

基于改进 YOLOv5s 的混凝土轨枕裂缝智能识别方法研究

熊律¹, 许欢纯¹, 宁善平¹, 肖赐麟¹, 宋斌²

(1. 广东交通职业技术学院 轨道交通学院, 广东 广州 510650;

2. 新疆交通职业技术大学 机电工程学院, 新疆 乌鲁木齐 831401)

摘要: 针对混凝土轨枕裂缝形状的不规则性, 采用标注工具 Labelme 对裂缝数据集进行多边形标注, 实现裂缝标注的精细调整。同时, 基于 YOLOv5s 模型, 在其 Backbone 层植入 CA (Coordinate Attention) 注意力机制, 建立 YOLOv5s-CA 模型, 实现对混凝土轨枕裂缝的精确定位。实验结果表明, YOLOv5s-CA 目标检测模型参数选取合理, 模型训练结果准确。YOLOv5s-CA 模型的检测精度和召回率训练结果为 92.68% 和 84.12%, 相比于 YOLOv5s 模型, 分别提升了 1.8% 和 4.44%。平均检测精度 mAP@0.5 和 mAP@0.5:0.95 达到了 90.41% 和 72.88%, 相比于 YOLOv5s 模型, 分别提升了 3.54% 和 9.39%, 验证了本文提出的 YOLOv5s-CA 模型对混凝土轨枕裂缝检测具有较好的效果。

关键词: 轨枕裂缝; Labelme; YOLOv5s; CA 注意力机制; 检测精度

中图分类号: U213.3

文献标识码: A

Research on Intelligent Identification Method of Concrete Sleeper Cracks Based on Improved YOLOv5s

XIONG Lü¹, XU Huanchun¹, NING Shanping¹, XIAO Cilin¹, SONG Bin²

(1. School of Rail Traffic, Guangdong Communication Polytechnic, Guangzhou 510650, Guangdong, China;

2. College of Electromechanical Engineering, Xinjiang Jiaotong Vocational and Technical University, Urumchi 831401, Xinjiang, China)

Abstract: In response to the irregularity of crack shapes in concrete sleepers, the labeling tool Labelme is used to perform polygon labeling on the crack dataset, achieving fine adjustment of crack labeling. At the same time, based on the YOLOv5s model, a CA (Coordinate Attention) mechanism is implanted in its Backbone layer to establish the YOLOv5s CA model and achieve precise localization capability for cracks in concrete sleepers. The experimental results show that the parameter selection of YOLOv5s-CA object detection model is reasonable, and the model training results are accurate. The detection accuracy and recall of the YOLOv5s CA model after training reached 92.68% and 84.12%, respectively, which improved by 1.8% and 4.44% compared to the YOLOv5s model. The average detection accuracy mAP@0.5 and mAP@0.5:0.95 achieved 90.41% and 72.88%, respectively, which improved by 3.54% and 9.39% compared to the YOLOv5s model. This verifies that the YOLOv5s CA model proposed in this paper has a good effect on crack detection of concrete sleepers.

Key words: sleeper crack; Labelme; YOLOv5s; CA attention mechanism; detection accuracy

依据国家铁路局发布的《2024 年铁道统计公报》^[1] 数据, 截至 2024 年底, 我国高速铁路营业

里程已突破 4.8 万千米。轨枕作为安装钢轨的基础, 其性能状态直接关系到高速动车组的安全运

收稿日期: 2025-10-27

作者简介: 熊律, 副教授, 博士, 研究方向为大型装备健康监测与智能维护决策

基金项目: 广东省科技创新战略专项资金项目 (编号: pdjh2025bk335); 广东省普通高校重点科研平台和项目 (编号: 2024KCXTD084, 2024KTSCX382); 广东交通职业技术学院大学生科技创新项目 (编号: GDCP-ZX-2023-040-N6)

营。然而,轨枕由于长期处于露天状态,受雨水侵蚀、太阳曝晒、线路沉降等外部因素影响,容易导致轨枕产生裂缝。此外,伴随着我国第一条现代化高速铁路——京津城际的开通,我国高速铁路投入运营已16年有余,早期开通的高速铁路线路将逐渐进入检修维护周期,因此,我国庞大高速铁路网络的性能状态检测工作量巨大,这为人工智能技术赋能高速铁路性能检测提供了广阔的应用前景。

实际上,人工智能赋能目标检测在诸多领域已有广泛应用。为了更加精确且有效地从图像中提取熔池和飞溅的信息,毛杨坤等^[2]提出了一种基于YOLOv5的目标检测模型,实现了对成形过程图像中飞溅与熔池的实时定位与捕获。该网络具有良好的检测精度的同时,网络模型的参数量极少,更加符合工业应用的需求。为了解决海面舰船检测过程中存在的小目标模糊、目标尺度差异大、类别不均衡等问题,马淦等^[3]设计了一种动态“复制-粘贴”的数据增强方式,将其嵌入YOLOv5框架,提出了一种改进YOLOv5s的海面目标检测算法。实验表明,所提出的方法可有效缓解类别、尺寸不均衡问题。针对传统带钢表面缺陷检测技术落后、效率不高及小目标识别能力不足等问题,付强等^[4]提出了一种改进的YOLOv5s-Tiny目标检测模型,在保持模型较小计算量的同时提升带钢表面缺陷检测速度和识别精度。此外,基于人工智能技术的目标检测技术在农业^[5-7]、医学^[8]和航空^[9-10]领域也有深入应用。近年来,随着我国高速铁路运维周期的集中到来,对以人工智能为代表的智能巡检技术的需求愈发迫切,诸多学者开始聚焦人工智能技术在铁路领域的赋能运用。崔晓宁等^[11]采用将计算机视觉技术应用于双块式轨枕裂缝的智能识别研究,建立了基于YOLOv5的裂缝目标检测模型,通过图像标注、参数选取、模型构建、模型训练及模型预测等步骤完成轨枕裂缝目标检测试验。针对难以捕获隧道裂缝的全局特征的问题,周中等^[12]以YOLOv5网络框架为基础,通过融合自注意力机制和卷积神经网络提出了一种高精度隧道衬砌裂缝智能检测算法ST-YOLO。YOLO系列算法在工程实践领域的应用极大程度地提高了工作效率,且在桥梁裂缝检测^[13]等方面的应用显著改善了工程师的工作条件。然而,如何提高模型预测

的准确性和可信度仍是目前存在的主要问题,这促使学者们在改进和优化相关算法结构上做了诸多探索。以添加注意力机制为例,SE注意力机制^[14]在通道引入注意力机制,以自动学习的方式,获取特征图的每个通道的重要性,然后重点关注某些特征通道,进而提升有用的特征图通道,抑制无用的特征图通道。ECA注意力机制^[15]是通道注意力机制的一种实现形式,可以看作是SE注意力机制的改进版;ECA注意力机制模块直接在全局平均池化层之后使用 1×1 卷积层,去除了全连接层。该模块避免了维度缩减,并有效捕获了跨通道交互,且ECANet只需少数参数就能达到很好的效果。CBAM注意力机制^[16]由通道注意力机制(channel)和空间注意力机制(spatial)组成。传统基于卷积神经网络的注意力机制更关注对通道域的分析,局限于考虑特征图通道之间的作用关系。CA注意力机制是由Hou等^[17]在2021年提出的注意力模块,CA机制除了关注通道信息外,还捕获了方位和位置敏感信息,它将图像中的目标位置信息嵌入通道注意力中,以便检测模型能够更好地定位目标并进行检测,此外,CA机制相比传统的注意力机制参数量更少、更加轻量化,对设备的算力要求更低。

本文采用Labelme标注软件实现对轨枕裂缝的多边形标注,提高数据集标注的精准性,并基于YOLOv5s模型,将CA注意力机制植入其Backbone层,提高模型目标检测的准确性。最后,通过实验分析,验证了所提出的YOLOv5s-CA模型对混凝土轨枕裂缝检测的有效性。

1 算法结构

为进一步提高模型的检测精度,本文在YOLOv5s模型中植入了CA注意力机制模块。

1.1 CA注意力机制

人类在观察事物时,视觉系统会自动选择关注事物的关键信息,筛除相对不重要的部分。机器视觉技术中的注意力机制类似于人类的视觉系统,致力于捕获检测目标的某些重要特征及其位置信息。不同于仅关注局部区域的普通卷积或通道注意力(如SE),CA注意力机制通过将空间注意力分解为垂直与水平两个方向的特征编码,能够远程捕捉裂缝沿特定方向的连续性与空间长

程依赖,这对捕获蜿蜒、斜向的不规则裂缝的整体形态至关重要。

CA 模块由坐标信息嵌入和坐标注意力生成两部分组成。CA 模块将输入特征张量分解为水平和垂直两个维度的特征编码,使模型能够捕获单一方向上通道之间的长距离依赖关系,并留存另一方向的位置信息。此时,目标位置信息被保存到生成的注意力图中,然后通过乘法运算添加到输入特征图中,以进行下一步的卷积操作^[18]。

图1为CA注意力机制的结构框图,对 $C \times H \times W$ 形状的输入特征图各通道执行 avgpool,其中 C 、 H 、 W 分别为特征图的通道数、高度、宽度。使用 $(H, 1)$ 和 $(1, W)$ 的池化核,分别沿 X 轴和 Y 轴方向进行池化以完成各通道编码,最终输出 $C \times H \times 1$ 和 $C \times 1 \times W$ 形状的特征图^[19]。

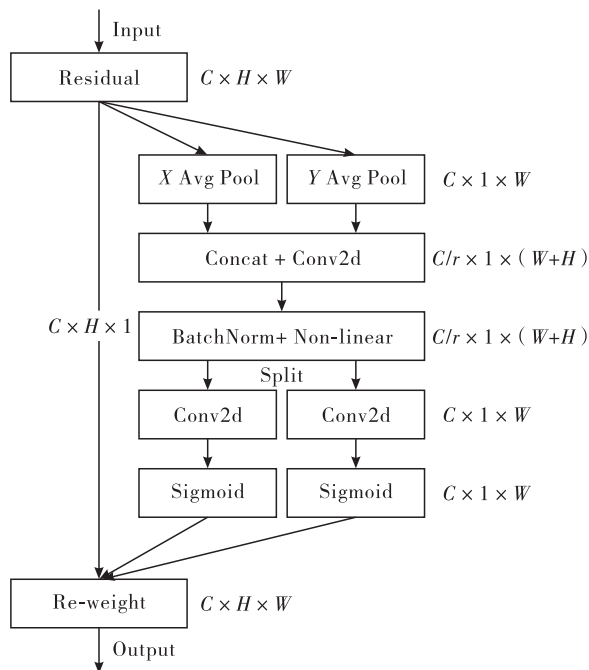


图1 CA注意力机制结构框图

1.2 改进的YOLOv5s 算法结构

YOLOv5s 模型主要包含 Backbone、Neck 和 Head 三大模块。其中,Backbone 是提取特征的网络,其作用是提取图片中的信息;Neck 层可利用 Backbone 提取的特征,提高模型的鲁棒性;Head 层可基于前面提取的特征做出预测。本文将 CA 注意力机制模块嵌入 YOLOv5s 的 Backbone 层,有助于更精细地处理和利用特征信息,进而提升目标检测的准确率。改进后的 YOLOv5s 模型网络结构如图2所示。

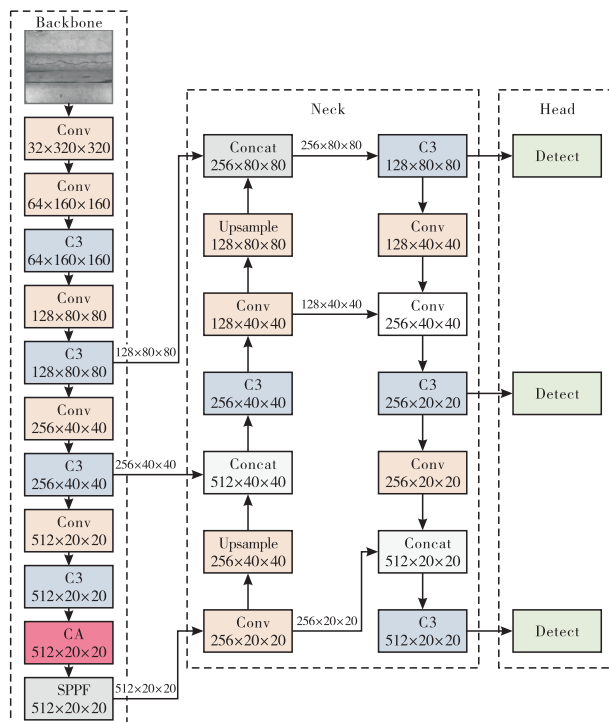


图2 改进的YOLOv5s网络结构图

CA 注意力机制是一种结合了卷积操作和注意力机制的视觉注意力变体,在计算机视觉任务中表现出色。首先,CA 机制利用卷积操作从输入序列中提取局部特征信息,这些卷积滤波器可以有效捕获输入数据中的空间局部结构。然后,生成的特征图被输入到一个注意力机制中。在注意力机制中,每个特征图都被赋予一个权重,此权重表示该特征图相对于其他特征图的重要性。这种加权过程能够使得模型更加关注重要的局部特征,并降低对无关信息的敏感度。最终,将特征图与其对应的注意力权重相乘,再进行池化或求和操作,以将局部信息与全局信息结合起来,此时得到的加权特征表示同时包含了局部和全局的重要信息。

CA 注意力机制结合卷积运算和注意力机制,能够高效捕捉序列数据中的局部和全局信息,以此提升模型性能和泛化能力。

2 实验与结果分析

为了验证本文改进的 YOLOv5s 模型对混凝土轨枕裂缝检测的有效性的检测效果,将添加 CA 注意力机制的模型与原模型进行对比实验。

2.1 数据集

由于高速铁路轨枕均为混凝土材质,采用的

数据集包含1 000张高铁轨枕裂缝和混凝土裂缝图片,数据集按照6:2:2划分为训练集、验证集和测试集。图3是部分具有代表性的轨枕裂缝和混凝土裂缝图片。

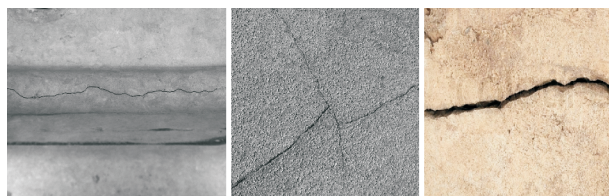


图3 轨枕裂缝和混凝土裂缝图片

2.2 轨枕裂缝标注

采用Labelme创建多边形标注框进行轨枕裂缝的标签制作,为轨枕裂缝检测提供数据基础,标注过程见图4。要注意的是,使用Labelme标注完成后软件会自动生成json格式文件(每张图对应一个json文件);后续要将其转换为txt格式方能用于模型训练。

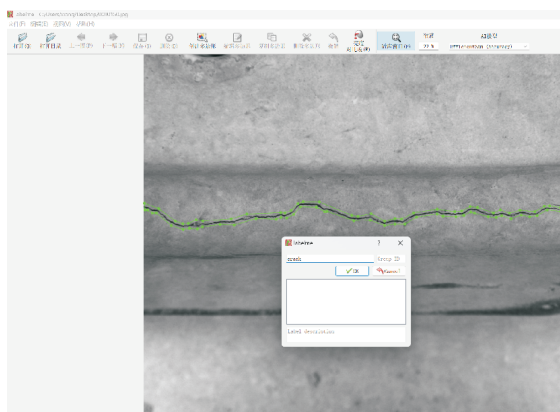


图4 轨枕裂缝标注

2.3 评价指标

根据目标真实标注框和模型预测框的相对关系,目标检测模型的预测结果分为四种:真阳性(TP)、假阳性(FP)、真阴性(TN)、假阴性(FN)。以检测精度(PA)、召回率(Recall)、mAP@0.5和mAP@0.5:0.95作为模型的评价指标,PA、Recall、mAP的计算公式如下:

$$m_{PA} = \frac{m_{TP} + m_{TN}}{m_{TP} + m_{FP} + m_{FN} + m_{TN}}$$

$$p = \frac{m_{TP}}{m_{FA}}$$

$$m_{Recall} = \frac{m_{TP}}{m_{TP} + m_{FN}}$$

$$m_{mAP} = \frac{\sum_{i=1}^C \sum_{k=1}^N p(k) m_{Recall}(k)}{C}$$

式中, p 为查准率; C 为目标检测任务中的类别总数; i 为类别编号; k 为图片编号; N 为图片总数。

2.4 实验环境

实验环境参数如表1所示。

表1 实验环境参数

配置	参数
CPU	Intel i9-13900HX, 2.2GHz
内存	16G
GPU	NVIDIA GeForce RTX 4060
GPU 显存	8GB
GPU 算力值	15TFLOPS
编程语言	Python 3.8
	Pytorch 1.12.1
	PyYAML 5.3.1
深度学习框架及依赖库	Numpy 1.23.5
	Tqdm 4.64.0
	Tensorboard 2.14.0

2.5 实验结果

图5为YOLOv5s模型与YOLOv5s-CA模型的mAP@0.5在各Epoch上的训练结果,表2为两模型各评价指标的结果对比。由图5和表2可知:

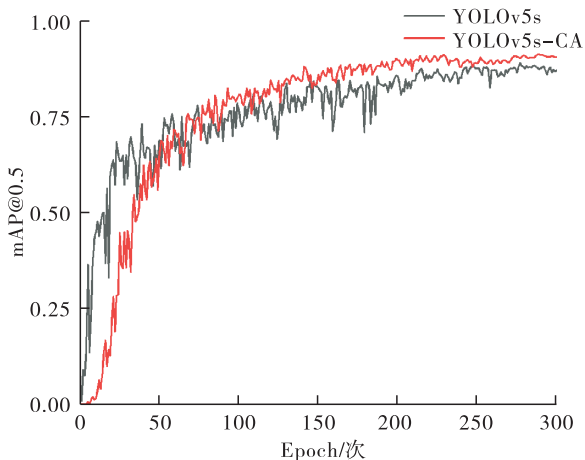


图5 两模型mAP@0.5结果对比

表2 不同模型检测结果对比

模型	检测精度/%	召回率/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv5s	90.88	79.68	86.87	63.49
YOLOv5s-CA	92.68	84.12	90.41	72.88

(1) YOLOv5s与YOLOv5s-CA模型在训练初期波动较大,训练200次之后逐步趋于稳定,且YOLOv5s-CA模型对混凝土轨枕裂缝的平均检测精度高于YOLOv5s。

(2) YOLOv5s 模型对混凝土轨枕裂缝的检测精度为 90.88%, 平均检测精度为 86.87%。加入 CA 注意力机制模块后, 模型的特征提取能力提升, 轨枕裂缝的检测精度提高了 1.8%, mAP@0.5 提高了 3.54%, mAP@0.5 : 0.95 提高了 9.39%, 召回率提高了 4.44%。

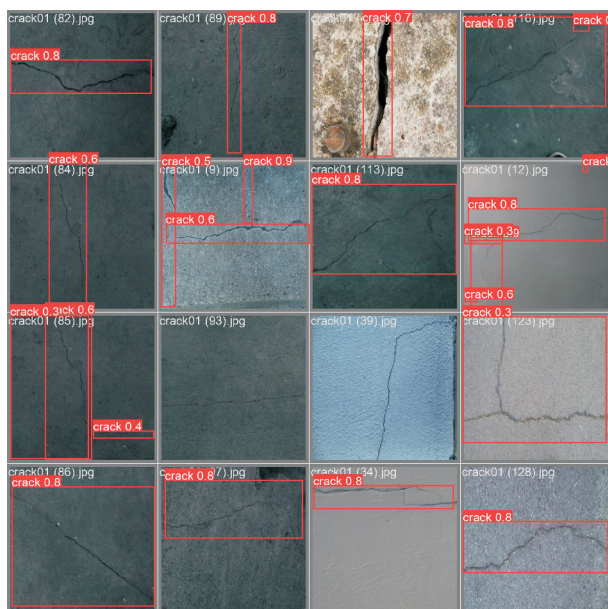
(3) 通过添加 CA 注意力机制, 实现了对 YOLOv5s 模型各评价指标的全面提升, 说明改进后的模型对轨枕裂缝目标检测具有较好的适用性。

2.6 实验结果可视化

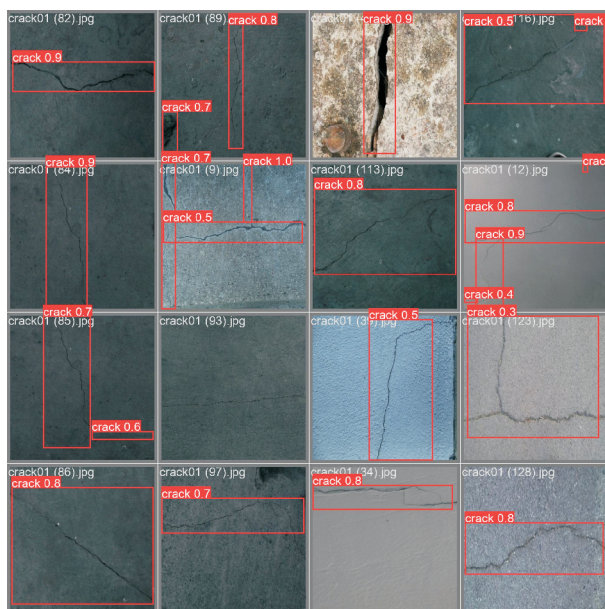
为了体现模型的泛化能力和强鲁棒性, 选取

不同光照和不同颜色下的混凝土裂缝检测结果进行展示, 其中, 图 6 (a) 为 YOLOv5s 模型的轨枕裂缝检测结果, 图 6 (b) 为 YOLOv5s - CA 模型的轨枕裂缝检测结果。

由图 6 可知: YOLOv5s 模型在不同条件下的原始裂缝检测的置信度偏低且存在漏检和过检现象; 而 YOLOv5s - CA 模型对轨枕裂缝的检出率和置信度更高, 体现了 YOLOv5s - CA 模型的良好泛化能力和鲁棒性, 该模型可在实际混凝土轨枕裂缝检测与智能识别中发挥重要作用。



(a) YOLOv5s 模型轨枕裂缝检测结果



(b) YOLOv5s - CA 模型轨枕裂缝检测结果

图 6 不同模型检测结果对比

3 结论

本研究针对轨枕裂缝形状的不规则性, 采用 Labelme 对数据集进行多边形标记, 并基于 YOLOv5s 模型, 在其网络结构 Backbone 层植入 CA 注意力机制, 进而提高了模型的检测精度。实验表明, YOLOv5s - CA 模型能够准确、快速地检测出不同光照和颜色条件下的混凝土轨枕裂缝。实际应用中检测精度达到 92.68%, 召回率达到 84.12%, mAP@0.5 达到 90.41%, mAP@0.5 : 0.95 达到 72.88%, 性能高于 YOLOv5s 模型, 对混凝土轨枕裂缝检测具有较高的实际应用价值。下一步将基于 YOLOv5s 模型研究 SE、ECA、CBAM 等注意力机制对轨枕裂缝智能识别的效果, 并对模型做轻量化设计以降低计算量和内存需求。

参考文献:

- [1] 国家铁路局. 2024 年铁道统计公报[R/OL]. (2025-06-06). <https://www.nra.gov.cn/xwzx/zlzx/hytj/202506/P020250606511097164786.pdf>.
- [2] 毛杨坤, 段现银, 林昕, 等. 基于目标检测的选区激光熔融成形过程熔池与飞溅监测[J]. 机械工程学报, 2022, 59(9):335-348.
- [3] 马淦, 谷雨, 彭冬亮. 结合改进 YOLOv5s 和动态数据增强的海面舰船检测[J]. 计算机工程, 2025, 51(9):294-305.
- [4] 付强, 朱传军, 梁泽启. 基于机器识别的带钢表面缺陷检测研究[J]. 机床与液压, 2024, 52(10):194-200.
- [5] 郭阳阳, 洪文浩, 丁屹, 等. 基于坐标注意力机制和 YOLOv5s 模型的山羊脸部检测方法[J]. 农业机械学报, 2023, 54(7):313-321.

- [6] WANG X, LIU J. Multiscale parallel algorithm for early detection of tomato gray mold in a complex natural environment [J]. *Frontiers in Plant Science*, 2021, 12: 620273.
- [7] WANG Y, SUN Y, LIU Z, et al. Dynamic graph CNN for learning on point clouds [J]. *ACM Transactions on Graphics*, 2019, 38(5): 1-12.
- [8] NAIRA E, WZEL A, MOHAMMED E. A multi-class brain tumor grading system based on histopathological images using a hybrid YOLO and RESNET networks [J]. *Scientific Reports*, 2024, 14(1), 4584-4596.
- [9] 李子豪, 王正平, 贺云涛. 基于自适应协同注意力机制的航拍密集小目标检测算法 [J]. *航空学报*, 2023, 44(13): 244-254.
- [10] HE W, LI C, NIE X, et al. Recognition and detection of aero-engine blade damage based on Improved CascadeMask R-CNN [J]. *Applied Optics*, 2021, 60(17): 5124-5133.
- [11] 崔晓宁, 王起才, 李盛, 等. 基于 YOLO-v5 的双块式轨枕裂缝智能识别 [J]. *铁道学报*, 2022, 44(4): 104-111.
- [12] 周中, 闫龙宾, 张俊杰, 等. 基于自注意力机制与卷积神经网络的隧道衬砌裂缝智能检测 [J]. *铁道学报*, 2024, 46(9): 182-192.
- [13] LI T, LIU G, TAN S. Superficial defect detection for concrete bridges using YOLOv8 with attention mechanism and deformation convolution [J]. *Applied Sciences*, 2024, 14(13): 1-18.
- [14] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, 7132-7141.
- [15] WANG Q, WU B, ZHU P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, 11534-11542.
- [16] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018: 3-19.
- [17] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021: 13713-13722.
- [18] 黄承孝, 刘娉婷. 基于 CA-YOLOv5s 的水下目标识别方法研究 [J]. *海洋技术学报*, 2024, 43(3): 17-24.
- [19] 仲重光. 集成电路晶圆图像针印缺陷检测研究 [D]. 镇江: 江苏科技大学, 2024.

(上接第 15 页)

(1) 软土路基扩建时, 扩建范围内路基土作为新增附加荷载是导致路基沉降的主要因素。

(2) 软土路基扩建并运营 15 年后, 对比分析不同方案处理的总沉降、工后沉降和差异沉降, 得出地基采用管桩复合地基处理的效果要优于水泥搅拌桩。扩建路基填料采用气泡轻质土对差异沉降的控制效果要优于路基填土。

(3) 珠三角地区软基段高速公路路基扩建时, 采用管桩复合地基桩 + 气泡轻质土路基填筑的处理方案, 可以有效控制总沉降、工后沉降和差异沉降, 可为类似工程提供技术参考。

参考文献:

- [1] 廖朝华, 王家强. 高速公路扩建工程主要技术问题及对策 [C]// 中国公路学会. 2005 年全国公路勘察设计技术交流会议论文集. 武汉: 中交第二公路勘察设计研究院, 2005: 15-22.
- [2] 胡志强. 路基加宽不均匀沉降及其对路面结构应力影响分析 [D]. 长沙: 湖南大学, 2019.
- [3] 曹海莹, 窦远明, 刘熙媛, 等. 影响软基沉降因素的分解与研究 [J]. *河北工业大学学报*, 2005(6): 82-86.
- [4] 路炜. 高速公路扩建工程差异沉降控制技术 [J]. *交通世界*, 2020(25): 76-77.
- [5] 于恒, 汪益敏, 王兆阳. 旧路软基处治欠佳道路扩建后路面裂缝形成机理研究 [J]. *中外公路*, 2017, 37(2): 18-23.
- [6] 黄先刚. 公路改扩建中新旧路基不均匀沉降的有限元模拟及处治研究 [J]. *公路交通科技 (应用技术版)*, 2020, 16(1): 74-75.
- [7] 杨丽珍. 软土地区高速公路扩建工程复合地基差异沉降数值模拟 [D]. 广州: 华南理工大学, 2010.
- [8] 叶红英, 张再武, 张艺. 软土地基上高速公路扩建工程的三维有限元模拟 [J]. *交通科技*, 2017(2): 25-27.
- [9] 中华人民共和国交通运输部. 公路路基设计规范: JTG D30—2015 [S]. 北京: 人民交通出版社股份有限公司, 2015.
- [10] 周旺. 泡沫轻质土在道路扩建工程中的应用分析 [J]. *西部交通科技*, 2023(5): 42-44, 78.
- [11] 王鹏程. 泡沫轻质土在道路扩建工程中的应用 [J]. *工程建设与设计*, 2022(16): 188-190.