

DOI:10.3969/j.issn.1003-5060.2025.01.005

基于 EfficientNet 深度学习网络的 遥感影像地物分割方法

姜克儒¹, 刘军², 谢枫³, 盛金马¹, 刘耀中³, 许水清⁴

(1. 国网安徽省电力有限公司经济技术研究院, 安徽 合肥 230022; 2. 国网安徽省电力有限公司建设分公司, 安徽 合肥 230071; 3. 中国能源建设集团安徽省电力设计院有限公司, 安徽 合肥 230093; 4. 合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009)

摘要:在电网工程建设中,自动提取遥感影像中包含的地物信息对实现电力规划自动化具有重要意义。文章提出一种基于深度学习的遥感影像地物分割方法,为解决遥感影像目标丰富、尺度多样的问题,以 UNet 网络为基本架构,选择 EfficientNet 网络作为主干网络,并加入特征融合;为解决遥感影像类别不均衡和泛化能力问题,采用联合损失函数和先进的数据增强方法。实验结果表明,所提方法能够有效提高遥感图像地物分割精度,对小目标和大目标均有较好的分割效果。

关键词:遥感图像;地物分割;深度学习;特征融合;损失函数

中图分类号:TP751.1 **文献标志码:**A **文章编号:**1003-5060(2025)01-0032-05

Ground object segmentation method for remote sensing images based on EfficientNet deep learning network

JIANG Keru¹, LIU Jun², XIE Feng³, SHENG Jinma¹, LIU Yaozhong³, XU Shuiqing⁴

(1. Economic and Technical Research Institute, State Grid Anhui Electric Power Co., Ltd., Hefei 230022, China; 2. Construction Company, State Grid Anhui Electric Power Co., Ltd., Hefei 230071, China; 3. China Energy Engineering Group Anhui Electric Power Design Institute Co., Ltd., Hefei 230093, China; 4. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

Abstract:In power grid project construction, automatic extraction of ground object information from remote sensing images is very important for realizing automation of power planning. In this paper, a ground object segmentation method for remote sensing images based on deep learning is proposed. In view of the multiple targets and diversified scales of remote sensing images, the UNet is used as the basic framework, EfficientNet is selected as the backbone network, and feature fusion is added. In order to solve the problem of classification unbalance and generalization ability of remote sensing images, combined loss function and advanced data enhancement method are adopted. Experiments show that the proposed method can effectively improve the ground object segmentation accuracy of remote sensing images, and has good segmentation effect for both small and large targets.

Key words: remote sensing image; ground object segmentation; deep learning; feature fusion; loss function

随着高分辨率遥感卫星的出现,通过复杂的遥感影像获取有用的地理信息、超高分辨率的影

像对国土资源利用、城乡规划、自然资源保护等方面有重要的意义^[1]。以往提取地物信息的方法依

收稿日期:2021-11-25;修回日期:2022-05-19

基金项目:安徽省能源互联网联合重点基金资助项目(2008085UD03)

作者简介:姜克儒(1993—),男,河南固始人,国网安徽省电力有限公司经济技术研究院工程师;

许水清(1991—),男,安徽太和人,博士,合肥工业大学副教授,硕士生导师,通信作者,E-mail:xsqanhu@hfut.edu.cn.

靠人工手动标绘,目视解译虽然精度较高,但是需要耗费大量的人力、物力和时间,因此遥感影像中丰富的地物信息对电力设计规划也十分重要。

早期的图像分割方法大多基于图形学算法。文献[2]提出一种基于直方图区域生长的遥感图像阈值分割算法,通过合并区域减少阈值数目,降低图像分割的时间复杂度;文献[3]利用多结构元素形态学重构边缘检测,并结合区域生长实现遥感图像的水体分割;文献[4]提出改进的分水岭分割方法,结合光谱信息分量和形态学原理提取遥感图像水体。这些传统的分割方法,只能对遥感图像实现前景和背景的分割^[5],无法充分利用到遥感图像中包含的丰富地理信息。

随着深度学习技术的不断发展,利用深度学习网络能够由浅到深地学习图像中的低级细节特征和高级语义信息,并且在速度和精度上有极大提升。现有的深度学习语义分割模型有全卷积神经网络(fully convolutional network, FCN)、UNet、PSPNet 和 DeepLab 等。FCN^[6]将图像分类网络中的全连接层改成上采样,将学习到的特征映射到每个像素上,但其精度较低。UNet 网络^[7]是对称的编解码架构,经过 4 次下采样和 4 次上采样形成 U 型结构,其主要贡献是引入跃层连接,在有更多语义信息的高层特征中补充低层特征中的细节信息;UNet 网络的参数量少、操作性强,但不利于特征较多的目标分割。PSPNet 模型^[8]引入金字塔池化模块,该模块通过不同尺度的池化操作来融合多尺度特征,但其上采样步幅较大,可能影响小目标的分割。DeepLabV3+ 模型^[9]引入空洞空间金字塔池化模块,通过多分支的并行结构和不同大小的空洞卷积率进行多尺度融合,并且使用 Xception 作为主干网络,精度较高,但同时计算量也很大。

遥感影像的分辨率高,包含目标种类多、尺度多样^[10],并且存在类别不均衡的问题,为了更好地分割出遥感图像中的地物,本文使用操作性强的 UNet 为网络架构,以精度高、速度快的 EfficientNet 作为主干网络,针对类别不均衡问题和泛化问题,使用联合损失函数和 Mixup 等先进的数据增强方法。

1 网络设计方法

1.1 网络结构

本文提出基于深度学习的遥感影像地物分割方法,其思路是基于 UNet 网络架构,以 Effi-

cientNet-B3 作为主干网络,在上采样阶段的跃层连接中融入深层特征,用于处理遥感影像的多尺度问题,同时使用联合损失函数缓解遥感图像类别不均衡问题。

EfficientNet 网络是近年来提出的一种新型卷积网络,比以往的模型具有更快的训练速度和更少的参数数量。在卷积神经网络中,有的是通过堆叠更多的层结构即增加网络的深度来提高模型性能^[11],有的是通过增加网络的通道数来提高网络性能^[12],有的会改变送入网络的图像分辨率来提高网络性能^[13]。但若过度依赖其中某一种方法,则会在提升网络性能的同时带来网络训练困难、计算量大等困难。EfficientNet 通过使用混合缩放方法统一缩放网络宽度、网络深度以及输入网络的分辨率来提升网络的性能。

EfficientNet 网络主要由 MBConv 模块堆叠而成,网络结构见表 1 所列。MBConv 使用倒残差结构(inverted residual),采用 Swish 激活函数。Swish 激活函数类似于 ReLU 和 Sigmoid 的结合^[14],同时具备这 2 种激活函数的一些优点。该模块还加入了通道注意力机制 SE(squeeze-and-excitation),且在最后加入 Dropout 层用于生成随机深度。模块的网络结构如图 1 所示。

表 1 EfficientNet-B3 网络结构

步骤	卷积操作	输入尺寸	通道数	层数
1	Conv3 * 3	300 × 300	40	1
2	MBConv1, $k * 3$	150 × 150	16	2
3	MBConv6, $k3 * 3$	150 × 150	32	3
4	MBConv6, $k5 * 5$	75 × 75	48	3
5	MBConv6, $k5 * 5$	38 × 38	96	5
6	MBConv6, $k5 * 5$	19 × 19	136	5
7	MBConv6, $k5 * 5$	19 × 19	232	6
8	MBConv6, $k5 * 5$	10 × 10	384	2

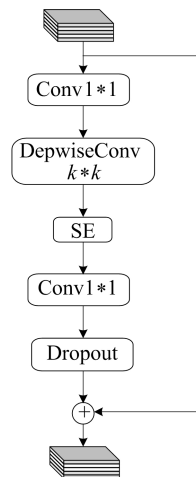


图 1 MBConv 模块

UNet 网络采用完全对称的编解码结构。其主要特点是在解码网络中采用跃层连接,将解码路径中经过上采样得到的不同尺寸的特征图与对应分辨率大小的编码部分的有效特征进行特征融合,通过浅层特征包含的细节信息^[15],对有更多语义信息的深层特征进行补充。原 UNet 网络中的跃层连接的特征比较单一,本文在 Decode-2、Decode-3 跃层连接特征前,将编码网络的浅层特征和上一个 Decode 层得到的特征图进行通道堆叠(concatenate),如图 2 所示,在保证原始浅层细节特征的同时丰富了语义信息,再和对应的解码网络中的特征进行融合。

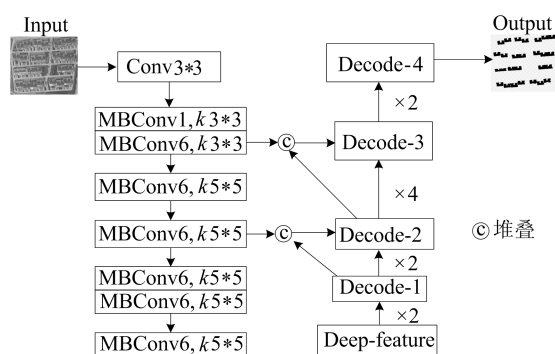


图 2 本文方法网络结构图

1.2 损失函数

损失函数是深度学习网络的重要组成部分^[16]。在分割任务中,最常用的是交叉熵损失函数(cross entropy loss, CE Loss)。该函数先单独对每个像素的类预测进行评估,再对所有像素求平均值,其表达式为:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \left[y_{i1} \ln(\hat{y}_{i1}) + y_{i2} \ln(\hat{y}_{i2}) + \dots + y_{im} \ln(\hat{y}_{im}) \right] = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^m [y_{ij} \ln(\hat{y}_{ij})] \quad (1)$$

其中: N 为像素个数; m 为类别数; $y_i = (y_{i1}, y_{i2}, \dots, y_{im})$ 为真实类别; $\hat{y}_i = (\hat{y}_{i1}, \hat{y}_{i2}, \dots, \hat{y}_{im})$ 为预测类别。

CE Loss 用于网络训练时,每个像素被平等地学习。在类别不均衡的情况下,由于训练会被占据较多像素的类主导,从而对小目标的分割不利。在对遥感图像数据分析时,发现其各类别分布非常不均衡,因此需要选择合适的损失函数。

D_{ice} 系数差异函数(Dice Loss)是一种区域相关的损失函数。Dice 系数用于衡量 2 个样本的重叠程度,其表达式为:

$$D_{ice} = \frac{2|A \cap B|}{|A| + |B|} \quad (2)$$

其中, A, B 分别表示真实图像和预测图像。

Dice 系数越大表示效果越好,因此使 Dice Loss 的值取 $1 - D_{ice}$ 。具体定义如下:

$$L_{Dice} = 1 - \frac{1}{m} \sum_{j=1}^m \frac{2 \sum_{i=1}^N y_{ij} \hat{y}_{ij}}{\sum_{i=1}^N y_{ij} + \sum_{i=1}^N \hat{y}_{ij}} \quad (3)$$

其中: N 为像素个数; m 为类别数; $y_i = (y_{i1}, y_{i2}, \dots, y_{im})$ 为真实类别; $\hat{y}_i = (\hat{y}_{i1}, \hat{y}_{i2}, \dots, \hat{y}_{im})$ 为预测类别。

Dice Loss 与 CE Loss 不同,某像素点的损失不仅与该点的标签值和预测值有关,还与其他点的标签和预测值相关,因此 Dice Loss 在一定程度上可以缓解样本不平衡,然而直接使用 Dice Loss 会降低网络训练的稳定性。实验选择 CE 与 Dice Loss 联合损失函数,最终损失函数定义为:

$$L = L_{CE} + L_{Dice} = 1 - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^m [y_{ij} \ln(\hat{y}_{ij})] - \frac{1}{m} \sum_{j=1}^m \left[\frac{2 \sum_{i=1}^N y_{ij} \hat{y}_{ij}}{\sum_{i=1}^N y_{ij} + \sum_{i=1}^N \hat{y}_{ij}} \right] \quad (4)$$

2 实验数据

2.1 数据说明

本文使用的实验数据来源于安徽省的高分辨率遥感影像,将其切割成 111 张 $4\,000 \times 4\,000$ 像素点的图像,使用 QGIS 软件制作矢量语义分割标签,原始影像和标签图如图 3 所示。

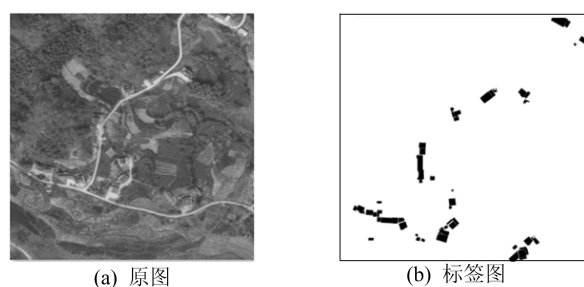


图 3 遥感数据示例

通过滑动窗口裁剪方法将遥感图像裁剪成 300×300 像素点的深度学习样本,获得 18 759 张样本数据。将制作好的数据随机划分为训练集和验证集,2 个数据集数量分别为 14 063、4 687 张。

在地物要素的选择上,本文主要依据输电线路设计规范以及电力工程规划选站需要考虑的因素(自然条件、基础设施、社会经济条件和环境制约条件),重点选择了道路、树木、湖水和建筑作为

提取对象。

2.2 数据分析

制作好数据集后,需要了解数据集中的类别分布情况。根据标签图中标注的目标区域像素点,分析地物要素类别像素数目情况,如图 4 所示。从图 4 可以看出,道路、树木、湖水和房屋的分布很不均衡,树木类像素占比较多,而建筑和道路类占比较少,若要提高遥感图像的分割精度,则需解决遥感图像类别分布不均衡的问题。

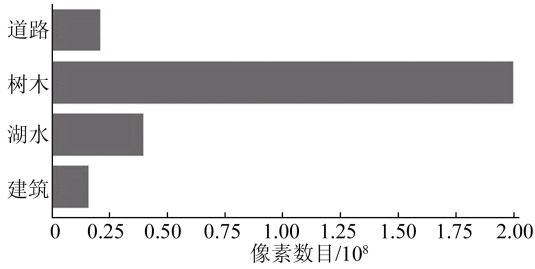


图 4 地物要素类别数目

2.3 数据增强

为了增强模型的泛化能力,需要对训练数据进行数据增强。常规数据增强策略有随机水平、垂直、对角翻转^[17],可以使得模型旋转不变形;通过图片模糊化、添加噪声可避免模型学习到无用噪声,抑制过拟合;通过缩放、错切可以保证尺度不变性。本文除了使用随机旋转 90°、水平翻转和高斯噪声等常规数据增强策略,还通过进行插值来混合构建新样本,从而使得模型在处理样本和样本之间的区域时表现为线性,使用 Mixup 数据增强方法构建虚拟样本。Mixup 是将随机的 2 个样本及其标签按比例混合得到新样本,从而减少在预测数据集以外数据的不适应性,提高模型泛化能力。此外为了模拟变域,针对验证数据集还进行了随机 1%、2% 的线性拉伸。

3 实验

3.1 实验方案

本实验硬件配置为处理器 Intel i9-10900x,显卡 GeForce RTX 3090;训练环境为 Ubuntu 18.04, Tensorflow-gpu 2.4, Python 3.7。

在基于深度学习方法的遥感影像地物分割实验过程中,将批处理大小设为 16,初始学习率设为 0.001,按指数衰减调整学习率,学习动量设为 0.9,权重衰减率设为 0.0005,总 Epoches 数设为 100,最大迭代次数为 30 000。

实验评价指标选择的平均交并比(mean in-

tersection over union, MIoU)是对每一类的交并比(intersection over union, IoU)求和再平均得到的。交并比是真实值与预测值两个集合的交集和并集的比值,该比值越大,表明预测结果越接近真实值。MIoU 的计算公式为:

$$M_{IoU} = \frac{1}{m+1} \sum_{i=0}^m \frac{y_{ii}}{\sum_{j=0}^m y_{ji} + \sum_{j=0}^m y_{ij} - y_{ii}} \quad (5)$$

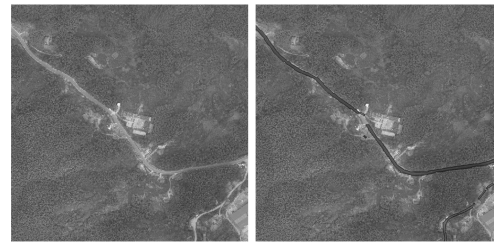
其中: $m+1$ 为图像分割的类别总数; j 为真实类别数; i 为预测类别数。

3.2 实验结果

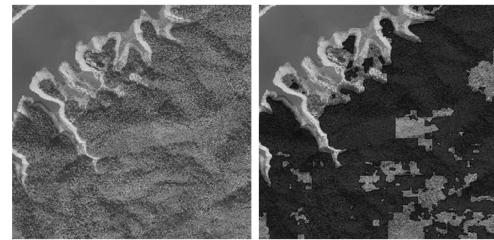
建筑、道路、湖水和树木的预测分割效果如图 5 所示。



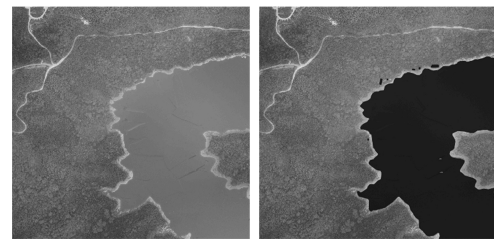
(a) 建筑分割图



(b) 道路分割图



(c) 树木分割图



(d) 湖水分割图

图 5 本文方法的预测分割效果图

此外,本文还进行了 3 组对比实验,采用 UNet 网络、CE Loss 和常规数据增强方法,通过对比实验进一步评估本文方法。不同方法得到的

预测图像 MIoU 见表 2 所列。

表 2 不同方法得到的预测图像 MIoU %

方法	本文方法	UNet	CE Loss	常规数据增强
MIoU	81.1	65.2	76.1	80.1

由表 2 可知,本文方法的 MIoU 可达 81.1%。本文方法比 UNet 网络提升了 15.9%,使用联合损失函数比只使用 CE Loss 提升了 5.0%的准确率,说明联合损失函数在一定程度上可以缓解类别不均衡的问题,而使用 Mixup 等数据增强方法可提升 1.0%的准确率,表明本文方法是有效的。

地物各类别 IoU 对比情况见表 3 所列。由表 3 可知,本文方法对小物体的识别效果也较好,其中房屋的准确率高达 90.9%,道路的准确率为 78.2%,而树木虽由于高度的影响准确率为 70.1%,但也有一定程度的提升。因此本文所提方法在地物分割的准确率和完整性上均表现良好。

表 3 地物各类别 IoU 对比 %

类别	本文方法	UNet	CE Loss	常规数据增强
建筑	90.9	72.2	83.3	89.3
道路	78.2	63.6	79.8	77.6
湖水	85.4	67.3	72.3	84.2
树木	70.1	57.5	68.8	69.5

4 结 论

为推动电力设计规划的自动化,充分高效地利用遥感影像的地物信息,本文提出了一种基于深度学习的遥感影像地物分割方法。选用 EfficientNet 网络作为主干网络,提高特征提取能力,在编码网络的跃层连接中加入特征融合,并使用联合损失函数缓解类别不均衡的问题,使用数据增强方法提高模型泛化能力。实验结果表明,使用本文方法分割遥感图像有较高的准确率。进一步研究将解决遥感影像中同物异谱的问题,从而减小地物分割的误识率。

[参 考 文 献]

[1] QIU C, MOU L, SCHMITT M, et al. Local climate zone-based urban land cover classification from multi-seasonal Sentinel-2 images with a recurrent residual network[J]. ISPRS Journal of Photogrammetry and Remote Sensing,

2019,154:151-162.

- [2] 刘思言,李玲,特日根,等.基于直方图区域生长的遥感图像阈值分割算法[J].测绘通报,2021(2):25-29.
- [3] 冯朝晖,李芹,韩留生,等.改进分水岭方法下的遥感图像水体提取[J].测绘通报,2019(6):11-15.
- [4] 王小鹏,文昊天,王伟,等.形态学边缘检测和区域生长相结合的遥感图像水体分割[J].测绘科学技术学报,2019,36(2):149-154,160.
- [5] 方茂青.基于 Deeplabv3+ 的遥感图像城市道路分割算法研究[D].武汉:华中科技大学,2019.
- [6] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2015:3431-3440.
- [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. [S. l.]: s. n., 2015: 234-241.
- [8] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2017: 2881-2890.
- [9] 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. [S. l.]: s. n., 2018:801-818.
- [10] 韩潇冰.高分辨率遥感影像“像素—目标—场景”的深度理解方法研究[D].武汉:武汉大学,2018.
- [11] 刘然,杨鼎鼎,王天真.基于 MobileNet 的海流透平机叶片附着物图像诊断[J].控制工程,2022,29(10):1810-1815.
- [12] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2015:1-9.
- [13] 冯秋晨,彭冬亮,谷雨. SAR 变体目标识别的卷积神经网络法[J].中国图象图形学报,2019,24(2):258-268.
- [14] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[EB/OL]. (2017-10-16)[2024-7-30]. <http://arxiv.org/pdf/1710.05941>.
- [15] 张斌,张晓静,赵丙辰,等.一种改进的 DenseASPP 网络在遥感图像分割中的应用[J].计算机应用与软件,2021,38(7):46-52.
- [16] 龙法宁,朱晓妹,甘井中.基于卷积神经网络的臂丛神经超声图像分割方法[J].合肥工业大学学报(自然科学版),2018,41(9):1191-1195,1296.
- [17] 王慧勇,殷明.基于 EfficientDet 的高效目标检测算法研究[J].合肥工业大学学报(自然科学版),2021,44(7):900-908.

(责任编辑 李 凯)