

一种基于编解码器的建筑工地无人机绿网识别算法

李进强¹ 吕俊杰¹ 叶延磊¹ 刘轩朋¹ 孙国梁¹ 姜楠¹
LI Jinqiang LÜ Junjie YE Yanlei LIU Xuanpeng SUN Guoliang JIANG Nan

摘要

针对建筑工地绿网识别精度差的问题,文章设计了一种基于编解码器结构的建筑工地绿网识别算法。首先,为获取高级语义信息,在编码器中设计了四个空洞卷积和一个池化网络,将五个特征维度进行融合获取编码信息;然后,在解码器中,将低级语义特征与编码特征进行融合,获取充分语义表达;最后,利用插值算法得到预测结果。在不同场景下,分别与R-CNN、SDD、FCN算法进行了对比实验,PA和MIoU分别为0.916、0.775。结果表明,基于编解码器结构的绿网识别检测算法能较好地对城市建筑工地绿网进行分割和检测,有利于节约人力和物力资源,具有实用价值。

关键词

建筑绿网识别; 编解码器; 深度学习; 神经网络

doi: 10.3969/j.issn.1672-9528.2026.04.013

0 引言

随着我国数字化和智能化进程的发展,智慧城市建设步伐也不断加快,然而在智慧城市建设中,建筑施工扬尘、拆迁区域地表裸露等现象普遍发生,这对周围居民生活环境造成了不好的影响,对大气环境也产生了一定的污染。国家虽然采取了一定的措施来解决这一问题,比如在扬尘、拆迁区域、裸露地表覆盖绿网,这样可以有效防止扬沙带来的环境污染,同时对空气质量和城市容貌景观也有很大提升,但是很多施工单位为了节省资金或监管不力,裸露地表绿网铺设往往形同虚设,如何采取有效手段对城市防尘绿网进行动态检测具有重要的现实意义。

针对上述存在的现实问题,构建了一种利用卷积模型对无人机拍摄影像进行绿网识别的算法模型,此模型将无人机拍摄、图像图形学和卷积网络提取技术进行结合,对绿网模型进行不断迭代和优化,模型精度已达0.916,能够满足对城市防尘绿网进行动态检测的要求,可以实现对绿网数量和分布情况进行快速获取,为城市规划、生态环境建设和城市精细化管理提供数据和技术支撑。

1 方法

1.1 数据获取

绿网识别首先要获取绿网数据,高质量的数据对模型识别结果至关重要,直接影响算法的精度。与东营市执法局进行协商,选用M30大疆无人机及低空平台进行影像数据获取。

M30无人机与大疆司空平台进行配合,可快速开展实时云端建图、点线面实时同步、飞行信息同步、直播查看、多机组织状态查看、航线任务实时同步协作,实现空地协同一体化、作业管理高效化,数据获取流程如图1所示。

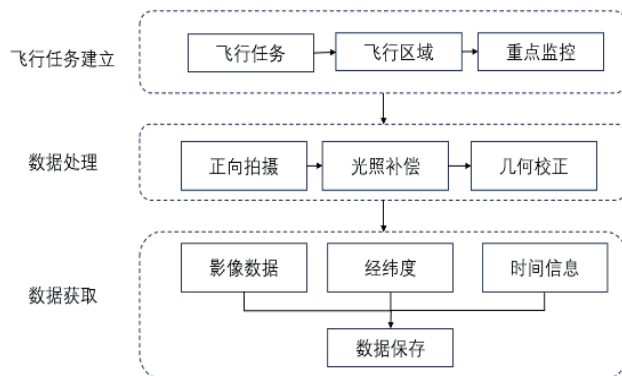


图1 数据获取流程图

具体步骤为:大疆司空平台划定飞行区域,并划定重点监控区域,无人机由地面站起飞,实时获取飞行影像数据,当无人机到达指定飞行区域时,进行影像正向拍摄,并将影像数据、经纬度、时间信息实时返回监测站点并保存记录。其中正射、几何校正环节由大疆无人机平台自主完成。绿网标签确定是由人工完成的,选择上述重点监控区域的集中连片覆盖影像,对绿网影像进行标签标定,制作了一个适用于绿网识别的图像数据库,数据集包含6 000个训练样本、1 000个测试样本,图像分辨率为1 920 px×2 080 px,数据集中包含建筑工地和施工道路两种场景。如表1所示。

1. 北京正元数科信息技术有限公司 北京 101300

表 1 数据集构成

场景	分辨率	训练样本	测试样本	总样本
建筑工地	1 920 px×2 080 px	2 500	500	3 000
施工道路	1 920 px×2 080 px	2 500	500	3 000

1.2 模型结构

如图 2 所示, 绿网分割模型架构由两部分组成: 编码器和解码器。编码器负责将绿网影像数据信息进行高度抽象编码^[1-3]。解码器部分负责将编码器抽象结果解构, 获得与原始绿网分辨率一样大小的预测图, 预测图中的绿网目标用绿色突出表示。

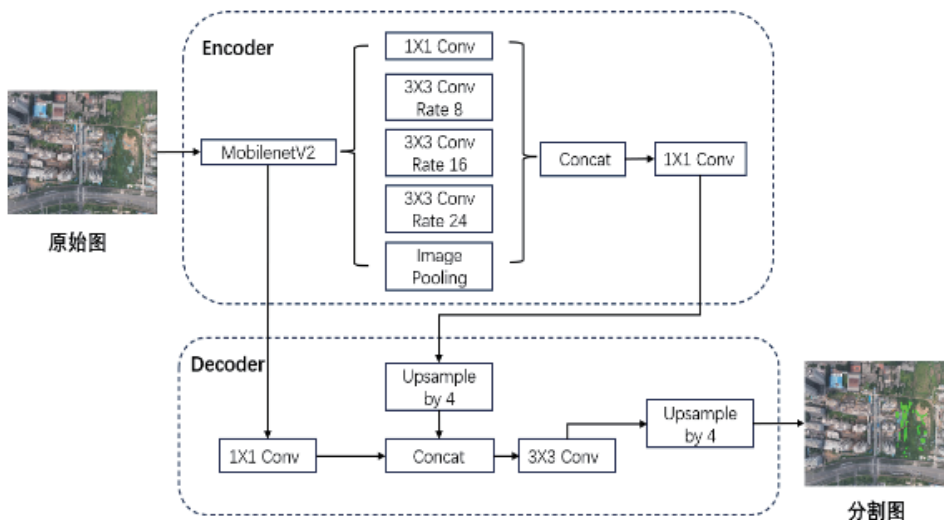


图 2 编解码结构绿网分割模型

为了在相同的计算条件下减少绿网信息损失, 同时加大绿网特征感受范围, 让每个卷积输出都包含较大范围的绿网信息, 在绿网编码器结构中选用了空洞卷积^[4-5], 利用 4 个空洞特征图和 1 个全局平均池化图构建编码器^[6]。具体实现是, 将 MobilenetV2 模型的中间层替换成空洞模块, 输出得到编码器低维抽象特征 1, 将输出结果经过 4 个空洞特征图和 1 个全局平均池化图后得到 5 个相同分辨率的抽象编码, 将抽象编码图进行空间拼接后送入一个卷积块, 得到输出抽象图。解码器核心组件主要由低维特征和线性插值模块构成。将获得低维抽象特征 1 输入到 1×1 低维网络进行降维输出得到特征 3, 将抽象特征 2 利用插值算法对特征图周围像素进行加权平均上采样后与特征 3 进行组合, 并送入 3×3 网络进行处理^[7-8]; 最后再次利用插值算法得到与原图分辨率大小相同的预测结果^[9]。

1.3 插值算法

绿网原图大小为 $n \times m$, 目标特征图大小为 $a \times b$ 。则目标

特征的第 (i, j) 个像素点可以通过计算得出^[10-11]。

$$f(i, j) = \frac{(y_1 - y)(x_1 - x)}{(y_1 - y_0)(x_1 - x_0)} f(x_0, y_0) + \frac{(y_1 - y)(x_1 - x)}{(y_1 - y_0)(x_1 - x_0)} f(x_1, y_0) + \frac{(y - y_0)(x_1 - x)}{(y_1 - y_0)(x_1 - x_0)} f(x_0, y_1) + \frac{(y - y_0)(x_1 - x)}{(y_1 - y_0)(x_1 - x_0)} f(x_1, y_1) \quad (1)$$

$$x = \frac{(i + 0.5) \times n}{a} - 0.5 \quad (2)$$

$$y = \frac{(j + 0.5) \times m}{b} - 0.5 \quad (3)$$

式中: f 为对应图像的像素值; (x, y) 分别为目标坐标 (i, j) 在原始图像中的映射位置; $(x_0, y_0)(x_1, y_0)(x_0, y_1)(x_1, y_1)$ 分别为 (x, y) 对应坐标最近的四个像素点。

1.4 评价指标

为了更好地评价算法的精度, 本文采用逐像素标记的精度标 PA 和 MIoU。PA 表示像素是否正确预测的精度, MIoU 表示真实像素和预测像素的平均交并集。

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \quad (4)$$

$$MIoU = \frac{1}{k + 1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (5)$$

式中: k 表示类别数, 本文共有 2 类 (其中一类是背景, 另一类是绿网); P_{ij} 表示本属于绿网类但被预测为背景类的像素数量; P_{ji} 表示本属于背景类但被预测为绿网类的像素数量, 无论是 P_{ii} 还是 P_{ij} 都属于预测错误的数量; P_{ii} 表示本属于绿网类预测也为绿网类的像素数量, 是模型正确预测出绿网的数量^[12-13]。

1.5 实验配置

本文算法在 Ubuntu 系统上基于 TensorFlow 深度学习框架上进行, 实验平台 GPU 为 NVIDIA GeForce RTX 3080。模型基于 TensorFlow2.0.0 深度学习框架进行搭建, 所用编程语言为 Python3.8。训练时将图像大小统一放到

800 px×800 px，初始学习率设置为 0.001，每次送入网络训练的图像批次为 2，迭代总次数为 50 次，每迭代 10 次保存一次训练细节。

1.6 损失函数

使用二分类交叉熵作为损失函数，优化器选择 Adam 优化器动态调整每个参数的学习率。正例为绿网，负例背景，采用交叉熵损失函数：

$$y_j = \frac{e^{x_j}}{\sum_{i=1}^{N_{\text{class}}} e^{x_i}} \quad (6)$$

$$\text{Loss} = - \sum_{j=1}^{N_{\text{class}}} t_j \log(y_j) \quad (7)$$

式中： x 、 y 为输入输出层的特征； i 、 j 为输出层的索引值； N_{class} 为类别的总数目； t_j 为第 j 类的真实类别； y_j 为模型输出预测概率的结果。

2 实验结果

在不同场景下分别与 R-CNN、SDS、FCN 方法进行对比实验分析：建筑工地场景下结果如表 2 所示，PA 分别提高 2.9%、5.1%、3.7%，MIoU 分别提高 4.0%、5.2%、3.1%；施工道路场景下结果如表 3 所示，PA 分别提高 5.8%、4.7%、7.1%，MIoU 分别提高 9.3%、6.9%、9.3%；组合场景下结果如表 4 所示，组合场景得分 PA 值 0.916，MIoU 为 0.775，PA 分别提高 6.2%、6.7%、7.2%，MIoU 分别提高 6.6%、6.0%、3.9%。

表 2 建筑工地场景下不同检测算法对比

场景	算法	PA	MIoU
建筑工地	R-CNN	0.864	0.724
	SDS	0.842	0.712
	FCN	0.856	0.733
	本文	0.893	0.764

表 3 施工道路场景下不同检测算法对比

场景	算法	PA	MIoU
施工道路	R-CNN	0.844	0.693
	SDS	0.855	0.717
	FCN	0.831	0.738
	本文	0.902	0.786

表 4 组合场景下不同检测算法综合对比

场景	算法	PA	MIoU
建筑工地 + 施工道路	R-CNN	0.854	0.709
	SDS	0.849	0.715
	FCN	0.844	0.736
	本文	0.916	0.775

在建筑工地及施工道路单一场景下，数据单一，无法获取有效的信息表示，在组合场景下，能获取更丰富绿网信息表达，有利于绿网区分，因此 PA 值相比单一场景下有提升；但是组合场景下随着场景变多，无关因素增多，精准率下降，因此 MIoU 值相比施工场景下有一定下降。

各模型算法识别效果图如图 3 所示：其中场景 1 和场景 2 是在不同飞行高度下无人机拍摄的影像数据，场景 1 是正在施工建筑工地，场景 2 是已经施工完毕建筑工地，场景 1 中绿网覆盖土地面积是大范围的，场景 2 中绿网分布范围分散且周围有绿植干扰，绿植的覆盖范围大于绿网，从场景 1 可以看出，相较于其他算法，本文设计网络模型 PA 和 MIoU 识别效果都有提升，但是仍存在不足，在有大范围绿植情况下，MIoU 值仅为 0.764，表示大范围绿植的存在，对绿网识别影响较大。场景 3 和场景 4 是无人机在某一段施工道路上的拍摄影像数据，路段狭长，绿网分布在整个影像的中间位置，分布比较集中，相比场景 1 和场景 2，PA 和 MIoU 都有提高。

3 结论

设计了一种基于编解码结构的网构模型对城市施工场地绿网进行检测，详细介绍了模型的构建原理及方法，并在建筑工地和施工道路两个不同场景下进行了实验分析，与 R-CNN、SDS、FCN 方法相比，PA 分别提高 6.2%、6.7%、7.2%，MIoU 分别提高 6.6%、6.0%、3.9%，验证了本文方法的有效性。虽然本文方法 PA 和 MIoU 都有显著提高，但是 MIoU 最大值 0.775，还有提升空间，后期考虑从插值算法和数据标注质量两个方面进一步研究。

参考文献：

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [2]HARIHARAN B, ARBELA'EZ P, GIRSHICK R,et al. Simultaneous detection and segmentation[EB/OL].(2014-07-07)[2026-04-10].<https://doi.org/10.48550/arXiv.1407.1808>.
- [3] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE transactions on pattern analysis and machine intelligence,2017,39(4):640-651.
- [4] 郭翰林, 林建, 张翔. 基于 HSV 空间再生稻植株与土壤背景图像分割 [J]. 农机化研究, 2017, 39(7): 169-174.
- [5] 陈天鹏, 胡建文. 基于改进 FCOS 的遥感图像舰船目标检

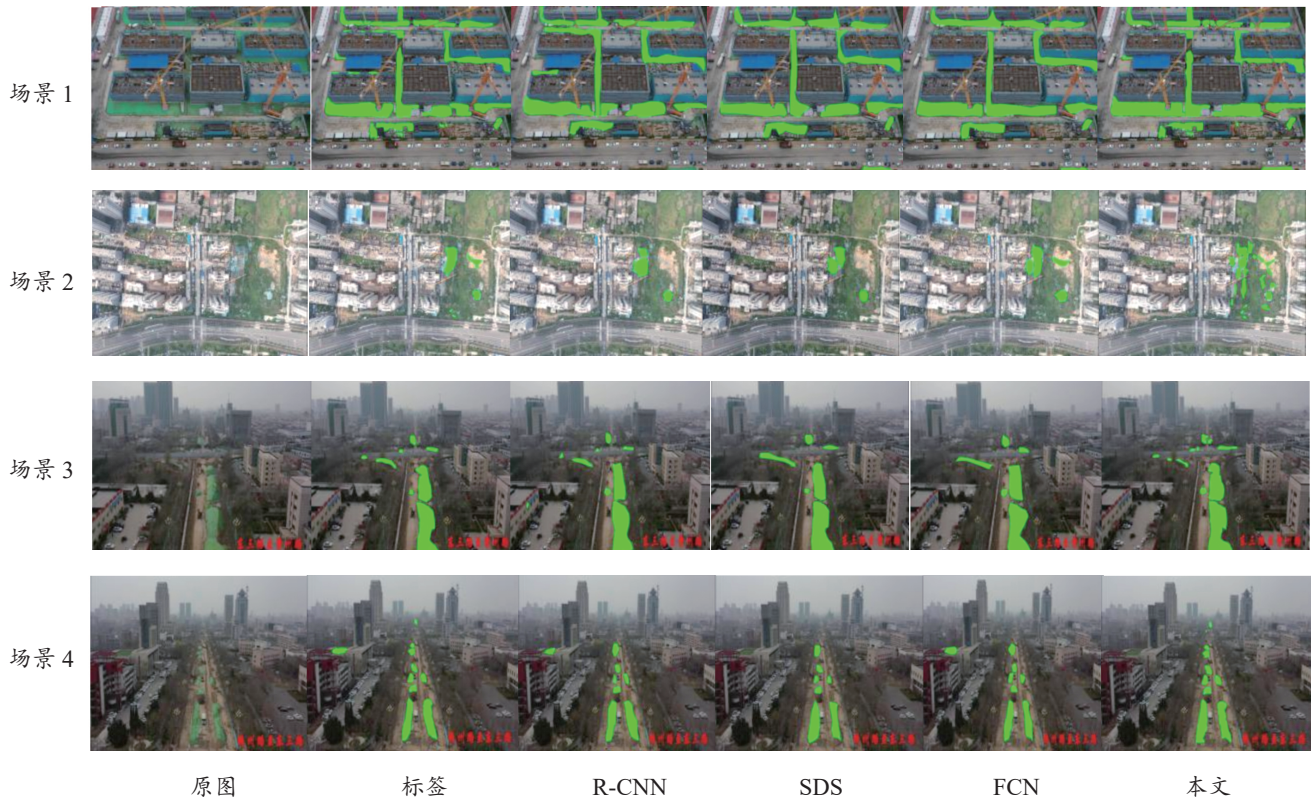


图3 不同场景下各模型算法识别效果图

测[J]. 计算机科学, 2024, 51(S1): 479-485.

- [6] 陈善继, 刘天禹, 刘鹏宇, 等. 基于计算机视觉的公路边坡裂缝监测方法[J]. 北京工业大学学报, 2024, 50(6): 702-710.
- [7] 沙宇洋, 陆京涛, 杜浩凡, 等. 适用于导盲场景的多尺度特征融合轻量化道路图像分割算法[J]. 计算机工程, 2025, 51(7): 314-325.
- [8] 田婷婷, 高瑞贞, 张京军. 基于区域插值的视差图平滑算法及三维重建[J]. 山东农业大学学报(自然科学版), 2020, 51(5): 893-898.
- [9] 张蕊, 张业荣. 基于双线性插值与 K-means 聚类算法结合的位置指纹定位改进算法[J]. 光通信技术, 2022, 46(5): 45-49.
- [10] 周跃进, 丁家益. 一种边缘梯度插值的感兴趣区域池化算法[J]. 吉林大学学报(理学版), 2024, 62(3): 643-654.
- [11] 张文翔, 董宏林. 基于 OTSU 的图像插值算法[J]. 计算机与数字工程, 2022, 50(1): 186-189.
- [12] SVANSTRÖM F, ENGLUND C, ALONSO-FERNANDEZ F. Real-time drone detection and tracking with visible, thermal and acoustic sensors[C]// 2020 25th International Conference on Pattern Recognition (ICPR). Piscataway: IEEE, 2021: 7265-7272.

- [13] BODDHU S K, MCCARTNEY M, CECCOPIERI O, et al. A collaborative smartphone sensing platform for detecting and tracking hostile drones[C]// Proceedings of SPIE - The International Society for Optical Engineering. Berlin: Springer, 2013: 293-303.

【作者简介】

李进强(1987—), 男, 山东济南人, 硕士研究生, 高级工程师, 研究方向: GIS 应用、低空经济。

吕俊杰(1991—), 男, 山东济南人, 硕士研究生, 工程师, 研究方向: 模式识别。

(收稿日期: 2026-02-27 修回日期: 2026-04-16)