

基于改进 YOLOv5s 模型的水工建筑物裂缝检测

李怡静, 邓佳蕙, 袁世虹

(南昌大学工程建设学院)

摘要:针对水工建筑物在复杂自然环境和运行工况下易出现水渍、析钙等劣化现象,进而增大表面裂缝检测难度的问题,提出了一种融合可变形卷积网络(DCNv2)和缩放解耦头(SEH)的裂缝检测模型 YOLO-SEH。该模型在颈部引入高效通道注意力机制,以抑制复杂背景中的噪声干扰;在骨干部分采用 DCNv2 替换 YOLOv5 模型的 C3 模块,以增强模型对裂缝形状的适应能力;在模型头部引入 SEH,以提高模型对裂缝特征的敏感度与辨识能力。将该模型用于无人机采集的水闸和大坝影像裂缝检测任务中,其 mAP@0.5 指标分别达到 86.2%和 84.9%;在消融实验中,与 YOLOv5 模型相比,YOLO-SEH 模型的 mAP@0.5 在大坝和水闸数据集上分别提升了 3.1、4.1 个百分点,精确率分别提升了 4.8、6.6 个百分点,召回率分别提升了 3.4、4.7 个百分点;与 Faster R-CNN、SSD300、YOLOv5、YOLOv7 和 YOLOv8 等模型相比,YOLO-SEH 模型在水工建筑物裂缝检测任务中表现出更优的检测性能。

关键词:水工建筑物;裂缝检测;深度学习;目标检测;YOLOv5s;可变形卷积网络;解耦头

Crack detection for hydraulic structures based on an improved YOLOv5s model//Li Yijing, Deng Jiahui, Yuan Shihong(*School of Infrastructure Engineering, Nanchang University*)

Abstract: Aiming at the problem that hydraulic structures are prone to deterioration such as water stains and calcium precipitation under complex natural environments and operating conditions, thereby complicating surface crack detection, this paper proposes a crack detection model YOLO-SEH integrating a deformable convolutional network (DCNv2) and scaling efficient decoupled head (SEH). An efficient channel attention mechanism is introduced into the model neck to suppress noise interference from complex backgrounds. In the backbone, DCNv2 is adopted to replace the C3 module of YOLOv5, improving the model's adaptability to irregular crack shapes. SEH is incorporated into the detection head to strengthen the model's sensitivity and discriminability to crack features. The proposed model was applied to crack detection on sluice and dam images captured by unmanned aerial vehicles (UAVs), achieving mAP@0.5 values of 86.2% and 84.9%, respectively. Ablation experiments show that compared with the YOLOv5 model, the mAP@0.5 of the YOLO-SEH model increased by 3.1 and 4.1 percentage points, the precision increased by 4.8 and 6.6 percentage points, and the recall increased by 3.4 and 4.7 percentage points for the two datasets, respectively. Furthermore, comparative experiments with the Faster R-CNN, SSD300, YOLOv5, YOLOv7, and YOLOv8 models demonstrate that the YOLO-SEH model achieves superior detection performance for crack detection of hydraulic structures.

Key words: hydraulic structures; crack detection; deep learning; object detection; YOLOv5s; deformable convolutional network; decoupled head

水工建筑物在运行过程中可能因结构形变、地震活动及水流冲刷而产生裂缝,进而面临安全风险。通过分析裂缝的几何形态及外观特性,可评估其劣化程度并推测潜在的成因,为结构健康监测提供科学依据。人工检测裂缝不仅耗时耗力,且存在较大的作业安全风险^[1]。目前,无人机摄影和计算机图像处理已成为裂缝检测的主要手段。然而,传统图

像处理方法(如阈值分割^[2]、边缘检测^[3-5]、区域生长^[6-7]等)过度依赖参数设定,普适性差,难以适应多样化的裂缝形态。相较于传统方法,机器学习方法(如支持向量机^[8]、决策树^[9]等)能够利用训练数据学习到更具代表性的特征。但该类方法严重依赖人工设计的特征,需要大量的专业知识和经验,且难以全面准确地描述裂缝特性,导致模型的泛化能力

不足。

深度学习方法的兴起为裂缝检测带来了新的突破,该方法能够自主地从图像数据集中学习特征,避免了传统方法的局限性^[10]。例如:Zhu 等^[11]将一种轻量级注意力机制的特征融合策略应用于单发多框探测器中,实现了混凝土裂缝的有效检测;刘万军等^[12]基于 Faster R-CNN 融合多尺度特征增强与密集连接结构,遥感识别特征提取性能提升,但推理效率偏低,难以满足实时应用需求;杨富强等^[13]利用迁移学习对 YOLOv3 模型进行了优化,结果表明,相较于直接使用原数据集进行训练的 YOLOv3 模型,改进模型的平均精度指标提升了 13.16 个百分点,显著提高了桥梁裂缝检测的准确率,这一改动虽提升了病害特征的代表能力和检测效率,但网络模型的参数复杂度仍然维持在较高水平;Li 等^[14]基于 YOLOv5 模型正则化优化策略去除了深度神经网络中存在的冗余参数,在自制的拱坝裂缝数据集上取得了较好的成果。

尽管深度学习中的目标检测算法已在裂缝识别领域取得了显著进步,仍存在对特定复杂背景检测效果不佳的问题。同时,由于水工建筑物容易出现水渍、析钙等现象,表面呈现灰白色、黑色、斑块状或条带状分布等劣化现象,使得裂缝与背景之间差异不明显,给裂缝检测带来了更高的挑战。YOLOv5 模型在满足工业领域对效率与成本双重考量方面,展现出了较高的潜力与价值。本文基于 YOLOv5 模型提出了一种融合可变形卷积网络(DCNv2)^[15]与缩放解耦头(SEH)的 YOLO-SEH 模型,并与其他模型进行了对比。

1 模型构建

1.1 YOLOv5s 模型优化策略

YOLOv5 模型主要由骨干网络(Backbone)、颈

部(Neck)和预测头(Head)组成。骨干网络由空间金字塔池化(SPPF)模块和多个卷积(Conv)模块、自适应聚合模块 C3 组成,负责提取输入图像的特征;颈部采用了上采样(Upsample)和拼接(Concat)操作,负责对骨干网络提取的特征进行多尺度特征融合,并把这些特征传递给预测头;预测头则基于这些特征完成目标检测的分类、回归和置信度预测任务。YOLOv5 目标检测框架根据模型深度和宽度参数的不同,提供了 4 个可扩展的版本,按照计算复杂度从低到高依次为:YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。YOLOv5s 作为 YOLOv5 系列中轻量化程度最高的版本,具有参数量少、推理速度快、硬件部署门槛低等优势,可适配水工场景实时监测需求,因此本文基于 YOLOv5s 模型展开研究。

为提高水工建筑物的裂缝检测精度,本文对 YOLOv5s 进行了如下优化(图 1):在骨干网络部分引入可变形卷积技术(DCNv2),将卷积神经网络(CNN)中的卷积核固定采样模式优化为随感受野变化的动态调整模式,以更好地捕捉水工建筑物中形态多样的裂缝细节;在颈部引入高效通道注意力机制(ECA)^[16],使模型自适应地强化裂缝特征,抑制水渍、析钙等背景噪声的干扰;提出一种缩放解耦头(scaling efficient decoupled head, SEH),将预测头部分的分类和定位任务分离,以更好地适应各自关注的特征,同时根据模型规模自动调整通道数,用少量计算代价提升模型对复杂背景下裂缝特征的敏感度与辨识能力。

1.2 DCNv2 可变形卷积

传统二维卷积核在处理输入的影像特征图时,常遵循网格 R 的固定采样模式,并在每个采样点应用权重 k 进行加权求和。然而,这种方法受限于其深度特征的固有特性,其感受野范围有限,难以充分捕捉裂缝区域的全局上下文信息。因此,本文采用

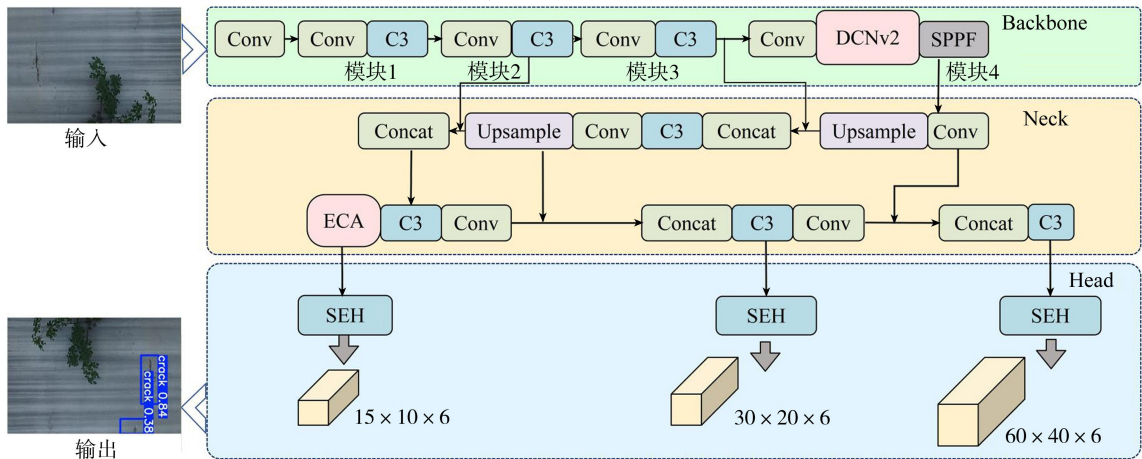


图 1 YOLO-SEH 网络的总体结构

DCNv2,旨在提升模型对裂缝几何形变的适应能力^[17],从而学习到对水工建筑物裂缝中几何形变更为稳健的深度特征表示,进而增强模型的鲁棒性和适用性。

在对水工建筑物影像中的裂缝进行特征提取时,传统标准卷积采样如图 2(a)所示,可变形卷积采样如图 2(b)所示,其中圆点符号代表从图像特征图中提取的激活单元。这些圆点经由箭头指向的后续单元,各自对应于先前特征图上突出标记的单元位置。经过可变形卷积的处理后,所得单元点与物体空间位置之间的关联性增强,从而能够更高效地利用目标的特征信息。

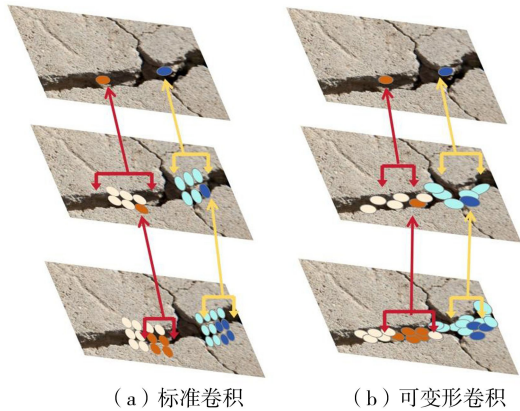


图2 标准卷积和可变形卷积对裂缝特征的提取对比

1.3 ECA 注意力机制

针对原模型在特征表达及关键信息关注方面的不足,以及通道间信息交互受限导致的误检等问题(如将水工建筑物表面的水渍、析钙等劣化现象误判为裂缝)^[18],本文将 ECA 模块嵌入网络结构中,以抑制背景噪声的干扰,同时增强对裂缝特征的提取能力。ECA 采用一维卷积增强通道间局部交互,避免降维造成的信息损失,最大程度地保留裂缝的细微信息,降低漏检和误检风险。

骨干网络和颈部网络是 YOLOv5s 特征提取与多尺度融合的核心结构,二者的结构设计直接决定模型检测性能。鉴于原模型的骨干和颈部结构引入 ECA 可能存在不确定性,因此本文探讨了在这两层结构的不同位置嵌入 ECA 模块的影响。为避免注意力机制过早介入导致特征提取出现偏差,将注意力模块分别部署在骨干网络和颈部网络的末端层。由于骨干部分特征图分辨率高,插入 ECA 可增强全局上下文捕捉能力,提升目标整体识别效果,但可能忽略局部细节,对小目标检测效果有限。颈部部分特征图分辨率低但细节丰富,插入 ECA 可以增强模型对小目标的敏感性,在裂缝检测任务中表现更优。通过后续实验验证上述分析,进一步明确 ECA 模块的最佳插入位置为颈部。

1.4 SEH

YOLOv5 模型采用的耦合头设计未对分类任务和回归任务进行分离,而是将融合多层的特征图同时输出分类、回归和目标性信息。这种耦合头设计在裂缝检测任务中易引发任务干扰,且对复杂目标的适应性不足,影响了模型的检测性能。

YOLOv6 模型^[19]的混合通道策略提高了解耦头的效率,其中 3×3 卷积层被精简为一层。然而,尽管加入了解耦头,模型的计算复杂度依然较高,并且解耦头的大小不具备随模型尺寸调整的灵活性,这对注重检测效率的轻量级物体检测网络而言是不利的^[20]。

为解决上述问题,本文对 YOLOv6 模型的解耦头进行了改进。YOLOv6 模型的头部采用固定通道配置,而 YOLOv5s 模型虽具备通道调整机制,但其骨干与颈部结构并未直接受通道数影响。因此,本文提出通道缩放因子(CSR)概念,CSR 是针对高效解耦头设计的模型尺寸自适应超参数,具体由以下公式实现:

$$K_{CSR} = C_{SEH} / C_{FPN} \quad (1)$$

其中

$$C_{FPN} = C_{FPN-base} \omega_m$$

式中: C_{SEH} 为 YOLO-SEH 模型解耦头中卷积层的实际输出通道数; C_{FPN} 为颈部特征金字塔网络(FPN)的实际通道数; $C_{FPN-base}$ 为 FPN 的基准通道数; ω_m 为通道宽度乘数。

如图 3 所示,YOLO-SEH 模型的头部网络通过结合骨干与颈部网络的 CSR 动态调整通道数,从而能够针对不同模型尺寸优化检测头部设计。随着 CSR 的变化,特征图通道数相应改变,使网络能够根据模型大小灵活调节特征图通道数。在水工建筑物裂缝检测中,小尺寸模型如 YOLOv5 模型可通过该方法减少参数量,避免因引入大尺寸解耦头导致参数冗余和性能下降问题,从而实现模型的轻量化与高效性。

2 实验分析

2.1 实验环境

实验环境配置为 Windows11 64bit 操作系统,CPU 为 Intel(R) Core(TM) i7-8565U@1.80GHz,GPU 为 NVIDIA GeForce MX230,7.9 G 内存。深度学习工具使用 PyTorch1.11,图形处理器驱动为 CUDA11.5,cuDNN8,编程语言使用 Python 3.11.7,开发环境为 PyCharm。参数设置上,本文采用随机梯度下降优化器,算法训练采用周期性衰减学习率策略,初始学习率设定为 0.01,学习率衰减周期为 200 个训练轮次,批量尺寸为 8,权重衰减系数取 0.0005。

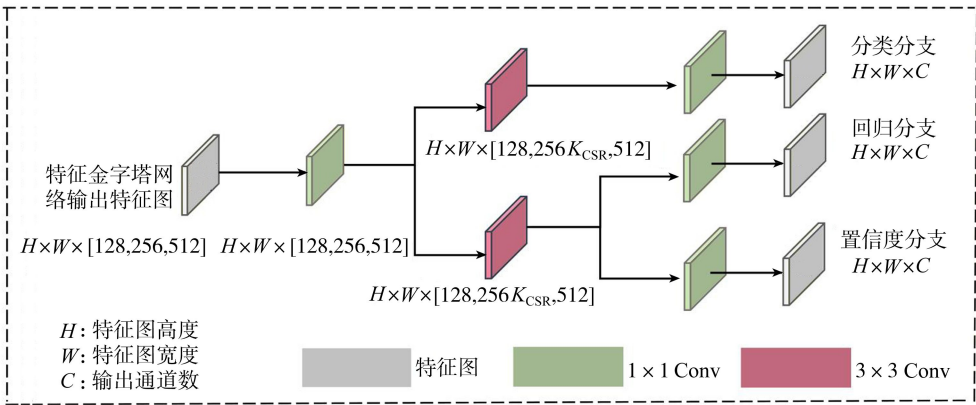


图3 YOLO-SEH 的解耦头结构

2.2 数据来源

本文构建了两套数据集用于 YOLO-SEH 模型的训练和测试。第一套数据集来自某进水闸的无人机序列影像和公开数据集 DeepCrack^[21]。无人机影像大小为 5472×3648 像素,数量为 400 张。挑选可用的 20 张图片进行镜像、旋转、色彩变换等方式增强数据,再加入 537 张大小为 544×384 像素的 DeepCrack 图像进行扩充,最后形成 1000 张大小为 480×320 像素的裂缝图像,如图 4 所示。

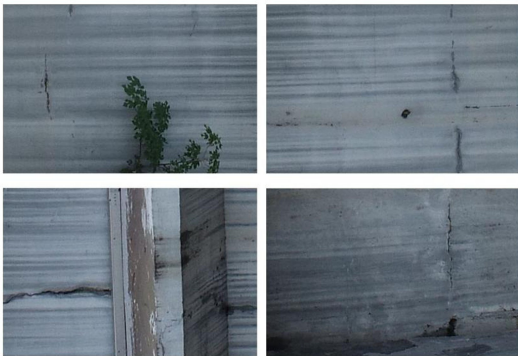


图4 水闸裂缝数据集样例

第二套数据集为本文独立构建的裂缝数据集和部分文献数据集^[22]。获得的大坝裂缝影像数量为 3800 张,选出可用图片 500 张,经过数据增强、裁剪后形成 416×416 像素与 200×200 像素的 2000 张裂缝图像,如图 5 所示。采用 Labellmg 软件对两个数



图5 大坝裂缝数据集样例

据集中的裂缝实施精确标注,并依据 7:2:1 的比例随机分配这些图像与标签至训练集、验证集及测试集中。

2.3 模型性能评价指标

为了评估 YOLO-SEH 模型的性能,采用均值平均精度(mean average precision, mAP, 本文交并比阈值取 0.5,记作 mAP@0.5)、准确率(P)以及召回率(R)等关键指标进行量化评价。

2.4 实验结果与分析

2.4.1 ECA 位置对模型性能的影响

在 YOLOv5s 模型的骨干与颈部分别引入了 ECA 模块后的实验结果如表 1 所示。在两个数据集的 YOLOv5s 裂缝检测中,将 ECA 融合至骨干网络时 mAP@0.5 涨幅较小,而将其融合至颈部部分 C3 模块的最后一层时检测效果表现更佳,能更好地适应裂缝检测的任务需求。骨干网络主要负责提取基础的图像特征,它所处理的是相对底层、宽泛的图像信息。此时引入 ECA,虽然能在一定程度上调整通道注意力,但由于基础特征的广泛性和一般性,对最终检测效果的提升较为有限。而颈部网络处于模型的中间过渡位置,起着承上启下的关键作用,它接收来自骨干网络的基础特征,并将其进一步整合,传递给检测头用于目标检测。将 ECA 模块加入颈部网络后,ECA 能够对这些中间特征进行更具针对性的通道注意力调整,从而让模型输出更贴合任务需求的特征表示,提升了模型在水工建筑物裂缝检测场景下的适应性和检测性能。

表1 ECA 模块不同引入位置时模型在水闸和大坝裂缝数据集上的评价指标

引入位置	水闸裂缝数据集			大坝裂缝数据集		
	$P/\%$	$R/\%$	mAP@0.5/ $\%$	$P/\%$	$R/\%$	mAP@0.5/ $\%$
未引入	79.2	79.0	82.1	79.0	78.8	81.8
骨干网络	81.1	79.6	82.7	79.4	79.0	82.2
颈部网络	82.3	80.2	84.2	81.1	80.6	83.1

2.4.2 注意力机制对模型性能的影响

为验证融合 ECA 模块的 YOLOv5s 模型在水工建筑物表面检测裂缝的优势,本文基于 YOLOv5s 模型分别加入了坐标注意力机制 CA^[23]、经典的空间通道注意力机制 CBAM^[24]和经典通道注意力机制 SE^[25]。水闸数据集的具体结果如表 2 所示,加入 ECA 模块后的 mAP@0.5 提升了 2.1 个百分点,精确率和召回率分别提升了 3.1 和 1.2 个百分点;大坝数据集中加入 ECA 模块后的 mAP@0.5 提高了 1.3 个百分点,精确率和召回率分别提升了 2.1 和 1.8 个百分点。从表 2 可知,YOLOv5s 并不适宜无差别地整合所有注意力机制。引入 SE 模块后,模型的漏检率反而呈现上升趋势,说明模型在处理深层信息方面仍存在局限性,未能有效提升其捕捉能力。引入传统的 CA 及 CBAM 模块后,模型大体上保持了原有的精度水平,这一结果表明,这两种注意力机制并未使网络性能产生实质上的提升。相较于其他 3 种典型的注意力机制,ECA 在水闸裂缝识别任务中表现优越,有效验证了其在提高模型检测精度方面的能力。

表 2 融合不同注意力机制时模型在水闸和大坝裂缝数据集上的性能对比

注意力机制	水闸裂缝数据集			大坝裂缝数据集		
	P/ %	R/ %	mAP@0.5/ %	P/ %	R/ %	mAP@0.5/ %
原模型	79.2	79.0	82.1	79.0	78.8	81.8
CA	81.9	78.6	81.9	80.7	79.8	82.2
CBAM	79.8	79.4	83.2	80.4	79.0	81.9
SE	78.8	77.9	80.2	81.9	79.9	82.8
ECA	82.3	80.2	84.2	81.1	80.6	83.1

2.4.3 CSR 值对模型性能的影响

为量化不同 CSR 取值对模型检测性能与计算

表 3 不同 CSR 取值时模型在水闸和大坝裂缝数据集上的性能对比

CSR 值	水闸裂缝数据集					大坝裂缝数据集				
	P/%	R/%	mAP@0.5/%	参数量/10 ⁶	GFLOPs	P/%	R/%	mAP@0.5/%	参数量/10 ⁶	GFLOPs
0.50	78.4	76.2	74.1	5.1	15.3	75.4	74.1	76.5	5.1	15.3
0.75	82.3	79.8	78.5	6.8	19.7	79.1	77.9	80.5	6.8	19.7
1.00	85.9	81.9	84.3	8.9	26.3	80.8	79.5	82.9	8.9	26.3
1.25	85.9	79.2	81.9	11.1	31.2	81.2	79.2	81.5	11.1	31.2

表 4 各改进模型在水闸和大坝裂缝数据集上的评价指标

模型	ECA	DCNv2	SEH	参数量	GFLOPs	水闸裂缝数据集			大坝裂缝数据集		
						P/%	R/%	mAP@0.5/%	P/%	R/%	mAP@0.5/%
1	×	×	×	7022326	15.9	79.2	79.0	82.1	79.0	78.8	81.8
2	√	×	×	7022329	15.9	82.3	80.2	84.2	81.1	80.6	83.1
3	×	√	×	7084817	15.5	83.1	81.0	83.8	81.4	80.7	83.5
4	×	×	√	8901110	26.3	83.9	81.9	84.3	80.8	79.5	82.9
5	√	√	×	7084820	15.5	84.2	81.5	83.9	82.3	81.4	84.2
6	√	×	√	8901113	25.5	84.4	82.3	84.1	82.1	78.4	82.3
7	×	√	√	8963601	25.8	84.6	82.6	85.9	83.2	81.9	83.7
8	√	√	√	8963604	25.8	85.8	83.7	86.2	83.8	82.2	84.9

注:“√”表示引入该模块,“×”表示未引入该模块。

效率的影响规律并确定最优 CSR 取值,以 YOLOv5s 融合 SEH 解耦头为基础,根据水工建筑物裂缝检测场景对模型精度与轻量化的双重需求,结合通道缩放的合理区间,设置 CSR 为 0.50、0.75、1.00、1.25 共 4 个水平,基于水闸与大坝裂缝数据集开展对比实验,以覆盖通道压缩、适配、过度扩张 3 种典型状态。

表 3 为不同 CSR 取值时模型在水闸和大坝裂缝数据集上的性能对比。由表 3 可知,CSR 为 0.50 时 SEH 解耦头通道过度压缩,细粒度特征丢失,模型检测性能明显下降;CSR 为 0.75 时 SEH 解耦头通道资源略有补充,性能较 CSR 为 0.50 时提升,但仍无法充分捕获复杂裂缝特征;CSR 为 1.25 时 SEH 解耦头通道冗余,导致计算量大幅上涨,浮点运算总量(GFLOPs)较 CSR 为 1.00 时增加了 4.9,基于水闸、大坝数据集的 mAP@0.5 指标较最优值分别低 2.4、1.4 个百分点;YOLOv5s 融合 SEH 解耦头的最优取值为 CSR=1.00,在该取值下,模型基本实现了检测性能峰值,同时参数量与 GFLOPs 处于小模型边缘部署合理区间,既保证了水工建筑物裂缝的精准检测,又避免了通道冗余带来的计算效率损失。

2.4.4 消融实验

本文设计了 8 组基于 YOLOv5s 模型的消融实验,并在统一的超参数配置与训练框架下实施,确保每组条件的一致性,评价指标如表 4 所示。以水闸裂缝数据集为例,模型 2 为了增强模型对重要特征的提取能力,在网络颈部部分最后一层加入了 ECA 注意力机制,mAP@0.5 提升了 2.1 个百分点,在几乎没有增加网络参数且 GFLOPs 与原模型一致的前提下,增强了网络的定位精度与边界框预测效能;为更好地捕捉裂缝的复杂形态,模型 3 在网络骨干部

分的最后一层用 DCNv2 可变形卷积替换了 C3 模块,mAP@0.5 提升了 1.7 个百分点;为使模型在检测头中分别处理分类、定位和置信度检测任务以防任务互相干扰,模型 4 用 SEH 模块替换 YOLOv5 耦合头,mAP@0.5 提升了 2.2 个百分点;模型 5、6、7 分别加入两个模块,网络性能依然保持良好,这表明它们之间未产生相互干扰,验证了模块融合的有效性。模型 8 在轻量化的基础上加入了可变形卷积和解耦头,mAP@0.5 提升了 4.1 个百分点,可见,将这两种元素共同整合进网络架构中,能更好地提升模型的特征提取能力。

图 6 展示了水闸影像的部分裂缝检测结果,橙色标记为漏检的裂缝,红色标记为误检的裂缝。该组数据存在较多背景噪声,裂缝特征不明显,YOLO-SEH 模型相较于 YOLOv5s 模型的检测结果表现出更高的完整性、置信度和更低的漏检率。

图像 2 的裂缝特征不明显,YOLOv5s 模型的检测完整性不足,而 YOLO-SEH 模型通过提取更丰富的特征信息,降低了漏检率。在裂缝灰度值与水闸材质很相似且有析钙的图像 3 中和存在水渍的图像 4 中,YOLOv5s 模型将水渍和析钙误检为裂缝,且存在漏检现象,优化后的所有模型纠正了这些误检,并

大幅提高了识别精度。对于图像 1 和图像 3,优化后模型的置信度和裂缝识别率有较大提升。综合分析表明,YOLO-SEH 模型在水闸裂缝检测中的性能明显提升,验证了其有效性。

图 7 为 YOLOv5s 和 YOLO-SEH 模型在不同数据集上的 $P-R$ 曲线,通过计算曲线下面积可以发现,曲线与坐标轴围成的区域面积与模型性能呈正相关关系,其中覆盖面积最大的曲线所对应的模型具有最优的综合分类能力。在水闸和大坝裂缝数据集上,YOLO-SEH 模型曲线覆盖面积更大,这表明其在精确率和召回率之间达到了更好的平衡,整体性能优于原模型。

图 8 对比了 YOLOv5s 与 YOLO-SEH 模型在 mAP@0.5 指标下的性能曲线。当训练迭代次数达到 75 次时,改进后的 YOLO-SEH 网络逐渐趋于稳定。通过对 YOLO-SEH 模型的性能曲线进行深入分析,可以发现其收敛特性较原模型具有优势:一方面,改进后算法的收敛速度更快;另一方面,在达到稳定状态后,其 mAP@0.5 较原模型有较大提升,这一结果验证了所提出改进方法的有效性。

2.4.5 模型性能对比

为保障对比实验的严谨性,本文在统一超参数

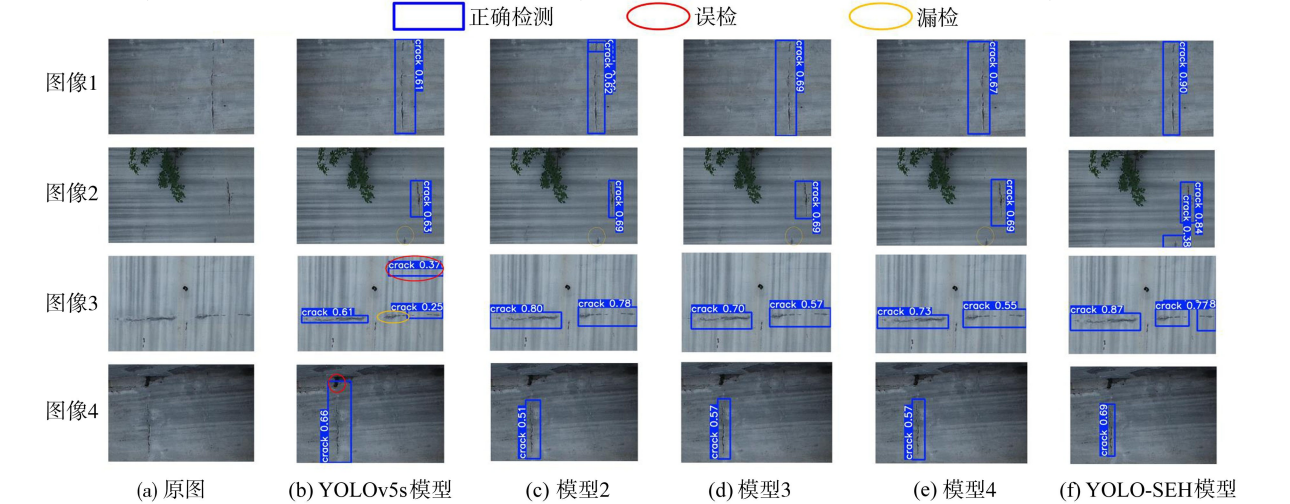


图 6 各改进模型对水闸裂缝图像的检测结果

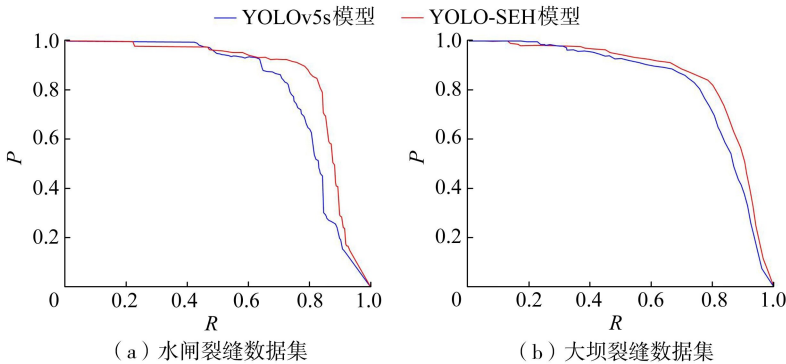


图 7 模型在水闸和大坝裂缝数据集上的 $P-R$ 曲线

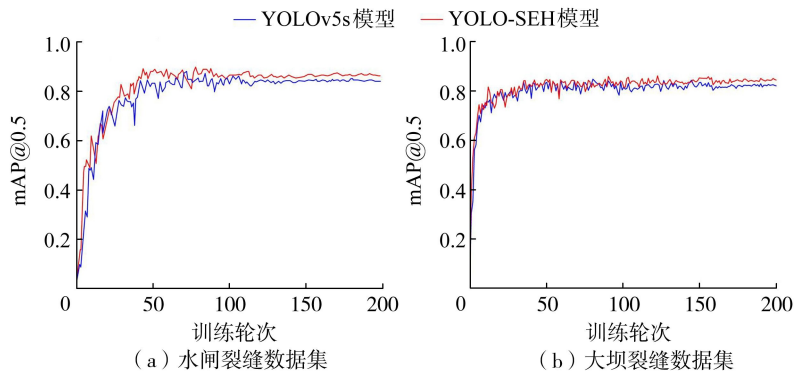


图 8 模型在水闸和大坝裂缝数据集上的 mAP@0.5 曲线

配置与训练框架的基础上,选取 Faster R-CNN^[26]、SSD300^[27]、YOLOv3^[28]、YOLOv5、YOLOv7^[29]、YOLOv8^[30]共 6 种目标检测模型,与本文提出的改进模型 YOLO-SEH 进行性能对比。评估指标聚焦于水闸和大坝两类裂缝数据集的检测精度。由表 5 和表 6 可知,相较于主流检测模型,YOLO-SEH 模型在核心精度指标上优势显著,在水闸裂缝数据集中,其精度、召回率和 mAP@0.5 分别达到 85.8%、83.7%、86.2%,较 YOLOv8 模型分别提升 3.1、5.9、1.7 个百分点,较 YOLOv5 模型分别提升 6.6、4.7、4.1 个百分点,有效增强了复杂场景下目标的识别与捕获能力,降低了漏检与误判风险。同时,该模型在精度跃升的前提下实现了计算量的可控增长,其参数量为 8.96×10^6 、GFLOPs 为 25.8,相较 YOLOv8 模型的参数量 (10.66×10^6) 和 GFLOPs (28.7) 分别降低了 16% 和 10.1%,仅略高于 YOLOv7、YOLOv3 等轻量化模型的参数规模,且保持了模型实时推理速度 (FPS) 为 31.41,较 YOLOv8 和 YOLOv7 模型分别提升了 12.6% 和 22.2%,仅略低于 YOLOv5 模型的推理速度。

表 5 不同网络中在水闸裂缝数据集的实验结果

模型	P/%	R/%	mAP@0.5%	FPS	参数量/ 10^6	GFLOPs
Faster R-CNN	72.9	53.5	72.0	2.74	41.18	10.4
SSD300	76.1	54.0	75.6	15.40	26.29	31.3
YOLOv3	77.2	51.8	76.2	20.13	8.30	13.1
YOLOv5	79.2	79.0	82.1	35.71	7.00	15.6
YOLOv7	80.6	76.5	82.8	25.70	5.90	13.1
YOLOv8	82.7	77.8	84.5	27.90	10.66	28.7
YOLO-SEH	85.8	83.7	86.2	31.41	8.96	25.8

表 6 不同网络中在大坝裂缝数据集的实验结果

模型	P/%	R/%	mAP@0.5%	FPS	参数量/ 10^6	GFLOPs
Faster R-CNN	73.8	55.7	73.2	3.54	41.18	10.4
SSD300	77.1	57.3	78.1	16.70	26.29	31.3
YOLOv3	78.2	56.7	79.7	22.05	8.30	13.1
YOLOv5	79.0	78.8	81.8	36.12	7.00	15.6
YOLOv7	81.2	75.4	82.4	29.97	5.90	13.1
YOLOv8	83.4	79.9	83.2	26.53	10.66	28.7
YOLO-SEH	83.8	82.2	84.9	33.22	8.96	25.8

图 9 为各模型对水闸裂缝部分图像的检测结果。在裂缝区域的检测效能评估中,尤其针对裂缝特征较为隐蔽的图像 2 与图像 3,尽管 Faster R-CNN、SSD300、YOLOv3 以及 YOLOv5 模型均能够实现对裂缝的基本识别,但它们在检测完整性及准确性方面存在不足。图像 3 中树枝与裂缝的灰度值存在相似性,导致 YOLOv3 模型错误地将部分树枝的灰度值识别为裂缝。相比之下,YOLO-SEH 模型不仅成功检测出了所有的裂缝,同时有效地避免了将树枝误判为裂缝的情况。

图 10 为各模型对大坝裂缝部分图像的检测结果。在所示的场景中,前几个模型都未能检测出图像 5 靠近边缘的部分裂缝以及在图像 7 主裂缝上的细小裂缝,而 YOLO-SEH 模型在特征提取方面展现出了更强的能力,能够识别出以往模型难以检测到的裂缝细节,从而显著降低了漏检率。由图像 4 和图像 6 可以看出,改进后的模型在识别先置信度较低的目标方面,展现出了更优的性能。

2.5 讨论

本文针对水闸与大坝两类裂缝数据集设计并实施了一系列对比实验。实验结果表明,YOLO-SEH 模型在裂缝识别任务中表现优异。在检测精度方面,该模型不仅能够更准确地定位裂缝区域,还有效降低了漏检率和误检率,且在复杂背景干扰下表现出优异的抗噪性能;在计算效率方面,YOLO-SEH 与 YOLOv5s 模型的推理速度相当,且显著快于其他对比模型,水工建筑物单张图像推理耗时约 0.031 s,展现了精度与效率间的良好平衡。YOLO-SEH 模型凭借其良好的特征表征能力和背景噪声抑制效果,能够高效地检测复杂表面水工建筑物的裂缝形态。

由于 YOLO-SEH 模型的推理速度低于 YOLOv5s 模型,为了优化检测性能与速度之间的权衡,未来将专注于研发更为轻量化的解耦头。本文仅对水闸和大坝两类水工建筑物的裂缝进行检测,而实际水工建筑物中的劣化现象复杂多样,下一步

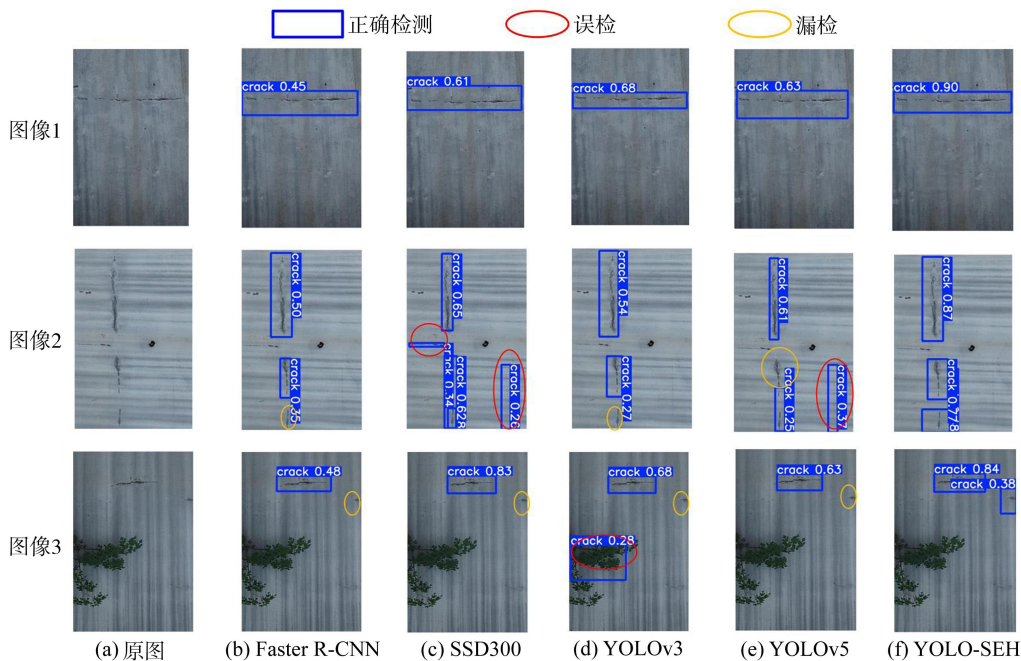


图9 各模型对水闸裂缝图像的检测效果

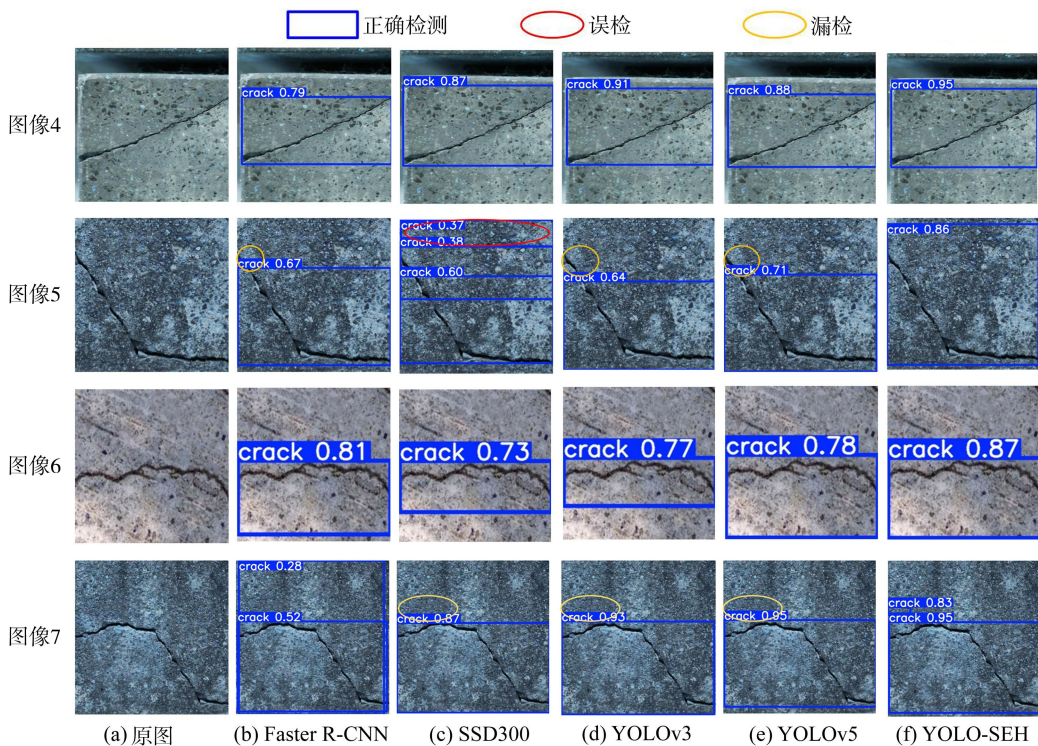


图10 各模型对大坝裂缝图像的检测效果

可采用多模态采集手段获取不同工况下典型水工建筑物的劣化样本,并将 YOLO-SEH 模型的检测目标从单一裂缝拓展至混凝土碳化、渗漏、冻融破坏等多类典型水工劣化类型,进一步提升模型的泛化能力和检测精度,使其更好地适应水工建筑物复杂多变的实际环境。此外,后续研究可考虑采用更高性能的图像采集设备,以获取分辨率更高、质量更优的数据样本,从而进一步提升模型对水工建筑物裂缝的识别能力。

3 结 语

本文针对水工建筑物表面水渍、析钙等现象干扰裂缝检测精度的问题,提出了基于可变形卷积与缩放解耦头的目标检测模型 YOLO-SEH。模型骨干部分引入可变形卷积 DCNv2,有效强化了复杂背景下裂缝的识别能力,为后续特征提取奠定坚实基础;颈部集成 ECA 注意力机制,能够精准区分裂缝与水渍、析钙的特征差异,显著降低了误检风险;头部引

入自适应缩放解耦头 (SEH), 可根据检测网络的具体规模灵活调整通道数量, 在仅产生轻微速度损失的前提下, 大幅提升裂缝检测精度, 各部分协同作用形成了适配复杂水工场景的高效检测架构。实验结果表明, 相较于 YOLOv5 模型, YOLO-SEH 模型在大坝和水闸裂缝数据集上均具有更优异的检测性能, mAP@0.5 分别提升了 3.1、4.1 个百分点, 精确率分别提升了 4.8、6.6 个百分点, 召回率分别提升了 3.4、4.7 个百分点。在与 Faster R-CNN、SSD300、YOLOv3、YOLOv7 等主流检测模型的对比实验中, YOLO-SEH 模型取得最优检测结果。该模型能够在水工建筑物复杂背景干扰下保持稳定的检测效能, 实现检测精度与速度的良好平衡, 为水工建筑物裂缝检测提供切实有效的技术解决方案。

参考文献:

[1] 王一兵,包腾飞,高治鑫. 基于 LabVIEW+VDM 的混凝土坝裂缝检测方法[J]. 水利水电科技进展, 2021, 41 (5): 76-82. (Wang Yibing, Bao Tengfei, Gao Zhixin. Crack detection method of concrete dams based on LabVIEW+VDM[J]. Advances in Science and Technology of Water Resources, 2021, 41(5): 76-82. (in Chinese))

[2] 徐辉,贺俊吉,方亮. 基于图像显著性的大坝裂缝检测[J]. 工业控制计算机, 2018, 31(9): 19-20. (Xu Hui, He Junji, Fang Liang. Dam crack detection based on image saliency[J]. Industrial Control Computer, 2018, 31 (9): 19-20. (in Chinese))

[3] 陈伟,范新南,李敏,等. 基于 Gabor 算子的人工蜂群算法大坝裂缝检测[J]. 微处理机, 2015, 36 (4): 32-38. (Chen Wei, Fan Xinnan, Li Min, et al. Dam crack detection of artificial colony algorithm based on Gabor operator[J]. Microprocessors, 2015, 36 (4): 32-38. (in Chinese))

[4] 张大伟,许梦钊,马莉,等. 水下大坝裂缝图像分割方法研究[J]. 软件导刊, 2016, 15 (9): 170-172. (Zhang Dawei, Xu Mengzhao, Ma Li, et al. Research on underwater dam crack image segmentation method [J]. Software Guide, 2016, 15(9): 170-172. (in Chinese))

[5] 张来胜,李亚萍. 融合视觉处理的堤坝水下裂缝检测方法[J]. 软件导刊, 2022, 21 (6): 188-193. (Zhang Laisheng, Li Yaping. Underwater crack detection method of dam with visual processing[J]. Software Guide, 2022, 21 (6): 188-193. (in Chinese))

[6] 张小伟,包腾飞,高兴和. 基于计算机视觉的混凝土坝裂缝检测方法[J]. 水利水电科技进展, 2021, 41 (5): 83-88. (Zhang Xiaowei, Bao Tengfei, Gao Xinghe. Crack detection method of concrete dams based on computer vision[J]. Advances in Science and Technology of Water Resources, 2021, 41(5): 83-88. (in Chinese))

[7] 张小伟,包腾飞. 基于局部大津阈值与区域生长的坝面细小裂缝识别分割算法[J]. 水电能源科学, 2022, 40 (2): 97-100. (Zhang Xiaowei, Bao Tengfei. Identification and segmentation algorithm of small cracks on dam surface based on local OTSU threshold and regional growth [J]. Water Resources and Power, 2022, 40 (2): 97-100. (in Chinese))

[8] Liu Zhiwei, Suandi S A, Ohashi T, et al. Tunnel crack detection and classification system based on image processing [C]//Machine Vision Applications in Industrial Inspection X. San Jose:SPIE, 2002:145-152.

[9] 杨吉才. 基于 ID3 决策树分类模型的自动评估学生技能水平的算法[J]. 无线互联科技, 2024, 21 (18): 17-22. (Yang Jicai. Algorithm for automatically evaluating student skill levels based on ID3 decision tree classification model [J]. Wireless Internet Science and Technology, 2024, 21(18): 17-22. (in Chinese))

[10] 牛慧余,包腾飞,李扬涛,等. 基于改进 Mask R-CNN 的混凝土坝裂缝像素级检测方法[J]. 水利水电科技进展, 2023, 43 (1): 87-92. (Niu Huiyu, Bao Tengfei, Li Yangtao, et al. Pixel-level crack detection method of concrete dam based on improved Mask R-CNN [J]. Advances in Science and Technology of Water Resources, 2023, 43(1): 87-92. (in Chinese))

[11] Zhu Wei, Zhang Hui, Eastwood J, et al. Concrete crack detection using lightweight attention feature fusion single shot multibox detector [J]. Knowledge-Based Systems, 2023, 261:110216.

[12] 刘万军,高健康,曲海成,等. 多尺度特征增强的遥感图像舰船目标检测[J]. 自然资源遥感, 2021, 33 (3): 97-106. (Liu Wanjun, Gao Jiankang, Qu Haicheng, et al. Ship detection based on multi-scale feature enhancement of remote sensing images [J]. Remote Sensing for Natural Resources, 2021, 33(3): 97-106. (in Chinese))

[13] 杨富强,余波,赵嘉彬,等. 基于改进 YOLOv3 的桥梁底部裂缝目标检测方法[J]. 中国科技论文, 2022, 17 (3): 252-259. (Yang Fuqiang, Yu Bo, Zhao Jiabin, et al. Crack target detection method of bridge bottom based on YOLOv3[J]. China Sciencepaper, 2022, 17(3): 252-259. (in Chinese))

[14] Li Yangtao, Bao Tengfei. A real-time multi-defect automatic identification framework for concrete dams via improved YOLOv5 and knowledge distillation[J]. Journal of Civil Structural Health Monitoring, 2023, 13 (6/7): 1333-1349.

[15] Zhu Xizhou, Hu Han, Lin S, et al. Deformable convnets v2: More deformable, better results [C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 9308-9316.

[16] Wang Qilong, Wu Banggu, Zhu Pengfei, et al. ECA-Net:

efficient channel attention for deep convolutional neural networks [C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle:IEEE,2020;11534-11542.

[17] 施慧慧,徐雁南,滕文秀,等. 高分辨率遥感影像深度迁移可变形卷积的场景分类法 [J]. 测绘学报,2021,50 (5):652-663. (Shi Huihui,Xu Yannan,Teng Wenxiu,et al. Scene classification of high-resolution remote sensing imagery based on deep transfer deformable convolutional neural networks [J]. Acta Geodaetica et Cartographica Sinica,2021,50 (5):652-663. (in Chinese))

[18] 何旺,钮新强,田金章,等. 基于 MobileNetV2-DeepLabv3+的混凝土坝水下裂缝语义分割模型 [J]. 水利水电科技进展,2024,44 (6):106-112. (He Wang,Niu Xinqiang,Tian Jinzhang,et al. Semantic segmentation model for underwater cracks in concrete dams based on MobileNetV2-DeepLabv3+[J]. Advances in Science and Technology of Water Resources,2024,44 (6):106-112. (in Chinese))

[19] Li C Y,Li L L,Jiang H L,et al. YOLOv6:a single-stage object detection framework for industrial applications [J/OL]. arXiv: 2209. 02976, 2022. <https://doi.org/10.48550/arXiv.2209.02976>.

[20] 叶栢铖,朱尤攀,周永康,等. 轻量级目标检测算法综述 [J]. 红外技术,2025,47 (3):289-298. (Ye Baicheng,Zhu Youpan,Zhou Yongkang,et al. Review of lightweight object detection algorithms [J]. Infrared Technology,2025,47 (3):289-298. (in Chinese))

[21] Liu Yahui,Yao Jian,Lu Xiaohu,et al. DeepCrack:a deep hierarchical feature learning architecture for crack segmentation [J]. Neurocomputing,2019,338:139-153.

[22] 高文飞. 基于 CW-YOLOv5 网络的大坝裂缝检测研究 [J]. 水利技术监督,2024 (12): 30-33. (Gao Wenfei. Research on dam crack detection based on CW-YOLOv5 network [J]. Technical Supervision in Water Resources,2024 (12):30-33. (in Chinese))

[23] Hou Qibin,Zhou Daquan,Feng Jiashi. Coordinate attention for efficient mobile network design [C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13713-13722.

[24] Woo S,Park J,Lee J Y,et al. CBAM:convolutional block attention module [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer,2018;3-19.

[25] Hu Jie, Shen Li, Sun Gang. Squeeze-and-excitation networks [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE,2018;7132-7141.

[26] Ren Shaoqing, He Kaiming, 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,2017,39 (6):1137-1149.

[27] Liu Wei, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [C]//European Conference on Computer Vision. Amsterdam:Springer,2016;21-37.

[28] Farhadi A, Redmon J. YOLOv3: an incremental improvement [C]//Computer Vision and Pattern Recognition. Berlin:Springer,2018,1804:1-6.

[29] Wang C Y,Bochkovskiy A,Liao H Y M. YOLOv7;trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver:IEEE,2023;7464-7475.

[30] Ultralytics. YOLOv8 (2023) [EB/OL]. (2023-04-01). <https://github.com/ultralytics/ultralytics>. (收稿日期:2025-03-25 编辑:俞云利)

(上接第 99 页)

[17] 欧斌,张才溢,陈德辉,等. 基于改进 EMD-LSTM 的混凝土坝变形预测模型 [J]. 水利水电科技进展,2024,44 (6):93-99. (Ou Bin,Zhang Caiyi,Chen Dehui,et al. Deformation prediction model for concrete dams based on improved EMD-LSTM [J]. Advances in Science and Technology of Water Resources,2024,44 (6):93-99. (in Chinese))

[18] 赵宇,郑东健,冉成,等. 基于 QR-KOA-ITransformer-BiLSTM 的混凝土坝变形和变形区间预测模型 [J]. 河海大学学报 (自然科学版),2026,54 (2):102-108. (Zhao Yu,Zheng Dongjian,Ran Cheng,et al. Concrete dam deformation and deformation interval prediction model based on QR-KOA-ITransformer-BiLSTM [J]. Journal of Hohai University (Natural Sciences),2026,54 (2):102-108. (in Chinese))

[19] 王译玲,丁勇,李登华. 耦合 DWT-IFLOF 的大坝监测数据异常检测算法 [J]. 中国农村水利水电,2024 (7):203-209. (Wang Yiling,Ding Yong,Li Denghua. Anomaly detection algorithm for dam monitoring data coupled with DWT-IFLOF [J]. China Rural Water and Hydropower,2024 (7):203-209. (in Chinese))

[20] 李元梦,李登华,丁勇. 基于 DBSCAN 的大坝安全监测异常数据检测算法 [J]. 水电能源科学,2024,42 (1):149-152. (Li Yuanmeng,Li Denghua,Ding Yong. Abnormal data detection algorithm for dam safety monitoring based on DBSCAN [J]. Water Resources and Power,2024,42 (1):149-152. (in Chinese))

(收稿日期:2025-03-18 编辑:俞云利)