

张在岩,宋伟东,邬嘉晨.改进YOLOv5的复杂场景下水泥路面病害检测[J].辽宁工程技术大学学报(自然科学版),2025,44(1):102-112.doi:10.11956/j.issn.1008-0562.20240026

ZHANG Zaiyan, SONG Weidong, WU Jiachen. Disease detection of cement pavement based on improved YOLOv5 in complex scenarios[J]. Journal of Liaoning Technical University(Natural Science), 2025, 44(1): 102-112. doi: 10.11956/j.issn.1008-0562.20240026

# 改进YOLOv5的复杂场景下水泥路面病害检测

张在岩<sup>1,2</sup>, 宋伟东<sup>1\*</sup>, 邬嘉晨<sup>2</sup>

(1. 辽宁工程技术大学 测绘与地理科学学院, 辽宁 阜新 123000; 2. 黑龙江科技大学 矿业工程学院, 黑龙江 哈尔滨 150022)

**摘要:** 针对国内水泥路面病害检测数据集缺乏、规模小、场景单一, 以及深度学习算法在复杂场景下泛化能力不足的问题, 提出一种基于改进YOLOv5的路面病害检测算法。收集并构建包含11 862张图像的水泥路面病害检测数据集, 覆盖9类场景下的3类最常见病害类型; 通过融合以IoU度量的K-Means聚类算法和遗传算法获取模型训练的先验锚框; 在特征增强阶段, 引入轻量级上采样模块(CARAFE), 减少特征重组过程中的信息损失; 引入顾及通道、高度和宽度维度的多维协同注意力模块(MCA), 增强多尺度病害特征的辨别力。实验结果表明: 所提算法在保持较快推理速度的前提下,  $F_1$ 分数和平均精确率(mAP)分别达到75.5%和81.6%, 优于5种主流的目标检测算法。实例分析表明: 基于改进YOLOv5的路面病害检测算法能够满足大规模水泥路面病害智能检测与破损状况评价的实际需求。

**关键词:** 水泥路面; 深度学习; ISTD-PDD3数据集; 病害检测; ISTD-YOLO模型

中图分类号: P258

文献标志码: A

文章编号: 1008-0562(2025)01-0102-11

## Disease detection of cement pavement based on improved YOLOv5 in complex scenarios

ZHANG Zaiyan<sup>1,2</sup>, SONG Weidong<sup>1\*</sup>, WU Jiachen<sup>2</sup>

(1. School of Geomatics, Liaoning Technical University, Fuxin 123000, China; 2. School of Mining Engineering, Heilongjiang University of Science and Technology, Harbin 150022, China)

**Abstract:** Aiming at the problems of lack of domestic cement pavement disease detection data sets, small scale, single scene, and insufficient generalization ability of deep learning algorithms in complex scenes, a pavement disease detection algorithm based on improved YOLOv5 is proposed. A cement pavement disease detection dataset containing 11 862 images was collected and constructed, covering 3 most common disease types in 9 scenarios. The prior anchor frame of model training is obtained by combining the K-Means clustering algorithm measured by IoU and genetic algorithm. In the feature enhancement stage, a lightweight upsampling module (CARAFE) is introduced to reduce the information loss in the feature recombination process. A multi-dimensional collaborative attention module (MCA) considering channel, height and width dimensions is introduced to enhance the discrimination of multi-scale disease features. The experimental results show that under the premise of maintaining fast inference speed, the average of  $F_1$  score and average precision (mAP) of the proposed algorithm reach 75.5% and 81.6%, respectively, which are better than the five mainstream target detection algorithms. The example analysis shows that the pavement disease detection algorithm based on improved YOLOv5 can meet the actual needs of large-scale cement pavement disease intelligent detection and damage condition evaluation.

**Key words:** cement pavement; deep learning; ISTD-PDD3 data set; disease detection; ISTD-YOLO model

收稿日期: 2024-01-25 修回日期: 2024-04-19 接受日期: 2024-04-24 责任编辑: 焦丽

基金项目: 国家自然科学基金项目(42071343); 黑龙江省省属本科高校基本科研业务费(2020-KYYWF-0690)

作者简介: 张在岩(1987-), 男, 山东菏泽人, 博士研究生, 主要从事遥感影像信息智能识别与提取方面的研究。E-mail: zzy\_258911@163.com

通信作者: 宋伟东(1965-), 男, 内蒙古赤峰人, 博士, 教授, 主要从事遥感影像信息提取与三维重建方面的研究。E-mail: lntu\_sw\_d@163.com

## 0 引言

水泥混凝土路面作为路网的重要组成部分之一, 具有强度高、稳定性好、使用寿命长、维护成本低等特点。实际运营过程中, 受各种载荷和自然因素的影响, 各类路面病害逐渐产生, 并严重影响路面的功能性和安全性<sup>[1]</sup>。因此, 定期开展路面状况调查、病害检测和修复, 对延长道路使用寿命至关重要。然而, 现有结合路面图像的人工视觉检测方法存在检测周期长、花费高、主观性强等问题, 难以满足快速、大范围路面技术状况自动检测与评价的需求<sup>[2]</sup>。

早期, 路面病害自动检测方法主要采用基于数字图像处理(如阈值分割<sup>[3]</sup>、边缘检测<sup>[4]</sup>、形态学<sup>[5]</sup>、最小路径选择<sup>[6]</sup>、区域生长<sup>[7]</sup>等)和机器学习技术(如支持向量机<sup>[8]</sup>(SVM)、浅层人工神经网络<sup>[9]</sup>(ANN)等)。此类算法依赖预先设立的特征, 但受到多源噪声影响, 容易导致漏检或误检<sup>[10]</sup>。随着人工智能算法的快速发展, 基于深度学习模型的路面破损检测成为研究的焦点<sup>[11]</sup>。基于深度学习的目标检测算法可以分为两大类, 第一类为包括 R-CNN 算法、SPP-Net 算法、Fast R-CNN 算法、Faster R-CNN 算法等两阶段方法, 第二类为以 YOLO 算法、SSD 算法等为代表的单阶段方法<sup>[12]</sup>。两阶段检测器在小目标检测中表现良好, 具有较高的平均精度(mAP), 但计算速度慢。单阶段方法可以在快速检测的同时准确检测和定位小目标, 因此被广泛采用<sup>[13]</sup>。近年来, 国内外研究者基于不同深度学习算法实现了路面病害的自动检测。文献[14]采用 Faster R-CNN 算法检测沥青路面的多种裂缝类型, 包括坑槽和龟裂, 但推理速度较慢。文献[15]将深度可分离卷积和 Focal loss 应用于 YOLOv4 算法, 并使用了迁移学习的训练策略, 有效提高了沥青路面多类型病害检测的速度和准确率。文献[16]基于全球道路损伤大赛提供的数据集 RDD 算法, 比较了 YOLOv5 算法、CenterNet 算法和 EfficientDet 算法在路面病害检测性能上的差异, 结果表明 YOLOv5 算法综合性能最佳, 但在横向裂缝检测方面精度较低。

目前, 路面病害检测的研究主要集中在沥青路面, 而对水泥路面的病害检测研究相对较少<sup>[17]</sup>。原因如下: 首先, 缺乏足够的水泥路面病害数据集, 这制约了最有效算法的研发; 其次, 复杂场景下水泥公路路面图像中的病害有 3 个显著特点:

图像质量受光照强度、路面干燥度、阴影及其他干扰噪声影响较大; 防滑槽、伸缩缝与裂缝具有高度的灰度相似性, 易产生假阳性预测; 不同病害拓扑结构更复杂, 灰度特征差异较小<sup>[18]</sup>。在图像识别中, 路面颜色差异将直接影响病害检测算法的准确性<sup>[19]</sup>。因此, 在缺乏大规模、多类型、复杂场景的水泥路面病害数据集的情况下, 病害检测具有极大挑战性。

综上, 筛选并构建专注于复杂场景下的水泥多类型路面病害目标检测任务的数据集 ISTD-PDD3, 以此为数据基础, 研究先验锚框的优化选择策略。基于易于部署的 YOLOv5 框架, 引入顾及输入特征的轻量上采样算法和多维协同注意力机制模块, 设计平衡检测精度和效率的多类型路面病害检测网络(ISTD-YOLO 模型), 实现复杂场景下路面病害的检测与破损状况自动评价。

## 1 数据集构建与分析

### 1.1 数据集构建

为构建具有复杂场景和显著样本差异性的数据集, 依托辽宁工程技术大学交通时空大数据研究中心(institute of spatiotemporal transportation data, ISTD), 采用辽宁省多年份路面图像数据库。通过专家筛选, 从该数据库中挑选出 11 862 张水泥路面病害图像, 构建 ISTD-PDD3 数据集。数据集图像具有不同的清晰度、光照强度、路面干燥程度, 以及包含各种干扰物的情况。此处干扰物包括伸缩缝、防滑槽、阴影、污渍和路面杂物等元素, 其目的是提高模型的鲁棒性。病害类型包含常见的水泥裂缝(SC)、破碎板(PSB)和坑洞(KD)。样本图像如图 1 所示。使用 LabelImg 进行精准的标注, 并保存为 Pascal VOC 的数据格式。新数据集按 8 : 1 : 1 的比例划分训练集、验证集和测试集。无病害的负样本不进行标注处理。样本图像见图 1, 标注样例见图 2。

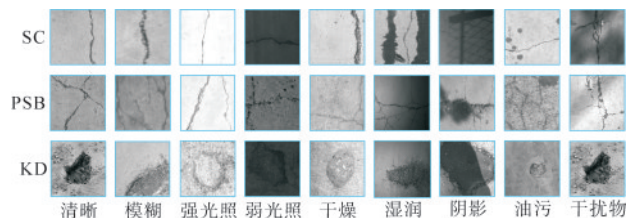


图 1 复杂场景下水泥路面病害样本图像

Fig.1 sample image of cement pavement disease in complex scene

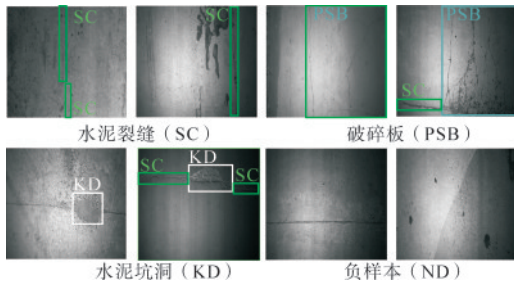


图 2 ISTD-PDD3 数据集标注样例

Fig.2 sample ISTD-PDD3 data set annotation

### 1.2 数据集对比分析

变异系数 (coefficient of variation, CV) 可反映不同数据集样本总体的离散程度, 选择图像像素灰度的平均值、标准差和信息熵 (information entropy, IE) 3 个度量指标的变异系数分析不同数据集图像场景的复杂度<sup>[20]</sup>。图像像素灰度平均值可定量反映图像的平均亮度; 图像像素灰度的标准

差可反映图像像素灰度相对于灰度均值的离散程度, 也可反映图像对比度的强弱; IE 从信息论的角度衡量图像中灰度分布的聚集特征所包含的平均信息量。

将 ISTD-PDD3 数据集与 3 个公开数据集进行比较分析。这 3 个公开数据集包括 2 个基于前方景观图像的数据集 CNRDD<sup>[21]</sup>和 RDD<sup>[16]</sup>, 1 个基于大视角的谷歌街景图像数据集 PID<sup>[22]</sup>。不同数据集的定性比较如图 3 所示。不同数据集 3 项对比指标的变异系数见图 4。由图 3 可知, 与其他 3 个数据集相比, ISTD-PDD3 数据集中的样本图像更关注路面场景, 且为经过正射校正后的图像, 更符合国内道路养护部门进行路面破损状况定量评价的业务场景。由图 4 可知, ISTD-PDD3 数据集在三项指标的上的变异系数均最高, 说明其样本复杂度和离散度最大。

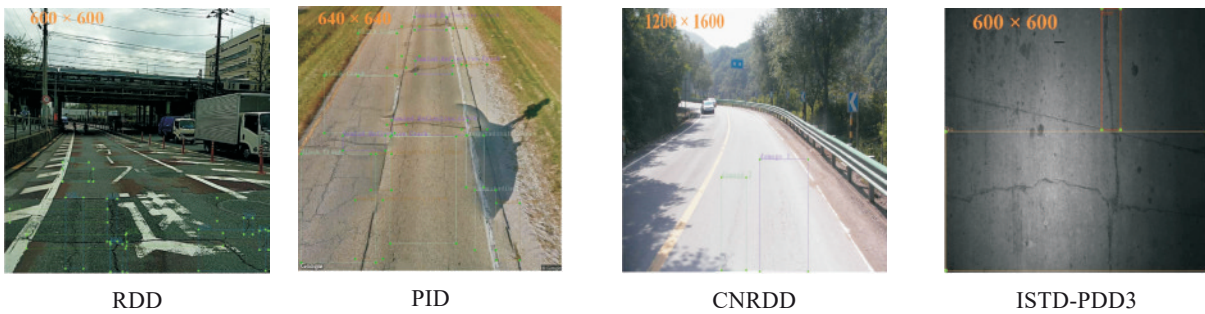


图 3 不同数据集定性比较

Fig.3 qualitative comparison of different data sets

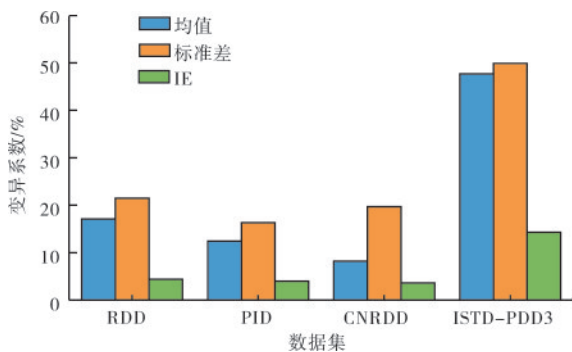


图 4 不同数据集 3 项对比指标的变异系数

Fig.4 variation coefficients of 3 comparison indicators in different data sets

### 1.3 先验锚框优化选择

对于 YOLO 系列的单阶段目标检测算法, 先验锚框的作用类似于网络的初始化权重, 样本先

验锚框是否合适将直接影响模型的准确度。YOLOv5 采用基于欧式距离 (euclidean distance, ED) 的 K-means 算法聚类训练集的先验锚框, 与较小的标注框相比, 这种度量样本与聚类中心的方法会导致较大的标注框产生更多的误差<sup>[23]</sup>。为此, 本文以先验锚框与标签样本框间的交并比 (IoU) 作为度量指标, 即以  $d(\text{bbox}, \text{anchor}) = 1 - \text{IoU}(\text{bbox}, \text{anchor})$  代替欧式距离。改进的 K-means 聚类算法获取 9 先验锚框后使用遗传算法 (genetic algorithm, GA) 优化锚框的宽、高, 最后获取与训练数据集具有最高适应度的先验锚框。表 1 给出了基于不同样本距离聚类及是否采用遗传算法的先验锚框计算结果。可以看出, 采用 “K-means(1-IoU)+GA” 方法生成的锚框与 ISTD-PDD3 数据集适应度最佳。

表 1 不同锚框聚类方法结果

Tab.1 results of different anchor frame clustering methods

方法	锚框	适应度
K-means(ED)	[39,83] [37,202] [42,331] [44,476] [235,109] [49,629] [584,151] [319,600] [610,597]	0.646 6
K-means(ED)+ GA	[16,127] [49,89] [24,288] [213,69] [30,601] [58,554] [475,166] [207,556] [611,615]	0.701 8
K-means(1-IoU)	[14,73] [20,194] [94,88] [24,425] [40,637] [82,629] [553,103] [264,637] [637,637]	0.716 7
K-means(1-IoU)+ GA	[16,85] [22,232] [70,100] [24,554] [39,590] [69,556] [448,101] [235,515] [613,613]	0.728 3

## 2 路面病害检测算法

### 2.1 ISTD-YOLO概述

ISTD-YOLO 网络结构见图 5。ISTD-YOLO 由输入端、主干特征提取网络、特征增强网络和检测头组成。输入端使用 1.3 节获取的最优先验锚框，并将输入图像的分辨率大小统一转换为 640×640。保留原始的骨干特征提取网络，以更好地使用迁移学习来训练模型，提取不同尺度的病害特

征信息。在特征增强网络中，使用顾及输入特征的轻量上采样模块 CARAFE<sup>[24-25]</sup>替换原始的最近邻上采样，以增强特征重组时病害信息的恢复能力；引入多维协同注意力模块 MCA<sup>[26-27]</sup>，提升多尺度路面病害的检测性能。最后输出特征图的分辨率分别为 20×20、40×40、80×80，并馈送至检测头，经过非极大似然抑制，输出包含病害名称和概率的边界框。

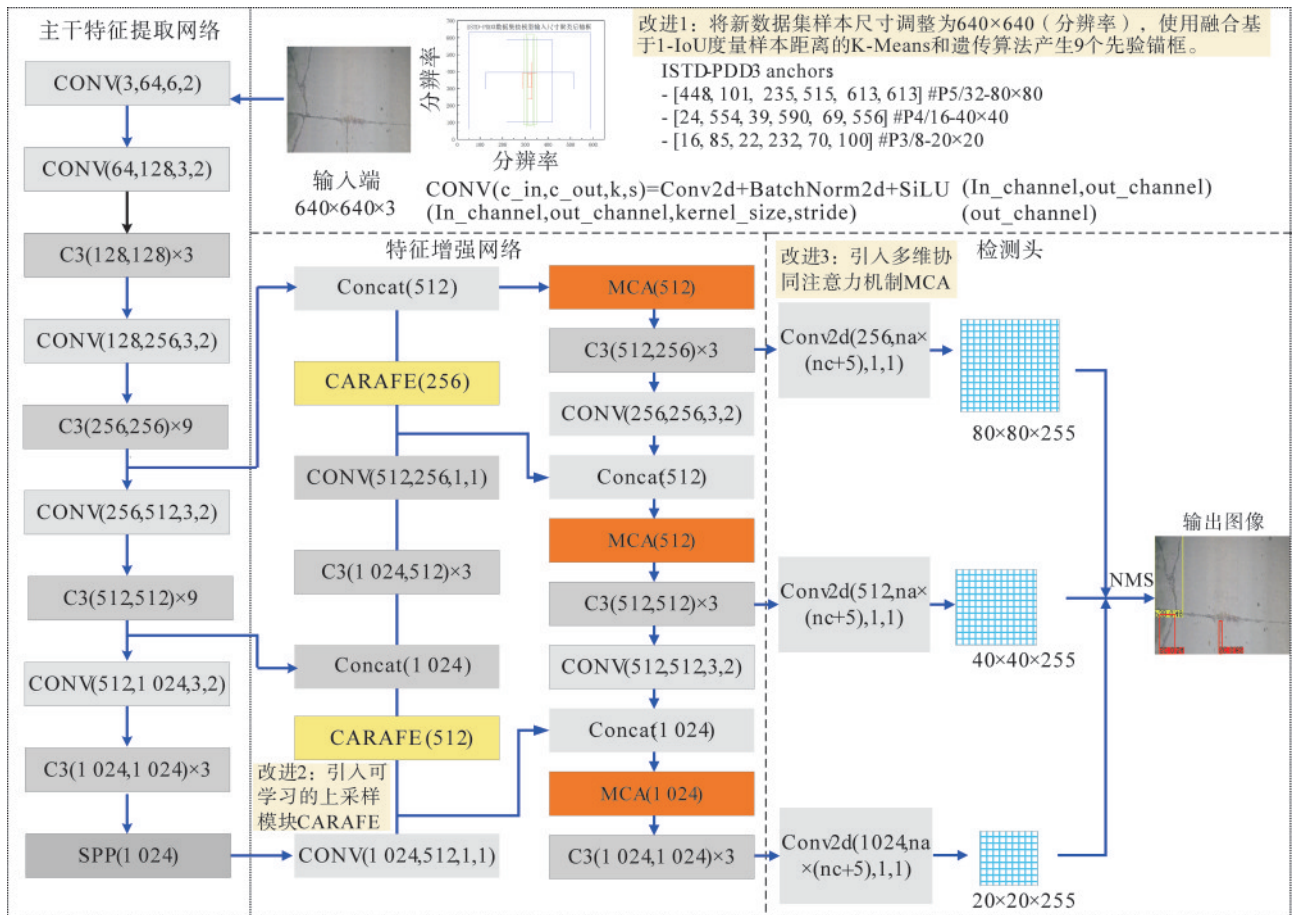


图 5 ISTD-YOLO 网络结构

Fig.5 network architecture of the ISTD-YOLO

### 2.2 顾及输入特征的轻量上采样模块

原始 YOLOv5 模型在特征增强阶段使用最近邻算法进行特征上采样，该算法未利用特征图的

语义信息进行特征重组，且感受野较小 (1×1)。为此，在特征增强阶段引入具有较大感受野且能顾及输入特征图像语义信息的轻量级上采样模块

CARAFE<sup>[24]</sup>, CARAFE上采样算子见图6。假定输入特征图  $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$ , 设上采样倍率为  $\sigma$ , CARAFE首先根据每个目标位置的内容预测一个重组核, 进而依据预测的核对输入特征进行重组, 最终生成大小为  $\mathbf{X}' \in \mathbf{R}^{C \times \sigma H \times \sigma W}$  的新特征图。

如图6所示, 内核预测模块负责以内容感知的方式生成重组内核。 $\mathbf{X}$ 上的每个位置对应于 $\mathbf{X}'$ 上大小为 $\sigma^2$ 的目标区域。每个目标区域都需要一个大小为 $k_{up} \times k_{up}$ 重组内核。首先使用一个 $1 \times 1$ 的卷积, 将 $\mathbf{X}$ 的通道压缩至 $C_m$  ( $C_m=64$ ), 以减少后续计算量; 进而采用 $k_{encoder} \times k_{encoder}$  ( $k_{encoder}=3$ )的卷积

层预测上采样核, 输出尺寸为 $H \times W \times \sigma^2 k_{up}^2$  ( $k_{up}=5$ )的特征图, 并将其通道维在空间维展开生成一个形状为 $\sigma H \times \sigma W \times k_{up}^2$ 的上采样核, 最终采用softmax函数进行归一化处理, 使其权重和为1。

特征重组模块首先将输出特征图中的每个位置映射回原输入特征图, 并取出以之为中心的 $k_{up} \times k_{up}$ 的区域, 进而与预测出的该点的上采样核, 并作点积运算, 得到最终输出值。特征重组过程中相同位置的不同通道将共享同一个上采样核。重组后的特征图的语义比原始特征图的语义更强, 因为来自局部区域中相关点的信息可以得到更多关注。

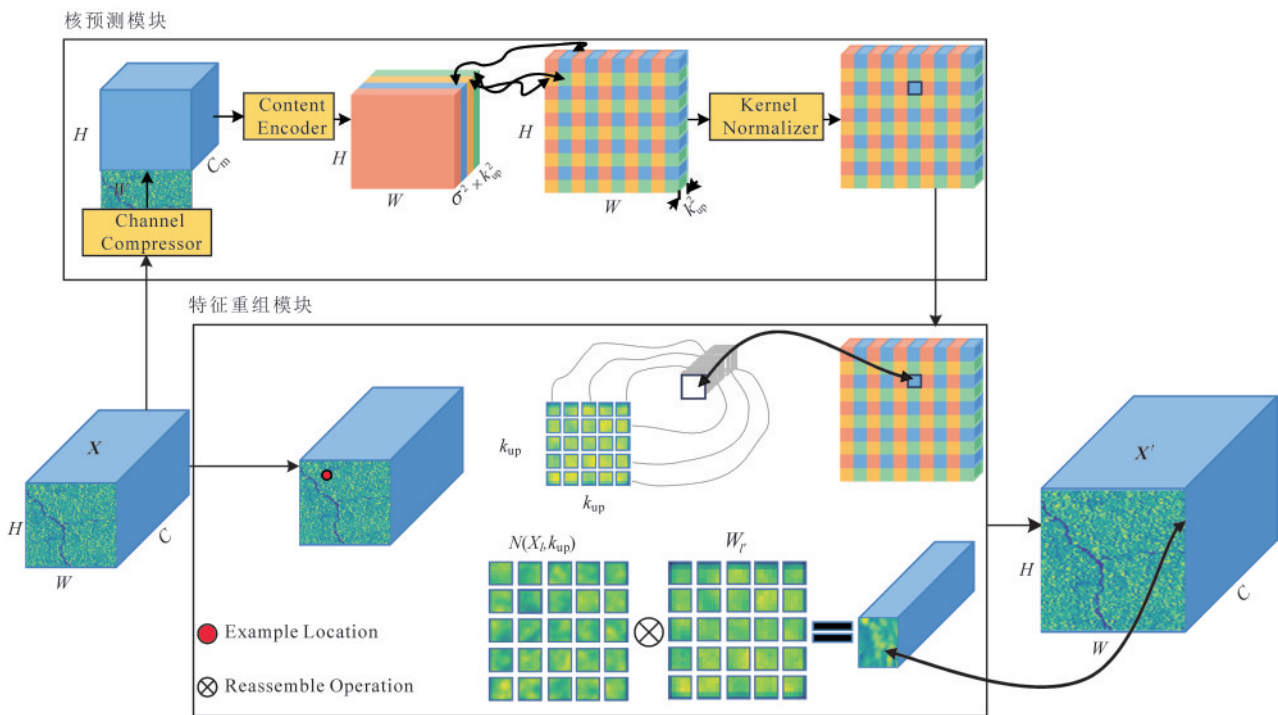


图6 CARAFE上采样算子  
Fig.6 CARAFE upsampling operator

### 2.3 轻量级多维协同注意力机制模块

注意机制可以在提高目标特征表达的同时, 抑制无用特征的表达, 从而提升目标检测的性能<sup>[25]</sup>。最近的研究表明, 对于YOLO系列的模型, 将注意力机制模块应用于特征增强阶段最能提升模型的目标检测精度<sup>[26]</sup>。考虑到模型的计算成本和不同病害类型的多尺度特征, 本文拟在基线模型的特征增强阶段引入轻量级的多维协同注意力机制模块MCA, 它不是单独的计算空间和通道注意, 而是强调捕获跨域交互的重要性, 从而丰富区域特征表示<sup>[27]</sup>。MCA模块由3个并行分支组成, 其

结构见图7, 其中分支1、2分别负责捕获空间维度中的特征相互依赖性, 分支3用于捕获通道间的交互。

具体而言, 对输入的特征向量  $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$ , 首先将其传递给多维协同注意力模块中的3个分支。在分支1中,  $\mathbf{X}$ 沿 $H$ 轴被逆时针旋转 $90^\circ$ , 得到  $\hat{\mathbf{X}}_w \in \mathbf{R}^{W \times H \times C}$ 。然后, 为构建高度维度 $H$ 和通道维度 $C$ 之间的交互, 对 $\hat{\mathbf{X}}_w$ 挤压变换得到特征图  $\hat{\mathbf{X}}_w \in \mathbf{R}^{W \times 1 \times 1}$ , 再通过激励变换捕获通道维度 $C$ 和空间维度 $W$ 之间的交互关系, 得到 $W$ 方向上的特征权重  $\tilde{\mathbf{X}}_w \in \mathbf{R}^{1 \times 1 \times W}$ 。接下来, 在 $W$ 维度上, 通过对 $\tilde{\mathbf{X}}_w$ 的矩阵转换和sigmoid函数生成特定的注意力权

重  $\lambda_w \in \mathbf{R}^{W \times 1 \times 1}$ ; 将  $\lambda_w$  与  $\hat{X}_w$  逐元素相乘得到具有空间维度  $W$  与通道维度  $C$  交互关系的特征映射矩阵  $X'_w \in \mathbf{R}^{W \times H \times C}$ ; 最后, 将  $X'_w$  沿  $H$  轴顺时针旋转  $90^\circ$ , 得到与原始输入相同形状的特征图  $X''_w \in \mathbf{R}^{C \times H \times W}$ 。

在分支 2 中, 输入特征向量  $X$  沿  $W$  轴被逆时针旋转  $90^\circ$ , 得到  $\hat{X}_H \in \mathbf{R}^{H \times C \times W}$ 。之后, 类似于分支 1 中的操作, 实现空间维度  $H$  与通道维度  $C$  交互关系的建模, 最终生成一个原始输入相同形状的特征  $X''_H \in \mathbf{R}^{C \times H \times W}$ 。

分支 3 包含通道注意力机制, 主要负责从通道的角度赋予图像不同位置不同的权重, 以获得更重要的特征信息。首先, 通过恒等映射 (identity mapping) 生成与  $X$  相同的特征图  $\hat{X}_C \in \mathbf{R}^{C \times H \times W}$ 。然后, 将  $\hat{X}_C$  依次输入到挤压及激励模块中, 获得通道特征权重  $\lambda_C \in \mathbf{R}^{C \times 1 \times 1}$ ; 接着对特征图  $X$  进行赋值, 获得特征映射图  $X''_C \in \mathbf{R}^{C \times H \times W}$ 。最后, 对不同维度的注意力权重做平均聚合, 并重新校准三个分支的所有输出, 即可推导出更加细化的病害信息特征图  $X''' \in \mathbf{R}^{H \times W \times C}$ , 其表达式为

$$X''' = \frac{1}{3} \otimes (X''_w \oplus X''_H \oplus X''_C), \quad (1)$$

式中:  $\otimes$  表示两特征图逐点相乘;  $\oplus$  表示两特征图逐点相加。

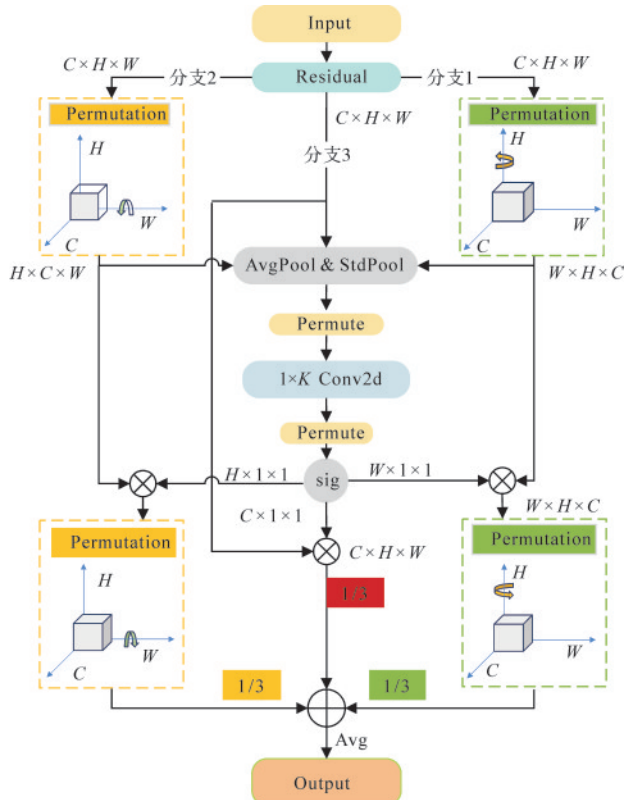


图 7 MCA 结构

Fig.7 architecture of MCA

## 3 实验与分析

### 3.1 实验细节设置

(1) 实验平台与训练参数。实验采用 Window10 操作系统, Intel(R) Core(TM) i7-8700 CPU 核心处理器, NVIDIA GTX 2070 图形处理器。程序开发框架为 Python3.8.5、CUDA11.0.221, Pytorch1.7.1。迁移学习采用各模型在 COCO 数据集上的预训练权重。训练过程采用 Mosaic 数据增强对输入图像进行处理, 使用 SGD 优化器对模型进行优化, 设学习率 Init\_lr 为  $1 \times 10^{-2}$ , 输入图像分辨率为  $640 \times 640$ , 模型训练批处理量为 4, 学习率为  $1 \times 10^{-4}$ , 训练 Epoch 次数为 500, 动量和权重衰减分别为 0.937 和  $5 \times 10^{-4}$ 。

(2) 对比算法。本文将 ISTD-YOLO 模型与 5 种主流目标检测模型进行对比, 包括基于候选区域的 Faster R-CNN 模型<sup>[28]</sup>, 基于边框回归的 SSD 模型<sup>[29]</sup>、基于中心点回归的无锚框 CenterNet 模型<sup>[30]</sup>、单阶段检测的 YOLOv5 模型与 YOLOv7 模型<sup>[31]</sup>。所有模型使用相同的训练策略, 并在 ISTD-PDD3 的训练和测试集上进行模型训练和性能评估。

(3) 评估指标。为评估 ISTD-YOLO 模型在路面病害检测中的性能, 采用  $F_1$  得分、评价精度 AP (average precision)、mAP 作为定量评价指标。

$F_1$  得分是精确率  $P$  和召回率  $R$  的调和平均值, 可以衡量模型的综合性能, 可表示为

$$F_1 = \frac{2PR}{(P+R)}. \quad (2)$$

平均精度 AP 是通过每个类别的  $P$ - $R$  (精确率-召回率) 曲线进行积分得到, 可表示为

$$AP = \int_0^1 P(R) dR. \quad (3)$$

在多类型目标检测任务中, 通过计算所有类型的 mAP 来评估模型的检测精度, 如式 (4) 所示,  $C$  取 3。

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i. \quad (4)$$

此外, 通过参数量 Par (单位: MB) 和浮点运算次数 (floating point operations, FLOPs) 表示运算复杂度。参数量大, 导致训练和推理时间延长。以单张图像的平均推理时间 (单位: ms) 衡量推理效率。

### 3.2 消融实验

为揭示所提 3 项改进方法对基准模型的影响,

设计如下消融实验：将先验锚框优选策略（1-IoU）、CARAFE 和 MCA 分别添加到 YOLOv5 模型中进行实验。消融实验评价指标统计见表 2。

表 2 消融实验评价指标统计

Tab.2 statistics of experimental evaluation indexes for ablation

模型	$F_1$ /%	mAP/%	浮点运算次数/ $10^9$	参数量/MB	平均推理时间/ms
YOLOv5	68.2	73.5	48.245	20.879 4	15.6
YOLOv5+(1-IOU)	73.5	77.8	48.245	20.879 4	15.6
YOLOv5+ CARAFE	71.4	77.2	48.257	21.157 2	17.0
YOLOv5+MCA	72.6	78.5	49.319	20.879 5	16.3
ISTD-YOLO	75.5	81.6	49.559	21.231 0	17.3

由表 2 可知，与原始的 YOLOv5 模型相比，仅采用本文所提出的初始锚框选择策略， $F_1$  和 mAP 分别提高了 5.3 个百分点和 4.3 个百分点，说明结合具体任务场景使用 1-IoU 代替欧式距离来获取锚框聚类结果，可提升病害检测性能；此外，仅引入轻量级上采样算子 CARAFE， $F_1$  和 mAP 分别提高了 3.2 个百分点和 3.7 个百分点，证明顾及输入内容和较大感受野的上采样可减少上采样的信息损失，进而

有效提高模型的检测性能；再者，仅在特征增强阶段引入 MCA 注意力机制模块，计算量有微小增加，但  $F_1$  和 mAP 分别提高了 4.4 个百分点和 5.0 个百分点，说明该模块能够有效捕捉病害目标特征中的通道和空间之间的长距离依赖关系，有利于提高预测精度；最后，将 3 项改进同时融入 YOLOv5 模型中，模型的浮点运算次数增加了  $1.354 \times 10^9$ 、参数量增加了 0.351 6 MB，推理速度降低了 1.4 ms，但  $F_1$  和 mAP 分别提高了 7.3 个百分点和 8.1 个百分点，在平衡检测速度的同时，有效提升了检测精度。

基于不同网络的不同病害类别及平均值的  $P-R$  曲线见图 8。由图 8 (a) 可知，对于单一病害的  $P-R$  曲线，ISTD-YOLO 模型的优化先验锚框选择策略对裂缝灾害的检测性能最优。由图 8 (b)、图 8 (c) 可知，单独引入 MAC 注意力机制模块后，破碎板和坑洞的检测性能得到最大提升。综上可见，本文提出的 3 项针对性改进措施显著提高了水泥路面多尺度病害的检测精度，但检测速度仅减慢 1.7 ms。由图 8 (d) 可知，在同一召回率下，融合 3 项改进措施算法的平均路面病害检测精确度优于其他组合算法，证明本文算法的综合检测性能最优。

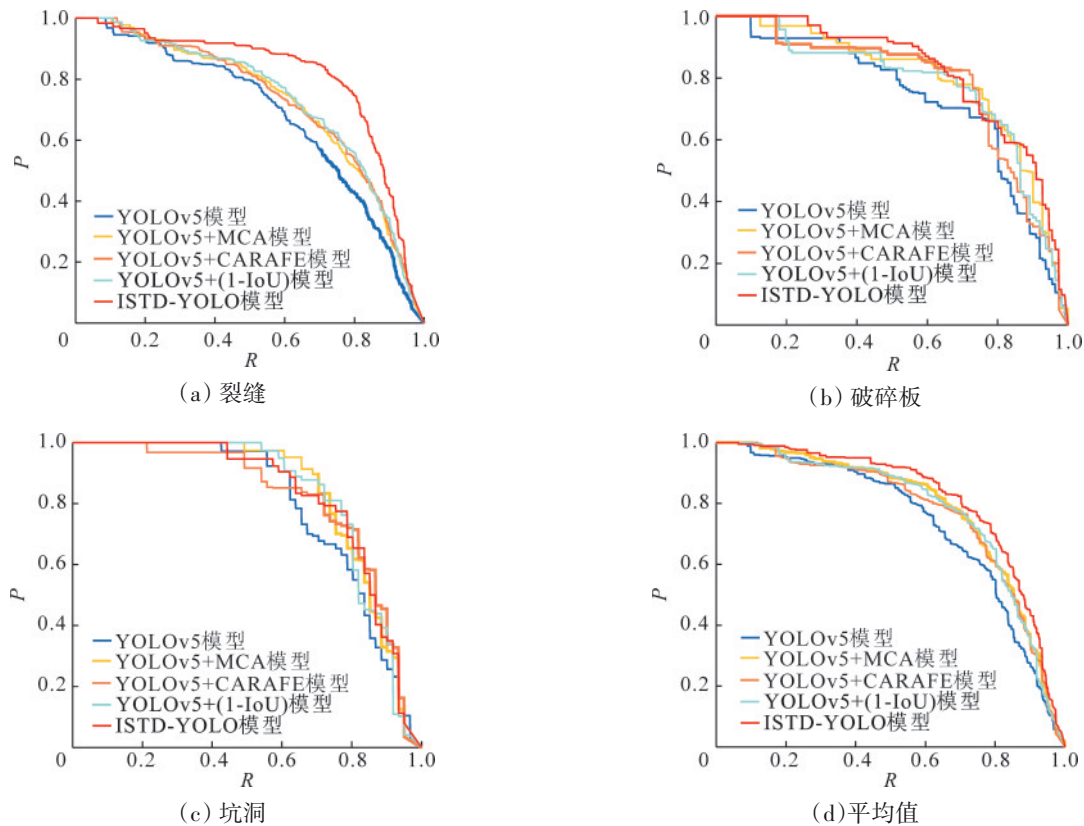


图 8 不同改进方法下 3 类病害及均值的 P-R 曲线

Fig.8 P-R curves of the three diseases and their mean values under different improvement method

### 3.3 不同算法对比实验与分析

表 3 给出了 ISTD-YOLO 模型与 5 种主流目标检测模型在测试集上的定量比较结果。分析可知, 与其他 5 种模型相比, 本文算法在保持较低计算复杂度、参数量和较快检测速度的同时, 取得了最佳的病害检测性能。相较于 Faster R-CNN 模型, ISTD-YOLO 模型的 mAP 提升了 17.1 个百分点, 检测效率提升约 4 倍。SSD 模型舍弃了区域提议

(region proposal) 阶段, 推理速度明显提升, 但其 mAP 仅为 65.3%。CenterNet 模型去除了复杂的 Anchors 和非极大似然抑制操作, 推理速度为 20.5 ms, 但检测精度较低。相较于 YOLOv7 模型, ISTD-YOLO 模型的 mAP 提升了 7.8 个百分点, 单张图像的推理时间降低了 15.3 ms。此外, 仅增加模型的深度会降低细小裂缝的检测性能。对比结果表明, ISTD-YOLO 模型更适合计算资源受限的任务场景。

表 3 不同模型性能对比

Tab.3 performance comparison of different models

模型	不同类别灾害的评价指标/%						属性			
	$F_1$ (SC)	AP(SC)	$F_1$ (PSB)	AP(PSB)	$F_1$ (KD)	AP(KD)	浮点运算次数/ $10^9$	参数量/MB	平均推理时间/ms	mAP/%
Faster R-CNN	39.9	64.7	58.5	73.1	34.6	56.5	948.2	28.3	86.7	64.5
SSD	52.9	63.9	70.6	71.5	59.8	61.3	274.9	23.9	42.4	65.3
CenterNet	48.1	75.2	32.7	61.2	27.6	60.3	109.7	32.7	20.5	65.5
YOLOv5	63.5	68.2	70.6	73.3	70.3	79.1	48.3	20.9	15.6	73.5
YOLOv7	46.8	63.3	61.8	76.9	71.4	81.2	105.2	37.2	32.6	73.8
ISTD-YOLO	76.9	80.9	72.4	81.7	75.9	82.2	49.5	21.2	17.3	81.6

不同的网络模型对 3 类水泥路面病害的检测效果展示见图 9。由图 9 可知, ISTD-YOLO 模型在不同亮度条件、不同病害共存及阴影干扰的场景中, 获得了最佳的检测效果。对比最后两列检

测结果可见, 相比基线网络其他模型, ISTD-YOLO 模型在对灾害的检查中更能顾及细小裂缝的定位和分类。

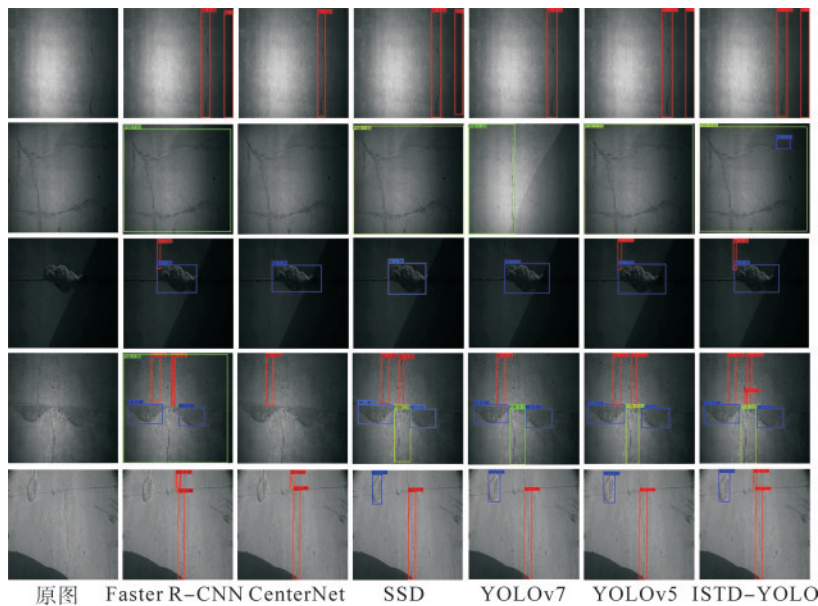


图 9 不同模型的水泥路面病害检测结果

Fig.9 cement pavement diseases detection results of different models

### 3.4 实例分析

(1) 基于 ISTD-YOLO 模型输出结果的路面破损状况指数计算方法。公路路面状况的定量评价

是基于图像自动病害检测手段能够走向实用化的关键<sup>[32]</sup>, 评价指标选用路面破损状况指数 (pavement condition index, PCI), 可表示为

$$PCI = 100 - a_0 \left( 100 \times \sum_{i=1}^n w_i A_i / A \right)^{a_1} \quad (5)$$

式中： $A_i$ 为第*i*类路面病害损坏的累计面积， $m^2$ ； $w_i$ 为第*i*类病害的权重； $A$ 为调查的路面面积的评价单元，取  $100 m^2$  或  $1\ 000 m^2$ ； $a_0$ 、 $a_1$  分别取 10.66、0.461。

由 ISTD-YOLO 模型的预测结果可转换为图像中病害矩形框 (bbox) 左上角坐标  $(x_1, y_1)$ ，右下角坐标  $(x_2, y_2)$ ，各坐标可表示为

$$\begin{cases} x_1 = Wx_c - 0.5Ww_ \\ y_1 = Hy_c - 0.5Hh_ \\ x_2 = Wx_c + 0.5Ww_ \\ y_2 = Hy_c + 0.5Hh_ \end{cases} \quad (6)$$

式中： $x_c$ 为 bbox 中心点坐标  $x_c$  与图像宽  $W$  的比； $y_c$  为 bbox 中心点坐标  $y_c$  与图像高  $H$  的比； $w_$  为 bbox 的  $w$  与  $W$  的比； $h_$  为 bbox 的  $h$  与  $H$  的比，则病害矩形框实际的宽  $W_T$ 、高  $H_T$ 、对角线长度  $L_T$  分别为

$$\begin{cases} W_T = 3.2 \times \sqrt{(x_1 - x_2)^2} / W \\ H_T = 2.0 \times \sqrt{(y_1 - y_2)^2} / H \\ L_T = \sqrt{W_T^2 + H_T^2} \end{cases} \quad (7)$$

实验所使用的路面图像经过正射校正后的  $W \times H$  为  $3\ 517 \times 2\ 193$ ，对应实际路面的宽和高分别为  $3.2 m$

和  $2.0 m$ ；根据式 (6) 和式 (7) 可得单张图像上第 *i* 类病害矩形框的面积为

$$S_{Ti} = W_{Ti} H_{Ti},$$

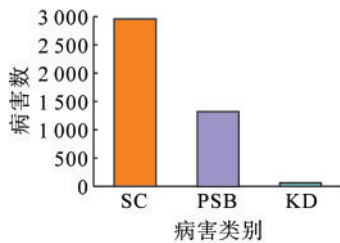
对角线长度为

$$L_{Ti} = \sqrt{W_{Ti}^2 + H_{Ti}^2},$$

式中， $W_{Ti}$ 、 $H_{Ti}$  分别为单张图像中第 *i* 类病害的检测框对应的实际宽和高， $m$ 。

进而可计算各评价单元的每类病害的总损坏面积  $A_i$ ，并由式 (6) 计算评价单元的 PCI 指标。

(2) 综合应用性能分析。为验证 ISTD-YOLO 模型在工程中的应用价值，选取湖北省十堰市、天门市 4 条不同路面破损程度的水泥公路，共 116 个评价单元 (评价单元面积为  $100 m^2$ )，5 800 张图像，3 类病害的数量统计及采集现场见图 10。分别由 3 位路面检测养护领域的专家通过目视判别的方式对各评价单元的路面破损状况评分，并统计耗时。使用 ISTD-YOLO 模型完成对路面病害的自动检测及 PCI 计算，并统计模型计算耗时。以检测的准确率、漏检率、误检率和耗时评估 ISTD-YOLO 模型的综合检测性能。准确率为检测正确病害数和实际病害数的比；漏检率为未检出病害数与实际病害数的比；误检率为错误检测的病害数和总检测病害数的比。以 PCI 差值绝对值的均值、中误差评价单元占比评估本文算法在路面破损状况评价方法的应用价值。



(a) 病害数量统计

(b) 采集现场

图 10 验证数据相关信息

Fig.10 verify data related information

ISTD-YOLO 模型在 116 个评价单元的综合性能展示见图 11。对于图像中面积占比较大的 PSD 和 KD 病害，ISTD-YOLO 模型的检测准确率达到 93.32%，而对 SC 病害的检测中出现了较高的漏检率和误检率。观察检测结果发现，漏检现象主要出现在具有细小裂缝的图像中，误检现象则多发生在含有边缘不明显的伸缩缝图像中。ISTD-YOLO 模型的平均漏检率为 6.91%，误检率为 0.60%。与人工目视判别法相比，ISTD-YOLO 模型的计算耗时仅为 13.3 min，而人工方法的计算耗时为 464 min，对比可见，ISTD-YOLO 模型的计算

效率提高了约 35 倍。

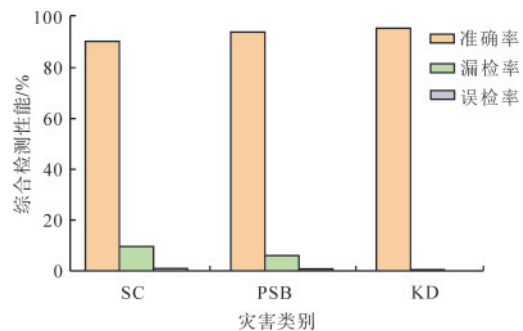


图 11 病害检测性能统计

Fig.11 disease detection performance statistics

ISTD-YOLO 模型与人工目视判别在各评价单元上得到的 PCI 对比结果见图 12。由图 12 可见, 两种方法所得的 PCI 较为吻合, 其中差值绝对值最小为 0.01, 最大为 6.19, 差值绝对值均值为 1.65, 中误差为 1.47。这表明在大规模水泥公路路面破损状况自动评价方面, 本文提出的 ISTD-YOLO 模型有一定的工程应用价值。

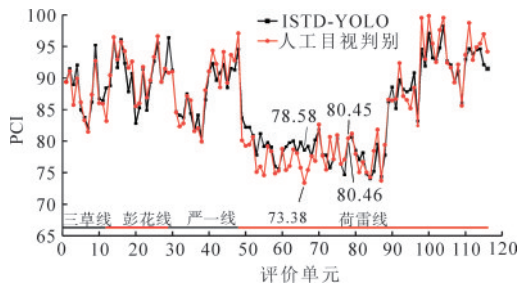


图 12 ISTD-YOLO 与人工目视判别所得 PCI 对比

Fig.12 comparison of PCI obtained by ISTD-YOLO and manual visual discrimination

## 4 结论

本文提出一种改进的 YOLOv5 模型用于复杂场景下水泥路面病害检测, 实现路面破损状况指标的鲁棒性计算, 并获得如下结论。

(1) 基于辽宁省实测路面图像数据库, 创建了 ISTD-PDD3 数据集, 用于多类型水泥公路路面病害目标检测。解决了国内水泥混凝土路面病害智能检测数据集规模小、场景少的问题, 推动了智能检测算法的工程化应用。

(2) 采用交并比度量的 K-Means 聚类算法与遗传算法对 ISTD-PDD3 进行先验锚框选择, 有效提升裂缝灾害的检测精度。在特征增强阶段, 使用 CARAFE 和 MCA 模块, 避免特征重组过程中细小目标的丢失, 并获取了丰富的病害区域特征表示。

(3) 与 5 种主流目标检测模型对比, 改进的 YOLOv5 模型对 3 类路面病害的评估指标均得分最高, 证明该模型在复杂场景下水泥路面病害检测方面的优越性。

(4) 综合应用性能分析结果表明, 改进的 YOLOv5 模型在检测准确率上表现优秀, 且与人工目视判别法的 PCI 一致性较高, 但该模型仍有改进空间, 未来的研究可着重于扩展更多病害类型的样本库, 以及提升模型在病害检测方面的实时性。

### 参考文献(References):

[1] HAO X L, SHA A M, SUN Z Y, et al. Laser-based measuring method for mean joint faulting value of concrete pavement[J]. Optik, 2016,

127(1):274-278.

- [2] 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.
- [3] CHENG H D, SHI X J, GLAZIER C. Real-time image thresholding based on sample space reduction and interpolation approach[J]. Journal of Computing in Civil Engineering, 2003, 17(4): 264-272.
- [4] AYENU-PRAH A, ATTOH-OKINE N. Evaluating pavement cracks with bidimensional empirical mode decomposition[J]. EURASIP Journal on Advances in Signal Processing, 2008, 2008(1): 861701.
- [5] HE Y Q, QIU H X, WANG J, et al. Studying of road crack image detection method based on the mathematical morphology[C]//2011 4th International Congress on Image and Signal Processing, October 15-17, 2011, Shanghai, China. IEEE, 2011: 967-969.
- [6] AMHAZ R, CHAMBON S, IDIER J, et al. Automatic crack detection on two-dimensional pavement images: an algorithm based on minimal path selection[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2718-2729.
- [7] ZHANG D J, LI Q Q, CHEN Y, et al. An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection[J]. Image and Vision Computing, 2017, 57: 130-146.
- [8] LI N N, HOU X D, YANG X Y, et al. Automation recognition of pavement surface distress based on support vector machine[C]//2009 Second International Conference on Intelligent Networks and Intelligent Systems, November 1-3, 2009, Tianjin, China. IEEE, 2009: 346-349.
- [9] 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(1): 7419058.
- [10] 李章维, 胡安顺, 王晓飞. 基于视觉的目标检测方法综述[J]. 计算机工程与应用, 2020, 56(8): 1-9.  
LI Zhangwei, HU Anshun, WANG Xiaofei. Survey of vision-based object detection methods[J]. Computer Engineering and Applications, 2020, 56(8): 1-9.
- [11] DU Z Y, YUAN J, XIAO F P, et al. Application of image technology on pavement distress detection: a review[J]. Measurement, 2021, 184: 109900.
- [12] RAJESHWARI P, ABHISHEK P, VINOD P S, et al. Object detection: an overview[J]. International Journal of Trend in Scientific Research and Development, 2019, 3(3): 1663-1665.
- [13] LIU Y C, LIU F, LIU W, et al. Pavement distress detection using street view images captured via action camera[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(1): 738-747.
- [14] LI J Q, ZHAO X F, LI H W. Method for detecting road pavement damage based on deep learning[C]//Health Monitoring of Structural and Biological Systems XIII. April 4-7, 2019, Denver, Colorado, United States. Bellingham, Washington: SPIE, 2019: 517-526.
- [15] 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): 336-344.  
LUO Hui, JIA Chen, LI Jian. Road surface disease detection algorithm based on improved YOLOv4[J]. Laser & Optoelectronics Progress, 2021, 58(14): 336-344.
- [16] 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. December 10-13,2020,Atlanta,GA,USA.IEEE,2020:5577-5583.
- [17] GAO M X,WANG X,ZHU S L,et al.Detection and segmentation of cement concrete pavement pothole based on image processing technology[J]. Mathematical Problems in Engineering,2020,2020: 1360832.
- [18] SADATI S,CHAGAS BRITO DA SILVA L E,WUNSCH D C I I, et al.Artificial intelligence to investigate modulus of elasticity of recycled aggregate concrete[J]. ACI Materials Journal,2019,116 (1):51-62.
- [19] 惠冰,郭牧,王洲,等. 基于三维激光技术的路面坑槽多维度指标检测[J]. 同济大学学报(自然科学版),2018,46(1):60-66.  
HUI Bing, GUO Mu, WANG Zhou, et al. Multi-dimensional index detection of potholes based on 3D laser technology[J]. Journal of Tongji University(Natural Science),2018,46(1):60-66.
- [20] 郭小英,李文书,钱宇华,等. 可计算图像复杂度评价方法综述 [J]. 电子学报,2020,48(4):819-826.  
GUO Xiaoying, LI Wenshu, QIAN Yuhua, et al. Computational evaluation methods of visual complexity perception for images[J]. Acta Electronica Sinica,2020,48(4):819-826.
- [21] SU B Y,ZHANG H,WU Z H,et al.FSRDD:an efficient few-shot detector for rare city road damage detection[J].IEEE Transactions on Intelligent Transportation Systems,2022,23(2):24379-24388.
- [22] MAJIDIFARD H,JIN P,ADU-GYAMFI Y,et al.Pavement image datasets: a new benchmark dataset to classify and densify pavement distresses[J]. Transportation Research Record: Journal of the Transportation Research Board,2020,2674(2): 328-339.
- [23] REDMON J,FARHADI A.YOLO9000: better, faster, stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. July 21-26,2017,Honolulu,HI,USA.IEEE,2017:6517-6525.
- [24] WANG J Q, CHEN K, XU R, et al. CARAFE: content-aware ReAssembly of FEatures[C]//2019 IEEE/CVF International Conference on Computer Vision. October 27- November 2, 2019, Seoul,Korea.IEEE,2019:3007-3016.
- [25] CAO Y,XU J R,LIN S, et al.GCNet: non-local networks meet squeeze-excitation networks and beyond[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop. October 27-28, 2019, Seoul, Korea. IEEE, 2019: 1971-1980.
- [26] YAO H,LIU Y H,LI X,et al.A detection method for pavement cracks combining object detection and attention mechanism[J]. IEEE Transactions on Intelligent Transportation Systems,2022,23 (11): 22179-22189.
- [27] YU Y,ZHANG Y,CHENG Z Y, et al. MCA: multidimensional collaborative attention in deep convolutional neural networks for image recognition[J]. Engineering Applications of Artificial Intelligence ,2020, 126: 107079.
- [28] REN S Q,HE K M,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.
- [29] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//2016 European Conference on Computer Vision, October 8-16, 2016, Amsterdam, Netherlands. Berlin: Springer International Publishing,2016:21-37.
- [30] DUAN K W,BAI S,XIE L X,et al.CenterNet:keypoint triplets for object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. October 27-November 2, 2019, Seoul, Korea. IEEE,2019:6568-6577.
- [31] WANG C Y,BOCHKOVSKIY A,LIAO H M.YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 17-24,2023,Vancouver,BC,Canada. IEEE,2023:7464-7475.
- [32] 徐鹏,祝轩,姚丁,等. 沥青路面养护智能检测与决策综述[J]. 中南大学学报(自然科学版),2021,52(7):2099-2117.  
XU Peng, ZHU Xuan, YAO Ding, et al. Review on intelligent detection and decision-making of asphalt pavement maintenance [J]. Journal of Central South University (Science and Technology), 2021,52(7): 2099-2117.