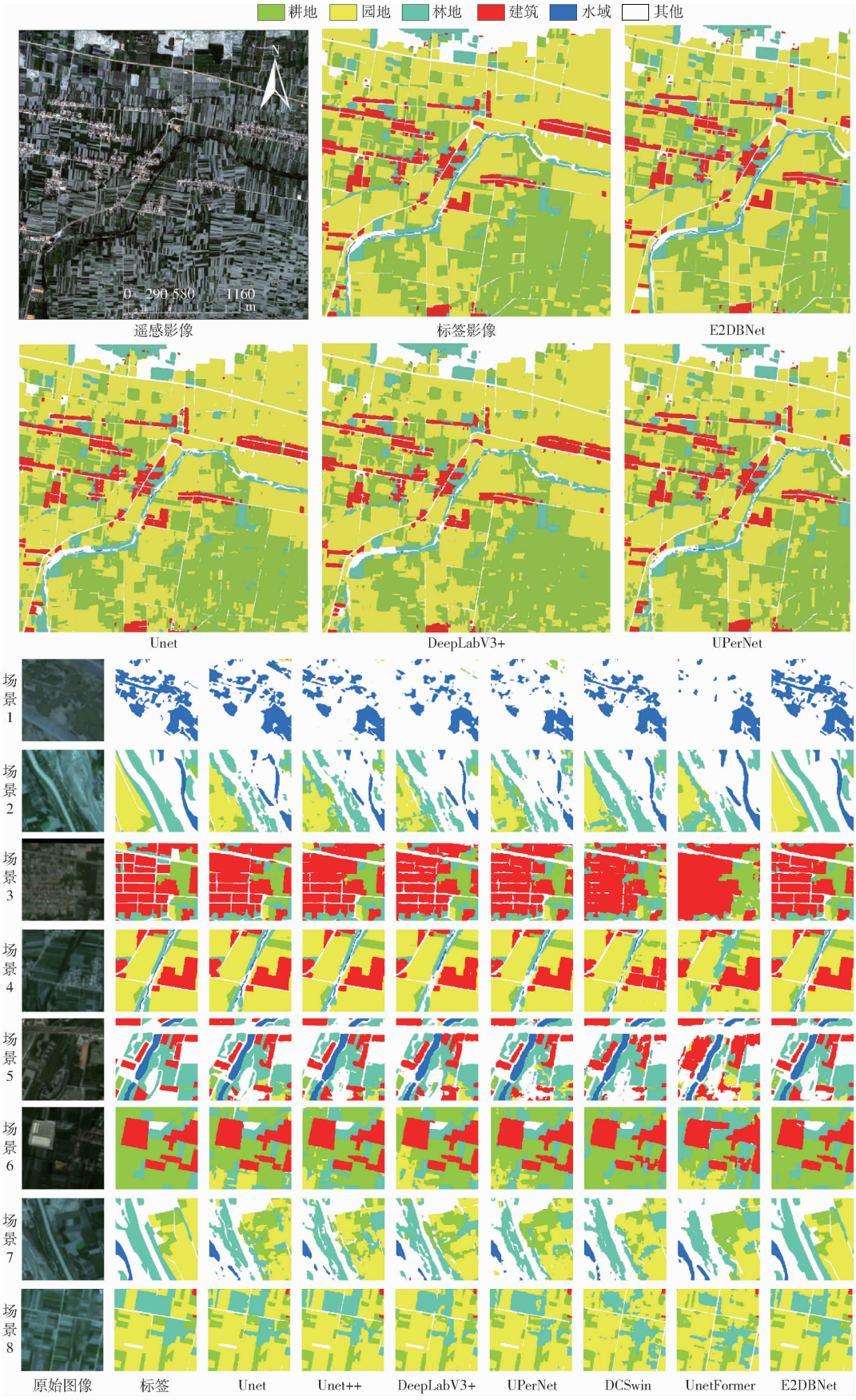


图 8 杨凌区土地覆盖分割数据集分割结果对比

Fig. 8 Comparison of model test results on Yangling segmentation dataset

布着一些林地,这些零散林地为山地生态系统提供了重要的支撑。其他类型土地利用形式则比较杂乱,零散分布在整个研究区域内,形成了复杂的背景土地利用格局。

表 3　杨凌区土地覆盖分割数据集分割方法实验结果对比

Tab. 3　Comparison of experimental results of segmentation methods on Yangling segmentation dataset

方法	F1 分数/%						总体指标/%				参数量
	耕地	园地	林地	建筑	水域	其他	交并比	召回率	F1 分数	精确率	
Unet	88.94	88.06	83.79	91.75	95.99	81.19	65.52	92.97	88.29	89.27	3.051×10^7
Unet + +	89.55	90.65	83.09	92.64	96.26	80.06	65.84	93.28	88.71	89.42	3.353×10^7
DeepLabV3 +	85.43	89.05	81.99	89.36	95.54	76.46	61.20	89.80	86.31	87.92	2.495×10^7
UPerNet	88.55	90.00	84.43	91.96	96.37	80.35	63.99	91.58	88.61	90.17	2.743×10^7
DCSwin	90.31	90.64	85.86	92.08	98.75	79.66	65.37	92.94	89.55	90.37	6.380×10^7
UnetFormer	72.88	76.38	70.30	88.63	91.46	85.26	68.58	80.80	80.82	80.90	2.308×10^7
E2DBNet	93.18	93.11	91.74	93.48	98.68	89.30	67.73	95.09	93.25	94.05	2.809×10^7

3.2　消融实验

为验证本文提出的 E2DBNet 中各个模块的有效性，分别去除 EEM、FAM 和 CLCM 模块进行充分消融实验，对实验结果进行定量分析，并且探究了不同损失函数对于最终结果的影响，消融实验定性结果如图 9 所示。

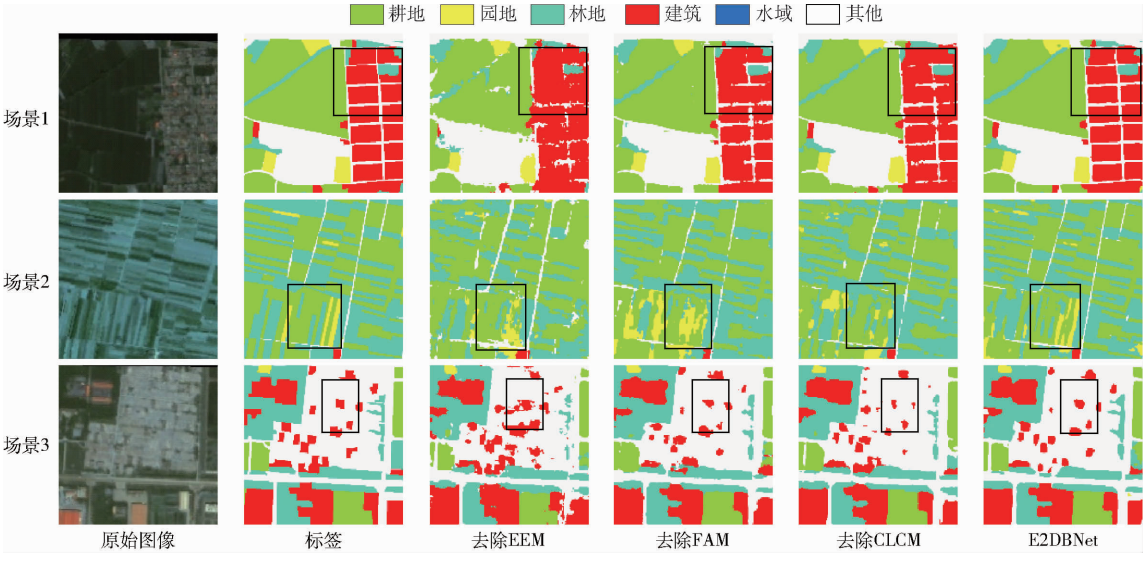


图 9　杨凌区土地覆盖分割数据集消融实验对比

Fig. 9　Comparison of ablation experiments on Yangling segmentation dataset

(1)EEM 模块效果

EEM 模块通过将边缘增强信息引入双分支网络中，使得网络可以更清晰地划分不同地物类型的边界，从而提高地块分割的效果。为验证其有效性，进一步评估了去除 EEM 模块的性能，如图 9 场景 1 所示，去除 EEM 模块后检测结果的边缘明显较差，表明边缘信息的辅助可以有效提高土地覆盖分割的精确性，使得分割结果边缘更加连续。如表 3、4 所示，E2DBNet 相较于去除 EEM 模块，其交并比、召回率、F1 分数、精确率分别高出 5.52、5.65、7.81、7.4 个百分点，在 3 个模块中的提升最高，证明了 EEM 模块在本文方法中的有效性。

(2)FAM 模块效果

FAM 模块有效聚合了语义特征及边缘特征，通过不同感受野的叠加确保全局语义与局部细节的增强，如图 9 场景 2 所示，去除 FAM 模块后检测结果中误检测明显增多，园地与林地之间并没有很好地区分开，而 E2DBNet 则很好地避免了此类问题，证明 FAM 模块的引入使得语义和边缘特征得到充分融合和利用。如表 4 所示，E2DBNet 相较于去除 FAM 模块，4 个指标分别高出 4.63、3.5、3.48、2.3 个百分点，证明了 FAM 模块在本文方法中的有效性。

表 4　E2DBNet 模型消融实验结果

Tab. 4　Results of ablation of E2DBNet model

模块组合			总体指标/%			
EEM	FAM	CLCM	交并比	召回率	F1 分数	精确率
×	×	×	58.53	83.92	72.30	74.04
×	√	√	62.21	89.44	85.44	86.65
√	×	√	63.10	91.59	89.77	91.75
√	√	×	65.11	93.13	92.03	93.35
√	√	√	67.73	95.09	93.25	94.05

注：√表示采用此模块，×表示未采用此模块。下同。

(3)CLCM 模块效果

CLCM 模块通过空间注意力模块，有效连接不同层次的融合特征，并强调其中关键的空间位置，进

一步增强模型对于不同尺寸地物的感知能力,如图 9 场景 3 所示,去除 CLCM 模块后模型对于微小建筑的分割能力明显下降,识别的建筑主体并不完整。如表 4 所示,与移除 CLCM 模块的情况相比,E2DBNet 4 个指标分别高出 2.62、1.96、1.22、0.7 个百分点,虽然在 3 个模块中的提升最小,但 CLCM 模块的引入进一步增强了模型捕捉细节的能力,证明了其在改善地物分割准确度方面的重要作用。在去除所有模块后,模型分割精度达到最低,交并比相较于 E2DBNet 下降 9.2 个百分点,充分证明了模块组合的有效性,极大提升了网络的分割能力,检测结果更加精确。与此同时,本文也进行了不同损失函数的探究实验,如表 5 所示,仅依赖 Dice Loss 或 Focal Loss 进行训练,均会在一定程度上限制模型的训练表现,进而影响地物的分割质量。采用混合损失函数的策略,结合 Dice Loss 的空间信息优化能力与 Focal Loss 对类别不平衡敏感度的特点,很好地平衡并提升模型在各个方面的性能,4 个指标均明显提升。

表 5 损失函数消融实验结果

Tab. 5 Results of ablation of loss function

损失函数		F1 分数/%						总体指标/%			
Dice Loss	Focal Loss	耕地	园地	林地	建筑	水域	其他	交并比	召回率	F1 分数	精确率
√	×	89.93	91.19	88.09	91.64	96.60	85.74	63.03	91.67	90.53	92.22
×	√	86.26	89.22	86.17	89.56	97.48	83.85	61.86	90.42	88.76	91.32
√	√	93.18	93.11	91.74	93.48	98.68	89.30	67.73	95.09	93.25	94.05

4 结论

(1)针对土地覆盖分割任务中存在的地物类型多样和边界区分不清晰的问题,本文构建了一种双分支语义分割网络,通过修改编码器结构,并引入额外的边缘辅助信息以提高相邻地物边界的连续性,并融合语义及边缘特征以加强信息交互,同时修改损失函数缓解类别不平衡带来的影响。通过将该方法与其他深度学习方法进行比较,证实了本研究所提出的 E2DBNet 模型在处理各种复杂遥感影像时的有效性。

(2)通过在编码器解码器结构中引入额外的边缘辅助信息,可以有效提高相邻地物边界的连续性;通过特征融合模块,可以有效融合语义及边缘特征,以加强信息交互,提高模型的分割精度;通过修改损失函数,可以缓解类别不平衡带来的影响。最终本文提出的 E2DBNet 在杨凌区土地覆盖分割数据集上取得了较好的分割效果。

参 考 文 献

[1] 常晗, 郭树欣, 张海洋, 等. 基于改进 DeepLab V3 + 的苹果种植面积提取[J]. 农业机械学报, 2023, 54(增刊 2): 206 - 213.
CHANG Han, GUO Shuxin, ZHANG Haiyang, et al. Apple planting area extraction based on improved DeepLab V3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2): 206 - 213. (in Chinese)

[2] 宁纪锋, 倪静, 何宜家, 等. 基于卷积注意力的无人机多光谱遥感影像地膜农田识别[J]. 农业机械学报, 2021, 52(9): 213 - 220.
NING Jifeng, NI Jing, HE Yijia, et al. Convolutional attention based plastic mulching farmland identification via UAV multispectral remote sesing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 213 - 220. (in Chinese)

[3] KOTARIDIS I, LAZARIDOU M. Remote sensing image segmentation advances: a meta-analysis[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 173: 309 - 322.

[4] CHEN S, SUN T, YANG F, et al. An improved optimum-path forest clustering algorithm for remote sensing image segmentation[J]. Computers & Geosciences, 2018, 112: 38 - 46.

[5] WANG Y, LI D, WANG Y. Realization of remote sensing image segmentation based on K-means clustering[J]. IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2019, 490(7): 072008.

[6] XIE B, ZHANG Y, MEI S, et al. Spectral variation augmented representation for hyperspectral imagery classification with few labeled samples[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1 - 12.

[7] 张宏鸣, 张国良, 朱珊娜, 等. 基于 U - Net 的葡萄种植区遥感识别方法[J]. 农业机械学报, 2022, 53(4): 173 - 182.
ZHANG Hongming, ZHANG Guoliang, ZHU Shanna, et al. Remote sensing recognition method of grape planting regions based on U - Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 173 - 182. (in Chinese)

[8] KUFFER M, PFEFFER K, SLIUZAS R, et al. Extraction of slum areas from VHR imagery using GLCM variance[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(5): 1830 - 1840.

[9] DAN Z, SANG N, HE Y, et al. An improved LBP transfer learning for remote sensing object recognition[J]. Optik, 2014, 125(1): 482 – 485.

[10] 武锦龙, 吴虹麒, 李浩, 等. 基于改进 DeepLabV3 + 的荞麦苗期无人机遥感图像分割识别方法研究[J]. 农业机械学报, 2024, 55(5): 186 – 195.
WU Jinlong, WU Hongqi, LI Hao, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV3 + at seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 186 – 195. (in Chinese)

[11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84 – 90.

[12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv, abs/1409.1556, 2014.

[13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770 – 778.

[14] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 – 4708.

[15] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640 – 651.

[16] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-assisted Intervention—MICCAI 2015: 18th International Conference, 2015: 234 – 241.

[17] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481 – 2495.

[18] CHEN L, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[J]. arXiv, abs/1412.7062, 2014.

[19] CHEN L, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(4): 834 – 848.

[20] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder – decoder with atrous separable convolution for semantic image segmentation[C]// ECCV, 2018: 801 – 818.

[21] 张宏鸣, 沈寅威, 阳光, 等. 融合注意力机制与多尺度信息的葡萄种植区变化检测[J]. 农业机械学报, 2024, 55(5): 196 – 206.
ZHANG Hongming, SHEN Yinwei, YANG Guang, et al. Change detection of grape growing areas based on integrating attention mechanism and multiscale information[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 196 – 206. (in Chinese)

[22] 顾兴健, 刘子儒, 任守纲, 等. 基于注意力机制和边缘感知的田梗提取模型[J]. 农业机械学报, 2023, 54(5): 210 – 218.
GU Xingjian, LIU Ziru, REN Shougang, et al. Ridge extraction model based on attention mechanism and edge perception[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(5): 210 – 218. (in Chinese)

[23] 杨蜀秦, 宋志双, 尹瀚平, 等. 基于深度语义分割的无人机多光谱遥感作物分类方法[J]. 农业机械学报, 2021, 52(3): 185 – 192.
YANG Shuqin, SONG Zhishuang, YIN Hanping, et al. Crop classification method of UVA multispectral remote sensing based on deep semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 185 – 192. (in Chinese)

[24] 孙志同, 朱珊娜, 高郑杰, 等. 基于波段增强的 DeepLabv3 + 多光谱影像葡萄种植区识别[J]. 农业工程学报, 2022, 38(7): 229 – 236.
SUN Zhitong, ZHU Shanna, GAO Zhengjie, et al. Recognition of grape growing areas in multispectral images based on band enhanced DeepLabv3 + [J]. Transactions of the CSAE, 2022, 38(7): 229 – 236. (in Chinese)

[25] 李亚飞, 董红斌. 基于卷积神经网络的遥感图像分类研究[J]. 智能系统学报, 2018, 13(4): 550 – 556.
LI Yafei, DONG Hongbin. Remote sensing images classification based on convolutional neural network[J]. CAAI Transactions on Intelligent Systems, 2018, 13(4): 550 – 556. (in Chinese)

[26] HE C, LI S, XIONG D, et al. Remote sensing image semantic segmentation based on edge information guidance[J]. Remote Sensing, 2020, 12(9): 1501.

[27] 杨军, 于茜子. 结合空洞卷积的 FuseNet 变体网络高分辨率遥感影像语义分割[J]. 武汉大学学报 (信息科学版), 2022, 47(7): 1071 – 1080.
YANG Jun, YU Xizi. Sematic segmentation of high-resolution remote sensing images based on improved FuseNet combined with atrous convolution[J]. Germatics and Information Science of Wuhan University, 2022, 47(7): 1071 – 1080. (in Chinese)

(下转第 491 页)

Vision and Pattern Recognition. IEEE, 2016:779－788.
[25] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision,2016:21－37.
[26] REIS D, KUPEC J, HONG J,et al. Real-time flying object detection with YOLOv8[J]. arXiv Preprint,arXiv:2305.09972, 2023.
[27] TANG Yehui, HAN Kai, GUO Jianyuan, et al. GhostNetV2: enhance cheap operation with long-range attention[J]. arXiv preprint arXiv:2211.12905,2022.
[28] TAN Mingxing, PANG Ruoming, QUOC V. Le. EfficientDet: scalable and efficient object detection[J]. arXiv Preprint,arXiv: 1911.09070,2019.
[29] ZHANG Hao, XU Cong, ZHANG Shuaijie. Inner－IoU: more effective intersection over union loss with auxiliary bounding box [J]. arXiv preprint, arXiv:2311.02877,2023.
[30] ZHOU Mingle, XING Rui, HAN Delong, et al. PDT: UAV target detection dataset for pests and diseases tree[J]. arXiv Preprint,arXiv:2409.15679,2024.
[31] HORVAT M, JELEČEVIĆ L, GLEDEC G. A comparative study of YOLOv5 models performance for image localization and classification[C]// Central European Conference on Information and Intelligent Systems. Faculty of Organization and Informatics Varazdin, 2022:349－356.
[32] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-Excitation networks[J]. arXiv Preprint,arXiv:1709.01507,2017.
[33] 林文树,张金生,何乃磊.基于改进 YOLO v4 的落叶松毛虫侵害树木实时检测方法[J].农业机械学报,2023,54(4):304－312,393.
LIN Wenshu, ZHANG Jinsheng, HE Nailei. Real-time detection method of dendrolimus superans-infested *Larix gmelinii* trees based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(4):304－312,393. (in Chinese)
[34] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS 2017). New York: Curran Associates Inc., 2017: 4768－4777.
[35] 仝召茂,陈学海,汪本福,等.一种基于改进 YOLOv5s 的麦穗检测计数方法[J].南京农业大学学报,2024,47(6):1202－1211.
TONG Zhaomao, CHEN Xuehai, WANG Benfu, et al. A wheat ear detection and counting method based on improved YOLOv5s[J]. Journal of Nanjing Agricultural University,2024,47(6):1202－1211. (in Chinese)
[36] 罗显志,汪航.跨尺度特征融合的无人机小目标检测算法[J].计算机工程与应用,2025,61(14):135－147.
LUO Xianzhi, WANG Hang. Small target detection algorithm for UAV based on cross-scale feature fusion[J]. Computer Engineering and Applications, 2025,61(14):135－147. (in Chinese)

(上接第 447 页)

[28] 叶沅鑫, 谭鑫, 孙苗苗, 等. 基于增强 DeepLabV3 网络的高分辨率遥感影像分类[J]. 测绘通报, 2021, 5(4): 40－44.
YE Yuanxin, TAN Xin, SUN Miaomiao, et al. High-resolution remote sensing image classification based on improved DeepLabV3 network[J]. Bulletin of Surveying and Mapping, 2021, 5(4): 40－44. (in Chinese)
[29] WANG L, LI R, ZHANG C, et al. UNetFormer: a UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 196－214.
[30] WANG X, HU Z, SHI S, et al. A deep learning method for optimizing semantic segmentation accuracy of remote sensing images based on improved UNet[J]. Scientific Reports, 2023, 13(1): 7600.
[31] 冯权泷,牛博文,陈泊安,等.土地利用/覆被深度学习遥感分类研究综述[J]. 农业机械学报,2022,53(3):1－17.
FENG Quanlong, NIU Bowen, CHEN Boan, et al. Review for deep learning in land use and land cover remote sensing classification[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 1－17. (in Chinese)