

基于改进 YOLOv5s 的复杂环境道路坑洼检测方法

魏武¹, 龚家元^{1,*}, 车凯², 朱政泽¹

¹ (湖北汽车工业学院汽车工程师学院, 十堰 442002)

² (电子科技大学光电科学与工程学院, 成都 611731)

摘要: 为使自动驾驶系统能对复杂环境道路坑洼进行有效检测与定位, 对现有 YOLOv5 目标检测算法进行改进。首先, 用 MobileNetV3 替换原模型的骨干部分, 减少模型参数量, 轻量化网络模型; 此外, 在 YOLOv5 的颈部引入 BiFPN 模块, 提升模型在多尺度特征融合、信息传递、特征表达能力以及检测精度方面的性能, 同时保持模型的轻量级设计; 然后, 引入生成对抗神经网络 (GAN) 的图像风格迁移的思想, 采用 PaddleGan 进行数据处理, 丰富数据集的多样性; 最后, 在自制的数据集上进行模型的训练与验证, 实验结果表明: 改进的 YOLOv5 算法准确率和 mAP 值分别提升了 3.5%、0.9%, 检测速度提升了 5.8 帧/s。改进的算法更加的轻量化, 提高了检测的精度, 为复杂环境下坑洼检测提供技术参考。

关键词: BiFPN; YOLOv5; 坑洼检测; 风格迁移; MobileNetV3

Road Pothole Detection Method in Complex Environment Based on Improved YOLOv5s

WEI Wu¹, GONG Jiayuan^{1,*}, CHE Kai², ZHU Zhengze¹

¹ (Institute of Automotive Engineers, Hubei University of Automotive Technology, Shiyan, 442002, China)

² (College of Optoelectronic Science and Engineering, University of Electronic Science and Technology, Chengdu, 611731, China)

Corresponding Author: GONG Jiayuan. Institute of Automotive Engineers, Hubei University of Automotive Technology, Shiyan, 442002, China. Email: jyong@huat.edu.cn.

Abstract: To enable autonomous driving systems to effectively detect and locate road potholes in complex environments, improvements have been made to the existing YOLOv5 object detection algorithm. Firstly, MobileNetV3 is employed to replace the original backbone of the model, reducing the parameter count and achieving a more lightweight network design. Additionally, a BiFPN (Bidirectional Feature Pyramid Network) module is introduced in the neck of YOLOv5, significantly enhancing the model's performance in multi-scale feature fusion, information propagation, feature representation, and detection accuracy, while maintaining the lightweight nature of the architecture. Furthermore, the concept of image style transfer from Generative Adversarial Networks (GANs) is incorporated, utilizing PaddleGAN for data augmentation to increase the diversity of the dataset. Finally, experiments conducted on a custom dataset revealed that the improved YOLOv5 algorithm achieved a 3.5% increase in accuracy, a 0.9% improvement in mAP, and an enhancement in detection speed by 5.8 frames per second (fps). The proposed algorithm is more lightweight and enhances detection precision, providing a valuable technical reference for pothole detection in complex environments.

Keywords: BiFPN; YOLOv5; Pothole Detection; Style Transfer; MobileNetV3

Funding: This work is supported by Natural Science Foundation of Hubei Province of China (Joint Fund for

Innovation and Development of Shiyao) (2024AFD116); Key Project of Science and Technology Research Plan of Hubei Provincial Department of Education (D20231805); Doctoral Research Start-up Fund of Hubei University of Automotive Technology (BK202307, BK201604); Natural Science Foundation of Hubei Province of China (2023AFB481)

1 引言

随着我国道路里程和机动车保有量的持续增长,交通事故的发生频率也在不断增加,据《中国统计年鉴-2022》数据显示,过去十年交通事故数量增长了近30%,每年导致超过30万人死伤。其中,糟糕的道路环境是造成交通事故的一个重要因素,包括坑洼、裂缝、下沉等,这些都会影响行车的平稳性和安全性。坑洼会降低人们的出行舒适度,加剧汽车的损耗,不利于行人和汽车的安全,增大交通事故出现的风险,如汽车磨损、爆胎甚至是车祸,造成不必要的经费支出。当前,许多学者已经对道路坑洼检测进行了深入研究,并提出了多种检测算法。然而,在复杂环境条件下,这些算法的识别率和精度仍然面临诸多挑战。因此,如何在复杂环境下高效、准确地检测道路坑洼,已成为一个关键且迫切需要解决的技术难题。

Zhang^[1]等学者引入了依赖于深度卷积神经网络的路面缺陷探测方法。该方法针对路面缺陷检测中的复杂场景有一定的检测性能,解决了干扰检测的噪声问题与裂缝难区分问题。张旭^[2]提出了一种小型化的裂缝检测模型 Tiny-YOLOv3,它基于 YOLOv3 和轻量级网络 Tiny-Darknet,通过 OpenVINO 平台优化模型并加速,既能够提高检测精度,又能够提升运行速度;Guo 等^[3]学者用 MobileNetV3 替换 YOLOv5s 中的基础网络,引入坐标注意力机制,经实验,改进后的网络 mPA 值提高了 2.5%,参数量减少了 1.66 倍,但模型泛化能力还要提升。张果荣^[4]对收集到的裂缝图像进行了滤波处理,并对经典的 U-Net 网络模型进行了改进,引入了空洞卷积技术,在减少模型参数数量的同时,提高了算法检测的效率。于海洋^[5]等人基于 U-Net 神经网络模型,结合 CBAM 注意力机制和残差网络,进一步提升了道路裂缝的检测能力。陈鹏^[6]通过采用多种数据增强方法来丰富坑洼数据集,构建了一个多类别的数据集,分别找到了最佳的分类网络,最终将改进后的多场景道路坑洼分类算法与 Faster RCNN 结合起来,形成了一个综合性能较好的检测网络。Hu 等^[7]学者将 C2f 结构的 Decoupled Head 集成,对 Silu 激活函数和 CIoU 损失函数进行优化,算法的检测速度达到 85FPS,但在低光照下的检测效果欠佳。Zhou 等^[8]学者用双向特征金字塔网络替换 PANet,实验结果在 Road Damage 数据集上 mAP50 提升了 5.4%。学者杨雷^[9]等人提出一种面线自适应曲线拟合的坑槽提取方法。该方法能够准确提取出坑槽的面积信息和深度信息,并且有较小的误差,但该方法对于较小的坑槽效果一般。在已有的大量对坑洼检测的研究中可以看出,检测精度和模型参数量的大小还不能很好的平衡,并且大多都是针对正常天气情况下的坑洼检测。缺乏对雨、雪、雾等特殊天气情况下以及夜间照明不足的复杂道路条件下的研究。

针对以上问题,本文提出了一种基于 YOLOv5s 改进的坑洞检测方法,以提高检测准确率,并提高模型的泛化能力,同时减少模型参数大小,使检测算法更易于在嵌入式系统上部署。本文的主要工作包括对 YOLOv5s 模型的 Backbone 部分与 Neck 部分进行改进,使用基于风格迁移的思想对数据集进行数据增强,创建一个增强后的数据集。其中,对原始模型的主干部分的改进为使用 MobileNetV3 替换原模型的骨干部分,对颈部进行的改进为引入 BiFPN 模块。

本文的结构安排如下:第 1 节研究国内外现状与提出问题;第 2 节介绍本文改进的坑洼检测算法;第 3 节介绍数据集增强的方法;第 4 节进行实验与结果分析;第 5 节对本文方法进行总结。

2 复杂环境道路坑洼检测网络模型构建

2.1 YOLOv5s 网络模型

YOLOv5^[10]是一种广泛应用的目标检测模型，由三个部分组成，分别是主干网络（Backbone Network）、颈部网络（Neck Network）和检测模块（Detect）。主干网络在整个模型中起着至关重要的作用，负责提取坑洼图像的特征，是网络处理的初始环节。颈部网络的功能是将主干网络提取的特征进行融合，而检测模块则通过生成边界框来完成坑洼的检测任务。

YOLOv5 使用 CSPDarknet53 作为特征提取网络^[11]。其特征融合的颈部网络由快速空间金字塔池化（SPPF）网络、特征金字塔网络（FPN）^[12]和像素聚合网络（PAN）^[13]共同组成。YOLOv5 提供了四种不同规模的模型，参数量从小到大依次为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。其中，YOLOv5s^[14]作为计算资源需求最低、检测速度最快的版本，特别适合资源有限的设备使用，即便是在嵌入式平台上也能实现较高的预测效率。因此，选择 YOLOv5s 作为改进后的坑洼检测基础模型。

YOLOv5s 的损失函数由三个部分组成：分类损失、定位损失和置信度损失。分类损失采用二值交叉熵（BCE）损失，用于指示锚框是否与先前校准的分类匹配。定位损失通过使用完全交并比（Complete-IOU, C-IOU）^[15]衡量预测框与校准框之间的差异。此外，置信度损失同样通过 BCE 损失表示网络的置信度误差。

2.2 替换 Backbone 层

为了降低 YOLOv5s 模型的大小，替换 Backbone 为 MobileNetV3。相较于原始 YOLOv5 的主干网络，MobileNetV3 能够大大降低模型的计算开销和参数量，这对实时道路检测系统尤为重要，与此同时，MobileNetV3 通过优化结构保持了较高的特征提取能力，使得它在有效捕捉道路坑洼的关键细节时性能较好。

MobileNetV3^[16]继承了 MobileNetV1 的深度可分离卷积和 MobileNetV2 的线性瓶颈残差结构，引入了基于 MobileNetV2 的挤压激励(Squeeze-and-Excitation, SE)通道注意力结构，并将 MobileNetV2 网络端部最后阶段修改为了图 3 所示的结构，是一种为移动和边缘设备优化的卷积神经网络架构。为了更高的准确性，使用了一种新的激活函数，称为 Swish，以替代 ReLU 函数。这个函数的计算公式如式(1)所示。

$$\text{swish } x = x\sigma(x) \quad (1)$$

其中， x 为输入值； σ 为 sigmoid 函数。

然而，swish 公式中的 sigmoid 函数可能使得在移动设备上耗费大量的计算资源。为了解决这个问题，MobileNetV3 的作者们使用 ReLU6 函数来近似 swish 公式中的 sigmoid 函数，并产生了一个称为 h-swish 的 swish 函数的近似值，其计算公式为：

$$\text{h-swish}[x] = x \frac{\text{ReLU6}(x+3)}{6} \quad (2)$$

图 1 显示了 MobileNetV3 的网络模型结构，分为三部分：第一部分由 1 个卷积层组成，通过 3×3 卷积提取特征；第二部分包含多个卷积层，针对两个不同的版本有不同的层级和参数，小型版本有 13 个卷积层，大型版本有 15 个；第三部分是为了减少参数和计算，提前进行了 Avg Pooling，用两个 1×1 卷积层替代了整个全连接层，最后输出分类结果。

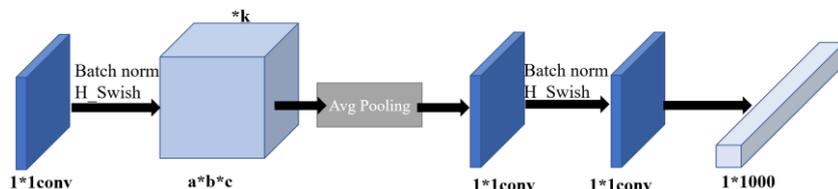


图 1 MobileNetV3 网络模型结构

Fig. 1 MobileNetV3 network model architecture

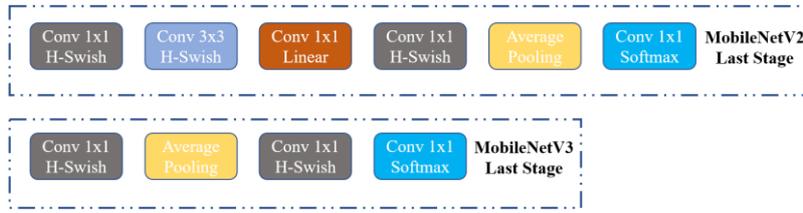


图 2 MobileNetV2 与 MobileNetV3 网络端部最后阶段对比

Fig. 2 Comparison of the final stages of MobileNetV2 and MobileNetV3 network endpoints

使用 MobileNetV3 替换 YOLOv5 的主干部分，主要是为了实现模型轻量化、提高推理速度、适应资源受限的移动和嵌入式设备，从而在保持相对较高检测精度的同时，优化实时应用的性能。

2.3 改进 Neck 层

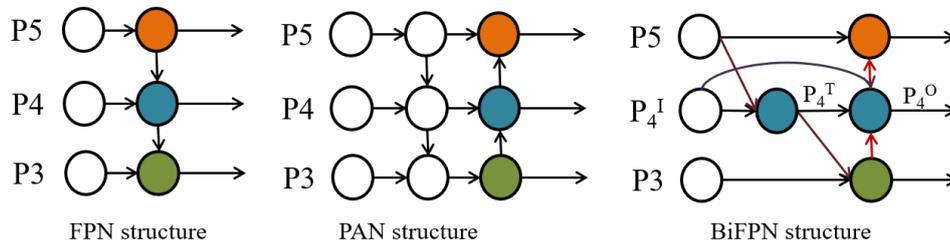


图 3 FPN, PAN 和 BiFPN 结构

Fig. 3 FPN, PAN, and BiFPN structure

在坑洼检测任务中，YOLOv5 的目标是准确地识别和定位不同尺寸的坑洼，而传统的特征金字塔网络（FPN）虽然能够处理不同尺度的目标，但其自顶向下的结构受到单向信息流的限制，可能无法充分利用各尺度之间的特征信息。^[17,18]为了解决这个问题，采用的路径聚合网络（PAN）增加了自底向上的聚合路径，提高了准确性，但也导致了参数量和计算复杂度的增加。因此，本文选择了加权双向（自顶向下加自底向上）的 BiFPN^[19]。BiFPN 的核心思想是通过高效的双向跨尺度连接和加权特征融合实现路径增强。其过程首先进行自顶向下的特征融合，然后进行自底向上的特征融合。其结构如图 3 所示。

具体来说，BiFPN 模块的引入通过多次特征融合，增强了模型对细节信息和大尺度坑洼的感知能力，从而提升了检测精度。同时，BiFPN 的加权特征融合方式能够根据特征的重要性进行动态调整，避免了传统 FPN 中可能出现的信息冗余问题。

此外，BiFPN 通过双向信息流增强了上下文信息的传递，帮助网络更好地理解坑洼的整体布局和局部细节，尤其在复杂的道路环境中，坑洼的形态、大小和位置变化较大。通过添加 BiFPN 模块使得 YOLOv5 在多尺度特征融合上具有更大的优势。

2.4 改进的 YOLOv5s 网络结构

通过在 YOLOv5s 模型中将主干网络替换为 MobileNetV3，并在 Neck 部分引入 BiFPN 模块，旨在提升模型的特征提取能力和多尺度特征融合性能，进而提高检测精度和效率。首先，MobileNetV3 作为轻量化的卷积神经网络，具有较低的计算复杂度和较小的参数量，在保持较高准确度的同时，显著提升了模型的推理速度。这一改进尤其适合嵌入式系统或资源受限的环境，能够加速目标检测任务，尤其是在处理复杂背景时，减少计算负担。

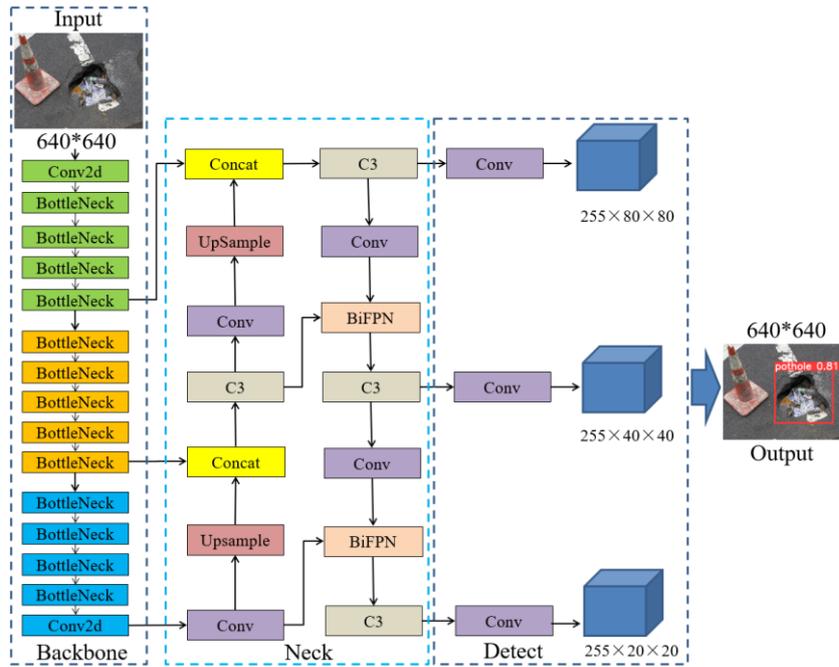


图 4 改进的 YOLOv5s 网络结构

Fig. 4 The improved YOLOv5s network structure

在 Neck 部分，BiFPN 模块的引入进一步优化了模型对不同尺度目标的特征融合能力。BiFPN 通过加权双向特征融合（自顶向下和自底向上的信息流动）有效融合不同尺度的特征信息，增强了模型对多尺度目标的感知能力，尤其对于像道路坑洼这类尺寸差异较大的目标，能显著提升检测效果。

改进后的网络模型如图 4 所示，主干替换为 MobileNetV3 后的网络能够更好地平衡模型性能与效率，而 BiFPN 的加入能够进一步提升多尺度目标的检测精度，减少冗余信息，提高模型的鲁棒性和实时性。

3 道路坑洼数据集的构建

3.1 数据图片收集与标注

目前已有公开的道路坑洼数据集，如同济大学 MIAS Group 提出的 UDTIRI(Urban Digital Twins Intelligent Road Inspections)道路坑洞数据集，该数据集包括 1000 幅在不同光照和天气条件下的道路坑洞图像；公开数据集 RDD2018 包含 9053 张图像，这些图像主要采集自日本、印度和捷克共和国等不同国家的道路；以及 kaggle 竞赛平台上公开的数据集，这些数据集包含道路裂缝和坑洼数据。上述公开数据集中的坑洼比较单一，大多数数据图像都采集于晴天、光照条件好的情况，缺少如雨、雪、雾以及夜间照明不足条件下的图像，对此，为使模型能够在特殊复杂的环境下有更好的识别效果，本研究将使用自建的数据增强后的数据集来进行模型训练。

为了能够使训练后的模型有更高的检测精度和泛化能力，通过实地拍摄和在网络查找挑选，一共挑选了不同大小、深度、光照、形状、遮挡程度的高质量道路坑洼照片 1398 张，如图 5 所示，保证了数据集的多样性。将收集好的图片通过标注工具“labelimg”进行手动标注，构建了多样性的道路坑洼数据集。



图 5 道路坑洼数据集部分图片

Fig. 5 Part of the pothole dataset images

3.2 基于风格迁移进行数据增强

图像风格迁移是一种图像处理方法，旨在将图像的语义内容以不同的风格呈现^[20]。近年来，算法如 DualGAN、CycleGAN、Pix2Pix 和 SSIM-GAN 等使得风格迁移技术逐渐成熟。例如，2018 年，伯克利 BAIR 实验室与 Adobe 合作提出的 Multi-Content GAN 通过仅用少量字母生成大多数未见过的字母，并能够复制原始样本的颜色和纹理。

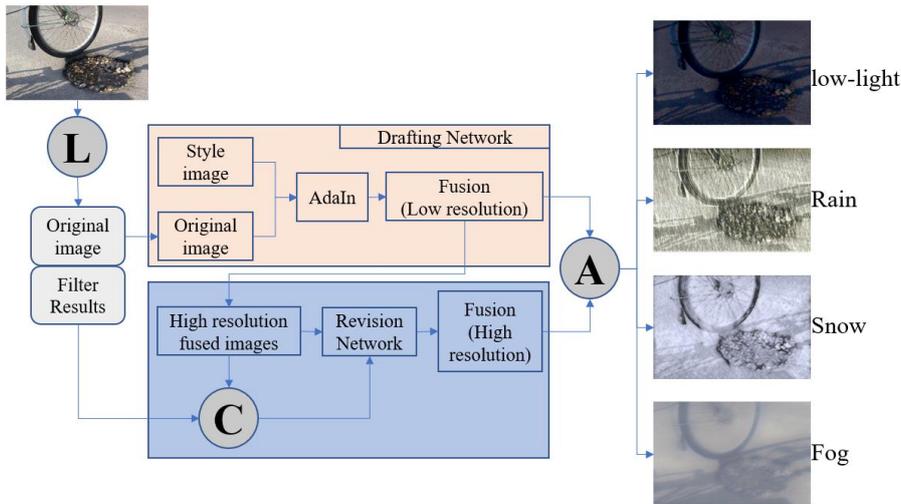


图 6 LapStyle 算法流程

Fig. 6 LapStyle algorithm process

为了能得到复杂情况下的道路坑洼数据集，本文借鉴了生成对抗网络（GAN）的理念，并采用了风格迁移技术，将雾、雨、雪、夜间环境的特征融合到原始图像中，同时保持原始图像的信息不变，从更深层次的图像语义和环境风格上进行增强，扩展数据集的规模，增加更多不同的语义和特征信息，这大大提高了训练模型的泛化能力。

在算法选择方面，采用了 PaddleGAN 中的 LapStyle 算法。该算法能够保持图像语义内容不变的情况下，有效地增强数据集的多样性，并提升模型在各种环境下的泛化能力。

图 6 展示了 LapStyle 算法的流程，其中 L、C 和 A 分别代表拉普拉斯、连结和聚合操作。首先，利用 Drafting Network 传递低分辨率的全球风格图案。该网络在低分辨率模式下融合原始图像和样式图像，并将融合结果在通过上采样后传递到 Revision Network。在修正网络中，使用高分辨率进行连续修正。最后，将修正后的图像与低分辨率的融合图像进行聚

合操作，以生成具有真实风格融合的图片。

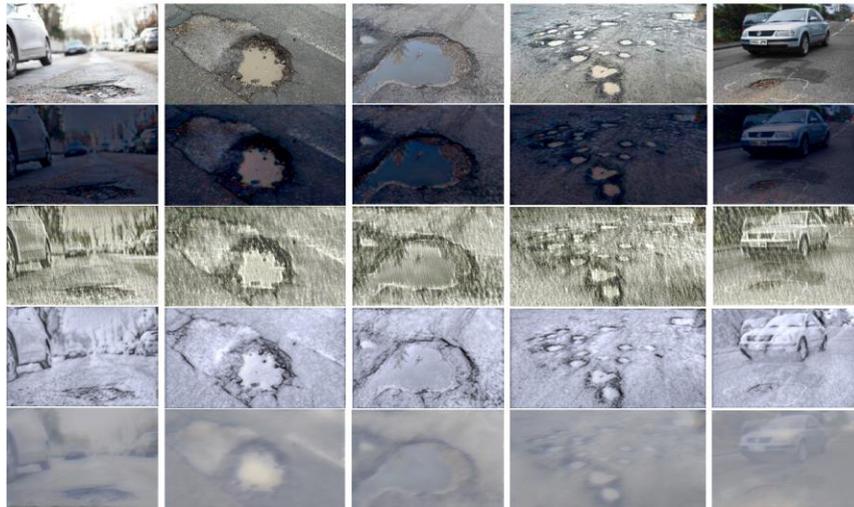


图 7 经 LapStyle 增强后的数据集部分图片

Fig. 7 Some images from the dataset enhanced by LapStyle

将前面提到的 1398 张图片都使用 LapStyle 算法后，得到共计 6990 张复杂环境下的道路坑洼数据集，数据集中的部分图片如图 7 所示。最后将数据集划分为训练集、验证集和测试集。其中训练集 6290 张，验证集 600 张，测试集 100 张。

3.3 数据增强效果验证

为验证经过 3.2 节中使用风格迁移进行数据增强的效果，使用本文改进的网络模型分别在数据增强前的 1398 张图片上与改进后的 6990 张图片上进行训练，模型训练结果如表 1 所示：

表 1 数据增强前后模型训练结果对比

Table 1 Comparative analysis of model training results before and after data augmentation

数据集	P	mAP@0.5
增强前数据集	0.737	0.759
增强后数据集	0.798	0.763

从表 1 内容可以看出，进行数据增强后的数据集，训练出的网络模型精确率与 mAP@0.5 的值均更高，数据增强后，精确率提升了 6.1%。图 8 与图 9 显示了使用数据增强前后训练的结果进行泛化能力测试结果，从图中可以看出，数据增强后训练出的模型在面对复杂环境，如夜间、雪天、雨天、雾天时，检测效果更好，模型的泛化能力有所增强。

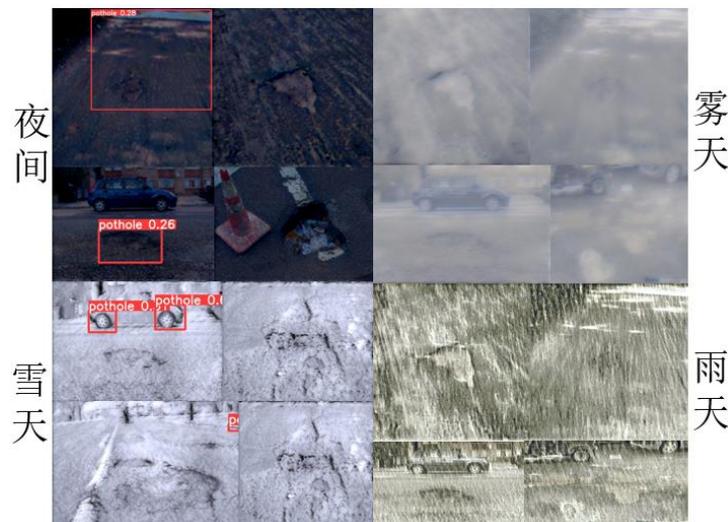


图 8 原始图像数据集训练结果检测效果

Fig. 8 Detection performance of training results on the original image dataset

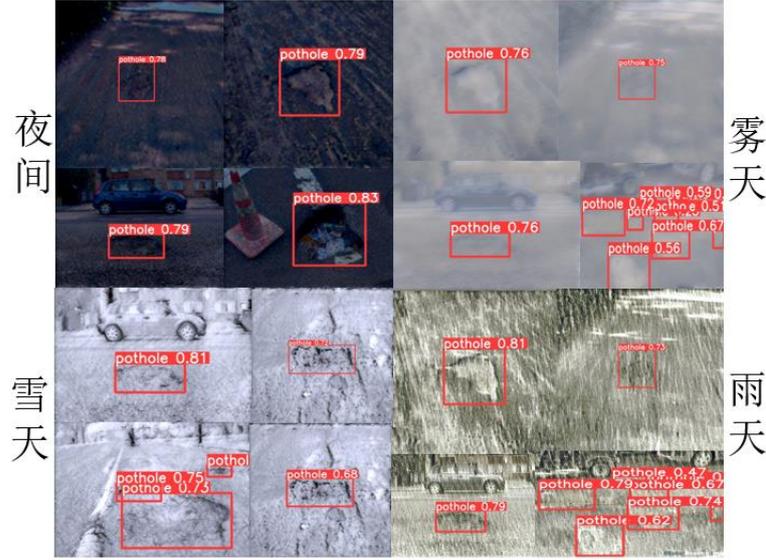


图 9 数据增强后数据集训练结果检测效果

Fig. 9 Detection performance of training results on the augmented dataset

4 实验与结果分析

4.1 实验环境及其配置

为评估改进模型的性能，分别进行了消融实验与对比实验，其中，本研究使用的算法模型超参数设置与其他配置如表 2 所示：

表 2 实验环境配置与基本超参数设置

Table 2 Experimental setup and basic hyperparameter configuration

名称	参数
操作系统	Ubuntu 18.04
GPU	NVIDIA RTX 2080(8GB)*1
CPU	Intel® Xeon® CPU E5-2620 v4 @ 2.10GHz
Python	Python 3.9
深度学习框架	PyTorch 1.12.0
输入图像尺寸	640x640
训练周期	200
初始学习率	0.01
优化器	Adam
Bath-Size	16

4.2 模型评估

本文根据精确率 P (式 3)、召回率 R (式 4)、以及 0.5 的置信度阈值的平均精度(mAP50)来评估模型。这些指标对于确定目标识别模型的有效性及其准确识别对象的能力至关重要^[21]。F1 得分通常用于评估分类任务的性能，但在本文中，目标种类只有 1 类，因此， P 被用作主要评估指标。

$$P = \frac{TP}{TP + FP} = \frac{TP}{All\ Detections} \quad (3)$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{All\ Ground\ Truths} \quad (4)$$

其中, TP 代表真正例, 表示模型做出的正确正预测的数量; FP 代表假正例, 表示模型做出的错误正预测的数量; FN 代表假负例, 表示模型未正确预测的正实例的数量。

$$AP = \sum_{i=0}^{y-1} [R(i) - R(i+1)]P(i) \quad (5)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (6)$$

其中, n 代表类别的数量; y 代表使用的阈值数量。

P 代表被预测为正的样本的比例, 而 R 代表被预测为真的真正样本的比例, AP 代表不同 R 值下各种 P 值的平均值, 本质上是衡量 PR 曲线下的面积。它使用方程 (5) 计算。 $mAP50$ 值表示在 IoU 阈值为 0.5 时所有类别的平均准确率, 并且经常作为目标检测模型的评估标准。较高的 $mAP50$ 值表示模型的识别和检测能力更强, 从而提升了性能。 mAP 使用方程 (6) 计算。

4.3 对比实验

将本文模型与原 YOLOv5s、YOLOv5n、YOLOv5m、YOLOv3-tiny、YOLOv7-tiny、YOLOv10s、与 Faster-RCNN 模型进行对比实验, 使用相同的参数和数据集, 最后将训练得出的性能指标进行对比。表 3 展示了不同模型的算法性能。

表 3 其他模型与本文方法性能指标

Table 3 Performance metrics comparison between other models and the proposed method

Model	P	mAP@0.5	检测速度/(帧)	GFLOPs	Weight Size/(MB)
YOLOv5s	0.763	0.754	48.3	15.9	13.7
YOLOv5n	0.746	0.741	56.4	4.2	3.71
YOLOv5m	0.78	0.775	39.9	48.2	40.2
YOLOv3-tiny	0.732	0.724	50.3	13.0	17.4
YOLOv7-tiny	0.79	0.732	49.3	13.2	12.3
YOLOv10s	0.759	0.761	50.2	24.4	16.5
Faster-RCNN	0.697	0.655	17	181.2	108
本文方法	0.798	0.763	54.1	7.4	8.66

从表 3 可以看出, YOLOv10s 模型在本数据集上相较于 YOLOv5s 性能和准确率上表现较差。两阶段目标检测模型 Faster-RCNN 在检测精度和检测速度上都明显劣于 YOLO 系列, 并且 Faster-RCNN 的 GFLOPs 明显高于其他一阶段目标检测模型, 不利于部署到边缘计算设备。YOLOv5n 有最高的检测速度, 但损失了检测精度。改进后的模型相较于原始模型, 提高了复杂环境下道路坑洼检测的精度, 其中精确率 P 显著提高了 3.5%, 平均精度 $mAP@0.5$ 提高了 0.9%。相较于原始 YOLOv5, 模型参数量降低了接近一半, 这使得模型更加轻量级, 并减少内存的占用, 降低计算资源需求。同时 GFLOPs 的显著减少, 将会让模型有更高的运行效率, 模型将会在资源受限的设备上更加实用, 权重文件的减少使得模型更容易在存储空间受限的环境中部署。因此, 改进后的模型在提升检测精度的同时, 能更加快速的移植到其他平台上进行实时检测。

4.4 消融实验

为了验证所提出改进模型中使用每个模块的有效性, 本文对两个不同的改进部分进行了消融实验。以 YOLOv5s 作为基准进行对照, 并建立了仅使用 MobileNetV3 替换主干网络的模型 (YOLOv5s+ MobileNetV3)、仅添加了 BiFPN 模块的模型 (YOLOv5s+BiFPN) 和本文同时修改了主干和添加 BiFPN 模块的模型 (YOLOv5s+ MobileNetV3+BiFPN) 进行比较。每种改进方法及其相应的实验数据详见表 4。

表 4 消融实验比较

Table 4 Comparison of Ablation Experiments

模型	准确率 P	召回率 R	mAP50	检测速度 (帧)	耗时/(h)
YOLOv5s	0.763	0.780	0.754	48.3	1.982
YOLOv5s+ MobileNetV3	0.759	0.764	0.749	55.6	1.885
YOLOv5s+BiFPN	0.781	0.774	0.757	46.6	1.873
YOLOv5s+ MobileNetV3+BiFPN (本文方法)	0.798	0.772	0.763	54.1	1.964

根据表 4 所示的内容可以看出，在检测速度方面，最快的为 YOLOv5s+MobileNetV3，但损失了较多的精度。与原始 YOLOv5s 相比，本文方法的检测速度更快，有 5.8 帧/s 的提升。结合 MobileNetV3 和 BiFPN 的 YOLOv5s 模型在准确率、召回率和 mAP50 等指标上表现优于原始 YOLOv5s 模型。引入 MobileNetV3 后，虽然模型的准确率、召回率和 mAP50 稍微下降，但计算效率显著提高，耗时略有减少。通过加入 BiFPN，模型的准确率和 mAP50 得到了进一步提升，尽管召回率略有下降，但整体性能明显优于仅使用 MobileNetV3 或 BiFPN 的单独模型。结合两者的优势，本文同时修改了主干和添加 BiFPN 模块的模型在检测精度和计算效率上达到了较好的平衡，提供了更加准确和高效的复杂环境道路坑洼检测模型。图 10 展示了本文改进的 YOLOv5 模型与原始 YOLOv5s 模型在不同复杂环境下的检测结果。从图中可以看出，相较于原始 YOLOv5s 网络模型，本文所提方法的检测效果更好，有更低的漏检率和误检率以及更高的置信度。

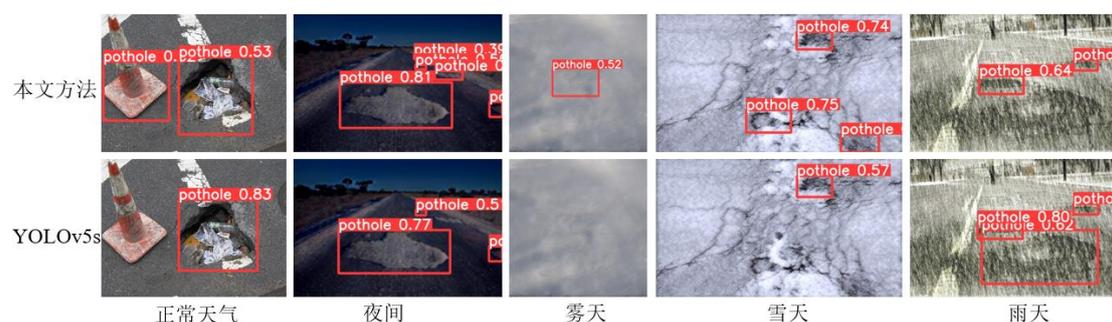


图 10 本文模型与原始 YOLOv5s 模型检测效果对比

Fig. 10 Comparative analysis of detection performance between the proposed model and the original YOLOv5s Model

5 总结

本文针对复杂环境下道路坑洼的检测与定位问题，对 YOLOv5 目标检测算法进行了有效改进。通过引入生成对抗神经网络（GAN）的图像风格迁移技术，解决了公共道路坑洼数据集中缺乏夜间、雨天、雾天和雪天等复杂场景数据的问题，显著增强了数据集的多样性。使用 MobileNetV3 替换原始 YOLOv5s 的骨干部分，这在基本保证模型性能的同时，大大减小了模型的参数量，同时，添加 BiFPN 模块使得 YOLOv5 在多尺度特征融合上具有更大的优势，能够有效提升道路坑洼的检测性能。实验结果证明，改进后的 YOLOv5 算法在自制的增强道路坑洼数据集上表现出更优的检测效果，相较于原始算法，精度提升了 3.5%，mAP50 值提升了 0.9%，检测速度提升了 5.8 帧/s，模型参数减小了一半。在保持检测速度的前提下，大大减少了模型的参数量，有效提高了检测精度。接下来需要继续优化数据集和模型，让其能有更高的精度和更容易在嵌入式设备上运行，将坑洼检测用于车载环视保障泊车与行车安全。

参考文献

- [1] Liu Z, Zhong J, Lyu Y, Liu K, Liu W. Location and fault detection of catenary support components based on deep learning[C]. IEEE 2018 International Instrumentation and

Measurement Technology Conference, 2018.

- [2] 张旭, 陈绪君, 刘瑞康, 张才裕. 基于 OpenVINO 模型优化的智能道路裂缝检测系统[J]. 信息技术, 2020, 44(07): 62-68.
- [3] Guo, G, Zhang, Z. Road damage detection algorithm for improved YOLOv5. *Sci Rep* 12, 15523 (2022).
- [4] 张果荣. 基于深度学习的道路裂缝检测研究[D]. 南京信息工程大学, 2021.
- [5] 于海洋, 景鹏, 张文涛, 等. 基于残差与注意力机制的道路裂缝检测 U-Net 改进模型[J]. 计算机工程, 2023, 49(06):265-273.
- [6] 陈鹏. 基于深度学习的道路坑洼检测[D]. 上海师范大学, 2020.
- [7] H. Hu, Z. Li, Z. He, L. Wang, S. Cao, W. Du, Road surface crack detection method based on improved YOLOv5 and vehicle-mounted images, *Measurement* 229 (Apr. 2024) 114443.
- [8] Zhou Z, Yu X, Iwahori Y, et al. Intelligent Detection of Road Cracks Based on Improved YOLOv5[J]. *Embedded Selforganising Systems*, 2023, 10(7): 9-13.
- [9] 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(01): 66-71.
- [10] Wang Z, Jin L, Wang S, et al. Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system[J]. *Postharvest Biology and Technology*, 2022, 185: 111808.
- [11] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020: 390-391.
- [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*. 2017: 2117-2125.
- [13] Wang W, Xie E, Song X, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[C]//*Proceedings of the IEEE/CVF international conference on computer vision*. 2019: 8440-8449.
- [14] Yin H, Chen M, Fan W, et al. Efficient smoke detection based on YOLO v5s[J]. *Mathematics*, 2022, 10(19): 3493. [15] Xue J, Cheng F, Li Y, et al. Detection of farmland obstacles based on an improved YOLOv5s algorithm by using CIoU and anchor box scale clustering[J]. *Sensors*, 2022, 22(5): 1790.
- [15] Xue J, Cheng F, Li Y, et al. Detection of farmland obstacles based on an improved YOLOv5s algorithm by using CIoU and anchor box scale clustering[J]. *Sensors*, 2022, 22(5): 1790.
- [16] Qian S, Ning C, Hu Y. MobileNetV3 for image classification[C]//*2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. IEEE, 2021: 490-497.
- [17] 孙剑, 张数数. 基于改进 YOLOv5s 的火焰烟雾检测方法[J/OL]. 信阳师范学院学报(自然科学版), 1-7[2024-12-07].
- [18] He L, Wei H, Wang Q. A new target detection method of ferrography wear particle images based on ECAM-YOLOv5-BiFPN network[J]. *Sensors*, 2023, 23(14): 6477.
- [19] Zhang Y, Wang W, Li Z, et al. Development of a cross-scale weighted feature fusion network for hot-rolled steel surface defect detection[J]. *Engineering Applications of Artificial Intelligence*, 2023, 117: 105628.
- [20] 赵文清. 基于风格迁移的图像卡通化算法研究与系统实现[D]. 武汉纺织大学, 2024.

- [21] 刘义艳, 郝婷楠, 贺晨, 等. 基于 DBBR-YOLO 的光伏电池表面缺陷检测[J]. 图学学报, 2024, 45(05):913-921.