

LRSS-Net: 轻量级遥感地物分割网络^①



唐琼霜, 何青, 戴思璇, 洪巍

(长沙理工大学 电气与信息工程学院, 长沙 410114)

通信作者: 何青, E-mail: 254739027@qq.com

摘要: 针对目前基于深度学习的高分辨率遥感图像分割模型由于参数量大、计算复杂而导致高延迟、低响应的问题, 提出了一种轻量级遥感地物分割方法, 较好的平衡了速度和精度. 该方法使用 MobileNetV2 进行特征粗提取, 通过构建空间信息嵌入分支实现不同尺度的特征细提取, 不同层次之间引入密集连接以获取密集的上下文信息. 解码端设计特征融合优化策略逐层融合不同尺度的特征增加对细粒度特征的感知, 同时以反卷积与双线性插值交替的上采样方式减少图像边缘信息丢失. 最后采用交叉熵损失与 Dice 损失结合的方式加快网络收敛速度. 为了验证所提方法的有效性, 与几种常用的语义分割方法进行了对比实验. 实验结果表明, 所提算法的分割准确率为 93.7%, *MIoU* 为 88.01%, 可以实现地物的有效分割.

关键词: 遥感图像; 轻量级; 地物分割; 密集连接; MobileNetV2; 深度学习

引用格式: 唐琼霜, 何青, 戴思璇, 洪巍. LRSS-Net: 轻量级遥感地物分割网络. 计算机系统应用, 2023, 32(5): 227-233. <http://www.c-s-a.org.cn/1003-3254/9071.html>

LRSS-Net: Lightweight Remote Sensing Segmentation Network

TANG Qiong-Shuang, HE Qing, DAI Si-Xuan, HONG Wei

(School of Electrical & Information Engineering, Changsha University of Science and Technology, Changsha 410114, China)

Abstract: The current high-resolution remote sensing image segmentation model based on deep learning has the problems of high delay and low response caused by a large number of parameters and complex calculations. Considering the problems, this study proposes a lightweight remote sensing feature segmentation method, which can better balance speed and accuracy. This method uses MobileNetV2 for rough feature extraction, constructs spatial information embedding branches to achieve fine feature extraction on different scales, and introduces dense connections between different levels to obtain dense contextual information. The decoding end designs the feature fusion optimization strategy to fuse the features of different scales layer by layer to increase the perception of fine-grained features. Meanwhile, upsampling with alternating deconvolution and bilinear interpolation is employed to reduce the image edge information loss. Finally, the cross-entropy loss is combined with the Dice loss to accelerate network convergence. Comparative experiments are carried out with several commonly used semantic segmentation methods to verify the effectiveness of the proposed method. The experimental results show that the segmentation accuracy of the proposed algorithm is 93.7%, and the *MIoU* is 88.01%, which can achieve effective segmentation of ground objects.

Key words: remote sensing image; lightweight; segmentation of ground objects; dense connection; MobileNetV2; deep learning

^① 基金项目: 国家自然科学基金 (62103063)

收稿时间: 2022-10-12; 修改时间: 2022-11-14; 采用时间: 2022-11-30; csa 在线出版时间: 2023-02-17

CNKI 网络首发时间: 2023-02-17

近年来,我国对遥感技术的研究力度加大,遥感技术飞速发展,遥感图像的可用性和普及性得到了极大的提高^[1]。遥感图像通常包含道路、建筑、水体、植被等具有丰富细节信息的地物目标,对各类地物目标信息的提取为城市规划、环境保护、城市动态监测、灾后救援、森林资源监测等领域提供了重要的支撑^[2]。遥感图像由于光谱分辨率和空间分辨率提升,不仅所包含的信息量大大增加,同时各种语义特征更加复杂,可以说高分辨率遥感图像语义分割成为图像分割领域最具挑战的难题之一。

深度学习在视觉应用中表现优异,利用深度学习的自动特征学习方法已经成为一种可行的遥感图像语义分割方法。Long等^[3]首次提出了以端到端形式进行语义分割的全卷积神经网络(fully convolutional networks, FCN),该模型通过卷积进行特征提取,并使用反卷积对最后一个特征图进行上采样使其恢复到输入图像大小。但是连续的下采样导致细节丢失并且没有考虑像素与像素之间的联系,分割效果不理想。虽然FCN不够完美,但是其全新的思路开辟了一个新的语义分割方向,对这个领域的影响是十分巨大的。此后出现的许多经典网络都是受FCN的启发设计的,例如SegNet^[4]、U-Net^[5]、PSPNet^[6]、Deeplabv3+^[7]等系列网络,这些网络在地物分割领域也得到了广泛的应用。苏健民等^[8]提出了一种基于U-Net改进的高分辨率遥感图像语义分割方法,增加卷积层的数量以实现更加精确的像素分类问题,后端采用集成学习策略优化分割结果,然而在操作上略显繁琐。张哲晗等^[9]对SegNet进行了改进,去除了部分池化减少空间信息的损失,通过加深过滤器深度增强语义分割的时效性,实现了农村遥感图像中地物的精准提取,但边缘不太平滑。朱戎等^[10]采用Canny算法提取物体边界并将其作为网络训练的监督约束,使地物分割中的边缘分割更加精细。上述方法都取得了不错的精度,然而庞大的架构和众多的参数都导致网络很难真正部署在要求低延迟、高响应率的实际应用中,许多研究者们也在研究分割的实时性。经典的轻量语义分割算法ENet^[11]使用非对称的编码-解码结构,通过限制下采样幅度极大地减少了计算量,然而感受野太小不足以观察到全局目标信息。DFANet^[12]通过轻量级网络Xception^[13]显著地降低了计算复杂度,将高层特征重复利用,增强了特征的表现能力。上述轻量级网络推理速度快、参数量少,但是精度依然

和非轻量级网络有差距,还有很大的提升空间。

针对以上问题,本文在保证参数量小、模型轻便的前提下,以提高模型分割效果为目标,提出了一种为高分辨率无人机遥感影像设计的轻量级遥感地物分割网络LRSS-Net(lightweight remote sensing segmentation network)。

1 LRSS-Net 模型

本文提出的LRSS-Net主要包括3个部分:初步特征提取模块、特征聚合模块、特征解码器模块,模型基础架构如图1所示。

为了降低常规分类网络带来的爆炸式增长的计算量,本文使用轻量级网络MobileNetV2^[14]进行初步特征提取,去除最后的 1×1 卷积层、平均池化层和全连接层,为了减少计算成本以及边缘形状等空间信息的丢失只进行4次下采样。特征聚合模块采取密集连接将初步特征提取网络的不同深度的细节与语义信息层层结合,产生阶段级的细化特征,提高模型的学习能力的同时增大了感受野,得到了丰富的上下文信息。特征解码器模块,使用双线性插值与反卷积交替进行上采样逐次还原图像的分辨率,在这个过程中与特征聚合模块对应大小的特征通过特征融合优化策略进行通道级别的特征融合。

1.1 初步特征提取模块

MobileNetV2在2018年被Google团队所提出^[14],它是专门为图像分类任务而设计,可以用较少的运算量得到较高的精度,因此本文将它用来进行初步特征的提取,大幅度减少计算量。

如图1的初步特征提取模块,由一个二维卷积和17个Bottleneck层组成。其中的Bottleneck称之为倒残差卷积块,包含3个操作,首先将输入通过一个卷积核大小为 1×1 的卷积进行升维,再经过 3×3 的深度可分离卷积(depthwise separable convolution, DWConv)提取特征,最后使用 1×1 的卷积降维。相比于普通卷积,深度可分离卷积在减少了计算成本和参数量的同时,还保持了较好的性能。它的核心思想是把一个普通的卷积分成了逐深度卷积与逐点卷积。逐深度卷积是对每个通道单独卷积,输入输出通道数不变。逐点卷积就是执行 1×1 卷积。每个卷积后面紧跟着批归一化(batch normalization, BN)以及ReLU6激活函数,使网络模型快速收敛。

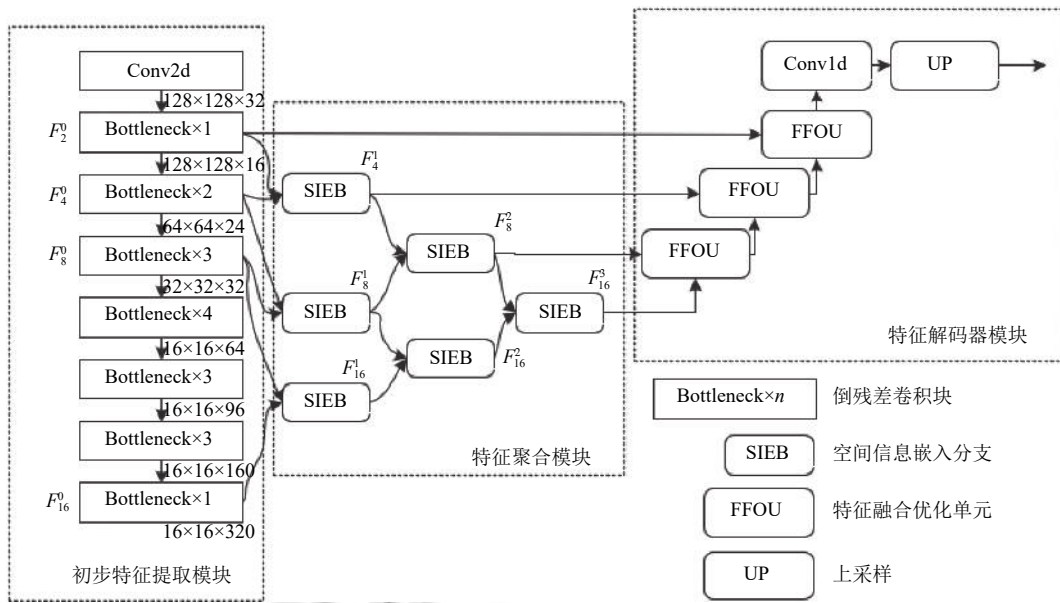


图1 LRSS-Net 模型基础架构

根据 256×256 的输入, 得到 4 个不同语义级别的初步特征图 F_2^0 、 F_4^0 、 F_8^0 和 F_{16}^0 , 其中 2、4、8、16 代表下采样的倍数, 所以其空间分辨率分别为 128、64、32、16, 通道数分别为 16、24、32、320。

1.2 聚合特征模块

低级特征具有丰富的空间细节信息, 但是缺乏语义信息, 高级特征语义信息丰富却又缺少空间细节信息, 多次下采样不可避免地会损失部分空间细节信息, 导致无法判断大规模的各种物体和精确的结构边缘。为了解决这个问题, 受文献 [15] 启发构建了空间信息嵌入分支 (spatial information embedding branch, SIEB), 将低级特征空间信息细节嵌入高级特征, SIEB 的详细设计如图 2 所示。

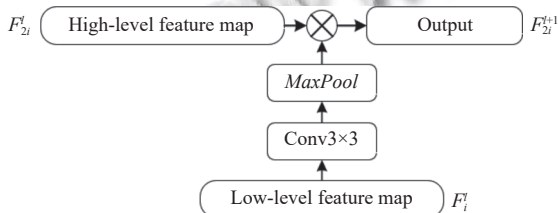


图2 空间信息嵌入分支

该单元有两个输入, 浅层特征图 F_l^i 和深层特征图 F_{2l}^l , 尺寸分别为 $C_1 \times H_1 \times W_1$ 和 $C_2 \times H_2 \times W_2$, C_1 、 C_2 表示通道数, H_1 、 H_2 表示高度, W_1 、 W_2 表示宽度, H_1 、 W_1 分别为 H_2 、 W_2 的两倍。首先 F_l^i 经过一个卷积

核大小为 3×3 的卷积进一步加强边缘和轮廓信息, 然后通过最大池化下采样去除图像中的视觉冗余信息, 所得到的输出与 F_{2l}^l 进行点乘操作, 最终输出特征图 F_{2l}^{l+1} , 尺寸为 $C_2 \times H_2 \times W_2$ 。计算公式如式 (1) 所示:

$$F_{2l}^{l+1} = F_{2l}^l \times \text{MaxPool}(\text{Conv}(F_l^i)) \quad (1)$$

其中, i 表示下采样的倍数, l 表示细化后的第几层特征, MaxPool 表示卷积核为 2×2 的最大池化, Conv 表示 3×3 卷积。

特征聚合模块如图 1 所示, 首先将初步特征提取模块得到的 4 个特征映射 F_2^0 、 F_4^0 、 F_8^0 、 F_{16}^0 以相邻的每两个特征图输入 SIEB, 分别得到 3 个新的语义级别的特征图 F_4^1 、 F_8^1 、 F_{16}^1 , 接着又将其相邻的每两个特征输入 SIEB 得到两个新的语义级别的特征图 F_8^2 、 F_{16}^2 , 然后将其输入 SIEB 得到最终的高级语义特征图 F_{16}^3 。经过层层的空间信息嵌入, 相比起缺乏空间细节信息的 F_{16}^0 , F_{16}^3 不仅拥有丰富的语义信息还恢复了大部分的空间细节信息, 这不仅有利于之后的特征融合, 而且直接有助于分割性能。

1.3 特征解码模块

解码器是一个十分重要的环节, 某些网络的解码部分是将编码器的输出直接上采样到原图大小, 这样的做法导致噪声被放大留下了更多的细节, 但是也带来了更大的负面干扰, 影响最终的分割效果, 因此本文以 2 倍为跨度来恢复空间维度。如图 1 中最右端虚线框所

示, 特征解码模块由多个特征融合优化单元 (feature fusion optimization unit, FFOU) 和一个 1×1 的卷积层组成, 是融合和优化的多次反复, 直至恢复原图大小。

U-Net 网络的特点之一是采用了跳跃连接方式, 将深层的语义信息上采样后与浅层的位置信息按深度连接在一起。本文借鉴了这种方法, 然后提出了一种优化方法, 更深层次的细化了每一层级的融合特征, 并在此基础上减少了计算量。特征融合优化单元的结构如图 3 所示。

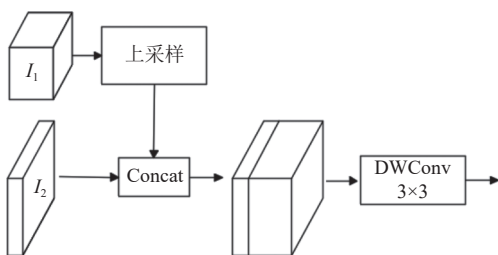


图 3 特征融合优化单元

假设尺寸大小为 $W/2 \times H/2 \times C1$ 的特征图 I_1 和 $W \times H \times C2$ 的特征图 I_2 是相邻层级所提取的特征, 首先对特征图 I_1 采取上采样操作使其与特征图 I_2 的分辨率一致, 其输出与特征图 I_2 在通道维度上进行串联融合, 然后使用 3×3 的深度可分离卷积对串联后的特征进行操作, 减少计算量的同时解耦特征深度信息与空间信息。

1.4 损失函数

假设一个合适的损失函数不仅能加快训练速度使网络快速收敛, 还能获得更好的精度。通过多次试验 (见第 2.4.2 节), 损失函数最终选择多类别交叉熵 (cross entropy loss, CE loss) 和 Dice loss 的组合。CE loss、Dice loss 和最终的损失函数分别由式 (2)–式 (4) 所示。

$$L_{CE} = - \sum_{c=1}^M y_c \log(p_c) \quad (2)$$

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N y_c p_c}{\sum_{i=1}^N (y_c^2 + p_c^2)} \quad (3)$$

$$Loss = \frac{\sum_{i=1}^N L_{CE_i}}{N} + \frac{\sum_{c=1}^M L_{Dice_c}}{M} \quad (4)$$

其中, 每个像素点对应的交叉熵损失定义为 L_{CE} , 每个类别的 Dice loss 定义为 L_{Dice} , 最终的损失定义为 $Loss$,

y_c 为某个类别的预测值, p_c 为某个类别的标签值, N 为一张图像的总像素点, M 为类别总数。

2 实验

2.1 数据源与预处理

为了验证算法的有效性, 在 2015 年中国南方某城市无人机遥感图像数据集上展开了研究, 数据来源于 CCF (China computer federation) 卫星影像的 AI 分类与识别大赛, 由 5 张带标注的高分辨率遥感图像组成, 其中尺寸最大为 7969×7939 , 最小为 4011×2470 。空间分辨率为亚米级, 光谱为可见光频段。将数据分为 5 类: 背景、植被、建筑、水体和道路。

由于该数据集的分辨率太高无法直接送入网络进行训练, 所以利用滑动窗口的思想对数据进行切割, 将图像以及标注图像切割成 256×256 大小, 使用数据增强策略增加数据量学习更多的特征, 提高网络的泛化能力, 噪声数据的增加也提高了网络的鲁棒性。本文使用的数据增强方法包括旋转、镜像、加噪、饱和度增强、亮度增强、模糊操作等。经过数据增强之后, 最终得到 15 000 张尺寸为 256×256 的高分辨率遥感数据集。实验将数据集按照 9:1 的比例进行划分, 即 13 500 张图像作为训练集, 1 500 张图像作为测试集。

2.2 实验平台及训练策略

训练基于 PyTorch 深度学习框架, GPU 是 NVIDIA TESLA P100。

在训练数据上进行 100 次迭代, 经过多次试验, 最终设置初始学习率为 0.000 1, 批尺寸为 8; 使用 Adam 优化器进行权重更新, 动量 (momentum) 设置为 0.9, 权重衰减 (weight decay) 设置为 0.000 5; 学习率调整策略采取 StepLR 有序调整方法, 每 4 个 epoch 进行一次学习率衰减, 衰减率为 0.94。在每一轮的迭代中利用验证集的损失和精度来衡量模型是否过拟合, 当验证集精度收敛时训练结束, 并且保存精度最好的模型。

2.3 评价指标

本文使用以下 3 个指标对模型进行评估。

交并比 (IoU) 表示对某一类别真实值和预测值的交集与并集的比值。如式 (5) 所示:

$$IoU = \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (5)$$

平均交并比 ($MIoU$) 表示每一类别的 IoU 先求和再取均值. 如式 (6) 所示:

$$MIoU = \frac{1}{k} \sum_{i=0}^k IoU_i \quad (6)$$

像素准确率 (PA) 表示正确预测类别的像素数与总像素的比率. 如式 (7) 所示:

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \quad (7)$$

其中, k 表示待预测的种类数, p_{ii} 表示将第 i 类预测成第 i 类的像素数量, p_{ij} 表示将第 i 类预测成第 j 类的像素数量, p_{ji} 表示将第 j 类预测成第 i 类的像素数量.

2.4 实验结果

2.4.1 特征融合方式的选择

各种网络常用 add 和 $Concat$ 来整合特征图信息, add 是特征图直接相加, 维度不会发生变化, 只是每一特征下的信息量增加了. 而 $Concat$ 是通道数合并, 特征图的特征数增加了, 每一特征下的信息没有增加. 为探究不同融合方式的影响做了以下实验, 结果如表 1 所示.

表 1 不同融合方式的影响

融合方式	$MIoU$ (%)	PA (%)	GFLOPs	参数量 (M)
add	86.24	92.63	1.48	2.52
$Concat$	88.01	93.7	14.01	3.48

注: 每列粗体为最优

从表 1 可以看出, $Concat$ 的分割精度优于 add , 而使用了 add 的模型比 $Concat$ 更轻便, 各有各的优点. 相比于 add , $Concat$ 保存了更多特征信息, 为特征重用提供了可能, 所以本文最终选择了分割精度更高的 $Concat$ 融合方式.

2.4.2 损失函数的选择

为了验证不同的损失函数对分割效果的影响, 做了以下实验, 不同损失函数对比结果如表 2 所示.

表 2 不同损失函数对比结果 (%)

损失函数	$MIoU$	PA
Focal loss	75.50	86.38
CE loss	85.17	92.31
Focal loss + Dice loss	84.70	91.58
Dice loss + CE loss	88.01	93.7

注: 每列粗体为最优

在表 2 中, 可以很明显的看到, 当损失函数为 Dice loss 和 CE loss 的组合时, $MIoU$ 和 PA 优于其他损失函数, 这是因为二者的结合能够兼顾像素级的分类精度和图像前景的分割效果, 使模型可以更加稳定的训练, 从而获得更加精确的预测结果.

2.4.3 结构优越性对比实验

基于 CNN 的语义分割的优点之一便是骨干网络可以随意更换, 不同的骨干网络提取特征的效果不一样. 为了定量、定性的评估, 将本文方法与 U-Net、PSPNet、Deeplabv3+ 进行对比, 所有模型使用同一骨干网络, 即 ImageNet 上预训练的 MobileNetV2. 表 3 为各模型在相同的骨干网络下的性能对比实验. 在表 3 中, LRSS-Net 的 IoU 、 $MIoU$ 、 PA 均高于以 MobileNetV2 为骨干网络的其他网络, 这表明 LRSS-Net 结构的有效性, 在遥感图像地物分割中具有很好的应用能力.

表 3 结构优越性对比实验结果 (%)

方法	IoU					$MIoU$	PA
	背景	植被	建筑	水体	道路		
U-Net	77.89	93.01	80.86	90.97	77.77	84.10	91.38
PSPNet	78.83	93.63	81.30	91.58	78.82	84.81	91.77
Deeplabv3+	78.65	93.19	81.54	90.77	78.67	84.56	91.69
Ours	83.49	95.03	85.41	93.04	83.05	88.01	93.7

2.4.4 与各种经典网络对比实验

为进一步探究 LRSS-Net 与各种原始经典网络性能对比, 做了以下实验, 使用 IoU 、 $MIoU$ 、 PA 、GFLOPs、参数量进行评估, GFLOPs 表示浮点运算次数, 用来衡量模型复杂度, 各模型评价指标结果如表 4 所示.

从表 4 中可以看到, LRSS-Net 在 GFLOPs 和参数量仅为 14.01 和 3.48M 的情况下, 在测试集上取得了 88.01% 的 $MIoU$ 和 93.7% 的 PA , 相比其他几个网络它表现出了更好的效果. 这是因为 MobileNetV2 网络在参数量、计算量方面做出了巨大贡献, 特征聚合模块能够帮助网络获得更多的上下文信息, 在保证计算量和参数量的前提下改善 MobileNetV2 网络因结构简单带来的性能损失. 虽然本文算法在 GFLOPs 和参数量上和 ENet 比较分别低了 13.48 和 3.14M, 但是在 $MIoU$ 和 PA 方面比 ENet 高了 14.53% 和 9.57%, 显然 LRSS-Net 在速度和精度之间做到了最大化平衡.

将实验结果进行可视化, 通过图 4 可以更直观地感受到不同网络的分割效果. LRSS-Net 基本能够准确识别出图像中各类别并且准确进行分割, 由于采用了高效的空信息嵌入分支, 减少了空信息的损失, 因

此各个类别的边缘轮廓清晰流畅, DFANet、SegNet、ENet 出现严重错误分割, 轮廓边缘线条粗糙, 整体分割效果较差. PSPNet 和 Deeplabv3+对道路的预测出现部分断续情况. 所有网络对于道路的预测效果在所有类

别中是最差的, 这是因为道路是细长的跨度很大, 而感受野是正方形的, 正方形的感受野无法很好地契合目标形状, 不可避免地会引入很多无用的特征, 因此很容易将道路错分割为其他目标.

表4 各模型评价指标

方法	骨干网络	IoU (%)					MIoU (%)	PA (%)	GFLOPs	参数量 (M)
		背景	植被	建筑	水体	道路				
DFANet	Xception	56.88	81.75	61.86	81.60	57.19	67.85	80.22	0.53	2.35
SegNet	VGG16	61.42	82.82	66.48	82.80	63.09	71.32	82.72	40.16	29.43
ENet	—	62.36	85.18	69.18	85.27	65.42	73.48	84.13	0.53	0.34
PSPNet	ResNet50	80.93	94.21	83.27	92.05	80.85	86.26	92.68	15.37	49.01
Deeplabv3+	Xception	82.07	94.41	84.14	91.96	81.98	86.95	92.86	20.69	54.52
U-Net	VGG16	81.83	94.89	84.22	93.39	81.65	87.2	93.16	56.30	24.89
Ours	MobileNetV2	83.49	95.03	85.41	93.04	83.05	88.01	93.7	14.01	3.48

注: 每列粗体为最优, —表示无

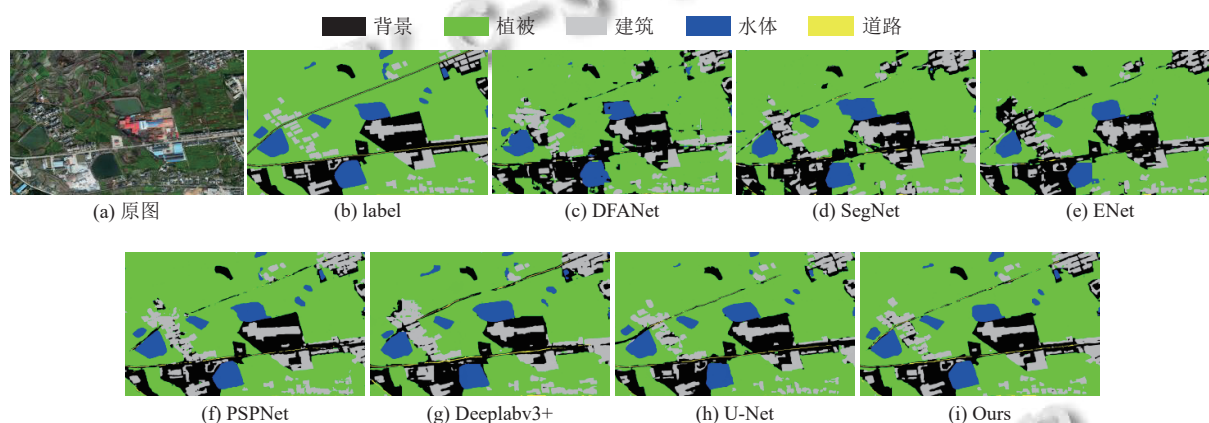


图4 遥感图像语义分割结果

3 结束语

遥感图像分割是遥感图像分析中的一个重要步骤, 为了提高分类的效率以及准确率, 本文提出了一种轻量级遥感影像地物分割网络. 利用 MobileNetV2 轻便的特点, 极大地减少了参数量, 又引入 SIEB 增强了高级语义特征中的空间信息, 不同尺度特征密集连接得到丰富的上下文信息, 解码端与编码部分输出的特征逐层连接增强特征信息的传递性和利用率, 采取双线性插值、反卷积交替的上采样方式, 显著地提升了分割效果. 从实验结果可以看出, 道路、建筑这类目标分割困难, 分割效果相比于其他类别还有很大的上升空间. 后续将会对道路、建筑物进行针对性的研究, 为地物分割提供更有效的手段.

参考文献

1 巫兆聪, 胡忠文, 张谦, 等. 结合光谱、纹理与形状结构信

息的遥感影像分割方法. 测绘学报, 2013, 42(1): 44–50.

2 徐誉远, 胡爽, 王本洋. 无人机遥感在我国森林资源监测中的应用动态. 林业与环境科学, 2017, 33(1): 97–101. [doi: 10.3969/j.issn.1006-4427.2017.01.018]

3 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640–651. [doi: 10.1109/tpami.2016.2572683]

4 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481–2495. [doi: 10.1109/tpami.2016.2644615]

5 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015. 234–241. [doi: 10.1007/978-3-319-24574-4_

- 28]
- 6 Zhao HS, Shi JP, Qi XJ, *et al.* Pyramid scene parsing network. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2881–2890. [doi: [10.1109/cvpr.2017.660](https://doi.org/10.1109/cvpr.2017.660)]
- 7 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 833–851. [doi: [10.1007/978-3-030-01234-2_49](https://doi.org/10.1007/978-3-030-01234-2_49)]
- 8 苏健民, 杨岚心, 景维鹏. 基于 U-Net 的高分辨率遥感图像语义分割方法. 计算机工程与应用, 2019, 55(7): 207–213. [doi: [10.3778/j.issn.1002-8331.1806-0024](https://doi.org/10.3778/j.issn.1002-8331.1806-0024)]
- 9 张哲晗, 方薇, 杜丽丽, 等. 基于编码-解码卷积神经网络的遥感图像语义分割. 光学学报, 2020, 40(3): 0310001.
- 10 朱戎, 叶宽, 杨博, 等. 基于改进 DeeplabV3+ 的地物分类方法研究. 计算机科学, 2021, 48(S2): 382–385. [doi: [10.11896/jsjcx.201100184](https://doi.org/10.11896/jsjcx.201100184)]
- 11 Paszke A, Chaurasia A, Kim S, *et al.* ENet: A deep neural network architecture for real-time semantic segmentation. arXiv:1606.02147, 2016.
- 12 Li HC, Xiong PF, Fan HQ, *et al.* DFANet: Deep feature aggregation for real-time semantic segmentation. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 9514–9523.
- 13 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 1800–1807. [doi: [10.1109/cvpr.2017.195](https://doi.org/10.1109/cvpr.2017.195)]
- 14 Sandler M, Howard A, Zhu ML, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4510–4520. [doi: [10.1109/cvpr.2018.00474](https://doi.org/10.1109/cvpr.2018.00474)]
- 15 Zhang ZL, Zhang XY, Peng C, *et al.* ExFuse: Enhancing feature fusion for semantic segmentation. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 273–288. [doi: [10.1007/978-3-030-01249-6_17](https://doi.org/10.1007/978-3-030-01249-6_17)]

(校对责编: 牛欣悦)