

DOI:10.19651/j.cnki.emt.2517954

改进 DeepLabV3+ 的乡村道路语义分割方法研究<sup>\*</sup>郑锐宇<sup>1,2</sup> 李兆飞<sup>1,2</sup> 黄 惟<sup>1,2</sup> 曾熙涵<sup>1,2</sup> 马 润<sup>1,2</sup>

(1. 四川轻化工大学自动化与信息工程学院 宜宾 644000; 2. 智能感知与控制四川省重点实验室 宜宾 644000)

**摘 要:** 针对遥感影像中乡村道路像素占比小、形态不规则、阴影遮挡多以及边缘模糊等问题,本文提出一种改进的 DeepLabV3+ 遥感乡村道路语义分割模型以提升乡村道路小目标和单目标分割精度。该模型以 MobileNetV3 为主干网络,减少参数量的同时提升分割精度;引入全局注意力机制,增强模型对全局信息的提取并提升泛化能力;ASPP 模块中采用深度可分离卷积替代普通卷积,降低信息损失和计算量。在自建的卫星遥感道路影像数据集上,本文提出的改进模型取得显著成果,MIoU 和 MPA 分别达到 84.45% 和 92.32%,相比原始 DeepLabV3+ 模型,分别提升 4.63% 和 6.48%。同时,模型参数量大幅减少至  $6.30 \times 10^6$ 。在公开 CHN6-CUG 数据集上的实验结果同样验证改进模型的有效性,MIoU 和 MPA 较原始 DeepLabV3+ 分别提升 3.05% 和 5.54%,达到 79.64% 和 88.13%。以上结果表明,本研究提出的改进模型在实现轻量化的同时,能够有效提升乡村道路的分割精度和效率。

**关键词:** 乡村道路;DeepLabV3+;语义分割;MobileNetV3;全局注意力机制

**中图分类号:** TP751;TN20 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Research on improving DeepLabV3+ for semantic segmentation of rural roads

Zheng Ruiyu<sup>1,2</sup> Li Zhaoifei<sup>1,2</sup> Huang Wei<sup>1,2</sup> Zeng Xihan<sup>1,2</sup> Ma Run<sup>1,2</sup>

(1. School of Automation and Information Engineering, Sichuan University of Science &amp; Engineering, Yibin 644000, China;

2. Intelligent Perception and Control Key Laboratory of Sichuan Province, Yibin 644000, China)

**Abstract:** This paper addresses the challenges of segmenting rural roads in remote sensing images, including small pixel proportion, irregular shapes, shadow occlusions, and blurred edges. To improve the segmentation accuracy of small and single-object rural roads, we propose an improved DeepLabV3+ semantic segmentation model. We employ MobileNetV3 as the backbone for parameter reduction and enhanced accuracy. A global attention mechanism is incorporated to improve global information extraction and generalization. Depthwise separable convolutions replace standard convolutions in the ASPP module to minimize information loss and computational cost. Experiments on a self-built satellite remote sensing road image dataset demonstrate significant improvements, achieving an MIoU of 84.45% and MPA of 92.32%, outperforming the original DeepLabV3+ by 4.63% and 6.48%, respectively, with a parameter size of only  $6.30 \times 10^6$ . Validation on the public CHN6-CUG dataset confirms the model's effectiveness, showing MIoU and MPA improvements of 3.05% and 5.54% to reach 79.64% and 88.13%, respectively. These results indicate that our lightweight, improved model effectively enhances rural road segmentation accuracy and efficiency.

**Keywords:** rural road; DeepLabV3+; semantic segmentation; MobileNetV3; global attention mechanism

## 0 引言

在信息技术时代,遥感技术已经成为地理信息获取的重要手段之一<sup>[1]</sup>,尤其在地形复杂、交通不便的乡村地区,传统道路测绘方法难以满足快速更新和精确管理的需求。乡村道路是连接城乡的重要纽带,对区域经济发展和居民

生活改善至关重要。然而,乡村道路狭窄、曲折且形态不规则,易受自然环境影响,传统测绘方法难以实现全覆盖和高精度提取。遥感影像以其宏观、快速和动态的特点,为乡村道路的提取、管理和规划提供了新的解决方案<sup>[2-3]</sup>。高分辨率卫星或航空影像可清晰识别道路的位置、走向和宽度等信息,实现快速提取和更新。随着遥感技术进步,高分辨

收稿日期:2025-01-17

<sup>\*</sup> 基金项目:四川省自然科学基金(2024NSFSC0770)、国家自然科学基金(42405145)、四川轻化工大学科研基金(2020RC32)项目资助

率、多光谱、多时相遥感数据为乡村道路提取提供更精确的数据支持。结合地理信息系统<sup>[4]</sup>和深度学习等技术,可进一步提高道路提取的自动化程度和准确性,为道路维护管理、智能交通、城市规划、灾害监测等领域提供数据支持。

目前,公开遥感道路数据集多基于城市道路构建,针对乡村道路的数据集构建及语义分割研究相对较少。高分辨率遥感影像中的乡村道路呈现复杂且不规则的特征,宽度仅为几个像素到十几个像素,属于小目标语义分割问题。同时,乡村道路与农田、建筑物、植被、山系等地物紧密相连,且遥感影像易被阴影遮挡,属于单目标语义分割问题。复杂光照条件和天气变化易降低影像质量,产生边缘模糊,从而降低分割精度。

传统的图像目标分割方法受限于手工设计特征的能力,难以处理复杂的地物分布和光照变化。深度学习算法虽然能够自动学习图像中的特征信息,提升分割效果,但现有方法在乡村道路遥感影像的语义分割方面仍存在局限性。Long 等<sup>[5]</sup>提出全卷积神经网络<sup>[6]</sup>实现图像分割端到端学习,通过全卷积操作,保留图像的空间信息的同时,可以处理任意尺寸的图像。但模型在上采样时,会丢失图像的特征信息。Ronneberger 等<sup>[7]</sup>提出 U-Net 网络<sup>[8]</sup>,通过由多个卷积层组成的编码器对图像的特征进行提取,再由反卷积层组成的解码器恢复图像的细节,其中在编码器和解码器之间通过跳跃连接保留低层次信息,提高分割精度。Badrinarayanan 等<sup>[9]</sup>提出 SegNet 网络,其通过编码器-解码器<sup>[10]</sup>结构实现对图像的分割,在编码器部分,使用最大池化技术来降低特征图的分辨率,从而有效减少计算量,并捕捉更高层次的特征,在解码器部分,由反池化层将低分辨率的特征图恢复到高分辨率。但在进行编-解码时,可能导致图像出现信息丢失。Zhao 等<sup>[11]</sup>提出 PSPNet 网络,使用卷积神经网络作为特征提取器的基础上引入金字塔池化模块,在不同尺度上对特征图进行池化,以捕捉全局上下文信息。Wang 等<sup>[12]</sup>提出 HRNet 网络,通过多分辨率并行分支和跨分辨率信息交换,模型在不同尺度上进行特征学习,让高分辨率细节信息保留,同时兼顾低分辨率语义信息。Chen 等<sup>[13]</sup>提出 DeepLabV1 网络,其中引入空洞卷积在不损失分辨率的情况下扩大感受野,以捕捉更多的上下文信息。此外在输出层后添加全连接条件随机场,对分割结果进行精细化处理,减少边界模糊和误分类问题。Chen 等<sup>[14]</sup>提出 DeepLabV2 网络,在 DeepLabV1 的基础上引入空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)模块,通过在不同膨胀率下应用空洞卷积,捕捉多尺度的上下文信息,提高分割结果的准确性。Chen 等<sup>[15]</sup>提出 DeepLabv3 网络,对 ASPP 模块进行优化,在模块中引入一个全局平均池化层,用于捕捉全局上下文信息。采用多尺度特征融合,有效提高分割结果的准确性。Chen 等<sup>[16]</sup>提出 DeepLabv3+ 网络,在 DeepLabv3 基础上进行优化改进,结合编码器-解码器结构对语义特征进行提取,进

一步提升分割的准确性。针对农村道路道路蜿蜒细长、易受阴影遮挡等问题,何士俊等<sup>[17]</sup>提出在 DeepLabv3+ 模型上使用 MobileNetV2 作为主干网络和在 ASPP 模块中串联 CBAM 以减少参数量和增强特征提取,添加 Dice Loss 损失函数以克服样本不平衡。针对道路分割边界模糊和遮挡等问题,张银胜等<sup>[18]</sup>提出在 DeepLabv3+ 模型基础上引入 MobileNetV3 网络和高效通道注意力机制<sup>[19]</sup>减少模型参数数量的同时,关注连续道路特征信息,采用多级上采样保留细节信息。谢国波等<sup>[20]</sup>提出 L-DeepLabV3+ 网络,将原始主干网络替换为 MobileNetV2 网络,在空洞空间卷积池化金字塔模块中引入通道空间并联注意力模块和 YOLOF 模块增强特征表达,替换深度可分离卷积减少模型参数量,结合 Dice Loss 和 Focal Loss 作为损失函数克服样本不平衡。

本文针对现有模型难以有效提取乡村道路等小目标特征、易受阴影遮挡和边缘模糊干扰,以及模型参数量过高的问题,提出一种改进 DeepLabv3+ 的乡村道路语义分割模型。在基于 GF-2 卫星和 planet 卫星遥感影像自建的乡村道路数据集以及公开 CHN6-CUG 道路数据集上展开实验。首先引入 MobileNetV3\_Large 网络替换 Xception 网络<sup>[21]</sup>作为 DeepLabV3+ 的主干网络,减少参数量,实现模型轻量化,提高模型分割的效率和准确性;其次,引入全局注意力机制(global attention mechanism, GAM)<sup>[22]</sup>,捕捉图像中全局关系的同时,整合不同位置的特征,增强特征表示的鲁棒性和丰富性,使其在面对遮挡或边缘模糊的输入数据时仍能保持较好的性能;最后,在 ASPP 模块中引入  $3 \times 3$  深度可分离卷积(depthwise separable convolution, DSC),减少传统卷积操作中信息的损失,实现高效特征提取的同时,减少参数量。通过改进,模型能更好地提取单目标乡村道路遥感语义特征信息、关注全局关系,有效应对乡村道路分割中边界模糊和阴影遮挡难以区分的问题。

## 1 改进 DeepLabV3+ 模型结构

### 1.1 模型总体结构

在遥感影像乡村道路语义分割任务中,获取乡村道路多尺度的特征和网络轻量化是提升模型性能和实际应用可行性的关键。改进的 DeepLabV3+ 模型在结构上融合 MobileNetV3 网络、GAM 模块和 DSC-ASPP 模块,增强对道路的分割能力,并降低计算量,其结构如图 1 所示。改进模型以轻量化 MobileNetV3\_Large 作为主干网络,该主干网络被广泛应用于移动端和实时场景<sup>[23]</sup>,其通过一系列卷积层和非线性激活函数提取多层次语义特征,输出高层特征图,有效减少模型参数量和计算量,为后续多尺度处理模块提供丰富的上下文信息。之后,模型引入 GAM 模块,利用全局池化操作来获取整个特征图的全局信息,随后通过学习得到的权重系数对各个通道的特征进行重新加权,能够更精确地关注具有重要语义的关键区域,抑制背景噪声

和无关特征对道路分割结果的影响,使后续空洞卷积模块从精炼的特征图中提取信息。最后,用深度可分离卷积替换 ASPP 模块中  $3 \times 3$  普通卷积,改进后的 DSC-ASPP 模块包含  $1 \times 1$  卷积分支、空洞率分别为 6, 12, 18 的  $3 \times 3$  深度

可分离卷积以及全局图像池化。深度可分离卷积将标准卷积分解为深度卷积和逐点卷积,进一步扩大卷积的感受野,使得 ASPP 模块能够捕捉到不同尺度特征,不仅能减少计算量,还能有效保留空间结构信息。

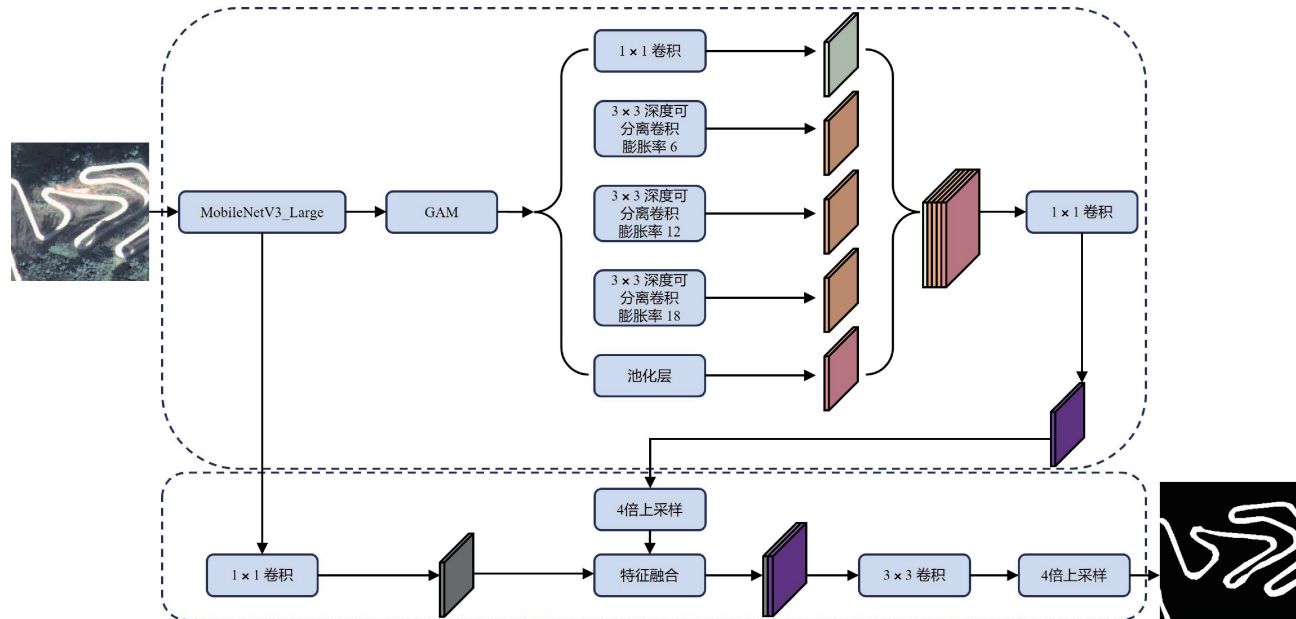


图 1 改进 DeepLabV3+ 的遥感语义分割模型结构图

Fig. 1 Architecture diagram of the improved DeepLabV3+ remote sensing semantic segmentation model

## 1.2 MobileNetV3 网络

MobileNetV3 是 Google 团队提出的一种轻量化模型,由于传统的卷积神经网络,参数量和运算量大,无法适配移动设备和嵌入式设备,为使网络模型更好地适配轻量化要求,MobileNet 网络应运而生。MobileNetV3 有 Small 与 Large 两个版本,分别对应计算和存储要求低和高的版本。在 MobileNetV2 的基础上,MobileNetV3 引入神经架构搜索、H-Swish 激活函数以及压缩和激励模块。本文选用 MobileNetV3\_Large 替换 DeepLabV3+ 模型原有的主干网络 Xception,在降低模型参数量的同时能够保证模型对乡村道路的分割精确度。MobileNetV3\_Large 网络结构如图 2 所示。

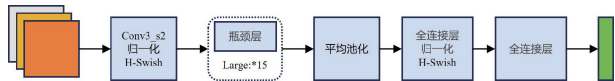


图 2 MobileNetV3 网络结构图

Fig. 2 MobileNetV3 network architecture diagram

在 MobileNetV3 网络中使用了  $H-Swish$  和  $ReLU$  两种非线性激活函数。其中用  $H-Swish$  激活函数替代原来的  $Swish$  激活函数, $H-Swish$  函数中的  $ReLU 6(x+3)$  部分可在前向传播中计算并共享,通过归一化输出至  $[0, 1]$ ,降低计算量。 $H-Swish$  和  $ReLU$  激活函数公式如式(1)和(2)所示。

$$H-Swish(x) = x \cdot \frac{ReLU6(x+3)}{6} \quad (1)$$

$$ReLU(x) = \max(0, x) \quad (2)$$

式中: $x$  表示输入值; $ReLU 6$  是  $ReLU$  函数的变体; $\max$  表示负数输入置 0,正数输入保持不变。

## 1.3 GAM 模块

在网络模型中引入注意力机制能够增强模型对重要信息的关注,减少信息的丢失,同时增强解释性来提高模型的表现。现有注意力机制常用于空间和通道作用域。其中,卷积块注意力机制(convolutional block attention module, CBAM)<sup>[24]</sup>是由通道注意力和空间注意力机制组成,实现从通道到空间的顺序注意力结构,注重局部细节的关联性特征提取,对模型的性能起到提升作用。本文引入 GAM 模块,注重全局的一致性,从而提升乡村道路遥感影像语义分割模型性能,GAM 模块结构如图 3 所示。

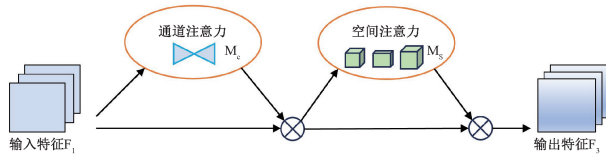


图 3 GAM 模块结构图

Fig. 3 GAM module architecture diagram

在 GAM 模块中,通道注意力子模块使用 3D 排列策



略以有效保留信息,同时通过一个两层的感知机增强跨维度的通道与空间相关性,从而显著提升特征表示能力。结构如图 4 所示。

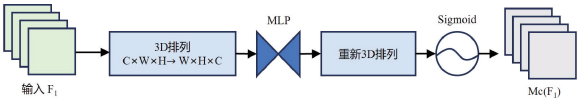


图 4 通道注意力子模块结构图

Fig. 4 Channel attention sub-module architecture diagram

在 GAM 模块中,空间注意力子模块使用两个  $7 \times 7$  卷积层来实现空间信息的深度融合,并去除最大池化操作,以有效避免因池化操作可能导致的信息丢失,从而更好地保留和利用空间特征。结构如图 5 所示。

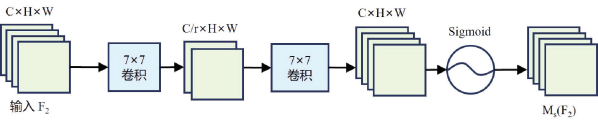


图 5 空间注意力子模块结构图

Fig. 5 Spatial attention sub-module architecture diagram

1.4 DSC-ASPP 模块

空洞空间金字塔池化是一种用于图像分割任务的特征提取模块。ASPP 模块通过不同空洞率的卷积操作,在不同尺度上捕捉图像的特征。ASPP 模块中包含  $1 \times 1$  卷积分支用于捕捉局部特征,多个空洞卷积分支用于捕捉不同尺度的上下文信息,全局平均池化分支对输入特征图进行全局平均池化,提取全局上下文信息。在除  $1 \times 1$  卷积分支外,其他 3 个分支( $3 \times 3$  卷积分支)引入深度可分离卷积。深度可分离卷积是一种轻量级的卷积操作,将标准卷积分解为深度卷积和逐点卷积,从而减少计算量和参数数量。深度可分离空洞空间金字塔池化模块结构如图 6 所示。

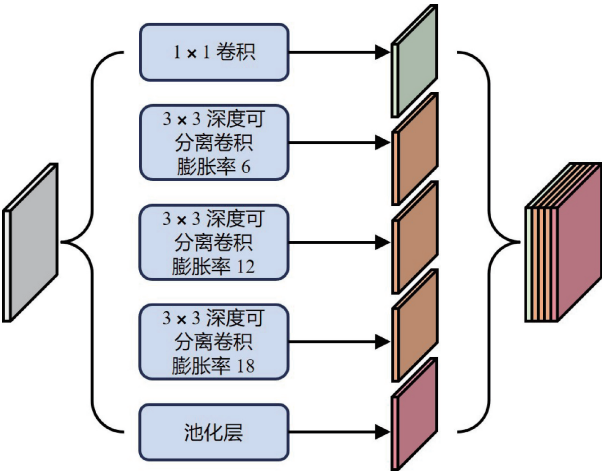


图 6 DSC-ASPP 模块结构图

Fig. 6 DSC-ASPP module architecture diagram

2 实验过程

2.1 实验数据

本文实验卫星遥感影像来源区域为四川省宜宾市。宜宾市地处四川盆地南部,地形复杂多样,山地、丘陵、河谷等地貌类型交错分布,这使得宜宾乡村道路呈现出地形复杂、等级较低、季节性影响明显等特点,为遥感影像道路提取研究提供了理想的实验区域。实验数据包括来自 Planet 卫星和 GF-2 卫星的遥感影像,共计 5 幅不同时相、不同空间分辨率及不同卫星的影像数据,影像数据参数如表 1 所示。

表 1 影像数据参数

Table 1 Image data parameters

影像数据	影像时间	影像大小
高分 2 号	2021 年 9 月	29 635×25 009,42 375×60 636
高分 2 号	2023 年 6 月	32 266×27 712,32 321×27 712
Planet	2023 年 7 月	53 336×50 604

这些影像数据涵盖了宜宾市城区及周边地区,能够有效反映不同地形、地貌和植被覆盖条件下的乡村道路特征。可用于样本标注的数字正射影像数据如图 7 所示。

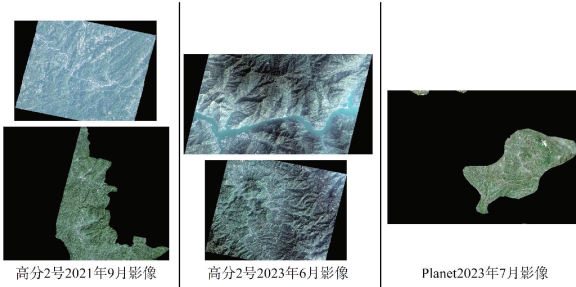


图 7 数字正射影像数据

Fig. 7 Digital orthophoto map data

本文使用 ArcGIS 软件对遥感影像中的乡村道路进行标注。标注完成后的图像被用于模型训练和验证。具体标注流程如图 8 所示。

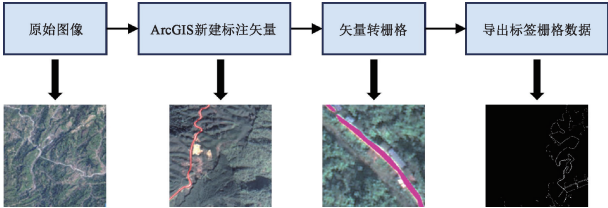


图 8 标注流程图

Fig. 8 Annotation workflow diagram

将标注完成后的图像数据通过规则网格裁剪法和滑动窗口裁剪法进行裁剪得到  $256 \text{ pixel} \times 256 \text{ pixel}$  的数据图像。但裁剪后的图像并不是所有都包含乡村道路信息,



剔除不含有乡村道路信息的图像及对应的标签,得到乡村道路样本数据 12 707 张。对裁剪后的样本数据进行旋转,翻转,添加椒盐噪声等增强操作得到最终样本数据 19 425 张。划分为 15 540 张训练集图像,1 943 张测试集图像和 1 942 张验证集图像。数据增强效果如图 9 所示。

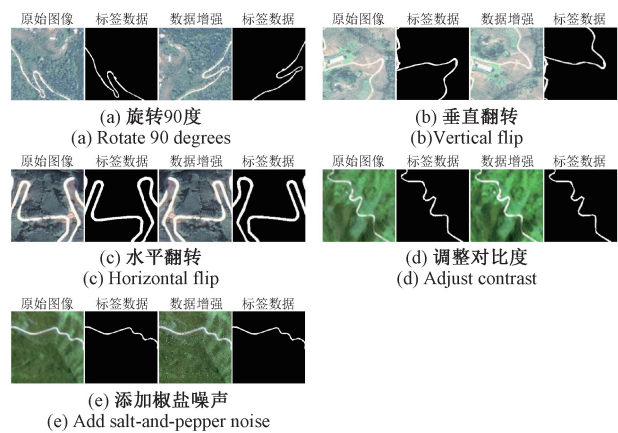


图 9 数据增强操作效果图

Fig. 9 Data augmentation operation effect images

CHN6-CUG 道路数据集<sup>[25]</sup>包含中国 6 个不同城市的卫星遥感影像道路数据,标记的道路包括铁路、公路、城市道路和农村道路等。数据集包含 4 511 张 512 pixel × 512 pixel 的标注图像,划分为 3 609 张训练集图像,451 张测试集图像和 451 张验证集图像。

2.2 实验环境

本文实验方法实现基于 Windows11 操作系统,使用 Pytorch 深度学习框架,版本为 1.12.1,编程语言为 Python 3.8。CPU 为 Intel(R) Core(TM) i5-12600KF, GPU 为 NVIDIA GeForce RTX4060Ti(16 GB)。初始学习率为 0.01,优化器选用 SGD,每次训练时的图像批量 batch 大小为 16,总共进行 200 轮迭代训练。

2.3 评价指标

本文在基于混淆矩阵的基础上对卫星遥感影像道路数据的分割性能进行评估,选取平均交并比(mean intersection over union, MIoU)和类别平均像素准确率(mean pixel accuracy, MPA)作为评价指标。

MIoU 用于衡量预测结果与真实标签之间的重叠程度,计算每个类别的交并比,并对所有类别的交并比取平均值,来评估模型的整体分割性能,公式如式(3)所示。

$$MIoU = \frac{1}{N} \sum_{i=1}^N \frac{P_{ii}}{\sum_{j=1}^N P_{ij} + \sum_{j=1}^N P_{ji} - P_{ii}} \tag{3}$$

MPA 用于衡量预测结果与真实标签匹配的像素所占的比例,计算每个类别的准确率,并对所有类别的准确率取平均值,公式如式(4)所示。

$$MPA = \frac{1}{N} \sum_{i=1}^N \frac{P_{ii}}{\sum_{j=1}^N P_{ij}} \tag{4}$$

式中: N 表示类的数量; P<sub>ij</sub> 表示将类 i 预测为类 j 的像素数量; P<sub>ii</sub> 表示被正确预测的像素数量; P<sub>ji</sub> 表示将类 j 预测为类 i 的像素数量。

3 实验结果分析

3.1 消融实验

消融实验结果如表 2 所示。在传统 DeepLabV3+ 模型的基础上,将原有主干网络 Xception 替换为轻量化 MobileNetV3 网络。MobileNetV3 在 MobileNetV2 的基础上对网络结构进行精简,采用 H-Swish 和 ReLU 两种非线性激活函数,模型的参数量从 54.71 × 10<sup>6</sup> 减小到 5.90 × 10<sup>6</sup>, MIoU 和 MPA 分别提升 2.09% 和 3.29%, 将原有主干网络替换后实现模型的轻量化和性能提升。引入 GAM 模块使得模型能够通过全局的上下文信息增强特征图的关键区域,同时此注意力机制能够动态地为不同的特征通道和空间位置分配权重,在模型的分割精度提升的同时,参数量也会略微增加。由实验 1、4 和 2、5 对比可知,将 ASPP 模块中的普通卷积替换为深度可分离卷积后,虽然模型的参数量会显著减少,但是模型分割精度轻微降低。MobileNetV3 网络和 GAM 模块同时作用,模型分割精度得到进一步提升。MobileNetV3 网络和 DSC-ASPP 模块同时作用时,模型的参数量仅为 3.74 × 10<sup>6</sup>, 但模型分割精度并不理想。改进 DeepLabV3+ 模型在 MIoU 和 MPA 分别达到了 84.45% 和 92.32%, 相对于原 DeepLabV3+ 模型分别提升 4.63% 和 6.48%, 参数量减小到 6.30 × 10<sup>6</sup>, 模型在分割性能和分割效率得到提升。

表 2 消融实验结果

Table 2 Ablation experiment results

序号	MobileNetV3	GAM	DS-ASPP	MIoU/%	MPA/%	参数量/(×10 <sup>6</sup> )
1				79.92	85.84	54.71
2	✓			82.01	89.13	5.90
3		✓		81.92	88.49	57.07
4			✓	79.84	85.37	42.18
5	✓	✓		83.05	90.95	8.26
6	✓		✓	81.97	88.86	<b>3.74</b>
7	✓	✓	✓	<b>84.45</b>	<b>92.32</b>	6.30

3.2 对比实验

在相同环境配置下,不同语义分割模型和本文模型在自建卫星遥感影像乡村道路数据集进行训练和验证,结果如表 3 所示。改进后的轻量化 DeepLabV3+模型的 MIoU 和 MPA 分别达到 84.45% 和 92.32%,参数量减小到  $6.30 \times 10^6$ 。PSPNet 模型进行训练时选择 MobileNetV2 作为主干网络,此时模型参数量为  $5.80 \times 10^6$ ,但模型的分割精度最低,并不适用于高分辨率遥感影像的乡村道路分割。HRNet 和 SegNet 虽然相较于 PSPNet 在精度有小幅提升,但模型的分割精度并未达到理想的效果,同时参数量增加,导致模型训练时间较长。U-Net 在分割精度和模型参数量有所改善,但道路分割的准确性依旧提升不多。传统 DeepLabV3+采用 Xception 作为主干网络,模型参数量达到  $54.71 \times 10^6$ ,MIoU 和 MPA 仅为 79.92% 和 85.84%,模型参数量大导致计算复杂,训练时间长。文献[17]提出改进后的 DeepLabV3+和 L-DeepLabV3+在分割精度和模型参数量方面取得了进步,但分割准确率仍有提升空间。本文提出的改进 DeepLabV3+模型选择 MobileNetV3 作为主干网络,引入 GAM 模块和深度可分离卷积,实现模型轻量化,提高模型对乡村道路遥感影像语义分割的精度。

通过选取 planet 和 GF-2 卫星遥感影像,将不同模型在自建乡村道路遥感影像数据的语义分割效果进行对比,乡村道路分割可视化效果如图 10 所示。PSPNet、HRNet、SegNet 和 U-Net 基础模型乡村道路分割中都表现出一定程度的不足,尤其是在道路连通性和边界清晰度方面,同

表 3 不同语义分割模型对比实验结果

Table 3 Comparative experimental results of different semantic segmentation models

模型	MIoU/ %	MPA/ %	参数量/ ( $\times 10^6$ )
PSPNet	76.35	81.33	5.80
HRNet	77.90	82.49	29.54
SegNet	78.26	82.76	11.56
U-Net	79.87	85.58	24.89
DeepLabV3+	79.92	85.84	54.71
文献[17]	81.75	87.31	<b>3.51</b>
L-DeepLabV3+	82.67	89.39	5.63
本文模型	<b>84.45</b>	<b>92.32</b>	6.30

时易受到植被遮挡影响造成道路分割不连续。对于乡村道路宽度变化适应性较差,易出现过度分割。这些模型缺少对全局上下文信息捕捉和有效的特征提取,无法准确对乡村道路的细节信息进行提取,难以有效处理乡村道路场景的复杂性和多样性。DeepLabV3+以及改进的模型在道路连通性和边界清晰度方面有所提升,同时具备一定的全局上下文信息捕捉能力,但在处理复杂场景时仍存在局限性,无法完全解决乡村道路场景下的挑战。本文改进的模型在道路连通性、边界清晰度和道路宽度一致性方面均优于其他模型,能够有效地应对植被遮挡、道路蜿蜒曲折和道路宽度变化等挑战。

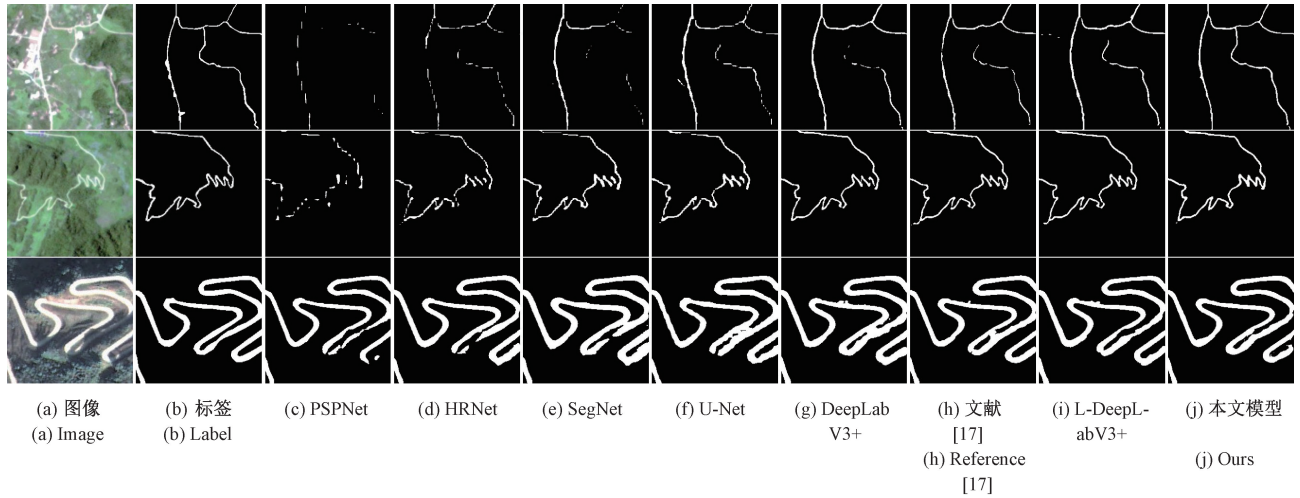


图 10 不同模型分割效果对比

Fig. 10 Comparison of segmentation outputs from different models

在相同环境配置下,不同语义分割模型和本文模型在CHN6-CUG 道路数据集进行训练和验证,结果如表 4 所示。改进后的轻量化 DeepLabV3+模型的 MIoU 和 MPA 分别达到 79.64% 和 88.13%,相较于原 DeepLabV3+模型分别提升了 3.05% 和 5.54%。PSPNet、HRNet、SegNet

和 U-Net 模型的 MIoU 和 MPA 值相对较低,这些模型在处理复杂的道路场景时,分割精度还有待提高。DeepLabV3+以及改进的模型在 MIoU 和 MPA 值上均优于基础模型,其中的空洞卷积和 ASPP 模块能更好地捕捉多尺度上下文信息方面。本文改进模型采用 MobileNetV3 的

轻量化设计降低计算复杂度,引入 GAM 注意力机制则提升模型对全局上下文信息的理解,有利于更精确地分割。深度可分离卷积在减少计算量的同时,维持模型的分割精度。

表 4 CHN6-CUG 数据集上不同语义分割模型对比实验结果

Table 4 Comparative experimental results of different semantic segmentation models on the CHN6-CUG dataset		
模型	MIoU/%	MPA/%
PSPNet	72.78	77.33
HRNet	73.90	78.49
SegNet	74.26	79.76
U-Net	76.12	81.58
DeepLabV3+	76.59	82.59
文献[17]	79.03	85.09
L-DeepLabV3+	79.18	87.76
本文模型	79.64	88.13

CHN6-CUG 道路数据集分割可视化效果如图 11 所示。PSPNet、HRNet、SegNet 和 U-Net 基础模型在城市道路的分割中,容易丢失很多细节信息。清晰的道路交叉口被模糊化,道路边缘呈现粗糙。在建筑阴影下,易产生误判。在对乡村道路的分割中,同样存在连通性问题,容易受到植被阴影遮挡的影响,道路边缘模糊。DeepLabV3+ 以及改进的模型在基础模型上有所改进,细节信息保留相对更好,对阴影遮挡的分割效果也更强,但分割效果仍有提升空间。在乡村道路分割中,道路连通性和边界清晰度上有所提升,但植被阴影遮挡的干扰依旧存在。本文改进模型在城市道路的分割中,能够提取到更多的细节信息,道路交叉口清晰,能够较为准确地地区分道路和建筑。在乡村道路分割中,道路连通性、边界清晰度和抗植被阴影遮挡干扰的能力均优于其他模型。能够准确分割道路的走向和边界,即使在植被遮挡的情况下也能保持较好的分割效果。

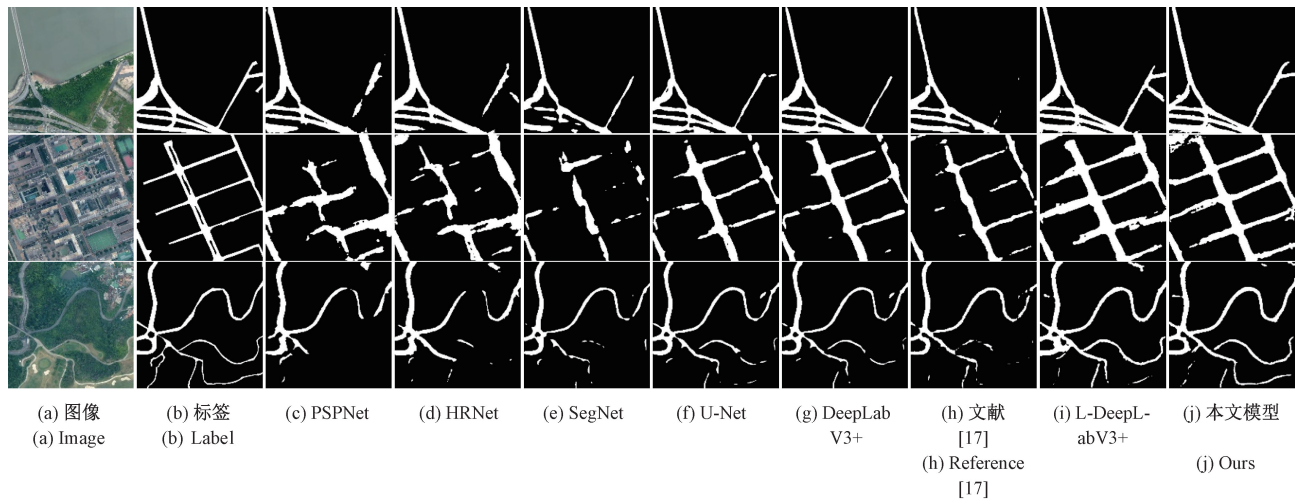


图 11 CHN6-CUG 数据集上不同模型分割效果对比  
Fig. 11 Comparison of segmentation results from different models on the CHN6-CUG dataset

4 结 论

针对乡村道路遥感影像存在狭窄、不规则、阴影遮挡和边缘模糊等特性,以及传统 DeepLabV3+ 语义分割模型精度不高和参数量大,对乡村道路小目标和单目标语义分割的效果不佳等问题。在 DeepLabV3+ 模型的基础上引入 MobileNetV3\_Large 网络作为主干网络,在大幅度减少参数量的同时能够提升模型的分割精度。引入 GAM 注意力模块增强模型对于全局信息的提取,整合不同位置的特征,模型的泛化能力得到提升。在 ASPP 模块,将 3×3 普通卷积替换为深度可分离卷积,在减少传统卷积操作信息损失的同时,还能显著降低模型的参数量和计算量,使得模型更加轻量化和高效。改进模型在自建乡村道路卫星遥感影像数据集上对乡村

道路进行分割效果显著,MIoU 和 MPA 分别为 84.45% 和 92.32%,模型参数量为  $6.30 \times 10^6$ ,模型轻量化同时,分割精度和效率得到提升。此外,改进模型也在 CHN6-CUG 数据集上有分割效果不错,MIoU 和 MPA 分别达到 79.64% 和 88.13%。实验结果表明,即使遥感图像中道路狭窄和存在阴影遮挡时,改进模型也能对道路进行有效分割。

由于数据标注不仅耗时且成本高昂,并且在标注过程中易出现错标、误标,后期正考虑将深度学习中的自监督学习和弱监督学习策略运用在数据标注过程,降低乡村道路数据集制作的成本,提高数据标注的效率和精度,进一步提升遥感图像语义分割能力。另外,后续也考虑将不同分辨率的卫星及航空遥感影像数据进行融合,实现对乡村道路更精确语义分割。



## 参考文献

- [1] 张静, 景海涛, 樊盛华. 基于海岸线数据库的遥感图像海陆分割[J]. 电子测量技术, 2020, 43(23): 115-120.  
ZHANG J, JING H T, FAN SH H. Sea-land segmentation for remote sensing imagery based on coastline database [J]. Electronic Measurement Technology, 2020, 43(23): 115-120.
- [2] 张晋赫, 秦育罗, 张在岩, 等. 复杂场景下农村道路裂缝分割方法[J]. 测绘通报, 2022(5): 74-78, 88.  
ZHANG J H, QIN Y L, ZHANG Z Y, et al. Rural road crack segmentation method in complex scene[J]. Bulletin of Surveying and Mapping, 2022(5): 74-78, 88.
- [3] 赵营刚. 遥感影像技术在地图编制中的应用[J]. 智能建筑与智慧城市, 2024(2): 49-51.  
ZHAO Y G. Application of remote sensing image technology in map compilation[J]. Intelligent Building & Smart City, 2024(2): 49-51.
- [4] 陈金宝. 遥感和地理信息系统技术在地理环境中的应用研究[J]. 工程建设与设计, 2024(17): 155-157.  
CHEN J B. Research on the application of remote sensing and geographic information system technology in geographical environment [J]. Construction & Design for Engineering, 2024(17): 155-157.
- [5] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [6] 吉茹, 张银胜, 杨宇龙, 等. 基于多尺度特征融合的改进型云图分割方法[J]. 国外电子测量技术, 2022, 41(11): 37-44.  
JI R, ZHANG Y SH, YANG Y L, et al. Improved cloud image segmentation method based on multi-scale feature fusion [J]. Foreign Electronic Measurement Technology, 2022, 41(11): 37-44.
- [7] 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, 2015: 234-241.
- [8] KATSAMENIS I, DOULAMIS N, DOULAMIS A, et al. Simultaneous precise localization and classification of metal rust defects for robotic-driven maintenance and prefabrication using residual attention U-Net [J]. Automation in Construction, 2022, 137: 104182.
- [9] 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.
- [10] ZHU Y T, LONG L H, WANG J J, et al. Road segmentation from high-fidelity remote sensing images using a context information capture network [J]. Cognitive Computation, 2022, 14(2): 780-793.
- [11] ZHAO H SH, SHI J P, QI X J, et al. Pyramid scene parsing network[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.
- [12] WANG J D, SUN K, CHENG T H, et al. Deep high-resolution representation learning for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(10): 3349-3364.
- [13] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [J]. Computer Science, 2014(4): 357-361.
- [14] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CFRs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [15] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. ArXiv preprint arXiv: 1706. 05587, 2017.
- [16] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. European Conference on Computer Vision, 2018: 801-818.
- [17] 何士俊, 肖提荣, 夏既胜. 基于改进 DeepLabv3+模型的农村道路提取方法研究[J]. 云南大学学报(自然科学版), 2024, 46(3): 486-495.  
HE SH J, XIAO T R, XIA J SH. Research on rural road extraction method based on improved DeepLabv3+ model [J]. Journal of Yunnan University (Natural Sciences Edition), 2024, 46(3): 486-495.
- [18] 张银胜, 单梦姣, 钟思远, 等. 基于改进 DeeplabV3+的遥感图像道路分割模型[J]. 国外电子测量技术, 2024, 43(1): 189-198.  
ZHANG Y SH, SHAN M J, ZHONG S Y, et al. Road segmentation model of remote sensing image based on improved DeeplabV3+ [J]. Foreign Electronic Measurement Technology, 2024, 43(1): 189-198.
- [19] WANG Q L, WU B G, ZHU P F, et al. ECA-Net:

Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.

[20] 谢国波, 何林, 林志毅, 等. 基于 L-DeepLabv3+ 的轻量化光学遥感图像道路提取[J]. 激光杂志, 2024, 45(3): 111-117.

XIE G B, HE L, LIN ZH Y, et al. Lightweight optical remote sensing image road extraction based on L-DeepLabv3+[J]. Laser Journal, 2024, 45(3): 111-117.

[21] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.

[22] LIU Y CH, SHAO Z R, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions[J]. ArXiv preprint arXiv: 2112.05561, 2021.

[23] 张乐. 遥感图像道路提取模型改进及轻量化研究[D]. 西安:西安邮电大学, 2023.

ZHANG L. Study on improvement and lightweight of road extraction model in remote sensing image[D]. Xi'an:Xi'an University of Posts & Telecommunications, 2023.

[24] WOO S Y, PARK J CH, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision, 2018: 3-19.

[25] ZHU Q Q, ZHANG Y N, WANG L Z, et al. A global context-aware and batch-independent network for road extraction from VHR satellite imagery[J]. ISPRS Journal of Photogram Metry and Remote Sensing, 2021, 175(12): 353-365.

作者简介

郑锐宇, 硕士研究生, 主要研究方向为图像处理。

E-mail:1621070801@qq.com

李兆飞(通信作者), 博士, 副教授, 主要研究方向为数据驱动的装备故障诊断与目标检测。

E-mail:lzf825@suse.edu.cn