

DOI:10.19651/j.cnki.emt.2313039

改进 YOLOv5 的路面裂缝检测模型研究^{*}

沈思远 华 蓓 黄汝维

(广西大学计算机与电子信息学院 南宁 530000)

摘要: 针对传统的路面裂缝检测方式耗时耗力、成本高、主观性强等问题,提出了一种基于 YOLOv5 的路面裂缝检测模型 YOLOv5-Crack。首先在主干部分处引入坐标注意力机制并优化成 CA-plus 结构以提高裂缝特征关注度;其次提出一种全新的特征融合网络 ESPP,降低部分计算量的同时提升特征融合能力;然后,在颈部网络中使用重影混洗卷积替代传统卷积,尽可能保持通道语义信息的同时降低计算成本;最后,整体引入 SIoU 损失函数提升回归精度。为验证改进模型 YOLOv5-Crack 的有效性,在数据集 GRDDC 2020 上进行对比实验,结果表明其 F1 分数分别为 58.43% 和 58.21%,与原 YOLOv5 模型相比分别提升了 4.05% 和 3.93%,并且降低了 7.8 GFLOPs 的计算量,FPS 达到 37.9,有效解决了路面裂缝检测的弊端;同时与主流目标检测算法相比,所提出的 YOLOv5-Crack 模型在路面裂缝检测方面更具有优越性。

关键词: 路面裂缝;坐标注意力机制;ESPP 结构;重影混洗卷积;SIoU 损失函数

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Research on improved YOLOv5 pavement crack detection model

Shen Siyuan Hua Bei Huang Ruwei

(School of Computer and Electronic Information, Guangxi University, Nanning 530000, China)

Abstract: In response to the problems of traditional road crack detection methods, such as time-consuming, labor-intensive, high cost, and subjectivity, a YOLOv5-based road crack detection model, named YOLOv5-Crack, is proposed. Firstly, a coordinate attention mechanism is introduced in the backbone and optimized as a CA-plus structure to enhance the crack feature focus. Secondly, a novel feature fusion network ESPP is proposed to reduce some computational costs while improving the feature fusion capability. Then, the heavy Ghost-Shuffle convolution is used in the neck network to replace the traditional convolution, which can keep the channel semantic information as much as possible while reducing computational costs. Finally, the SIoU loss function is introduced to improve the regression accuracy. To validate the effectiveness of the improved YOLOv5-Crack model, comparative experiments are conducted on the GRDDC 2020 dataset, and the results show that the F1 scores are 58.43% and 58.21%, respectively, which are 4.05% and 3.93% higher than those of the original YOLOv5 model, and the computational cost is reduced by 7.8 GFLOPs, with an FPS of 37.9, effectively addressing the shortcomings of road crack detection. Furthermore, compared with mainstream object detection algorithms, the proposed YOLOv5-Crack model has superior performance in road crack detection.

Keywords: pavement cracks; coordinate attention mechanism; ESPP structure; Ghost-Shuffle convolution; SIoU loss function

0 引 言

随着社会的发展,公路交通网越来越庞大,根据中华人民共和国交通运输部发布的《2021 年交通运输行业发展统计公报》^[1],截至 2022 年,全国公路总里程达 528.07×10^4 km,

较上年末增加 8.26×10^4 km,在社会的发展和建设中有举足轻重的地位。大量的货运和客运需求,使得每日的交通量不断增大,导致我国公路网的维护难度剧增,进而导致公路的养护财政支出也在持续上涨。如果不能及时修复路面裂缝,可能会导致路面受到进一步严重的破坏,增加行车

收稿日期:2023-03-08

^{*} 基金项目:国家自然科学基金(62062009)项目资助

的危险性,对交通安全造成影响。此外,还可能会导致车辆受损并增加维修成本的风险。目前,大多数公路部门主要通过人工进行实地检查和测量,效率低、精度不高、成本较大,并且易受人为因素的干扰,难以应对大规模路网的检测需求。因此,针对高检测精度、实时的路面裂缝目标检测方法的探究,不仅能提高行车舒适度,也能进一步确保提高交通安全和道路使用寿命,还能优化路网维护成本和管理效率,节省国家财政支出,实现可持续的道路交通发展。

目前,基于深度学习和计算机视觉的目标检测技术因其卓越的特征提取能力,已经得到了广泛的进展和应用,其精确性和抗噪声性都优于传统的图像形态学方法。当前的目标检测算法包括由YOLO^[21]、SSD^[31]、RetinaNet^[41]代表的一阶段算法和由R-CNN、FPN^[53]代表的二阶段算法。

在传统方法方面,唐伟等^[6]改进Mask匀光与K-means聚类算法进行裂缝的提取,相较于人工提升了检测速度,但是效率仍然不高。赵芳等^[7]改进了Canny算子,为实时道路裂缝检测提供了实验依据,但是对于路面上细小的裂缝的识别仍然存在不足。Quintana等^[8]利用支持向量机基于局部特征描述符训练分类器用于路面裂缝的识别。Dihao等^[9]针对光照环境差、具有阴影等问题,提出利用像素强度和多尺度领域等信息。Fernandes等^[10]提出使用裂缝的图片特征对路面裂缝进行有效描述,但最终的识别效果仍然不理想。

在基于深度学习的二阶段目标检测算法中,李良福等^[11]用改进的窗口滑动算法将裂缝图像分割成较小的裂缝和背景图像之后,建立了一个卷积神经网络的DBCC(deep bridge crack classify)分类模式,增强了识别的效用性和泛化性。Li等^[12]提出了把基于像素的自适应加权交叉熵(weighted cross-entropy, WCE)损失与杰卡德距离(Jaccard distance)相结合,再将其应用于二阶段算法中,促进了高质量的像素级路面裂缝检测。徐康等^[13]将FasterRCNN模型与CBAM(convolutional block attention module)注意力机制结合,但是对细微裂缝的漏检问题仍然存在。Nguyen等^[14]提出一种新的二阶段算法框架,在第一阶段消除噪声或伪影,并将潜在裂缝分离到一个小区域,在第二阶段进行学习检测裂缝,该方法局限于特定的路面裂缝图片且运行时间偏长。Yu等^[15]开发上下文增强胶囊特征金字塔网络,便于提取不同级别的裂缝特征,但其动态路由的特点,导致整个模型的执行效率较低。

近年来,以边框回归法为基础的一阶段目标检测算法兴起,通过对网络中提取的特征进行回归来预测裂缝类别和位置。瞿中等^[16]基于优化的VGG-16,通过使用注意力机制和空洞卷积获得更多细节信息并扩大感受野,既提高了算法的鲁棒性,又提高了检测的准确度,但是整个模型缺乏应有的泛化能力。张诗慧等^[17]提出基于改进RetinaNet的算法,提升了表面裂缝的定位精度,但是网络计算量有所上升。Mandal等^[18]使用YOLOv2技术,实现了对路面裂

缝的自动检测,但是精度不足。Ma等^[19]将YOLOv3与PCGAN对抗网络相结合,同时结合中值流算法,加快了检测速度的同时也提升了一定的精确度,但是该算法对于龟裂状裂缝的检测效果不佳。郝巨鸣等^[20]在YOLOv4主干网络中引入Ghost模块,然后在输出端引入ECA(efficient channel attention)注意力机制,提高检测速度的同时也保证了一定的精度。Yan等^[21]提出了一种可变形SSD的新型网络,并将其与VGG-16相结合,增强了其在复杂环境中的适应性,但是对于图像边缘的裂缝特征关注度有欠缺。

综上所述,传统路面裂缝检测方法存在对图像质量要求较高的问题,限制了其适用范围。其次,这些方法提取的裂缝特征不够精确,导致检测的精度较低;二阶段路面裂缝检测算法运行时候选区域过多,虽然能保证一定的精确度,但是由于其计算量大,难以保障实时性;一阶段检测算法缺乏一定精度。为更加关注裂缝的位置信息,保证实时性的同时提高精确度,本文基于一种工业级目标检测算法YOLOv5^[22],提出改进的路面裂缝检测模型YOLOv5-Crack,本文的主要贡献如下:

- 1)引入高效轻量的坐标注意力机制^[23](coordinate attention, CA)考虑了通道间的相互作用,还考虑了裂缝特征空间的位置信息,在此基础上优化为CA-plus结构,使其能更有效地集中于有用信息;
- 2)在SPPF^[24]的基础上,提出一种全新的特征融合网络ESPP(efficient spatial pyramid pooling),既可增强特征融合能力,又能降低模型计算量;
- 3)通过使用轻量级重影混洗卷积(ghost-shuffle Conv)^[25],在保证对模型学习贡献能力的同时又降低计算成本;
- 4)引入考虑连续性的边框损失函数SIoU^[26],提高回归速度和精度。

在实验方面采用对比和消融实验来验证YOLOv5-Crack模型在路面裂缝检测方面的准确性和实时性。该模型可以帮助政府和相关道路部门机构及时检测出路面裂缝,以便采取适当的措施,提供更好的道路交通环境。

1 模型改进

1.1 YOLOv5 基本网络模型

YOLOv5是一种典型的一阶段目标检测算法,在检测过程中不产生图像备选框,一步到位直接生成检测目标的坐标和类别,在目标检测领域有着举足轻重的地位。YOLOv5目前有着多种网络宽度和深度的版本,包括YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。鉴于路面裂缝在道路病害检测中的重要性,并综合考虑检测速度和检测精度,本文选择了YOLOv5l模型。其主体结构由输入端、特征提取、颈部网络和预测端组成。

输入端主要采用自适应性锚框计算和Mosaic数据强化。Mosaic数据强化是使用4张图片,对其进行随机缩

放、剪裁、排布的方式加以拼接,并凭其结果获得最新的图像加以训练,该方法可以使训练所得的模型具有更强的泛化能力以及鲁棒性,如图 1 所示。

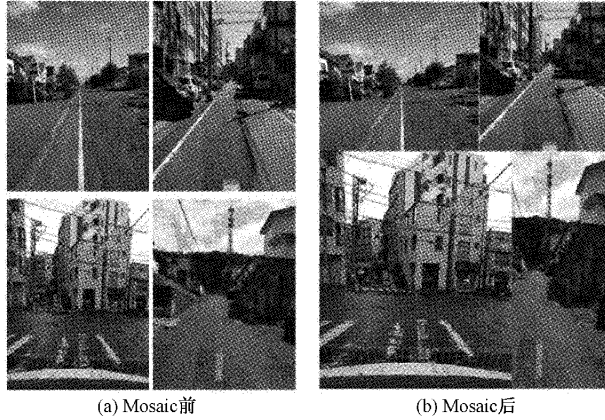


图 1 Mosaic 数据增强对比

特征提取网络中包括了 Conv 模块处理下采样,减少计算的复杂度;C3 结构作为一种残差模块的组合,把特征图根据通道维度分为两个分支,一个分支通过传统卷积模块,另一个分支直接与其映射拼接;SPPF 快速空间金字塔池化用于多尺度的融合,并且提高了运算速度。

颈部网络则是使用路径聚合 PAN 结构,自底向上将特征提取网络中所提取的不同阶段的特征进行上采样和池化操作后融合。

预测端则有 3 种不同标准的分支,依次对应大、中、小目标的检测输出。YOLOv5l 基本架构如图 2 所示。

1.2 YOLOv5-Crack 模型

本文提出的路面裂缝检测模型 YOLOv5-Crack,如图 3 所示,改进之处使用虚线框进行标注。本文的改进具体如下:

1)在 Backbone 层中,使用 CA-plus 结构提取路面裂缝特征信息;

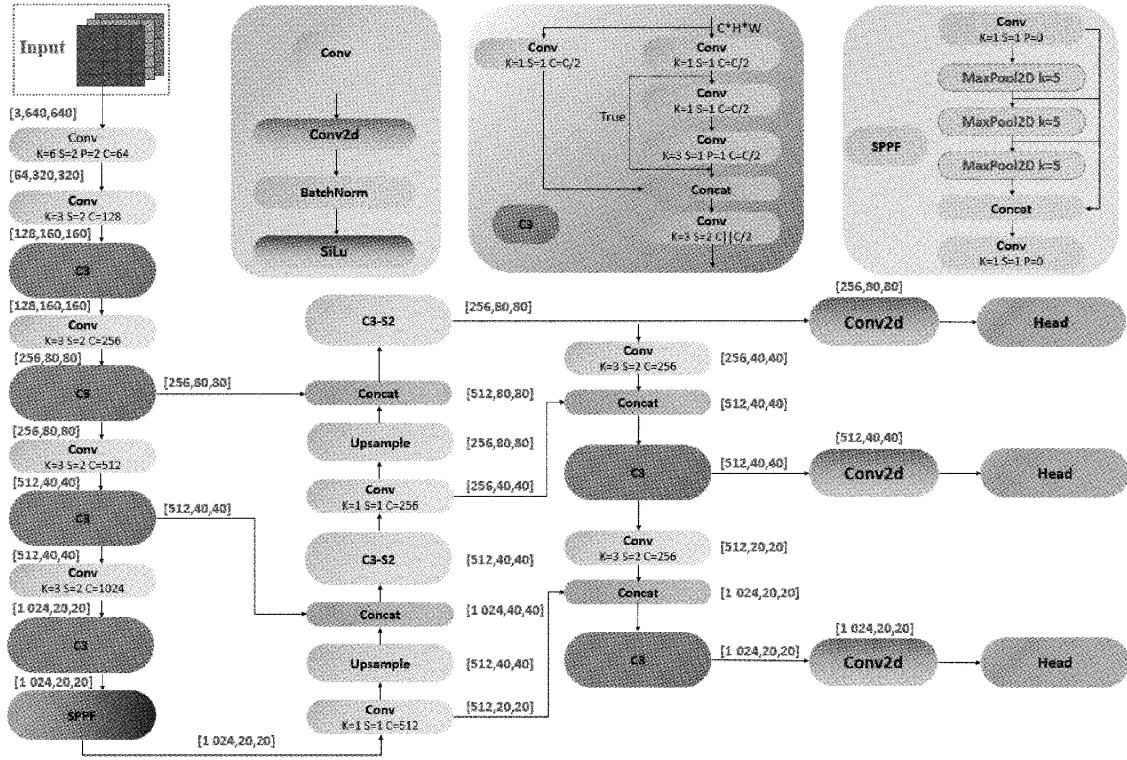


图 2 YOLOv5l 基本网络结构图

2)在 Backbone 层中,用全新的 ESPP 结构替换原来的 SPPF 结构;

3)在 Neck 层中,通过使用轻量级重影混洗卷积来代替 Neck 部分的传统卷积进行上采样;

4)将 YOLOv5 中默认的边框损失函数 CIoU 替换为 Siou。

1.3 CA 注意力机制

由于路面裂缝的隐藏性以及多重性,目标检测算法经

常无法准确观察到具体裂缝的信息。注意力机制可以让网络聚焦于局部信息,找到有用的裂缝信息。传统的通道注意力较少关注位置信息,空间注意力较少关注通道信。

CA 注意力机制在关注图像通道内容的同时,又整合了位置信息,使网络在扩大全局感受野的基础上,又能捕获到空间方向上的长程依赖。为了能够捕获路面裂缝的精确位置信息,CA 注意力机制模块如图 4 所示:

CA 注意力机制仅仅对高度 H 、宽度 W 两个维度编

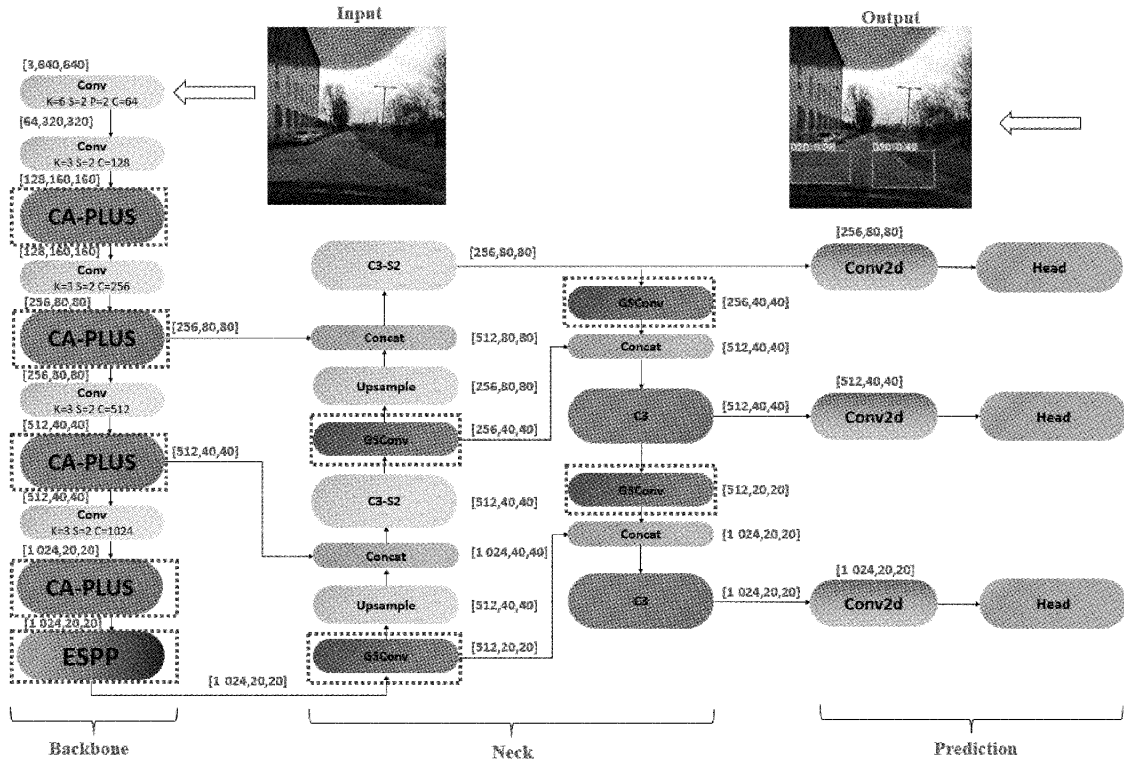


图3 YOLOv5-Crack网络结构图

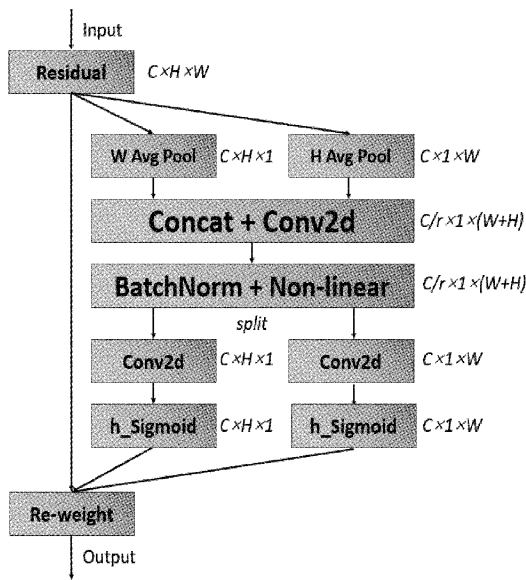


图4 CA注意力机制模块

码,其输入 X 通过残差模块后,分别进行水平与垂直的一维自适应平均池化,得到两个不同空间方向的特征,高度 h 处的第 c 个通道的输出如式(1):

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (1)$$

同理,宽度为 w 处的第 c 个通道的输出如式(2):

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (2)$$

对这两个一维特征,在其宽度上执行拼接操作,再经过卷积、归一化处理、非线性回归得到包含横向和纵向空间信息的特征 s ,其定义如式(3):

$$s = \delta(F([\mathbf{z}^h, \mathbf{z}^w])) \quad (3)$$

之后将特征 s 在通道上进行拆分,通过卷积和 Sigmoid 激活函数得到两个特征图,使其维度与输入 X 相同:

$$y^h = h_\sigma(s_h(s^h)) \quad (4)$$

$$y^w = h_\sigma(s_w(s^w)) \quad (5)$$

这两个特征图通过一定的权重调整,互补地应用于输入特征图来增强对于路面裂缝的注意力,即将原特征与在宽度与高度方向上有权重的特征进行融合,输出 o 定义如式(6):

$$o_c(i, j) = x_c(i, j) \times y_c^h(i) \times y_c^w(j) \quad (6)$$

如图5所示,针对CA注意力机制的种种优势,传统的CA结构仅仅将CA注意力模块放在深层,即只对深层特征图执行特征操作。本文对传统的CA结构进行改进,将CA注意力模块嵌入C3模块,构成CA-plus结构,即对每个特征层提取的特征图都进行处理,最终得到感受视野大、感知方向独立的特征图。相比传统CA结构,CA-PLUS结构多使用了3次CA注意力模块,仅在提升少部分复杂度的情况下,获得更多的特征语义,大大增加了特征提取的精度。CA-plus模块具体实现伪码如图6所示。

1.4 ESPP结构

YOLOv5中Conv模块采取的SiLU激活函数负半轴

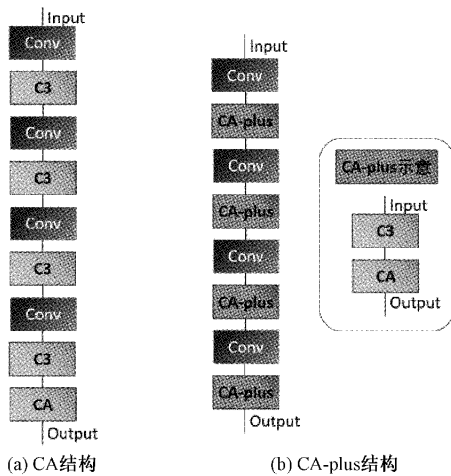


图 5 CA、CA-PLUS 结构对比图

Algorithm 1 CA-plus 模块

Input: 输入特征图 (维度为 $h*w*c$)

Output: 输出特征图

- 1: for $m = 1 \rightarrow 2$ do
- 2: 经过 Conv 卷积模块;
- 3: end for
- 4: 将两者所得结果进行 Sequential 操作;
- 5: 进行 Concat 拼接操作;
- 6: 将拼接结果进行 Conv 操作;
- 7: for $i = 1 \rightarrow h$ do
- 8: 提取 h 向坐标信息;
- 9: end for
- 10: for $i = 1 \rightarrow w$ do
- 11: 提取 w 向坐标信息;
- 12: end for
- 13: 将位置坐标信息和输入特征图拼接;
- 14: 将位置坐标的权重与原始特征图相乘, 得到加权特征图;
- 15: 对加权特征图进行卷积操作;
- 16: return 经 CA-plus 所提取的特征信息

图 6 CA-plus 伪码实现

值较大,使得在训练过程中,产生的正则化效果及泛化能力较差,故提出一种新的 MConv 卷积模块。其采用 Mish^[27] 激活函数,无上界的同时并且非饱和,能有效地加快训练速度,而且该激活函数具有连续可微、非单调的性质,有效地避免了奇异点,其优秀的性能使得 MConv 卷积模块在下采样提取裂缝特征时精度更高,速度更快,鲁棒性更强。SiLU 激活函数与 Mish 激活函数的对比如图 7 所示。

SPPF 作为进行特征融合的结构,将下采样得到的特征图依次输入至 3 个 5×5 的池化层,将每一层的最大池化结果执行拼接操作,有效避免了对图像区域裁剪、缩放操作导致的图像失真问题,实现了 Neck 层之前的输入特征图尺度统一,能更好保证结果的准确性。然而,SPPF 中的最大池化操作只考虑特征图上最大值的位置,会忽略某些

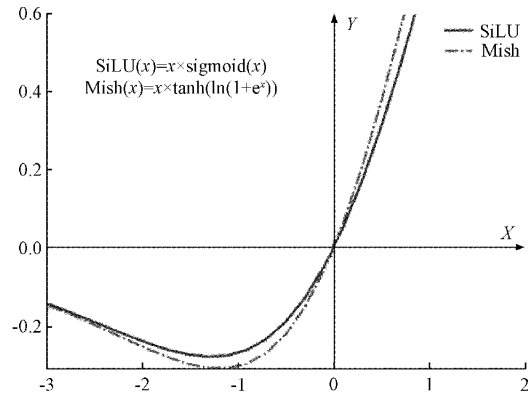


图 7 SiLU 和 Mish 激活函数对比图

裂缝特征信息,增加了过拟合的风险,也会引入噪声导致裂缝信息损失。因此,结合 MConv 卷积,提出一种全新的空间金字塔池化结构 ESPP 如图 8 所示。相比于原本的 SPPF 模块,ESPP 使用了 MConv,具有更好的平滑性,使得特征信息能够更好地进入特征融合网络。在小感受野层面,ESPP 仍使用最大池化;而在大感受野层面,ESPP 采用 9×9 和 13×13 的平均池化层,平均池化操作会将特征图上的值平均化,受噪声和异常值的影响相对较小,更加适用于环境复杂的路面裂缝检测。同时,平均池化更能保留空间信息,减少了部分过拟合的风险。

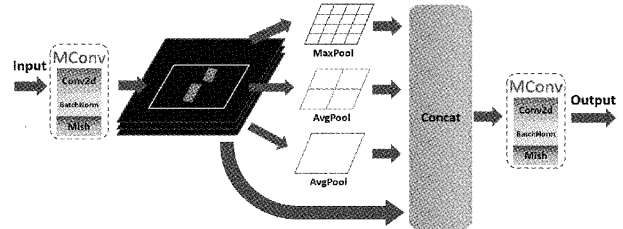


图 8 ESPP 结构

1.5 重影混洗卷积 GSConv

由于路面裂缝检测对于实时性具有一定要求,所以在保证精度的同时,需要加快计算速度,在颈部网络 Neck 层中引入重影混洗卷积 GSConv 来替代标准卷积 Conv。在卷积神经网络中,所有反馈图像都需要在 Backbone 中将提取的特征向通道传输。在传输过程中,特征图的宽度和高度特征的压缩和通道扩展会引起局部裂缝图像语义信息丢失。GSConv 尽可能地保持语义信息,即尽可能地保存了每个通道之间的隐藏连接的同时,又最大程度地降低了计算量,其结构如图 9 所示:

令输入的通道数为 C_{inp} ,输出的通道数为 C_{outp} 。输入在通过第 1 个卷积后会产生通道数为 $C_{outp}/2$ 的特征张量,然后再通过 1 个卷积,得到另一个通道数为 $C_{outp}/2$ 的特征张量,接着对这两个特征张量进行拼接,之后对其数据尺寸进行通道混洗,即打乱原特征图通道排列顺序,最后再对其执行拼接操作。

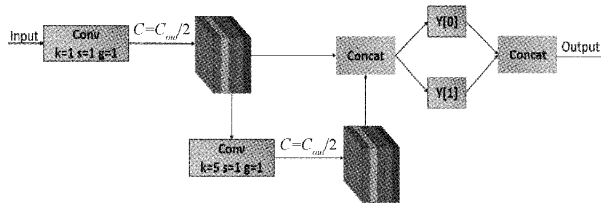


图9 GSConv 结构图

如果在YOLOv5网络的每个阶段都用GSConv替代Conv,将会引起网络层数的大幅增加,从而使得信息流的传播阻力增大。而在Neck层中,特征信息经ESPP结构融合,通道维度达到最大,宽度和高维度都达到最小,并且此刻的冗余信息最少,使用GSConv效果会更好。

1.6 SIoU 损失函数

YOLOv5算法中,其损失函数分别由分类损失、定位损失和置信度损失3部分通过加权相加构成。对于目标检测算法,最重要的就是通过边界框具体位置数据预测出需要检测的目标,因此提出了IoU的损失计算方法。YOLOv5中采用的是CIoU(complete intersection over union)损失函数,其定义如式(7)~(10):

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (9)$$

$$Loss_{CIoU} = 1 - IoU + \frac{\rho_{(b,b^{gt})}^2}{d^2} + \alpha v \quad (10)$$

其中,参数A和B分别表示真实边界框和预测边界框的面积, b 和 b^{gt} 分别表示预测边界框与实际边界框的质心, ρ 则表示两边界框之间的欧氏距离, α 是权重函数, v 用来表示纵横比的相似性。然而,CIoU的计算依赖于边界框回归指标的聚合,而未顾及到所需边界框和预测边界框之间的方向不匹配,以至于大大降低了网络的训练速度和预测精确度。因此,本文提出使用更为全面的SIoU(scylla intersection over union)损失函数,如图10所示,它由4部分损失组成,分别为角度、距离、形状和IoU损失。

角度损失定义如式(11)~(12):

$$\Delta = 1 - 2 \times \sin^2 \left(\arcsin(x) - \frac{\pi}{4} \right) \quad (11)$$

$$x = \sin(\alpha) = \frac{c_h}{\sigma} \quad (12)$$

其中, σ 为预测边界框与真实边界框的欧氏距离, c_h 为两边界框的高度差。该角度损失的计算可以使预测框能更好得回归到与真实框同一水平或垂直线上,加速了模型收敛。

距离损失定义如式(13):

$$\Delta = 2 - e^{-\frac{(b_x^{gt} - b_x)^2}{c_w}} - e^{-\frac{(b_y^{gt} - b_y)^2}{c_h}} \quad (13)$$

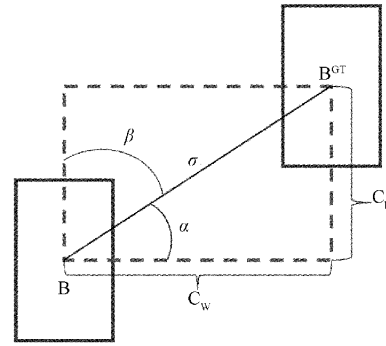


图10 SIoU 角度损失

其中, b_x^{gt} 为地面裂缝真实边界框的质心的横坐标, b_y^{gt} 为地面裂缝真实边界框的质心的纵坐标, b_x 为地面裂缝预测边界框的质心的横坐标, b_y 为地面裂缝预测边界框的质心的纵坐标, c_w 为地面裂缝真实边界框与地面裂缝预测边界框最小外接矩形的宽, c_h 为地面裂缝真实边界框与地面裂缝预测边界框最小外接矩形的高。

形状损失定义如式(14)~(16):

$$\Omega = \sum_{z=w,h} (1 - e^{-\omega_z})^\theta \quad (14)$$

$$\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (15)$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (16)$$

其中, θ 定义的是形状的成本,且对于每个数据集都是唯一的,通过遗传算法求得。 w 为地面裂缝预测边界框的宽, h 为地面裂缝预测边界框的高, w^{gt} 为地面裂缝真实边界框的宽, h^{gt} 为地面裂缝真实边界框的高。

综上,SIoU引入了回归所需的角度、距离以及形状损失,相较于CIoU降低了回归的自由度,并且使得网络收敛更快,进而能提升回归的精度。

2 实验

2.1 实验环境

操作系统为Ubuntu20.04,深度学习框架为Pytorch 1.11.0,编程语言为Python 3.8,计算机内存为75 GB,CPU型号为Intel(R) Xeon(R) Platinum 8255C,显卡是RTX3090×2张,显存共48 GB。

2.2 数据集

GRDDC 2020^[28](global road damage detection challenge 2020)是一个专业的路面裂缝数据集。该数据集涵盖了3个不同国家(日本、捷克、印度)的多种来自高速公路、乡村道路等的路面裂缝图片,具有很高的专业性和实用性。其裂缝分别为D00(纵向裂缝)、D10(横向裂缝)、D20(龟裂裂缝)和D40(坑洞裂缝),共21 041张图片,按照6:2:2的形式分割了训练集、测试集和验证集。官方也给了两个测试集Test1和Test2,图片数量分别为2 631张和2 664张。

2.3 实验过程与评价指标

以在 COCO 数据集上经过训练得到的 YOLOv5l 权重作为预训练权重,其中训练轮次设置为 300 epoch,同时为了防止过拟合,将 YOLOv5 的耐心值设置为 30,即模型如果在 30 轮内没有提升,就立刻停止训练模型。

令 TP (true positive)、TN (true negative)、FP (false positive)、FN (false negative) 分别是实际为正样本且预测成正样本、实际为负样本且预测成负样本、实际为负样本但预测成正样本、实际为正样本但预测成负样本的数量, m 是样本类别数量。

精度 (precision, P) 指所有预测为正样本的结果中,预测正确的百分比。召回率 (recall, R) 指所有正样本中被预测正确的百分比。AP 是 P-R 曲线下的面积。mAP 是所有类别 AP 的平均。计算公式如式(17)~(20):

$$P = \frac{TP}{TP + FP} \quad (17)$$

$$R = \frac{TP}{TP + FN} \quad (18)$$

$$AP = \int_0^1 PdR \quad (19)$$

$$mAP = \frac{\sum_{i=1}^m AP_i}{m} \quad (20)$$

同时比赛方又给出两个测试集: Test1 和 Test2,其结果通过“F-Measure”进行评估,其仅当 IoU 大于等于 0.5 时,且预测标签和真实标签相互对应时,推理才被认为是正确的,其指标名称为 $F1$,计算公式如式(21):

$$F1 = 2 \times \frac{P \cdot R}{P + R} \quad (21)$$

计算量 GFLOPs 代表每秒 10 亿次的浮点运算数, FPS 代表模型每秒可以处理图片的数量,用于衡量模型的检测速度。

本实验在过程中采用精度 P 、召回率 R 、所有类别平均精度 $mAP@0.5$ 、GFLOPs、FPS 作为评价指标,在最终比较时,采用两个数据集的 $F1$ 分数作为评价指标。

2.4 注意力模块对比实验

为了验证 CA 注意力模块对于裂缝检测的有效性,本

文分别引入 SE、ECA、CBAM 注意力模块进行对比实验,实验结果如表 1 所示。

由表 1 的注意力模块对比实验可知,除了 CA,其他注意力模块对于各项指标或多或少都有所降低,而且 CA 注意力模块对于 $mAP@0.5$ 的提升高达 0.7%,而且计算量的提升也是最低,因此,引入 CA 注意力模块对于裂缝细节信息提取具有实际意义。同时,为了验证优化后的 CA-plus 结构的有效性,进行了实验,其结果如表 2 所示。

由表 2 的注意力结构对比实验可知,CA-plus 结构仅以 0.1 GFLOPs 的计算量为代价,全面提升了评价指标的值。综上所述,本文提出的 CA-plus 结构对于检测性能提升具有作用。

表 1 注意力模块对比实验

Model	P/%	R/%	mAP@0.5/%	GFLOPs
Baseline	58.1	53.4	54.7	107.7
+SE	58.4 (+0.3)	53.0 (-0.4)	54.4 (-0.3)	107.9 (+0.2)
+ECA	57.4 (-0.4)	53.2 (-0.2)	54.2 (-0.5)	107.9 (+0.2)
+CBAM	58.0 (-0.1)	53.5 (+0.1)	54.8 (+0.1)	108.2 (+0.5)
+CA	58.2 (+0.1)	53.8 (+0.4)	55.4 (+0.7)	107.9 (+0.2)

表 2 注意力结构对比实验

Model	P/%	R/%	mAP@0.5/%	GFLOPs
CA	58.2	53.8	55.4	107.9
CA-plus	58.3 (+0.1)	54.2 (+0.4)	55.6 (+0.2)	108.0 (+0.1)

2.5 ESPP 模块对比实验

为验证改进的 ESPP 结构对于裂缝检测的有效性,本文将原模型与更换为 ESPP 的模型进行了对比实验,结果如表 3 所示。

表 3 ESPP 对比实验

Model	P/%	R/%	mAP@0.5/%	GFLOPs	训练时间
Baseline	58.1	53.4	54.7	107.7	15 h 21 min
+ES	58.5(+0.4)	53.8(+0.5)	55.2(+0.5)	105.7(-2.0)	14 h 15 min(-1 h 6 min)

由表 3 的验证实验可知,由于 MConv 模块和平均池化的优越性,在 ESPP 中的使用可以让模型整体的计算量降低,提升了回归的速度。同时带来的检测精度的上升,证明了特征信息更好地进入了融合网络,提升了模型的泛化能力。训练时间下降了 7.2%,也验证了 ESPP 在提升

训练速度方面的有效性。

2.6 GSConv 模块对比实验

为验证本文提出的重影混洗卷积 GSConv 的有效性,本文将基线模型与把 Neck 中的 Conv 替换为 GSConv 的模型进行对比实验,结果如表 4 所示。

表4 GSConv对比实验

Model	P/%	R/%	mAP@0.5/%	GFLOPs
Baseline	58.1	53.4	54.7	107.7
+GSconv	58.3 (+0.2)	53.0 (-0.3)	55.0 (+0.3)	105.7 (-2.0)

由表4的验证实验可知,将Conv替换为GSConv后,计算量会有小幅度的下降,以R值降低为代价,能够提升P值和mAP@0.5值。

2.7 损失函数对比实验

为验证本文提出的将YOLOv5中的CIoU损失函数计算替换为SIoU损失函数,对其进行对比实验,实验结果如表5所示。

表5 损失函数对比实验

Loss	P/%	R/%	mAP@0.5/%	GFLOPs
CIoU	58.1	53.4	54.7	107.7
SIoU	58.2 (+0.1)	54.9 (+1.5)	55.3 (+0.6)	107.7

由表5可知,通过替换模型中的损失函数计算方式,在计算量不变的情况下,训练出的模型的召回率R大幅提升1.5%,证明了更改SIoU对于回归精度的提升具有作用。

2.8 曲线对比分析

在自测数据集中,YOLOv5-Crack和Baseline的损失收敛曲线和精度收敛曲线分别如图11和12。

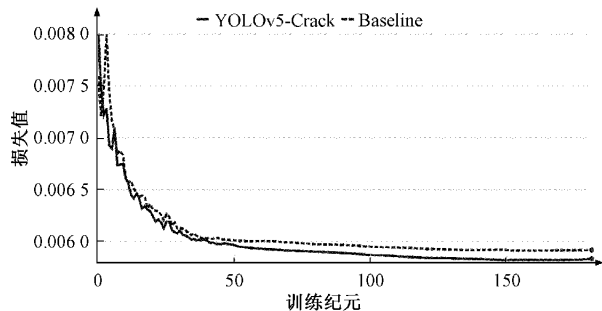


图11 损失收敛曲线对比图

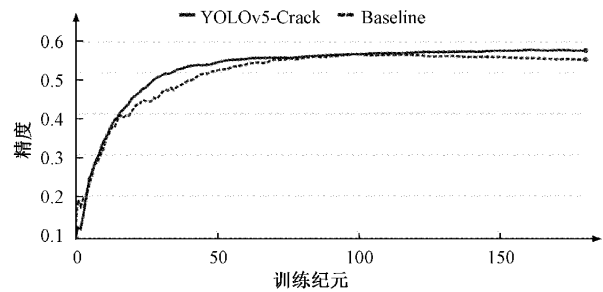


图12 精度收敛曲线对比图

从损失收敛曲线对比图中可以看出,YOLOv5-Crack的损失下降相较之前的Baseline更为平滑,并且其损失值在最终的值更低,说明YOLOv5-Crack在学习裂缝的特征过程中稳定性更好,能够更好地泛化测试数据。

从精度收敛曲线对比图中可以看出,YOLOv5-Crack的平均精度值明显高于Baseline,且前者的收敛时间更迟,没有出现明显的过拟合的现象,说明其泛化能力较强,可以适应更广泛的路面裂缝场景,具有更好的实际应用价值。

2.9 消融实验

通过官方给出的测试集Test1和Test2进行消融实验,并以F1得分作为评价指标,由表6的实验结果可得,相比于基线Baseline,序号5同时改进了CA-plus与ESPP,其F1分数在两个测试集上分别提升2.13%和1.21%,证明了CA-plus与ESPP同时改进的有效性;序号6中,在CA-plus的基础上又将Conv替换为GSConv,其F1分数的提升均高于单项,证明了GSConv在提升检测精度和加快检测速度方面有效果;序号7,在原有CA-plus的基础上修改了损失函数为SIoU,相比序号1中的FPS提升0.8,证明了SIoU损失函数对于回归精度的提升有帮助,并且也一定程度上提升了检测速度。对于最终模型YOLOv5-Crack,降低了7.8 GFLOPs的计算量,F1分数分别提高4.05%和3.93%,同时其检测速度FPS方面只付出降低1.3的代价,其中计算速度的降低主要来自于CA-plus模块,但是将4项改进综合在一起,所取得的成效完全优于各模块单独相添加,即在路面裂缝检测上取得了更优的检测成效。

表6 消融实验

Number	CA-plus	ESPP	GSConv	SIoU	Test1-F1	Test2-F1	GFLOPs	FPS
Baseline					54.38	54.28	107.7	39.2
1	✓				56.14(+1.76)	55.24(+0.96)	107.9(+0.2)	36.8(-2.4)
2		✓			55.66(+1.28)	55.38(+1.10)	105.7(-2.0)	38.9(-0.3)
3			✓		54.71(+0.33)	55.61(+1.33)	105.7(-2.0)	38.9(-0.3)
4				✓	55.86(+1.48)	54.67(+0.39)	107.7	39.4(+0.2)
5	✓	✓			56.51(+2.13)	55.49(+1.21)	107.7(+0.2)	37.6(-1.6)
6	✓		✓		55.47(+1.09)	55.73(+1.45)	105.7(-2.0)	37.7(-1.5)
7	✓			✓	56.31(+1.93)	56.42(+2.14)	107.9(+0.2)	37.6(-1.6)
本文	✓	✓	✓	✓	58.43(+4.05)	58.21(+3.93)	99.9(-7.8)	37.9(-1.3)

2.10 主流算法对比实验

为验证本文提出的 YOLOv5-Crack 模型的优越性,将本文改进模型与 YOLOv4、YOLOv5m、YOLOv7、EfficientDet、Faster R-CNN 进行了对比,实验结果如表 7 所示。

由表 7 可得,相较于主流二阶段目标检测算法和多种主流一阶段目标检测算法,本文提出的改进算法在 F1 分数表现更具有优越性。为了使改进效果更加明显,本文对改进前后模型进行热力图可视化分析,实验结果如图 13 所示。

通过对热力图的分析,改进的 YOLOv5-Crack 模型的热力图覆盖更加明显,并且对于预测输出结果的影响更大,即对于裂缝目标的定位精确度大幅提升,意味着改进

表 7 主流算法对比实验

Algorithm name	Test1-F1/%	Test2-F1/%
YOLOv4	56.36	57.07
YOLOv5m	55.36	54.12
YOLOv7	54.86	54.13
EfficientDet	56.50	54.70
Faster R-CNN	54.13	54.30
YOLOv5-Crack	58.43	58.21

后的算法更加关注于裂缝的特征目标,进而证明了改进的有效性。同时为了使对比结果更加直观,展示算法模型的检测效果如图 14 所示。

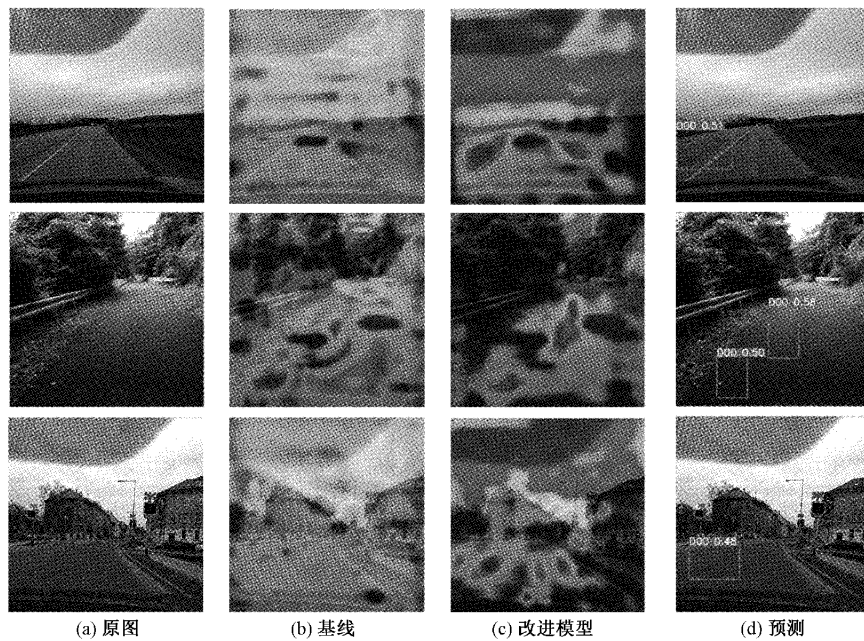


图 13 热力图对比

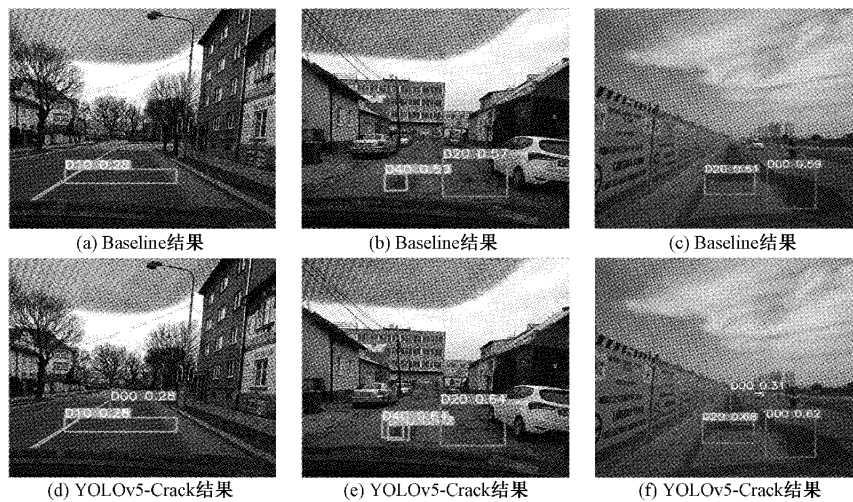


图 14 实验对比图

3 结 论

本文针对路面裂缝检测问题,提供了一种高效且实用的目标检测模型YOLOv5-Crack,该模型旨在解决路面裂缝检测精度低,实时性差等问题,在YOLOv5l的基础上,增添了CA注意力模块的同时并将其改进为CA-plus结构,提升了特征关注能力;提出ESPP结构,有效降低噪声干扰并提升特征融合能力;用轻量化GS-Conv替换Conv,有效提高检测速度;采用Siou损失函数提升回归精度。通过消融实验证明,改进的模型在测试集1上的F1分数为58.43%,测试集2上的分数为58.21%,计算量下降7.8GFLOPs,而且相比主流目标检测算法也具有一定的优越性,但是仍然存在一定的缺陷:模型不够轻量化,且数据集中只包含4种主要道路裂缝,对于其他类型的裂缝检测不足。因此,对模型进行轻量化操作(如知识蒸馏、剪枝),扩大数据集等问题还有待继续研究。

参考文献

- [1] 中华人民共和国交通运输部. 2021年交通运输行业发展统计公报[EB/OL]. https://xxgk.mot.gov.cn/2020/jigou/zhghs/202212/t20221228_3730414.html.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 779-788.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [4] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE international conference on computer vision, 2017: 2980-2988.
- [5] 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, 2017: 2117-2125.
- [6] 唐伟, 余波, 赵嘉彬, 等. 改进 MASK 匀光与 K-means 聚类结合的桥梁裂缝提取 [J]. 电子测量技术, 2021, 44(22): 128-133, DOI: 10.19651/j.cnki.emt.2107371.
- [7] 赵芳, 周旺辉, 陈岳涛, 等. 改进的 Canny 算子在裂缝检测中的应用 [J]. 电子测量技术, 2018, 41(20): 107-111. 1801773.
- [8] QUINTANA M, TORRES J AND MENÉNDEZ J M. A simplified computer vision system for road surface inspection and maintenance [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(3): 608-619.
- [9] DIHAO A, GUIYUAN J, LAM S K, et al. Automatic pixel-level pavement crack detection using information of multi-scaleneighborhoods [J]. IEEE Access, 2018, 8(6): 24452-24463.
- [10] FERNANDES K, CIOBANU L. Pavement pathologies classification using graph-based features [C]. 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 2014.
- [11] 李良福, 马卫飞, 李丽, 等. 基于深度学习的桥梁裂缝检测算法研究 [J]. 自动化学报, 2019, 45(9): 1727-1742.
- [12] LI K, WANG B, TIAN Y J, et al. Fast and accurate road crack detection based on adaptive cost-sensitive loss function [J]. IEEE Transactions on Cybernetics, 2023, 53(2): 1051-1062.
- [13] 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.
- [14] NGUYEN N H T, PERRY S, BONE D, et al. Two-stage convolutional neural network for road crack detection and segmentation [J]. Expert Systems With Applications, 2021, 186.
- [15] YU Y, GUAN H, LI D, et al. CCapFPN: A context-augmented capsule feature pyramid network for pavement crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 23(4): 3324-3335.
- [16] 瞿中, 王彩云. 基于注意力机制和轻量级空洞卷积的混凝土路面裂缝检测 [J]. 计算机科学, 2023, 50(2): 6.
- [17] 张诗慧, 罗晖, 裴莹玲, 等. 基于改进 RetinaNet 的高铁无砟轨道板表面裂缝检测 [J]. 计算机工程与应用, 2023, 59(6): 310-317.
- [18] MANDAL V, UONG L, ADU-GYAMFI Y. Automated road crack detection using deep convolutional neural networks [C]. 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018.
- [19] MA D, FANG H, WANG N, et al. Automatic detection and counting system for pavement cracks based on PCGAN and YOLO-MF [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 22166-22178.
- [20] 郝巨鸣, 杨景玉, 韩淑梅, 等. 引入 Ghost 模块和 ECA 的 YOLOv4 公路路面裂缝检测方法 [J]. 计算机应用, 2023, 43(4): 1284-1290.
- [21] 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.
- [22] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. 2017 IEEE

- International Conference on Computer Vision, New York: IEEE Press, 2017: 2999-3007.
- [23] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway: IEEE, 2021: 13708-13717.
- [24] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [25] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles [J]. ArXiv Preprint, 2022, ArXiv:2206.02424.
- [26] GEVORGYAN Z. SIoU Loss: More powerful learning for bounding box regression [J]. ArXiv Preprint, 2022, ArXiv:2205.12740.
- [27] DIGANTA M. Mish: A self regularized non-monotonic activation function. [J]. ArXiv Preprint, 2019, ArXiv:1908.08681.
- [28] MAEDA H, SEKIMOTO Y, SETO T, et al. Road damage detection and classification using deep neural networks with smartphone images[J]. Computr-Aided Civil Infrastructure Engineering, 2018, 33(12): 1127-1141.

作者简介

沈思远,主要研究方向为图像识别、计算机视觉等。

E-mail:2010677484@qq.com

华蓓(通信作者),硕士,副教授,主要研究方向为图像识别、信息安全等。

E-mail:huabei111@qq.com

黄汝维,博士,副教授,主要研究方向为云计算、同态加密、人工智能等。

E-mail:ruweih@126.com