DOI:10.14188/j.1671-8844.2025.0262 文章编号:1671-8844(XXXX)XX-0001-11

扩散模型驱动的多任务地块信息提取方法

许松涛[1],钟晨[2],杨光义[2]

(1. 武汉大学测绘遥感信息工程全国重点实验室,湖北 武汉 430079; 2. 武汉大学电子信息学院,湖北 武汉 430072)

摘要:针对遥感耕地信息提取任务中存在的数据稀缺、边界模糊和地块内部同质性差等问题,提出一种扩散模型驱动的多任务耕地提取方法。通过条件扩散模型集成边缘形态、文本语义与风格控制,生成多样化合成耕地影像,解决样本稀缺问题;同时设计边缘-语义-距离多任务协同网络,利用任务互补性强化地块提取结果在空间分布上的连续性,降低特征漂移影响。实验表明,本文方法在AI4Boundary数据集上表现良好,mIoU指标较基线方法提升3.8%,在引入合成数据增强后达到85.9%的优异性能。本文不仅为耕地资源监测提供了更精准的技术手段,通过合成数据生成有效缓解了样本不足的问题,其提出的边缘引导增强策略和多源复合损失等方法也为其它地物分割任务提供了有价值的参考,特别适用于标注数据匮乏地区的耕地监测需求。

关键词:耕地地块;扩散模型;遥感样本生成;深度学习;边缘特征

中图分类号:TP 79 **文献标志码:**A

Multitask Agricultural Parcel Extraction via Diffusion Model-Guided Data Synthesis

XU Songtao[1], ZHONG Chen[2], YANG Guangyi[2]

(1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079;
2. Electronic Information School, Wuhan University, Wuhan 430072)

Abstract: To address the challenges of data scarcity, boundary ambiguity, and poor intra-parcel homogeneity in remote sensing cropland extraction tasks, this study proposes a diffusion model-driven multi-task cropland extraction method. By integrating edge morphology, text semantics, and style control into a conditional diffusion model, diverse synthetic cropland images are generated to mitigate sample scarcity. Additionally, an edge-semantic-distance multi-task collaborative network is designed to enhance spatial continuity in parcel extraction results through task complementarity, reducing feature drift effects. Experiments demonstrate that our method achieves strong performance on the AI4Boundary dataset, with a 3.8% improvement in mIoU over baseline methods and an outstanding accuracy of 85.9% after synthetic data augmentation. This paper not only provides a more precise technical solution for cropland monitoring—effectively alleviating sample insufficiency via synthetic data generation—but also offers valuable references for other land cover segmentation tasks

收稿日期:2025-07-21
作者简介:许松涛(1998-),男,硕士研究生,主要从事遥感图像分析与应用等研究,E-mail: 2021206190022@whu.edu.cn。
通信作者:杨光义(1983-),男,教授级高级实验师,主要从事高频电路教学和图像处理等研究,E-mail: ygy@whu.edu.cn。
基金项目:科技部重点研发计划课题(编号:2022YFB3903605),应急管理部重点科技计划课题(编号:2024EMST070706)。



本文引用格式:许松涛,钟晨,杨光义.扩散模型驱动的多任务地块信息提取方法[J].武汉大学学报(工学版),XXXX,XX(XX):1-11.

Citation:XU Songtao, ZHONG Chen, YANG Guangyi. Multitask Agricultural Parcel Extraction via Diffusion Model-Guided Data Synthesis[J]. Engineering Journal of Wuhan University, XXXX, XX(XX):1-11.

through its edge-guided enhancement strategy and multi-source composite loss, particularly suited for cropland monitoring in annotation-scarce regions.

Key words: cropland plots; diffusion model; remote sensing sample generation; deep learning; edge features

耕地作为人类生存和发展的物质基础,是国家粮食安全的重要保障[1]。在我国,现存土地资源总量有限并且人均占有量不足,国家提出18亿亩耕地红线政策,成为维护国家粮食安全的关键举措[2]。随着城镇化与工业化进程的加速推进,耕地保护面临着"非农化"和"非粮化"的双重挑战[3]。在此背景下,构建科学、高效的耕地统计管理体系,对于合理利用土地资源以及保障国家粮食安全具有重要意义。

长期以来,实地调查方法是国家获取耕地信息的主要手段,但在大范围区域分析中暴露出效率低下的问题。卫星遥感技术凭借高分辨率、低成本及高度自动化作业的优势,在国土资源调查中脱颖而出,成为地块信息统计的有效技术手段。早期研究中,已有学者提出多尺度均值漂移算法用于农林地块边界的精准提取[4]。亦有大量研究尝试基于山脊、道路等地表特征或图像梯度信息等边缘分界线划分耕地地块[5,6],但此类方法依赖手工设计特征,在大范围地块信息提取任务中存在提取精度不足、泛化能力较弱的问题。

随着深度学习技术的蓬勃发展,相关领域的耕地信息提取方法逐渐成为了研究热点,其中以语义分割网络研究最为广泛[7-9]。张海天等[7]提出改进的Unet++网络模型,有效实现了复杂丘陵地区的耕地提取任务。部分研究聚焦于地块边缘特征,探索基于边缘信息的耕地轮廓提取方法[10-12],也取得了较好的成效。然而,当前研究仍存在一定局限性:多数工作采用边缘特征或语义特征等单一模态信息进行信息提取[13,14],缺乏对多维度空间信息的融合处理。并且,此类方法过度依赖大规模标注训练数据,难以在跨区域、跨时相的耕地遥感影像中保持稳定性能。

针对遥感地物信息的样本依赖问题,研究者们逐渐将目光转向样本依赖较小的解决方案:半监督学习与弱监督学习[15-18]。丁一鹏等[19]通过对抗学习方法实现跨数据集域适应,将在ISPRS Vaihingen数据集上训练的模型迁移至ISPRS Potsdam数据集,使迁移后分类精度提升约7%。栾晓梅等[16]提出一种基于中间特征图调整的弱监督语义分割方法,无需耕地语义标注即可实现信息提取,以其他任务形式的标注信息来引导分割。然而,尽管这些方法在一定程度上缓解了样本依赖问题,但此类方法要么高度依赖边缘特征相似、仅风格色调存在差异的前置数据集[19];要么对目标对象的空间分布特征有严格要求,需目标具备清晰的范围界定[16],严格的应用前提条件制约了其通用性与适用场景。

为解决半监督学习和弱监督方法带来的通用性不足问题,基于生成模型的数据增强技术为小样本学习提供了新的方向[20,21]。Sastry等[20]提出采用OSM(open street map)地理数据、SAM(segment anything mask)语义掩码及Canny边缘特征进行控制的手段,以实现符合设计预期的高质量遥感影像生成。Khanna等[25]设计DiffusionSat基础模型,作为将扩散模型引入遥感的基础大模型,实现时空感知的卫星图像生成,同时支持卫星图像超分辨率重建与修复任务。龚健雅等[22]将ControlNet技术[23-24]引入遥感数据增强,结合语义分割样本生成高质量合成遥感图像,以用于遥感语义分割任务,使mIoU整体精度提升3.25%。项俊武[14]通过融合耕地纹理先验知识生成高逼真度的合成影像,并将其应用于耕地信息提取任务,有效提升地块边界的提取精度。

尽管以深度学习和生成模型为代表的新兴技术在耕地信息提取领域进步显著,但数据依赖问题与应用场景适配性差仍制约其实际落地。具体表现为:1)样本稀缺性;耕地地块在地形地貌、气候条件等方面存在显著区域差异,实际应用中可用标注样本相对匮乏。而样本覆盖不足问题则限制了耕地信息提取方法的适用范围。2)模型适应性;由于耕地内部作物种类多样且存在季节性物候变化,导致遥感影像中普遍存在同物异谱现象。这使得现有模型难以实现跨季节、跨作物类型的稳定特征表达,制约了其对复杂地块场景的提取性能。

为此,本文提出一种扩散模型驱动的多任务地块信息提取方法。首先,通过构建条件扩散生成模型,学习生成包含多样化的模拟地块图像,并以模拟地块图像结合文本描述信息和图像风格信息,构建多模态控制条件,实现耕地图像的可控合成。其次,针对地块内部纹理异质性及季节性变化导致的特征漂移问题,本文设计了一种多任务特征增强网

络,结合边缘提取、语义分割和距离预测的多任务学习结构,强化模型的鲁棒表征能力,显著提升模型性能。本文的主要贡献包括:

(1)提出一种基于扩散模型的耕地样本生成方法,通过集成多元边缘形态约束与文本语义信息,实现多场景耕地地块的高质量合成,为样本稀缺区域提供低成本的数据增广解决方案;

(2)构建一种多任务协同学习方法,通过边缘-语义-距离的联合特征学习扩展模型表征维度,利用任务间的互补性降低单一任务的学习复杂度,有效提升模型对跨区域、跨场景遥感影像的适应性。

1 基于扩散模型增强的地块提取方法

扩散模型驱动的多任务地块信息提取方法(multitask agricultural parcel extraction via diffusion model-guided data synthesis, MAPES-DM)主要包含两大核心模块:合成数据生成模块和多任务信息提取模块。合成数据生成模块包含整体语义信息生成、地块模拟图生成、合成数据生成和合成数据筛选四个部分,其核心是获取图像所处的场景、色调、地块形态。作为图像生成的全局的语义信息输入,训练使模型具备不同条件下的地块生成能力。多任务信息提取模块主要为多任务耕地地块提取。模块所含多任务地块提取模型通过完成距离、像素、语义和边缘联合训练,提高地块信息提取的准确性和鲁棒性。具体技术流程如图1所示。

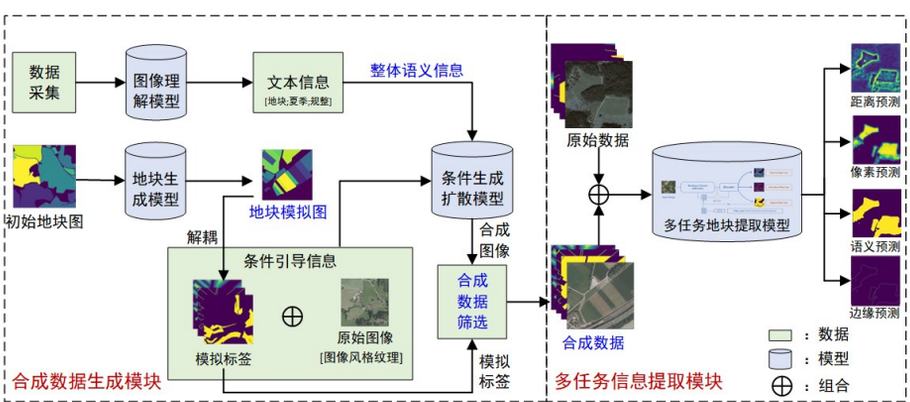


图 1 多任务地块信息提取流程图

Fig.1 Flow chart of multi-task plot information extraction

1.1 合成数据生成模块

条件扩散模型通过在生成过程中引入额外的条件信息(如文本、边缘图、深度图等),让图像生成变得精准可控,相较于一般扩散模型,保留了原有的前向扩散过程和反向去噪过程,基于噪声的添加或去除,在前向过程向图像添加噪声,直到变成纯噪声;反向过程则从噪声中逐步去噪以生成图像,在此基础引入条件指导生成,ControlNet是条件扩散模型的一种形式,其核心设计思想在于不破坏预训练大模型原有生成能力的前提下,通过权重分离和零卷积来实现高效、鲁棒的条件控制,以图像作为引导生成新的原图像。本文借鉴这种思想,提出一种基于扩散模型的合成数据生成模块,其核心采用ControlNet架构实现可控图像生成,旨在生成多样化场景的地块数据。该模块包含地块生成模型和条件生成扩散模型两个关键组成。其中地块生成模型创新性地提出一种扭曲因子计算方法,将其作为控制条件引入训练过程以调控高斯噪声生成地块模拟图。

相较于传统方法仅通过风格调整模拟不同时期地块影像的局限性,本方法通过扭曲因子实现了对生成地块形态的多样化控制。在条件生成扩散模型中则是以文本信息和条件引导信息(含地块分割图、距离图、边缘图及纹理特征图)作为输入。在条件生成扩散模型前向传播过程中,文本信息与条件引导信息所包含多模态图像被编码为特征向量共同注入网络,驱动生成具有新颖空间形态的合成数据。生成的合成图像与条件引导信息构成初始“图像-标签对”,后经过合成数据筛选组成最终合成数据,筛选过程主要通过对比语言-图像预训练模型(contrastive language-image pre-training, CLIP)抽取所生成合成图像高维特征,随后对比条件引导信息多模态图像和合成图像的高维特征分布情况,选出一致性较好的“图像-标签对”构成合成数据集。最

终将所得合成数据集与原始训练数据集合并,构建多元形态数据集合。整体流程如图2所示。

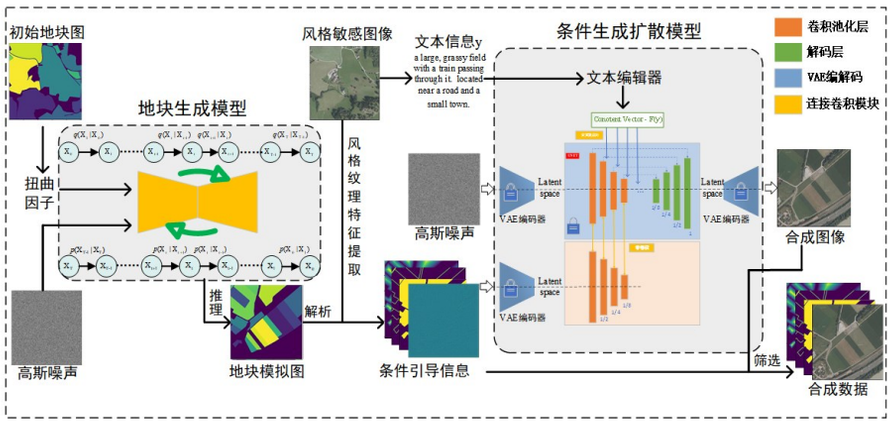


图2 合成数据集构建流程图

Fig.2 Synthetic dataset construction flow chart

在地块模拟图的生成过程中,本文通过噪声图扩散生成地块模拟图,其间加入扭曲因子来控制生成地块边界的形态和数量。扭曲因子定义为几何图形周长与面积比值的归一化指标:

$$Q=\frac{\sum_i\left(\frac{P_i}{A_i}\right)}{\max\left(\frac{P_i}{A_i}\right)} \tag{1}$$

其中,P_i表示第i个地块的周长,A_i为地块面积。训练时将扭曲因子作为条件输入,预测阶段通过调节扭曲因子控制生成地块的形态复杂度。地块形态越复杂,地块数目越多,扭曲因子的值越大。

1.2 多任务信息提取模块

多任务信息提取模块旨在解决遥感地块提取任务中地块边界模糊和地块内部同质性差的问题,其核心是本文提出的空间增强多任务学习网络(spatial enhancement multi-task learning network,SEMLNet)。如图3所示,SEMLNet包括四个组成部分,分别是边缘引导增强子模块(boundary-guided augmentation submodule,BAM)、风格敏感子模块(style-sensitive submodule,SSM)、多任务联合学习子模块(multi-task joint learning submodule,MJLS)和主干网络。主干网络采用Deeplabv3+[32]空洞卷积架构作为基本骨干网络,通过风格敏感子模块和边缘引导增强子模块构建跨域鲁棒性的风格表征,并引入边缘强约束抑制内部噪声干扰,从而解决地块边缘模糊和内容同质性差的问题。另外,SEMLNet采用多任务联合学习子模块引导模型学习泛化性特征,显著缓解单一任务噪声过拟合风险。

图3中,边缘引导增强子模块用于从原始图像中提取边缘特征,并对主干网络的语义特征进行引导增强;风格敏感子模块于融合低维和高维特征,以增强地块特征;多任务联合学习子模块将来自多个子模块的多源特征进行融合,得到最终结果。具体流程包括:利用拉普拉斯边缘卷积核(LapConv)进行卷积,得到初步边缘结果图;随后从主干网络分支中抽取低维特征进行结合,以增加边缘分支的特征丰富度;经过通道注意力机制(channel attention mechanism,CAM)进行特征精炼后得到边缘特征,即变换后的空间特征层;最后经过两次3×3卷积结构进行特征融合,得到边缘轮廓图。

长距离门控卷积注意力结构(long-range gated convolution-based attention,LGBA)是风格敏感子模块进行特征增强的核心内容,其将输入特征经过多层次多尺寸门控卷积结构动态学习通道与空间位置的特征权重,以对网络中间特征进行增强,并且采用多维度门控特征提取结构结合长短距离特征动态融合局部与全局语义依赖。具体来说,多维度门控特征提取结构通过上下文编码提取上下文信息,并通过门控交互投影进行信息融合;长短距离特征则通过上下并行卷积来进行特征抽取,从多个维度对网络特征层进行增强,以提升模型对边界颜色变化的敏感度。在风格敏感子模块的末端,4个LGBA所提取特征信息经过两次1×1卷积进行特征融合,以及一次1×1卷积对齐特征通道,最终整体输出至主干网络。

多任务联合学习子模块将来自多个子模块的多源特征进行融合,并构建多任务联合损失函数支持模型训练,最终输出中心距离、语义分割、像素损失结果。这些输入特征包含主干网络分支中空洞

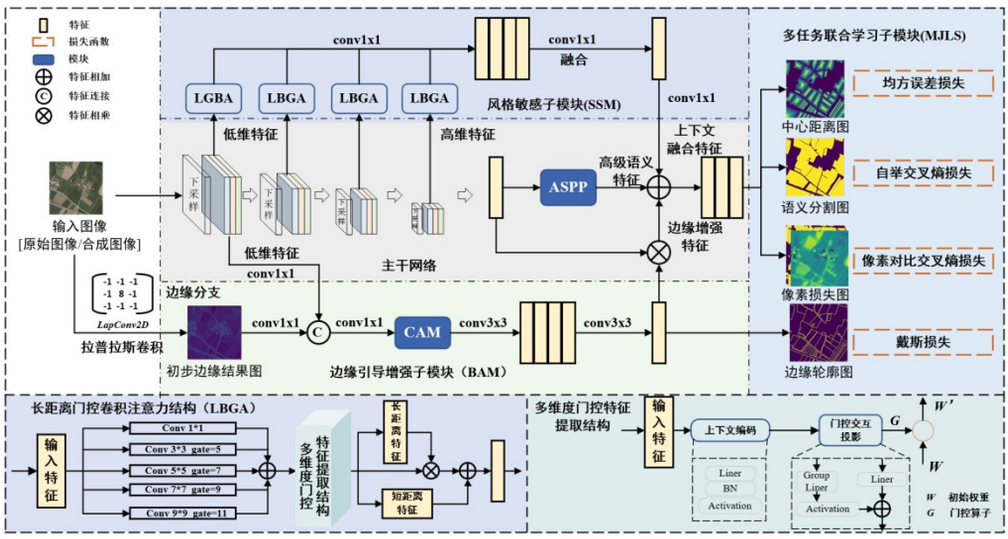


图 3　空间增强多任务学习网络

Fig.3　Structural diagram of the spatially enhanced multi-task learning network

空间金字塔池化结构(atrous spatial pyramid pooling,ASPP)输出的高层次语义特征、边缘引导增强子模块输出的边缘增强特征和风格敏感子模块输出的上下文融合特征。

多任务联合学习子模块基于中心距离图、语义分割图、像素损失图和边缘轮廓图构建多分支联合损失函数。语义分割损失函数采用自举交叉熵损失(BoostStropLoss)平衡耕地与背景的像素级分类损失，具体可见公式(3)；像素对比交叉熵损失(PixelContrastLoss)用于统计像素损失以增强地块语义结果的空间一致性，具体可见公式(5)；边缘分支采用戴斯损失(DiceLoss)处理稀疏边缘样本的梯度失衡问题，具体可见公式(4)；中心距离预测采用均方误差损失(MSELoss)约束距离图的回归精度，具体可见公式(6)；联合损失聚合语义、像素、边缘及距离损失，以提升多任务学习的准确性，具体可见公式(2)：

$$L_{sum}=\lambda_1 L_{seg}+\lambda_2 L_{pixer}+\lambda_3 L_{bnd}+\lambda_4 L_{dst} \tag{2}$$

$$L_{seg}=-\frac{1}{N}\sum_{i=1}^{N}(1-\beta)*y_i*\log(p_i) \tag{3}$$

$$L_{bnd}=1-\frac{2\sum_{i=1}^{N}p_i y_i}{\sum_{i=1}^{N}p_i+\sum_{i=1}^{N}y_i} \tag{4}$$

$$L_{pixer}=-log\frac{\sum_{i,j}e^{\cos(x_i,x_j')/\tau}}{\sum_{i,j}e^{\cos(x_i,x_j')/\tau}+\sum_{i,k}e^{\cos(x_i,x_k')/\tau}} \tag{5}$$

$$L_{dst}=\frac{1}{N}\sum_{i=1}^{N}(y_i-p_i)^2 \tag{6}$$

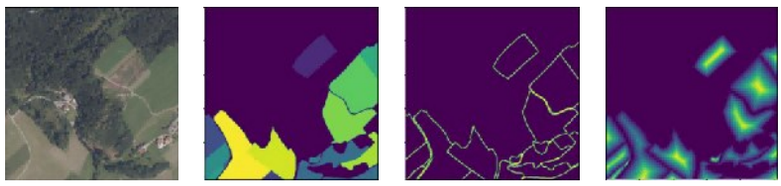
在语义分割损失L_{seg}、边缘损失L_{bnd}以及距离损失L_{dst}中，y_i表示真实标签，p_i表示预测的结果，N为像素点总数，β是类别权重。本文为地块提取任务，实践中背景与地块占比大致相当，故β值设置为0.5。在像素损失函数L_{pixer}中，x_i表示选中像素特征向量，x_j'表示正样本特征向量，x_k'表示负样本特征向量，τ为超参数。另外，在联合损失L_{sum}中，综合考量各损失权重综合的数量级差异及对地块提取任务重要性，设定$\lambda_1=1$，$\lambda_2=\lambda_3=0.4$，$\lambda_4=0.0005$。

2　结果与分析

2.1　数据集与实验设置

AI4Boundaries耕地数据集(简称AI4耕地数据集)[26]于2022年公开发布，是国际上广泛关注的地块信息提取基准数据集。数据集包含用于区域性规模分析的1米分辨率正射影像，提供开放获取渠道，数据可从欧盟数据门户(http://data.europa.eu/89h/0e79ce5d-e4c8-4721-8773-59a4acf2c9c9)下载。数据集由7831个512×512像素的样本构成，按照70%、15%、15%的比例划分为训练集(5319个)、验证集(1140个)和测试集(1139个)。AI4耕地数据集提供原始图像对应的分割、边缘和中心距离标签，其切片划分示例如图4所示。

实验分为两部分，一是合成数据有效性验证，该部分主要通过对比加入合成数据的地块提取效果，验证合成数据对训练的增益作用；二是多任务网络性能验证，通过与主流方法对比验证整体框架有效性。本文将OCRNet[27]、BiseNetv2[28]、DANet[29]、Unet[30]、Segnext[31]和Deeplabv3+[32]等方法与本文所提出的空间增强多任务学习网络(SEMLNet)进行横向对比，其中Unet、Deeplabv3+



(a)遥感影像原图 (b)分割标签图 (c)边缘标签图 (d)中心距离标签

图 4 AI4Boundaries耕地数据集切片示例

Fig.4 AI4Boundaries cropland dataset slice map

为早期的经典语义分割任务框架,Unet为经典的编码器-解码器结构,配合跳跃连接,可以在有限标注数据下,也能有效学习并生成精细的地块分割图;Deeplabv3+结构空洞空间金字塔池化与解码器结合,捕获多尺度上下文,感受野大,能充分理解地块所在的宏观环境;OCRNet、BiseNetv2、DANet和Segnext则是近年来遥感领域常用的网络结构,OCRNet提对象上下文表示模块,显式建模像素与物体区域的关系,BiSeNetV2采用双分支结构,可兼顾细节与语义信息,可以更好地保留地块轮廓细节,而Segnext采用多尺度卷积注意力模块,融合不同尺度卷积核特征,DANet则采用了双注意力机制,这两个方法特殊的注意力结构设计使其可以更好地捕获全局上下文信息。在此基础之上,加入补充合成数据训练的扩展实验,最后通过消融实验量化各子模块(边缘引导增强子模块、风格敏感子模块和多任务联合学习子模块)的贡献。

实验环境配置为Ubuntu 18.04系统,搭载GTX 4090D显卡。超参数设置如下:输入图像尺寸512×512,批量大小8,训练轮次50,采用SGD(stochastic gradient descent)优化器(初始学习率1e-4)。训练过程中保留测试集精度最优的模型,最终在验证集上完成性能评估,所有方法均采用了相同的训练集、测试集和验证集进行实验,相关超参数也同样保持一致。

评估体系包括合成数据集质量评价与地块提取精度评估两部分任务。在合成数据集质量评价中,采用峰值信噪比[14](peak signal-to-noise ratio,PSNR)作为核心指标。针对地块提取精度的评价,本文沿用遥感语义分割任务的主流评估体系,选取平均交并比(mean intersection over union,mIoU)[14]和F1分数[14]作为关键指标,并以精确率(Precision)和召回率(Recall)配合进行精度分析,增加分析的全面性。

2.2 合成数据效果验证

合成数据效果验证包括扩散模型的生成效果分析以及合成数据对模型训练的效果分析两个环节。首先是对扩散模型的生成效果分析,本文比较不同生成条件下的生成图像质量和标签一致性情况,分别在不同轮廓条件和不同的文本控制下进行图像生成实验,评估不同场景下的模型生成效果和图像生成质量。合成数据效果验证方面,本文横向对比OCRNet、DANet、Unet、Deeplabv3+以及本文所提出的SEMLNet网络的基础效果,并纵向对比了加入合成数据后的实验效果。

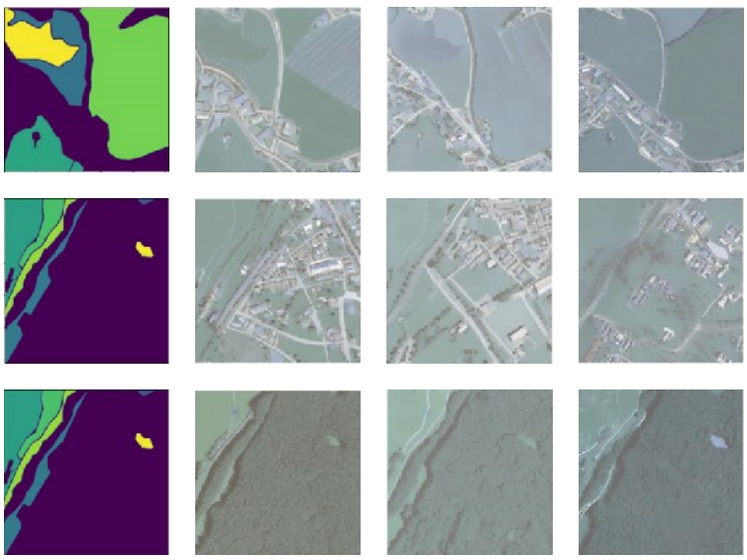
2.2.1 扩散模型的生成效果分析

为了验证扩散模型的生产效果,选取AI4耕地数据集中的三张初始地块图,设置不同的控制条件,实验结果如图5所示。图5中,(b)(c)(d)三列代表不同条件下的生成结果。

第1行与第2行图像提供不同的初始地块轮廓,但给予相同的文本提示,文本提示均为The image depicts a rural area with a mix of agricultural land and residential structures(由耕地和住宅建筑组成的农村地区)。不难发现,部分生成图像的地块边界不够显著,如第1行的图像(d),其左上角被包围的地块边界线难以辨别。第2行图像(d)将耕地部分生成为蓝灰色,此类色调一般不会在耕地中出现。

第3行的图像则采用与第2行相同的地块轮廓,但修改给出的文本提示,文本提示改为The image depicts a lush green forest with a mountainous terrain(一片郁郁葱葱的森林和山峦起伏的地形)。可以看出,两行图像对应的生成图像不仅在色调上存在一些差异,而且孤立地区的地块生成也有较大差异,如第2行的图像(d),预期的耕地地块区域生成了城镇水泥路面,而第3行的图像(d)则表现出较好的自然景观特征,森林轮廓和边界阴影均有较好的呈现。

此外,本文对比原始数据定量评估了生成数据的效果,PSNR值均在28左右。在图像生成任务中PSNR在20-30之间被认为是失真较小,可见采用本文思路可以得到质量较好的遥感合成图像。



(a)地块图像 (b)合成图像1 (c)合成图像2 (d)合成图像3

图5 不同控制条件地块图像生成结果

Fig.5 Image generation results for plots with different control conditions

2.2.2 合成数据对模型训练的效果分析

除了要求合成数据的质量,更为重要的是合成数据能否有效提升地块提取效果。本文基于训练数据得到大量合成数据,并对比不同模型中是否加入合成数据的提取效果。为验证合成数据的效果,实验中以原始数据:合成数据=1:1的关系进行训练和测试,具体数据如表1所示。

表1 合成数据补充精度对比分析(mIoU指标)

Table 1 Comparative analysis of synthetic data supplement accuracy

模型方法	原始数据集	加入合成数据	提升比例
OCRNet	0.754	0.773	**1.9%**
DANet	0.776	0.786	1.0%
Unet	0.789	0.795	0.6%
Deeplabv3+	0.814	0.826	1.2%
SEMLNet	**0.843**	**0.859**	1.8%

可以看出,引入合成数据后,所有对比模型的地块提取精度均得到提升,提升幅度在0.6%-1.9%之间,表明合成数据的加入可有效增强模型对多样化地块的识别能力。其中,基于HRNet骨干网络的OCRNet初始精度最低,但精度提升最为显著(提高1.9%),可能的原因是OCRNet较深的网络层次使其在数据量增加时,展现出更强的表征学习能力。SEMLNet的精度提升优于除OCRNet外的其他方法,且在所有模型中取得了最好的效果,充分验证了本文方法在地块提取任务中的应用潜力。

2.3 多任务学习地块提取实验

多任务学习地块提取实验分析旨在比较不同模型对同一原始数据的提取效果,对比方法包括常用的地块提取方法:OCRNet、BiseNetv2、DANet、Unet、Segnext及Deeplabv3+,这些方法均采用ImageNet预训练模型作为网络骨架,且使用相同的地块分割损失函数——自举交叉熵损失(BoostStrapLoss)。

2.3.1 定性分析

选取AI4耕地数据集中的三张原始图像作为模型输入,得到的输出结果如图6所示。五张原始图像分别代表规整农田地块、城市周边零散农田地块、郊区附近较规整的农田地块、山区大面积破碎地块和林区不规则地块。

可以看出,SEMLNet凭借边缘引导增强子模块,对窄小地物边界(如地块间田间路)和不规则地块轮廓的提取效果显著优于其它方法。对于规整农田地块,所有算法均得到了较好的结果;对于城市周边零散农田地块,SEMLNet能输出相对更连续和完整的耕地信息结果;对于郊区附近较规整的农田地块,SEMLNet较好地保留了地块边缘细节,并且有效抑制了孤立噪声误分现象的发生;对于山区大面积破碎地块,SEMLNet能较好排除边缘噪点;对于林区不规则地块,SEMLNet仍能较好保留孤立地块零碎信息,体现了本文方法在复杂地块轮廓环境下的鲁棒性。

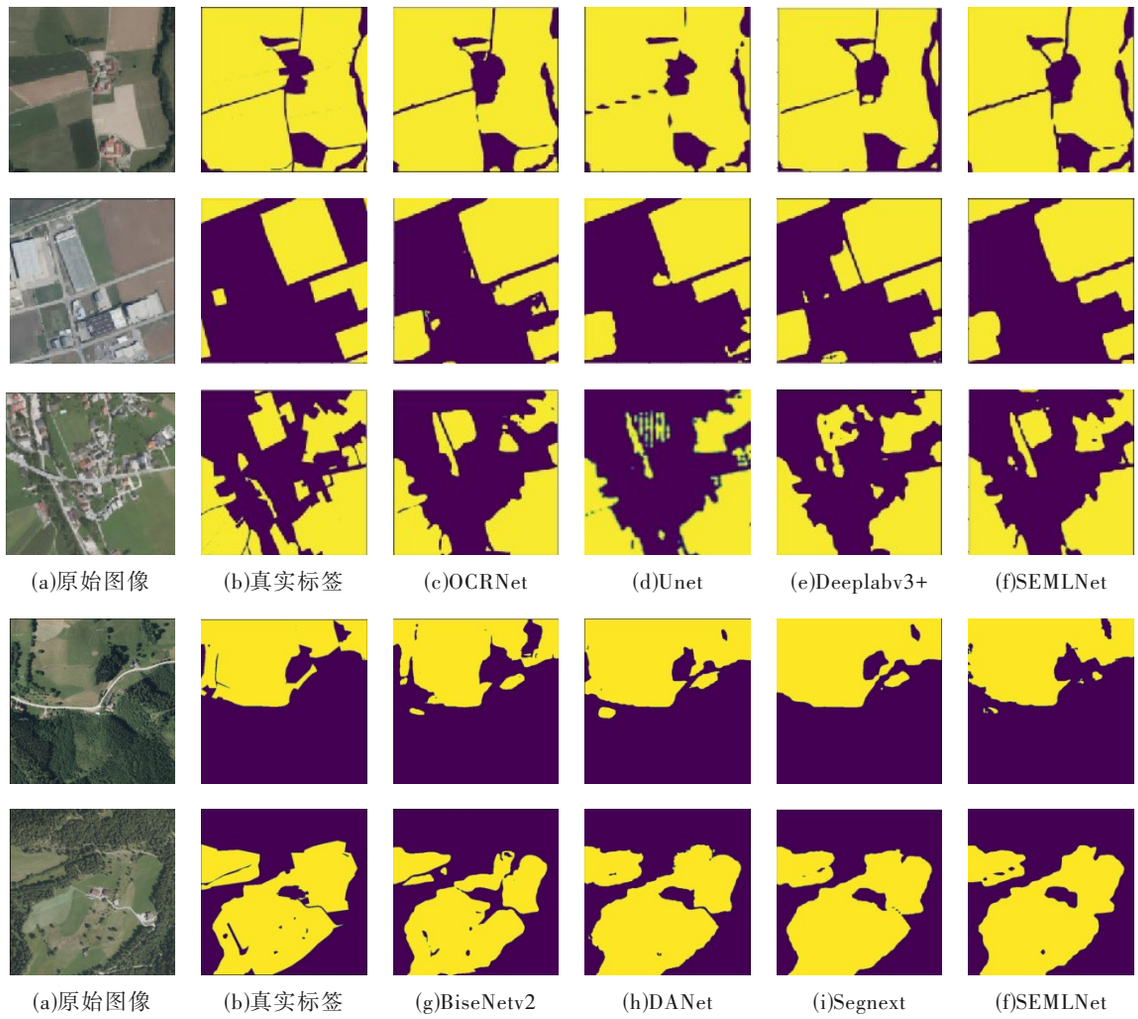


图 6 地块提取定性结果

Fig.6 Qualitative results of parcel extraction

2.3.2 定量分析

定量实验结果见表2,表中SEMLNet++表示增加合成数据的SEMLNet模型。实验表明,除本文所提出的SEMLNet外,DeepLabv3+与Segnext在地块提取任务中表现较好(mIoU>80%),其共性在于利用空洞卷积技术扩大模型感受野,有效捕捉了地块的大尺度空间分布特征。具体而言,DeepLabv3+通过多尺度特征融合模块优化上下文建模,Segnext则侧重空域特征建模与跨尺度特征交互,二者的结构设计均与耕地目标的宏观特性高度契合。

本文提出的SEMLNet借鉴DeepLabv3+和SegNext的特殊网络结构设计,结合特征子模块SSM与BAM,显著提升空间特征提取能力与上下文语义融合能力;主干网络引入的空洞空间金字塔池化结构ASPP有效扩大了感受野,扩展了网络特征的深度和广度。在未使用合成数据时,SEMLNet取得84.7%的mIoU与91.5%的F1分数,验证了SEMLNet网络结构设计的有效性;加入合成数据后,SEMLNet++进一步提升至85.9%的mIoU与92.4%的F1分数,表明合成数据与模型结构具有协同优化效应,通过引入多场景的地块合成数据,可以显著提升模型在复杂地块场景中的准确性与鲁棒性。

2.4 消融实验

为探究各子模块对模型性能的贡献,本节在原始的AI4耕地数据集上开展消融实验,重点分析边缘引导增强子模块(BAM)、风格敏感子模块(SSM)和多任务联合学习子模块(MJLS)的作用。共设计四组对照实验:1)基线模型:仅保留主干网络架构,通过单任务分割提取特征并解码得到最终的语义输出结果;2)SSM增强模型:在基线基础上添加风格敏感子模块,加强上下文特征增强能力,仍维持单任务分割;3)MJLS+SSM增强模型:在

表2 地块提取定量结果

Table 2 Quantitative results of parcel extraction

模型方法	精确度	召回率	F1	mIoU
OCRNet	0.868	0.860	0.864	0.754
BiseNetv2	0.874	0.861	0.864	0.762
DANet	0.875	0.869	0.873	0.776
Unet	0.886	0.883	0.884	0.789
Segnext	0.897	0.892	0.8950	0.811
Deeplabv3+	0.897	0.900	0.898	0.814
SEMLNet	0.916	0.912	0.915	0.843
SEMLNet++	**0.923**	**0.924**	**0.924**	**0.859**

SSM基础上引入多任务联合学习子模块,通过联合损失函数与注意力机制优化多任务输出,共享主干网络并生成多头任务结果;4)完整模型:在MJLS+SSM基础上加入边缘引导增强子模块,实现边缘特征与主干网络多尺度特征的协同优化。训练参数与多任务学习地块提取实验设置保持一致,将验证集的精度作为评估指标,结果如表3所示。

表3 SEMLNet消融实验

Table 3 SEMLNet ablation experiments

SSM	MJLS	BAM	F1	mIoU	F1提升比例	mIoU提升比例
×	×	×	0.887	0.805	—	—
√	×	×	0.894	0.811	0.7%	0.6%
√	√	×	0.910	0.836	2.3%	3.1%
√	√	√	**0.915**	**0.843**	2.8%	3.8%

相比于基线模型,风格敏感子模块的引入使F1分数和mIoU分别提升0.7%和0.6%;多任务联合学习子模块进一步推动F1分数和mIoU提升2.3%和3.1%;边缘引导增强子模块加入后,F1分数和mIoU再提升2.8%和3.8%。其中,多任务联合学习子模块对mIoU的贡献最为显著。实验结果充分验证了多尺度多源特征融合在大范围地块信息提取任务中的有效性,特别是融合地块边缘特征与网络上下文特征后,模型整体性能提升效果突出。

3 结论

本文提出一种扩散模型驱动的多任务地块信息提取方法(MAPES-DM),以解决现有遥感地块信息提取方法中存在的数据稀缺、边界模糊和地块内部同质性差等问题。常规数据增强方式无法引入额外信息,通过预训练的生成大模型可有效增广原有数据集,利用扩散模型实现可控地块图像合成,以边缘形态约束与文本语义控制生成多元场景数据,突破样本稀缺限制;在地块分割实验中,本文发现增大感受野的网络结构设计能取得较为显著的效果,利用多任务信息提取模块融合边缘引导增强与多任务多尺度联合学习机制,实现空间结构与语义特征的深度耦合,克服边界模糊带来的挑战,采用多尺度语义特征融合地块上下文特征,有效扩大了感受野,抑制了孤立噪声误分现象的发生,取得了较好的提取效果。此外,从实验结果来看,本文在引入了多任务模块后,整体精度有了显著提升,可见在地块分割任务网络模块设计时,增加边缘特征分支是一个有效的策略。综上,本文从数据增强和网络模块设计方向着手,为耕地资源监测提供了高效可靠的技术范式,提升了模型在复杂场景下的解译精度与鲁棒性,也为其它地物分割任务提供了有价值的参考。

参考文献:

[1] 赵晓宇, 孙春强, 崔荣国, 等. 基于粮食安全的中国耕地保护阈值研究[J]. 资源科学, 2024, 46(12): 2355-2366.
ZHAO Xiaoyu, SUN Chunqiang, CUI Rongguo, et al. China′s cultivated land threshold from the perspective of food security[J]. Resources Science, 2024, 46(12): 2355-2366.

[2] 张擂, 周煜明, 董杰谋, 等. 中国耕地非农化与非粮化时空分异及其治理策略[J]. 干旱区研究, 2025, 42(2): 372-383.
ZHANG Lei, ZHOU Yuming, DONG Jiemou, et al. Spatiotemporal differentiation of nonagricultural and nongrainfarmland in China and its management strategies[J]. Arid Zone Research, 2025, 42(2): 372-383.

[3] 张育超, 刘健, 姚志. 农业强国视域下中国耕地"非粮化"的动因与治理综述[J]. 安徽农业科学, 2024, 52(24): 10-16+22.
ZHANG Yuchao, LIU Jian, YAO Zhi. Review of the Motivation and Management of "Non-grain" Cultivated Land in China from the Perspective of Agricultural Powel[J]. Journal of Anhui Agricultural Sciences, 2024, 52(24): 10-16+22.

[4] 沈占锋, 骆剑承, 吴炜, 等. 采用并行均值漂移算法实现农林地块边界的精确提取[J]. 中国图象图形学报, 2011, 16(9): 1689-1695.
SHEN Zhanfeng, LUO Jiancheng, WU Wei, et al. Agricultural and forestry land boundary precise segmentation from remotesensing images by parallel mean shift algorithm[J]. Journals of Image and Graphics, 2011, 16(9): 1689-1695.

[5] TURKER M, KOK E H. Field-based sub-boundary

extraction from remote sensing imagery using perceptual grouping[J]. Isprs Journal of Photogrammetry & Remote Sensing, 2013, 79(5): 106-21.

[6] XI J, ZHANG J Z. Edge Detection from Remote Sensing Images Based on Canny Operator and Hough Transform[C]//Advances in Computer Science and Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 807-814.

[7] 张海天, 高懋芳, 任超. 基于改进Unet++的丘陵地区耕地地块深度分割与提取[J]. 航天返回与遥感, 2022, 43(4): 36-45.
ZHANG Haitian, GAO Maofang, REN Chao. Deep Segmentation and Extraction of Cultivated Land in Hilly Areas Based on Improved Unet[J]. Spacecraft Recovery & Remote Sensing, 2022, 43(4): 36-45.

[8] 马海荣, 沈祥成, 罗治情, 等. 基于Psi-Net深度学习网络的高空间分辨率遥感影像地块尺度的耕地提取[J]. 湖北农业科学, 2024, 63(11): 197-202.
MA Hairong, SHEN Xiangcheng, LUO Zhiqing, et al. High spatial resolution remote sensing image plot scale farmland extraction based on Psi-Net deep learning network [J]. Hubei Agricultural Sciences, 2024, 63(11): 197-202.

[9] 孙涛. 基于遥感时空特征融合的农田地块分割方法研究[D].南京:华中农业大学, 2024.
SUN Tao. Research on farmland plot segmentation method based on remote sensing spatio-temporal feature fusion[D]. Nanjing: Huazhong Agricultural University, 2024.

[10] 张海东, 田婷, 张青, 等. 基于GF-1影像的耕地地块破碎区水稻遥感提取[J]. 遥感技术与应用, 2019, 34(4): 785-792.
ZHANG Haidong, TIAN Ting, ZHANG Qing, et al. Study on Extraction of Paddy Rice Planting Area in Low Fragmented Regions based on GF-1 WFV Images [J]. Remote Sensing Technology and Application, 2019, 34(4): 785-792.

[11] 朱昱, 潘耀忠, 张杜娟. 基于深度卷积网络和分水岭分割的耕地地块识别方法[J]. 地球信息科学学报, 2022, 24(12): 2389-2403.
ZHU Yu, PAN Yaozhong, ZHANG Dujuan. An Agriculture Parcel Identification Method based on ConvolutionaNeural Network and Watershed Segmentation[J]. Journal of Geo-Information Science, 2022, 24(12): 2389-2403.

[12] 刘炜. 土地利用/覆被变化信息遥感图像自动分类识别与提取方法研究[D]. 杨凌:西北农林科技大学, 2012.
LIU Wei. Automatic Classification, Auto-identity classfication technology for lucc information based on remote sensing data[D]. Yangling: Northwest A&F University, 2012.

[13] LIU S, LIU L, XU F, et al. A deep learning method for individual arable field (IAF) extraction with cross-domain adversarial capability[J]. Computers and Electronics in Agriculture, 2022, 203: 107473.

[14] 项俊武. 基于深度多任务学习的高分辨率遥感影像耕地信息提取[D]. 济南:山东建筑大学,2024
XIANG Junwu. Extraction of cropland information from high-resolutionremote sensing imagery based on deep multi-task learning[D]. Jinan: Shandong Jianzhu University, 2024.

[15] 白坤, 慕晓冬, 陈雪冰, 等. 融合半监督学习的无监督遥感影像场景分类[J]. 测绘学报, 2022, 51(05): 691-702.
BAI Kun, MU Xiaodong, CHEN Xuebing, et al. Unsupervised remote sensing image scene classification based on semi-supervised learning[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(5): 691-702.

[16] 栾晓梅, 刘恩海, 武鹏飞, 等. 基于边缘增强的遥感图像弱监督语义分割方法[J]. 计算机工程与应用, 2022, 58(20): 188-196.
LUAN Xiaomei, LIU Enhai, WU Pengfei, et al. Weakly-Supervised Semantic Segmentation Method of Remote Sensinglmages Based on Edge Enhancement [J]. Computer Engineering and Applications, 2022, 58(20): 188-196.

[17] 王本礼, 王也. 基于深度学习的遥感影像地类信息获取技术现状研究[J]. 国土资源导刊, 2022, 19(04): 74-80.
WANG Benli, WANG Ye. Research on the Status Quo of Remote Sensing lmage Land TypeInformation Acquisition Technology Based on Depth Learning[J]. Land and Resources Herald, 2022, 19(4): 74-80.

[18] 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(01): 20-41.
YANG Xing, FANG Leyuan, YUE Jun. Advances in semi-supervised classification of hyperspectral remote sensing images[J]. National Remote Sensing Bulletin, 2024, 28(1): 20-41.

[19] 丁一鹏, 赵璐. 遥感图像语义分割中的弱监督域自适应算法[J]. 计算机工程与应用, 2022, 58(22): 195-202.
DING Yipeng, ZHAO Lu. Weakly-Supervised Domain Adaptive Algorithm for Remote Sensinglmage Semantic Segmentation[J]. Computer Engineering and Applications, 2022, 58(22): 195-202.

[20] Sastry S, Khanal S, Dhakal A, et al. Geosynth:

Contextually-aware high-resolution satellite image synthesis[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition . Seattle, WA, USA: IEEE, 2024: 460-470.

[21] PODELL D, ENGLISH Z, LACEY K, et al. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis[J/OL]. arXiv preprint arXiv:2307.01952, 2023 [2025-07-18]. arxiv.org/abs/2307.01952.

[22] 龚健雅, 刘青珮, 张觅, 等. RS-SegDif:基于扩散模型的遥感语义分割样本合成方法[J]. 城市勘测, 2025(1): 1-7.
GONG Jianya, LIU Qingyu, ZHANG Mi, et al. RS-SegDif: A Diffusion Model-Based Method for Remote SensingSemantic Segmentation Sample Synthesis[J]. Urban Geotechnical Investigation & Surveying, 2025, (1): 1-7.

[23] 李克文, 蒋衡杰, 李国庆, 等. 基于扩散模型的遥感图像变化检测方法[J]. 计算机工程与设计, 2025, 46(2): 337-344.
LI Kewen, JIANG Hengjie, LI Guoqing, et al. Diffusion model based change detection method for remote sensingimages[J]. Computer Engineering and Design, 2025, 46(2): 337-344.

[24] 左宪禹, 田展硕, 殷梦晗, 等. 基于残差扩散模型的遥感超分辨率图像生成研究[J]. 河南师范大学学报(自然科学版), 2025, 53(3):58-65+178.
ZUO Xianyu, TIAN Zhanshuo, YIN Menghan, et al. Remote sensing super-resolution image generation based on residual diffusion model[J]. Journal of Henan Normal University (Natural Science Edition), 2025, 53(3):58-65+178.

[25] KHANNA S, LIU P, and ZHOU L, et al. Diffusionsat: A generative foundation model for satellite imagery[J/OL]. arXiv preprint arXiv:2312.03606, 2023 [2025-07-18]. arxiv.org/abs/2312.03606.

[26] D'ANDRIMONT R, CLAVERIE M, and KEMPENEERS P, et al. AI4Boundaries: an open AI-ready dataset to map field boundaries with Sentinel-2 and aerial photography[J]. Earth System Science Data, 2023, 15(1): 317-329.

[27] Yuan Y, Chen X, Wang J. Object-contextual representations for semantic segmentation[C]//Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23 – 28, 2020, Proceedings, Part VI 16. Springer International Publishing, 2020: 173-190.

[28] Yu C, Gao C, Wang J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. International journal of computer vision, 2021, 129(11): 3051-3068.

[29] Fu J, Liu J, Tian H, et al. Dual attention network for scene segmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Piscataway, NJ: IEEE, 2019: 3146-3154.

[30] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015: 234-241.

[31] GUO M-H, LU C-Z, HOU Q, et al. Segnext: Rethinking convolutional attention design for semantic segmentation[J]. Advances in neural information processing systems, 2022, 35: 1140-1156.

[32] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision (ECCV). Munich, Germany: Springer, 2018: 801-818.