

文章编号: 1005-0523(2024)02-0056-08



基于注意力机制的YOLOv5路面裂缝检测与识别

周双喜^{1,2}, 杨丹², 潘远², 丁建新¹, 丁杨³

(1. 广州航海学院土木与工程管理学院, 广东广州 510765; 2. 华东交通大学土木与建筑学院, 江西 南昌 330013;
3. 浙大城市学院土木工程学院, 浙江 杭州 310015)

摘要:【目的】针对传统路面裂缝检测实时性差、准确率低的问题。【方法】利用深度学习网络在目标检测方面的优势, 提出一种改进的YOLOv5算法, 称为YOLOv5s-attention, 以实现路面裂缝自动化检测与识别。首先, 对采集到的裂缝图片用LabelImg标注软件进行手工标记, 然后通过改进YOLOv5网络训练得到网络模型参数。最后, 利用所建立的模型对裂缝进行验证和预测。除此之外, 采用综合评价指标(F_1 -measure, F_1)和平均精度均值(mean average precision, mAP)这两个指标来比较原YOLOv5s、YOLOv5s-attention模型在路面裂缝上检测与识别的性能。【结果】经YOLOv5s与YOLOv5s-attention比较发现, YOLOv5s-attention检测准确率(Precision)提高1.0%, F_1 提高0.9%, mAP 提高了1.8%。【结论】由此可知, 该网络在实现道路裂缝自动化识别上具有一定的现实意义。

关键词: 道路养护; 路面裂缝; 目标检测; YOLO; 注意力机制; 图像处理

中图分类号: TU528; U416

文献标志码: A

本文引用格式: 周双喜, 杨丹, 潘远, 等. 基于注意力机制的YOLOv5路面裂缝检测与识别[J]. 华东交通大学学报, 2024, 41(2): 56-63.

Detection and Recognition of YOLOv5 Pavement Cracks Based on Attention Mechanism

Zhou Shuangxi^{1,2}, Yang Dan², Pan Yuan², Li Jianxing¹, Ding Yang³

(1. School of Civil Engineering and Management, Guangzhou Maritime University, Guangzhou 510765, China; 2. School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330013, China; 3. School of Civil Engineering, Hangzhou City University, Hangzhou 310015, China)

Abstract: 【Objective】Aiming at the problem of poor real-time performance and low precision of traditional pavement crack detection. 【Method】This paper uses the advantages of deep learning network in target detection, and proposes an improved YOLOv5 algorithm, which is called YOLOv5s-attention in this paper, to realize the automatic detection and recognition of pavement cracks. Firstly, the collected crack images are manually labeled with LabelImg annotation software, and then the network model parameters were obtained by improving the YOLOv5 network training. Finally, the model is used to verify and predict the cracks. In addition, F_1 and mAP are used to compare the performance of the original YOLOv5s and YOLOv5s-attention models in detecting and identifying pavement cracks. 【Result】The comparison between YOLOv5s and YOLOv5s-attention showed that the precision of YOLOv5s attention increased by 1.0%, F_1 increased by 0.9%, and mAP increased by 1.8%. 【Conclusion】It can

收稿日期: 2024-05-16

基金项目: 国家自然科学基金项目(51968022); 江西省主要学科学术和技术带头人培养计划(20213BCJL22039)

be seen that the network has certain practical significance in realizing the automatic recognition of road cracks.

Key words: maintenance; pavement cracks; object detection; YOLO; attention mechanism; image processing

Citation format: ZHOU S X, YANG D, PAN Y, et al. Detection and recognition of YOLOv5 pavement cracks based on attention mechanism[J]. Journal of East China Jiaotong University, 2024, 41(2): 56–63.

【研究意义】随着计算机技术的迅速发展,机器学习(machine learning, ML)已经成为当今的热门,许多领域都与其相结合研究,特别是最近几年也多应用于道路裂缝检测上,产生了许多基于机器视觉自动化检测算法。根据最近流行的机器学习技术,基于深度学习结合神经网络,本研究提出基于注意力机制的YOLOv5路面裂缝检测与识别算法,该研究具有提高道路维护效率、降低交通事故风险、推动智慧交通发展,并促进深度学习算法的发展等重要的研究意义。

【研究进展】基于深度学习的目标检测分为两大类,第一类是单阶段目标检测,不使用区域候选框进行筛选,而是直接通过主干提取网络提取出目标物体的类别和位置信息,代表的网络有仅看一次(you only look once, YOLO)系列和单步多框目标检测(single shot multibox detector, SSD)等;第二类是两阶段目标检测,第一步是通过卷积神经网络提取物体区域候选框,第二步是对区域候选框进行卷积神经网络(convolutional neural network, CNN)分类识别,代表网络有RCNN系列和SPPNet等^[1-3]。

Faster-RCNN目前是应用在路面裂缝中最多的一个两阶段目标检测法,徐康等^[4]对沥青路面因复杂背景难以检测提出了用VGG16、Mobile Net-V2和ResNet50替换主干特征提取网络,并在添加进的ResNet50中引入了注意力机制,实验表明,改进后的Faster-RCNN裂缝检测精准度达到85.64%,且能识别出沥青路面的细微裂缝。Hao等^[5]以机场跑道裂缝为研究背景,采用Faster-RCNN网络检测跑道的裂缝,并进入MobileNet2替换Faster-RCNN网络的主干网络,以及在里面加入注意力机制,实验表明,改进后的Faster-RCNN网络准确率提高了6.4%。此类算法精度虽然达到80%以上,但训练速度小于单步目标检测算法^[6]。

YOLO系列是最流行的单级检测方法之一,在一阶检测模型中实现了快速检测速度和高检测精度之间的良好平衡,并且已经广泛应用到农业、地质、遥

感和医学等领域。此外,它还广泛应用于交通运输,如交通标志、车流、路面坑洼、视觉裂缝等检测^[7-10]。

【创新特色】本文研究基于深度学习建立卷积神经网络来实现路面裂缝自动化检测与识别。**【关键问题】**为了解决提高裂缝检测速度的问题,本文选择YOLOv5s算法应用到路面裂缝检测与识别上。

1 初始准备工作

1.1 数据集处理

本次研究目的是能利用最新的目标检测算法来自动识别道路裂缝,因此需要采集比较多的裂缝。本文的裂缝图片取自公开文献中提供的裂缝数据^[11-12],制作成一个新的数据集,新的数据集收集了3 022个样本,将自建的数据集以7:2:1的比例随机分为训练集、验证集和测试集,得训练集2 115张、测试集605张、验证集302张。像素密度分别为300 dpi和96 dpi,数据集中图片的尺寸不一,但由于YOLO网络中输入照片尺寸一般设置为32的倍数。因此,在搭建网络时通过程序将裂缝图片尺寸自动修改成640×640像素。在输入裂缝图片之前,我们将数据集中所有的照片用标注软件LabelImg进行标注,类别为HCrack(横向裂缝)、30°Crack、45°Crack、LCrack(纵向裂缝)、HoneyCrack(龟裂裂缝),标注类型保存为txt格式。标注步骤为先建立数据集存放的文件夹,一共在文件夹datasets下建立两个子文件夹,分别为images、labels,且images和labels文件夹中又包括train、test、val三个文件夹。其中images用于存放需要使用的裂缝图片,labels中存放对应的标注文件和索引文件classes。然后在Pytorch-GPU环境下打开LabelImg标注软件,对裂缝照片画矩形框进行标注。

1.2 设备与软件准备

用于搭建基于YOLOv5算法的深度学习道路裂缝自定义识别模型的硬件CPU为英特尔Intel(R) Xeon(R) Platinum 8260,内存为128 G,显卡(GPU)为NVIDIA GeForce RTX 2080 Ti,内存为75 G。

2 YOLOv5算法介绍

2.1 目标检测算法

目标检测模型主要分为两大类：

1) 第一类是One-stage,是直接基于anchors进行分类和同时调整边界框。主要的网络有SSD、YOLO系列。

2) 第二类是Two-stage,首先通过专门的模块生成区域候选框(region proposal network, RPN),寻找前景(Foreground)以及调整基于anchors的边界框,然后基于之前生成的区域候选框RPN进一步分类以及调整基于Proposals的边界框。这一类主要的网络有Faster-RCNN。

这两类网络都有自己各自的优点,One-stage检测速度快,Two-stage检测更准确。本文是研究关于道路裂缝的自动识别,在前人研究了许多神经网络模型的情况下,对于道路识别精度已经达到比较好的一个精度了,因此这次为了使得检测速度更加快一点,选择将YOLO算法与道路裂缝结合在一起进行研究。

2.2 YOLOv5网络算法

YOLOv5算法有4个组成部分,分别是:输入(Input),主干特征提取网络(Backbone),特征金字塔网络(feature pyramid networks, FPN)和检测头(YOLOHead)。

YOLOHead是YOLOv5的分类器与回归器。详细信息如图1所示。

与广泛应用的YOLOv3相比,YOLOv5在以下几个部位进行了改进：

1) 数据增强,其中包括Mosaic数据增强和Mosaic利用了4张图片进行拼接实现数据增强。它最大的优点是可以丰富检测物体的背景,且当进行标准归一化BN计算的时候可以直接计算出四张图片的数据。

2) 多正样本匹配。在之前的YOLO系列里面,在训练时每一个真实框对应一个正样本,即在训练时,每一个真实框仅由一个先验框负责预测。YOLOv5中为了加快模型的训练效率,增加了正样本的数量,在训练时,每一个真实框可以由多个先验框负责预测。

2.3 注意力机制

注意力机制是基于人类视觉研究出来的机器视觉,主要有Soft attention和Hard attention两类。选择与通道、空间有关的Soft attention中的SENet(squeeze-and-excitation networks)和CBAM。

SENet是通过Squeeze和Excitation两个模块实现对特征通道加入注意力机制,先用Squeeze对每个特征图进行全局池化,平均成一个实数值,然后通过Excitation操作,此时网络输出了 $1 \times 1 \times C$ 大小的特征图,此Excitation步骤提升了重要特征,抑制了不重要特征。如图2。图2中的 H 表示输入图片的高度, W 表示输入图片的宽度, C 表示通道数, X 表示输入值, \bar{X} 随之研究者又提出与SENet注意力机制类似的通过学习的方式自获取每个特征通道的重要程度的轻量的注意力模块CBAM,且CBAM还具有通过类似的学习方式自获取每个特征空间的重要程度。CBAM的通道注意原理和空间注意原理如图3和图4。

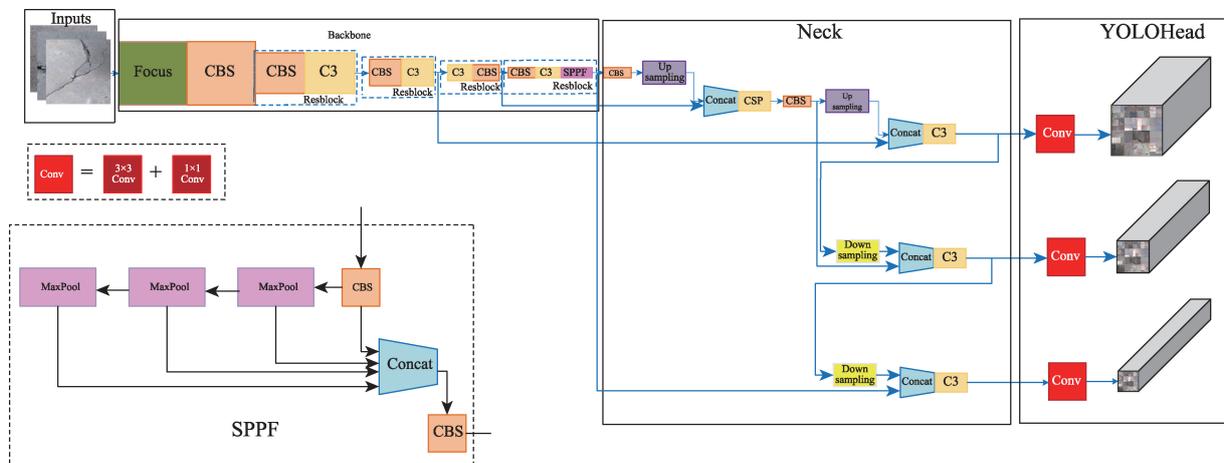


图1 YOLOv5网络流程
Fig. 1 YOLOv5 network process

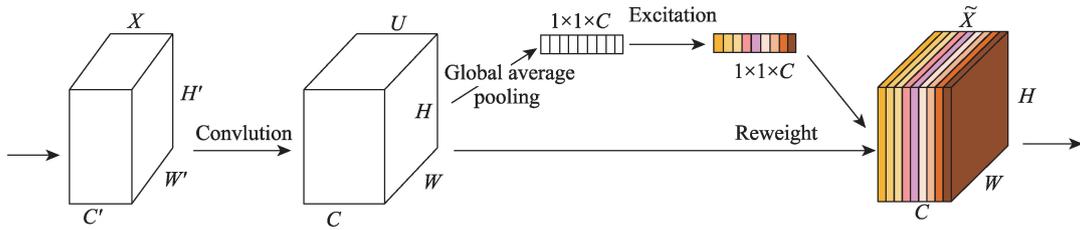


图2 SENet注意力模块
Fig. 2 SENet attention module

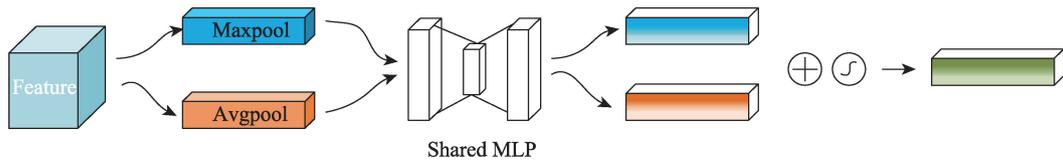


图3 通道注意力模块
Fig. 3 Channel attention module

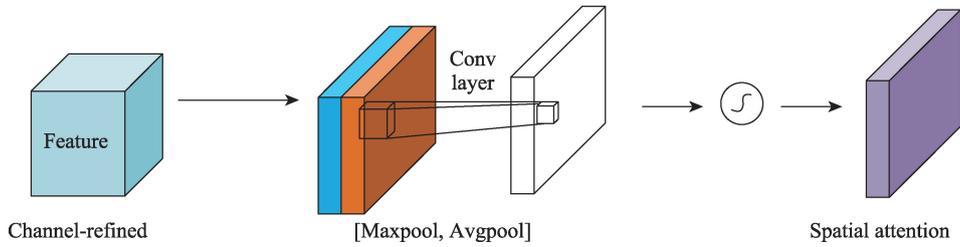


图4 空间注意力模块
Fig. 4 Spatial attention module

2.4 改进主干特征提取网络

CSPLayer-Darknet53是YOLOv5s的主干特征提取网络,其中使用到的CSPLayer模块记作C3。C3中使用了3个Conv_BN_SiLU模块,首先图片输入进主干网络,经过Focus层,Focus层是在裂缝图片上每隔一个像素值取一个特征值,且原本的裂缝照片为4×4×3的像素组成,现经过Focus层后,获得4个独立的特征层,并将这4个独立的特征层进行堆叠,此时关于高和宽的信息就集中到了通道上,使得通道数扩充了4倍,拼接起来的特征层从原来的3通道变为12通道,也就是图片信息从原始的高宽集中到通道上,然后输入C3模块,在此模块中,首先通过两条通道进行卷积、标准归一化和激活函数处理,其中一条通道还会经过bottleneck模块,然后将提取到的特征堆叠,最后将堆叠后的结果再一次卷积、标准归一化和激活函数处理,结合C3模块处理流程,可以在Bottle-Neck模块中添加注意力机制。图5是原C3模块和改进后的C3CBAM模块,图6是改进YOLOv5s-attention主干结构。

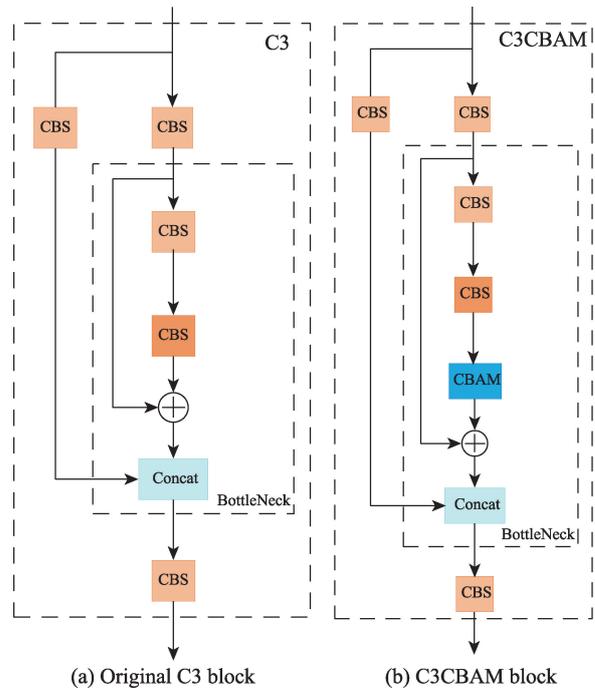


图5 原C3模块和改进后的C3CBAM模块
Fig. 5 Original C3 block and C3CBAM block

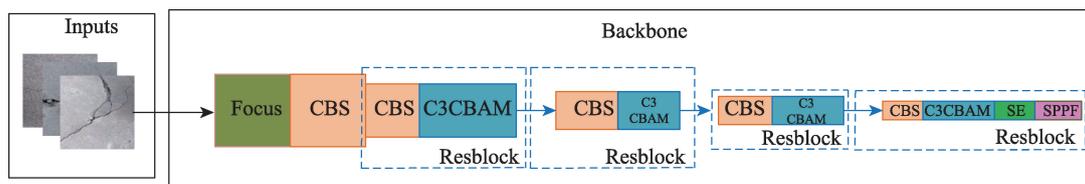


图6 YOLOv5s-attention 主干结构
Fig. 6 YOLOv5s-attention backbone

2.5 基于改进YOLOv5s的路面裂缝检测

网络学习的具体操作为:首先将尺寸为 640×640 的裂缝图片输入到主干结构C3CBAMDarknet中;然后在主干结构中提取到的特征传入Neck结构中进行特征融合;再次提取传入的裂缝特征;最后传入YOLOHead中进行预测。图7是改进后的YOLOv5s-attention网络流程图。

3 训练结果和评价

3.1 模型训练

首先将处理好的裂缝数据集放置在根目录中,此时裂缝数据集的形式是txt格式。运行train.py,开始网络训练,训练的数据集就是处理好之后的裂缝图片。训练得到的权重会自动放在run文件夹内。对YOLOv3、原YOLOv5s、YOLOv5s-attention以及其他改进的YOLOv5s模型都不加YOLOv5s的预权重分别训练300 epoch,然后得到相对应的值。

3.2 评价指标

综合评价指标 F_1 是用来平衡准确率和召回率的调和平均值。准确率是检测出一类特征的数量与检测出所有特征数量之间的比率,是用来评价此

模型是否能准确检测出目标,即,查准率;召回率是检测出一类特征的数量和数据集中所有的该类特征数量的比率,衡量的是模型检测系统的查全率。准确率 P 和召回率 R 、综合评价指标 F_1 的计算式如下

$$P = \frac{T_p}{T_p + F_p} \quad (1)$$

$$R = \frac{T_p}{T_p + F_N} \quad (2)$$

式中: T_p 是网络模型预测是裂缝的正样本; F_p 是网络模型预测为裂缝的负样本; F_N 是网络模型预测不是裂缝的正样本,即误检。

$$F = \frac{2 \times P \times R}{\alpha^2(P+R)} \quad (3)$$

式中: α 为权重因子。且当 $\alpha=1$ 时,表示准确率和召回率的权重相同,即, F 为 F_1 ,如下

$$F_1 = \frac{2 \times P \times R}{P+R} \quad (4)$$

一般情况下, F_1 较高时说明模型更有效。

平均精度均值 mAP 是检测到所有目标类别平均精度(average precision, AP)的均值,平均精度 AP 是用来衡量某一类检测效果的好坏,在某一类裂缝检测中即为识别精度,而平均精度均值 mAP 是衡量

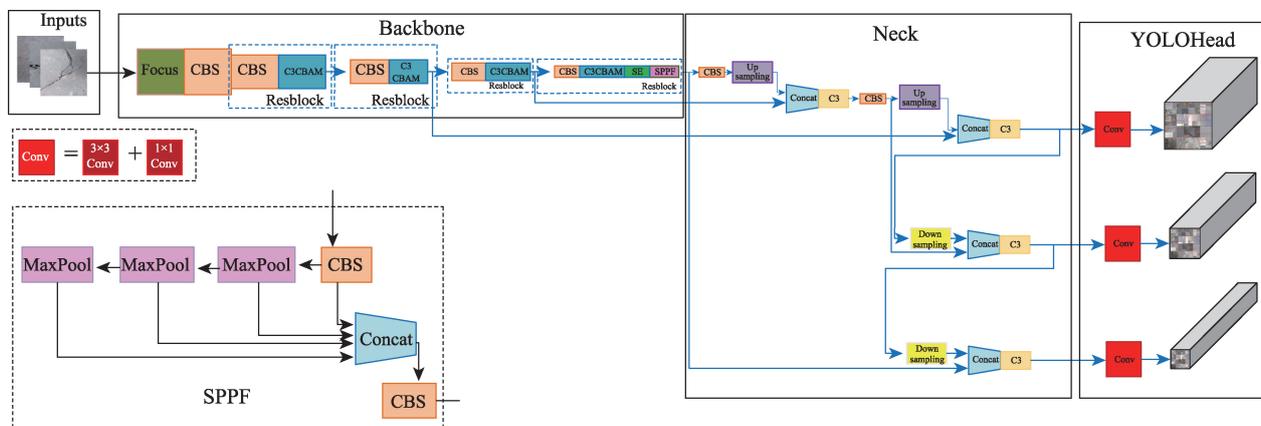


图7 YOLOv5s-attention 网络流程
Fig. 7 YOLOv5s-attention network process

多个类别的检测效果,即所有裂缝识别效果。

$$AP = \sum_{i=0}^{n-1} (r_{i+1} - r_i) P_i(r_i + 1) \quad (5)$$

其中: r_1, r_2, \dots, r_n 是按升序排列的 Precision 插值段第一个插值处对应的 Recall 值; P_i 表示第 i 个检测到的精度。

$$mAP = \frac{\sum AP}{N(\text{classes})} \quad (6)$$

式中: $N(\text{classes})$ 表示类的种类个数。

3.3 训练结果评估

3.3.1 综合评价指标 F_1

用 F_1 评估 YOLOv3、YOLOv5s、YOLOv5s-attention 以及其他改进的 YOLOv5s 模型是否能准确识别出这 5 类裂缝。取 $\text{batch_size}=1$, 计算出 F_1 值, 见表 1。

由表 1 可以看出, YOLOv5s 的准确率为 55.3%, 召回率为 59.0%, YOLOv5s-C3CBAM-SimSPPF 的准确率为 58.0%, 召回率为 57%, YOLOv5s-attention 的准确率为 56.4%, 召回率为 59.8%。由此可知, 经过改良的 YOLOv5s-C3CBAM-SimSPPF 模型精确度相对较高, 但是召回率比 YOLOv5s 模型低, 也就是在正确的检测出的结果中召回正确的个数少, 使得 F_1 值小于 YOLO 模型, 在 YOLOv5s-C3CBAM-SimSPPF 模型引入了 YOLOv6 中的 SimSPPF 模块, 既不增加计算参数, 还减少了运算速度, 但是经训练表现出召回率不太高, 导致检错率上升, 因此去掉了 SimSPPF 模块, 将 SPPF 层前添加对特征通道自学习的注意力模块 SENet, 从而得到 YOLOv5s-attention 模型, 经验证可知, 此模型精确度比原 YOLOv5s 模型提高了 1.0%, 召回率提高了 0.7%。

3.3.2 平均精度均值 mAP

用 mAP 衡量检测这 5 类裂缝的整体检测效果, 以用于评估 YOLOv5 系列模块的性能。下表是各个模型的 mAP 值。

从表 2 可知, YOLOv5s-SE 模型是在 YOLOv5s 的主干特征提取网络第 9 层添加了一个特征通道自学习的注意力机制 SENet, mAP 达到 56.9%, 但是此模型的召回率较低, 导致裂缝的检错率较高, 而 YOLOv5s-attention 模型各方面都表现得比 YOLOv5s 好, mAP 提高了 1.8%。

3.4 图片预测结果分析

图 8 分别是 YOLOv5s, YOLOv5s-C3CBAM-SimSPPF, YOLOv5s-attention 三个模型在自学习道

表 1 YOLO 网络 P, R, F_1 比较

Tab. 1 Comparison of YOLO networks $P, R,$ and F_1 %

Module	P	R	F_1
YOLOv3	43.0	50.7	46.5
YOLOv5s	55.3	59.0	57.1
YOLOv5s-SE(1)	56.9	57.1	57.0
YOLOv5s-SE(4)	57.3	54.7	56.0
YOLOv5s-C3CBAM-SimSPPF	58.0	57.0	57.5
YOLOv5s-attention(Ours)	56.4	59.8	58.1

表 2 YOLO 网络的 mAP

Tab. 2 The mAP of YOLO

Module	$mAP/\%$
YOLOv3	30.7
YOLOv5s	53.7
YOLOv5s-SE(1)	56.9
YOLOv5s-SE(4)	54.1
YOLOv5s-C3CBAM-SimSPPF	53.9
YOLOv5s-attention(Ours)	55.7

路裂缝特征后裂缝预测结果。图 8(a) 是使用 YOLOv5s 网络来识别的裂缝, 此算法可以识别出龟裂裂缝、横向裂缝和 45° 裂缝, 然而出现漏检裂缝的现象; 图 8(b) 是 YOLOv5s-C3CBAM-SimSPPF 网络识别裂缝结果, 从图中可知虽然能够识别出三种裂缝, 然而对横向裂缝识别的检测框不唯一, 也就是不能识别出此条裂缝的类别, 且仍存在漏检的裂缝; 图 8(c) 是使用 YOLOv5s-attention 网络识别裂缝的结果, 可识别出龟裂裂缝、横向裂缝、 45° 裂缝和 30° 裂缝, 能别出其他两个网络未识别出的 30° 裂缝, 且标签框能够准确识别出裂缝的类别, 漏检的裂缝能够识别到的原因是在 YOLOv5s-attention 网络主干结构中添加了关于通道的注意力机制 SENet, 以及在 C3 模块中添加了 CBAM 模块, 两者同时在通道和空间上作用, 其召回率相对于 YOLOv5s 和 YOLOv5s-C3CBAM-SimSPPF 网络分别提高了 0.7% 和 2.4%, 因此能够识别出漏检的 30° 裂缝, 另外 YOLOv5s 和 YOLOv5s-attention 网络使用了 K-means 聚类算法, 使得标签框能够准确识别出裂缝的类别。

由图 8(d)~图 8(f) 可以看出, 原始 YOLOv5s 网络和改进后的 YOLOv5s-C3CBAM-SimSPPF 网络存在误检的现象, 且不能检测到小裂缝目标, 而 YOLOv5s-attention 网络即没有把划痕误以为是裂缝, 也识别到图片中右下角的裂缝小目标。

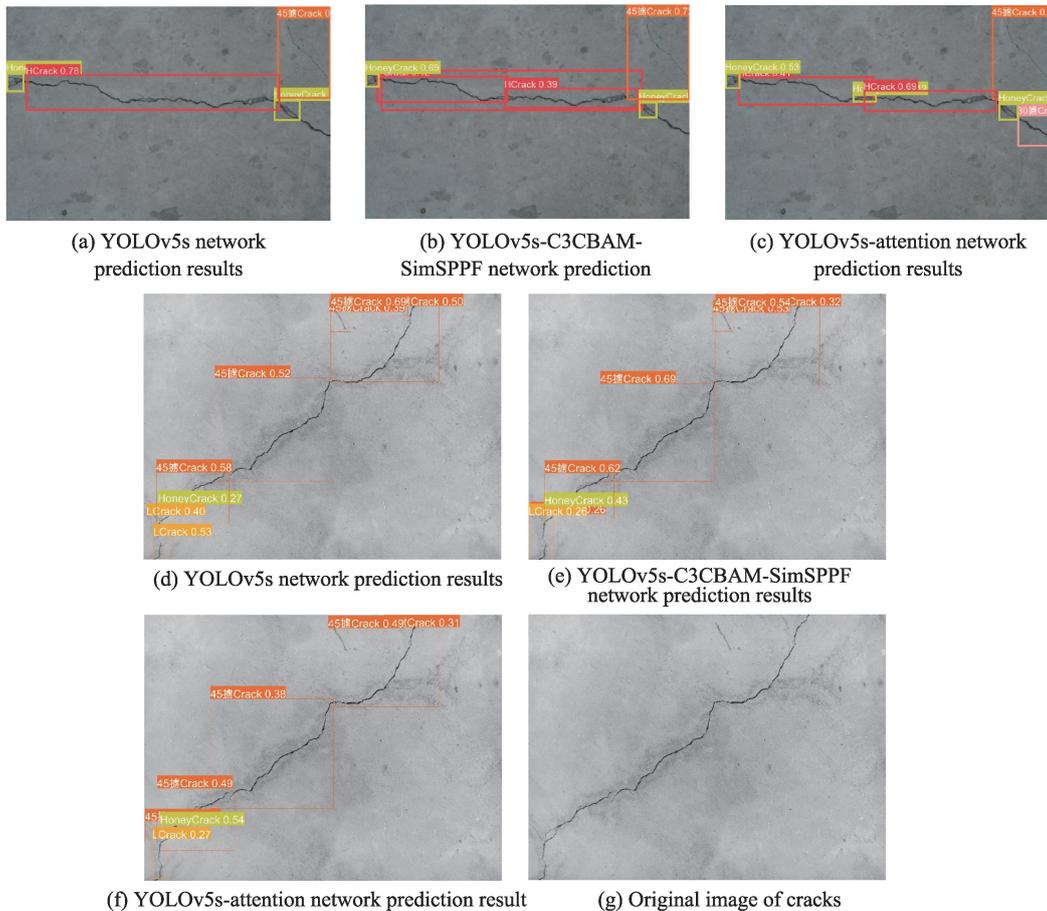


图8 裂缝预测

Fig. 8 Crack prediction results

4 结论

1) 与原 YOLOv5s 训练时长相比,添加注意力机制后对训练时长影响不大,平均 3 h 训练 300 epoch。

2) 新的 YOLOv5s-attention 网络对识别 30° 裂缝的准确率提高,且可以预测出原 YOLOv5s 网络未识别到的 30° 裂缝。

3) 新的 YOLOv5s-attention 网络在识别裂缝能力上与原 YOLOv5s 网络相比, F_1 上升 0.9%, mAP 提高 1.8%。

4) 进一步改良 YOLOv5 网络,使得改良后的 YOLOv5 模型更加占用的内存更小,即更加轻量化,检测速度更快,并且可实现安装在小型移动设备上,进行实时检测路面裂缝,随有随补,避免道路持续开裂。

参考文献:

[1] MARQUES A, CORREIA P L. Automatic road pavement

crack detection using SVM[J]. Electrical and Computer Engineering, 2012, 21(4): 157-170.

[2] YAN K, ZHANG Z. Automated asphalt highway pavement crack detection based on deformable single shot multi-box detector under a complex environment[J]. IEEE Access, 2021, 9: 150925-150938.

[3] 李鹏程, 孙立双, 谢志伟, 等. 基于改进 MobileNet-SSD 的路面裂缝图像检测算法[J]. 激光杂志, 2022, 43(7): 123-127.

LI P C, SU L S, XIE Z W, et al. Improved pavement crack image detection algorithm based on MobileNet-SSD [J]. Laser Journal, 2022, 43(7): 123-127.

[4] 徐康, 马荣贵. 基于改进 Faster-RCNN 的沥青路面裂缝检测[J]. 计算机系统应用, 2022, 31(7): 341-348.

XU K, MA R G. Crack detection of asphalt pavement based on improved Faster-RCNN[J]. Computer Systems & Applications, 2022, 31(7): 341-348.

[5] HAO S, SHAO L, WANG S. A Faster-RCNN airport pavement crack detection method based on attention

- mechanism[J]. Academic Journal of Science and Technology, 4(2):129-132.
- [6] HU N, YANG J, JIN X, et al. Few-shot crack detection based on image processing and improved YOLOv5[J]. Journal of Civil Structural Health Monitoring, 2023, 13(1): 165-180.
- [7] GUO G, ZHANG Z. Road damage detection algorithm for improved YOLOv5[J]. Scientific Reports, 2022, 12(1): 1-12.
- [8] 周清松,董绍江,罗家元,等.改进YOLOv3的桥梁表观病害检测识别[J].重庆大学学报,2022,45(6):121-130.
ZHOU Q S, DONG S J, LUO J Y, et al. Bridge apparent disease detection based on improved YOLOv3[J]. Journal of Chongqing University, 2022, 45(6): 121-130.
- [9] ZHANG J, QIAN S, TAN C. Automated bridge crack detection method based on lightweight vision models [J]. Complex & Intelligent Systems, 2023, 9(2): 1639-1652.
- [10] HU G X, HU B L, YANG Z, et al. Pavement crack detection method based on deep learning models[J]. Wireless Communications and Mobile Computing, 2021, 220: 1-13.
- [11] LIU Y H, YAO J, LU X H, et al. Deep crack:A deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing, 2019, 338(21): 139-153.
- [12] YU G, ZHOU X. An improved YOLOv5 crack detection method combined with a bottleneck transformer[J]. Mathematics, 2023, 11(10): 2377.
- [12] LI L F, MA W F, LI L, et al. Research on detection algorithm for bridge cracks based on deep learning[J]. ACTA Automatica Sinica, 2019, 45(9): 1727-1742.



第一作者:周双喜(1973—),男,教授,博士,博士生导师,南昌市521学术技术带头人,研究方向为混凝土材料与检测技术。E-mail:green.55@163.com。



通信作者:杨丹(1997—),女,硕士研究生,研究方向为混凝土材料与检测技术。E-mail:gryangdan@163.com。

(责任编辑:姜红贵)