

基于改进 YOLOv5 的路面坑洼检测算法研究

王哲兴¹, 李 军¹, 谭 倩²

(1. 重庆交通大学 机电与车辆工程学院, 重庆 400074; 2. 重庆交通大学 交通运输学院, 重庆 400074)

摘要:路面坑洼作为路面结构的重要缺陷之一,对于保障自动驾驶车辆的行驶安全或移动机器人运行具有重大意义。在处理路面坑洼检测时,面临着挑战性的计算机视觉任务,要求在不同的工况下对多样化的数据样本进行处理。诸如雾、雨、雪等天气因素会对道路图像的质量和可见性产生负面影响,进而增加了数据预处理和特征提取的难度。传统的目标检测算法通常难以有效适应这些场景变化,导致训练数据集无法充分反映道路坑洼的多样性和复杂性,从而降低了目标检测模型的泛化能力和准确性。在实际应用中,这些方法容易导致漏检和误检的错误,对道路状况识别和评估的效率和质量造成影响。本文提出了一种基于 YOLOv5 的改进的路面坑洼检测算法,通过引入一个概念简单、功能强大但非常新颖的注意力机制(BiFPN),并替换适合的激活函数与损失函数,提升了模型的检测精度以及查全率,同时降低计算参数,简化检测模型。经实验验证,本文改进的算法与原模型相比准确率(Precision)提升了7.2%,召回率(Recall)提高了5.5%,平均准确率(mAP)达到了80.8%,较原 YOLOv5s 模型提升了2.1%。综上,与常用的一类传统算法相比,本文的改进算法能在几乎不牺牲运行速度的情况下较为明显地提高检测精度,降低漏检率,且不影响检测时的帧率,具有较好的移动端部署价值与对相应研究的参考价值。

关键词:YOLOv5; 图像识别; 坑洼; 计算机视觉

中图分类号:TN958.98; TP391.41 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2024.05.022

Research on pavement pothole detection algorithm based on improved YOLOv5

WANG Zhe-xing¹, LI Jun¹, TAN Qian²

(1. School of Mechanical, Electrical and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2. School of Transportation, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: As one of the important defects of pavement structure, pavement potholes are of great significance for ensuring the driving safety of autonomous vehicles or the operation of mobile robots. When dealing with pavement pothole detection, challenging computer vision tasks are faced, requiring diverse data samples to be processed under different working conditions. Weather factors such as fog, rain, and snow can negative affect the quality and visibility of road images, which in turn increases the difficulty of data preprocessing and feature extraction. Traditional target detection algorithms are usually difficult to effectively adapt to these scenario variations, resulting in training datasets that do not adequately reflect the diversity and complexity of road potholes, which reduces the generalization ability and accuracy

基金项目:国家自然科学基金项目(No. 51305472);重庆市研究生联合培养基地项目(No. JDLHPYJD2018003)资助。

作者简介:王哲兴(1998-),男,硕士研究生,主要研究方向为无人驾驶、环境感知。

通讯作者:谭倩(1999-),女,硕士研究生,研究方向为交通运输规划与管理。E-mail:1540936528@qq.com

收稿日期:2023-06-05; **修订日期:**2023-07-30

of the target detection model. In practical applications, these methods are prone to lead to omission and misdetection errors, which have an impact on efficiency and quality of road condition identification and evaluation. In this paper, an improved pavement pothole detection algorithm based on YOLOv5 is proposed, which improves the detection accuracy as well as the recall rate of the model. By introducing a simple, powerful but very novel attention mechanism (BiFPN) and replacing the appropriate activation function and loss function, while the calculation parameters are reduced and the detection model is simplified. The experimental results show that the improved algorithm in this paper improves the accuracy (Precision) by 7.2 % compared with the original model, the recall rate (Recall) by 5.5 %, and the average accuracy (mAP) by 80.8 %, which is 2.1 % higher than the original YOLOv5 model. In summary, compared with the commonly used traditional algorithms, the improved algorithm in this paper can significantly improve the detection accuracy and reduce the missed detection rate without sacrificing the running speed, which has a better value for mobile deployment and reference value for corresponding research.

Keywords: YOLOv5; image recognition; pothole; computer vision

1 引言

道路交通伤害给人们的生命和财产造成了巨大的损失,已经成为全球性的公共安全问题。根据世卫组织的统筹数据,全世界每年几乎有 130 万人因交通事故死亡,平均每天有超过 3500 人死于道路交通伤害,而由于恶劣天气影响^[1]与夜间行车导致的车祸占了近 7 成,给个人、家庭和国家带来巨大的损失,其中道路路面问题所占比例突出。当前对道路坑洼进行检测的研究较少,并且坑洼检测容易受到环境因素^[2]的影响。例如在缺少光源、阴雨天、雾天的成像效果差^[3],由于道路环境的多样性和复杂性,对抗洼的识别率和精度难以保证,因此在车辆行驶过程中实时检测坑洼的技术仍未成熟。在汽车智能驾驶的技术中,对抗洼的目标检测是一个关键的问题。然而,现有的检测方法在稳健性、鲁棒性和实时性方面仍存在不足,难以适应各种复杂的道路环境。因此,如何提高坑洼检测的定位精度和降低检测延时,是保证自动驾驶车辆能够及时避开坑洼,提高行驶安全和舒适性的重要研究目标^[4]。

Zoysa K D^[5]等学者利用震动法,利用加速度传感器的信号反馈作为坑洼病害的检测依据,是一种成本低、速度快的方法。然而,这种方法也存在一定的局限性,如信号干扰、阈值设定等因素可能导致坑洼的误测和漏检,不能满足精确提取路面坑洼病害的要求。学者杨雷等人^[6]通过对车载激光扫描点云数据进行研究,为了有效地检测城市道路上的坑洼病害,提出了一种基于点云数据的坑洼自动提取方法。该方法利用点云的三维空间分布特征,采用剖面线自适应曲线拟合的技术,能够准确地识别和

定位坑洼区域,提高坑洼检测的效率和精度。但是此类方法基于车载激光,故对于面积较小的坑洼提取会受到点云纵向密度影响,且成本较高。学者焦双健等人^[7]介绍了基于深度学习网络的嵌入式检测系统,该系统的核心思想是利用大量的样本数据训练一个深度学习网络模型,得到一个能够准确识别路面坑洼的最优模型。然后,将其嵌入到英伟达 TX2 平台上运行。最后,通过车载摄像头采集路面图像,输入到系统中,实现路面坑洼的自动检测,但其检测模型的精度与召回率还有提升的空间,未来将尝试降低系统的模型大小,提升检测速度。高建贞^[8]等学者采用图像处理方法来检测路面坑洼,具体步骤如下:首先,对路面坑洼图像进行图像增强^[9]处理,提高其对比度、清晰度;其次,对处理过的图像施行阈值分割以区分坑洼区域及背景区域;然后,提取出坑洼区域的线性特征,如边缘、长度、宽度等;最后,根据线性特征识别出坑洼目标,并计算其位置和大小。该方法能够较高地提高坑洼检测的准确度,但是由于图像处理过程较为复杂,无法满足实时性的检测要求。Wang^[10]等学者开发出一种路面坑洼识别方法,其核心原理是小波能量域与马尔科夫随机场,即通过小波能量域提取路面图像的特征,得到坑洼区域的能量分布特征,然后利用马尔科夫随机场对能量分布特征进行建模和分类,实现坑洼区域的识别和定位。Jo^[11]等学者提出一种低花费的坑洼识别系统,该系统的核心思想是利用图像处理方法来检测路面坑洼,具体步骤如下:首先,采用 Otsu 阈值法将原始图像二值化,把兴趣区域集中于车道线间;其次,区域内的纹理以及几何特征,如灰度、梯度、面积等,提取出坑洼候选区域;最后,对

坑洼候选区域进行分析和判断,识别出坑洼的位置和大小。该系统的优点是时间花费较低,能够实现实时检测。Hoang^[12]等学者提出了一种基于方向可调滤波和监督学习的坑洼识别方法。该方法的核心思想是利用方向可调滤波对路面图像进行预处理,检测出路面上有缺陷的部分,再通过积分投影提取出其中的形状特征,最后利用监督学习模型对形状特征进行分类,实现坑洼图像的识别。以上部分本文综述了文献中的一些路面坑洼识别方法,这些方法都能够实现较高的识别效果,但是仍然存在一些局限性。主要的问题是环境因素的干扰,如光线、阴影、水滩等,影响了图像的质量和特征的提取,导致识别精度受到负面影响。赵珊珊等学者^[13]开发出基于卷积神经网络的路面缝隙探测模型,该方法的核心思想是利用卷积神经网络对路面图像进行特征提取和分类,实现坑洼的识别。首先对路面图像进行图像增强、分割等预处理,处理后的图像会输入至卷积神经网络模型,得到探测结果。该方法通过实验验证了卷积神经网络对于特征明显的坑洼识别具有较强的敏感性,但是容易受到车道线与井盖等杂物的影响,导致其精度不高,缺失对裂缝较大而形成的水坑进行检测。祝一帆等人^[14]提出的一种基于 Crack U Net 的路面坑洼检测方法,该方法的核心思想是利用 Crack U Net 模型对路面图像进行特征提取和分类,实现坑洼的识别。该方法的优点是模型较轻量化,能够部署于移动设备,实现移动检测。但检测精度仅为 59.4%。

综上,为了提高智能汽车、摩托车等移动设备的安全性和舒适性,即增强其对路面坑洼的检测能力,亟需一种能够适应移动设备的灵活部署、并且具有高实时性和高准确性的算法。本文针对路面坑洼的检测问题,提出了一种基于 YOLOv5s 的深度学习方法。本方法在以 YOLOv5s 为基础,对其网络结构、注意力机制、激活、损失函数等部分优化,以适应坑洼的密集分布的特征需求。通过该方法,可以实现路面坑洼的实时检测,并提高检测的效率与精度。本文通过消融实验,验证出本文改进的检测模型的有效性和优越性。实验结果显示,与 YOLOv5s 的基础模型相比,本文检测模型在检测精度方面有了显著的提高,达到了坑洼检测与部署于移动端的基本要求。

2 材料与方法

2.1 数据集

为了验证道路坑洼实时检测的原 YOLOv5 算法与改进后 HDF-YOLO 算法之间的差别以及算法的有效性和实时性,且现有的公开的数据集中并没有大型的坑洼数据集,因此通过在网上查找和实地拍摄的方式获得数据集。为了建立本实验所需的数据集,我们在 python 库中选择了“labelimg”对图像执行手动标注,完成了数据集的构建。

本实验构建的数据集涵盖了交通路面上的各类缺陷,如坑洼、坑洞、水洼等,这些缺陷具有不同的大小、数量和形状,且存在重叠、阴影、遮挡等复杂情况。数据集共包含 665 张图像,从不同的角度和距离拍摄,其中 68 张作为验证集用于加强模型的训练效果。

2.2 坑洼目标检测方法

由于真实移动设备运行环境中的检测的道路坑洼是处于运动状态的,对道路坑洼的目标检测模型必须具有较高的实时性与精准性,YOLOv5 具有速度快、精度高等特点,能满足道路坑洼检测要求故以 YOLOv5 为基本框架模型。

Hardswish-DIoU-BiFPN-YOLOv5 是基于 YOLOv5 改进的模型,简称为 HDF-YOLO,模型结构如图 1 所示。YOLOv5 的基本结构包括 Input、Backbone、Neck 以及 Prediction 4 部分。Input 即输入端包含对图像的 Mosaic 数据增强、自适应锚框计算、自适应图片缩放。Backbone 即骨干网络是一种图像特征提取器,它的主要功能是从输入图像中获取不同层次与维度的特征数据,可以为 Neck 层与 Prediction 层的后续检测任务提供有效的输入。Neck 层即颈部网络是一种特征融合^[15]网络,这层可以有效融合骨干网络提取器提取的低层与高层的特征,进而获得具体细节、语义信息以及高定位信息的数据特征。Prediction 层即 Head 层,是预测网络层也为输出层。是模型检测层,可以对不同尺度的特征图上的多尺寸物体根据融合的特征图来预测,再通过 nms 即非极大值抑制筛选框获得最后的预测的结果。

为了改善 YOLOv5 模型在实际检测情况中存在因运动或环境、天气等因素导致的道路坑洼识别精度差、错误识别或识别率低的情况。HDF-YOLO 在网络原有的基础上的 CBS 卷积模块对激活函数进

行替换,以此提升模型的检测精度与运算效率以及模型的表征能力。利用 BiFPN 替换掉 YOLOv5 特征融合网络中 PANet,可以减少道路坑洼特征信息丢失,以此改善多尺度特征融合能力提升检测精度。利用 DIoU 替换原始损失函数 CIoULOSS 来减少损失模型的运算复杂度、增强训练速度,提升检测精度加快回归精度。

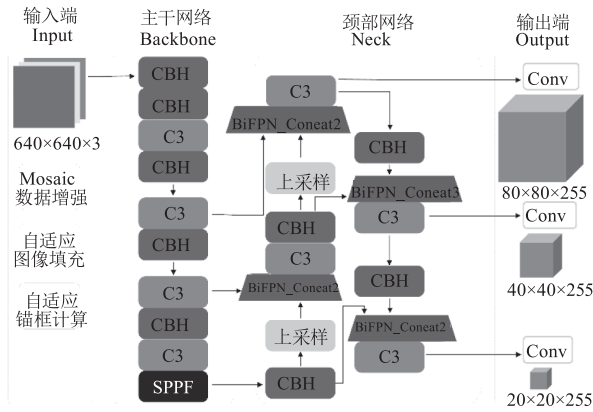


图1 YOLOv5 网络结构图

Fig. 1 YOLOv5 network structure diagram

2.2.1 激活函数改进

在神经网络领域,若全部特征都接受来自上一层的网络输入函数的线性变化,模型的验证的难度都会降低,导致整个网络的逼近能力较低,意味着网络只能解决线性可分问题但不包含非线性问题。激活函数的神经元部分含有非线性成分,可以让深层的神经网络的表达能力更强并且可以应用到非线性模型中。故选取合适的激活函数可以显著提升模型的表达、计算能力以及运行速度。YOLOv5 中使用的是 SILU 函数,此函数是基于 Swish 激活函数改进所得的,Swish 激活函数的表达式如下:

$$\text{Swish}(x) = x \cdot \text{Sigmoid}(\beta x) \tag{1}$$

$$\text{Swish}(x) = \frac{1}{1 + \exp(-x)} \tag{2}$$

当式(1)中的 β 取值为 1 时,Swish 激活函数就变为了 SiLU 激活函数,SiLU 的表达式如下:

$$\text{Swish}(x) = x \cdot \text{Sigmoid}(x) \tag{3}$$

但 SiLU 激活函数容易在输入值接近正负无穷时,函数的导数会接近于 0,这意味着模型的梯度将接近于消失,即阻止梯度的更新,使得模型收敛的难度较大。而 Hardswish 激活函数是一个分段的线性函数,其特点是无上界、有下界,并且具有单调性,能够让模型输入的值趋于无穷时也能维持梯度存在,

避免信息损失,并且相对于其他激活函数,Hardswish 激活函数在计算复杂度与近似性能方面具有优势,相较于 ReLU 与 SiLU 激活函数能提供更好的非线性拟合能力,进而增强模型的表达能力,故本文卷积块的激活函数选用 Hardswish 激活函数。原 CBS 模块修改后的网络结构为 CBH 模块,修改后的网络结构图 2 所示。



图2 CBH 模块结构图

Fig. 2 CBH module structure diagram

Hardswish 激活函数(图 3)的数学表达式如式(4)所示,它根据输入值 x 的不同,分为 3 种情况进行计算:

$$\text{Hardswish} = \begin{cases} 0, & x \leq -3 \\ x, & -3 < x < 3 \\ \frac{x(x+3)}{6}, & \text{otherwise} \end{cases} \tag{4}$$

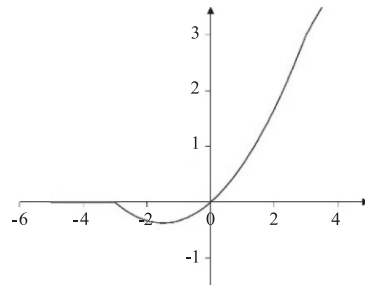


图3 Hardswish 函数图像

Fig. 3 Hardswish function image

2.2.2 特征融合网络改进

在 Backbone 层的对卷积块的改进会改善整个模型对数据集的特征的提取效果,故提取特征变多,则对 Neck 层也会有更多要求。其中,Neck 层的 PANet,即路径聚合网络,其基本功能是通过下采样由下往上的传递目标的定位数据,保证深层的特征图能得到浅层大量的细节数据,因此对提取特征增多则对特征融合网络的要求也会提升。YOLOv5 中的 PANet 的特征传递操作需要在各个阶段之间进行多次信息传递和合并,这导致了较大的通信开销。这可能会增加模型的计算复杂度和推理时间。PANet 采用自上而下和自下而上的特征融合策略,但由于不同层级之间的特征差异,可能存在信息传递不完全或信息损失的问题。这可能影响到对目标的准确表示和检测性能。BiFPN 全称为 Bidirectional

Feature Pyramid Network 加权双向(自顶向下+自低向上)特征金字塔网络^[16], BiFPN 网络通过逐层的特征融合和上下采样操作,能够有效地融合来自不同尺度的特征信息^[17]。这种多尺度的特征融合有助于提升目标检测算法在不同大小目标上的性能。并且 BiFPN 使用了一种自适应的特征融合机制,可以根据不同特征图的重要性动态调整权重。这种机制能够有效地捕捉目标物体的上下文信息和语义信息,减少有用特征的丢失,提升特征融合能力,即提升了目标检测算法的建模能力。此外, BiFPN 网络通过删除没有参与特征融合的节点,简化了网络结构,减少了不必要的参数量和计算量。这使得 BiFPN 网络在实际应用中更具可行性和高效性。

如图4所示,其中“↘”箭头代表由上向下的通道,能够传递高层特征的语义信息;“↓”的箭头代表从下向上的通道,能够传递底层特征的位置数据;曲线箭头为同层中在输入节点与输出节点间新加的一条通道。此外,每个节点都有多个输入和输出通道。而为了判断是否需要增加额外通道,需要计算输入通道之间的相似性分数,即根据特征图的内容和位置信息来进行计算,以此来判断输入通道之间的相关性,若某个输入通道与其他输入通道有比较高的相似性,则说明此通道中含有比较重要特征信息,能直接传输至输出通道。反之,若某通道与其他输入通道的相似性较低,则说明需要增添额外的通道以提供更多的特征信息。而额外通道的数量则需要根据具体任务需求与网络结构来确定,通常额外通道的数量可以通过实验与调优来决定。通过适当增加或删减通道,可以为模型提供更强的特征表达能力,进一步改善目标检测性能。如图4(b)中 P7 层只有一条输入边的节点被删除,因为只有一条输入边,却没有特征融合,那么这个节点在网络在融合多种特征的时候特征信息贡献就会很小。删除此节点对网络的影响较小,同时还能简化双向网络。经

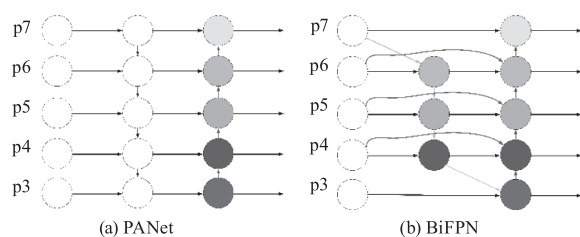


图4 PANet 与 BiFPN 网络结构对比图

Fig. 4 Comparison of PANet and BiFPN network structure

过实验,对于 P4、P5、P6 三层, BiFPN 在输入节点与输出节点中增添了条额外通道,这能使网络在不增加过多额外的计算和存储开销的前提下补充、融合更多网络特性。故本文引用改进的 BiFPN 替换特征融合网络中的 PANet。

2.2.3 损失函数的改进

损失函数存在于检测网络,即 Head 层中。YOLOv5 自带的 CIoU 引入了更多的约束条件,考虑了预测框和真实框之间的重叠区域,虽然可以更好地避免由于预测框和真实框相交而导致 IoU 计算不准确的问题,在计算目标之间的距离时,能考虑目标的大小和位置,可以更加准确地区分不同类别之间的目标,但会因此反而会导致仅需要单独识别目标的模型的精度与召回率受到影响,且 CIoU 有更高的计算复杂度,训练速度相对较慢等特性。其次 DIoU 损失函数在计算边界框之间的距离时,考虑了目标的尺寸信息。这有助于减少尺寸较小的目标对损失函数的影响,从而提高目标检测模型对小尺寸目标的鲁棒性。此外, DIoU 损失函数能够更好地引导模型学习准确的目标定位,因为它在计算损失时考虑了目标之间的距离。这有助于减少定位误差,提高目标检测的定位准确性。因此提出使用距离交并比(Distance Intersection over Union, DIoU)代替 CIoU 的方案。

DIoU 引入的距离惩罚项,还能惩罚预测框与真实框之间的距离差异,可提高检测精度,加快回归精度。另外 DIoU 损失函数在计算时更加稳定,可以有效地规避训练导致的梯度爆炸或消失的情况,让模型的训练更稳固。DIoU 公式:

$$DIoU = \frac{|C \cap D|}{|C \cup D|} - \frac{\rho^2(b, b^{gt})}{c^2} \quad (5)$$

式中: c 为预测框与真实框所在最小闭合区域中的对角线长; $\rho^2(b, b^{gt})$ 为两框的中心点间的欧几里得距离^[18]。 b 与 b^{gt} 为预测矩形框与标注矩形框的中心点, C 为预测矩形框, D 为标注矩形框。则 DIoU 的损失函数公式为:

$$L_{DIoU} = 1 - \frac{|C \cap D|}{|C \cup D|} + \frac{\rho^2(b, b^{gt})}{c^2} \quad (6)$$

综上,本文提出的一种改进的模型,称为 HDF-YOLO,其网络结构和组件结构如图5和图6所示。本文修改后的模型网络整体结构如表1所示。

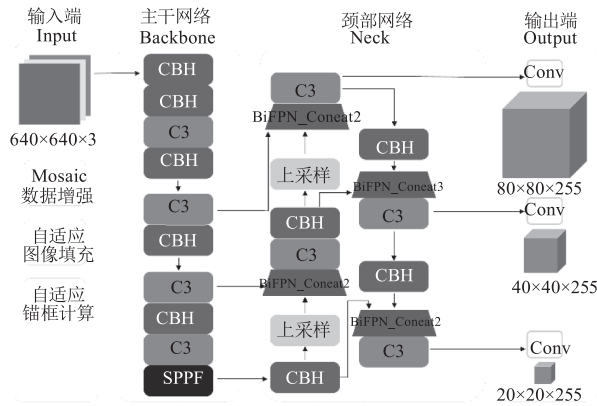


图 5 HDF-YOLO 网络结构图

Fig. 5 HDF-YOLO network structure diagram

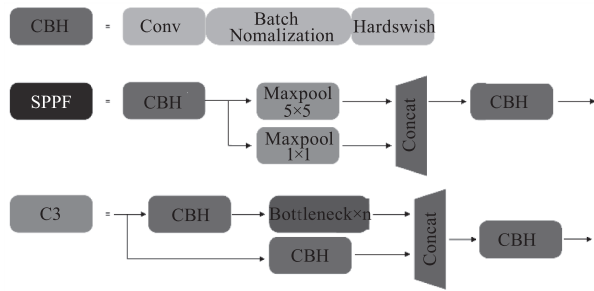


图 6 HDF-YOLO 网络组件结构图

Fig. 6 HDF-YOLO network component structure diagram

2.3 试验平台与评估指标

2.3.1 实验平台与模型训练参数

本文的实验环境基于 Windows11 操作系统, CPU 型号为 12th Gen Intel (R) Core (TM) i7 - 12700H 2.30 GHz, 16G RAM, GPU 型号为 GPU NVIDIA RTX3060 6GB, 显存为 6 GB。网络训练的调参如下:原始输入图片尺寸为 640 × 640, 初始学习率设为 0.01, 周期学习率为 0.1, 权重衰减系数为 0.0005, 动量系数为 0.8, batch size 设置为 16, epochs 为 300 轮。

2.3.2 评估指标

为实现评测模型的检测能力的目的, 本文选用以下指标作为评判标准: 准确率 (Precision, P)、召回率 (Recall, R)、平均精度 (mean Average Precision, mAP)、参数量 (parameters)、每秒浮点运算次数 (Floating Point Operations per Second, FLOPs)、帧率 (Frames Per Second, FPS)。其中, P 表示检测正确的目标占有检测结果的比例, R 表示检测正确的目标占有真实目标的比例, mAP 代表不同类别的 P 与 R 的平均值。 P 、 R 、 mAP 的计算公式分别如式 (7)、(8)、(9) 所示:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (9)$$

在本实验中, 我们使用 TP (True Positive) 表示检测框正确地预测了正例, 即水坑被成功地识别出来。我们使用 FP (False Positive) 表示检测框错误地预测了负例, 即非水坑的区域被误判为水坑。 FN (False negatives) 为没有被预测出的坑洼。 mAP 表示 PR 曲线下所围成的面积, 它是对 P 和 R 的一种综合性评价指标, $mAP@0.5$ 表示 IOU (Intersection over Union) 为 0.5 时的 mAP , 多数情况默认 IOU 阈值为 0.5。即大于 0.5 的认为是识别目标。这种方法得到的 mAP 称为 $mAP@0.5$, $mAP@0.5:0.95$ 为 IOU 取值位于 0.5 到 0.95 的区间同时其步长为 0.05 时的 mAP 的平均值, N 代表类别总数。FPS 则是代表模型每秒传输的图片帧数, 其数值越大, 表示检测的画面流畅度越高, 也说明网络的检测速度越快, 实时性越强。

表 1 HDF-YOLO 网络结构列表

Tab. 1 List of HDF-YOLO network structures

	from	n	params	module	arguments
0	-1	1	3520	models. common. Conv	[3, 32, 6, 2, 2]
1	-1	1	18560	models. common. Conv	[32, 64, 3, 2]
2	-1	1	18816	models. common. C3	[64, 64, 1]
3	-1	1	73984	models. common. Conv	[64, 128, 3, 2]
4	-1	2	115712	models. common. C3	[128, 128, 2]
5	-1	1	295424	models. common. Conv	[128, 256, 3, 2]
6	-1	3	625152	models. common. C3	[256, 256, 3]

(续表)

	from	n	params	module	arguments
7	-1	1	1180672	models. common. Conv	[256,512,3,2]
8	-1	1	1182720	models. common. C3	[512,512,1]
9	-1	1	656896	models. common. SPPF	[512,512,5]
10	-1	1	131584	models. common. Conv	[512,256,1,1]
11	-1	1	0	torch. nn. modules. upsampling. Upsample	[None,2,'nearest']
12	[-1,6]	1	65794	models. common. BiFPN_Add2	[256,256]
13	-1	1	296448	models. common. C3	[256,256,1,False]
14	-1	1	33024	models. common. Conv	[256,128,1,1]
15	-1	1	0	torch. nn. modules. upsampling. Upsample	[None,2,'nearest']
16	[-1,4]	1	16514	models. common. BiFPN_Add2	[128,128]
17	-1	1	74496	models. common. C3	[128,128,1,False]
18	-1	1	295424	models. common. Conv	[128,256,3,2]
19	[-1,13,6]	1	65795	models. common. BiFPN_Add3	[256,256]
20	-1	1	296448	models. common. C3	[256,256,1,False]
21	-1	1	590336	models. common. Conv	[256,256,3,2]
22	[-1,10]	1	65794	models. common. BiFPN_Add2	[256,256]
23	-1	1	1051648	models. common. C3	[256,512,1,False]

4 试验设计与结果分析

4.1 消融实验

为验证本研究对 YOLOv5 改进的有效性,设计消融试验内容如下:依次为 YOLOv5 模型更改 Hardswish 激活函数、BiFPN 模块 DIoU 损失函数来训练模型进而对比实验情况,模型检测对比都使用相同的数据集,训练批次与轮数等实验设定都相同,以验证不同算法的检测性能,并且都使用 YOLOv5s-pt 作为预训练权重模型,进行三百轮训练,训练完成后,选取训练效果最好的一轮所产生的模型,在测试集上进行实验评估各个算法的性能指标,训练过程所得的 YOLOv5 与 HDF-YOLO 两个模型曲线变化图 7 所示。三角形图标曲线表示改进模型,正方形图标的曲线表示原模型。由图 7 可知,两指标曲线在一百轮前呈现不断增加的趋势,在一百轮训练之后趋于稳定,YOLOv5s 原模型会提前出现训练拟合的情况,而改进的 YOLOv5s 坑洼检测模型则不会,且 mAP 值与 precision 值较原 YOLOv5s 模型均有比较明显的提升。

本文又通过使用 YOLOv5s、YOLOv5-Bot、YOLOv5-Hardswish、YOLOv5-Diou 和 HDF-YOLO 等

算法进行消融实验,以验证不同算法的检测性能,均在前文介绍的数据集上进行了模型训练,并以 YOLOv5s-pt 作为预训练权重模型。训练完成后,在测试集上进行实验评估了各个算法的性能指标,结果如表 2 所示。

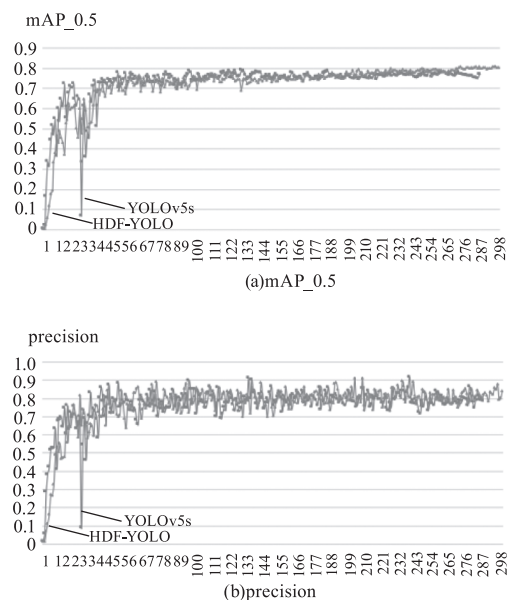


图 7 初始模型与改进模型的实验指标变化曲线

Fig. 7 The experimental index change curves of the initial model and the improved model

表 2 HDF-YOLOv5 实验对比表

Tab.2 HDF-YOLOv5 experimental comparison table

模型	准确率 $P/\%$	召回率 $R/\%$	平均精度均值 $mAP/\%$	parameters	FLOPs/B	FPS/ $(f \cdot s^{-1})$
YOLOv5s	77.30	69.60	78.70	7012822	15.8	96.1538
YOLOv5-BiFPN	76.90	71.80	78.40	7161311	16.4	96.1538
YOLOv5-Hardswish	76.50	71.80	77.40	7012822	15.8	98.0392
YOLOv5-DIoU	77.00	70.10	77.60	7012822	15.8	92.5926
HDF-YOLO	84.50	75.10	80.80	7161311	16.4	97.0874

从表 2 中可以看出,虽然参数量增加了 148489,计算量增加了 0.6 B,但 HDF-YOLO 模型的 P 、 R 、 $mAP@0.5$ 等多个指标与初始模型 YOLOv5s 的对应指标相比,均出现了一定提升,准确率 P 提高了 7.2%,召回率提高了 5.5%, $mAP@0.5$ 提升了 2.1%,FPS 几乎维持不变。说明在准确度、查全率大幅提升的情况下,模型部署能力与检测速度几乎没有受到影响。YOLOv5-BiFPN 模型的召回率 R 比原模型提升了 2.2%,但是准确率、平均精度均有所下降,且参数量与运算量有一定提升,说明 BiFPN 注意力机制能够有效提升模型的查全率,但是会牺牲一部分检测精度、增大模型体积与降低运行性能;YOLOv5-Hardswish 模型 FPS 仅仅提升了 1.89,却降低了模型检测的准确性与平均精度,使得模型检测能力不够。同样的,YOLOv5-DIoU 网络会提升模型的查全率,但是却牺牲了检测精度,即减少了漏检的情况,但会增加误检错检的概率,从而导致模型的检测效果会变差。以上模型的实验数据说明,HDF-YOLO 模型的改进发挥出了一加一大于二的作用,对模型的检测性能提升有一定帮助。

3.2 模型对比

为了更好地验证改进的 HDF-YOLO 模型相较于原模型 YOLOv5s 的提升效果,本文使用不同环境下道路坑洼检测数据集对模型进行测试。图 8 为 YOLOv5s 原模型与 YOLOv5s-BHD 模型对相同图片进行检测的结果对比,图 8 中(a)、(c)为原 YOLOv5s 模型测试图,图 8 中(b)、(d)为改进的 YOLOv5s 模型即 HDF-YOLO 的检测效果。HDF-YOLO 模型用 DIoU 损失函数替换了 CIoU 损失函数,增强了模型对尺寸较小目标的识别能力与定位能力,提升了检测精度。图(a)在图中央有明显的道路坑洼的情况下出现了漏

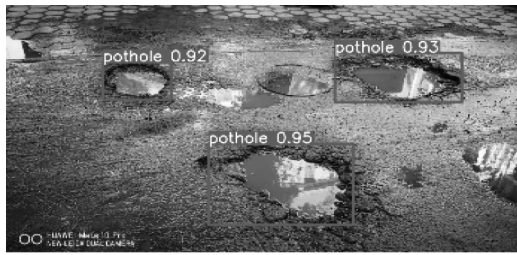
检,以及对图右下角的有阴影且颜色不同的平坦区域误检为坑洞。图(b)则成功检测出了汽车视角下图中央的坑洼,也没有对图右下角的阴影部分错误识别,且对道路坑洼检测的有着更高的置信度,说明 HDF-YOLO 模型在干旱道路上有少量坑洼的情况下增加了对坑洼识别的查全率。改进模型替换了 BiFPN 结构,即增强了注意力集中机制,关注更多目标特征以及位置信息,减少对无关信息的注意,提升识别精度,使得模型识别精度与查全度得到加强,识别效果更好。图(c)显示,在潮湿的道路上坑洼较为密集的情况下漏检了许多特征较为明显的坑洼目标,图(d)则检测到了特征明显的目标,明显减少了漏检的情况,HDF-YOLO 模型增强了对图中间与边缘的目标的识别率。同样图(d)所展示的图片中,HDF-YOLO 的检测都有更高的置信度,能为智能汽车的决策提供更有力的数据信息支撑。综上所述,本文所改进的模型拥有更好的检测效果,更适合完成道路坑洼检测任务。



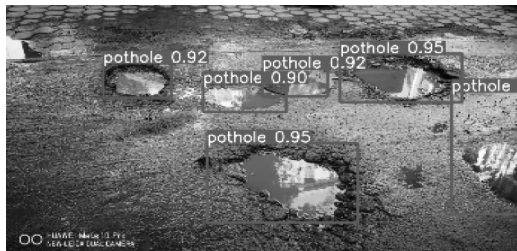
(a)YOLOv5s干燥道路测试效果



(b)HDF-YOLO干燥道路测试效果



(c)YOLOv5s潮湿道路测试效果



(d)HDF-YOLO潮湿道路测试效果

图8 模型检测效果对比

Fig. 8 Comparison of model detection effects

4 结论

目前对于道路坑洼进行检测的算法的研究与改进较少,受实操环境与各种工况的影响现有的检测准确度与查全能力都不够完善,实际运行能力有待加强,参数量与计算量使得算法在机器的部署也会受到影响,检测与运行会存在误差干扰,本文针对目前研究较少的坑洼检测领域,对YOLOv5坑洼检测算法进行检测能力的改进,增强模型的表达能力,通过引入BiFPN网络来强化网络捕捉目标特征的能力,增进特征融合能力,简化网络结构,并对无关特征信息进行抑制。使用DIoU提高目标的检测精度与定位精度,从而增强模型的鲁棒性。结果表明,HDF-YOLO算法在开源的坑洼数据集上表现出色,相较于YOLOv5s算法,在模型的检测精度提升了7.2%,召回率提高了5.5%,平均精度提升了2.1%的情况下,FPS几乎没有改变,即代表检测速度几乎没有受到影响,而对道路的坑洼检测更准确,具有较强的扩展性以及鲁棒性,满足在小型边缘设备上的部署要求,对交通安全方面具有重要的实际应用,对相关研究也有一定帮助^[19]。

参考文献:

[1] Yang Feng, Yang Qin, Chen Xiaolin, et al. Simulation analysis of the influence of bad weather on the performance of frequency modulated continuous wave lidar [J]. Laser

& Infrared, 2023, 53(5):663-669. (in Chinese)

杨峰, 阳琴, 陈孝林, 等. 恶劣天气对调频连续波激光雷达性能影响仿真分析[J]. 激光与红外, 2023, 53(5):663-669.

[2] Huang Zhiqiang, Li Jun. Research on environmental perception technology of driverless cars [J]. Equipment Machinery, 2021, (1):1-6. (in Chinese)

黄志强, 李军. 无人驾驶汽车环境感知技术研究[J]. 装备机械, 2021, (1):1-6.

[3] Huang Zhiqiang, Li Jun. Research on traffic sign recognition based on spatial channel attention mechanism and multi-scale fusion [J]. Journal of Nanjing University of Posts and Telecommunications: Natural Science Edition, 2022, 42(2):93-102. (in Chinese)

黄志强, 李军. 基于空间通道注意力机制与多尺度融合的交通标志识别研究[J]. 南京邮电大学学报:自然科学版, 2022, 42(2):93-102.

[4] Shi Yancheng, Li Jun. Sensor fusion technology in the field of autonomous driving [J]. Equipment Machinery, 2021(3):1-6, 12. (in Chinese)

石晏丞, 李军. 汽车自动驾驶领域的传感器融合技术[J]. 装备机械, 2021(3):1-6, 12.

[5] Zoysa K D, Keppitiyagama C, Seneviratne G P, et al. A public transport system based sensor network for road surface condition monitoring [C]//Workshop on Networked Systems for Developing Regions, 2007.

[6] Yang Lei, Liu Rufe, Lu Xiushan, et al. An automatic extraction method of pavement pits in vehicle-borne laser scanning point cloud [J]. Surveying and Mapping Engineering, 2020, 29(1):66-71. (in Chinese)

杨雷, 刘如飞, 卢秀山, 等. 一种车载激光扫描点云中路面坑槽自动提取方法[J]. 测绘工程, 2020, 29(1):66-71.

[7] Jiao Shuangjian, Du Fujun. Design of pavement pit detection system based on deep learning [J]. Single Chip Microcomputer and Embedded System Application, 2022, 22(7):10-13. (in Chinese)

焦双健, 杜福君. 基于深度学习的路面坑洼检测系统设计[J]. 单片机与嵌入式系统应用, 2022, 22(7):10-13.

[8] Gao Jianzhen. Research on automatic detection of road diseases based on image analysis [D]. Nanjing: Nanjing University of Science and Technology, 2003. (in Chinese)

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.

- [9] Huang Xu,Zhang Shiyi,Li Jun. Review of image segmentation technology[J]. Equipment Machinery,2021,(2):6-9. (in Chinese)
黄旭,张世义,李军. 图像分割技术研究综述[J]. 装备机械,2021,(2):6-9.
- [10] Wang P,Hu Y,Dai Y, et al. Asphalt pavement pothole detection and segmentation based on wavelet energy field [J]. Mathematical Problems in Engineering,2017.
- [11] Jo Y,Ryu S. Pothole detection system using a black-box camera[J]. Sensors,2015,15(11):29316-29331.
- [12] Hoang N D. An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature extraction[J]. Advances in Civil Engineering,2018,4:7419058. 7419051-7419058. 7419012.
- [13] Zhao Shanshan,He Ning. Pavement crack detection based on convolutional neural network [J]. Sensors and Microsystems,2017,36(11):135-138. (in Chinese)
赵珊珊,何宁. 基于卷积神经网络的路面裂缝检测[J]. 传感器与微系统,2017,36(11):135-138.
- [14] Zhu Yifan,Wang Haitao,Li Ke, et al. A high-precision pavement crack detection network structure:Crack U-Net [J]. Computer Science,2022,49(1):204-211. (in Chinese)
祝一帆,王海涛,李可,等. 一种高精度路面裂缝检测网络结构:Crack U-Net[J]. 计算机科学,2022,49(1):204-211.
- [15] Zhang Liguo,Ma Zijian,Jin Mei, et al. Pedestrian detection method for lightweight infrared image based on YOLO [J]. Laser & Infrared,2022,52(11):1737-1744. (in Chinese)
张立国,马子荐,金梅,等. 基于 YOLO 的轻量级红外图像行人检测方法[J]. 激光与红外,2022,52(11):1737-1744.
- [16] Kang Zixiang,Wang Shengzhe,Cui Yuyong, et al. A voxelized laser point cloud target detection algorithm based on transformer[J]. Laser & Infrared,2023,53(2):202-207. (in Chinese)
康自祥,王升哲,崔雨勇,等. 基于 Transformer 的体素化激光点云目标检测算法[J]. 激光与红外,2023,53(2):202-207.
- [17] Wang Xiaohong,Chen Zheqi. Research on pedestrian detection in infrared images based on YOLOv5 algorithm [J]. Laser & Infrared,2023,53(1):57-63. (in Chinese)
王晓红,陈哲奇. 基于 YOLOv5 算法的红外图像行人检测研究[J]. 激光与红外,2023,53(1):57-63.
- [18] Xia Zhenghong,Wei Ruxiang,Li Yandong. Moving target detection of airport surface under improved YOLOv3 algorithm [J]. Chinese Journal of Safety Science,2023,33(2):82-88. (in Chinese)
夏正洪,魏汝祥,李彦冬. 改进 YOLOv3 算法下通航机场场面运动目标检测[J]. 中国安全科学学报,2023,33(2):82-88.
- [19] Zou Peng,Yang Kaijun,Liang Chen. Improve the lightweight non-standard driving behavior real-time detection of YOLOv5 [J]. Computer Engineering and Application,2023,13:186-193. (in Chinese)
邹鹏,杨凯军,梁晨. 改进 YOLOv5 的轻量级不规范驾驶行为实时检测[J]. 计算机工程与应用,2023,13:186-193.