

引文格式:肖力炀,李伟,袁博,等.一种基于改进实例分割模型的路面裂缝检测方法[J].武汉大学学报(信息科学版),2023,48(5):765-776.DOI:10.13203/j.whugis20210279



Citation: XIAO Liyang, LI Wei, YUAN Bo, et al. A Pavement Crack Identification Method Based on Improved Instance Segmentation Model[J]. Geomatics and Information Science of Wuhan University, 2023, 48(5):765-776. DOI:10.13203/j.whugis20210279

# 一种基于改进实例分割模型的路面 裂缝检测方法

肖力炀<sup>1</sup> 李 伟<sup>2</sup> 袁 博<sup>2</sup> 崔逸群<sup>1</sup> 高 荣<sup>2</sup> 王文庆<sup>1</sup>

1 西安热工研究院有限公司,陕西 西安,710032

2 长安大学信息工程学院,陕西 西安,710064

**摘要:**为了解决现有裂缝识别算法准确度不高、检测与分割任务不能同时进行等问题,提出了一种基于改进型Mask R-CNN模型的路面裂缝识别方法。首先,建立裂缝数据集并进行标注,然后使用Mask R-CNN模型对裂缝数据集进行训练和测试,并对模型中锚点的长宽比进行调整,实现在裂缝定位的同时对生成的检测框内的裂缝像素进行分割;其次,针对Mask R-CNN模型生成的裂缝检测框不精准的问题,设计了C-Mask R-CNN多阈值检测方法,通过结合级联不同阈值的检测器来提高候选框质量,实现高阈值检测下的裂缝精准定位;最后,对改进后的模型进行一系列的优化参数和实验对比,并验证所提模型的有效性。实验结果表明,C-Mask R-CNN模型检测部分的平均准确率均值(mean average precision, mAP)达到0.954,与改进前模型相比提升了9.7%,分割部分的mAP达到0.935,与改进前相比提升了13.0%,识别效果较好。综上所述,C-Mask R-CNN模型可以较为完整地对裂缝进行定位及提取,识别精度较高。

**关键词:**路面裂缝识别;深度学习;Mask R-CNN模型;级联阈值检测器;道路工程

中图分类号:U418.6;P237

文献标识码:A

收稿日期:2021-12-07

DOI:10.13203/j.whugis20210279

文章编号:1671-8860(2023)05-0765-12

## A Pavement Crack Identification Method Based on Improved Instance Segmentation Model

XIAO Liyang<sup>1</sup> LI Wei<sup>2</sup> YUAN Bo<sup>2</sup> CUI Yiqun<sup>1</sup> GAO Rong<sup>2</sup> WANG Wenqing<sup>1</sup>

1 Xi'an Thermal Power Research Institute, Xi'an 710032, China

2 School of Information Engineering, Chang'an University, Xi'an 710064, China

**Abstract:** Objectives: To solve problems that the existing crack identification algorithms are not accurate and the detection and segmentation tasks cannot be performed simultaneously, we propose a pavement crack identification method via the improved Mask R-CNN model. Methods: First, the crack dataset is collected and labeled. And the crack dataset is trained and tested by Mask R-CNN model, and the aspect ratios of the anchor points in the model are adjusted to segment the crack pixels in the generated detection box while the crack is located. Second, to solve the problem that the crack detection boxes generated by Mask R-CNN model are inaccurate, C-Mask R-CNN is designed to improve the quality of crack region proposal by cascading multi-threshold detectors and achieve accurate positioning under high threshold. Finally, a series of optimization parameters and experimental comparison are carried out for the improved model, and the effectiveness of the proposed model is verified. Results: The experimental results show that the mean average precision (mAP) of C-Mask R-CNN model in the detection part is 0.954, which is 9.7% higher than that of the conventional model, and its mAP in the segmentation part is 0.935, which is 13.0% higher than that of the conventional model. It confirms that the C-Mask R-CNN model performs well in identi-

**基金项目:**国家重点研发计划(2018YFB1600202);国家自然科学基金(51978071);国家自然科学基金青年科学基金(51908059)。

**第一作者:**肖力炀,硕士,主要从事人工智能与深度学习领域的研究。liyang\_xiao@163.com

**通讯作者:**李伟,博士,教授。grandy@chd.edu.cn

fying cracks. **Conclusions:** In summary, the proposed C-Mask R-CNN model can locate and extract cracks with high identification accuracy.

**Key words:** pavement crack identification; deep learning; Mask R-CNN model; cascade threshold detector; road engineering

在公路基础设施建设中,混凝土起着非常重要的作用,而路面裂缝又是混凝土结构表面维护中很常见的一种道路病害,它对公路的使用期限及车辆的行车安全均会产生严重危害,并且还需要花费大量费用进行修补维护。因此,对路面裂缝进行检测及识别具有极其重要的意义。

近年来,国内外学者对路面裂缝识别的研究也取得了一定的成果。邹勤等<sup>[1]</sup>对裂缝全局几何特征进行分析,以实现裂缝检测。该方法采用最小生成树原理描述裂缝特征,对间断的裂缝也有较好的识别效果。Premachandra等<sup>[2]</sup>对裂缝图像像素进行方差分析,采用改进的判别分析法来提取裂缝。Oliveira等<sup>[3]</sup>提出了一套完整的基于块和像素方法的图像处理算法来对路面裂缝进行检测,但该方法对细裂缝的检测精度稍低。马荣贵等<sup>[4]</sup>提出了一种基于多窗口的滤波算法,用于解决路面裂缝中不均匀的光照现象,并通过使用灰度矫正、阈值分割等方法有效解决了裂缝不均匀成像对分割的影响。Yoo等<sup>[5]</sup>提出了一种通过消除噪声来检测裂缝的方法,将神经网络与二元 Logistic 回归分析结合起来,进而从噪声中智能识别裂缝。赵珊珊等<sup>[6]</sup>针对路面裂缝分布复杂的问题,提出了一种基于卷积神经网络(convolutional neural network,CNN)模型的裂缝检测方法,但由于噪声等干扰,模型的检测精度并不高。王丽苹<sup>[7]</sup>提出了一种基于 Faster R-CNN 的混凝土路面裂缝检测方法,实现了从粗到精的裂缝实时定位。Zou等<sup>[8]</sup>提出了结合多尺度信息的深度卷积网络来对裂缝进行分割,该方法在 SegNet<sup>[9]</sup>结构基础上构建了 DeepCrack 网络,并将编码-解码网络中产生的卷积特征以相同的尺寸成对融合。Maeda 等<sup>[10]</sup>提出了一种基于 SSD(single shot multibox detector)轻量级网络的裂缝检测方法,结合 MobileNet 和 Inception V2 网络对裂缝进行特征提取。侯越等<sup>[11]</sup>针对裂缝数据集小、标注成本高等问题,提出了先基于卷积自编码进行数据扩增,再使用 DenseNet 进行裂缝识别方法,该方法无须人工标注且为实现端到端的裂缝检测提供了可行性。Alipour 等<sup>[12]</sup>针对固定尺寸输入图像造成的精度缺失现象,提出了一种基于深度全卷积 CrackPix 网络的像素级别裂缝检测系统。

Nguyen 等<sup>[13]</sup>针对道路裂缝背景复杂等问题,提出了一种基于灰度图像的深度神经网络(deep neural network,DNN)的像素级裂缝识别方法。

从上述裂缝识别研究现状中可以看出,随着人工智能技术的发展,国内外学者对裂缝识别的研究重点从数字图像处理技术逐渐转移到深度学习,并取得了不错的成果。但现有的很多算法使用场景局限,完成任务单一,无法同时实现裂缝的检测与分割任务,并且实际工程中的路面情况往往比较复杂,因此依然存在多方面问题。基于此,本文提出了一种基于改进型 Mask R-CNN 深度学习模型的路面裂缝识别方法。该方法在实现裂缝高阈值定位的同时,能对生成的检测框内的裂缝像素进行分割,有效提高了路面裂缝识别的准确度,在路面裂缝的自动化识别和道路养护工作中具有重要的工程应用价值。

## 1 基于 Mask R-CNN 的裂缝识别方法

在路面裂缝的识别任务中,使用检测网络可以实现裂缝定位,但不能计算裂缝长宽、面积等参数;使用分割网络能对裂缝像素进行分割,但无法定位裂缝位置以及区别裂缝条数。这种使用单一任务的网络模型(如检测网络或分割网络)只能完成路面裂缝的检测或分割任务。当使用检测与分割的组合模型时,会出现一些由裂缝像素分割结果断裂引起的裂缝数目错判问题,或者是检测网络精度低造成检测框框不全裂缝,从而导致分割结果不完整的现象。针对以上问题,本文将 Mask R-CNN 用于路面裂缝的识别工作。如图 1 所示,Mask R-CNN 可以将裂缝的检测与分割任务融合在一个网络模型中,在实现目标定位的同时对生成的检测框内的裂缝像素进行分割。

### 1.1 Mask R-CNN 原理

#### 1.1.1 特征金字塔网络

Mask R-CNN 采用 CNN 加特征金字塔网络(feature pyramid network,FPN)的结构进行图像特征提取。通过将低层特征和高层特征进行多次融合,解决了多尺度预测的问题。首先将输入

图像从下到上进行下采样,特征图尺寸越来越小;然后从上到下对特征图上采样,使其越变越大;最后横向连接,将高语义信息和强位置信息不断相加,得到不同尺度的特征。FPN 的特点在于它可以实现精准定位,适合裂缝等细小目标的检测。

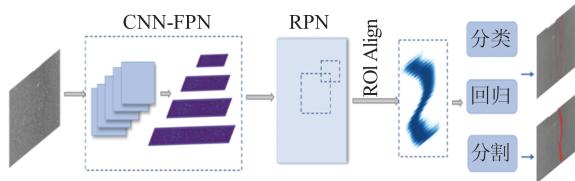


图 1 Mask R-CNN 裂缝识别模型示意图

Fig.1 Diagram of Crack Identification Model  
Based on Mask R-CNN

### 1.1.2 区域建议网络

区域建议网络(region proposal network, RPN)用于生成高质量的区域候选框,网络的输入为加入 FPN 的特征提取网络生成的特征图。由于本文检测的目标是裂缝,结合线性裂缝细长的特点,本文对锚点 anchor 的长宽比进行调整,调整比例为{1:1, 1:3, 3:1},以适应裂缝目标的检测。图 2(a)为长宽比调整前特征图中某个点映射至原始裂缝图像上生成的候选区域,图 2(b)为调整后生成的候选区域。

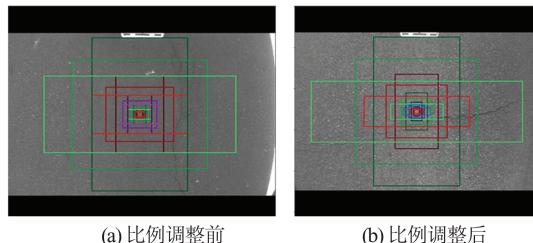


图 2 映射至原图的候选区域

Fig.2 Proposal Region Mapped to the Original Image

### 1.1.3 mask 掩膜分支

与 R-CNN 系列<sup>[14-16]</sup>其他算法相比,Mask R-CNN 最大的不同在于其增加了一个 mask 掩膜部分,通过连接全卷积网络来对特征图中的每个像素点进行分类<sup>[17]</sup>,从而实现像素级的目标分割。

Mask R-CNN 算法采用多任务的训练方式,其损失函数计算公式如下:

$$L = L_{rc} + L_{rb} + L_{cls} + L_{box} + L_{mask} \quad (1)$$

式中, $L_{rc}$  表示 RPN 分类损失; $L_{rb}$  表示 RPN 边界回归损失; $L_{cls}$  表示目标分类损失; $L_{box}$  表示边界框回归损失; $L_{mask}$  表示掩膜分割结果的损失。

Mask 分支采用逐像素 Sigmoid 来计算损失,每个像素对应一个预测向量,对于一个类别为  $n$

的兴趣区域, $L_{mask}$  代表在第  $n$  个 mask 上输出的损失(其余  $n-1$  个 mask 输出对整个损失无贡献),计算公式如下:

$$L_{mask} = -\frac{1}{a^2} \sum_{1 \leq i, j \leq a} [y_{ij} \log_2 \hat{y}_{ij}^n + (1 - y_{ij}) \log_2 (1 - \hat{y}_{ij}^n)] \quad (2)$$

式中, $y_{ij}$  表示  $a \times a$  掩膜区域内的像素  $(i, j)$  标签; $\hat{y}_{ij}^n$  代表在相同位置上第  $n$  层 mask 上像素的预测值。

### 1.2 数据准备与环境搭建

本文使用智能检测车设备来采集路面裂缝图像,见图 3,在校园内和高速公路上采用了不同设备以适应不同路面和场景。本文重点对高速公路沥青路面裂缝进行采集,采集的裂缝数据集包含横向裂缝、纵向裂缝和网状裂缝 3 种类别。本文用于模型训练和测试的样本数据集共 6 400 张,其中横向裂缝 2 200 张、纵向裂缝 2 200 张、网状裂缝 2 000 张,按照 9:1 的比例将数据随机划分为训练集(5 760 张)和测试集(640 张),每张图像尺寸为  $1280 \times 960$  像素,数据集链接为 [https://github.com/juhuyan/CrackDataset\\_DL\\_HY](https://github.com/juhuyan/CrackDataset_DL_HY)。图 4 展示了 3 种类别的部分裂缝图像,可以看出,本文使用的裂缝数据集中,裂缝与背景之间的像素较为相近,且存在阴影、路面污染等噪声。



图 3 采集裂缝数据

Fig.3 Collection of Crack Images

裂缝识别模型的开发框架为 MMDetection 检测库,相关的实验环境设备及软件配置如表 1 所示。MMDetection 是一个开源的深度学习工具箱,它包含 Faster R-CNN、Mask R-CNN 等多种深度学习框架,其优点是性能较高、训练速度快且所需显存小<sup>[18]</sup>。

### 1.3 模型训练结果

本文使用 ResNet101 和 ResNeXt101 两种特征提取网络分别对模型进行训练,表 2 为两种特征提取网络下 Mask R-CNN 模型的检测平均准确率(average precision, AP)与分割 AP 结果,其中,

mAP@50表示交并比(intersection over union,IoU)设为0.5时检测的平均准确率均值(mean average precision,mAP)。与ResNet相比,使用ResNeXt后,检测mAP从0.713提升至0.857,分割mAP从0.658提升至0.805,效果较好,因此本文选择ResNeXt101作为裂缝识别模型的特征提取网络。

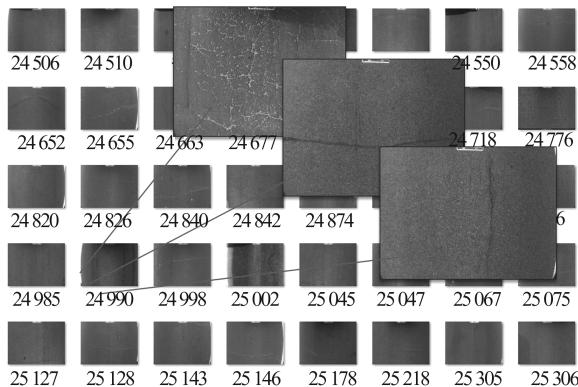


图4 部分裂缝数据集

Fig.4 Partial Crack Datasets

表1 实验环境配置说明

Tab.1 Details of Experimental Environment Configuration

名称	版本
操作系统	Ubuntu 18.04 LTS
GPU显卡	NVIDIA GeForce GTX 1080Ti
CUDA	Cuda-9.0.1-Linux
CUDNN	Cudnn-7.3.1-Linux
mmdet	1.0rc0
mmcv	0.4.3
Python	3.6
Pytorch	1.2.0
Anaconda	Anaconda3-6.0.1-Linux-x86_64

表2 两种特征提取网络下的模型对比

Tab.2 Model Comparison Under Two Feature Extraction Networks

类别	ResNet101		ResNeXt101	
	检测	分割	检测	分割
横向裂缝 AP	0.796	0.753	0.906	0.887
纵向裂缝 AP	0.764	0.712	0.909	0.873
网状裂缝 AP	0.579	0.508	0.756	0.654
mAP@50	0.713	0.658	0.857	0.805

图5展示了Mask R-CNN裂缝识别模型的精确-召回(precision-recall,Pr-Re)曲线。Pr-Re曲线下面积大小代表了模型准确度的高低(即AP值),面积越大,则准确度越高。图5中,C75表示当划分正负样本IoU阈值为0.75时的AP值;C50表示当IoU为0.5时的AP值;Loc表示当IoU为0.1时的AP值;Sim表示除去相似种类混淆后的AP值;Oth表示不同种类间混淆的AP值;BG表示除去背景上假正例的AP值;FN表示除去所有错误后的AP值。由图5(a)可知,当IoU=0.5时,检测AP为85.7%,IoU=0.75时,检测AP只有52.8%,检测精度骤降。

淆后的AP值;Oth表示不同种类间混淆的AP值;BG表示除去背景上假正例的AP值;FN表示除去所有错误后的AP值。由图5(a)可知,当IoU=0.5时,检测AP为85.7%,IoU=0.75时,检测AP只有52.8%,检测精度骤降。

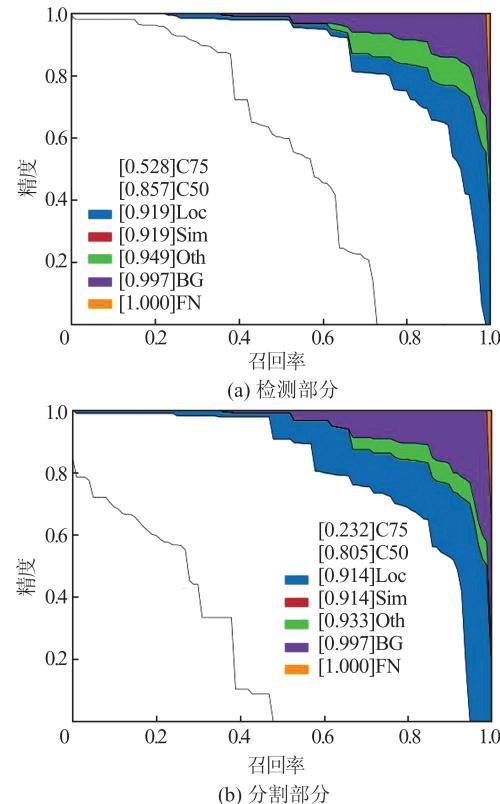


图5 Mask R-CNN模型的Pr-Re曲线图

Fig.5 Pr-Re Curves of Mask R-CNN Model

图6展示了3种裂缝类别下模型测试的部分效果图,可以看出,利用Mask R-CNN模型基本可以正确区分出裂缝,其对裂缝轮廓的分割较为细致,但对裂缝边缘部分的检测不是很精确,会存在一些检测框框不全裂缝的现象(见图6(c)和图6(f)),以及一条裂缝被误识别为两条裂缝的情况(见图6(a)和图6(d)),其对于网状裂缝的识别效果不是很好(见图6(g)~图6(i))。针对这些不足,本文将在下一节中改进。

## 2 基于C-Mask R-CNN模型的裂缝识别方法

### 2.1 级联多阈值检测器原理

通常通过某个IoU阈值下的AP(Pr-Re曲线下面积)来评价模型的性能。图5(a)出现的检测精度骤降现象表明,IoU阈值越高,对检测结果的要求就越高,相应的AP值会降低。检测结果不

精确会直接导致分割结果不完整,因此,实现高 IoU 阈值下的检测有助于获得高精度的分割效果。

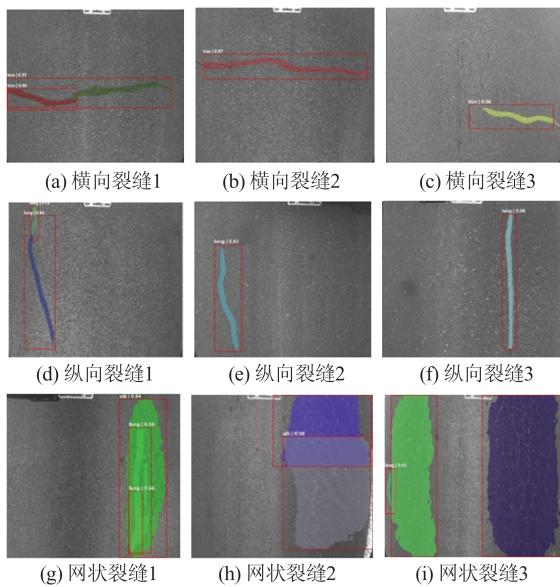


图 6 Mask R-CNN 模型测试裂缝识别效果图  
Fig.6 Crack Identification Results Obtained by Mask R-CNN Model

对于裂缝检测问题,由于裂缝的形状和尺寸大小不尽相同,即使是同一类别的裂缝,其长度也会不一样,有的裂缝较为细长,有的比较窄短,网状裂缝的面积往往又很大,这些形状大小不一的裂缝图像给检测任务带来了很大挑战。Mask R-CNN 采用的是单阈值检测器,在进行多分类任务时为了均衡正负样本,把划分正负样本的 IoU 阈值设为 0.5,但这种划分方法其实并不适合裂缝目标的检测,因为当 IoU 阈值为 0.5 时,候选区域内存在许多背景噪声,对于裂缝这种一般存在于候选框对角线位置上的细长形目标,会因为背景的干扰而仅使很少一部分裂缝目标作为正样本参与模型训练,导致检测结果出现大量误检。但直接提高 IoU 阈值又会使正负样本的分布不均衡,从而导致正样本数量减少而出现漏检现象<sup>[19]</sup>。针对 Mask R-CNN 这种单阈值检测的弊端,引入了 Cascade R-CNN<sup>[20]</sup>模型中多阈值检测的思想,通过级联多个不同的阈值检测器逐步提高 IoU 阈值,实现高阈值检测下的精准定位。

Cascade R-CNN 结构如图 7 所示,它是一种级联的网络结构,模型的输入由两部分组成,分别是通过 RPN 产生的候选区域  $B_0$  和通过 CNN 生成的特征图,以图 7 为例,它使用了 3 个检测器,每个检测器都包含有池化层、全连接层( $H$ )、分类部分( $C$ )和边界回归部分( $B$ ),每个检测器的

输入都是上一个检测器经过边界回归生成的候选框  $B$ ,即  $B_1$  连接第二个检测器的池化层,  $B_2$  连接第三个检测器的池化层,越后面的检测器设定的 IoU 值越大,这就可以使每个检测器只对某一个 IoU 阈值下的候选区域进行检测,从而提升候选框的质量和模型的训练效果。当只运行前向运算预测类别时,Cascade R-CNN 会将最后一层检测器输出的候选框送入前面每个检测器,各自计算分类,然后对各个检测器的分类结果进行集成计算,作为最终的结果输出。

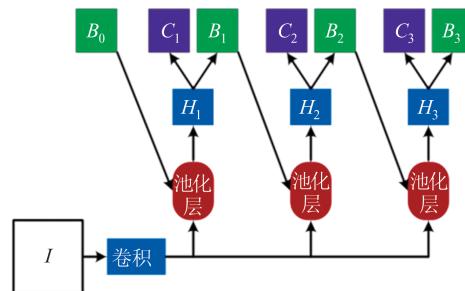


图 7 级联检测器结构  
Fig.7 Structure of Cascade Detector

级联检测器的模型架构如下:

$$f(x, b) = f_T, f_{T-1}, \dots, f_1(x, b) \quad (3)$$

式中,  $T$  表示级联的个数;  $b$  表示样本。每个检测器对应本阶段的样本  $b_T$ ,下一阶段得到的样本  $b_T$  会比前一阶段的样本  $b_{T-1}$  质量高。

## 2.2 C-Mask R-CNN 方法

Mask R-CNN 只包含了一种检测器,这大大限制了模型的检测性能,在指定的 IoU 阈值增大时,会出现 AP 值骤降的情况,从而对后续的裂缝分割部分造成一定影响,而 Cascade R-CNN 通过级联多个不同阈值的检测网络逐层提高候选区域的 IoU 阈值,通过这种方式训练出来的模型可以有效避免 AP 值骤降的问题,实现高阈值检测下的高精度定位。因此,本文设计了一种基于级联多阈值检测的 C-Mask R-CNN 方法,以完成路面裂缝的识别工作。该方法将 Cascade R-CNN 模型与 Mask R-CNN 模型结合起来,在实现裂缝目标高精度检测的基础上,对检测框内的裂缝进行像素分割,具有了两种模型的优点。

图 8 为加入级联多阈值检测器的 Mask R-CNN 模型结构。该模型在 RPN 生成候选区域之后级联了 3 个检测器来分别完成目标的分类和回归,池化方式摒弃了 Cascade R-CNN 中的 ROI Pooling 池化,而是选择保留小数信息的 ROI Align 池化。如图 8 中虚线箭头所示,前一个检测

器经过边界回归生成的候选框,会被送进下一个检测器,越靠后的检测器指定其划分的IoU阈值越高,生成候选框的质量也就越高,同时高质量的候选框反过来又会作为输入训练更高阈值下的检测器,通过这种逐步回归的检测方式来提高模型的检测精度。在模型最后一层检测器上加入mask掩膜分支部分完成裂缝分割任务。但该方法仅对Mask R-CNN的检测部分进行了阈值级联,并未涉及模型的mask分割部分,因此,尽管提升生成检测框的质量会对分割效果起到一定帮助,但这种没有与mask分支进行信息交互的级

联方法对裂缝的高精度分割来说效果不大。

基于此问题,本文在上述结构的基础上设计了图9所示的C-Mask R-CNN模型,在每个阶段都会生成相应的mask,将这些mask分支串联起来完成裂缝像素信息的交互。如图9中虚线箭头所示,当前阶段的 $mask_i$ 会与下一阶段的 $mask_{i+1}$ 进行卷积操作,与FPN中横向连接的思想类似,通过 $1 \times 1$ 的卷积核使 $mask_{i+1}$ 得到 $mask_i$ 的特征信息。这种对每一阶段输出的mask分支信息进行融合的方式可以使裂缝像素的分割更为细致,从而实现高精度的裂缝检测与分割任务。

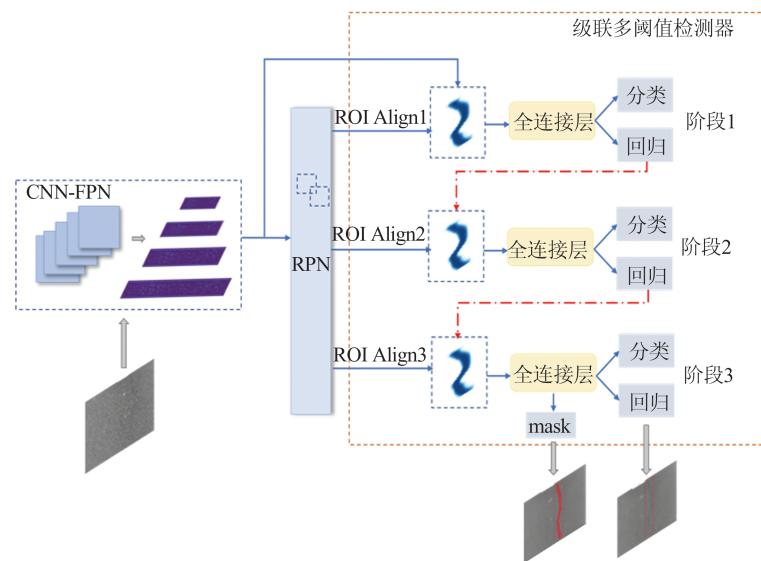


图8 级联阈值检测器模型结构

Fig.8 Model Structure of Cascade Threshold Detector

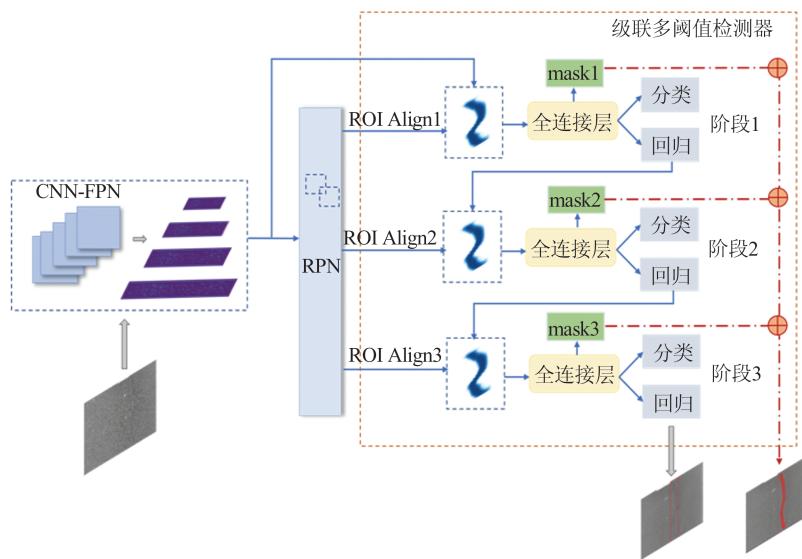


图9 C-Mask R-CNN模型结构

Fig.9 Structure of C-Mask R-CNN Model

当级联多个检测器时,越靠后的检测器生成候选框的质量越好,对目标的定位也就越准确,但当级联检测器数量足够多时,前后层训练出来

的检测器差异会越来越大,后检测器输出的特征与前检测器输出的特征之间差异越明显,检测效果也不一样。由于输出的检测结果是综合了各

个层检测器的结果,越靠后的检测器输出特征在使用较前的检测器进行检测时,会因前后特征的不匹配而造成准确度下降。因此,级联检测器数目过多时反而会出现检测性能降低的情况。本文分别级联 3、4、5 层检测器进行实验,实验结果如表 3 所示,其中,mAP@75 表示当 IoU 阈值设定为 0.75 时裂缝的平均检测精度,由结果可知,当检测器个数为 3 时,检测性能最好。

**表 3 级联不同个数的检测器**

Tab.3 Cascade Detectors with Different Numbers

检测器个数	IoU 阈值	mAP@75	mAP@50
3	0.5,0.6,0.7	0.860	0.951
4	0.5,0.6,0.7,0.8	0.853	0.948
5	0.5,0.58,0.66,0.74,0.82	0.839	0.945

检测器设定的 IoU 阈值与输入候选框 IoU 阈值越接近时,检测性能越好<sup>[21]</sup>,随着检测器逐层回归,候选框与真实标签之间的 IoU 也会有所提升,并且越靠后的检测器生成的候选框与真实标签之间的差距会越小,从而导致候选框的质量提升程度越来越低。因此,为了减弱候选框质量提升程度降低带来的影响,本文对级联 IoU 阈值采用线性增长和曲线增长两种方式进行实验对比,线性增长 IoU 阈值设为 0.5、0.6 和 0.7,每次固定递增 0.1 个阈值;曲线增长 IoU 阈值设为 0.5、0.62 和 0.72,每次递增阈值数依次减小。表 4 为两种增长方式的对比结果,可以看出,当检测器阈值的提升采用曲线增长方式时检测性能更好。综上所述,本文采用级联三层检测器,IoU 阈值设定分别为 0.5、0.62 和 0.72。

**表 4 不同 IoU 阈值增长方式对比**

Tab.4 Comparison of Different IoU Threshold Growth Ways

增长方式	IoU 阈值	mAP@75	mAP@50
线性增长	0.5,0.6,0.7	0.860	0.951
曲线增长	0.5,0.62,0.72	0.862	0.954

### 3 实验与结果分析

#### 3.1 模型优化对比

本文分别采用 ResNet101 和 ResNeXt101 两种网络作为裂缝检测与分割模型的特征提取网络进行训练,表 5 为两种特征提取网络下的 C-Mask R-CNN 模型的检测 AP 与分割 AP 对比结果,可以看出,使用 ResNeXt 特征提取网络,模型

的检测 mAP 从 0.778 提升至 0.887,分割 mAP 从 0.753 提升至 0.874;相比于 Mask R-CNN 结果,本文提出的 C-Mask R-CNN 模型表现效果更为优异。综上,本文采用 ResNeXt 作为裂缝的特征提取网络。

**表 5 两种特征提取网络下的模型对比**

Tab.5 Model Comparison Under Two Feature Extraction Networks

类别	ResNet101		ResNeXt101	
	检测	分割	检测	分割
横向裂缝 AP	0.851	0.829	0.915	0.903
纵向裂缝 AP	0.816	0.798	0.927	0.923
网状裂缝 AP	0.666	0.633	0.819	0.797
mAP@50	0.778	0.753	0.887	0.874

本文对模型优化器的选择也进行了实验,分别采用随机梯度下降(stochastic gradient descent, SGD)、SGD 加动量(SGD+Momentum)、自适应学习率优化器(AdaGrad、RMSProp 和 Adam)进行训练,训练 loss 曲线见图 10,可以看出,不同优化器下 loss 曲线的收敛速度不同,最终趋于稳定的 loss 值也不同,AdaGrad 的表现效果最差,Adam 和 RMSProp 走势基本相同,Adam 最终稳定的 loss 值最低,说明模型训练效果最好。表 6 为 5 种优化器下的模型训练结果,可以看出,单纯使用 SGD 训练出来的模型准确率最低,加入动量之后 AP 值有明显提升;3 种自适应学习率优化器中,AdaGrad 的表现结果最差,RMSProp 和 Adam 比较接近,但综合来说,Adam 算法效果更好,因此,本文选择 Adam 作为裂缝识别模型的优化器。

#### 3.2 模型训练

在模型训练中,本文采用动态调整的方式来设置学习率,将初始学习率设置为 0.0001,采用动态调整方式训练 20 个 epoch,将 batch\_size 设为 1,一共进行 17 500 次迭代。

在模型训练过程中,学习率变化曲线、训练准确度变化曲线以及 loss 曲线如图 11 所示。由图 11(a)可知,随着训练轮数的增加,学习率也在动态地不断降低;图 11(b)展示了训练过程中级联 3 层检测器的准确度的变化过程,其中, $s_0$ 、 $s_1$  和  $s_2$  分别代表第一、第二和第三阶段的检测器,随着 IoU 阈值的提升,对检测结果的要求会越高,三层检测器的准确度是依次下降的;图 11(c)展示了训练过程中 RPN 的分类损失 loss\_rpn\_cls 和回归损失 loss\_rpn\_bbox;图 11(d)展示了三层检测器  $s_0$ 、 $s_1$  和  $s_2$  的分类损失 cls、回归损失 bbox 和分割损

失 mask 的 loss 曲线;图 11(e)展示了 9 种 loss 相加得到的总 loss 曲线的变化过程。由这些 loss 曲线可以看出,loss 是逐渐下降并趋于平稳的。

### 3.3 实验结果与对比

本文采用 Pr-Re 曲线、AP 和 mAP 作为评价指标来衡量模型性能。图 12 展示了 C-Mask R-CNN 模型检测部分的 Pr-Re 曲线。图 13 展示了 C-Mask R-CNN 模型分割部分的 Pr-Re 曲线。

表 7 和表 8 分别为模型改进前后的检测部分与分割部分在几种指标评价下的对比结果,分析表 7 和表 8 可知,在 C50 指标下,对于检测部分,Mask R-CNN 模型 3 种裂缝类别的 mAP 为 0.857,C-Mask R-CNN 模型的 mAP 为 0.954,提升了 9.7%;对于分割部分,Mask R-CNN 模型 3 种裂缝类别的平均准确率 mAP 为 0.805,C-

Mask R-CNN 模型的 mAP 为 0.935,提升了 13.0%,分割效果有很大提升。

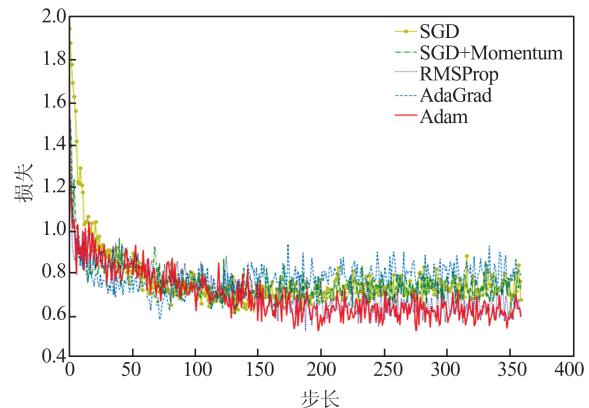


图 10 不同优化器下的 loss 曲线对比

Fig.10 Comparison of Loss Curves Under Different Optimizers

表 6 不同优化器下模型训练结果对比

Tab.6 Comparison of Model Training Results Under Different Optimizers

优化器	检测 AP				分割 AP			
	横向裂缝	纵向裂缝	网状裂缝	mAP	横向裂缝	纵向裂缝	网状裂缝	mAP
SGD	0.710	0.538	0.462	0.570	0.552	0.449	0.486	0.496
SGD+Momentum	0.915	0.927	0.819	0.887	0.903	0.923	0.797	0.874
AdaGrad	0.751	0.679	0.606	0.679	0.809	0.686	0.567	0.687
RMSProp	0.960	0.948	0.948	0.952	0.943	0.939	0.890	0.924
Adam	0.956	0.952	0.954	0.954	0.927	0.952	0.924	0.935

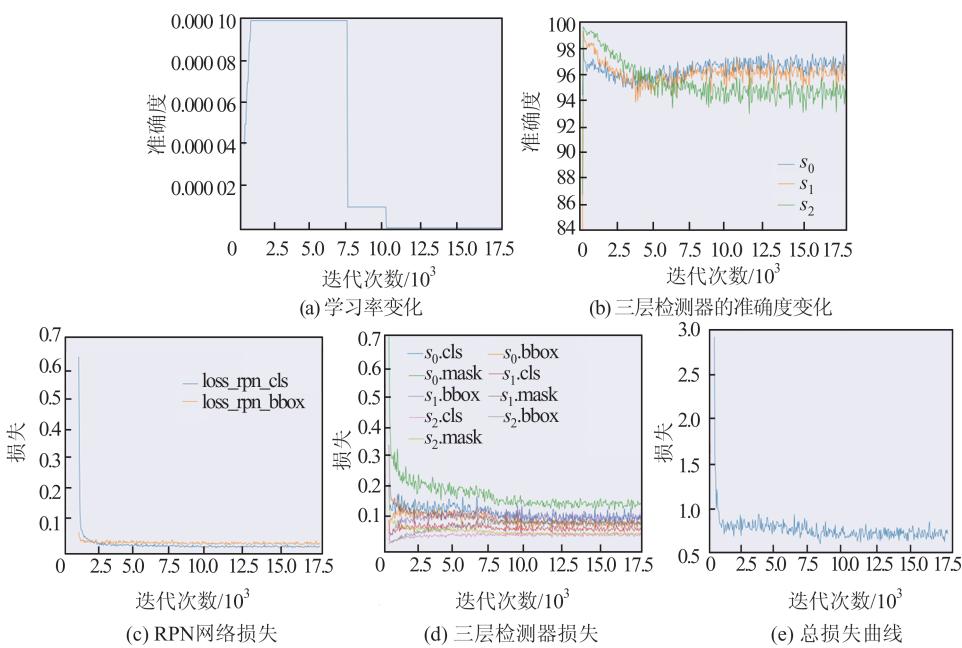


图 11 模型训练过程的变化曲线

Fig.11 Change Curves of Model Training Process

此外,当 IoU 阈值由 0.5 提高到 0.75 时,对于检测部分,Mask R-CNN 模型的 mAP 从 0.857 降

至 0.528,下降了 32.9%,C-Mask R-CNN 模型的 mAP 从 0.954 降至 0.862,只下降了 9.2%。同

理,对于分割部分,C-Mask R-CNN 模型的 mAP 值下降率比 Mask R-CNN 模型的 mAP 值下降率低 9.6%,这充分表明采用级联多阈值检测器的

方法可以有效解决在 IoU 阈值增大时出现的 AP 骤降现象,实现高 IoU 阈值下高质量的裂缝检测。

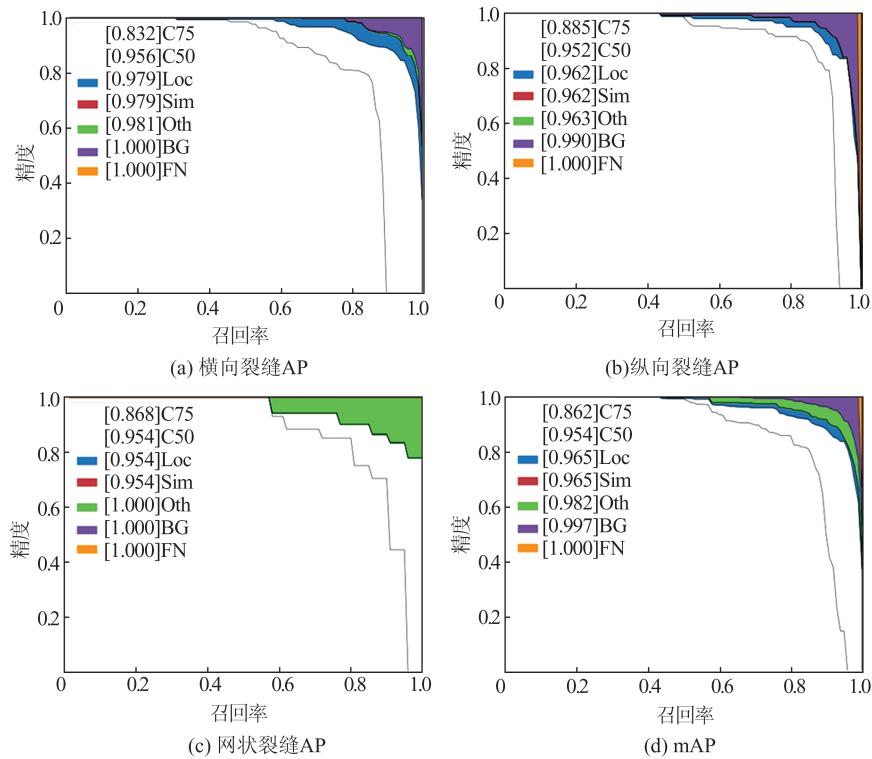


图 12 检测部分 Pr-Re 曲线  
Fig.12 Pr-Re Curves of Detection Part

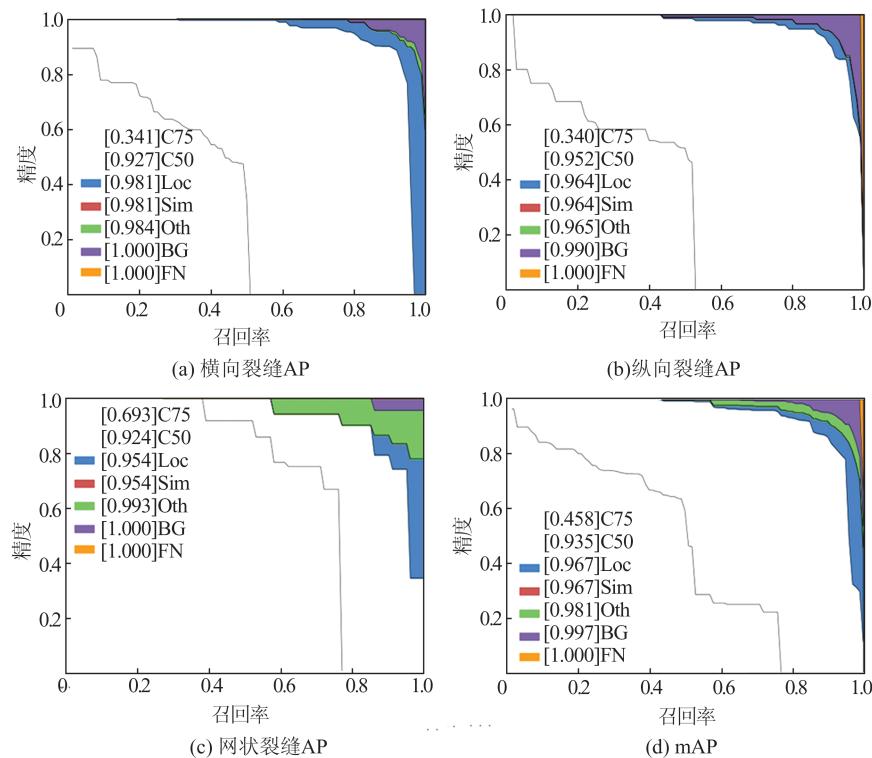


图 13 分割部分 Pr-Re 曲线  
Fig.13 Pr-Re Curves of Segmentation Part

表7 改进前后检测模型在各指标评价下的对比结果

Tab.7 Precision Comparison of Detection Models in Different Indexes Before and After Improvement

模型	类别	C75	C50	Loc	Sim	Oth	BG	FN
Mask R-CNN	横向裂缝 AP	0.506	0.906	0.964	0.964	0.969	1.00	1.00
	纵向裂缝 AP	0.630	0.909	0.945	0.945	0.948	0.990	1.00
	网状裂缝 AP	0.448	0.756	0.847	0.847	0.929	1.00	1.00
	mAP	0.528	0.857	0.919	0.919	0.949	0.997	1.00
C-Mask R-CNN	横向裂缝 AP	0.832	0.956	0.979	0.979	0.981	1.00	1.00
	纵向裂缝 AP	0.885	0.952	0.962	0.962	0.963	0.990	1.00
	网状裂缝 AP	0.868	0.954	0.964	0.964	1.00	1.00	1.00
	mAP	0.862	0.954	0.965	0.965	0.982	0.997	1.00

表8 改进前后分割模型在各指标评价下的对比结果

Tab.8 Precision Comparison of Segmentation Models in Different Indexes Before and After Improvement

模型	类别	C75	C50	Loc	Sim	Oth	BG	FN
Mask R-CNN	横向裂缝 AP	0.144	0.887	0.963	0.963	0.967	1.00	1.00
	纵向裂缝 AP	0.141	0.873	0.945	0.945	0.948	0.990	1.00
	网状裂缝 AP	0.411	0.654	0.834	0.834	0.885	1.00	1.00
	mAP	0.232	0.805	0.914	0.914	0.933	0.997	1.00
C-Mask R-CNN	横向裂缝 AP	0.341	0.927	0.981	0.981	0.984	1.00	1.00
	纵向裂缝 AP	0.340	0.952	0.964	0.964	0.965	0.990	1.00
	网状裂缝 AP	0.693	0.924	0.954	0.954	0.993	1.00	1.00
	mAP	0.458	0.935	0.967	0.967	0.981	0.997	1.00

图14为模型改进前后的裂缝识别对比图,可以看出,C-Mask R-CNN模型对裂缝的检测与分割更为细致,改善了使用Mask R-CNN时出现的裂缝边缘分割不准确、检测框框不全裂缝的问题(如图14(b)所示),以及一条裂缝被误识别为多条裂缝的现象(如图14(a)和图14(c)所示),在高质量的裂缝检测框内可以较好地对裂缝像素进行分割。

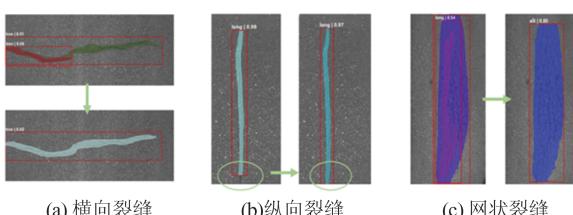


图14 模型改进前后裂缝识别对比图

Fig.14 Comparison of Crack Identification Effects Before and After Model Improvement

图15展示了使用C-Mask R-CNN模型对横向裂缝、纵向裂缝和网状裂缝识别的部分测试效果图,可以看出,该模型具有良好的裂缝识别能力。

图16展示了在当路面存在多种噪声干扰的

情况下,使用C-Mask R-CNN模型的识别效果,从图16中可以看出,当路面背景复杂,存在污渍、标志线、阴影干扰时,所提模型也能较为准确地识别裂缝。

表9为C-Mask R-CNN模型与目前几种主流模型在裂缝数据集上测试的准确度对照表,可以看出,C-Mask R-CNN模型的mAP值最高,这充分验证了其有效性<sup>[22]</sup>。

## 4 结语

为了解决现有裂缝识别算法中使用单一模型或组合模型后识别精度不高等问题,本文提出了一种基于Mask R-CNN模型的裂缝识别方法,在实现裂缝检测定位的同时,对检测框内的裂缝像素进行分割,并对模型中anchor的长宽比进行调整,以便更好地适应裂缝目标的检测。同时,针对Mask R-CNN模型中提高IoU阈值时出现的检测精度骤降问题,设计了一种C-Mask R-CNN多阈值检测方法,通过级联不同IoU阈值的检测器来提高生成候选框的质量,实现高阈值检测下的裂缝精准定位,同时通过对级联不同个数检测器和不同阈值增长方式来改善模型。

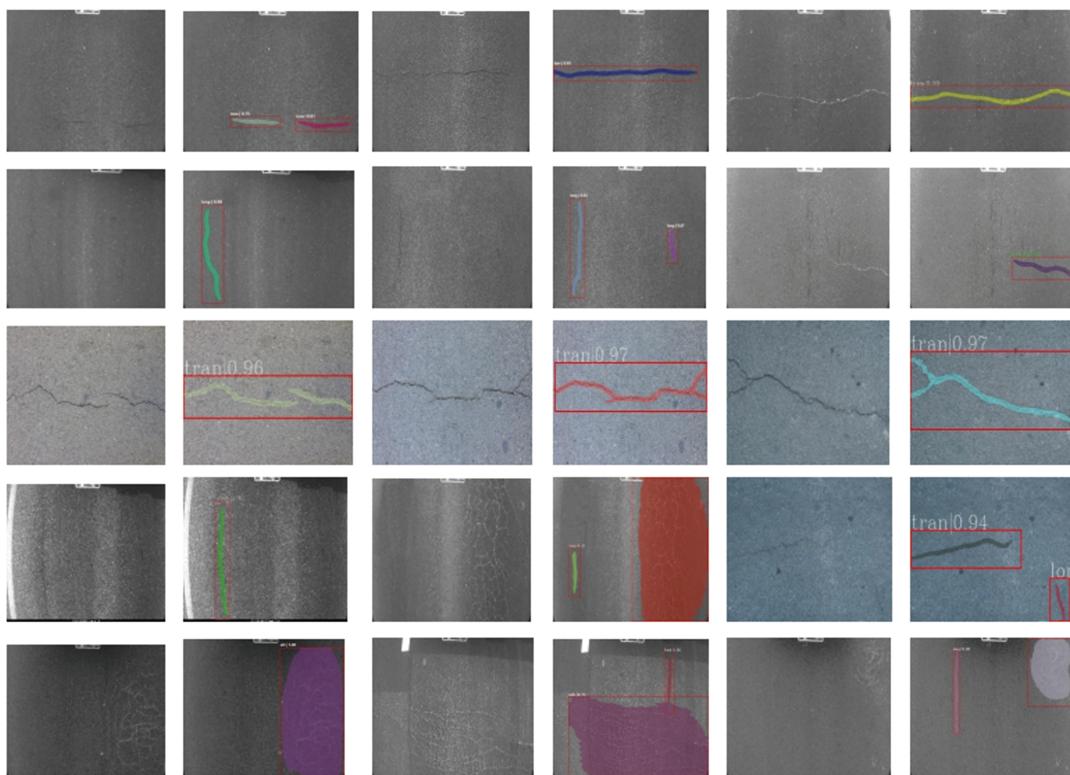


图 15 C-Mask R-CNN 模型测试裂缝识别效果图

Fig.15 Crack Identification Results Obtained by C-Mask R-CNN Model

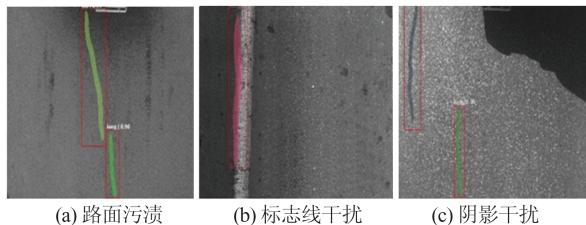


图 16 复杂路面背景下的裂缝识别效果

Fig.16 Crack Identification Effects Under Complex Backgrounds

表 9 不同模型在裂缝数据集上的对比

Tab.9 Comparison of Different Models on Crack Datasets

模型	主干网络	检测 mAP@50	分割 mAP@50
SSD	ResNet50	0.823	—
	MobileNet <sup>[10]</sup>	0.858	—
Faster R-CNN	ResNet101	0.865	—
	ZFNet <sup>[7]</sup>	0.897	—
U-Net	VGG16	—	0.752
	ResNet50	—	0.787
DeepCrack <sup>[8]</sup>	ResNet101	—	0.836
Mask R-CNN	ResNeXt101	0.857	0.805
C-Mask R-CNN	ResNeXt101	0.954	0.935

## 参 考 文 献

- [1] Zou Qin, Li Qingquan, Mao Qingzhou, et al. Tar-

get-Points MST for Pavement Crack Detection [J].

*Geomatics and Information Science of Wuhan University*, 2011, 36(1): 71-75. (邹勤, 李清泉, 毛庆洲, 等. 利用目标点最小生成树的路面裂缝检测 [J]. 武汉大学学报(信息科学版), 2011, 36(1): 71-75.)

- [2] Premachandra C, Waruna H, Premachandra H, et al. Image Based Automatic Road Surface Crack Detection for Achieving Smooth Driving on Deformed Roads [C]// IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 2014.
- [3] Oliveira H, Correia P L. CrackIT : An Image Processing Toolbox for Crack Detection and Characterization [C]// IEEE International Conference on Image Processing, Paris, France, 2015.
- [4] Ma Ronggui, Xu Kun, Liu Feifei. Highway Surface Crack Image Identifying Algorithm [J]. *Journal of Transport Information and Safety*, 2014, 32 (2) : 90-94. (马荣贵, 徐琨, 刘飞飞. 高速公路路面裂缝识别算法研究[J]. 交通信息与安全, 2014, 32 (2): 90-94.)
- [5] Yoo H S, Kim Y S. Development of a Crack Recognition Algorithm from Non-routed Pavement Images Using Artificial Neural Network and Binary Logistic Regression [J]. *KSCE Journal of Civil Engineering*, 2016, 20(4): 1151-1162.
- [6] Zhao Shanshan, He Ning. Road Surface Crack De-

- tection Based on CNN [J]. *Transducer and Microsystem Technologies*, 2017, 36(11): 135–138. (赵珊珊, 何宁. 基于卷积神经网络的路面裂缝检测[J]. 传感器与微系统, 2017, 36(11): 135–138.)
- [7] Wang Liping. Deep Learning-Based Crack Detection of Concrete Pavement [D]. Handan: Hebei University of Engineering, 2018. (王丽萍. 基于深度学习的混凝土路面裂缝检测[D]. 邯郸: 河北工程大学, 2018.)
- [8] Zou Q, Zhang Z, Li Q Q, et al. DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection [J]. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 2018, DOI: 10.1109/TIP.2018.2878966.
- [9] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder–Decoder Architecture for Image Segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495.
- [10] Maeda H, Sekimoto Y, Seto T, et al. Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2018, 33(12): 1127–1141.
- [11] Hou Yue, Chen Yihan, Gu Xingyu, et al. Automatic Identification of Pavement Objects and Cracks Using the Convolutional Auto-Encoder [J]. *China Journal of Highway and Transport*, 2020, 33(10): 288–303. (侯越, 陈逸涵, 顾兴宇, 等. 基于卷积自编码的沥青路面目标与裂缝智能识别[J]. 中国公路学报, 2020, 33(10): 288–303.)
- [12] Alipour M, Harris D K, Miller G R. Robust Pixel-Level Crack Detection Using Deep Fully Convolutional Neural Networks [J]. *Journal of Computing in Civil Engineering*, 2019, 33(6): 4019040.
- [13] Nguyen H T, Yu G H, Na S Y, et al. Pavement Crack Detection and Segmentation Based on Deep Neural Network [J]. *The Journal of Korean Institute of Information Technology*, 2019, 17 (9) : 99–112.
- [14] 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*, 2016, 39(6): 1137–1149.
- [15] Girshick R. Fast R-CNN [C]// IEEE International Conference on Computer Vision, Santiago, Chile, 2016.
- [16] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]// IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014.
- [17] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]// IEEE International Conference on Computer Vision, Venice, Italy, 2017.
- [18] Xiao Liyang, Li Wei, Gao Rong, et al. X-Ray Detection System for Foreign Bodies in Sole of Shoes Based on Deep Learning [J]. *Computer Systems & Applications*, 2021, 30(3) : 88–94. (肖力炀, 李伟, 高荣, 等. 基于深度学习的X射线鞋底异物检测系统[J]. 计算机系统应用, 2021, 30(3) : 88–94.)
- [19] Sun Jiachi, Zou Huanxin, Deng Zhipeng, et al. Oriented Inshore Ship Detection and Classification Based on Cascade RCNN [J]. *Systems Engineering and Electronics*, 2020, 42(9) : 1903–1910. (孙嘉赤, 邹焕新, 邓志鹏, 等. 基于级联卷积神经网络的港口多方向舰船检测与分类[J]. 系统工程与电子技术, 2020, 42(9) : 1903–1910.)
- [20] Cai Z W, Vasconcelos N. Cascade R-CNN: Delving into High Quality Object Detection [C]// IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018.
- [21] Shen Yuan. Research on Object Detection Algorithm of Remote Sensing Images Based on Deep Learning [D]. Shenzhen: Chinese Academy of Sciences, 2020. (申原. 基于深度学习的遥感影像目标检测算法研究[D]. 深圳: 中国科学院大学, 2020.)
- [22] Li Xiaodong, Chen Yapeng, Wang Baoying, et al. Application of Machine Learning in Condition Analysis of Power Plant Equipment [J]. *Thermal Power Generation*, 2020, 49(1): 129–133. (李晓东, 陈亚鹏, 王保营, 等. 机器学习在电站设备状态分析中的应用[J]. 热力发电, 2020, 49(1): 129–133.)