

DOI:10.13409/j.cnki.jdpme.20250430038

基于改进YOLOv8的隧道渗漏水轻量化检测*

田城航¹, 赵力萌³, 李永刚⁴, 张方剑⁴, 仇文岗^{1,2}, 孙伟鑫^{1,5}, 丁选明⁶

(1. 重庆大学土木工程学院, 重庆 400045; 2. 山区土木工程安全与韧性全国重点实验室, 重庆 400045; 3. 广州地铁建设管理有限公司, 广东 广州 510010; 4. 中铁十一局集团有限公司, 湖北 武汉 430061; 5. 重庆大学航空航天学院, 重庆 400045; 6. 湖南工程学院土木工程智能防灾减灾与生态修复湖南省重点实验室, 湖南 湘潭, 411104)

摘要: 针对隧道渗漏水检测中传统方法效率低、现有深度学习模型参数量大且实时性不足的问题, 提出一种基于改进YOLOv8n-seg的轻量化实例分割模型。通过引入Coordinate Attention(CA)注意力机制增强目标区域特征权重, 采用MobileNetV4替换主干网络以降低计算复杂度, 并结合EfficientHead分割头优化特征解码效率, 显著提升了模型在复杂背景下的检测精度与推理速度。实验基于三维激光扫描技术构建的隧道渗漏水数据集(包含3140张增强图像), 通过消融实验验证了各模块的有效性: CA机制使平均精度(AP)提升0.82%, MobileNetV4降低参数量43.2%的同时提升AP至81.21%, EfficientHead分割头进一步优化分割细节。联合改进后, 模型AP达83.21%, F1值提升至78.53%, 参数量仅1.96 M, 推理速度达355.2 FPS, 较原YOLOv8n-seg提升6.6%。对比实验表明, 改进模型在轻量化指标(参数量、GFLOPs)显著优于Mask R-CNN等主流模型, 且精度接近两阶段方法, 满足隧道渗漏水实时检测需求。研究为隧道结构健康监测提供了高效、可靠的轻量化解决方案, 具有工程应用价值。

关键词: 隧道病害; 深度学习; 病害识别; 实例分割; 轻量化检测

中图分类号: U231.94; U457 **文献标识码:** A **文章编号:** 1672-2132(2025)06-1421-13

Lightweight Detection of Tunnel Water Leakage Based on Improved YOLOv8 Model

TIAN Chenghang¹, ZHAO Limeng³, LI Yonggang⁴, ZHANG Fangjian⁴,
ZHANG Wengang^{1,2}, SUN Weixin^{1,5}, DING Xuanming⁶

(1. School of Civil Engineering, Chongqing University, Chongqing 400045, China; 2. State Key Laboratory of Safety and Resilience of Civil Engineering in Mountain Area, Chongqing 400045, China; 3. Guangzhou Metro Construction Management Co., Ltd, Guangzhou 510010, China; 4. China Railway 11th Bureau Group Co., Ltd, Wuhan 430061, China; 5. College of Aerospace Engineering, Chongqing University, Chongqing 400045, China; 6. Hunan Provincial Key Laboratory of Intelligent Disaster Prevention-Mitigation and Ecological Restoration in Civil Engineering, Hunan Institute of Engineering, Xiangtan 411104, China)

Abstract: To address the challenges of low efficiency in traditional methods for tunnel water leakage detection and the large number of parameters and insufficient real-time performance of existing deep learning models, this study proposed a lightweight instance segmentation model based on an improved YOLOv8n-seg. The coordinate attention (CA) mechanism was introduced to enhance feature represen-

* 收稿日期: 2025-03-27; 修回日期: 2025-06-09

基金项目: 广州地铁集团有限公司科研委托项目(JT204-100111-23001)、重庆铁路投资集团有限公司科研委托项目(CSTB2022TIAD-KPX0101)、中铁二局集团有限公司科研委托项目(N2023G045)资助

作者简介: 田城航(2001—), 男, 硕士研究生。主要从事城市地下工程研究。E-mail: tch2001@163.com

通信作者: 仇文岗(1983—), 男, 教授, 博士。主要从事城市地下工程研究。E-mail: zhangwg@cqu.edu.cn

tation in target regions. The backbone network was replaced with MobileNetV4 to reduce computational complexity, and the EfficientHead segmentation head was incorporated to improve the efficiency of feature decoding. These improvements collectively enhanced both detection accuracy and inference speed in complex environments. The experiments were conducted on a tunnel water leakage dataset constructed using 3D laser scanning (including 3 140 enhanced images). Ablation experiments were employed to validate the effectiveness of each module. The CA mechanism increased the average precision (AP) by 0.82%, MobileNetV4 increased the AP to 81.21% while reducing the number of parameters by 43.2%, and the EfficientHead further optimized segmentation details. After joint improvements, the model achieved an AP of 83.21%, an F1-score of 78.53%, a number of parameters of 1.96M, and an inference speed of 355.2 FPS, representing a 6.6% increase over the original YOLOv8n-seg. Comparative experiments demonstrated that the proposed model significantly outperformed mainstream models such as Mask R-CNN in lightweight indicators (number of parameters and GFLOPs), while achieving accuracy comparable to that of two-stage methods, thereby meeting the real-time detection requirement for tunnel water leakage. This study provides an efficient and reliable lightweight solution for structural health monitoring in tunnels, offering practical value for engineering applications.

Keywords: tunnel defect; deep learning; defect detection; instance segmentation; lightweight detection

0 引言

随着我国城市化进程的不断加快,隧道作为城市轨道交通和公路交通的重要组成部分,在缓解交通压力、促进区域经济发展等方面发挥着举足轻重的作用。然而,长期服役、环境侵蚀及运营条件的多重影响,使得隧道结构逐渐出现裂缝、渗漏水等多种病害问题,严重威胁到隧道的结构安全与正常运行^[1-2]。其中,隧道渗漏水不仅诱发混凝土碳化与钢筋锈蚀,在寒冷地区更易因冻融循环引发衬砌内部微裂纹扩展,如未及时发现并处理将导致严重后果。传统的隧道病害人工检测方法存在主观性强、效率低、易遗漏等问题,在复杂场景下难以实现高准确的识别^[3-4]。

近年来,深度学习作为人工智能领域的重要分支,凭借其在大规模数据学习、自动特征提取及高效识别等方面的优势,在图像处理、目标检测、自然语言处理等多个领域取得了显著成效。基于深度学习的方法能够自动挖掘数据中蕴含的隐含特征,打破传统手工设计特征的局限,因而在土木工程中得到了广泛推广^[5-7]。针对隧道病害检测任务,一些研究已尝试将卷积神经网络、目标检测和实例分割等技术应用于隧道衬砌病害的自动识别,并初步取得了较为理想的实验效果。Z.K. Ding等^[8]提出了一种基于无人机和VGG16算法的裂缝检测方法,

以量化裂缝的长度和宽度,并利用深圳地铁9号线某段的实测数据验证了模型的正确性。A.F. Zhu等^[9]选择VGG16和ResNet34构建自动识别模型对隧道缺陷进行识别,结果表明ResNet34比VGG16有更快的收敛速度,更高的准确率,更短的训练时间。A. Bani Mustafa等^[10]使用ResNet50搭建了裂缝视觉标志监测的系统,经训练该模型的准确性达到70%。

尽管深度学习网络在众多领域展示出强大的特征提取能力,但直接应用于隧道病害检测时仍面临诸多挑战。首先,隧道环境复杂,病害目标通常体积较小且形态多样,易被复杂背景所干扰;其次,现有网络在全局信息捕捉和局部细节识别上存在一定局限,难以全面适应隧道衬砌病害的实际情况^[11-13]。为了提升对隧道渗漏水的检测精度和鲁棒性,许多学者对传统深度学习网络结构进行有针对性的改进和优化。Y.D. Xue等^[14]在Mask R-CNN的基础上,结合迁移学习优化特征提取,并引入级联策略提高分割质量,改进模型识别准确率远超Mask R-CNN。王宝坤等^[15]利用残差单元和注意力机制提出新型的语义分割模型SU-ResNet++,对隧道多种表观病害完成了高精度分割。S. Zhao等^[16]通过优化主干网络、改进RPN、调整损失函数完成了对Mask R-CNN的改进,提升盾构隧道渗漏水检测的精度和分割质量。A.H. Ouyang等^[17]使用的改进DeepLab V3+语义分割模型,引入轻量级骨

干网络以提高检测效率,并采用平滑融合预测方法优化高分辨率图像推理,检测结果清晰展现了病害结构。L.J. Xiong等^[18]增强特征提取能力,并优化目标区域分割,采用改进的全卷积网络进行渗漏水像素级识别,提高了复杂背景下的检测精度。此外,近年来YOLO系列算法因其训练速度快,模型小等优势被用于隧道病害检测中,并取得较好的病害检测效果。宋娟等^[19]通过改进主干网络,注意力机制和卷积模块,使用改进的TDD-YOLO实现了对隧道病害的高精度检测。S.Q. Duan等^[20]提出了一种基于改进YOLOv5的隧道衬砌裂缝实时检测模型,在低光照、低对比度等高噪声环境中实现了高精度的快速检测,显著优于传统方法。朱家松等^[21]以YOLOv5模型为基础,结合坐标注意力和BiFPN,提升地铁隧道衬砌病害检测精度至89.2%,模型轻量化便于边缘部署。X.Q. Xu等^[22]提出基于YOLOv8算法和3D成像的隧道管片裂缝宽度识别方法,结合激光测距校正像素分辨率,检测误差率低于4%,显著优于传统方法。

上述学者从多方面对既有算法进行了优化与改进,实现了更高的病害识别性能。然而,当前研究在聚焦于模型高性能的同时往往会忽视模型的大小与训练速度,反而会出现超大模型推理的情

况。这种情况对训练硬件设备提出了更高的要求,也不利于病害检测系统与平台的开发。针对此点不足,基于快速分割的算法YOLOv8n-seg进行相关研究,通过加入注意力机制,替换高性能分割头与轻量化主干网络,以实现盾构隧道的精准识别与检测。

1 算法介绍

本章主要介绍提出的改进模型的关键构成模块及其工作原理,包括YOLOv8n-seg架构、Coordinate Attention注意力机制、MobileNet轻量化主干网络、EfficientHead分割头以及各模块的有机融合与改进策略。

1.1 YOLOv8n-seg

YOLOv8n-seg是YOLOv8系列中针对轻量级实例分割任务设计的模型,其整体架构主要包括预处理模块、主干特征提取网络、特征融合层以及输出模块(如图1所示)。与传统目标检测模型不同,该模型在兼顾实时推理性能与分割精度的基础上,对网络内部各模块进行了多处优化,从而使其在资源受限的边缘设备上具有更好的应用适应性。

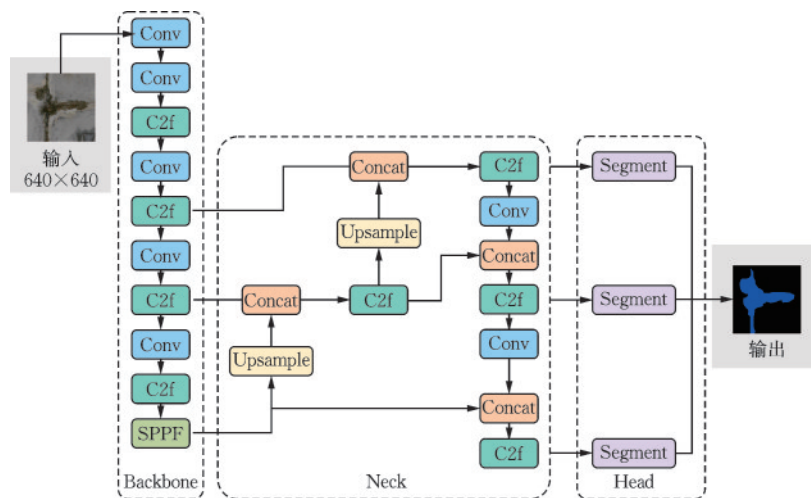


图1 YOLOv8n-seg网络结构
Fig.1 YOLOv8n-seg network structure

在输入阶段,原始图像首先经过归一化与尺寸调整,常用输入尺寸为416x416或640x640,以确保后续特征提取过程在统一尺度下进行。主干特征提取部分采用改进的C2f结构,通过跨阶段的特征融合显著提升了信息传递效率和梯度流动性。

与传统的CSP结构相比,C2f模块在降低重复计算的同时,有效增强了局部细节与全局语义的表达能力。经过多层卷积、批量归一化及SiLU激活操作后,主干网络能够输出多个尺度的特征图,为检测和分割任务提供了坚实的信息基础。

在特征融合阶段,模型对主干网络输出的多尺度特征进行进一步整合,采用了FPN与PAN相结合的策略。具体而言,FPN通过自顶向下的路径将高层语义信息传递给低层,从而强化了细粒度特征;同时,PAN则利用自底向上的路径聚合低层细节信息,提升了对小目标的识别能力。与传统多分支堆叠模块不同,YOLO v8n-seg采用了改良的FPN与PAN融合模块,在高效计算的前提下实现了各尺度特征的充分交互,并确保输出特征图尺寸与主干网络保持一致,针对实例分割任务,可细化融合后的特征图空间对齐。

在输出模块中,网络设计了并行的目标检测分支和实例分割分支。目标检测分支主要负责预测目标类别、置信度及边界框,通过多层卷积生成候选框,并借助非极大值抑制技术去除冗余检测结果;实例分割分支则在高分辨率特征图上进行上采样,并通过轻量化掩码头对目标边缘进行细化,从而获得精细的像素级分割结果。为进一步提升掩码预测的精度,该分支引入了动态卷积机制,以更好地适应目标形态的多样性^[23-24]。

综上所述,YOLOv8n-seg通过主干网络的轻量化改进、多尺度特征的高效融合以及检测与分割输出模块的协同设计,在复杂场景下能够有效提升对小目标和细微病害的识别能力,同时保证实时推理性能,为精细分割任务提供了坚实的技术支撑。

1.2 Coordinate attention 注意力机制

Coordinate Attention (CA) 注意力机制是一种结合了空间位置信息和通道信息的创新型机制。与传统的注意力机制(如CBAM和SE)采用全局最大池化或全局平均池化方式,导致空间信息丢失不同,CA注意力机制通过引入坐标信息来增强对空间位置的感知,从而能更有效地提升目标区域的权重。

如图2所示,CA注意力机制的基本结构包括两个并行的操作阶段。首先,输入特征图分别在高和宽两个方向上执行全局平均池化,提取每个方向上的全局特征信息。其中,在高度方向(Y轴)进行全局平均池化可保留图像在水平方向上的结构特征,而在宽度方向(X轴)执行全局平均池化则捕捉垂直方向的特征变化。这种将池化操作沿坐标轴分别展开的方式,相当于为每个像素位置编码了其在二维空间中的方向性位置信息。通过这种方式,CA

注意力机制实现了对图像中空间坐标特征的显式建模,从而能在后续注意力计算中更加精准地强化目标区域的响应。接着,这两个方向的信息会合并,并转置到相同维度上进行堆叠。通过卷积、标准化以及激活函数进一步处理特征,然后再分为两个并行阶段。

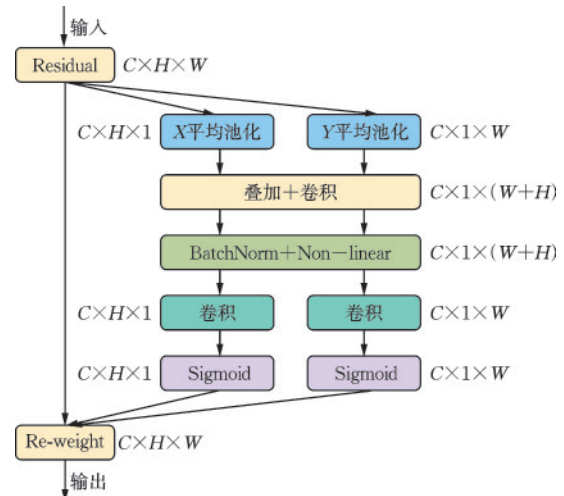


图2 Coordinate attention 注意力机制结构

Fig.2 Coordinate attention mechanism structure

然后, 1×1 卷积用来调整通道数,并通过 Sigmoid 激活函数获取高和宽方向上的注意力权重。这两个方向分别对应两个并行的注意力生成分支:宽度方向的注意力分支负责对水平方向上的特征进行建模,增强图像中目标在横向结构中的显著性;高度方向的注意力分支则强调垂直方向上的目标特征,从而在垂直结构上提升响应精度。此外,由于两个方向的注意力权重分别作用于不同的通道维度,它们在功能上也具备通道注意力的调节特性。因此,CA机制是一种融合了方向性空间注意力与通道注意力的复合注意力模块,能够同时关注目标的空间布局和语义通道分布,有效提升图像表征能力。

最终,高度方向和宽度方向上的注意力图分别与原始输入特征图进行广播并逐元素相乘,形成最终输出特征图。具体地,高度注意力图沿通道维度扩展至与输入特征图形状一致,作用于垂直方向的重要区域;宽度注意力图则沿通道维度和高度方向扩展,强化水平方向的显著特征。通过这种方式,两个方向的空间注意力得以同时发挥作用,且在计算结构中保持联动。因此,CA注意力机制通过方向性空间建模与通道级特征调节的并行结构,实现了空间注意力和通道注意力的深度融

合。这种设计不仅保留了空间位置依赖性,还强化了语义维度的辨别能力,显著提升了网络对局部与全局目标的识别精度。

综上,CA注意力机制通过巧妙设计的结构,在轻量化网络中引入了对空间坐标的感知能力,并融合了通道与空间注意力的优势,已广泛应用于图像分类、目标检测与语义分割等多种视觉任务中,取得了良好的性能表现。

1.3 MobileNetV4网络

MobileNetV4的整体结构可以分为三个主要部分:初始层(stem)、一系列通用倒置瓶颈模块(Universal Inverted Bottleneck Blocks),以及最后的全局聚合与分类头。输入图像首先经过一个标准的 $3 \times$

3卷积层,这一层不仅完成了最初的特征提取,还通过步幅实现了初步的下采样,从而为后续模块提供了足够的感受野和较高分辨率的特征表示。

如图3所示,在初始层之后,网络进入主体部分,即多个通用倒置瓶颈模块。每个倒置瓶颈模块通常由三个连续步骤构成:首先是扩展阶段,通过 1×1 卷积将低维输入扩展到较高的通道数,这一步骤旨在增加网络的表达能力;随后是深度卷积阶段,采用深度可分离卷积对扩展后的特征进行局部特征提取,这种分解方式大幅度降低了计算成本,同时能够捕捉细粒度的局部信息;最后,投影阶段通过另一组 1×1 卷积将特征映射回低维空间,并利用残差连接将输入直接与输出相加,从而实现梯度的有效传递,提升训练稳定性。

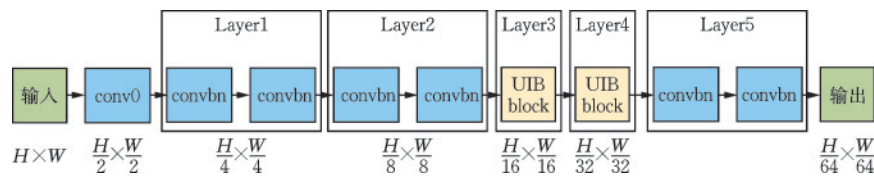


图3 MobileNetV4网络结构

Fig.3 MobileNetV4 network structure

为了弥补纯卷积网络在捕捉长程依赖和全局上下文信息方面的不足,MobileNetV4在部分倒置瓶颈模块中集成了轻量级的注意力机制,如多查询注意力模块和Squeeze-and-Excitation模块。这些注意力机制能够在低额外计算量下动态调整各通道的重要性,并捕捉到全局信息,从而进一步增强特征表达的丰富性和判别能力。

此外,MobileNetV4的设计过程中还采用了神经架构搜索(NAS)技术,对各个模块的参数和结构进行了自动化优化,使得整个网络在保持轻量级和高效计算的同时,依然能够达到较高的准确率和良好的泛化性能。总体而言,MobileNetV4通过巧妙地结合卷积、倒置瓶颈、注意力机制以及自动化搜索优化的策略,构建了一种既能高效捕捉局部细节,又能融合全局信息的新型网络架构,特别适用于资源受限的移动和嵌入式设备。

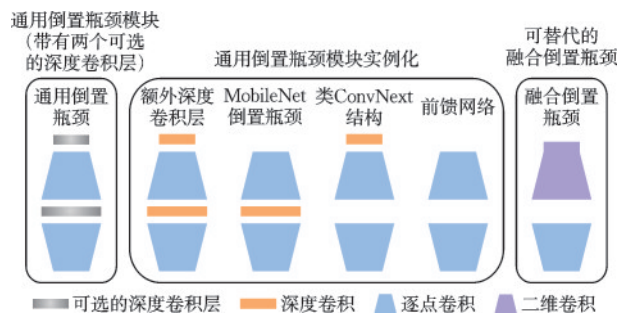


图4 倒置瓶颈模块

Fig.4 Universal inverted bottleneck (UIB) blocks

在经过多个倒置瓶颈模块后,网络将得到的特征图通过全局平均池化层进行空间降维,生成一个全局描述向量。最后,一个全连接(或线性)层将这一向量映射到目标类别空间,输出最终的分类结果或用于其他任务的预测结果。

1.4 EfficientHead分割头

EfficientHead分割头通过精简结构可以提升计算效率,同时保持较高的分割精度。该模块广泛应用于需要高效推理和实时响应的计算机视觉任务中,特别是在资源受限的环境中,能够有效提高性能。

其主要特点体现在轻量化卷积操作,采用小卷积核(如 3×3 或 1×1 卷积),有效减少了模型的参数量,同时降低了计算负担。轻量化的设计使得EfficientHead能够在保证精度的前提下,显著提升推理速度。此外,EfficientHead还具有高效的特征处理机制,通过优化特征提取流程,EfficientHead能够在较低计算成本下实现目标分割的准确性。这种

结构尤其适合应用于需要快速响应的任务,如视频流处理或嵌入式设备。

1.5 改进算法

以YOLOv8n-seg分割网络为基础,在网络中添

加CA注意力机制,借鉴MobileNetV4网络的轻量化设计替换主干网络,同时使用结构较为精简的EfficientHead分割头提升计算效率,保持分割精度,提出了一种轻量化设计的改进YOLOv8n-seg算法,该算法的网络结构如图5所示。

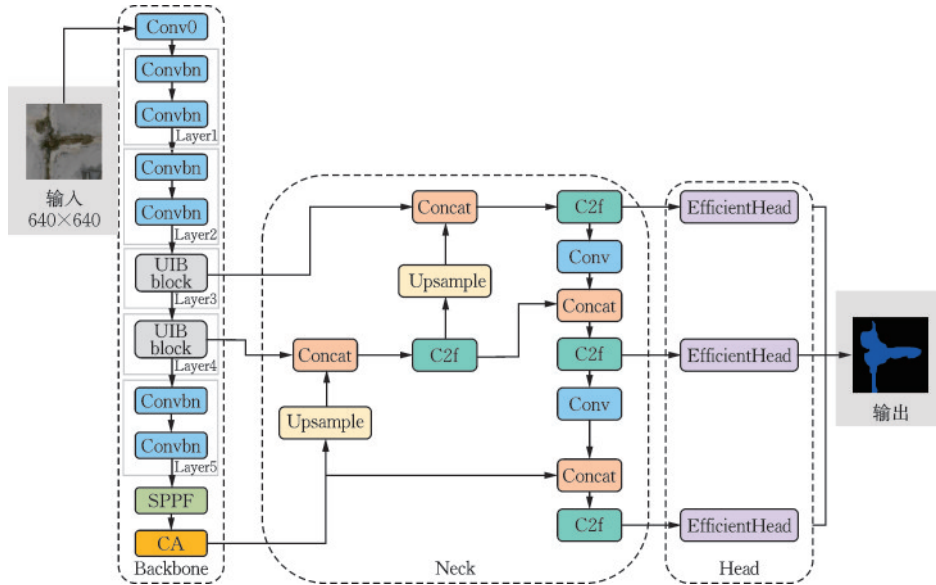


图5 改进YOLOv8网络结构

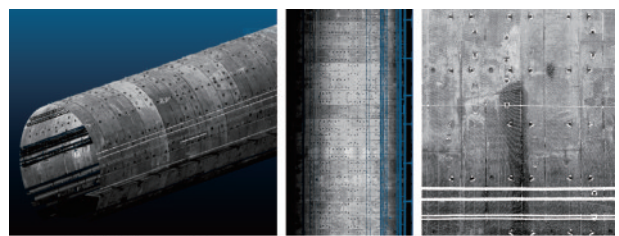
Fig.5 Improved YOLOv8n-seg network structure

2 病害数据集建立

2.1 数据获取与处理

本研究所构建的病害数据集主要来源于广州某地铁盾构隧道,利用三维激光扫描技术采集点云数据,最终生成3140张渗漏水病害灰度图像,灰度图像为统一固定尺寸640×640。如图6(a)所示,采集过程中,通过高精度激光扫描设备获取隧道内部的三维空间信息,形成密集的点云数据,为后续病害检测提供了可靠的物理依据。获得点云数据后,首先对其进行预处理,包括噪声滤除、坐标校正及异常值剔除等操作,以确保数据的准确性和连续性。随后,通过特定算法将处理后的点云数据转换为灰度图像,利用不同灰度级别来反映隧道内表面病害的分布和特征,图7列出了灰度图像生成的具体流程。采用统计滤波方法进行降噪处理,通过分析局部像素分布特征有效消除离散噪声点,同时应用自适应直方图均衡化技术增强图像对比度,显著改善了渗漏水区域的视觉显著性。所有图像统一

重采样至640×640像素的标准尺寸,确保输入数据规格一致。该预处理流程在保持图像关键特征的前提下,有效提升了后续检测的准确性和鲁棒性,为后续目标检测与实例分割模型的训练提供了坚实的数据支持。



(a) 原始点云三维图 (b) 点云展开 (c) 点云灰度图像

图6 盾构隧道点云数据

Fig.6 Shield tunnel point cloud data

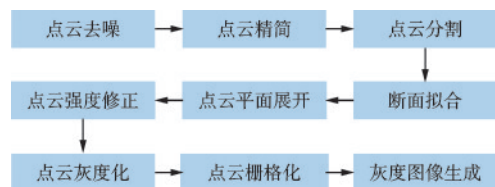


图7 灰度图像生成流程

Fig.7 Grayscale image generation process

2.2 数据集组成

通过点云图像处理得到了600多张图像。为了增强数据多样性,避免过度拟合和增加模型的鲁棒性,利用图像增强技术扩充数据集,如图8所示。数据集扩充有效解决因采集环境差异导致的样本不平衡问题。

根据8:1:1的比例将扩充后的数据集划分为训练集、验证集和测试集,划分后的渗漏水数据集数量见表1。

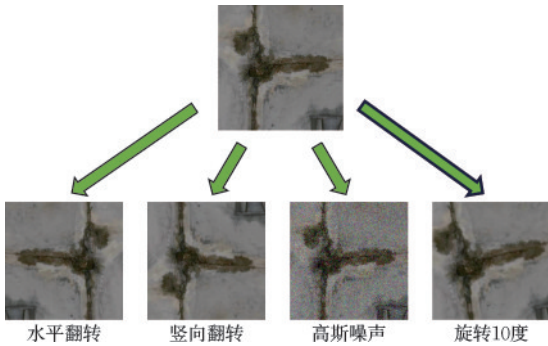


图8 病害数据集扩充方法

Fig.8 Methods for expanding defect datasets

表1 隧道渗漏水病害数据集分布

Table 1 Distribution of tunnel water leakage defect dataset

	N_{tot}	N_{tra}	N_{val}	N_{tes}
渗漏水	3 140	2 512	3 14	3 14

2.3 数据集标注

采用Labelme标注软件对隧道渗漏水病害图像进行标注,如图9所示,该软件生成的标注文件通常为COCO格式的JSON格式文件,每个文件包含了图像中对象的类别、边界框或多边形坐标等详细信息

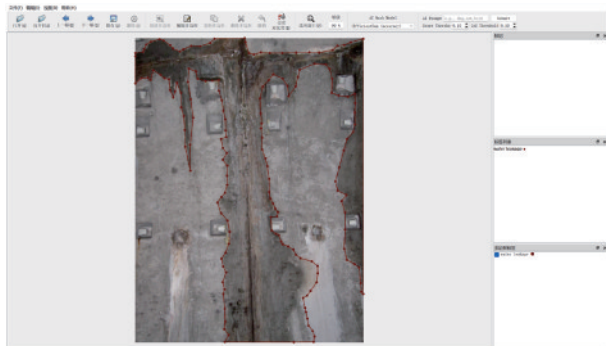


图9 Labelme标注软件界面展示

Fig.9 Labelme labeling software interface display

息,但仅能被部分算法使用。为了后续进行多种深度学习网络的训练,需对JSON格式标注文件进行转化以适应不同网络,例如将JSON格式标注文件转化为YOLO格式的TXT文件以满足本文主要使用的YOLO网络训练。

3 渗漏水病害分割实验

3.1 实验环境与训练参数

相关深度学习实验均基于Python3.9计算机语言和Pytorch模块,在配置有图形处理器(GPU)的工作站上执行(CPU为Intel(R) Xeon(R) W2295 @ 3.00 GHz, GPU为NVIDIA GeForce RTX3080Ti 12 GB),所采用的系统为Windows 10。

为了保证模型训练过程充分收敛,并在实际应用中达到较高的泛化能力,本研究对训练超参数进行了精细调优。训练总轮数设定为300轮,以确保模型有足够的迭代次数实现参数稳定;同时,将输入图像尺寸固定为 640×640 ,从而消除不同图像尺度对特征提取的干扰。在批次大小上,依据GPU显存条件,本研究选择了16作为批次大小,并启用8个工作线程进行数据加载,以加快数据预处理及输入速度。此外,指定使用默认GPU加速训练,以充分利用硬件计算资源。

在优化策略方面,本研究采用AdamW优化器,其在权重衰减方面具有较为突出的性能,有助于抑制过拟合现象。初始学习率设定为0.001,并利用余弦退火策略对学习率进行逐步衰减,直至终值0.000 01,以确保在训练后期能够细化参数调整,防止陷入局部最优。与此同时,动量参数设为0.95,权重衰减系数为0.000 5,这些参数均经过反复实验验证,旨在提高模型的稳定性和泛化能力。

为进一步提升网络对复杂多变环境的适应性,本研究在数据增强策略上采用了部分Mosaic和Mixup技术,两者均以0.5的概率启用,从而在训练过程中丰富样本多样性,降低因数据单一性引发的过拟合风险。训练完成后,基于验证误差最小原则选取最佳权重,作为后续实验分析的最终模型。总体而言,以上参数配置在保证训练效率的同时,有效提升了模型性能,为后续的实际应用提供了可靠的技术支持。网络训练的具体超参数见表2。

为了验证训练过程的收敛性,图10展示了改进模型在300轮训练中的损失变化曲线。可以观察

到,训练损失与验证损失均呈现稳定下降趋势:前50轮快速下降,中期(50-200轮)平稳过渡,200轮到300轮波动幅度较小处于可接受状态。损失曲线的平滑收敛特性表明,设定的300轮训练能确保模型参数充分稳定,满足工程应用对稳定性的要求。

表2 网络训练超参数设置

超参数	设置情况
训练轮数	300
输入尺寸	640
批次大小	16
数据加载线程数	8
优化器	AdamW
初始学习率	0.001
学习率终值	0.000 01
动量	0.95
权重衰减	0.000 5
Mosaic数据增强概率	0.5
保持Mosaic始终开启	True
Mixup数据增强概率	0.5
使用余弦退火策略	True

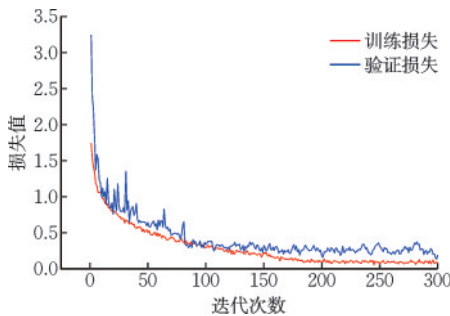


图10 训练损失与验证损失变化曲线

Fig.10 Training loss and validation loss variation curves

3.2 评估指标

为了全面评估本研究构建的网络性能,本文主要采用精确率(Precision, P)、召回率(Recall, R)、平均精度(Average Precision, AP)和F1-score进行量化。这些指标分别用于衡量模型在正确识别目标、覆盖真实目标以及两者平衡表现方面的能力。精确率 P 计算为检测正确的正样本占有所有检测为正样本的比例,反映了模型预测的准确性;召回率 R 则是检测正确的正样本占有所有真实正样本的比例,用以衡量模型对目标的检出能力。平均精度 AP 则通过计算不同召回率下精确率的积分,综合展示了模

型在多种阈值下的检测效果,而F1-score则是 P 与 R 的调和平均,能够直观反映模型整体检测性能。

此外,为了评估网络在轻量化设计方面的效果,本研究还使用FPS(每秒帧数体现推理速度)、Params(模型参数量)和Gflops(Giga floating-point operations per second,模型单次推理的浮点运算量,用于衡量计算复杂度)作为补充指标。这些轻量化指标不仅能体现模型在实际应用中对计算资源的需求,还能为资源受限的边缘设备部署提供参考,从而更好地平衡检测精度与实时性。各评价指标及其公式如表3所示。

表3 病害量化评价指标

指标	公式
精确率	$P = \frac{TP}{TP + FP}$
召回率	$R = \frac{TP}{TP + FN}$
平均精度	$AP = \int_0^1 P(R) dR$
F1值	$F1 = \frac{2 \times P \times R}{P + R}$
每秒帧数	$FPS = \frac{N_{picture}}{T_{total}}$

其中, TP 为正样本被检测正确的数量, FP 为负样本被检测为正样本的数量, FN 为正样本未被检测为正样本的数量, $N_{picture}$ 为分割图片的数量; T_{total} 为分割所有图片所用的总时间。

3.3 注意力机制与主干网络对比实验

为了验证不同注意力机制及主干网络对模型性能的影响,本研究以YOLOv8n-seg作为baseline,在其基础上分别添加不同的注意力机制和替换主干网络进行实验对比分析。其中,注意力机制部分则在baseline模型基础上分别添加了EMA(Efficient Multi-scale Attention)、ELA(Efficient Lightweight Attention)、CBAM(Convolutional Block Attention Module)、SCSE(Squeeze and Excitation with Spatial Attention)以及CA(Coordinate Attention),以分析不同注意力机制在强化关键特征、抑制无关背景信息以及提升检测与分割精度方面的作用。另一方面,主干网络部分选取了FasterNet、EfficientViT、Swin Transformer、RevCol和MobileNetV4五种结构进行替换,旨在探讨不同网络架构

在特征提取能力、计算复杂度以及轻量化方面的差异。

注意力机制对比实验的结果见表4,各注意力机制均能在一定程度上提高精确率和平均精度。与EMA、ELA、CBAM、SCSE这四种注意力机制相比,CA注意力机制在*P*、*AP*、*F1*三项指标上均取得最高值,说明它在抑制背景干扰、突出病害特征方面的综合能力更强,不仅能准确定位目标,也能在多种阈值下保持稳定的检测性能。反而在添加注意力机制后,召回率与基础模型相比出现不同程度的下降,这是因为注意力机制通过更精细的特征加权与背景抑制来提升模型的区分度,从而提高精确率,“误判”减少,但也因为其强化显著特征、忽略边缘特征的特性,可能导致部分真实目标被筛掉,造

表4 注意力机制对比实验结果

Table 4 Comparative test results of attention mechanisms

注意力机制	<i>P</i> /%	<i>R</i> /%	<i>AP</i> /%	<i>F1</i> /%
baseline	75.97	72.26	79.40	74.07
EMA	80.30	70.02	79.56	74.81
ELA	80.53	68.93	79.82	74.28
CBAM	79.46	70.16	79.22	74.52
SCSE	78.04	70.20	77.39	73.91
CA	81.79	72.11	80.22	76.65

成“漏判”增多。CA注意力机制的召回率较其他模型相比,减少程度较小,处于可接受范围内。总体而言,在添加各注意力机制后,模型的参数量,计算

量等方面没有显著增加的前提下,CA注意力机制在多项指标上表现最为突出,能够在较高精确率的同时保证良好的召回率,并获得最高的平均精度与*F1*值,因此选用CA作为YOLOv8n-seg的注意力机制。

主干网络对比实验的结果见表5,不同网络结构在检测精度和计算资源消耗上呈现出明显的权衡趋势。整体而言,MobileNetV4与Swin Transformer均表现出较高的检测精度,其中MobileNetV4在精确率和*F1*值上均达到了较优水平,而Swin Transformer在这两方面上则表现出更为明显的优势;相比之下,FasterNet的检测精度略低于前两者,而其他轻量化模型(如EfficientViT和RevCol)则表现出更低的检测指标,其原因可能在于这些网络在特征表达和信息融合上存在不足,导致对复杂背景中微弱目标的捕捉能力下降。然而,基于轻量化的目标,MobileNetV4由于其较低的参数量和计算量,能够在嵌入式及边缘设备上实现实时检测,同时在保持较高检测精度的同时,显著降低了部署成本;而Swin Transformer虽然在精度上表现卓越,但其基于窗口自注意力机制和全局信息交互设计导致参数量和计算量大幅上升,其中参数量与原网络参数量相比增加近一个数量级,从而限制了其在资源受限场景下的应用。综上所述,在兼顾检测精度与实时性的应用需求下,选用MobileNetV4作为YOLOv8n-seg的主干网络显然是更为合适的选择。

表5 主干网络对比实验结果

Table 5 Comparative test results of backbone networks

主干网络类型	<i>P</i> /%	<i>R</i> /%	<i>AP</i> /%	<i>F1</i> /%	Params	Gflops
baseline	75.97	72.26	79.40	74.07	3.259	12.0
FasterNet	80.62	70.78	80.04	75.38	4.429	14.6
EfficientViT	76.23	71.82	76.10	73.96	4.261	13.3
Swin Transformer	81.35	73.15	81.78	77.03	30.239	83.0
RevCol	76.81	68.97	77.48	72.68	2.529	10.1
MobileNetV4	79.84	72.61	81.21	76.05	1.852	7.8

3.4 消融实验

为了验证不同改进策略在YOLOv8n-seg模型中的有效性,逐一将CA注意力机制、MobileNetV4主干网络、EfficientHead分割头等模块集成到模型中进行消融实验,实验结果见表6。实验结果表明:

CA注意力机制对模型性能的提升也具有显著作用。单独添加CA模块后,模型的*AP*和*F1*分别提升了0.82%和2.58%。这一结果表明,CA模块能够有效地增强隧道病害目标区域的权重,从而在更复杂的场景中进一步提升网络的检测精度。MobileNetV4的引入显著提高了模型的性能,特别是在

AP和F1指标上表现突出,分别提高了1.81%和1.98%。这表明MobileNetV4作为主干网络能够更有效地提取隧道病害目标的特征,从而提高模型的识别精度,此外,随着主干网络替换为更简单的MobileNet系列网络,模型参数量和计算量也得到显著

减少。EfficientHead分割头模块的替换提高了模型在细粒度特征提取上的能力。实验结果显示,替换分割头后,模型的AP和F1分别提高了1.45%和1.15%。这一结果表明,分割头能够在增强模型的分割能力的同时,不显著增加计算负担。

表6 消融实验结果

Table 6 Results of ablation experiments

网络	CA	MobileNetV4	分割头	AP/%	F1/%	FPS	Params	Gflops
v8n-seg	×	×	×	79.40	74.07	333.3	3.259	12.0
v8n-seg+CA	√	×	×	80.22	76.65	344.8	3.265	12.0
v8n-seg+MbV4	×	√	×	81.21	76.05	375.1	1.852	7.8
v8n-seg+eHead	×	×	√	80.85	75.22	326.5	3.402	13.2
v8n-seg+CA+MbV4	√	√	×	82.13	77.30	360.9	1.858	7.9
v8n-seg+CA+eHead	√	×	√	81.35	77.11	335.4	3.408	13.3
v8n-seg+MbV4+eHead	×	√	√	82.46	77.22	340.6	1.955	8.5
v8n-seg+CA+MbV4+eHead	√	√	√	83.21	78.53	355.2	1.960	8.7

综上所述,单独使用这三种改进策略,模型在特征提取、识别精度和速度方面均表现出了不同程度的提升,证明了这些改进策略在隧道病害检测中的有效性。

在进一步测试联合改进策略时,将MobileNetV4、CA注意力机制和EfficientHead分割头同时应用于YOLOv8n-seg网络中,评估其对模型性能的综合影响。实验结果显示:MobileNetV4与CA注意力模块的联合使用,使得模型的AP和F1分别提高了2.73%和3.23%;CA注意力模块与EfficientHead分割头模块的联合使用,模型的AP和F1分别提升了1.95%和3.04%;MobileNetV4与EfficientHead分割头模块的组合提升了模型的AP和F1,分别提高了3.06%和3.15%;三种改进策略的综合应用,即CA注意力模块、MobileNetV4和EfficientHead分割头模块的联合使用,使得模型的AP和F1分别提升了3.81%和4.46%。这一结果验证了三者结合能够更好地互补各自的优势,极大提升了网络的特征提取、图像分割和识别精度。

如图11所示,展示了部分渗漏水病害的检测结果。相比于基线模型,所提出的改进模型在漏检与误检方面均表现出更高的检测精度。具体而言,改进模型能够更有效地识别第一、二张图中位于图像上部的狭长渗漏水区域,而基线模型在该区域存在明显漏检。在第三张图中,基线模型未能识别图像正上方的小面积渗漏水区域,且由于线缆遮挡,其对较大渗漏水区域的检测同样存在缺失,而改进模

型则在上述两方面均实现了更为准确的检测。此外,在第四张图中,基线模型出现了较多的误检现象,而改进模型显著减少了该类错误。综上所述,改进模型在复杂背景条件下展现出更强的渗漏水识别能力与鲁棒性。

此外,联合改进策略在每秒帧数(FPS)上也表现出了较好的提升,参数量与计算量都大幅减少,说明模型在提升检测精度的同时,仍能保持较高的计算效率,适应高效实时检测的需求。

3.5 对比实验

为了验证改进模型的综合性能,选取了若干网络进行对比训练,实验结果见表7。所有模型中表现最好的是Mask R-CNN,AP和F1分别达到了85.24%和81.89%,其次是本文提出的改进模型,AP和F1分别为83.21%和78.53%,改进模型的表现略低于Mask R-CNN。但对比两者的参数量与计算量,显然Mask R-CNN要高出数倍,这是因为Mask R-CNN基于两阶段框架,其结构天然适合高精度实例分割,但两阶段流程需多次特征提取与计算,导致参数量和计算量大幅增加。此外YOLO11n-seg在部分模块改进后的训练结果与v8版本相比有小幅提升。U-net则因其对称结构,参数量较大,训练表现一般。FCN尽管有着与改进模型相差无几的轻量化设计,但因其模型结构较为简单,训练结果远不及改进模型,在轻量化的基础上无法实现较高精度的病害分割。

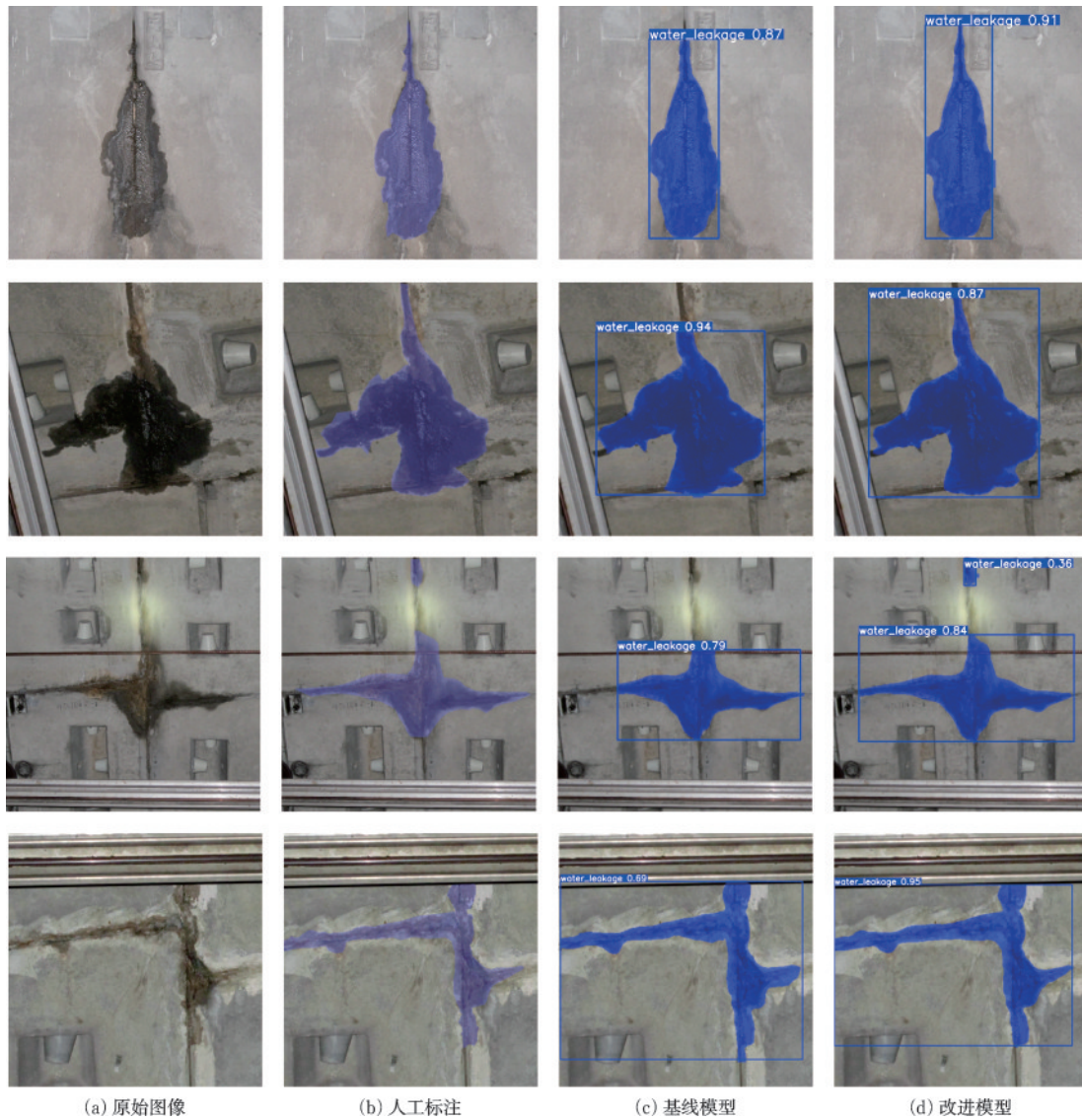


图 11 隧道渗漏水病害检测结果
Fig.11 Detection results of tunnel water leakage defects

表 7 模型对比实验结果

Table 7 Comparative test results of models

模型名称	$P/\%$	$R/\%$	$AP/\%$	$F1/\%$	Params	Gflops
v8n-seg	75.97	72.26	79.40	74.07	3.259	12.0
11n-seg	79.65	71.42	78.58	75.31	2.835	10.2
u-net	78.34	73.18	80.75	75.67	4.523	15.2
FCN	73.59	70.35	75.83	71.93	2.107	8.4
Mask R-CNN	83.12	80.65	85.24	81.87	15.354	32.3
Improved v8n-seg	82.68	74.78	83.21	78.53	1.960	8.7

因此,改进模型可以在满足轻量化检测的前提下,完成对病害的精准分割,显著提升了检测精度和分割效果。这使得模型能够适应实际应用中的实时检测需求,尤其在隧道病害检测等高要求场景中展现出良好的性能。

4 结 论

针对隧道渗漏水检测中传统方法效率低下以及现有深度学习模型计算资源消耗过高的问题,提

出了一种基于改进YOLOv8n-seg的轻量化实例分割模型,通过多项优化实现了性能的显著提升。具体改进与成果如下:

(1) 通过引入Coordinate Attention(CA)注意力机制增强目标区域的特征表达能力,采用MobileNetV4替换主干网络以优化特征提取效率,并结合EfficientHead分割头提升解码能力,有效平衡了检测精度与计算复杂度。

(2) 实验结果显示,改进模型在渗漏水数据集上的平均精度(AP)达到83.21%,F1值提升至78.53%,分别较基线模型提高3.81%和4.46%。同时,模型参数量减少至1.96 M,推理速度达到355.2 FPS,较原模型提升6.6%。消融实验进一步证明,各改进模块协同作用显著,其中CA机制增强了特征区分能力,MobileNetV4在减少43.2%参数量的同时保持高精度特征提取,EfficientHead通过轻量化设计优化了分割细节。

(3) 与Mask R-CNN、U-Net等主流模型相比,改进模型在参数量和GFLOPs等轻量化指标上具有显著优势,同时精度接近两阶段方法,能够满足隧道复杂场景下的实时检测需求,为隧道结构健康监测提供了高效、可靠的解决方案,尤其适用于资源受限的边缘计算设备部署。

(4) 本研究提出的轻量化模型可进一步优化以适应移动终端的部署需求,同时拓展至其他隧道病害类型的检测任务,实现更全面的隧道健康状态监测。通过持续改进模型架构和算法效率,在多病害复杂场景下进行训练,该轻量化模型将会在工程实践中得到广泛应用,为隧道智能化运维提供有力支持。

参考文献:

- [1] 谢森林,胡安峰,肖志荣,等. 渗漏病害加剧影响下隧道周围土体流变固结分析[J]. 岩石力学与工程学报, 2024, 43(10): 2570-2580.
Xie S L, Hu A F, Xiao Z R, et al. Rheological consolidation analysis of soil around tunnels under exacerbated leakage conditions[J]. Chinese Journal of Rock Mechanics and Engineering, 2024, 43(10): 2570-2580. (in Chinese)
- [2] Sun W X, Han F C, Zhang Y M, et al. Experimental assessment of structural responses of tunnels under the groundwater level fluctuation [J]. Tunnelling and Underground Space Technology, 2023, 137: 105138.
- [3] 许力之,王耀东,朱力强,等. 隧道表面图像多目标智能识别算法研究[J]. 铁道学报, 2022, 44(9): 154-162.
Xu L Z, Wang Y D, Zhu L Q, et al. Research on algorithm of intelligent recognition of multiple objects of tunnel surface images[J]. Journal of the China Railway Society, 2022, 44(9): 154-162. (in Chinese)
- [4] 周鸣亮,汪长松,黄宏伟,等. 融合机器视觉与性能分析的运营盾构隧道结构安全状态评价[J]. 应用基础与工程科学学报, 2023, 31(6): 1461-1476.
Zhou M L, Wang C S, Huang H W, et al. Safety state evaluation of operational shield tunnel structures by integrating computer vision and performance analysis [J]. Journal of Basic Science and Engineering, 2023, 31(6): 1461-1476. (in Chinese)
- [5] Wang S, Wang L Q, Zhang W G, et al. Fracture mode of water-immersed sandstone based on micro CT and statistical damage model under triaxial compression[J]. Engineering Fracture Mechanics, 2025, 315: 110781.
- [6] Wen S H, Park M, Tran D Q, et al. Automated construction safety reporting system integrating deep learning-based real-time advanced detection and visual question answering[J]. Advances in Engineering Software, 2024, 198: 103779.
- [7] Zhou Z P, Zhuo W, Cui J Q, et al. Developing a deep reinforcement learning model for safety risk prediction at subway construction sites[J]. Reliability Engineering & System Safety, 2025, 257: 110885.
- [8] Ding Z K, Luo L W, Wang X R, et al. An artificial intelligence-based method for crack detection in engineering facilities around subways [J]. Applied Sciences, 2023, 13(19): 11002.
- [9] Zhu A F, Chen S H, Lu F F, et al. Recognition method of tunnel lining defects based on deep learning[J]. Wireless Communications and Mobile Computing, 2021, 2021: 9070182.
- [10] Bani Mustafa A, Abdel Halim R, Bulkrock O, et al. Deep learning for assessing severity of cracks in concrete structures [J]. International Journal of Computers Communications & Control, 2023, 18(1): 17-17.
- [11] Zhou Z, Li S S, Yan L B, et al. Intelligent recognition of tunnel lining defects based on deep learning: Methods, challenges and prospects[J]. Engineering Failure Analysis, 2025, 170: 109332.
- [12] 耿鹏. 基于深度学习的隧道衬砌病害图像检测技术[D]. 石家庄: 石家庄铁道大学, 2024.
Geng P. Tunnel lining disease image detection technology based on deep learning[D]. Shijiazhuang: Shijiazhuang University of Railway Engineering, 2024.

- zhuang Tiedao University, 2024. (in Chinese)
- [13] Huang H W, Zhao S, Zhang D M, et al. Deep learning-based instance segmentation of cracks from shield tunnel lining images[J]. *Structure and Infrastructure Engineering*, 2022, 18(2): 183-196.
- [14] Xue Y D, Cai X Y, Shadabfar M, et al. Deep learning-based automatic recognition of water leakage area in shield tunnel lining[J]. *Tunnelling and Underground Space Technology*, 2020, 104: 103524.
- [15] 王宝坤, 王如路, 陈锦剑, 等. 基于深度学习的盾构隧道表观病害自动检测方法[J]. *上海交通大学学报*, 2024, 58(11): 1716-1723.
Wang B K, Wang R L, Chen J J, et al. Automatic detection method for surface diseases of shield tunnel based on deep learning[J]. *Journal of Shanghai Jiao Tong University*, 2024, 58(11): 1716-1723. (in Chinese)
- [16] Zhao S, Zhang D M, Huang H W. Deep learning - based image instance segmentation for moisture marks of shield tunnel lining[J]. *Tunnelling and Underground Space Technology*, 2020, 95: 103156.
- [17] Ouyang A H, Di Murro V, Cull M, et al. Automated pixel-level crack monitoring system for large-scale underground infrastructure: A case study at CERN[J]. *Tunnelling and Underground Space Technology*, 2023, 140: 105310.
- [18] Xiong L J, Zhang D L, Zhang Y. Water leakage image recognition of shield tunnel via learning deep feature representation[J]. *Journal of Visual Communication and Image Representation*, 2020, 71: 102708.
- [19] 宋娟, 贺龙喜, 龙会平. 基于深度学习的隧道衬砌多病害检测算法[J]. *浙江大学学报(工学版)*, 2024, 58(6): 1161-1173.
Song J, He L X, Long H P. Deep learning-based algorithm for multi defect detection in tunnel lining[J]. *Journal of Zhejiang University (Engineering Science)*, 2024, 58(6): 1161-1173. (in Chinese)
- [20] Duan S Q, Zhang M H, Qiu S L, et al. Tunnel lining crack detection model based on improved YOLOv5[J]. *Tunnelling and Underground Space Technology*, 2024, 147: 105713.
- [21] 朱家松, 郑澳, 雷占占, 等. 基于改进 Yolov5 的地铁隧道附属设施与衬砌表观病害检测方法[J]. *铁道科学与工程学报*, 2023, 20(3): 1008-1019.
Zhu J S, Zheng A, Lei Z Z, et al. Metro tunnel accessorial facilities and lining diseases detection method based on improved Yolov5[J]. *Journal of Railway Science and Engineering*, 2023, 20(3): 1008-1019. (in Chinese)
- [22] Xu X Q, Li Q, Li S E, et al. Crack width recognition of tunnel tube sheet based on YOLOv8 algorithm and 3D imaging [J]. *Buildings*, 2024, 14(2): 531.
- [23] Chen Y P, Dai X Y, Liu M C, et al. Dynamic convolution: Attention over convolution kernels[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020. Seattle, WA, USA: IEEE, 2020: 11027-11036.
- [24] Tian Z, Shen C H, Chen H. Conditional convolutions for instance segmentation[C]//Computer Vision-ECCV 2020. Cham, Switzerland: Springer, 2020: 282-298.

(本文编辑:周小潭)