

doi: 10.11835/j.issn.1000-582X.2023.101

# 改进 YOLOv5s 的桥梁表观病害检测方法

董绍江<sup>1</sup>, 谭浩<sup>1</sup>, 刘超<sup>1</sup>, 胡小林<sup>2</sup>

(1. 重庆交通大学机电与车辆工程学院, 重庆 400074; 2. 重庆工业大数据创新中心有限公司, 重庆 404100)

**摘要:**针对已有目标检测方法在混凝土桥梁表观病害检测的应用中识别精度低且伴随较多误检和漏检的问题,提出了一种改进的 YOLOv5s 桥梁表观病害检测方法。针对目前桥梁表观病害特征成分较复杂的问题,为了更有效地利用不同尺度的缺陷特征,在主干网中添加修改后的空间金字塔池化模块,提高了整体网络对缺陷特征信息的获取能力,同时减少了运算工作量;针对由病害图像中不同缺陷特征交叉分布导致的误检率、漏检率高的问题,在 YOLOv5s 网络中加入轻量化注意力模块;针对桥梁缺陷尺寸差异大、分类困难和数据集小而导致的边界回归不匹配的问题,采用考虑了向量角度的损失函数。实验证明,改进后的 YOLOv5s 检测器在桥梁表观病害目标检测识别任务中能够有效提高精度、降低误检率和漏检率。

**关键词:**病害检测; YOLOv5s; 特征融合; 平均精度

中图分类号: TP391.4

文献标志码: A

文章编号: 1000-582X(2024)09-091-10

## Apparent disease detection of bridges using improved YOLOv5s

DONG Shaojiang<sup>1</sup>, TAN Hao<sup>1</sup>, LIU Chao<sup>1</sup>, HU Xiaolin<sup>2</sup>

(1. School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, P. R. China; 2. Chongqing Industrial Big Data Innovation Center Co., Ltd., Chongqing 404100, P. R. China)

**Abstract:** To solve the problems of low accuracy, high false detection rate, and high missed detection rate in current target detection methods for apparent diseases in concrete bridges, an improved YOLOv5s method is proposed. To achieve more effective fusion of features at different scales and increase receptive fields, an improved spatial pyramid pooling module is added to the YOLOv5s network to enhance feature extraction capabilities and reduce computational cost; a light-weight attention module is incorporated into the YOLOv5s network to tackle the high false detection and missed detection rates caused by the cross-distribution of different defect features in disease images; and a loss function considering vector angles is adopted to solve the problems related to varying defect sizes, classification difficulties and small dataset-induced boundary box regression mismatches. Experimental results show that the improved YOLOv5s detector significantly improves accuracy while reducing false detection and missed detection rates in the task of detecting apparent diseases in bridges.

**Keywords:** disease detection; YOLOv5s; feature fusion; mean average accuracy

收稿日期: 2022-08-19 网络出版日期: 2023-04-03

基金项目: 国家自然科学基金资助项目(51775072); 重庆市科技创新领军人才支持计划项目(CSTCCCXLJRC201920); 重庆市高校创新研究群体(CXQT20019)。

Supported by National Natural Science Foundation of China (51775072), the Chongqing Science and Technology Innovation Leading Talents Support Program (CSTCCCXLJRC201920), and the Chongqing University Innovation Research Group (CXQT20019).

作者简介: 董绍江(1982—), 男, 教授, 博士, 主要从事机电一体化、智能机器人研究, (E-mail) dongshaojiang100@163.com。

通信作者: 谭浩, 男, 硕士研究生, (E-mail) tanhao999@outlook.com。

截至2021年底,中国的公路桥梁总共有96.11万座,全长73 802.1 km。桥梁的运营过程中,随之而来的是桥梁功能的退化和桥梁本体结构的损伤,因此对桥梁结构的表观检测是桥梁运维的重要工作之一。目前桥梁表观缺陷检测的主要方式为目测和简单仪器测量,但这些检测方法主观性较强,且受现场环境等综合因素影响,对一些存在病害的地方可能产生大量误检、漏检。为了解决这一问题,基于机器视觉的检测方法被提出,通过无人机和爬壁机器人<sup>[1]</sup>等设备来获取图像,将获取的图像导入计算机并且提取图像特征,进一步识别和分类,达到缺陷识别的目的,如Nishikawa等<sup>[2]</sup>在对桥梁表面裂缝识别中引入机器学习的方法。虽然机器学习方法相较于传统检测技术在检测效率、鲁棒性上有一定提升<sup>[3]</sup>,但这些方法的特征提取过程十分繁琐,且不能够一次性、端到端地解决缺陷识别问题。

近年来,基于深度学习的缺陷检测方法解决了传统机器学习方法检测速率低、漏检率高的问题,同时兼具检测速度快、鲁棒性强等优点。基于深度学习的检测方式按照其网络体系结构分为以下2类:第1类是以区域卷积神经网络<sup>[4]</sup>(regional convolutional neural networks, R-CNN)为代表的Fast R-CNN<sup>[5]</sup>、Faster R-CNN<sup>[6]</sup>等双阶段检测算法。R-CNN借用了滑动窗口来产生候选框域的思想,这种方法步骤复杂、计算量大。针对R-CNN较费时的选择性搜索,Faster R-CNN使用了一个区域候选网络来取代选择性搜索,如Cha等<sup>[7]</sup>将Faster R-CNN用于识别混凝土和钢铁裂缝相关的多种表面损伤。虽然大量研究证实了双阶段检测方法可用于钢筋混凝土等结构件表面缺陷的检测,但中间过程繁琐,且需要大量重复性实验,因此其检验效率不够好,无法适应现场检测要求。第2类是以SSD(single shot detector)<sup>[8]</sup>、YOLO(you only look once)<sup>[9]</sup>为代表的单阶段检测算法。它并不要求先给出目标候选框,而是直接把目标图像特征的提取和预测框的确定结合在一起进行图像的分类估计和区域坐标评估,极大地提高了检测效率。SSD算法的缺点明显,一是较低层级的特征非线性程度不够,导致训练精度不如YOLO;二是候选框尺寸需要人工经验判断。随着YOLOv3<sup>[10]</sup>、YOLOv4<sup>[11]</sup>、YOLOv5的先后提出,YOLO系列算法与其他算法比较在检测速率和精度上均表现出较大的优越性。Peng等<sup>[12]</sup>提出一种使用改进YOLO的双网络方法来提高算法对块状缺陷和裂缝病害的检测能力;Jang等<sup>[13]</sup>提出一种基于混合扩展卷积块的深度学习网络HDCB-Net,并结合YOLOv4用于实现像素级的裂纹检测。

现有基于YOLOv3、YOLOv4的桥梁表观病害检测识别方法在现场作业时不能满足实时检测需求,YOLOv5系列算法则在检测精度、速度上都优于YOLOv3,其在一些公共数据集如COCO<sup>[14]</sup>(common objects in context)数据集上有较好的表现,但在数据特征较为复杂的桥梁表观病害检测中尚有改进空间。本研究的目的是探讨对YOLOv5s算法的改进及其对桥梁表观病害检测任务的适用性。

## 1 YOLOv5s 模型框架

YOLOv5s是一种单阶段检测方法,结合特征提取和预测框设定对目标的分类概率和定位坐标进行评估。相对于其他算法,YOLOv5s具有模型小、检测速度快和检测精度高等优势。YOLOv5s结构如图1所示,可将YOLOv5s算法模型划分为4个部分。

第1部分是输入层,主要负责对已提供的图像数据进行预处理,包括数据强化、自适应锚框算法和自适应图像缩放。YOLOv5s使用Mosaic数据增强技术,通过对前4张图片随机压缩、剪切、重新排布并加以拼接,以加强对小目标的检测效果;自适应锚框主要是根据当前数据集选择与其长宽对应的锚框;自适应图像缩放不同于上一代YOLOv4,为图片增加最少像素的额外黑边,最大程度保持图片的原有性质,同时减少不必要的计算量。图1中Input为输入;Conv2d表示二维卷积;BN表示归一化层;Silu为激活函数;CBS表示顺序连接的Conv2d层、BN层和Silu层;C3表示包含3个卷积层的跳跃连接模块,C3的数量为模块中残差结构的数量;SPPF表示空间金字塔池融合模块;Concat为拼接;Upsample为上采样;Head为模型输出头。

第2部分为主干网络,主要作用是提取图像中目标的特征。在YOLOv4算法的基础上,使用了改进后的CSP-Darknet53结构,网络结构如图2所示。输入特征图通过2个维度进行计算,首先都使用 $1 \times 1$ 卷积(Conv)

进行变换,其中一路在经过步幅为 2 的 3×3 卷积后与另一路跨阶段层级拼接。这不仅保证了准确率,同时也解决了普通网络在信息传递时梯度重复的问题,因此参数更小,网络计算量更少,进一步节省内存开支。

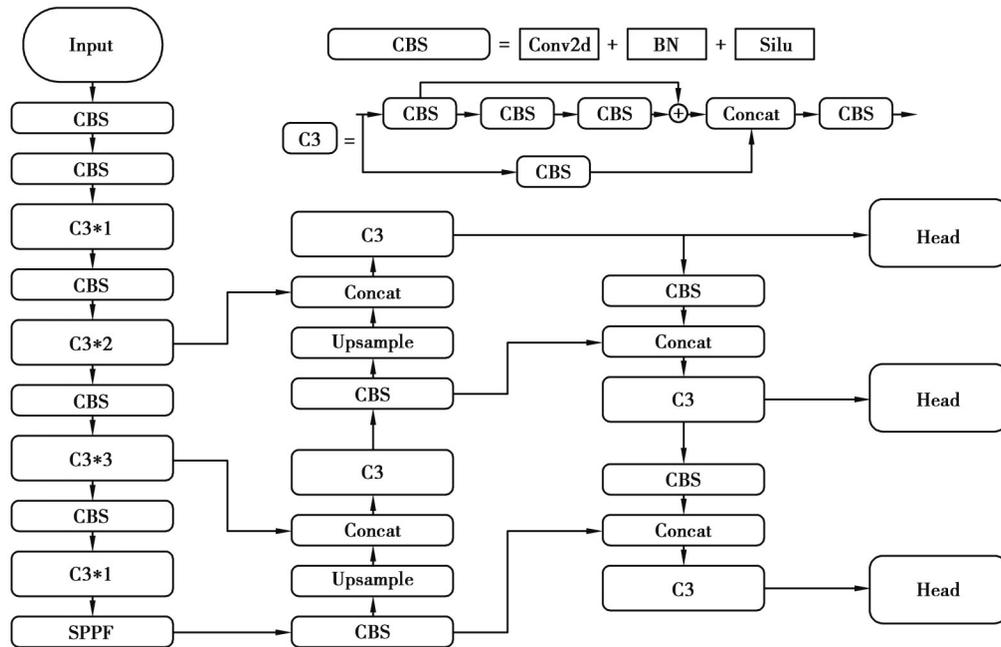


图 1 YOLOv5s 模型框架

Fig. 1 YOLOv5s framework

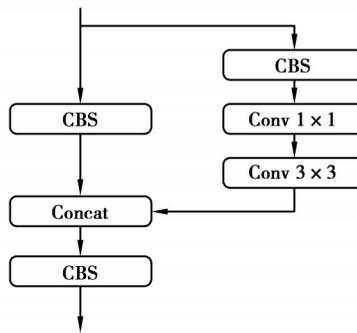


图 2 CSP-Darknet 网络结构

Fig. 2 CSP-Darknet structure

第 3 部分为颈部网络,作用为收集目标特征,其结构为特征金字塔网络加路径聚合网络。特征金字塔网络的作用是传递高层语义特征,路径聚合网络主要传递定位信息,对特征金字塔网络进行补充。

第 4 部分为检测层,主要用于预测信息损失部分。

## 2 改进的桥梁表观病害检测模型

本研究中将改进的快速空间金字塔池化(ReLu spatial pyramid pooling-fast, RSPPF)融入骨干网络,加入高效通道注意力<sup>[15]</sup>(efficient channel attention, ECA),提出用于桥梁表观病害检测的网络(YOLOv5s of bridge detection, YOLOv5s-BD)如图 3 所示,并使用 Siou<sup>[16]</sup>进行损失计算。

YOLOv5s-BD 中,将原骨干网络末端的金字塔池化块替换为改进的 RSPPF,并在网络的第 5 和第 8 层添加 RSPPF,在第 11 层添加 ECA 注意力模块。

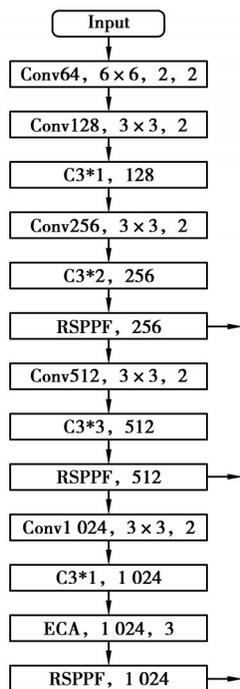


图 3 YOLOv5s-BD 骨干网络  
Fig. 3 YOLOv5s-BD backbone

2.1 空间金字塔池化

空间金字塔池化<sup>[17]</sup>(spatial pyramid pooling, SPP)网络结构是将输入特征先通过一个卷积,然后分别进行不同卷积核的池化后与输入拼接,再通过一个卷积,中间的卷积核大小分别为 $5 \times 5$ 、 $9 \times 9$ 和 $13 \times 13$ 。不仅在一定程度上避免了由区域裁剪、缩放导致的失真问题和重复特征提取的问题,同时也节省了计算成本。SPP能够将图片在不同维度上的特征信息融合到一个维度上。

YOLOv5s中的快速金字塔池化(spatial pyramid pooling-fast, SPPF)借鉴了SPP的思想,但通道间的传递关系改为顺序传递,YOLOv5s中的SPPF模块如图4所示,图中Maxpool表示最大池化层, $k$ 代表卷积核大小。RSPPF沿用SPPF结构,其中CBR为使用了Relu激活函数替换掉Silu激活函数的卷积操作,中间卷积核大小统一为 $5 \times 5$ ,在保证网络精度的同时减少模块计算量。Silu和Relu函数关系如图5所示。

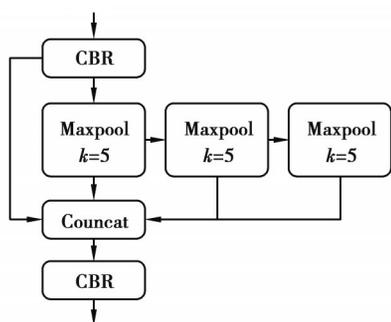


图 4 SPPF 结构  
Fig. 4 SPPF structure

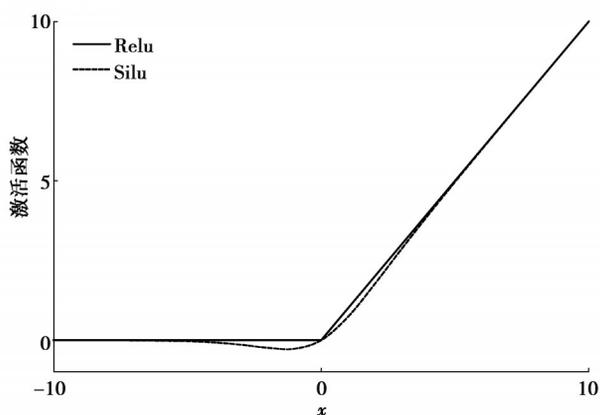


图 5 Silu 和 Relu 激活函数曲线  
Fig. 5 Activation functions Silu and Relu

Relu 函数表达式:

$$\text{Relu}(x) = \begin{cases} x, & x > 0; \\ 0, & x \leq 0. \end{cases} \quad (1)$$

Silu 函数表达式:

$$\text{Silu}(x) = \frac{x}{1 + e^{-x}} \quad (2)$$

显然 Silu 无上界有下界、平滑而不单调的特征使它在深层模型的效果上具有突出的优点,但其包含的指数运算和除法运算会占用大量计算量。Relu 为线性函数,计算量小,更适合特征选取,且能够有效避免反向传播带来的梯度弥散。

### 2.2 轻量级注意力机制 ECA

ECA 是由 Huang 等<sup>[18]</sup>提出的一个不降维的局部跨通道互动方法,克服了普通注意力中降维对学习的负面影响,实现了不同通道间的交互。ECA 网络模型如图 6 所示,其中  $W$ 、 $H$  分别为特征图的宽和高,  $C$  为通道维数,  $\sigma$  表示 Sigmoid 激活函数, GAP 为全局平均池化,  $\chi$  和  $\tilde{\chi}$  分别为输入张量和输出张量。

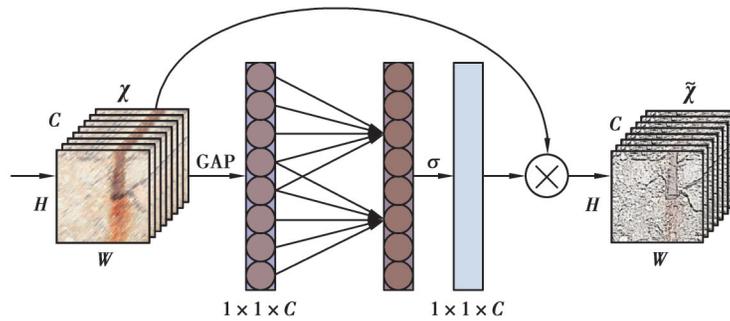


图 6 ECA 网络模型示意图

Fig. 6 Schematic diagram of efficient channel attention module

ECA 网络中利用矩阵  $W_k$  来学习通道注意力:

$$W_k = \begin{bmatrix} w^{1,1} & \dots & w^{1,k} & 0 & 0 & \dots & \dots & 0 \\ 0 & w^{2,2} & \dots & w^{2,k} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & w^{C,C-k+1} & \dots & w^{C,C} \end{bmatrix} \quad (3)$$

$W_k$  由  $k \times C$  个参数组成,避免了特征在不同通道上相互独立的问题。通道注意力  $\omega$  可通过下式计算:

$$\omega = \sigma(C1D_k(y)) \quad (4)$$

式中:  $C1D$  表示一维卷积,  $y$  为输出。通道维数  $C$  与卷积核大小  $k$  关系公式如下:

$$C = \phi(k) = 2^{\lceil \gamma(k-b) \rceil} \quad (5)$$

式中:  $\gamma$  和  $b$  为可调参数。

因此,卷积核尺寸表达式为:

$$k = \psi(C) = \left\lceil \frac{\log_2(C) + b}{\gamma} \right\rceil_{\text{odd}} \quad (6)$$

式中:  $\left\lceil \frac{\log_2(C) + b}{\gamma} \right\rceil_{\text{odd}}$  表示与  $\frac{\log_2(C) + b}{\gamma}$  值最相近的整数;  $k$  的取值会影响网络的整体运算速度及准确率,  $k$  值过小会导致交互的覆盖范围小,  $k$  值过大会增加计算量,在本网络中,设置  $k=3$ 。

### 2.3 损失函数

交并比(intersection over union, IoU)用于计算预测框损失,很大程度上决定了模型预测结果的准确程度。YOLOv5s 使用完全交并比(complete intersection over union, CIoU),减少了其他类 IoU 在计算中的发散现象。

CIoU计算公式为

$$\text{CIoU} = \text{IoU} - \frac{\rho^2(b, b^{\text{gt}})}{c^2} - \alpha v, \quad (7)$$

$$v = \frac{4^2}{\pi} \left( \arctan\left(\frac{w^{\text{gt}}}{h^{\text{gt}}}\right) - \arctan\left(\frac{w}{h}\right) \right)^2, \quad (8)$$

$$\alpha = \frac{v}{(1 - \text{IoU}) + v}. \quad (9)$$

式中: $\alpha$ 是权重系数; $v$ 用来度量宽高比的一致性; $w$ 、 $h$ 分别为候选框的宽和高; $\rho$ 为欧氏距离; $c$ 为预测框和标注框的对角线长度。

式(7)中的 $\rho^2(b, b^{\text{gt}})$ 表示预测框 $b$ 和真实框 $b^{\text{gt}}$ 中心点之间的欧式距离 $d$ 。 $c$ 代表它们的对角线距离,如图7所示。

在IoU、GIoU和CIoU等边界框回归指标中,预测框和真实框之间的方向不匹配时,会降低模型的收敛速度和效率。考虑回归向量角度并重新定义惩罚指标的SIoU损失函数由4个代价函数组成:角度函数、距离函数、形状函数和IoU函数。角度函数为

$$A = 1 - 2 \sin^2\left(\arcsin(x) - \frac{\pi}{4}\right). \quad (10)$$

式中: $x = \frac{\max(b_{c_x}^{\text{gt}}, b_{c_x}) - \min(b_{c_x}^{\text{gt}}, b_{c_x})}{\sqrt{(b_{c_x}^{\text{gt}} - b_{c_x})^2 + (b_{c_y}^{\text{gt}} - b_{c_y})^2}} = \sin(\alpha)$ ,这里的 $c_x$ 和 $c_y$ 分别为对角线距离的长和宽分量。

角度函数最大限度地减少了与距离相关的不匹配变量数量,收敛过程中首先最小化 $\alpha$ 。距离函数为

$$D = \sum_{t=x,y} (1 - e^{-\gamma \rho^t}). \quad (11)$$

式中: $\rho_x = \left(\frac{b_{c_x}^{\text{gt}} - b_{c_x}}{c_w}\right)^2$ ;  $\rho_y = \left(\frac{b_{c_y}^{\text{gt}} - b_{c_y}}{c_h}\right)^2$ ;  $\gamma = 2 - A$ 。

形状函数为

$$\Omega = \sum_{t=w,h} (1 - e^{-\theta t})^\theta. \quad (12)$$

式中: $\omega_w = \frac{|w - w^{\text{gt}}|}{\max(w, w^{\text{gt}})}$ ;  $\omega_h = \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})}$ ;  $\theta$ 值由数据集决定。

由SIoU定义的损失函数为

$$L_{\text{box}} = 1 - \text{IoU} + \frac{A + \Omega}{2}, \quad (13)$$

式中:IoU为预测框和真实框的交并比。

### 3 实验结果分析

#### 3.1 实验环境及数据集来源

实验在Windows操作系统环境下进行,使用CPU为AMD Ryzen7 5800H, RAM为16 GB随机存取内存, GPU为RTX 3060 Laptop, 6 GB显示内存。深度学习框架为pytorch, 整个训练过程设置epoch为300, Batch\_size为8, 训练时使用SGD优化算法进行参数优化, 初始学习率为0.01, 输入图片分辨率为640×640。

实验数据集为Mundt等<sup>[19]</sup>在2019年公开的用于桥梁混凝土缺陷检测的多目标具体分类数据集CODEBRIM(COncrete DEfect BRidge IMage Dataset)。该数据集包含裂缝、脱落、腐蚀、露筋和风化共5种桥梁病害,如图8所示,有效缺陷图像共1 052张。由于数据样本较少,且未考虑实际检测环境变化,因此使用数据增强技术对缺陷图像样本进行扩充,将原数据进行裁剪、镜像、亮度调节操作,得到有效缺陷图像共3 807张。将数据集按9:1的比例随机分为训练集、验证集。

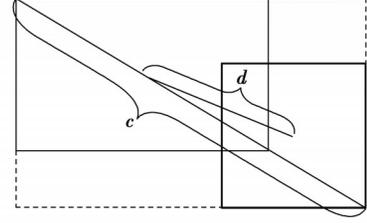


图7 预测框与标注框参数解释图  
Fig. 7 Explanatory diagram of parameters of prediction frame and anchor frame

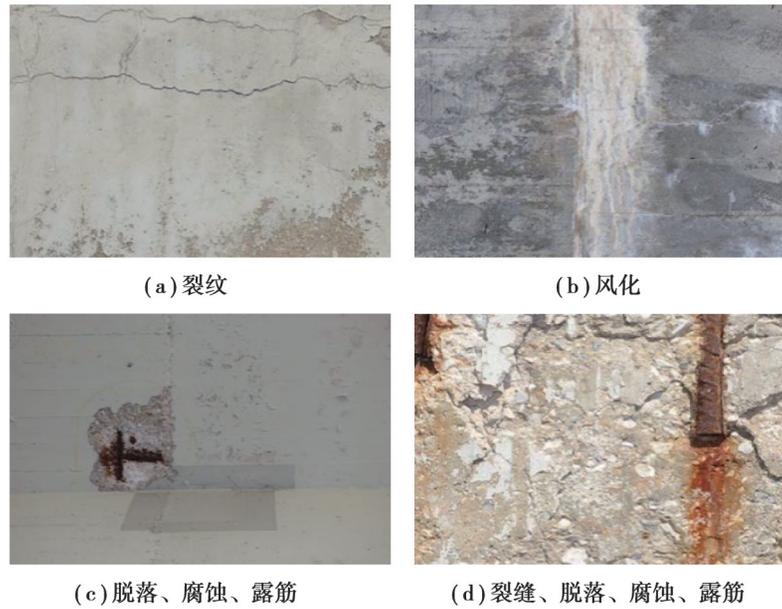


图 8 桥梁缺陷图片

Fig. 8 Pictures of various defects of bridges

### 3.2 评价指标

多标签图片划分任务中不止一种图像标签,因此不能用普通单标签图片划分任务的精度标准来作为评判指标。本研究中采用 mAP(mean average precision)作为评价指标。二元分类问题分类结果的混淆矩阵如表 1 所示,判断结果依据其标记类和预测类的组合可分为 4 类,分别为真正例、假正例、真负例和假负例,分别对应表 1 的 TP、FP、TN 和 FN。

表 1 二元分类混淆矩阵

Table 1 Confusion matrix for binary classification

标记	真实值	
	正 (Positive)	负 (Negative)
正 (Positive)	TP	FP
负 (Negative)	FN	TN

精确度 (precision,  $P$ ) 和召回率 (recall,  $R$ ) 计算公式如下:

$$P = \frac{TP}{TP + FP}, \quad (14)$$

$$R = \frac{TP}{TP + FN}. \quad (15)$$

精确度表示该模型预测为正例的样本中实际为正例的样本所占的比例;召回率表示实际为正例的样本中模型预测为正例的样本所占的比例。每个类的精度值 (AP) 就是  $P$ - $R$  曲线与坐标轴围成的面积。mAP 是这些缺陷分类的  $P$ - $R$  曲线下的面积取平均值, mAP 可以作为一个相对较好的度量指标。AP 和 mAP 公式如下:

$$AP = \frac{\sum_{i=1}^N P_i}{N}, \quad (16)$$

$$mAP = \frac{\sum_{j=1}^M AP_j}{M}. \quad (17)$$

式中:  $M$  表示用于检测的类别总数;  $N$  表示测试的图像数量。mAP 包含了  $mAP_{@0.5}$  和  $mAP_{@0.5:0.95}$ , 具体由设定的 IoU 阈值决定, 其表达式如下:

$$mAP_{@0.5} = \frac{\sum_{j=1}^M AP_{@0.5j}}{M}, \quad (18)$$

$$mAP_{@0.5:0.95} = \frac{\sum_{t=0}^9 mAP_{@(0.5+0.05t)}}{10} \quad (19)$$

由式(18)和式(19)可知,  $mAP_{@0.5}$  即 IoU 阈值为 0.5 时的平均精度, 揭示了精确度  $P$  和召回率  $R$  的变化趋势。  $mAP_{@0.5:0.95}$  是以 0.05 为步进, IoU 阈值从 0.50 到 0.95 的平均精度的平均值, 揭示了模型在不同 IoU 阈值下的综合表现,  $mAP_{@0.5:0.95}$  越高代表模型的边界回归能力越强, 预测框与标注框的拟合越精确。

### 3.3 实验结果与分析

本研究中使用的数据集缺陷不易于分类, 且某些缺陷(如裂缝)这样的小目标受输入图片分辨率的影响, 在检测时, 设置输入图片分辨率为  $640 \times 640$ 。为了验证提出的方法的有效性, 考虑到实时性检测的需求, 将使用了 YOLOv5s 的原生模型检测效果分别与改进空间金字塔池化的 YOLOv5s-RSPPF、加入 ECA 注意力模块的 YOLOv5s-ECA、使用 SIoU 进行损失计算的 YOLOv5s-SIoU 和本文中提出的改进骨干网络和损失函数的 YOLOv5s-BD 通过测试数据集进行对比, 结果见表 2, 其中 FPS 为每秒检测帧数(frames per second)。

表 2 改进的 YOLOv5s 和原版 YOLOv5s 实验表现

Table 2 Experimental results of improved YOLOv5s and original YOLOv5s

算法	AP/%					$mAP_{@0.5}/\%$	$mAP_{@0.5:0.95}/\%$	FPS/Hz
	裂缝	脱落	腐蚀	露筋	风化			
YOLOv5s	79.6	86.7	83.9	90.7	77.8	83.7	55.4	71.9
YOLOv5s-SIoU	79.9	87.1	85.0	90.6	80.0	84.5	56.0	70.4
YOLOv5s-RSPPF	80.9	87.7	84.7	90.9	79.9	84.8	56.8	78.0
YOLOv5s-ECA	81.3	87.6	85.0	91.6	81.9	85.5	57.2	70.4
YOLOv5s-BD	84.2	90.5	86.8	92.3	83.6	87.5	61.5	73.6

改进后的 YOLOv5s-BD 相较于 YOLOv5s 的原生模型对各种缺陷的检测效果均有提升。其中, 由于对预测框的损失函数做出改进, YOLOv5s-SIoU 相较于原模型  $mAP_{@0.5}$  提升了 0.8%,  $mAP_{@0.5:0.95}$  提升了 0.6%; 由于对原生模型的 SPPF 模块做出改进, YOLOv5s-RSPPF 相较于原模型  $mAP_{@0.5}$  提升了 1.1%,  $mAP_{@0.5:0.95}$  提升了 1.4%; 由于在特征提取网络尾部加入 ECA 注意力机制, YOLOv5s-ECA 相较于原模型  $mAP_{@0.5}$  提升了 1.8%,  $mAP_{@0.5:0.95}$  提升了 1.8%。结合以上几种方法, YOLOv5s-BD 相较于原模型  $mAP_{@0.5}$  提升了 3.8%,  $mAP_{@0.5:0.95}$  提升了 6.1%。虽然总体网络层数有所增加, 但由于改进的网络采用了更加轻量化的计算方式, 每秒检测帧数 FPS 也略微提升, 满足实时检测需求。

为了进一步验证本研究中的模型在用于混凝土桥梁表面缺陷检测时的优势, 在不改变数据集的情况下, 使用本研究改进网络 YOLOv5s-BD 与 Faster RCNN、SSD 和 YOLOv3 进行实验验证, 结果见表 3。

表 3 不同检测网络对比

Table 3 Comparison of different detection networks

算法	AP/%					$mAP/\%$	FPS/Hz
	裂缝	脱落	腐蚀	露筋	风化		
Faster RCNN	71.2	76.8	76.9	78.2	77.4	76.3	12.4
SSD	70.5	75.3	75.1	76.2	74.3	74.3	45.3
YOLOv3	69.9	83.9	85.3	86.1	80.4	81.1	29.6
YOLOv5s-BD	84.2	90.5	86.8	92.3	83.6	87.5	73.6

通过表 3 可以看出, 由于 YOLOv5s-BD 拥有对交叉分布病害特征更有效的提取能力和更先进的边界框回归指标, 不论是以 Faster RCNN 为代表的双阶段检测算法, 还是 SSD、YOLOv3 这类单阶段检测算法, 平均准确率和 FPS 都比 YOLOv5s-BD 算法低, 证明本研究中改进的网络模型在应用于混凝土桥梁表面病害检测方面有更大的优势。

### 3.4 缺陷检测的效果

对测试集中的样本进行随机抽检,其检测效果如图9所示。

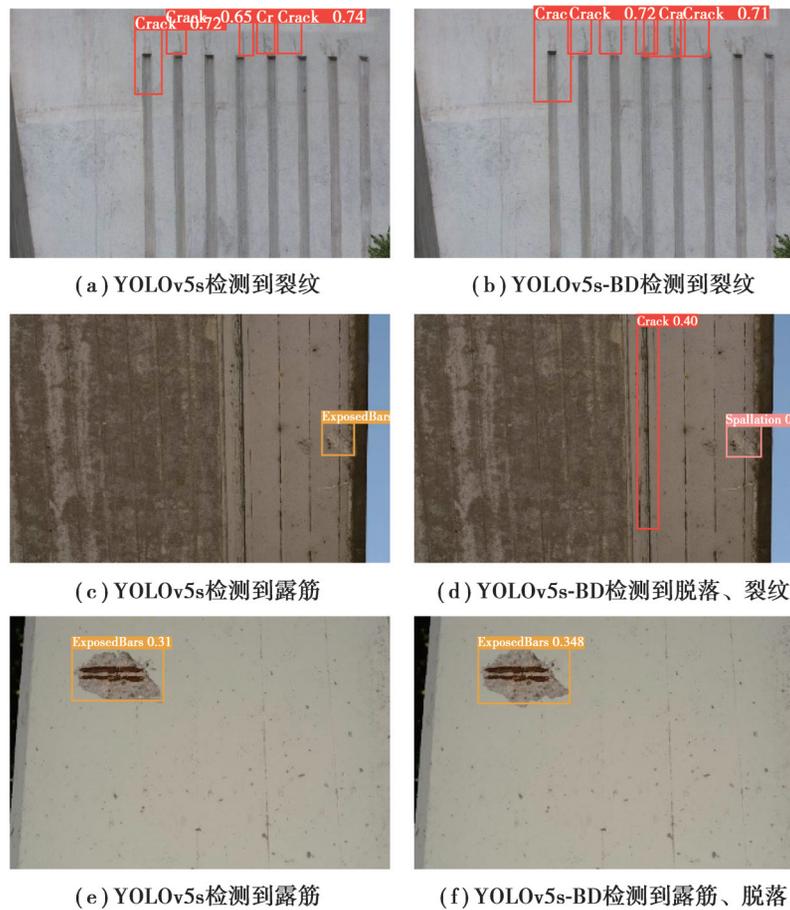


图9 YOLOv5s和YOLOv5s-BD检测效果

Fig. 9 Detect results of YOLOv5s and YOLOv5s-BD

图9(a)检测到部分裂纹(crack),图9(b)检测到了更多的裂纹,且没有误检;图9(c)将脱落(spallation)误检为露筋(exposed bars)且漏检了裂纹,图9(d)准确检测到脱落和裂纹;图9(e)检测到露筋但没有检测到脱落;图9(f)准确识别到露筋和脱落。综上所述,相比于YOLOv5s,YOLOv5s-BD模型平均精度 $mAP@0.5$ 提升了3.8%, $mAP@0.5:0.95$ 提升了6.1%,检测速度也略有提升,能够有效地检测到更多缺陷且拥有更低的误检率和漏检率,因此更适用于混凝土桥梁表面病害检测。

## 4 结束语

针对当前目标检测算法在应用于桥梁表面病害检测时精度低、误检率高、漏检率高和检测速率低的问题,提出一种用于桥梁表面病害检测的改进YOLOv5s检测网络;针对桥梁缺陷尺寸差异大、分类困难,以及数据集小的问题,为提高缺陷预测的精确度,引入新的预测框损失函数进行训练。实验结果表明,改进后的YOLOv5s模型在进行桥梁表面缺陷实时检测时,多种缺陷的检测平均精度均有明显提升,且满足实时检测需求,相比于原模型漏检率和误检率更低。下一步工作为:1)针对直接破坏桥梁结构的缺陷(如裂缝)进行尺寸检测工作;2)研究搭载YOLOv5改进算法的无人机或者爬壁机器人在进行桥梁病害检测中的应用,取代人工目视来达到病害精准识别和保证工人安全的目的。

### 参考文献

- [1] Ellenberg A, Kotsos A, Bartoli I, et al. Masonry crack detection application of an unmanned aerial vehicle[C]// 2014 International Conference on Computing in Civil and Building Engineering, June 23-25, 2014, Orlando, Florida, USA. Reston, VA, USA: American Society of Civil Engineers, 2014: 1788-1795.

- [ 2 ] Nishikawa T, Yoshida J, Sugiyama T, et al. Concrete crack detection by multiple sequential image filtering[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2012, 27(1): 29-47.
- [ 3 ] 周清松, 董绍江, 罗家元, 等. 改进YOLOv3的桥梁表观病害检测识别[J]. *重庆大学学报*, 2022, 45(6): 121-130.  
Zhou Q S, Dong S J, Luo J Y, et al. Bridge apparent disease detection based on improved YOLOv3[J]. *Journal of Chongqing University*, 2022, 45(6): 121-130. (in Chinese)
- [ 4 ] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 23-28, 2014, Columbus, OH, USA. IEEE, 2014: 580-587.
- [ 5 ] Girshick R. Fast R-CNN[C]// 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. IEEE, 2015: 1440-1448.
- [ 6 ] Ren S, He K, 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 .
- [ 7 ] Cha Y J, Choi W, Suh G, et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2018, 33(9): 731-747.
- [ 8 ] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [M]// *Computer Vision-ECCV 2016*. Cham: Springer International Publishing, 2016: 21-37.
- [ 9 ] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [ 10 ] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. 2018-04-08[2022-08-12]. <https://arxiv.org/abs/1804.02767>.
- [ 11 ] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. 2020-04-23[2022-08-12]. <https://arxiv.org/abs/2004.10934>.
- [ 12 ] 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 Y N, Liu M, Wan Z, et al. A dual deep network based on the improved YOLO for fast bridge surface defect detection[J]. *Journal of Automation*, 2022, 48(4): 1018-1032. (in Chinese)
- [ 13 ] Jiang W, Liu M, Peng Y, et al. HDCB-Net: a neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges[J]. *IEEE Transactions on Industrial Informatics*, 2020, 17(8): 5485-5494.
- [ 14 ] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context[C]// *European Conference on Computer Vision (ECCV)*, September 06-12, 2014, Zurich, Switzerland. ECCV, 2014: 740-755.
- [ 15 ] Wang Q, Wu B, Zhu P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 2575-7075.
- [ 16 ] Gevorgyan Z. SIOU loss: more powerful learning for bounding box regression [EB/OL]. 2022-05-25 [2022-08-12]. <https://arxiv.org/abs/2205.12740>.
- [ 17 ] Purkait P, Zhao C, Zach C. SPP-Net: deep absolute pose regression with synthetic views[EB/OL]. 2017-12-09[2022-08-12]. <https://arxiv.org/abs/1712.03452>.
- [ 18 ] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2020: 11534-11542.
- [ 19 ] Mundt M, Majumder S, Murali S, et al. Meta-learning convolutional neural architectures for multi-target concrete defect classification with the COncrete DEfect BRidge IMage Dataset[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. IEEE, 2019: 11188-11197.

(编辑 罗敏)