

基于改进 DeeplabV3+ 的遥感图像道路分割模型^{*}

张银胜^{1,2} 单梦姣¹ 钟思远² 陈戈¹ 童俊毅¹ 单慧琳^{1,2}

(1. 南京信息工程大学电子与信息工程学院 南京 210044; 2. 无锡学院电子信息工程学院 无锡 214105)

摘要:针对遥感图像道路分割边界模糊和遮挡难以区分的问题,提出了基于改进 DeeplabV3+ 的遥感图像道路分割模型。该模型在主干网络中引入 MobileNetV3 和高效通道注意力机制(ECA),减少了参数量并关注连续的道路特征信息。在解码过程中采用多级上采样,增强了编码器和解码器之间的紧密连接,全面保留了细节信息。同时,在 ASPP 模块中采用深度可分离膨胀卷积 DS-ASPP,显著减少了参数量。实验结果表明,该模型在 Massachusetts Roads 数据集上的交并比达到了 83.71%,准确率达到 93.71%,分割精度最优,模型参数量为 55.57×10^6 ,能够有效地避免边界模糊和遮挡导致的错漏检问题,在遥感道路分割中提高了精度和速度。

关键词:遥感图像;道路分割;DeeplabV3+ 模型;MobileNetV3 模型;多级上采样

中图分类号: TN2 **文献标识码:** A **国家标准学科分类代码:** 520.604

Road segmentation model of remote sensing image based on improved DeeplabV3+

Zhang Yinsheng^{1,2} Shan Mengjiao¹ Zhong Siyuan² Chen Ge¹ Tong Junyi¹ Shan Huilin^{1,2}

(1. College of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. School of Electronic Information Engineering, Wuxi University, Wuxi 214105, China)

Abstract: In response to the problem of fuzzy boundaries and difficulty in distinguishing occlusions in road segmentation of remote sensing images, this paper proposes a remote sensing image road segmentation model based on an improved DeeplabV3+. The model introduces MobileNetV3 and ECA attention mechanism in the backbone network to reduce parameter volume and focus on continuous road feature information. In the decoding process, multi-level upsampling is adopted to enhance the tight connection between the encoder and decoder, fully preserving detailed information. Meanwhile, deep separable dilated convolution (DS-ASPP) is used in the ASPP module to significantly reduce the number of parameters. The experimental results demonstrate that the model achieves an intersection over union (IoU) of 83.71% and an accuracy of 93.71% on the Massachusetts Roads dataset. With a parameter count of 55.57×10^6 , the model exhibits superior segmentation accuracy and effectively avoids errors and omissions caused by boundary blurring and occlusion. It enhances both precision and speed in remote sensing road segmentation.

Keywords: remote sensing images; road segmentation; DeeplabV3+ model; mobileNetV3 model; multi-level upsampling

0 引言

遥感图像语义分割是遥感图像研究中的一个重要问题,其目标是将图像划分为多个同质区域,并赋予它们相应的标签^[1]。准确的道路提取对于其他物体的检测也具

有重要影响,如海面舰船检测^[2]、土地覆盖分类^[3]、水稻倒伏^[4]和光伏足迹^[5]等。因此,在图像道路分割领域中,遥感技术的应用日益重要,其优势在于覆盖范围广且具备高空间分辨率,可以提供更多的细节信息,有助于准确地分割出目标边界和内部结构。但同时也伴随着大数据量的

收稿日期:2023-08-13

^{*} 基金项目:国家自然科学基金(62071240)、江苏省产教融合型一流课程(2022-133)、无锡学院 2023 年教改研究课题(XYJG2023010,XYJG2023011)项目资助

缺点,需要考虑处理大数据的模型参数量,以提高效率和准确性。但是遥感图像中道路的多样性、狭窄性、稀疏性以及易受外部环境干扰等特征为道路提取任务制造了困难^[6]。

随着深度学习技术的不断发展,卷积神经网络被广泛应用于语义分割领域^[7]。深度神经网络(deep neural network,DNN)具有多层次的神经网络结构^[8],为语义分割提供了分层表示^[9]。Long等^[10]提出全卷积神经网络(fully convolutional networks,FCN)^[11],是首个端对端像素级的图像语义分割模型^[12]。但模型直接采用上采样,忽视了遥感图像道路的细节特征信息。Ronneberger等^[13]提出了U-Net^[14]网络,该网络结合特征提取层提取的信息,在上采样过程中恢复原始图像,有助于准确地分割道路,但基于U形结构的方法不足以提取目标特征,导致对上下文信息的考虑不完整。Badrinarayanan等^[15]提出了SegNet网络,采用编码—解码架构(encoder-decoder)^[16]。编码器通过池化层逐渐减小输入特征映射的大小,解码器与编码器对称,通过上采样逐渐恢复图像细节和特征映射的大小。但编码器—解码器结构可能导致在遥感图像分割小目标时出现信息丢失,网络可能无法保留细节信息。Zhao等^[17]提出了PSPNet,它整合了不同地区的特征信息,采用金字塔的结构来获取图像多尺度特征。然而,上述网络未充分利用图像空间环境,导致无法更好的处理遥感图像中的细节特征。Chen等^[18]提出了DeepLabv3+网络,该模型使用DeepLabv3与编码器—解码器框架或空间金字塔池化元素进行语义分割^[19]。但DeepLab系列网络仍存在小目标分割不精确和边界信息较粗糙的问题。Wu等^[20]提出了密集—全局—残差网络DGRN的深度学习模型,该模型减少了空间信息的损失并增强了上下文感知。但在提取过程中由于更加关注全局信息而忽略一些细节信息,导致少量道路遗漏。何哲等^[21]在U-Net网络的基础上,提出了HSA-Unet(hybrid scale attention-UNet)网络模型,在编码网络使用注意力残差学习单元提升网络对道路特征的提取,增强对不同尺度道路信息的提取能力。但是对背景信息与显著信息的区分存在较大误差,应当增加模型对于图像中细节部分的处理。苏志鹏等^[22]提出了一种结合特征图切分模块和注意力模块遥感影像语义分割网络(AFSM-Net),引入通道和空间注意力模块,使模型更加关注显著的特征信息,而忽略无关的背景信息,同时结合膨胀卷积来捕获多尺度特征信息。但对于遥感道路中存在遮挡严重或相似场景时,模型难以区分道路和背景信息,易导致局部道路不连续的问题。

为了应对遥感道路图像在语义分割过程中,存在小目标信息缺失导致分割边界粗糙模糊和物体遮挡或相似场景难以区分道路和背景信息的问题,本文提出了一种改进的遥感图像道路分割模型。首先,引入了改进的 Mobile-

NetV3作为主干网络,并采用了高效通道注意力机制(efficient channel attention,ECA)。这样做有助于更好地捕捉输入数据的丰富特征,并具备较好的鲁棒性。其次,为了实现高效的特征提取,提出了多级上采样,以更彻底地保留目标的位置和边界信息,可以有效地提升模型对图像中小目标物体的分割能力。最后,在ASPP模块中,使用了 3×3 深度可分离膨胀卷积(dense separable dilated convolution,DSDC)来实现高效的特征提取。通过这些改进措施,使模型更加关注局部信息以及上下文相关性,从而解决了道路分割中边界模糊和阴影遮挡难以区分的问题,并且能够保证道路分割的精度和效率。

1 网络模型

1.1 总体结构

由于遥感图像在分割的过程中具有高分辨率和大范围覆盖的特点,为了能够更精准提取出图像中的边缘细节特征,需要更深的网络层次。然而,遥感图像的数据量会随着网络层次的加深,带来运算量的增加和分割精度下降的问题。为了快速且准确地对遥感图像进行分割,提出基于改进DeepLabV3+的遥感图像道路分割模型。如图1所示,采用编码—解码结构。其中编码器负责特征抽取,解码器负责恢复图像的空间信息并给每个像素分类^[23],以获取多尺度的上下文信息。

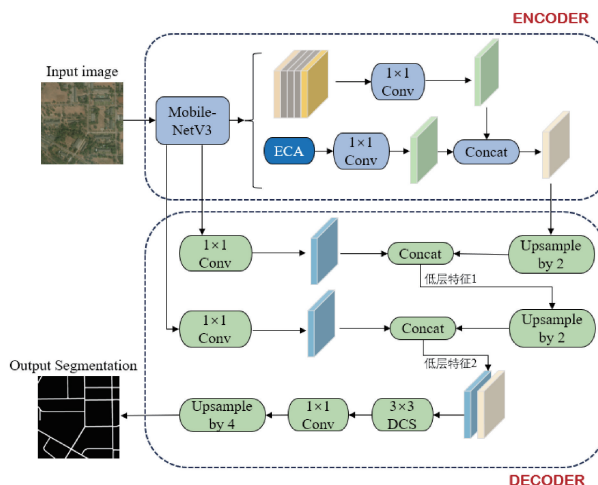


图1 改进的DeepLabV3+模型结构

引入带有注意力模块的MobileNetV3主干网络,以提取特征并输出原图 $1/4$ 、 $1/8$ 和 $1/16$ 大小的特征图。为了避免降低模型精度的可能性,在主干络中采用ECA注意力模块,并在编码结构中引入ECA。与ASPP模型提取的特征进行融合,将输出的高级特征输入解码结构中。

在解码模块中,使用 1×1 卷积来减少原图 $1/8$ 和原图 $1/4$ 的特征图的通道数。对于原图 $1/8$ 的特征图,将其与经过双线性两倍上采样的高级特征完全融合。同样地,对于原图 $1/4$ 的特征图,将其与经过双线性两倍上采样的

低层特征 1 完全融合。为了减少模型参数量,将改进的 DeepLabv3+ 模型中的卷积调整为深度可分离卷积。此外,模型还使用深度可分离卷积对低层特征 2 进行处理,以还原高级空间信息。通过多级上采样处理后,高层和低层特征相互补充,从而更好地进行特征融合,以获得更多有用的信息。这一方法明显提升了模型性能,提高了遥感图像道路分割的准确度。

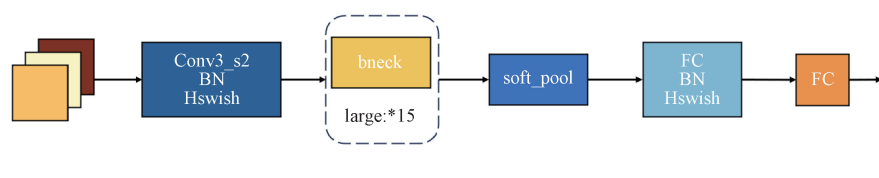


图2 MobileNetV3 基本单元

MobileNetV3 中使用了两种不同的非线性激活函数: ReLU 和 H-swish。其中, χ 表示输入的值, $\max(a, b)$ 是一个函数, 表示返回 a 和 b 中较大的值。ReLU 是一个函数, 表示进行 ReLU 操作。

ReLU 的公式如下:

$$\text{ReLU}[\chi] = \max(0, \chi) \quad (1)$$

H-swish 的公式如下:

$$\text{H-swish}[\chi] = \chi(\text{ReLU6}(x+3))/6 \quad (2)$$

式中: χ 表示输入的值; $\max(a, b)$ 是一个函数, 表示返回 a 和 b 中较大的值; ReLU 是一个函数, 表示进行 ReLU 操作。

1.3 ECA 注意力机制

注意力机制起到提取图像细节, 减少背景对目标特征的影响。为了解决道路边界模糊或难以区分的问题, 采用 ECA 注意力机制模块代替 SE 注意力模块。

ECA-MobileNetV3 网络结构如图 3 所示。在分割过程中, 该网络能够更加关注连续的道路特征信息, 抑制背景信息的干扰, 实现特征图的信道自适应目标。这样一来, 即使在有光照变化和复杂背景的情况下, 网络也能够有效地捕捉到道路的视觉特征, 使得道路边界更加清晰可见, 提高了道路分割的准确性。

ECA 模块如图 4 所示。该模块通过使用全局平均池化技术(global average pooling, GAP)来获得特征, 通过一维卷积核获取通道权重。在其中, H 、 W 和 C 分别表示了特征图的高度、宽度和通道数。

ECA 表达式为:

$$\tilde{x} = x \times \sigma \sum_{j=1}^k \omega^j y_i^j, y_i^j \in \Omega_i^k \quad (3)$$

$$y = \frac{1}{W \cdot H} \sum_{i=1}^W \sum_{j=1}^H x(i, j) \quad (4)$$

$$\omega = \sigma(C1D_k(y)) \quad (5)$$

式中: x 是 ECA 注意力机制的输入特征; y 是全局平均池化操作后得到的通道信息; $C1D$ 表示一维卷积; σ 是 sig-

1.2 MobileNetV3 主干网络

相比其他网络, MobileNetV3 具有更少的参数, 同时尽量避免丢失一些有用信息。基本单元如图 2 所示。为进一步提高网络的预测能力, 采用 ECA 注意力机制替换 SE(squeeze-and-excitation)模块, 并对 MobileNetV3 的网络末端做出进一步优化, 采用软池化(soft pool)来提高网络模型的检测性能。

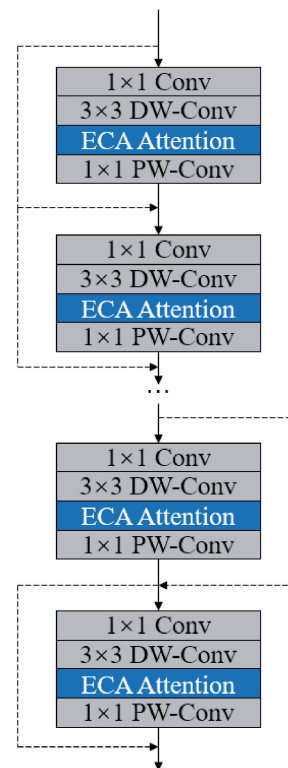


图3 ECA-MobileNetV3 网络结构

moid 激活函数; Ω_i^k 是通道为 y 的 k 个相邻通道集合; y_i^j 是通道集合的其中一个。

自适应 kernel_size:

$$C = \varphi(k) = 2^{(\gamma \cdot k - b)} \quad (6)$$

$$K = \psi(C) = \left\lceil \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (7)$$

式中: $\lceil X \rceil_{\text{odd}}$ 为选择最近的奇数; $\gamma=2$; $b=1$; C 为通道维度; K 为内核大小。

1.4 多级上采样模块

DeeplabV3+ 在解码过程中通过 4 倍上采样直接恢复

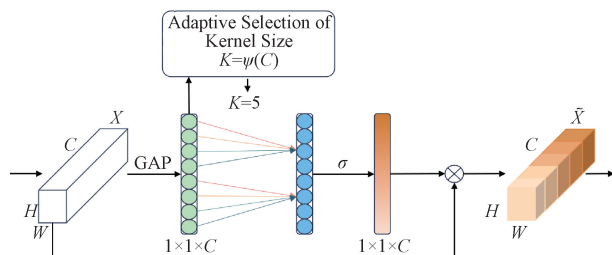


图4 ECA模块结构

特征图以获得高级语义特征。这种解码方法应用于普通图像时具有良好的性能,但对于遥感图像会丢失很多细节信息,从而使得网络的分割性能不够好。因此,本文在解码器模块中设计了多级上采样模块,如图5所示。

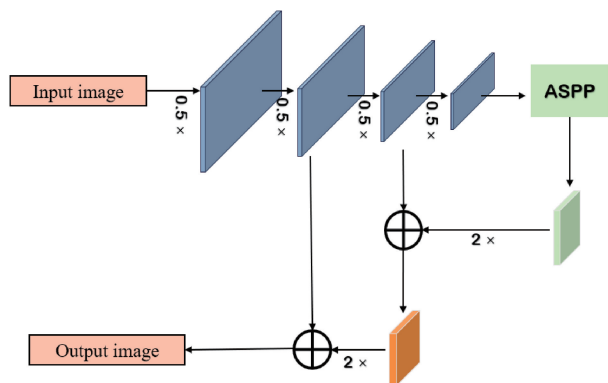


图5 解码一多级上采样模块

为了在编码器—解码器结构中提高特征融合的效果,本文设计了一个解码一多级上采样模块。该模块通过两次2倍上采样操作,将ASPP模块输出的高级语义特征进行上采样。每次上采样后,将语义特征与编码器中相同大小的低层特征图进行求和操作,从而增强编码器和解码器之间的连接。在特征融合过程中,这种连接求和操作能够保持通道数不变,同时减少参数量。而且,由于低层特征具有更多的位置信息和边界信息,通过与解码器的高级语义特征进行补偿,能够更好地保留目标的位置信息和边界信息。因此,解码一多级上采样模块的设计能够有效地实现特征融合,提高编码器和解码器之间的紧密连接,并且更彻底地保留目标的位置和边界信息。

1.5 ASPP 模块

在ASPP模块中,将3个 3×3 膨胀卷积调整为3个 3×3 深度可分离膨胀卷积(DSDC),膨胀系数分别为3、6和9。DS-ASPP模块如图6所示。

ASPP模块的目标是捕获不同尺度的上下文信息,以更好地理解图像中的对象。通过采用深度可分离膨胀卷积,调整膨胀系数,可以进一步增加感受野,从而使网络能够更好地理解输入图像的上下文信息。

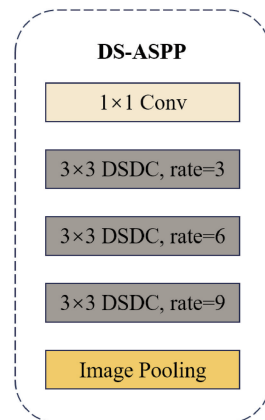


图6 DS-ASPP模块

2 实验结果与分析

2.1 实验数据及参数

本文实验数据来自 Massachusetts Roads^[24] 遥感道路语义分割数据集,该数据集来自马萨诸塞州的1171张航空影像,数据集中的影像原图及其相应标签示例如图7所示,包含影像原图和标签图。其中,红色框表示物体遮挡或狭窄道路,蓝色框表示背景与道路像素相近,绿色框表示小目标信息。影像原图分别在城市、郊区和乡村拍摄。图7(a)为城市中的遥感影像,主要包括街道、建筑物以及交通工具,城市道路宽阔清晰,但道路灰度与地面灰度非常接近,存在背景与道路相似难以区分的问题。图7(b)为在郊区所拍摄的影像,图片包括道路、住宅和湖泊等自然景观。道路种类更加多样化,存在背景和特征难以区分、树木阴影遮挡以及小目标信息易丢失的问题。图7(c)为乡村影像,包括小路、农田、森林和小镇。影像中多为窄小道路,道路颜色与地面颜色相近,存在大量的树木阴影的干扰^[25]。每张图像的大小为 1500×1500 pixels,覆盖面积为 2.25 km^2 ,并重新缩放为 1 pixel/m^2 。标签图包含两个类别:“道路”和“背景”。为了适应网络训练的要求,

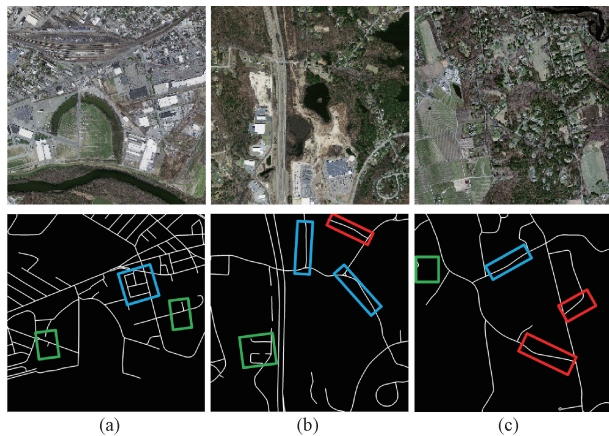


图7 数据集中的影像原图及其相应标签示例

原始图像被裁剪成 512×512 pixels 的大小,并将图像转换为 JPG 格式,标签图像为 PNG 格式。为了增加数据集的规模并提高模型的泛化能力,采用了多种图像增强技术,如翻转、平移、缩放和剪切等。最终得到了 12 452 张遥感道路图像以及对应的标签图像。然后将其按照 9:1 的比例划分为训练集和测试集。实验环境的配置信息如表 1 所示。

表 1 配置信息

配置名称	参数
CPU	Intel(R) Core(TM) i9-12900KF
GPU	NVIDIA GeForce RTX3090
操作系统	Windows10
框架	Pytorch1.12.0
Batch Size	16
Epochs	500
优化器	SGD
学习率	0.000 1
学习率下降方式	cos

2.2 评估标准

本文选取准确率(accuracy, P)、召回率(recall, R)、F1 分数(F1)、交并比(intersection over union, IoU)作为衡量模型精度高低的指标,选取模型参数量作为衡量模型大小的指标。

$$P = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$F1 = 2 \times \frac{P \times R}{R + P} \quad (10)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (11)$$

式中:TP、TN、FP 和 FN 分别表示真阳性、真阴性、假阳性和假阴性。

2.3 实验结果

通过在不同参数下实验并评估模型性能,使用 Massachusetts Roads 数据集对该模型进行训练,并使用准确率、召回率、F1 分数、交并比和参数量评估其定量性能。

道路分割结果如图 8~12 所示,可以看出模型的分割结果与遥感图像中的真实道路基本吻合。图 8 红色标注框表示标签图中未被正确标注的道路,在检测图中可以被正确分割出,图 9 蓝色标注框表示原图中道路被树木遮挡或存在光照阴影而导致道路不清晰的现象,在检测图中可以被正确分割出,图 10 黄色标注框表示标签图中将双行道的道路标注成单行道,在检测图中可以正确判断分割为双行道道路,图 11 绿色标注框表示标签图中错误标注了

原图中非道路的部分,在检测图中可以正确判断该部分为非道路。图 12 为图 8~11 的综合情况,可以看出模型检测过程中能够正确分割道路,证明了模型具有强大的道路学习能力,并且能够准确地分割图像特征。

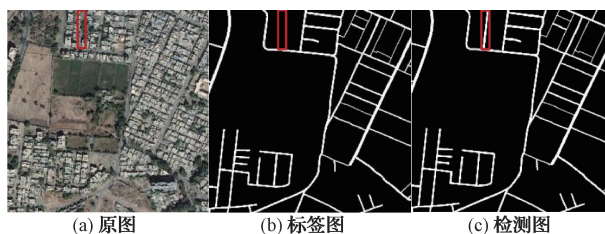


图 8 未正确标注

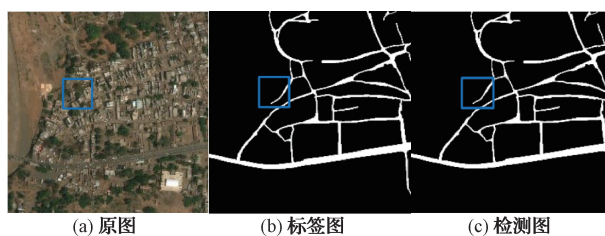


图 9 遮挡阴影严重

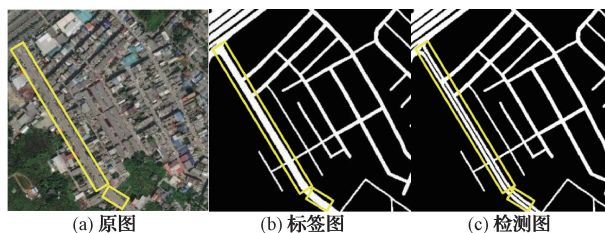


图 10 双行道检测

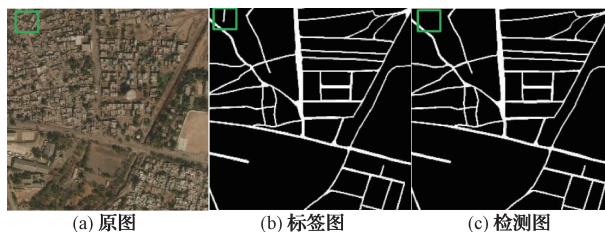


图 11 误标道路

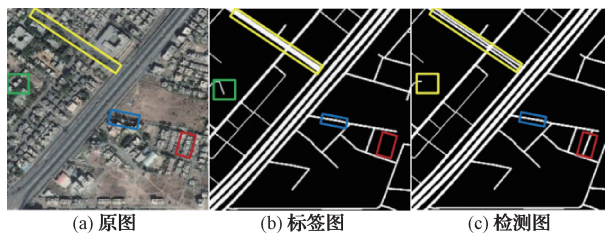


图 12 综合情况

2.4 消融实验

改进的 DeeplabV3+ 模型旨在解决遥感道路分割中边界模糊粗糙导致小目标难以识别与物体遮挡道路和背景与道路难以区分导致的错漏检的问题。如图 13 所示,其中,红色框表示道路边界模糊,蓝色框表示物体阴影遮挡,绿色框表示背景与道路像素相近难以区分,黄色框表示将背景错检为道路。

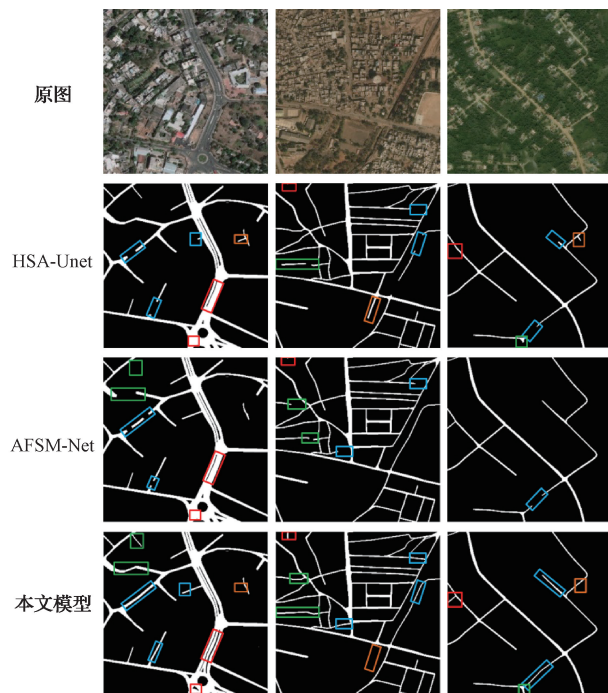


图 13 错漏检对比实验

错漏检对比实验结果表明,实验分为前两组复杂道路

和最后一组稀疏道路,在 HSA-Unet 的道路分割中,模型对较窄存在阴影遮挡的道路容易产生目标丢失,同时存在小目标丢失的问题,表明模型在细节方面有所缺失。在 AFSM-Net 的道路分割中,模型对背景与道路像素相近的地方易产生断裂不连续的问题。然而,本文提出的模型能够与真实道路相比,保持道路中间的连续性,避免出现中断或粘连情况,并能准确地分割出背景与道路相近的地方,避免错误检测的情况发生。因此,本文模型适用于提取和分割道路细节特征,具有更强的遮挡处理能力和更加准确的判别能力。

消融实验测试结果如表 2 所示,可知,本文提出的模型结构中的 3 个模块对于网络性能有着不同的影响。首先是 MobileNetv3-ECA 模块,它采用了轻量级的 MobileNetv3 模型和 ECA 注意力机制。减少参数数量和计算量的大小,以实现轻量化要求。同时能够更好地捕捉通道之间的关联信息,从而在高层特征提取方面更全面。其次是多级上采样模块,即 ASPP 模块输出的高级语义特征通过两次 2 倍上采样从特征图中恢复。这样可以更彻底地保留目标的位置和边界信息。但相应地也增加了参数量。最后是 DS-ASPP 模块,它利用深度可分离膨胀卷积来提升分割速度。但对模型的精度有一定的负面影响。综合来看,MobileNetv3-ECA 模块和多级上采样模块都能提升分割精准度,并且将两个模块融合在一起的分割精度指标明显优于基础网络。而 MobileNetv3-ECA 模块和 DS-ASPP 模块则显著提升了分割速度,同时还大幅降低了参数量。改进的 DeeplabV3+ 模型在准确率、召回率、F1 分数和交并比上分别达到了 93.71%、87.49%、90.49% 和 83.71%,参数量为 55.57×10^6 。总体来说,这些改进使得网络模型的分割性能得到了提升。

表 2 模块间的消融实验结果

序号	MobileNetv3-ECA	多级上采样	DS-ASPP	P/%	R/%	F1/%	IoU/%	参数量/ $(\times 10^6)$
1				92.45	87.00	89.64	81.06	93.42
2	✓			92.65	87.14	89.81	81.96	65.60
3		✓		93.86	87.57	90.67	82.79	95.77
4			✓	92.27	86.85	89.48	81.04	80.9
5	✓	✓		93.68	87.43	90.45	82.44	68.2
6	✓		✓	93.12	87.36	90.15	81.87	52.7
7	✓	✓	✓	93.71	87.49	90.49	83.71	55.57

2.5 对比实验

在 Massachusetts Roads 数据集上,模型 PSPNet、SegNet、U-Net、DeepLabV3+, DGRN、HSA-Unet、AFSM-Net 和本文提出的改进模型的对比实验结果如表 3 所示。所有模型都在相同的环境配置中进行训练。PSPNet 通过使用金字塔池化模块和空洞卷积等技术,能够有效地提取多尺度上下文信息。SegNet 使用编码-解码结构和

特征映射索引的方式,在遥感领域的图像分割任务中取得了显著的成果。U-Net 由一个编码器和一个解码器组成,通过跳跃连接将下采样获得的底层特征与上采样输入相结合^[26]。常被应用于医学影像分割领域,道路也可以类比为医学影像中的血管通道,因此 U-Net 可以用来分割遥感图像中的道路,主干网络采用 Xception 模型。密集-全局-残差网络(DGRN)利用残差网络和密集全局空间金

字塔池化提取特征并对上下文信息进行感知和聚合,在遥感道路提取方面有很好的效果。基于改进 U-Net 网络的 HSA-Unet 模型,构建了注意力残差学习单元来提高模型的特征提取能力。AFSM-Net 网络结合特征图切分模块和注意力机制模块来提升分割精度。

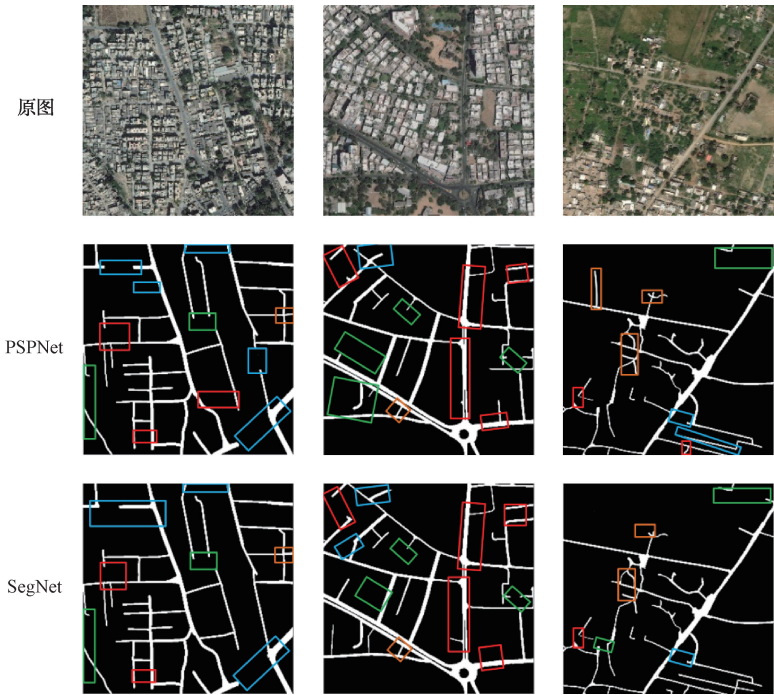
在相同的 Massachusetts Roads 数据集上,改进后的 DeepLabV3+模型相较于 PSPNet、SegNet、U-Net、DeepLabV3+、DGRN、HSA-Unet 和 AFSM-Net 模型具有更显著的效果。本文模型的准确率、召回率、F1 分数和交并比分别为 93.71%、87.49%、90.49% 和 83.71%,模型参数量为 55.57×10^6 。尽管 PSPNet 模型的主干网络使用了轻量化的 MobileNetV2,并且参数量只有 5.80×10^6 ,但其分割精度最低,因此不适用于高分辨率遥感影像的道路语义分割任务。SegNet、U-Net 和 DeepLabV3+ 模型随着参数量的增加,图像分割精度逐渐提高,但由于网络参数繁多、计算量庞大,导致分割预测的时间较长,无法满足时间上的要求。相比之下,DGRN、HSA-Unet 和 AFSM-Net 在精度和参数量上分别有所改善,但依旧无法满足分割的准确性。因此,本文提出的改进模型采用了轻量化的 MobileNetV3、ECA 注意力机制、多级上采样和 DS-ASPP 模块,既降低了参数量,又提高了网络的图像分割精度。

为了更直观地比较不同模型在遥感图像道路分割方面的效果,本文分别选择了 Massachusetts Roads 数据集中城市、郊区和乡村的道路影像进行对比,如图 14 所示。从图 14 可以看出,PSPNet 模型分割精度较低,对于高分辨率的遥感图像分割不太适用,它无法有效地提取道路细节。相比之下,SegNet 和 U-Net 模型的分割精度有所提高,但对于双车道的分割效果依旧不好。此外,在存在遮

挡和阴影的区域,模型的效果更不理想,这会影响整条道路的分割效果。由于 U-Net 模型具有更细化的检测能力,易将狭窄的背景误判为道路,导致错检的问题。DeepLabV3+模型不会产生错检的情况,但在遥感图像中存在树木或严重阴影遮挡的情况下,其分割效果较差,易产生道路局部中断的问题,同时模型存在错检的问题。DGRN 模型在树木和建筑的阴影并存,道路提取依然会产生断裂,同时存在小目标易丢失的问题。HSA-Unet 模型对背景信息与显著信息区分存在较大误差,导致对于较窄道路会出现分割错误的情况。AFSM-Net 模型当背景与道路像素相近时,会存在局部道路不连续的问题。而本文模型改进的 DeepLabV3+在应对道路分割边界模糊、物体阴影遮挡和场景像素相近等细节方面表现更好。因此,在道路分割准确性方面,本文模型的结果最为完整,避免了边缘粗糙和缺失等问题。不同模型试验结果对比如表 3 所示。

表 3 不同模型试验结果对比

模型	P/%	R/%	F1/%	IoU/%	参数量 ($\times 10^6$)
PSPNet	82.95	76.59	81.69	76.14	5.80
SegNet	89.26	82.46	85.73	78.26	46.24
U-Net	91.74	84.27	87.85	80.87	60.70
DeepLabV3+	92.45	87.00	89.64	81.06	93.42
DGRN	89.28	87.11	89.09	80.39	66.30
HSA-Unet	92.18	86.31	87.87	80.75	75.48
AFSM-Net	92.42	87.22	89.74	82.43	45.52
本文模型	93.71	87.49	90.49	83.71	55.57



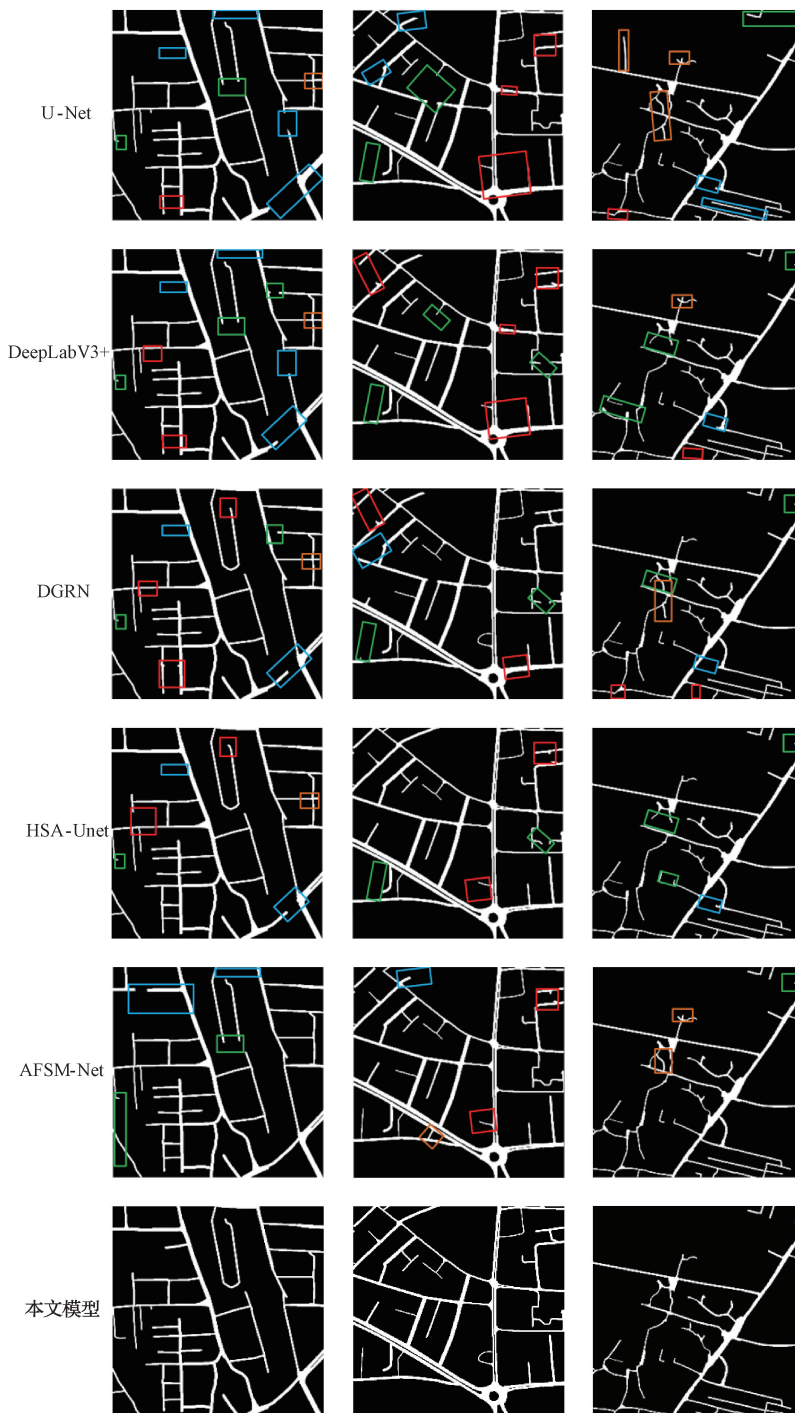


图 14 不同模型分割结果对比

3 结 论

针对高分辨率的遥感图像道路细节特征提取效果差,即边缘模糊、信息丢失和精度较低的问题本身提出基于改进 DeeplabV3+ 的遥感图像道路分割模型。该模型在主干网络引入轻量型的 MobileNetV3,并结合 ECA 注意力机制,降低了模型的参数量,使得模型能够更加关注连续

的道路特征信息,抑制背景信息干扰。在解码模块中,采用多级上采样的方式增强编码器和解码器之间的紧密连接,以更好地保留目标位置和边界信息。而在 ASPP 模块中,利用深度可分离膨胀卷积并设置不同的膨胀系数,即 DS-ASPP,显著减少了参数量,并在不同尺度上提取更丰富的特征,从而增强了网络对多尺度目标的检测和识别能力。通过在 Massachusetts Roads 遥感道路数据集上与其

他模型进行对比,实验结果表明,所提出的遥感道路分割模型的交并比可达83.71%、准确率为93.71%、召回率为87.49%、F1分数为90.49%,该模型参数量为 55.57×10^6 ,证明了模型在道路分割效果上取得了显著的提升,并且可以有效应用于遥感图像道路分割任务中。

DeepLabV3+模型通常需要大量高质量标记数据进行训练,虽然用语义分割模型在道路提取任务中能够取得良好的效果,但其缺少对道路整体拓扑结构的关注,且不易获得用于训练模型所需的大量全像素标签,结合监督学习的方式成为未来研究的重点。

参考文献

- [1] MA L, LIU Y, ZHANG X, et al. Deep learning in remote sensing applications: A meta-analysis and review[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2019, 152: 166-177.
- [2] 陈彦彤, 李雨阳, 陈伟楠, 等. 基于深度语义分割的遥感图像海面舰船检测研究[J]. 仪器仪表学报, 2020, 41(1): 233-240.
- [3] LV N, ZHANG Z, LI C, et al. A hybrid-attention semantic segmentation network for remote sensing interpretation in land-use surveillance [J]. International Journal of Machine Learning and Cybernetics, 2023, 14(2): 395-406.
- [4] SU Z B, WANG Y, XU Q, et al. LodgeNet: Improved rice lodging recognition using semantic segmentation of UAV high-resolution remote sensing images [J]. Computers and Electronics in Agriculture, 2022, 196: 106873.
- [5] WANG J X, CHEN X, JIANG W C, et al. PVNet: A novel semantic segmentation model for extracting high-quality photovoltaic panels in large-scale systems from high-resolution remote sensing imagery [J]. International Journal of Applied Earth Observation and Geoinformation, 2023, 119: 103309.
- [6] 李旭涛, 杨寒玉, 卢业飞, 等. 基于深度学习的遥感图像道路分割[J]. 山东大学学报(工学版), 2022, 52(6): 139-145.
- [7] DONG X J, ZHANG C S, FANG L, et al. A deep learning based framework for remote sensing image ground object segmentation [J]. Applied Soft Computing Journal, 2022, 130: 109695.
- [8] WANG J X, FENG Z X, JIANG Y, et al. Orientation attention network for semantic segmentation of remote sensing images [J]. Knowledge-Based Systems, 2023, DOI: 10.1016/j.knsys.2023.110415.
- [9] WURM M, STARK T, ZHU X, et al. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks [J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2019, 150: 59-69.
- [10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [11] 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.
- [12] 王凌云, 李婷宜, 李阳, 等. 基于 FEF-DeepLabV3+ 的电力金具锈蚀分割方法[J]. 电子测量与仪器学报, 2023, 37(7): 166-176.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, 2015: 234-241.
- [14] 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.
- [15] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoderdecoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39: 2481-2495.
- [16] ZHU Y T, LONG L H, WANG J J. Road segmentation from high-fidelity remote sensing images using a context information capture network [J]. Cognitive Computation, 2022, 14(2): 780-793.
- [17] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 6230-6239.
- [18] 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, 2018: 801-818.
- [19] ANILKUMAR P, VENUGOPAL P. An enhanced multi-objective-derived adaptive DeeplabV3 using G-RDA for semantic segmentation of aerial images[J]. Arabian Journal for Science and Engineering, 2023, 48: 10745-10769.
- [20] WU Q Q, LUO F, WU P H, et al. Automatic road extraction from high-resolution remote sensing images

- using a method based on densely connected spatial feature-enhanced pyramid [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 3-17.
- [21] 何哲, 陶于祥, 罗小波, 等. 基于改进 U-Net 的遥感图像道路提取[J]. 激光与光电子学进展, 2023, 60(16): 1628004.
- [22] 苏志鹏, 李景文, 姜建武, 等. 基于改进 DeeplabV3 的遥感影像语义分割方法[J]. 激光与光电子学进展, 2023, 60(6): 0628003.
- [23] 赵凌虎, 袁希平, 甘淑, 等. 改进 DeeplabV3 的高分辨率遥感影像道路提取模型[J]. 自然资源遥感, 2023, 35(1): 107-114.
- [24] ZHOU G D, CHEN W T, QIAN S G, et al. Split depth-wise separable graph-convolution network for road extraction in complex environments from high-resolution remote-sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-15.
- [25] 房玉品, 王小鹏, 李新娜. 基于自适应形态学的遥感图像道路提取[J]. 激光与光电子学进展, 2022, 59(16): 1610006.
- [26] 黄聪, 杨珺, 刘毅, 等. 基于改进 DeeplabV3+ 的遥感图像分割算法[J]. 电子测量技术, 2022, 45(21): 148-155.

作者简介

张银胜, 博士, 副教授, 主要研究方向为图像处理、人工智能。

E-mail: yorkzhang@nuist.edu.cn

单梦姣, 硕士研究生, 主要研究方向为图像处理、人工智能。

E-mail: 202212490576@nuist.edu.cn