

基于改进 YOLOv8 的路面坑洼检测

朱成杰,蔡子正,朱洪波

安徽理工大学 电气与信息工程学院,安徽 淮南 232001

摘要:目的 针对现有路面坑洼检测模型参数量大,检测实时性较差,路面环境复杂而导致检测准确度不高的问题,提出一种基于改进 YOLOv8 的路面坑洼检测算法。方法 首先,将颈部网络替换为双向金字塔网络 Bifpn(Bi-directional Feature Pyramid Network),进行不同尺度的特征融合;其次,为提高模型推理速度,使用多分支结构 DBB(Diverse Branch Block)替换 C2f 模块中的卷积层,提出了重参数化模块 C2f-DBB 并替换部分 C2f 模块;然后,在骨干网络末端引入通道先验注意力机制 CPCA(Channel Prior Convolutional Attention),在有效提取空间关系的同时保留通道先验,强化特征提取能力;最后,使用边界框回归损失函数 MPDIoU(MPD Intersection over Union)代替原模型的 CIoU(Complete Intersection over Union),进一步提高算法精度。结果 实验结果表明:改进后的算法在平均精度上较原始网络提高了 3.1%,而模型参数量仅为 5.9 M,计算量为 7.3 G,分别下降了 6.3%和 9.8%,每秒帧数可达 53.4。结论 通过实验结果可以看出:改进后的模型与当前主流模型相比在精度上提升显著,且满足实时性要求,对于实际应用具备一定的参考价值。

关键词:坑洼检测;YOLOv8;重参数化;注意力机制

中图分类号:TP391.41 **文献标识码:**A **doi:**10.16055/j.issn.1672-058X.2026.0002.012

Road Surface Pothole Detection Based on Improved YOLOv8

ZHU Chengjie, CAI Zizheng, ZHU Hongbo

School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, Anhui, China

Abstract: Objective To address the issues of large model parameters, poor real-time detection performance, and low detection accuracy due to complex road environments in existing road surface pothole detection models, an improved YOLOv8-based road surface pothole detection algorithm is proposed. **Methods** First, the neck network was replaced with a bi-directional feature pyramid network (BiFPN) to achieve multi-scale feature fusion. Second, to enhance the inference speed of the model, a diverse branch block (DBB) was used to replace the convolutional layers in the C2f module, and a reparameterization module, C2f-DBB, was introduced to substitute part of the C2f modules. Then, a channel prior convolutional attention (CPCA) mechanism was introduced at the end of the backbone network to retain channel priors while effectively extracting spatial relationships, thereby strengthening feature extraction capabilities. Finally, the boundary box regression loss function MPD intersection over union (MPDIoU) was used instead of the original model's complete intersection over union (CIoU), further improving the accuracy of the algorithm. **Results** Experimental results showed that the improved algorithm increased the average precision by 3.1% compared to the original network, with the model parameters being only 5.9 M and the computation being 7.3 G, representing reductions of 6.3% and 9.8%, respectively. The frame rate reached 53.4 frames per second. **Conclusion** The experimental results indicate that the

收稿日期:2024-05-09 **修回日期:**2024-07-16 **文章编号:**1672-058X(2026)02-0093-08

基金项目:国家自然科学基金项目资助(62003001)。

作者简介:朱成杰(1979—),男,博士,副教授,从事图像处理与模式识别研究。

通信作者:蔡子正(1999—),男,硕士研究生,从事目标检测研究。Email:1486544178@qq.com。

引用格式:朱成杰,蔡子正,朱洪波.基于改进 YOLOv8 的路面坑洼检测[J].重庆工商大学学报(自然科学版),2026,43(2):93-100.

ZHU Chengjie, CAI Zizheng, ZHU Hongbo. Road surface pothole detection based on improved YOLOv8[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(2): 93-100.

improved model significantly enhances accuracy compared to current mainstream models while meeting real-time requirements, providing valuable reference for practical applications.

Keywords: pothole detection; YOLOv8; reparameterization; attention mechanism

公路网络作为我国重要的基础设施,是经济发展的基石,交通运输系统不仅满足了人们的出行需求,而且促进了工商业的蓬勃发展^[1]。但公路路面在实际运营过程中,随着时间的推移产生坑洼等各种路面病害^[2]。如果这些坑洼得不到及时发现和处理,不仅会降低公路的使用年限,还会损坏高速行驶汽车的轮胎,导致交通事故的发生。因此,将目标检测技术应用于路面坑检测具有重要意义^[3]。

传统的路面坑洼检测方法大部分是通过人工提取目标特征信息,进而实现对路面坑洼的图像分割,例如,高建贞^[4]通过图像处理方法实现了路面坑洼检测,先对路面坑洼图像进行图像增强处理,再进行阈值分割、提取出线性特征,最后识别出坑洼目标。韩锬等^[5]提出了一种基于像素级特征的路面裂缝检测算法,然而以上算法都需要基于人工进行坑洼特征信息的提取,只能在特定场景下进行目标检测,在背景复杂的路面环境下检测精度较低、且鲁棒性较差。

近年来,随着计算机技术的发展和硬件的提升,基于深度学习的检测算法已成为路面坑洼检测的主流方法,可分为两类,分别为单阶段算法和双阶段算法^[6]。双阶段算法先生成候选框以确定坑洼区域,再对候选框进行检测,所以检测速度较慢,不适用于实时检测^[7]。韦正璐等^[8]利用改进的 Inception 模块替代 SSD 网络中的部分卷积层,有效提高了检测速度,但是当坑洼目标较小时其算法精度不高。单阶段算法可以直接对所有空间区域进行检测。苏佳等^[9]在 YOLOv5 的基础上提出一种自适应提取模块 AFE(Adaptive Feature Extraction),通过动态调整权重分配突出关键目标的特征表示,引入输出重构模块 ORS(Output Reconstruction),通过重构不同大小的目标检测层以实现小目标的精确定位,但在低光照条件下漏检率较高。王志等^[10]针对检测目标尺度差异较大等问题,将渐进金字塔网络 AFPN(Asymptotic Feature Pyramid Network)引入 YOLOv7 中,解决了多尺度融合的特征丢失问题。此外,采用高效聚合网络模块替换 AFPN 中的基础模块,进一步提高了网络的特征提取能力。但经过改进的网络计算量较大,不适用于实时性检测。罗磊等^[11]将 YOLOv8 的部分卷积层替换为幻影卷积 Ghostconv,并且引入全局注意力模块,有效捕捉输入特征图的全局上下文信息,使检测精度大幅提高,但改进后的模型体积比原模型增加了 8%。

上述研究虽然都对路面坑洼检测算法进行了改进,并取得了一些效果,但仍存在着不足。部分算法虽然提升了检测速度,但忽略了检测精度的提升。部分

算法虽然在常规环境下检测效果较好,但在背景复杂或检测目标大小不一时表现较差。为解决上述问题,本文提出了一种基于改进 YOLOv8 的路面坑洼检测方法。首先,针对检测目标尺度不同的问题,将颈部网络替换为双向金字塔网,在多个尺度进行特征融合。其次,基于重参数化的思想提出了 C2f-DBB 模块,使模型在预测阶段将多个分支结构转化为一个 3×3 卷积,可以有效提高检测速度。最后,通过引入注意力机制 CPCA 和将原损失函数替换为 MPDIoU,大幅提高了模型的检测精度。

1 YOLOv8 模型设计与算法改进

1.1 YOLOv8 模型介绍

YOLO 系列算法能够在保持高精度的同时进行实时目标检测,因此在目标检测领域应用广泛。Redmon 等首先受 CoogLeNet 网络结构启发于 2016 年提出了 YOLOv1,后续学者又提出了 YOLOv2、YOLOv3、YOLOv4、YOLOv5、YOLOv7、YOLOX 等一系列算法。YOLOv8 是 Ultralytics 公司于 2023 年推出的最新一代 YOLO 模型,支持多种视觉任务。YOLOv8 具有多个版本,YOLOv8n 是 YOLOv8 系列中最轻量级的版本,模型参数和计算量更少、文件更小、便于存储和传输。此外,YOLOv8n 对硬件资源的需求最低,可以在资源受限的设备(如嵌入式设备、移动设备)上高效运行,这使得其在边缘计算和移动应用中具有明显优势。所以,本文选用 YOLOv8n 网络模型进行改进。YOLOv8n 网络模型结构如图 1 所示。

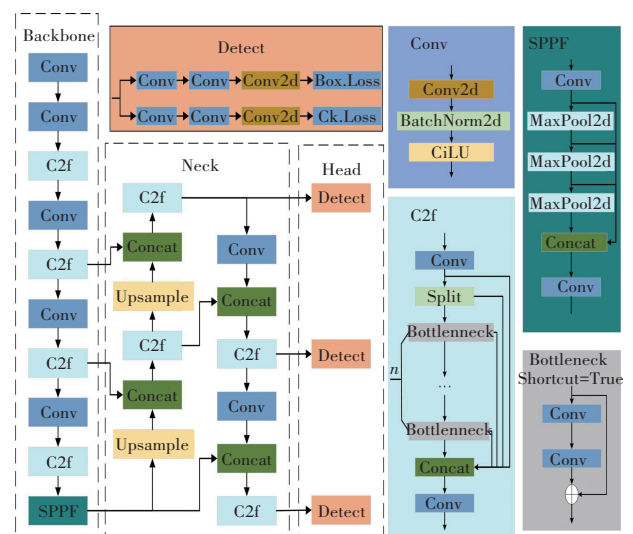


图 1 YOLOv8n 网络结构

Fig. 1 YOLOv8n network structure

YOLOv8n 的输入端采用了数据增强的方式,加强对小目标物体的检测效果;此外,还采用自适应锚框计算方法,可以将输入的图像调整为统一尺寸,加快了检测效率。

主干网络部分使用全新的 C2f 结构,可以将目标信息与上下文结合起来,以此完成对残差信息的学习。SPPF 模块又称空间金字塔池化,可以将不同大小的图像转化为统一大小,并将输出的结果进行 Concat 拼接。

颈部网络主要由两部分组成,分别为特征金字塔网络 FPN^[12]和路径聚合网络 PAN^[13]。FPN 自顶向下利用上采样实现不同层次特征的融合,使底层的特征图带有更丰富的语义信息;PAN 自底向上进行下采样,使不同层次的特征信息充分融合,保证了网络对不同尺寸目标的检测性能。

输出端凭借不同大小的输入信息来获得相应大小目标的特征信息,克服了顶端目标的局限性。

1.2 改进的 YOLOv8 算法

由于传统的 YOLOv8n 算法在对路面坑洼目标检测时存在漏检率高、参数量大等问题。因此,本文从以下 4 个方面对原模型进行改进。首先,将颈部网络替换为双向金字塔网络 BiFPN,增强了特征信息融合效率;为提高模型推理速度,减少模型参数量,在主干和颈部网络引入重参数化模块 C2f-DBB;为了强化模型特征的提取能力,在 SPPF 层后面引入 CPCA 注意力机制;最后,为了提高算法的边界框回归精度,使用 MPDIoU 代替 CIoU。本文所提算法模型的网络结构如图 2 所示。

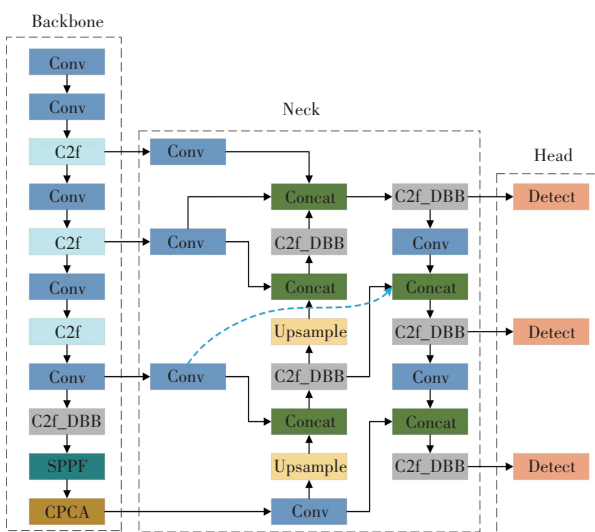


图 2 改进的 YOLOv8n 网络结构

Fig. 2 Improved YOLOv8n network structure

1.2.1 双向特征金字塔网络

Google Brain 团队的 TAN 等^[14]在 PANet 的基础上提

出了双向特征金字塔网络 BiFPN,公式如下:

$$O = \sum_i \frac{\omega_i}{\varepsilon + \sum_j \omega_j} \cdot I_i \quad (1)$$

式(1)中, i, j 表示节点输入层的个数, ω 表示权值, I 表示节点的输入特征图。

BiFPN 通过更高效的双向跨尺度连接和特征图融合增强了原算法。如果一个节点只有一个输入边,而且不存在特征融合的节点,那么此节点对网络的贡献相对较小,所以需要进行网络学习以增加其贡献度。BiFPN 引入了训练权重,根据不同特征层对特征融合的贡献度学习分配相应的权重,实现特征层的深度融合。

该算法还在同一层中的原始输入和输出节点之间添加了额外的特征分支,以实现集成更多特征。通过跨尺度加权连接和快速归一化方法对特征进行融合, BiFPN 通过优化特征融合路径和引入自适应加权融合机制,实现了高效的多尺度特征融合,能够保留更多的细节信息,从而提高小目标的检测准确性。反向特征传播路径的引入,使得高层次的语义信息能够传播回低层次特征图,提升了低层次特征对小目标的识别能力。BiFPN 结构如图 3 所示。

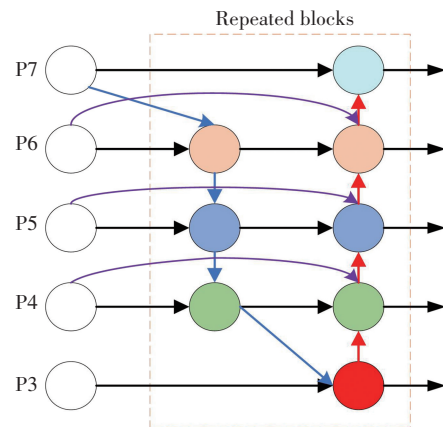


图 3 BiFPN 结构图

Fig. 3 BiFPN structure

1.2.2 基于重参数化的 C2f-DBB 模块

(1) 重参数化结构 DBB。DBB (Diverse Branch Block) 是 Xiaohan Ding 等^[15]于 2021 年基于结构重参数化思想推出的模块,是 ACNet^[16]的改进。DBB 是一种多分支结构,共有 4 条分支、5 种转换方法,分别为 BN 融合、串联卷积融合、平均池化层转换为卷积层、多尺度卷积转化为同一尺度与卷积加法的分支融合。通过组合这些 1×1 卷积、 $K \times K$ 卷积、平均池化层等可以构建多分支结构,提高网络的表现能力。虽然多分支结构在训练过程中会占用较多内存,且需要更长的训练

时间,但在预测时可以将多分支结构进行转换合并操作,使其融合为一个 3×3 卷积,加快推理速度,如图 4 所示。

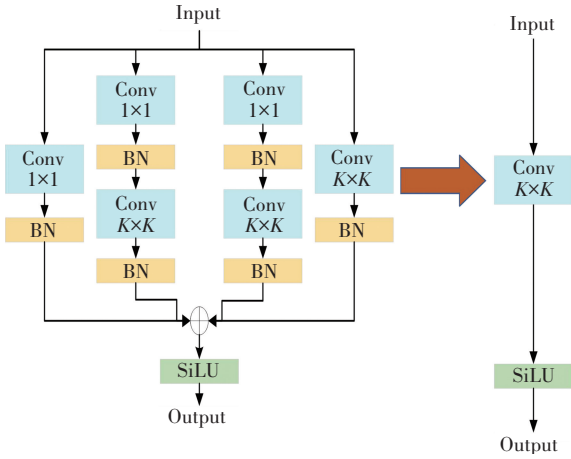


图 4 DBB 模块
Fig. 4 DBB module

(2) C2f-DBB 模块。本文融合重参数化思想与 C2f 模块,在此基础上提出了 C2f-DBB 模块。首先使用两个串行的 DBB 模块替换 Bottleneck 中的 3×3 卷积,组成 Bottleneck-DBB 模块,通过组合 1×1 卷积、 $K \times K$ 卷积、平均池化层等构建多分支结构,如图 5 所示。在训练阶段使用两个串行的 DBB 结构融合特征,由于 DBB 各分支感受野不同,可以获得不同位置的特征,拼接后的特征更丰富,最后融合各阶段提取的特征,而在推理时将多分支结构等价变换为 3×3 卷积。

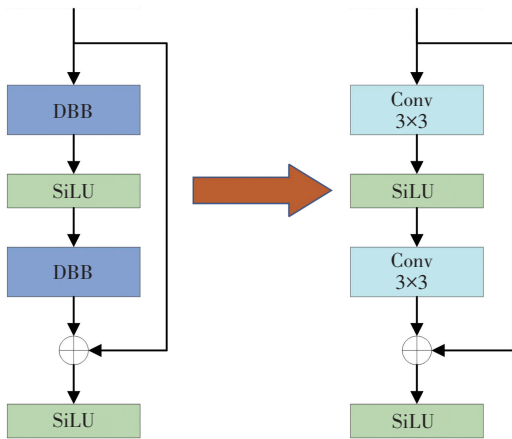


图 5 Bottleneck-DBB 模块
Fig. 5 Bottleneck-DBB module

将原 YOLOv8n 模型中 C2f 模块中的 Bottleneck 部分全部替换为 Bottleneck-DBB 模块,即可构成本文所提出的新模块 C2f-DBB。输入时首先经过一个 1×1 卷积处理,然后通过 Split 将输入分为两部分 part1 和 part2,将 part1 输入 Bottleneck-DBB 模块中,得到的结

果输入 Concat,重复 n 次,然后将 part2 也输入 Concat 进行特征融合,最后再通过一个 1×1 卷积得到最终输出,如图 6 所示。

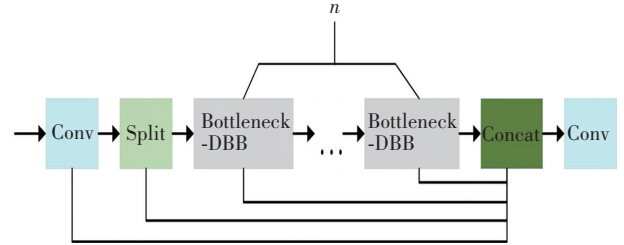


图 6 C2f-DBB 模块
Fig. 6 C2f-DBB module

1.2.3 CPCA 注意力机制

为了进一步增强检测模型对不同大小坑洼目标的检测能力,在 YOLOv8 骨干网络 Backbone 中加入通道优先卷积注意力模块(Channel Prior Convolutional Attention, CPCA)。CPCA 通过为每个通道分配不同的注意力权重,增强了小目标相关特征表示,从而提高了检测准确性。通道先验机制的引入,使得注意力分配更加合理,网络能够更好地选择与小目标相关的特征通道,提升了小目标的检测效果。此外,CPCA 在不同层级特征图中的应用,增强了小目标特征的细粒度性,进一步提高了检测精度。

通道优先卷积注意力机制结构如图 7 所示,和 CBAM 类似,采用先通道后空间的形式。首先给定一个中间特征图 $F \in \mathbf{R}^{C \times H \times W}$,通道注意力模块推理出一个一维通道注意力图 $M_c \in \mathbf{R}^{C \times 1 \times 1}$,将 M_c 逐元素与输入特征 F 相乘,然后沿空间维度广播通道注意力值,以获得具有通道注意力的精细特征 $F_c \in \mathbf{R}^{C \times H \times W}$ 。空间注意力模块对 F_c 进行处理,生成一个三维空间注意力图 $M_s \in \mathbf{R}^{C \times H \times W}$ 。最后的输出特征 $\hat{F} \in \mathbf{R}^{C \times H \times W}$ 是由 M_s 与 F_c 逐元素相乘得到的。总的注意力过程可以总结为

$$F_c = C_A(F) \otimes F \tag{2}$$

$$\hat{F} = S_A(F_c) \otimes F_c \tag{3}$$

其中,通道注意力沿袭了 CBAM 的方法,通过平均池化和最大池化,将空间信息传输到共享的 MLP 中。空间注意力利用可分离卷积获取输入目标之间的空间关系。最后再使用一个 1×1 的卷积进行通道混合。

$$C_A(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \tag{4}$$

$$S_A(F) = \text{Conv}_{1 \times 1}(\sum_{i=0}^3 \text{Branch}_i(\text{DwConv}(F))) \tag{5}$$

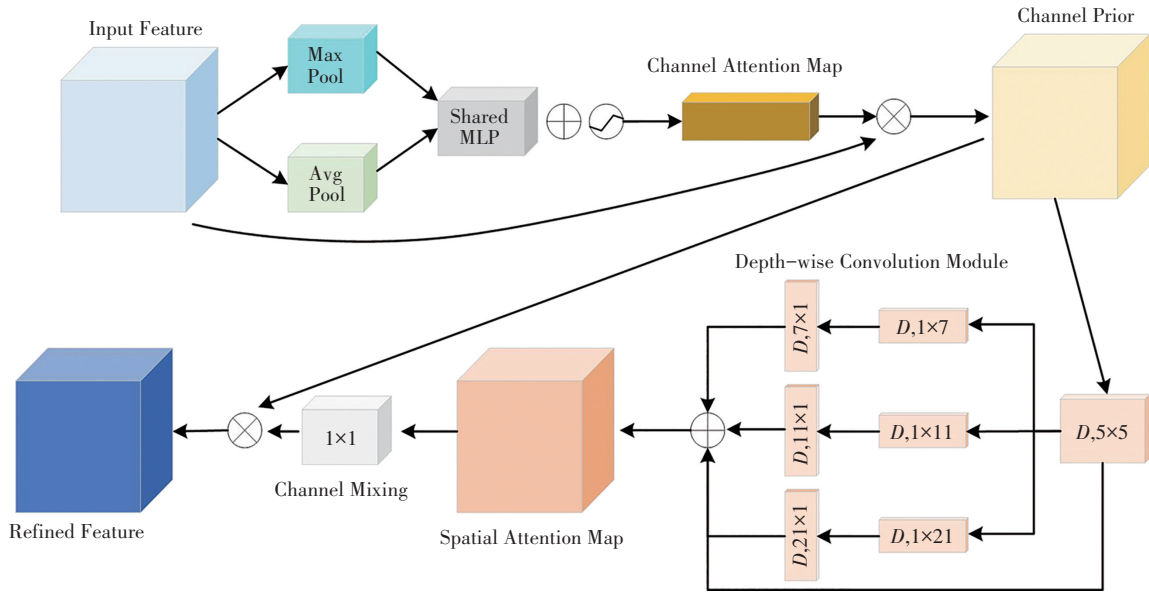


图 7 CPCA 注意力机制

Fig. 7 CPCA attention mechanism

1.2.4 MPDIoU

原始 YOLOv8 模型使用 CIoU 作为回归损失函数。然而,CIoU 损失函数在计算时只考虑了目标框在预测框的重合度,而忽略了目标框的方向信息,因此本文引入 MPDIoU 损失函数来替代原始损失函数。

根据矩形的数学定义,在矩形竖边垂直于水平面的情况下,仅依据左上点和右下点坐标可以确定一个唯一的矩形,基于这一推论,本文设计了一个新的 IoU 的度量 MPDIoU^[17],以最小化预测框和标注框左上和右下点之间的距离。定义两个任意矩形 $A, B \in \mathbf{R}^n$, 输入图像的宽度和高度为 w 和 h 。对于 A 和 B , (x_1^A, y_1^A) 、 (x_2^A, y_2^A) 表示 A 的左上和右下点坐标, (x_1^B, y_1^B) 、 (x_2^B, y_2^B) 表示 B 的左上和右下点坐标,MPDIoU 的计算公式如下:

$$d_1^2 = (x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2 \quad (6)$$

$$d_2^2 = (x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2 \quad (7)$$

$$f_{\text{MPDIoU}} = f_{\text{IoU}} \frac{d_1^2}{w^2 + h^2} \frac{d_2^2}{w^2 + h^2} \quad (8)$$

MPDIoU 的损失函数定义如下:

$$L_{\text{MPDIoU}} = 1 - f_{\text{MPDIoU}} \quad (9)$$

本文提出的 MPDIoU 简化了边界框回归过程中预测框与标注框之间的相似性比较,在预测框与标注框重叠或不重叠的情况下都能有效收敛。

2 实验与结果分析

2.1 实验环境及数据集

本文实验在云计算平台 AutoDL 上进行,深度学习框架为 PyTorch,使用 CUDA11.8 进行加速。CPU 使用的是

Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50 GHz,GPU 选用 NVIDIA 的 RTX 2080Ti,11 GB 显存。

本实验使用 Kaggle 网站上的公开路面坑洼数据集,提供了多种场景下的路面坑洼图片。数据集共有 2 083 张图片,训练集和测试集的比例划分为 9 : 1。

在训练过程中,首先结合数据集图像特征将输入图像尺寸设置为 640×640,然后在平衡显存利用率和训练稳定性后,选择将批量大小设置为 32,采用 SGD 优化器对模型进行优化,初始学习速率的选择对模型训练影响较大,对 3 种不同学习率的对比如图 8 所示,最终选择将初始学习速率设置为 0.01;为防止训练周期过长导致过拟合现象,将训练周期设置为 200。

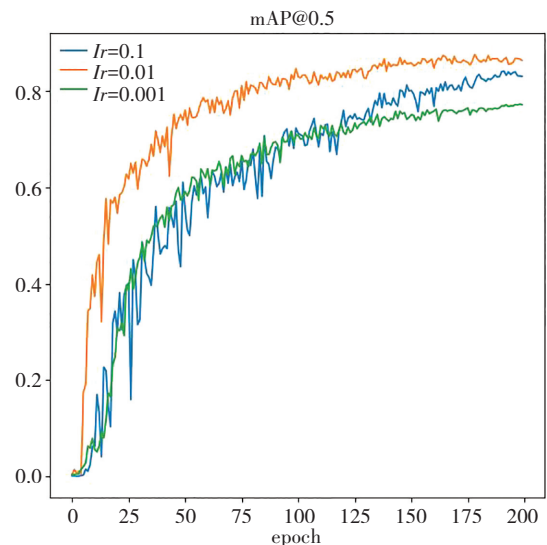


图 8 不同学习率对比曲线图

Fig. 8 Comparison curve of different learning rates

2.2 评价指标

本实验使用准确率($\xi_{\text{Precision}}$)、召回率(ξ_{Recall})、均值平均精度(ξ_{mAP})等指标来判别网络模型的性能,公式如下:

$$\xi_{\text{Precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (10)$$

$$\xi_{\text{Recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (11)$$

$$\xi_{\text{mAP}} = \frac{\sum P_A}{N_C} \quad (12)$$

其次,使用参数量来评价网络模型的权重大小,使用 Giga 浮点运算每秒(GFLOPs)来评价算法的复杂性。

2.3 实验分析

2.3.1 注意力机制对比实验

为了验证本文网络模型中引入 CPCA 注意力机制的有效性,选择 CBAM^[18]、GAM、SimAM^[19] 注意力模块与之进行对比实验,以显示不同的注意力机制给模型带来的影响。CBAM 将通道注意机制和空间注意机制融合在一起,但在各个通道上的空间注意力权重分布是一致的,这种一致的权重分布引入了噪声,并严重影响网络的检测性能。GAM 能够在减少信息弥散的情况下能放大全局维交互特征。SimAM 通过优化能量函数以挖掘神经元的重要性,无需额外参数为特征图推导出 3D 注意力权值。在 SPPF 层后加入各注意力机制后的结果如表 1 所示。

表 1 注意力机制对比实验

Table 1 Comparative experiments of attention mechanisms

Models	Precision	Recall	mAP@0.5	Params	FLOPs
		/%	/%	/M	/G
YOLOv8	78.9	78.8	84.8	2.87	8.1
YOLOv8+CBAM	85.5	74.5	85.4	2.93	8.1
YOLOv8+GAM	83.3	79.3	86.5	4.43	9.4
YOLOv8+SimAM	86.1	76.6	85.3	2.87	8.1
YOLOv8+CPCA	93.8	73.4	86.5	2.99	8.3

实验表明:引入 CPCA 注意力机制相比于 CBAM、SimAM 注意力模块,mAP@0.5 分别提高了 1.1%、1.2%,参数量分别只升高 0.06 M、0.12 M,计算量分别只升高 0.2 G、0.2 G。相比于 GAM 注意力模块,mAP@0.5 两者相同,但是参数量和计算量分别降低 1.44 M 和 1.1 G。综合以上分析,CPCA 注意力机制在增加极小

的参数量和计算量情况下,可以大幅提高检测精度。

2.3.2 消融实验

为了验证不同的改进方法对 YOLOv8 模型检测性能的影响,本文设计了一系列消融实验,如表 2 所示。其中“√”表明采用了相应的改进方法。为了公平比较,环境及参数设置均保持统一。Exp1 为原始 YOLOv8 实验结果,作为后 6 组实验的对比基准;Exp2 是将颈部网络替换为 Bifpn 结构;Exp3 是将原始网络的 C2f 模块全部替换为 C2f-DBB;Exp4 是在 SPPF 层后添加 CPCA 注意力机制;Exp5 是将 Bifpn 结构中的 C2f 模块替换为 C2f-DBB;Exp6 是在 Exp5 的基础上,添加 CPCA 注意力机制;Exp7 为本文提出的算法。

表 2 消融实验

Table 2 Ablation experiments

Index	Bifpn	C2f-DBB	CPCA	MPDIoU	mAP@0.5	Params	FLOPs
					/%	/M	/G
Exp1					84.8	2.87	8.1
Exp2	√				85.4	1.90	7.1
Exp3		√			85.7	2.87	8.1
Exp4			√		86.5	2.99	8.3
Exp5	√	√			85.9	1.90	7.1
Exp6	√	√	√		87.0	2.02	7.3
Exp7	√	√	√	√	87.9	2.02	7.3

结果表明:引入 Bifpn 结构后,使用的多尺度融合方法可以显著降低目标信息的损失,检测精度提升了 0.6%,参数量和计算量分别下降 0.97 M 和 1.0 G。将 C2f-DBB 替换 C2f 模块后,在训练阶段使用多分支结构提高网络表现能力,在推理阶段将多分支结构等价变换为一个 3×3 卷积,在保持参数量和计算量不变的情况下,检测精度提升了 0.9%。在添加 CPCA 注意力机制后,采用多尺度的深度可分离卷积模块构成空间注意力,可以在通道和空间维度上动态分配注意权重,使网络的检测精度提升 1.7%,而参数量和计算量仅增加 0.12 M 和 0.2 G。将上述改进和 MPDIoU 融合,最终使得网络模型的检测精度提升了 3.1%,参数量仅为 2.02 M,计算量仅为 7.3 G,分别降低了 29.6% 和 9.9%,有效平衡了精度与轻量化,为部署在移动设备上提供了可能。

2.3.3 不同模型对比实验

为了进一步验证本文算法的优越性,将本文算法

与 Faster R-CNN、YOLOv3、YOLOv3-tiny、YOLOv4、YOLOv5、YOLOv7、YOLOv7-tiny、YOLOv8、BFDS-YOLO^[20]采用相同的训练方式进行对比,结果如表 3 所示。

表 3 不同算法对比实验

Table 3 Comparison experiments of different algorithms

Models	Precision	Recall	mAP@0.5	Params	FLOPs
	/%	/%	/%	/M	/G
Faster R-CNN	54.7	78.6	75.8	28.31	108.2
YOLOv3	85.5	79.6	85.4	58.65	154.5
YOLOv3-tiny	69.9	69.1	71.7	8.27	12.9
YOLOv5	82.1	77.7	82.4	6.69	15.8
YOLOv5n	75.0	73.7	77.1	1.68	4.1
YOLOv7	79.2	78.0	83.2	34.79	103.2
YOLOv7-tiny	78.5	75.8	81.8	5.73	13
YOLOv8	78.9	78.8	84.8	2.87	8.1
BFDS-YOLO	89.1	72.6	85.4	4.80	6.9
Ours	89.1	73.9	87.9	2.02	7.3

从表 3 可以看出,单阶段检测模型 Faster R-CNN 精度较低。YOLOv3、YOLOv5s、YOLOv7 等主流目标检测模型虽然精确度较高,mAP@0.5 均超过了 80%,但是参数量和计算量都较大,不利于模型的轻量化部署。YOLOv3-tiny、YOLOv5n、YOLOv7-tiny 作为各自模型的轻量化版本,虽然参数量和计算量大幅降低,但是检测精度较低,mAP@0.5 最高的 YOLOv7-tiny 只能达到 81.8%。本文改进算法与原算法 YOLOv8 相比,Precision 和 mAP@0.5 分别提升了 10.2% 和 3.1%,参数量和计算量分别降低了 2.67 M 和 0.8 G。此外,BFDS-YOLO 虽然大大降低了模型计算量,但模型的准确度却提升较小,上述改进方法均没有在准确度和模型大小上取得较好的平衡。

为了进一步观察本文所提算法的实际检测效果,本文挑选了部分路面坑洼图像进行检测,如图 9 所示。从检测效果来看,改进前后的模型都能准确识别出较大的坑洼,但对于较小目标,原始 YOLOv8 模型出现了不同程度的漏检。本文改进算法在多种场景下均检测出所有坑洼目标,且相比其他模型,预测检测框与真实检测框的重合度更高;综合分析表明,本文算法在所有对比算法中综合性能表现最好。



图 9 检测效果对比

Fig. 9 Comparison of detection effects

3 结论

本文提出了一种改进 YOLOv8n 的路面坑洼检测算法,目的是解决传统 YOLOv8n 在实际应用场景下漏检率较高的问题。首先,将颈部网络替换为双向金字塔网络,进行不同尺度的特征融合,提高检测的准确度;其次,提出了重参数化模块 C2f-DBB,提高模型的推理速度;然后,在骨干网络末端引入通道先验卷积注意力机制 CPCA,在有效提取空间关系的同时保留通道先验,强化特征提取能力;最后,采用边界框回归损失函数 MPDioU 代替 CioU,提高边界框回归精度。

通过对路坑洼图像数据集进行实验,结果表明:改进后的算法在检测精度和模型轻量化方面均取得较好的效果,能够实时准确检测路面坑洼目标。但本文仍在某些方面存在不足,后续研究考虑加强数据收集和数据扩展工作,获取更多高质量的数据,以提升模型的泛化能力和稳定性,此外,考虑引入更高效的网络架构和剪枝技术,进一步降低模型的计算复杂度和显存需求,以适应更多的实际应用场景。

参考文献(References):

- [1] 李松,史涛,井方科.改进 YOLOv8 的道路损伤检测算法[J].计算机工程与应用,2023,59(23):165-174.
LI Song, SHI Tao, JING Fang-ke. Improved road damage detection algorithm of YOLOv8[J]. Computer Engineering and

- Applications, 2023, 59(23): 165–174.
- [2] PAN Y, ZHANG X, JIN X, et al. Road pavement condition mapping and assessment using remote sensing data based on MESMA[J]. IOP Conference Series: Earth and Environmental Science, 2016, 34: 012023.
- [3] 李良福, 孙瑞赟. 复杂背景下基于图像处理的桥梁裂缝检测算法[J]. 激光与光电子学进展, 2019, 56(6): 061002. LI Liang-fu, SUN Rui-yun. Bridge crack detection algorithm based on image processing under complex background[J]. Laser & Optoelectronics Progress, 2019, 56(6): 061002.
- [4] 高建贞. 基于图像分析的道路病害自动检测研究[D]. 南京: 南京理工大学, 2003: 4–10. GAO Jian-zhen. Research on automatic pavement distress detection based on image analysis[D]. Nanjing: Nanjing University of Science and Technology, 2003: 4–10.
- [5] 韩锟, 韩洪飞. 基于区域级和像素级特征的路面裂缝检测方法[J]. 铁道科学与工程学报, 2018, 15(5): 1178–1186. HAN Kun, HAN Hong-fei. Pavement crack detection method based on region-level and pixel-level features[J]. Journal of Railway Science and Engineering, 2018, 15(5): 1178–1186.
- [6] 王玲敏, 段军, 辛立伟. 引入注意力机制的 YOLOv5 安全帽佩戴检测方法[J]. 计算机工程与应用, 2022, 58(9): 303–312. WANG Ling-min, DUAN Jun, XIN Li-wei. YOLOv5 helmet wear detection method with introduction of attention mechanism[J]. Computer Engineering and Applications, 2022, 58(9): 303–312.
- [7] 程长文, 陈玮, 陈劲宏, 等. 改进 YOLO 的口罩佩戴实时检测方法[J]. 电子科技, 2023, 36(2): 73–80. CHENG Chang-wen, CHEN Wei, CHEN Jin-hong, et al. YOLO-improve detection method of real-time mask wearing[J]. Electronic Science and Technology, 2023, 36(2): 73–80.
- [8] 韦正璐, 王家晨, 刘庆华. 基于改进 SSD 算法的路面破损检测[J]. 电子设计工程, 2023, 31(3): 63–68. WEI Zheng-lu, WANG Jia-chen, LIU Qing-hua. Pavement damage detection based on improved SSD[J]. Electronic Design Engineering, 2023, 31(3): 63–68.
- [9] 苏佳, 秦一畅, 贾泽, 等. 基于 ATO-YOLO 的小目标检测算法[J]. 计算机工程与应用, 2024, 60(6): 68–77. SU Jia, QIN Yi-chang, JIA Ze, et al. Small object detection algorithm based on ATO-YOLO[J]. Computer Engineering and Applications, 2024, 60(6): 68–77.
- [10] 王志, 高林, 杨宇. 基于改进 YOLOv7 算法的风力涡轮机表面缺陷检测[J]. 湖北民族大学学报(自然科学版), 2024, 42(1): 75–80. WANG Zhi, GAO Lin, YANG Yu. Surface defect detection of wind turbines based on improved YOLOv7 algorithm[J]. Journal of Hubei Minzu University (Natural Science Edition), 2024, 42(1): 75–80.
- [11] 罗磊, 谢竹逵. 基于改进 YOLOv8 的交通标志检测算法[J]. 机电工程技术, 2024, 53(3): 205–210. LUO Lei, XIE Zhu-kui. Traffic sign detection algorithm based on improved YOLOv8[J]. Mechanical & Electrical Engineering Technology, 2024, 53(3): 205–210.
- [12] 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. Piscataway: IEEE Press, 2017: 936–944.
- [13] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 8759–8768.
- [14] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10778–10787.
- [15] DING X, ZHANG X, HAN J, et al. Diverse branch block: Building a convolution as an inception-like unit[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 10881–10890.
- [16] DING X, GUO Y, DING G, et al. ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 1911–1920.
- [17] 麻斯亮, 许勇. 最小点距离的边界框回归损失函数及其应用[J/OL]. 小型微型计算机系统, 2023, 44(12): 1–8. MA Si-liang, XU Yong. Bounding box regression loss function based on minimum point distance and its application [J/OL]. Journal of Chinese Computer Systems, 2023, 44(12): 1–8.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3–19.
- [19] GUO N, JIANG M, GAO L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks[C]//International Conference on Machine Learning, PMLR, 2021: 11863–11874.
- [20] 郑凯东, 舒心. 基于 BFDS 改进的 YOLOv8 疲劳驾驶检测算法[J]. 信息技术与信息化, 2024(2): 213–216. ZHENG Kai-dong, SHU Xin. Improved detection algorithm of YOLOv8 fatigue driving based on BFDS[J]. Information Technology and Informatization, 2024(2): 213–216.

责任编辑:代小红