

采用多尺度特征增强的路面病害检测模型

胡鹏¹, 夏晓华¹, 钟预全¹, 段智威¹, 姚运仕¹, 成高立²

(1. 长安大学道路施工技术与装备教育部重点实验室, 710064, 西安;

2. 陕西高速机械化工程有限公司, 710038, 西安)

摘要: 针对现有网络多尺度特征提取能力不足造成路面病害因尺寸差异难以完全识别的问题,提出了一种多尺度特征增强的路面病害检测模型。构建基于混合空洞卷积的快速空间金字塔池化模块,通过堆叠不同膨胀系数的空洞卷积进一步扩大网络感受野,以实现更大范围上下文信息的捕捉,并保留更多的空间信息;设计多路径特征融合网络,通过多分支和跳跃连接实现跨层级的特征捕捉,并减少特征融合过程中的信息丢失;采用 K-means 聚类算法结合交叉比获得合理的锚点框;在损失函数中,设计一种面积惩罚项并设置下降梯度,提高预测框回归精度与效率;通过引入跨通道交互的高效注意力实现模型重要通道间的交互。实验结果表明:所提模型的检测精度比原模型 YOLOv5s 提高了 4.0%;与 Faster R-CNN、CenterNet 等经典模型和 YOLOv8s、YOLOv7n-tiny 等先进模型相比,检测精度提高了 1.0%~17.9%。模型经 TensorRT 加速引擎优化加速后,在 NVIDIA Jetson TX2 与 NVIDIA Jetson Nano 平台上的检测速率提高近 1 倍,同时不影响检测精度。

关键词: 路面病害检测;多尺度特征增强;混合空洞卷积;特征融合网络;高效通道注意力;嵌入式平台

中图分类号: TP277 **文献标志码:** A

DOI: 10.7652/xjtuxb202502016 **文章编号:** 0253-987X(2025)02-0156-14

Pavement Disease Detection Model Based on Multi-Scale Feature Reinforcement

HU Peng¹, XIA Xiaohua¹, ZHONG Yuquan¹, DUAN Zhiwei¹, YAO Yunshi¹, CHENG Gaoli¹

(1. Key Laboratory of Road Construction Technology and Equipment of MOE, Chang'an University, Xi'an 710064, China;

2. Shaanxi Expressway Mechanization Engineering Co., Ltd., Xi'an 710038, China)

Abstract: To address the difficulty in fully identifying pavement defects of different sizes caused by insufficient multi-scale feature extraction capability of existing networks, a pavement disease detection model based on multi-scale feature reinforcement was proposed in this paper. Firstly, a fast spatial pyramid pooling module based on mixed dilated convolution was constructed, and by stacking dilated convolutions with different dilation coefficients, the network receptive field was further expanded to capture a larger range of contextual information and preserve more spatial information. Next, a multi-path feature fusion network was designed to achieve cross level feature capture and reduce information loss during the feature fusion process through multiple branches and skip connections. The K-means clustering algorithm was used together with the

收稿日期: 2024-06-21。 作者简介: 胡鹏(1999—),男,硕士生;夏晓华(通信作者),男,副教授,博士生导师。 基金项目: 陕西省重点研发计划资助项目(2024GX-YBXM-197);陕西省交通运输厅交通科研项目(24-74K);秦创原引用高层次创新创业人才项目(QCYRCXM-2022-352)。

网络出版时间: 2024-08-30

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20240829.1738.006>

Intersection over Union to obtain reasonable anchor boxes. In addition, a penalty term for area was designed in the loss function and a descent gradient was set up to improve the accuracy and efficiency of the predicted box regression. Finally, efficient attention through cross channel interaction was introduced to achieve interaction between important channels in the model. Experimental results show that in terms of detection accuracy, the proposed model was 4.0% higher than the original model YOLOv5s and 1.0% to 17.9% higher than classical models such as Faster R-CNN and CenterNet and advanced models such as YOLOv8s and YOLOv7n-tiny. After optimization with TensorRT acceleration engine, the detection speed on NVIDIA Jetson TX2 and NVIDIA Jetson Nano embedded platforms nearly doubled without compromising the detection accuracy.

Keywords: pavement disease detection; multi-scale feature reinforcement; mixed dilated convolution; feature fusion network; efficient channel attention; embedded platform

随着交通运输领域的快速发展,道路总里程不断增加,交通安全问题已经成为人们关注的焦点。路面的状况直接关系到车辆行驶的安全,而路面病害的存在会导致路面状况恶化,对车辆行驶安全和路面使用寿命产生不利影响。因此,对路面病害及时检测并修复对延长路面使用寿命和保障车辆行驶安全具有重要意义。为及时客观地检测到车辙、裂缝、坑槽等路面病害,从而进行有效的路面养护,需要采用准确的病害检测手段^[1]。

目前,路面病害的检测方法包括人工巡检、传统图像处理算法以及机器学习等。然而,传统的人工巡检存在人工疲劳、漏检和成本高等问题,且易受主观判断的影响,存在较大的误差^[2]。由于道路病害形状各异,且路面存在油斑、路面标志等强干扰,传统图像处理算法对病害的检测性能有限^[3]。Tang 等^[4]利用灰度与形态学等算法检测裂缝,存在对噪声敏感和计算耗时长的问題。Zhao 等^[5]改进 Canny 算子,并在路面裂缝检测中成功应用,但易受斑马线等干扰。

随着信息技术的发展和深度学习算法的不断优化,卷积神经网络(convolutional neural network, CNN)在图像识别领域取得了巨大成功^[6]。近年来,深度学习在道路表面病害检测领域被广泛应用。Zhang 等^[7]首次将深度学习算法应用于路面裂缝检测问题的研究中,该方法在裂缝尺寸分布不均匀、背景噪声复杂等情况下仍然表现出良好的检测效果。张伟光等^[8]使用三层卷积层和两层全连接层设计了一个卷积神经网络模型,实现了准确的裂缝识别。王丽苹等^[9]基于 AlexNet 网络架构设计了一种路面裂缝检测模型,该模型能够准确且快速地识别裂缝。但是,上述研究仅判断图像中是否存在裂缝,不能对裂缝定位。孙朝云等^[10]将 VGG16 网络与 Faster R-CNN 结合并改进裂缝锚点框的生成方法,

提高了对小尺度裂缝的定位精度,但感受野有限且多尺度特征未得到充分的融合。许贵阳等^[11]在 Faster R-CNN 中引入引导锚框,在不同尺度上生成区域建议,提高了裂缝检测的针对性,但缺乏对多尺度特征进行有效融合的机制。罗晖等^[12]设计了 YOLOv4+DC+FL 算法,通过引入损失函数和迁移学习策略提高对多尺度路面病害的检测精度,但未考虑网络池化层和特征融合层等方面对多尺度的影响。Roy 等^[13]在 YOLOv5 的特征融合阶段增加小目标检测层,并结合了 Swin-Transformer 模块,虽然在特征融合阶段考虑了多尺度病害特征的增强,但忽略了池化阶段感受野有限和存在信息损失的问题。Li 等^[14]基于 YOLOv5 设计了一种从边缘到客户端的道路损坏检测系统,该系统具有较好的实时性,但没有针对路面病害多尺度特性去优化网络。周中等^[15]在 YOLOX 算法的基础上提出了一种名为 YOLOX-G 的实时裂缝图像检测算法,其特征融合阶段采用卷积和上下采样对多尺度特征进行融合,存在信息损失从而不利于多尺度裂缝的检测。Xiang 等^[16]通过在 YOLOv5s 特征融合网络中引入 Space-to-depth-Conv 模块提高对小尺度病害的检测精度,但依旧未考虑多尺度信息在融合过程中的损失。尽管上述方法在整体检测精度或实时性方面取得了效果,然而它们通常只关注单个或少数病害类别,未充分考虑实际路面病害种类的多样性和病害在空间上的大小差异。病害在空间上的大小差异导致路面病害特征呈现出多尺度的复杂性,现有网络在应对这种多尺度的路面病害特征时,往往会暴露出多尺度特征处理能力不足的问题,造成病害因尺寸差异大而不能实现准确的识别。

针对上述存在的问题,本文首先优选包含 7 种病害类别的开源数据集,相较于以往仅关注单一或

少数病害类别的研究,更全面地考虑了路面病害类别及尺寸差异。其次,基于 YOLOv5s 提出多尺度特征增强的路面病害检测模型,充分提取与融合多尺度特征,以提高不同尺度病害的识别与定位精度。本文的主要创新如下:

(1)针对池化阶段感受野有限和不同尺度池化操作导致信息丢失的问题,提出一种基于混合空洞卷积的快速空间金字塔池化模块(mixed dilated convolution-SPPF,MdC-SPPF),通过混合空洞卷积增强相邻像素点之间的空间关系,有效地捕捉更大范围的上下文信息,同时减少信息丢失,从而使得模型在处理不同尺度的特征时更加鲁棒。为进一步有效捕捉复杂的多尺度特征,通过一种不降维跨通道交互的注意力,显式关注重要通道间的多尺度目标区域,从而提高模型的鲁棒性。

(2)针对多尺度信息融合不充分和原始多尺度特征融合网络中存在部分信息损失的问题,设计一种多路径特征融合网络(multi-path feature fusion,MF-CSP)。通过构建多条并行路径,分别提取并融合不同尺度的特征,从而能够更全面地捕捉到病害在各个尺度上的信息特征,以提高模型的检测精度和对各种尺度病害的鲁棒性。

(3)针对锚点框会影响模型对多尺度目标的检测效果,本文改用交叉比(intersection over union,

IoU)帮助网络更好地学习不同尺寸的病害。为进一步提高多尺度目标检测的效果和稳定性,建立基于面积惩罚项和带梯度变化的损失函数(SGCIoU_Loss),提高损失函数对形状和尺度差异较大的目标的敏感度。

(4)在开源数据集上证明本文模型优越的多尺度特征处理性能,再将训练好的模型分别部署在 NVIDIA Jetson TX2 和 NVIDIA Jetson Nano 两个不同算力的平台上进行测试,验证该模型在嵌入式设备上进行病害检测的性能及在实际工程中应用的可行性。

1 病害检测模型

1.1 YOLOv5s 模型

在深度学习模型选择方面,选择 YOLO 系列模型^[17-18]。其中,YOLOv5^[19]作为该系列中典型且成熟的目标检测算法,广泛应用于路面病害检测领域中。YOLOv5 的网络结构^[16]如图 1 所示。YOLOv5 包含 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 这 5 种网络宽度和深度逐步增大的变体模型以满足不同检测任务的需求。为保证所选取的模型能够有效地捕捉数据特征,同时考虑到实际计算资源有限,本文选用网络深度和宽度较适中的 YOLOv5s 作为基线模型。

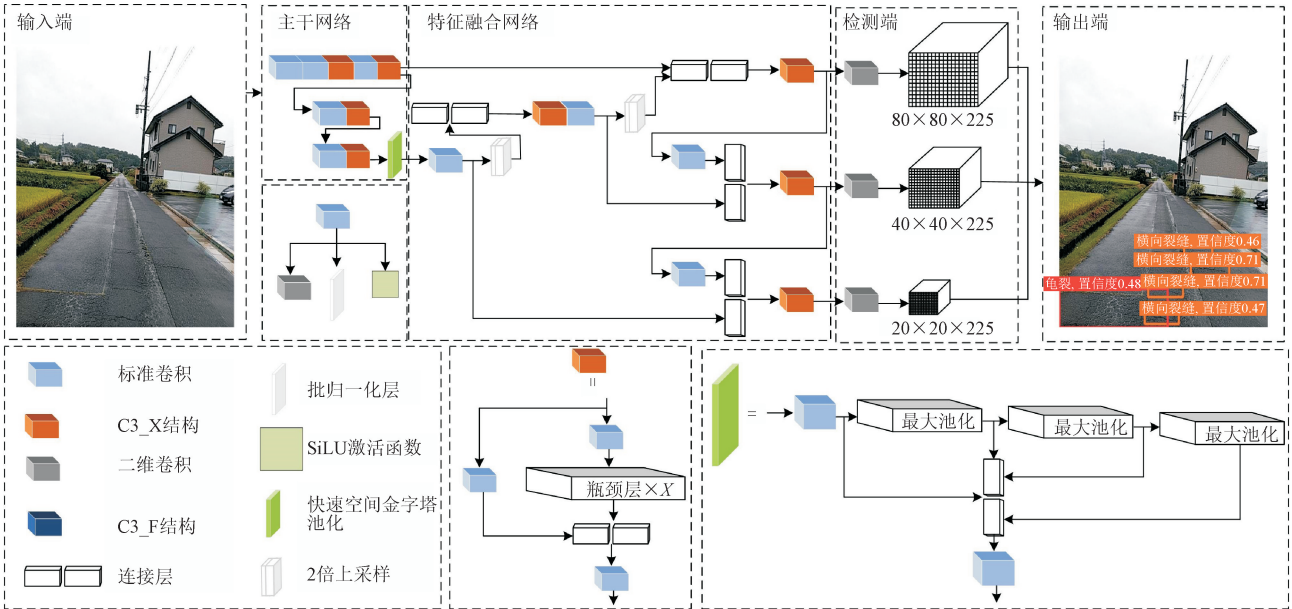


图 1 YOLOv5 网络结构^[16]
Fig. 1 The network structure of YOLOv5^[16]

1.2 多尺度特征增强的病害检测模型

本文提出的多尺度特征增强模型如图 2 所示,

主要结构包括主干网络、高效通道注意力、池化结构、多尺度特征融合网络及检测端。该模型以

YOLOv5s 模型作为主框架,采用本文提出的基于混合空洞卷积的快速空间金字塔池化作为模型的池化结构,通过高效通道注意力实现重要通道间的交互,利用所设计的多路径特征融合网络实现不同层

级间多尺度语义信息的融合, K -means 聚类算法结合 IoU 获得合理的锚点框。在损失函数方面,构建基于面积惩罚项和带梯度变化的损失函数,实现模型权重的不断优化。

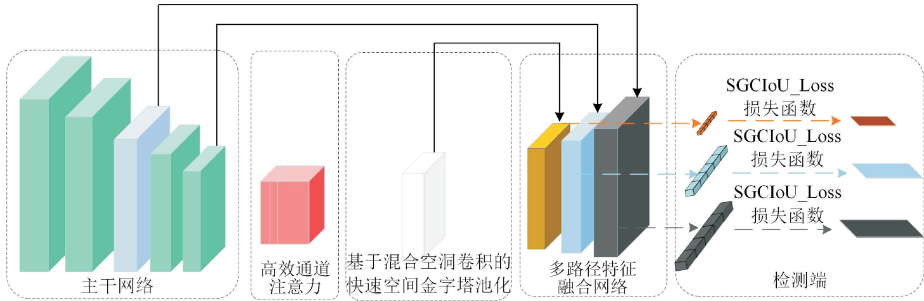
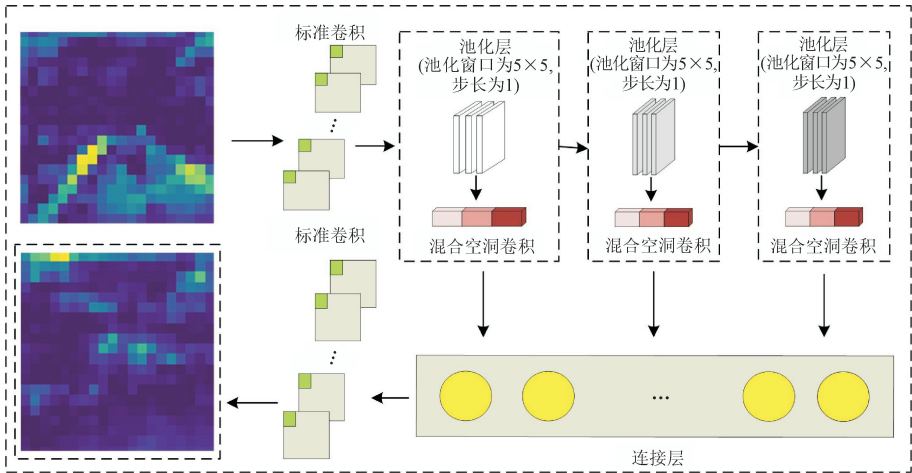
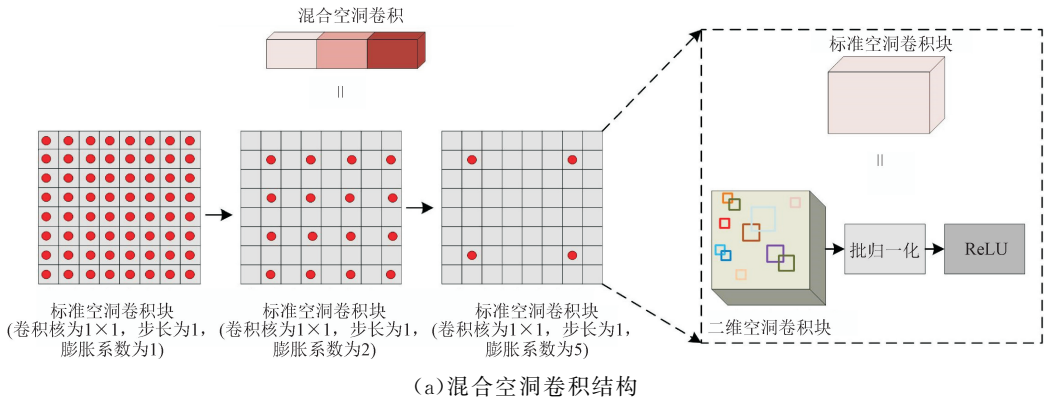


图 2 多尺度特征增强的病害检测模型
Fig. 2 Multi-scale feature enhancement model

1. 2. 1 MdC-SPPF 模块

SPPF 模块采用金字塔池化操作,将特征图经过 3 次池化,得到不同尺度的子特征图,后经连接层拼接成固定长度的特征向量,在保留空间信息的同时实现了多尺度的感受野,从而增强模型的鲁棒性

和泛化能力。但是,池化窗口大小和感受野仍然是有限的,并且池化运算后会造成特征尺度的下降,从而带来边界信息的缺失,不利于路面病害的检测。为此,本文提出基于混合空洞卷积的快速空间金字塔池化模块,结构如图 3 所示。



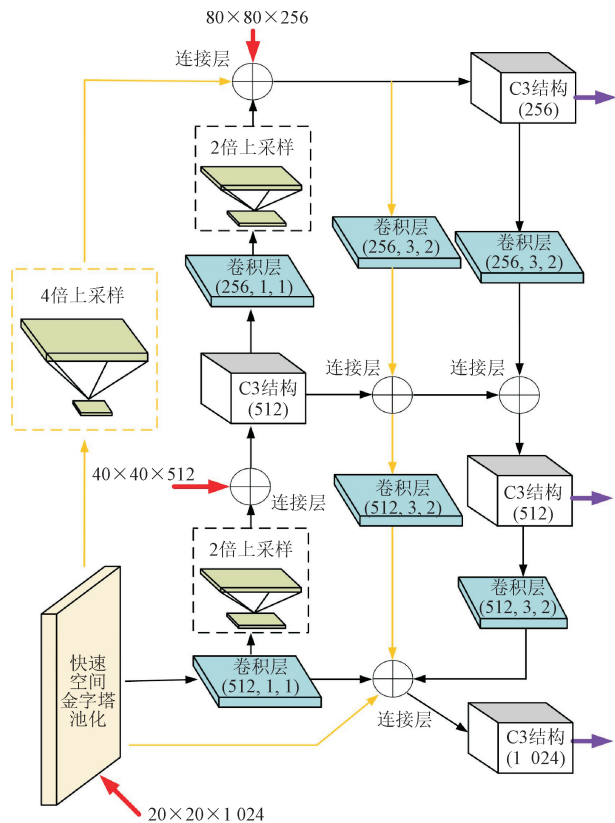
(b)MdC-SPPF 模块整体结构
图 3 MdC-SPPF 模块结构

Fig. 3 The structure of MdC-SPPF module

MdC-SPPF 模块中的具体实现步骤如下:首先,利用3个膨胀系数不同的标准空洞卷积^[20-21]搭建一个混合空洞卷积块,如图3(a)所示。其中,为使混合空洞卷积完全覆盖特征层区域,充分学习输入的信息,将膨胀系数设置为1、2、5。每个标准空洞卷积的卷积核大小、步长分别设置 1×1 和1,以高效地关注局部细节,并减少细节信息损失。接着,将该混合空洞卷积嵌入到每个池化层后,更好地捕捉大范围的语义信息,提高网络在处理多尺度路面病害特征时的性能。

1.2.2 多路径特征融合网络

YOLOv5 中的特征融合网络由路径聚合网络和特征金字塔网络构成,通过融合不同尺度的特征,以获取丰富的语义信息并更好地适应多尺度目标的检测任务。但是,不同尺度的特征在融合过程中会有一定的损失。为尽量减少这种损失,借助残差网络的思想,设计了一种多路径特征融合网络,以改善经卷积后可能导致语义信息丢失的问题,其结构如图4所示。首先,在 SPPF 结构后新增一个4倍上采样操作,将其连接到两个2倍上采样和卷积操作



注:红色箭头表示输入,维度为宽×高×通道数;卷积层中参数为(输出通道数,卷积核大小,步长);C3结构中参数为输出通道数。

图4 MF-CSP特征融合网络结构

Fig. 4 The structure of MF-CSP feature fusion network

的支路,以融合高分辨率和更高语义的特征。其次,在特征金字塔网络结构中,在每个C3结构处构建残差结构,维护信息的一致性,最小化信息的损失,有助于更好地捕获和融合不同层级的特征。最后,在SPPF结构后执行逐点卷积,将SPPF结构输出的特征与卷积块输出的特征进行拼接,增强特征的多样性,从而在路面病害检测任务中获得更好的性能。

1.2.3 瞄点框获取

原始算法通过K-means聚类算法和遗传算法来生成一组瞄点框。瞄点框定义了模型期望检测到的目标形状和大小,使模型可以更好地适应不同尺度的目标。然而,路面病害形状存在尺度变化较大、长宽比显著不同等特点^[22],导致初始聚类算法所采用的欧式距离公式不能很好地聚类出与病害形状相似的瞄点框,故在K-means聚类算法中结合IoU聚类出适合本文数据集的瞄点框。距离度量的定义如下

$$d = 1 - \varphi_{\text{IoU}}(g, a) \quad (1)$$

式中: d 表示两框间的距离; φ_{IoU} 表示两框交叉比函数; g, a 分别表示真实框、瞄点框。

1.2.4 SGCIoU_Loss 损失函数

损失函数作为CNN模型训练过程中的目标函数,决定模型权重的优化方向。最早使用IoU衡量预测框与真实框之间的重叠程度,即两框的交集面积除以它们的并集面积。在YOLOv5中改用回归性能较好的回归函数CIoU^[23],综合考虑了IoU、两框中心点间的距离与两框长宽比,使得预测结果具有较高的准确性和稳定性。IoU及CIoU的定义如下

$$R_{\text{IoU}} = \left| \frac{b_{\text{pred}} \cap b_{\text{true}}}{b_{\text{pred}} \cup b_{\text{true}}} \right| \quad (2)$$

$$R_{\text{CIoU}} = R_{\text{IoU}} - \frac{\rho^2(b, b^{\text{gt}})}{c^2} - \alpha\beta \quad (3)$$

式中: $b_{\text{pred}}, b_{\text{true}}$ 为预测框、真实框面积; b 和 b^{gt} 为预测框和真实框的中心点; c 为同时包含病害的预测框和真实框的最小区域对角线长度; $\rho^2(b, b^{\text{gt}})$ 为两框中心点间欧氏距离的平方; $\alpha\beta$ 为惩罚项,反映了预测框和真实框之间宽高比的差异。 β 和 α 分别为两框宽高比一致性参数和权衡量,定义如下

$$\beta = \frac{4}{\pi^2} \left(\arctan \frac{w_A}{h_A} - \arctan \frac{w_B}{h_B} \right)^2$$

$$\alpha = \frac{\beta}{1 - R_{\text{IoU}} + \beta}$$

式中: (w_A, h_A) 和 (w_B, h_B) 为真实框和预测框的宽高; β 由真实框的宽高和预测框的宽高动态调整,该值越小,表示两框宽高比越接近,使得模型更加注重

预测框的形状,提升目标检测的精度和鲁棒性; α 动态平衡 R_{IoU} 和 β ,确保在不同 R_{IoU} 水平下都有合适的梯度信号,提升模型的检测精度和可靠性。

病害形状通常长宽比差异较大,两框之间的重合度对两框中心点间欧氏距离较为敏感,不利于预测框的准确回归。为此,本文在 $CIoU$ 的基础上新增面积惩罚项,并以此建立新的回归函数 R_{SCIoU} ,定义如下

$$R_{SCIoU} = R_{CIoU} + s =$$

$$R_{CIoU} + \frac{2}{\pi} \arctan\left(\frac{w_A h_A + w_B h_B}{\max(w_A, w_B) \max(h_A, h_B)}\right) \quad (4)$$

$$s = \frac{2}{\pi} \arctan\left(\frac{w_A h_A + w_B h_B}{\max(w_A, w_B) \max(h_A, h_B)}\right) \quad (5)$$

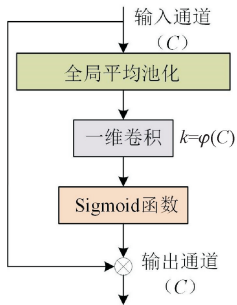
式中: s 为面积惩罚项。

为提升预测框回归精度以及提升模型的收敛速度,设计了基于面积惩罚项和带梯度变化的损失函数,定义如下

$$L_{SGIoU} = \lg 2 - \lg(1 + R_{SCIoU}) \quad (6)$$

1.2.5 高效通道注意力模块

在卷积神经网络中,常通过注意力机制突出目标特征。在计算注意力过程中,文献[24]证明了通道降维不利用通道间的学习,因此本文通过高效通道注意力模块[24]对重要的多尺度特征进行增强。该模块通过一维卷积,在不改变通道维度的同时实现跨通道交互,其结构如图5所示。通过关注通道间关系来进一步强化模型多尺度能力,而无需引入大量额外参数。一维卷积根据输入通道数自适应地选择其卷积核大小,确保通道注意力高效有效。



k ——一维卷积的卷积核大小; $\varphi(C)$ ——随通道数 C 变化的自适应函数。

图5 高效通道注意力模块结构

Fig. 5 The structure of efficient channel attention module

2 实验验证与结果分析

2.1 数据集准备与环境搭建

为验证本文改进方法的有效性及其模型多尺度性能,选取由全球道路损伤检测挑战赛提供的开源数

据集 Road Damage Detection Challenge 2020^[25], 包含捷克、印度和日本这3个国家的道路病害图像。本文优选了日本国家的数据,称作RDDJ数据集,因为该部分数据样本量最大且病害类型丰富,有助于评估模型的性能。RDDJ数据集包含10 506张分辨率为 600×600 像素的图像,含有D20(龟裂,6 199张)、D40(坑洞,2 243张)、D10(横向裂缝,3 979张)、D00(纵向裂缝,4 049张)、D44(白线模糊,3 995张)、D43(十字路口模糊,736张)、D50(井盖,3 553张)共7个病害类别。同时,数据集按照8:1:1的比例随机划分为训练集、验证集和测试集。

实验环境配置如下:操作系统为Windows10专业版,CPU型号为12th Gen Intel(R) Core(TM) i5-12400F,内存为32 GB。加速硬件方面,配备了一块 GeForce RTX 3060 Ti GPU,使用CUDA 11.3和cuDNN 8.2.1进行加速。编程语言选用Python 3.8,开发平台为Visual Studio Code 2019。此外,网络优化器为随机梯度下降,初始学习率为0.001,批次大小设置为16,训练轮数设置为300。

同时,结合迁移学习策略^[26]对本文模型进行训练,提高模型的泛化能力。将在大型开源数据集PASCAL Visual Object Classes Challenge 2007^[27]下训练得到的权重作为初始权重,并在路面病害数据集下对初始权重微调,最终得到本文的路面病害检测模型。

2.2 评价指标

本文采用在IoU为0.5下的平均精度均值(mean average precision, mAP@0.5)作为模型精度指标,该指标综合了召回率和准确率,其定义如下

$$mAP = \frac{1}{N} \int_0^1 p(r) dr \quad (7)$$

式中: N 为数据中的类别数; p 为准确率; r 为召回率。

模型体积、检测速率常用于衡量模型大小和检测实时性。为方便全面评价模型在实际应用中的可行性,本文使用综合评价指标 S 对模型进行评价,该指标综合考虑了模型体积、检测速率、mAP@0.5这3个指标,能够同时确保每个指标的贡献均衡。先对3个指标进行归一化处理,并对归一化后的结果求平均值,得到指标 S ,相关计算方法定义如下

$$S_V = 1 - \frac{V - V_{\min}}{V_{\text{avg}} V_{\text{std}}} \quad (8)$$

$$S_m = \frac{m_{AP} - m_{AP, \min}}{m_{AP, \text{avg}} m_{AP, \text{std}}} \quad (9)$$

$$S_F = \frac{F - F_{\min}}{F_{\text{avg}} F_{\text{std}}} \quad (10)$$

$$S = \frac{S_V + S_m + S_F}{3} \tag{11}$$

式中： S_V 、 S_m 、 S_F 分别为模型体积、检测精度、检测速率的归一化结果； V 、 m_{AP} 、 F 分别为模型体积、检测精度、检测速率；下标 min、avg、std 分别表示最小

值、平均值、标准差。

2.3 消融实验

为验证改进 K -means、SGCIOU_Loss、MdC-SPPF 和 MF-CSP 对病害检测性能的影响,共设计了 16 组消融实验,结果如表 1 所示。

表 1 消融实验结果
Table 1 Results of ablation experiment

基线模型	改进 K -means	SGCIOU_Loss	MdC-SPPF	MF-CSP	模型体积/MB	m_{AP}	检测速率/(帧·s ⁻¹)	S
✓					13.8	0.649	104	0.340
✓	✓				13.8	0.668	104	1.573
✓		✓			13.8	0.670	107	1.704
✓			✓		14.1	0.675	107	2.027
✓				✓	20.9	0.672	96	1.791
✓	✓	✓			13.8	0.672	103	1.832
✓	✓		✓		14.1	0.674	108	1.963
✓	✓			✓	20.9	0.676	97	2.051
✓	✓	✓	✓		14.1	0.677	106	2.156
✓	✓	✓		✓	20.9	0.666	96	1.402
✓	✓		✓	✓	21.2	0.679	92	2.241
✓		✓	✓	✓	21.2	0.676	94	2.048
✓		✓	✓		14.1	0.669	106	1.637
✓		✓		✓	20.9	0.675	96	1.985
✓			✓	✓	21.2	0.678	96	2.178
✓	✓	✓	✓	✓	21.2	0.683	94	2.502

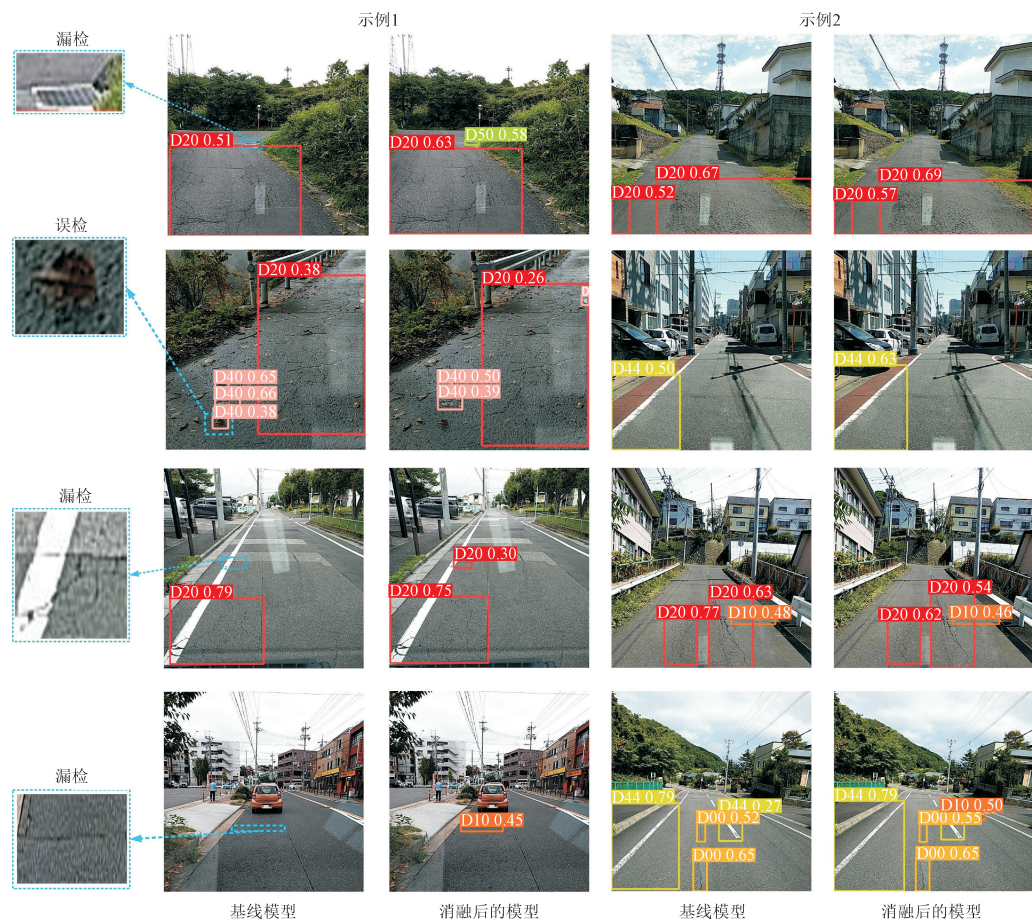
注:表中加粗结果为最优值。

由表 1 可看出,各改进方法均能提升模型的检测精度。首先,改进后的 K -means 聚类算法能够生成更符合病害尺度特点的锚点框,因此提高了模型的鲁棒性, $mAP@0.5$ 提高了 1.9%。其次,SGCIOU_Loss 损失函数有助于提高预测框回归的准确性,从而使模型的 $mAP@0.5$ 提高了 2.1%。引入 MdC-SPPF 模块扩大了感受野,且有效地改善了因池化导致分辨率减小造成的语义信息丢失问题,提高了模型对多尺度特征的表达能力, $mAP@0.5$ 提高了 2.6%。添加 MF-CSP 后,模型能够学习到更丰富的多尺度语义信息, $mAP@0.5$ 提高了 2.3%。这些结果证明了各改进方法在提高模型检测精度方面的有效性。

通过消融实验结果可知,改进的 K -means、SGCIOU_Loss 损失函数、MdC-SPPF 和 MF-CSP 这 4 种方法组合得到的模型,能够获得最高的 $mAP@0.5$,虽然模型的检测速率有所降低,但依旧

达到了 94 帧/s,满足实时性检测的要求。此外,该模型还获得最高的 S ,达到 2.502,较好地兼顾了模型体积、检测精度与检测实时性。

为进一步验证消融后的模型对路面病害的多尺度检测效果,将其与基线模型在测试集上进行了对比,其中部分典型结果如图 6 所示。可以看出,相较于基线模型,消融后的模型在多尺度特征提取方面表现出优越的性能,能更准确地检测路面病害。如图 6 中的示例 1 所示,基线模型对于远处的小尺度病害、与病害特征相近的干扰物(例如图 6 示例 1 中第二行中的落叶)会出现漏检和误检的情况,但消融后的模型在处理类似病害时表现出较高的可靠性。在示例 2 中,消融后的模型提高了路面病害检测的置信度,其在多尺度特征提取与融合方面的有效性得到了验证。综上可知,经消融后获得的模型在多尺度病害检测能力方面得到了提升。



注:检测结果以“病害类别 置信度”的形式呈现。

图 6 模型消融前后的效果对比

Fig. 6 Comparison of the effects before and after ablation of the model

2.4 注意力模块实验对比

为验证高效通道注意力模块对模型多尺度特征的表达能力,与通道注意力的典型代表压缩与激励模块(squeeze and excitation, SE)和卷积注意力模块(convolutional block attention module, CBAM)进行对比,结果如表 2 所示。可以看出,高效通道注意力模块的检测精度最高,其次是 SE,但添加 CBAM 降低了模型的检测精度。

表 2 注意力模块实验对比

Table 2 Experimental comparison of attention module		
模型	模型体积/MB	m_{AP}
消融后模型	21.2	0.683
消融后模型+SE	22.3	0.685
消融后模型+CBAM	22.3	0.681
消融后模型+高效通道注意力模块	22.2	0.689

注:表中加粗结果为最优值。

图 7 展示了嵌入各注意力模块后生成的梯度加

权类激活映射 (gradient-weighted class activation mapping, Grad CAM) 可视化结果。依据实验数据,并结合可视化结果,可进一步直观地分析注意力模块对多尺度病害区域的感知能力。可以看出,相比 SE 和 CBAM,高效通道注意力模块能有效地提高多尺度目标区域的感知效果,同时对背景有一定的抑制的作用。这是因为不降低通道维度更有利于各通道间的交互,进而突出多尺度目标特征。SE 和 CBAM 因为涉及到通道降维的过程,所以对多尺度特征的关注和对背景的抑制效果不如高效通道注意力模块。

2.5 优越性检验

为进一步验证本文模型的优越性,在相同的实验条件下进行对比实验。对比模型包括 Faster R-CNN、SSD 等常用模型以及 YOLOv8s、YOLOv6n、YOLOv7n-tiny 等较新的目标检测模型。此外,还与引言中的文献[13,15-16]这 3 个较新的方法进行对比,结果如表 3 所示。

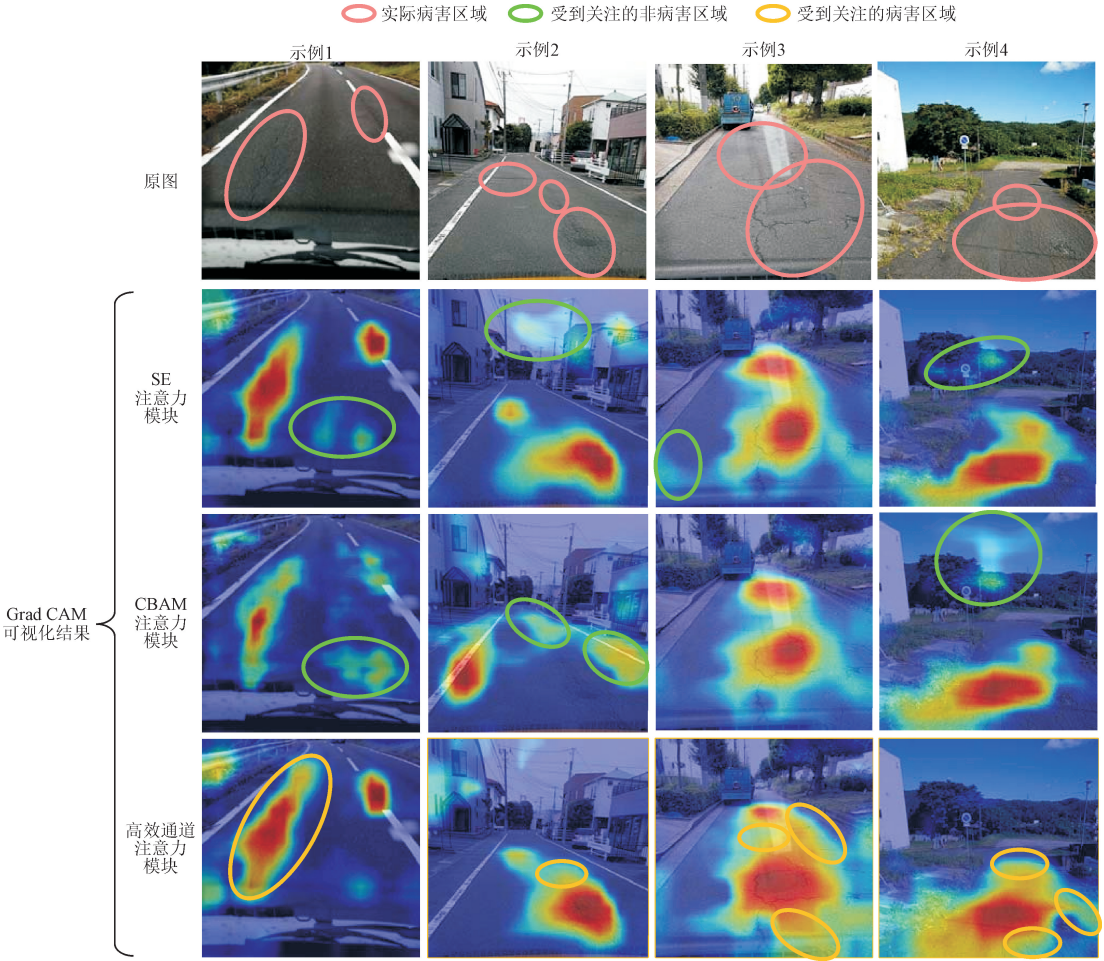


图 7 Grad CAM 可视化结果

Fig. 7 Grad CAM visualization results

表 3 本文模型与其他模型重要性能参数的对比

Table 3 Comparison of important performance parameters between proposed model and other models

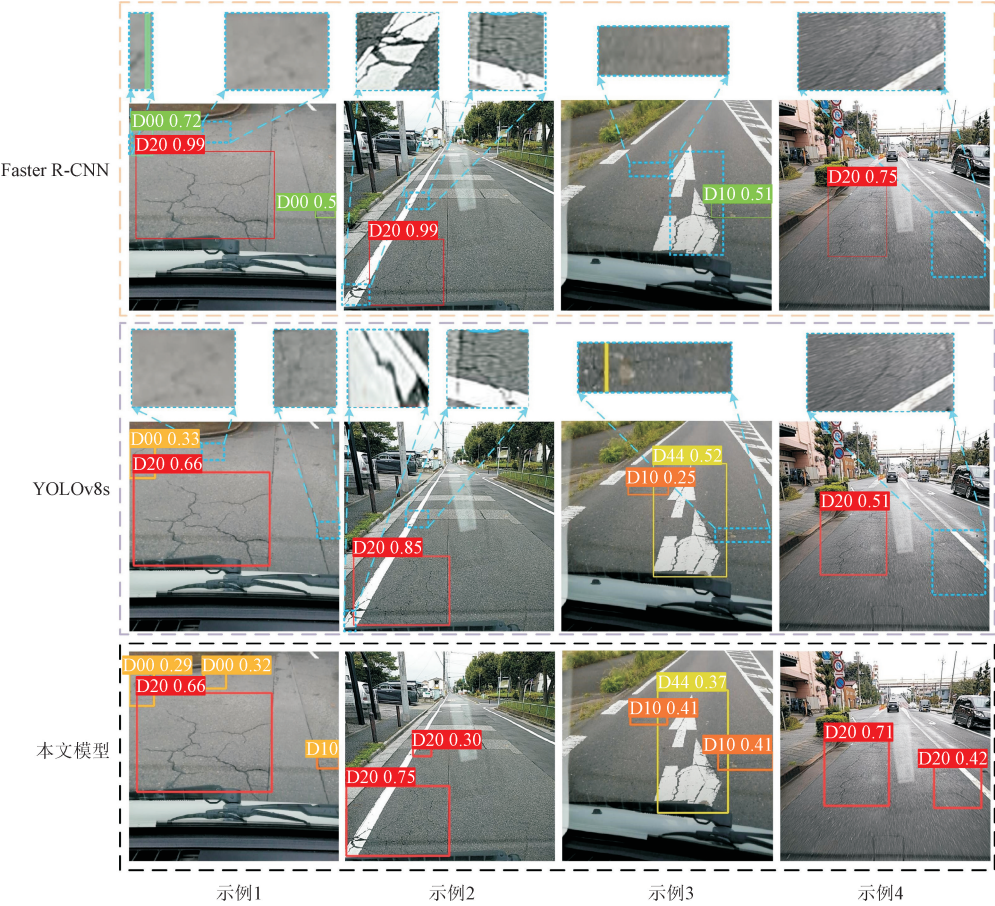
模型	模型体积/MB	m_{AP}	检测速率/(帧 · s ⁻¹)	S
YOLOv6n ^[28]	37.5	0.657	16	1.336
YOLOv7n-tiny ^[29-30]	11.7	0.653	19	1.299
Efficientdet ^[31]	25.6	0.519	6	0.083
Faster R-CNN ^[32-34]	108.0	0.520	6	0.099
YOLOX ^[35]	68.5	0.610	61	0.922
SSD ^[36-37]	192.0	0.520	10	0.108
CenterNet ^[38]	124.0	0.586	6	0.699
YOLOv5s	13.8	0.649	104	1.280
YOLOv5m	40.8	0.672	59	1.481
YOLOv5l	89.2	0.679	51	1.547
FOCS ^[39]	122.0	0.629	8	1.089
RetinaNet ^[40]	139.0	0.510	8	0.012
YOLOv8s ^[41]	22.5	0.662	98	1.397
文献[13]模型	21.7	0.582	54	0.663
文献[15]模型	120.5	0.576	8	0.608
文献[16]模型	11.0	0.648	116	1.273
本文模型	22.2	0.689	94	1.641

注：表中加粗结果为最优值。

从表 3 可知,本文模型的模型体积略大于 YOLOv5s、YOLOv7-tiny 和文献[16]中的模型,相较于其他模型仍较小,但 $mAP@0.5$ 相比于所有模型高出 1.0%~17.9%。在检测速率方面,本文模型虽然略低于 YOLOv5s 与 YOLOv8s 和文献[16]中的模型,但仍能达到 94 帧/s,高于其他模型,满足实时性的要求。从 S 来看,本文模型的 S 最高,表明本文模型在性能综合考量下兼顾了模型体积、检测精度以及检测速率的需求。

为进一步直观地评价本文模型在多尺度特征

提取方面的优势,对各模型在测试集上进行测试。图 8 展示了本文模型、对比模型中最新且精度最高的 YOLOv8s 及经典模型 Faster R-CNN 的检测效果。可以看出,Faster R-CNN 对小尺度病害或尺度较大但特征不明显的病害会出现识别不全或漏检的情况。虽然 YOLOv8s 相比 Faster R-CNN 有所改善,但同样会出现上述漏检情况。本文模型明显改善了识别不全和漏检的情况,证明了本文模型在面对多尺度病害时优越的特征提取能力和对实际多尺度路面病害检测的可靠性。



注:检测结果以“病害类别 置信度”的形式呈现。

图 8 部分对比模型的可视化结果

Fig. 8 Some of the visualization results in comparison models

3 模型的嵌入式移植与推理加速

模型训练在台式机上完成,为测试本文模型在实际工程应用中的可行性,将其经 TensorRT 优化加速后分别在 NVIDIA Jetson TX2 与 NVIDIA Jetson Nano 两个不同算力的嵌入式平台上部署测试。具体地,在模型推理阶段,通过 TensorRT 加速引擎代替 PyTorch 引擎,实现对模型的推理加速。

TensorRT 是 NVIDIA 推出的深度学习推理加速引擎,原理包括层次优化、精度混合等,内部程序通过 CUDA、C 语言编写^[42]。NVIDIA Jetson TX2 与 NVIDIA Jetson Nano 平台上的环境均一致,包括 Ubuntu18.04 操作系统、Python3.6.9、PyTorch1.10.1 深度学习框架、CUDA10.2、cuDNN7.6.5 以及各种必要的 Python 库等。本文模型在不同嵌入式平台上应用不同推理引擎的检测速率如表 4 所示。

表 4 本文模型在嵌入式平台上的检测速率

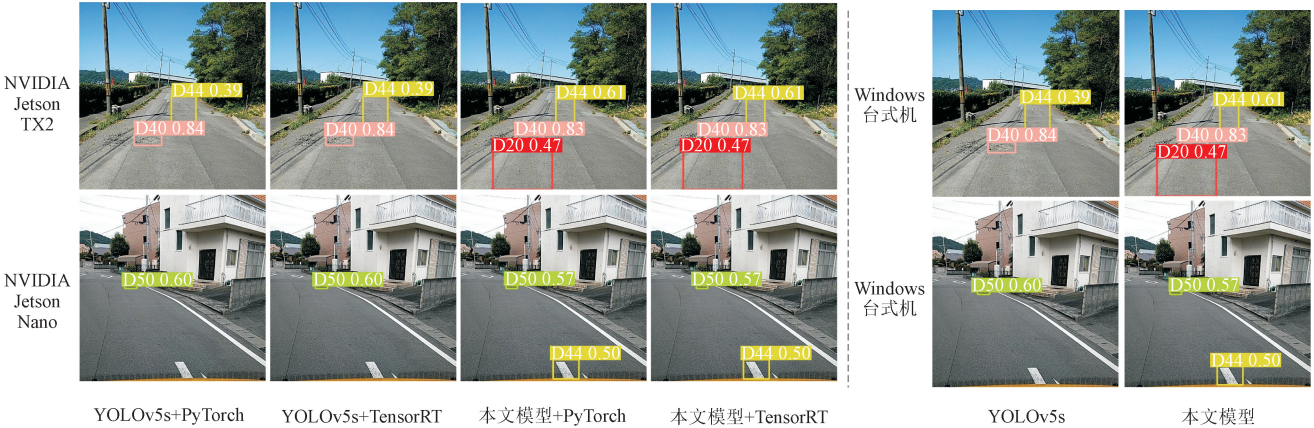
Table 4 Detection rate of the model on embedded platforms

嵌入式设备	推理引擎	检测速率/(帧·s ⁻¹)
NVIDIA Jetson TX2	PyTorch	8
	TensorRT	13
NVIDIA Jetson Nano	PyTorch	3
	TensorRT	6

注:表中加粗结果为最优值。

从表 4 可知,经过 TensorRT 优化加速后,本文

模型的检测速率得到了显著提升。在 NVIDIA Jetson TX2 上,检测速率从 8 帧·s⁻¹提高到了 13 帧·s⁻¹。在 NVIDIA Jetson Nano 上,检测速率从 3 帧·s⁻¹提高到了 6 帧·s⁻¹。检测速率在不同边缘设备上获得近 1 倍的提升。不同模型在不同嵌入式平台上的检测效果如图 9 所示。可以看出,两个边缘设备推理生成的目标检测框与在 Windows10 主机推理生成的几乎一致,基本没有精度损失,验证了本文模型在实际路面病害检测任务中的可靠性。



注:检测结果以“病害类别 置信度”的形式呈现。

图 9 不同模型在不同嵌入式平台上的检测效果

Fig. 9 Detection results on different embedded platforms of different models

4 结 论

为实现对实际路面病害准确地检测,本文基于 YOLOv5s 提出多尺度特征增强的路面病害检测模型,在数据集上训练测试并与 YOLOv8s、YOLOv7tiny、CenterNet 等模型及引言中部分较新的方法进行对比,最后经 TensorRT 框架分别部署在 NVIDIA Jetson TX2 与 NVIDIA Jetson Nano 两个不同算力的嵌入式平台上进行推理测试。本文结论如下。

(1)针对病害尺度多元化、背景复杂及原始池化阶段感受野有限等问题,提出嵌有混合空洞卷积的快速空间特征金字塔池化模块,有效地提高了模型对多尺度特征的提取能力。其次,在路径聚合网络和特征金字塔网络的基础上设计了多路径特征融合网络,获得了丰富的多尺度语义信息,提高了模型对病害的表征能力。接着,采用 K-means 聚类算法结合 IoU 聚类出适合病害尺度特点的锚点框,提高了模型的鲁棒性。接着,建立了带面积惩罚项和带梯度变化的损失函数,提高了模型的检测性能与收敛速度。通过消融实验证明了上述改进方法的有效

性。最后,通过不降维跨通道交互的高效通道注意力强化模型对重要通道间多尺度目标区域的关注度,进一步提高了模型对多尺度特征的感知能力。

(2)在 RDDJ 数据集上进行模型性能的评估实验。结果表明,本文模型的检测精度相比原模型 YOLOv5s 提高了 4.0%,并且具有较高的检测速率,有效地提升了模型在实际场景下对病害的识别效果。并且,在相同的实验条件下,本文模型相比引言中较新的方法及 YOLOv8、YOLOv6 等先进模型具有最高的检测精度,并保持一定的实时性,表明本文模型具有优越的检测性能。

(3)本文模型分别在 NVIDIA Jetson TX2 与 NVIDIA Jetson Nano 两个不同算力的嵌入式平台上部署测试,结果验证了其在实际路面病害检测中的实用性。

(4)尽管本文提出的模型取得了较好的检测效果,但在模型轻量化方面还不够先进。因此,下一步将结合结构化剪枝、知识蒸馏等模型压缩技术对本文的多尺度特征增强模型进行轻量化,使该模型在达到更高准确率的同时实现轻量化。然后,将

其部署在无人机平台上,在线实现准确的道路病害检测,为公路管理部门提供路面健康数据并及时进行修复,防止路面进一步恶化,保证路面车辆行驶安全。

参考文献:

- [1] 何铁军,李华恩. 基于改进 YOLOv5 的路面病害检测模型[J]. 土木工程学报, 2024, 57(2): 96-106.
HE Tiejun, LI Huaen. The pavement damage detection model based on improved YOLOv5 [J]. China Civil Engineering Journal, 2024, 57(2): 96-106.
- [2] 沙爱民,童峥,高杰. 基于卷积神经网络的路表病害识别与测量[J]. 中国公路学报, 2018, 31(1): 1-10.
SHA Aimin, TONG Zheng, GAO Jie. Recognition and measurement of pavement disasters based on convolutional neural networks [J]. China Journal of Highway and Transport, 2018, 31(1): 1-10.
- [3] 张跃飞,王敬飞,陈斌,等. 基于改进的 Mask R-CNN 的公路裂缝检测算法[J]. 计算机应用, 2020, 40(S2): 162-165.
ZHANG Yuefei, WANG Jingfei, CHEN Bin, et al. Pavement crack detection algorithm based on improved Mask R-CNN [J]. Journal of Computer Applications, 2020, 40(S2): 162-165.
- [4] TANG Jinshan, GU Yanliang. Automatic crack detection and segmentation using a hybrid algorithm for road distress analysis [C]//2013 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ, USA: IEEE, 2013: 3026-3030.
- [5] ZHAO Huili, QIN Guofeng, WANG Xingjian. Improvement of canny algorithm based on pavement edge detection [C]//Proceedings of the 2010 3rd International Congress on Image and Signal Processing. Piscataway, NJ, USA: IEEE, 2010: 964-967.
- [6] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [7] ZHANG Lei, YANG Fan, ZHANG Yimin, et al. Road crack detection using deep convolutional neural network [C]//2016 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2016: 3708-3712.
- [8] 张伟光,钟靖涛,于建新,等. 基于机器学习和图像处理的路面裂缝检测技术研究[J]. 中南大学学报(自然科学版), 2021, 52(7): 2402-2415.
ZHANG Weiguang, ZHONG Jingtao, YU Jianxin, et al. Research on pavement crack detection technology based on convolution neural network [J]. Journal of Central South University (Science and Technology), 2021, 52(7): 2402-2415.
- [9] 王丽苹,高瑞贞,张京军,等. 基于卷积神经网络的混凝土路面裂缝检测[J]. 计算机科学, 2019, 46(S2): 584-589.
WANG Liping, GAO Ruizhen, ZHANG Jingjun, et al. Crack detection of concrete pavement based on convolutional neural network [J]. Computer Science, 2019, 46(S2): 584-589.
- [10] 孙朝云,裴莉莉,李伟,等. 基于改进 Faster R-CNN 的路面灌封裂缝检测方法[J]. 华南理工大学学报(自然科学版), 2020, 48(2): 84-93.
SUN Zhaoyun, PEI Lili, LI Wei, et al. Pavement sealed crack detection method based on improved Faster R-CNN [J]. Journal of South China University of Technology (Natural Science Edition), 2020, 48(2): 84-93.
- [11] 许贵阳,张诗泉,白堂博. 基于改进 Faster R-CNN 的 CRTS II 型轨道板裂缝检测方法[J]. 中国铁道科学, 2023, 44(1): 106-113.
XU Guiyang, ZHANG Shiquan, BAI Tangbo. Crack detection method of CRTS II track slab based on Faster R-CNN improvement [J]. China Railway Science, 2023, 44(1): 106-113.
- [12] 罗晖,贾晨,李健. 基于改进 YOLOv4 的公路路面病害检测算法[J]. 激光与光电子学进展, 2021, 58(14): 328-336.
LUO Hui, JIA Chen, LI Jian. Road surface disease detection algorithm based on improved YOLOv4 [J]. Laser & Optoelectronics Progress, 2021, 58(14): 328-336.
- [13] ROY A M, BHADURI J. DenseSPH-YOLOv5: an automated damage detection model based on DenseNet and swin-transformer prediction head-enabled YOLOv5 with attention mechanism [J]. Advanced Engineering Informatics, 2023, 56: 102007.
- [14] LI Mingjue, WANG Hao, PENG Yuhang, et al. Edge-to-client real-time road damage detection system based on Yolov5 [C]//2021 China Automation Congress (CAC). Piscataway, NJ, USA: IEEE, 2021: 1221-1226.
- [15] 周中,闫龙宾,张俊杰,等. 基于 YOLOX-G 算法的隧道裂缝实时检测[J]. 铁道科学与工程学报, 2023, 20(7): 2751-2762.
ZHOU Zhong, YAN Longbin, ZHANG Junjie, et al. Real-time detection of tunnel cracks based on YOLOX-G algorithm [J]. Journal of Railway Science and Engineering, 2023, 20(7): 2751-2762.
- [16] XIANG Wann, WANG Haichen, XU Yuan, et al. Road

- disease detection algorithm based on YOLOv5s-DSG [J]. *Journal of Real-Time Image Processing*, 2023, 20(3): 56.
- [17] ALFARRARJEH A, TRIVEDI D, KIM S H, et al. A deep learning approach for road damage detection from smartphone images [C]//2018 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2018: 5201-5204.
- [18] MANDAL V, MUSSAH A R, ADU-GYAMFI Y. Deep learning frameworks for pavement distress classification: a comparative analysis [C]//2020 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2020: 5577-5583.
- [19] 张凡, 张鹏超, 王磊, 等. 基于 YOLOv5s 的轻量化朱鹮检测算法研究 [J]. *西安交通大学学报*, 2023, 57(1): 110-121.
- ZHANG Fan, ZHANG Pengchao, WANG Lei, et al. Research on lightweight crested ibis detection algorithm based on YOLOv5s [J]. *Journal of Xi'an Jiaotong University*, 2023, 57(1): 110-121.
- [20] 王保宪, 白少雄, 赵维刚. 基于特征增强学习的路面裂缝病害视觉检测方法 [J]. *铁道科学与工程学报*, 2022, 19(7): 1927-1935.
- WANG Baoxian, BAI Shaoxiong, ZHAO Weigang. Pavement crack damage visual detection method based on feature reinforcement learning [J]. *Journal of Railway Science and Engineering*, 2022, 19(7): 1927-1935.
- [21] 彭雨诺, 刘敏, 万智, 等. 基于改进 YOLO 的双网络桥梁表观病害快速检测算法 [J]. *自动化学报*, 2022, 48(4): 1018-1032.
- PENG Yunuo, LIU Min, WAN Zhi, et al. A dual deep net work based on the improved YOLO for fast bridge surface defect detection [J]. *Acta Automatica Sinica*, 2022, 48(4): 1018-1032.
- [22] 邹俊志, 杨建喜, 李昊, 等. 复杂背景下基于改进 YOLO v3 算法的桥梁表观病害识别 [J]. *铁道科学与工程学报*, 2021, 18(12): 3257-3266.
- ZOU Junzhi, YANG Jianxi, LI Hao, et al. Bridge apparent damage detection based on the improved YOLO v3 in complex background [J]. *Journal of Railway Science and Engineering*, 2021, 18(12): 3257-3266.
- [23] ZHENG Zhaohui, WANG Ping, REN Dongwei, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. *IEEE Transactions on Cybernetics*, 2022, 52(8): 8574-8586.
- [24] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 11531-11539.
- [25] ARYA D, MAEDA H, GHOSH S K, et al. RDD2020: an annotated image dataset for automatic road damage detection using deep learning [J]. *Data in Brief*, 2021, 36: 107133.
- [26] RIBANI R, MARENGONI M. A survey of transfer learning for convolutional neural networks [C]//2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T). Piscataway, NJ, USA: IEEE, 2019: 47-57.
- [27] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The PASCAL visual object classes challenge 2007 [EB/OL]. (2007-12-20) [2024-07-25]. <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/>.
- [28] NGUYEN K H, NGUYEN H V N, TRAN H N, et al. Combining autoencoder and Yolov6 model for classification and disease detection in chickens [C]//Proceedings of the 2023 8th International Conference on Intelligent Information Technology. New York, NY, USA: Association for Computing Machinery, 2023: 132-138.
- [29] OKRAN A M, ABDEL-NASSER M, RASHWAN H A, et al. Effective deep learning-based ensemble model for road crack detection [C]//2022 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2022: 6407-6415.
- [30] 倪昌双, 李林, 罗文婷, 等. 改进 YOLOv7 的沥青路面病害检测 [J]. *计算机工程与应用*, 2023, 59(13): 305-316.
- NI Changshuang, LI Lin, LUO Wenting, et al. Disease detection of asphalt pavement based on improved YOLOv7 [J]. *Computer Engineering and Applications*, 2023, 59(13): 305-316.
- [31] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 10778-10787.
- [32] KLUGER F, REINDERS C, RAETZ K, et al. Region-based cycle-consistent data augmentation for object detection [C]//2018 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2018: 5205-5211.
- [33] WANG Wenzhe, WU Bin, YANG Sixiong, et al. Road damage detection and classification with faster R-CNN [C]//2018 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2018: 5220-5223.

- [34] 晏班夫, 徐观亚, 栾健, 等. 基于 Faster R-CNN 与形态法的路面病害识别 [J]. 中国公路学报, 2021, 34(9): 181-193.
- YAN Banfu, XU Guanya, LUAN Jian, et al. Pavement distress detection based on Faster R-CNN and morphological operations [J]. China Journal of Highway and Transport, 2021, 34(9): 181-193.
- [35] GE Zheng, LIU Songtao, WANG Feng, et al. YOLOX: exceeding YOLO series in 2021 [EB/OL]. (2021-08-06) [2024-06-01]. <https://arxiv.org/abs/2107.08430>.
- [36] MAEDA H, SEKIMOTO Y, SETO T, et al. Road damage detection and classification using deep neural networks with smartphone images [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(12): 1127-1141.
- [37] WANG Yanbo, DING Ming, KAN Shichao, et al. Deep proposal and detection networks for road damage detection and classification [C]//2018 IEEE International Conference on Big Data (Big Data). Piscataway, NJ, USA: IEEE, 2018: 5224-5227.
- [38] DUAN Kaiwen, BAI Song, XIE Lingxi, et al. CenterNet: keypoint triplets for object detection [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 6568-6577.
- [39] 郑美俊, 田益民, 杨帅. 基于改进 FCOS 网络的遥感目标检测 [J]. 航天返回与遥感, 2022, 43(5): 133-141.
- ZHENG Meijun, TIAN Yimin, YANG Shuai. Remote sensing target detection based on improved FCOS network [J]. Spacecraft Recovery & Remote Sensing, 2022, 43(5): 133-141.
- [40] ANGULO A, VEGA-FERNÁNDEZ J A, AGUILAR-LOBO L M, et al. Road damage detection acquisition system based on deep neural networks for physical asset management [C]//Advances in Soft Computing. Cham: Springer International Publishing, 2019: 3-14.
- [41] LI Xiang, LI Li, LIU Zhigui, et al. Dam crack detection studies by UAV based on YOLO algorithm [C]//2023 2nd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC). Piscataway, NJ, USA: IEEE, 2023: 104-108.
- [42] 顾德英, 罗聿伦, 李文超. 基于改进 YOLOv5 算法的复杂场景交通目标检测 [J]. 东北大学学报(自然科学版), 2022, 43(8): 1073-1079.
- GU Deying, LUO Yulun, LI Wenchao. Traffic target detection in complex scenes based on improved YOLOv5 algorithm [J]. Journal of Northeastern University(Natural Science), 2022, 43(8): 1073-1079.

(编辑 陶晴)