

基于 Crack Mask R-CNN 模型的路面裂缝 像素级分割及测量算法应用分析

刘建华¹,董家修²,王念念^{2*},方宏远²

(1.广州市高速公路有限公司,广东 广州 510288;2.郑州大学 水利科学与工程学院,河南 郑州 450001)

摘要:针对现有路面裂缝分割存在的问题,该文提出一种 Crack Mask R-CNN 像素级分割算法。Crack Mask R-CNN 是一种用于路面裂缝图像的实例分割框架,其不仅可对图像中的裂缝进行检测,还可以对每一个裂缝的具体轮廓给出一个高质量的分割结果。首先对采集的道路裂缝大数据进行数据去噪和数据增强,构建用于模型训练、测试的数据集;其次,通过优化分割算法中锚框的比例和大小提高模型选择裂缝候选区域的准确度,并使用 IoU-guided 非极大值抑制(Non-Maximum Suppression, NMS)算法替代传统算法,以此提高道路裂缝分割精度。在模型学习超参数优化方面,通过训练多种组合算例,从中选择分割效果最优的超参数组合,最终训练出裂缝分割精度为 93.45% 的分割模型;最后,通过提取裂缝区域的拓扑特征信息,进一步实现裂缝像素级别尺寸信息的有效测量。

关键词:深度学习;Crack Mask R-CNN;路面裂缝;分割及测量

中图分类号:U418

文献标志码:A

0 引言

裂缝是路面常见的病害之一。路面出现裂缝病害时,随着车辆载荷的作用会导致裂缝两侧进一步破坏,造成路面大面积损坏,增加道路的维护费用。因此,快速、准确地找到道路裂缝的位置并及时修复,可防止路面结构进一步恶化,并且维护费用也会进一步减少^[1]。目前,该项工作已得到了社会各界的高度重视,各类先进的技术手段都已应用到道路裂缝检测中^[2]。

数字图像处理技术应用于路面的病害检测中。基于图像处理的算法包括阈值分割、区域增长及边缘检测的方法^[3-4]。Wang 等^[5]提出了一种基于多尺度局部最优阈值的分割算法,与其他阈值分割方法相比,该算法更有效、鲁棒性更好;Muduli 等^[6]结合双曲线正切(HBT)滤波和 Canny 边缘检测算法进行路面裂缝的检测,可以更加精确地检测到目标区域的边缘,故该方法提高了路面裂缝检测的精度;

Ayenu-Prah 等^[7]提出一种基于二维经验模式分解(BEMD)与 Sobel 边缘检测器检测路面裂缝的方法,该方法使用 BEMD 去除路面图像的噪声,然后使用 Sobel 边缘检测器进行裂缝的检测,以此提高裂缝检测的精度,但是,这种检测方法效率较低,并且对于背景复杂的图像检测效果不佳。

基于机器学习的检测方法已经应用于路面病害检测中。Marques 等^[8]提出了一种基于支持向量机的路面裂缝检测算法。该方法首先对图像进行预处理,然后使用支持向量机的方法检测每块中是否包含裂缝。该方法有效提取了裂缝区域,但是该方法对图像质量要求较高,不能有效提取模糊图像的裂缝轮廓。监督学习与非监督学习的最大不同是用于训练与测试的数据中有没有标签。Akagic 等^[9]提出了一种基于灰度直方图和 Ostu 阈值法的沥青路面裂缝无监督的检测方法,该方法在低信噪比的情况下取得了令人满意的性能。

随着深度学习的发展,深度学习逐渐应用到路面病害的检测与分割中。基于深度学习的路面病害

收稿日期:2022-09-22(修改稿)

作者简介:刘建华,男,硕士,工程师.E-mail:770657629@qq.com

*通信作者:王念念,女,博士,教授.E-mail:wnnian@163.com

检测与分割算法,大大提高了检测与分割的性能。这些方法可以分为非像素级别检测方法、像素级别分割方法两大类。非像素级别检测的目的是在图像中使用边界框定位目标对象并确定目标对象类型。为了提高检测的效率与精度,科研人员提出了许多深度学习模型,如YOLO^[10]、SSD^[11]等算法。Zhang等^[12]提出了一种基于卷积神经网络(CNN)的深度学习模型来进行混凝土裂缝的检测;Cha等^[13]提出了一种包含4个卷积层的卷积神经网络来提取混凝土子裂缝特征,并且取得了优异的效果。但是该方法无法提取裂缝的整体特征。而像素级别的分割方法为图像中的每个像素分别标签,可以提取出路面损害的具体轮廓。Zou等^[14]提出了一种Deep Crack模型,主要使用编码器—解码器将路面图像像素有效分割为背景区域与裂缝区域;David等^[15]提出了一种基于U-Net的分割网络,仅实现了对路面图像裂纹的像素级的有效分割;Yang等^[16]提出了一种基于全卷积神经网络的裂缝分割技术,该网络有效实现了裂缝的分类与逐像素的裂缝分割,并进一步测量了各类裂缝的长度、宽度。但是该方法对于细小裂缝无法进行有效的分割与测量。

基于以上研究存在的问题,为了进一步提高算法的精度与效率,本文提出一种Crack Mask R-CNN路面裂缝像素级实例分割算法。为了适应不同大小的裂缝,通过调整锚的比例和大小来优化RPN。使用IoU-guided NMS^[17]代替非极大值抑制算法(NMS),提高分割的精度。为了优化模型的超参数,通过多次试验,选取分割效果最优的超参数组合。然后与其他分割的模型进行了对比分析,证明本文所提出模型的分割有效性。最后,通过提取裂缝的拓扑特征,进一步测量裂缝的像素级别尺寸信息。

1 Crack Mask R-CNN 分割模型

Crack Mask R-CNN的基本结构如图1所示,主要由卷积神经网络、改进的区域建议网络(Region Proposal Network, RPN)、RoIAlign、分割网络4部分组成。通过预训练的卷积神经网络提取裂缝特征,结合特征金字塔网络(Feature Pyramid network, FPN)^[18]自顶向下进行特征融合,然后由改进的RPN

网络生成若干个候选区域(RoI),并采用IoU-guided NMS代替非极大值抑制算法(NMS)保留精确的候选区域,接着通过RoIAlign将保留的候选区域映射到固定维度,最后通过分割网络,输出路面损害的类别、像素级别的分割结果。

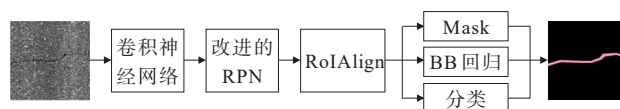


图1 Crack Mask R-CNN网络结构

(1) 卷积神经网络(Convolutional Neural Network, CNN)

Crack Mask R-CNN算法使用具有5个卷积层的ResNet101^[19]作为卷积神经网络。该网络还采用了特征金字塔网络(Feature Pyramid Network, FPN)自顶向下将高分辨率、低语义的底层特征和低分辨率、高语义的顶部特征进行融合,从而丰富了各个尺寸特征的语义信息。卷积的计算方式如式(1)所示:

$$S = X(i, j) * W(i, j) = \sum_m \sum_n X(m, n) W(i - m, j - n) \quad (1)$$

式中: $W(i, j)$ 为卷积核; $X(i, j)$ 为神经网络的输入;*为卷积操作。 i 代表第 i 行, j 代表第 j 列, m 代表裂缝图像的最大行, n 代表裂缝图像的最大列。

(2) 改进的区域建议网络(Region Proposal Network, RPN)

区域建议网络是在卷积神经网络提取特征的基础上获得多个获选区域(Region of Interest, RoI)。区域建议网络采用全卷积神经网络,使用上层网络输出的特征图作为该网络的输入,通过不同纵横比、不同尺度anchor在特征图中滑动,生成若干个候选区域,一次全卷积操作可以获得多个区域的特征。依据路面裂缝病害的特征,为了适应不同形状、不同大小的裂缝区域,本文采用尺度为 32×32 、 64×64 、 128×128 ,纵横比为1:1、1:2、2:1进行候选区域的提取。候选区域计算过程如式(2)所示:

$$d(P) = \omega^T X(P) \quad (2)$$

式中: $X(P)$ 为特征图的特征向量; $d(P)$ 为模型的预测区域值; ω^T 为参数矩阵。

通过上述方式,会产生很多候选区域,并且有冗余候选框。原始RPN网络采用非极大值抑制法(Non-Maximum Suppression, NMS)消除由RPN产

生的冗余候选框,该种方式首先会降序排列每一类别保留的候选区域并选取得分高的目标框,若其余目标区域与该区域重叠面积(IoU)大于一定的阈值时会删除该区域框。该种方式只考虑了分类置信度,但是候选区域框与位置也有密切的联系,故本文通过引入IoU-guided NMS引入位置置信度,解决了分类置信度与位置置信度不匹配的问题。

IoU-guided NMS首先取出位置置信度最大的候选区域框,然后过滤掉与该区域框重叠面积大于一定阈值的区域框,计算剩下候选区域框的分类置信度,并更新分类置信度最高的候选区域框,直至所有的获选区域框处理完毕。NMS是以候选区域框的分类置信度排序,然后按IoU过滤得到分类置信度最高的框。IoU-guided NMS则是以候选区域框的定位置信度排序,然后按IoU过滤,并更新分类置信度。以得到定位置信度最高的框,并且分类置信度。

(3) RoIAlign

RPN保留下来的候选区域和特征图在位置上具有对应关系,需要把这些候选区域映射到与其对应的特征图中。Crack Mask R-CNN采用RoIAlign实现了该过程,并将对应特征向量固定到一定维度。

(4) 分割网络

经过上述网络之后,采用包含Bounding-box回归网络、分类网络及Mask分支的分割网络,实现路面裂缝病害的定位、分类及分割。

2 数据集

2.1 数据集来源

深度学习训练需要大数据的支持,本文主要对裂缝病害的分割进行研究,故在高速公路上使用车载CCD相机实时拍摄高清道路图像,并进行筛选,收集具有裂缝病害的2000张图像作为本文试验的初始数据进行研究。

2.2 数据预处理

2.2.1 数据去噪

首先,在数据采集过程中,采集到的图像通常有各种噪声的干扰,比如拍摄光线干扰造成的图像昏暗、拍摄角度干扰造成的图像灰度分布不均匀等。为了使裂缝病害图像更加清晰,需要对图像进行去

噪处理。根据本文裂缝图像存在噪声的特点,使用高斯过滤器对图像进行光滑去噪,消除拍摄过程中产生的噪声;其次,拍摄到的裂缝图像,有可能尺寸不同,故需要对图像进行裁剪,本文试验将图像裁剪为 512×512 像素。

2.2.2 数据增强

可用于试验的裂缝图像只有2000张,故需要使用数据增强技术进行数据集扩充。本文首先采用翻转变换、平移变换,尺度变换及添加高斯噪声的方式,总计生成用于训练测试的图像6000张。这样不仅增加模型训练的数据量,而且,通过随机添加高斯噪声,提升了训练模型的泛化能力及鲁棒性。

2.3 数据标注

用于模型训练与测试的图像都需要相应的标签才能学习到语义信息,所以对图像进行标注。本文使用开源的labelme软件对图像的背景区域及裂缝区域进行标注,背景及裂缝区域分别记为0,1。按照3:1:1的比例将标注后的图像分为训练集、验证集、测试集。训练集用来训练本文所提出模型,使用验证集对训练的模型进行模型验证并进行超参数选择,测试集用来进行模型最终性能的评估。

3 试验环境及评价指标

3.1 试验环境

数据处理需要高性能计算的支持,本文的工作是基于linux系统中的TensorFlow框架,并且配备了Nvidia 2080Ti的GPU以及32GBRAM的Intel i9的CPU。本文使用CUDA和CUDNN进一步加快图像处理的速度。使用的计算机语言为Python3.6。

3.2 评价指标

图像分割中有许多标准来衡量算法的精度。在本文中采用像素精度来评估算法的分割效果。像素精度(Pixel Accuracy, PA)^[20]表示标记正确的像素占总像素的比例。如式(3)所示:

$$P_{PA} = \frac{\sum_{i=1}^n p_{ii}}{\sum_{i=1}^n \sum_{j=1}^n p_{ij}} \quad (3)$$

式中: P_{PA} 为像素精度; p_{ij} 为属于类*i*而被判定为类*j*的像素数量; p_{ii} 为判定正确的像素数量。

4 试验结果及分析

4.1 模型训练及验证

4.1.1 超参数的选择

在 Crack Mask R-CNN 中需要设置许多影响模型训练的超参数,其中常见且重要的超参数包括初始学习率、动量参数、权重衰减参数。初始学习率影响着模型的学习速率,若初始学习率太大将会导致模型无法收敛;若初始学习率太小,则会导致模型学习速度太慢。动量参数可以提高学习效率,使用权重衰减参数可以有效地防止模型过度拟合。根据 Crack Mask R-CNN 模型训练的需要,开始训练时应选择较大的学习率;训练一段时间后,应缩短参数更新步幅,减小学习率。在本试验中,最大迭代次数设置为 10 000,选择初始学习率为 0.001、0.005、0.000 1,并选择指数衰减方式,每隔 2 000 次迭代学习率衰减一次。为了选择最优的超参数组合,进行多次对比试验,并在验证集进行验证,试验结果如表 1 所示。故选择最优的超参数组合,初始学习率、动量参数、权重衰减参数分别为 0.005、0.85、0.000 2。此时像素精度为 92.98%。

表 1 超参数选择

初始学习率	动量参数	权重衰减参数	损失值	验证集(像素精度)/%
0.001	0.80	0.000 2	0.096 5	90.37
0.001	0.85	0.000 3	0.087 4	91.25
0.001	0.90	0.000 2	0.089 6	91.34
0.005	0.90	0.000 2	0.084 5	90.68
0.005	0.85	0.000 3	0.086 5	91.65
0.005	0.85	0.000 2	0.072 5	92.98
0.005	0.85	0.000 5	0.079 8	92.13
0.005	0.80	0.000 3	0.089 8	91.03
0.000 1	0.80	0.000 2	0.126 2	88.65
0.000 1	0.90	0.000 3	0.112 9	89.12
0.000 1	0.85	0.000 5	0.102 5	90.29

4.1.2 卷积神经网络对模型的影响

Crack Mask R-CNN 中使用了预训练的卷积神经网络进行裂缝特征的初步提取。为了验证卷积神经网络对 Crack Mask R-CNN 分割性能的影响,本文

采用 VGG16、ResNet50、ResNet101 共 3 种卷积神经网络进行试验。本试验对在验证集上的分割 PA 及分割效率进行分析。分割 PA 为在验证集上的像素精度,分割效率为在验证集上的平均分割时间。表 2 显示了基于 3 种预训练卷积神经网络的分割性能。从表 2 可以看出:随着 3 种预训练卷积神经网络深度的增加,分割效率有所下降,平均裂缝分割时间增加了 0.8 s,但是,路面裂缝像素精度越来越高,基于 ResNet101 的 Crack Mask R-CNN 的像素精度达 92.38%,增加了 3.73%。

表 2 3 种卷积神经网络的性能对比

网络	网络层数/层	分割像素精度/%	分割效率/(s·张 ⁻¹)
基于 VGG16 的 Crack Mask R-CNN	16	88.65	2.7
基于 ResNet50 的 Crack Mask R-CNN	50	90.26	2.9
基于 ResNet101 的 Crack Mask R-CNN	101	92.38	3.5

4.2 试验测试

为了验证 Crack Mask R-CNN 路面裂缝分割的有效性,本试验使用测试集中的 1 200 张裂缝图像进行 Crack Mask R-CNN、Mask R-CNN、FCN、边缘检测方法的定量与定性对比分析。图 2 为 4 种模型在混凝土及沥青路面裂缝分割示例。从图 2 可以看出:本文提出的 Crack Mask R-CNN 可以较为完整、准确地提取到裂缝轮廓,并且没有冗余信息,但是 Mask R-CNN、FCN 分割模型与 Crack Mask R-CNN 相比,分割效果有待提高。边缘检测方法裂缝分割效果较差,并且会产生一些冗余信息,这是由于边缘检测等数字图像处理方法对于目标区域与背景区域灰度值变化较大的检测效果较好,对于变化较小的检测效果有待提高。

表 3 展示了 4 种分割模型在测试集上的定量分割性能。可以看出:Crack Mask R-CNN 与其他 3 种分割模型相比具有最优的分割像素精度值,达 93.45%,分别增加了 3.18%、3.40%、12.24%。但是由于 Crack Mask R-CNN 的网络结构更加复杂,平均分割时间有所提高。边缘检测方法的路面裂缝分

割时间最短,但是分割效果最差,难以达到有效分割。

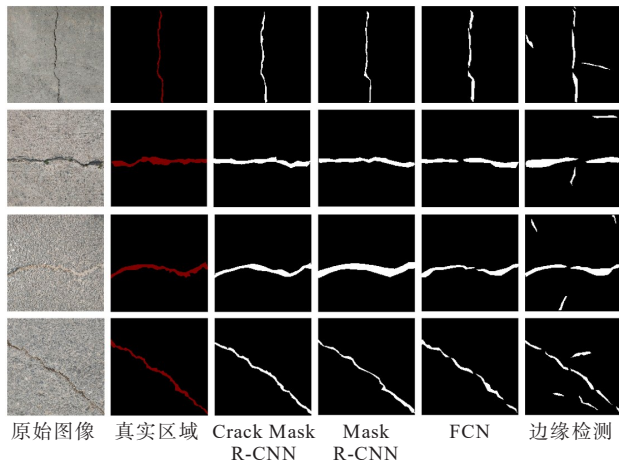


图2 4种模型分割效果对比示例

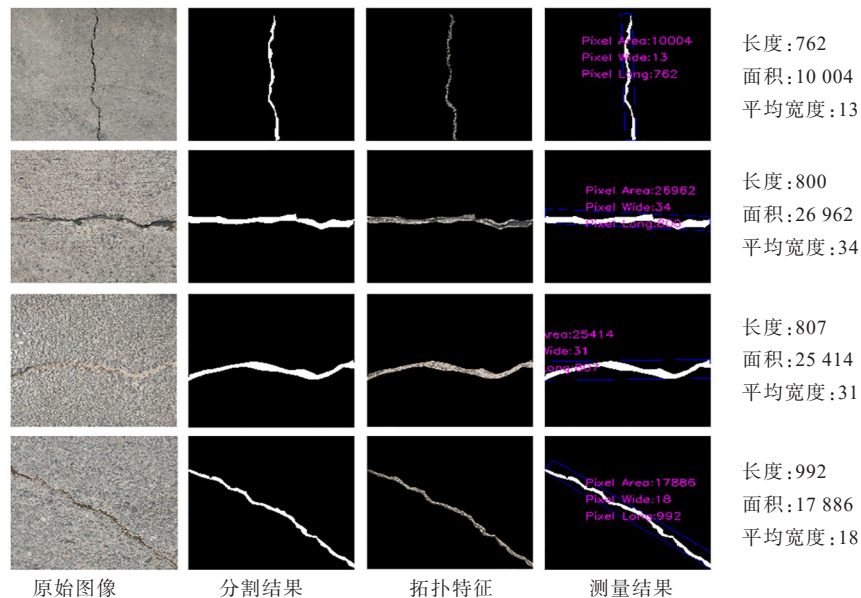


图3 路面裂缝像素级别尺寸信息测量示例

5 结语

针对路面裂缝的精确分割与测量,提出了一种Crack Mask R-CNN智能分割模型。本文通过预训练的卷积神经网络并结合特征金字塔网络提取裂缝特征,在RPN网络中通过优化锚框的比例和大小提取更加精确的候选区域,并采用IoU-guided NMS保留定位信度及分类置信度较高的候选区域,以此提高了分割的性能。试验结果表明:本文提出的Crack Mask R-CNN分割模型具有有效的分割效果,

表3 四种分割模型性能对比

模型	分割像素精度/ %	分割效率/ (s·张 ⁻¹)
Crack Mask R-CNN	93.45	3.2
Mask R-CNN	90.27	2.4
FCN	90.05	2.5
边缘检测	81.21	1.8

4.3 裂缝尺寸测量

为了给公路养护部门提供更好的帮助,本文在路面裂缝良好分割的基础上,对裂缝像素级别的长度、平均宽度及面积展开研究。首先基于分割得到的Mask掩码图像得到裂缝的拓扑特征,然后使用遍历算法统计裂缝的面积及长度像素级别尺寸信息,通过面积像素尺寸信息/长度像素尺寸信息,从而得到裂缝的平均宽度像素尺寸信息,如图3所示。

与Mask R-CNN、FCN、边缘检测方法进行对比,证明了本文提出方法具有更高的分割精度。在良好分割结果的基础上,实现了路面裂缝像素级别尺寸信息的测量,可为公路养护部门提供帮助。

参考文献:

- [1] CHEN J, WU J J, CHEN G, et al. Design and development of a multi-rotor unmanned aerial vehicle system for bridge inspection[C]//KUBOTA N, KIGUCHI K, LIU H, et al. International Conference on Intelligent Robotics and Applications. Cham: Springer, 2016: 498-510.
- [2] 周基,蔡强,田琼. 70年中国公路路基路面病害研究现状

- 与发展趋势:基于CNKI 1949—2019年文献的知识图谱分析[J].中外公路,2020,40(3):60-66.
- [3] CAO W M, LIU Q F, HE Z Q. Review of pavement defect detection methods[J]. IEEE Access, 2020, 8: 14531-14544.
- [4] ZOU Q, CAO Y, LI Q Q, et al. CrackTree: Automatic crack detection from pavement images[J]. Pattern recognition letters, 2012, 33(3): 227-238.
- [5] WANG S C, TANG W S. Pavement crack segmentation algorithm based on local optimal threshold of cracks density distribution[C]//International Conference on Intelligent Computing. Berlin, Heidelberg: Springer, 2011: 298-302.
- [6] MUDULI P R, PATI U C. A novel technique for wall crack detection using image fusion[C]//2013 International Conference on Computer Communication and Informatics. January 4-6, 2013, Coimbatore, India. IEEE, 2013: 1-6.
- [7] AYENU-PRAH A, ATTOH-OKINE N. Evaluating pavement cracks with bidimensional empirical mode decomposition[J]. EURASIP Journal on Advances in Signal Processing, 2008, 2008(1): 861701.
- [8] MARQUES A, CORREIA P L. Automatic road pavement crack detection using SVM[Z], 2012.
- [9] AKAGIC A, BUZA E, OMANOVIC S, et al. Pavement crack detection using Otsu thresholding for image segmentation [C]//2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics(MIPRO). May 21-25, 2018, Opatija, Croatia. IEEE, 2018: 1092-1097.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [12] ZHANG L, YANG F, ZHANG Y D, et al. Road crack detection using deep convolutional neural network[C]//2016 IEEE International Conference on Image Processing (ICIP). September 25-28, 2016, Phoenix, AZ, USA. IEEE, 2016: 3708-3712.
- [13] CHA Y J, CHOI W, BÜYÜKÖZTÜRK O. Deep learning-based crack damage detection using convolutional neural networks[J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [14] ZOU Q, ZHANG Z, LI Q Q, et al. DeepCrack: Learning hierarchical convolutional features for crack detection[J]. IEEE Transactions on Image Processing, 2019, 28(3): 1498-1512.
- [15] DAVID JENKINS M, CARR T A, IGLESIAS M I, et al. A deep convolutional neural network for semantic pixel-wise segmentation of road and pavement surface cracks[C]//2018 26th European Signal Processing Conference (EUSIPCO). September 3-7, 2018, Rome, Italy. IEEE, 2018: 2120-2124.
- [16] YANG X C, LI H, YU Y T, et al. Automatic pixel-level crack detection and measurement using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(12): 1090-1109.
- [17] JIANG B R, LUO R X, MAO J Y, et al. Acquisition of localization confidence for accurate object detection[C]//European Conference on Computer Vision. Cham: Springer, 2018: 816-832.
- [18] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: Piscataway, N.J. 2017: 2117-2125.
- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [20] 翁飘, 陆彦辉, 齐宪标, 等. 基于改进的全卷积神经网络的路面裂缝分割技术[J]. 计算机工程与应用, 2019, 55(16): 235-239, 245.