

doi:10.11835/j.issn.1000-582X.2021.107

改进 YOLOv3 的桥梁表观病害检测识别

周清松^{1a}, 董绍江^{1a,2}, 罗家元^{1a}, 秦悦^{1a}, 夏宗佑^{1a}, 杨建喜^{1b}

(1.重庆交通大学 a.机电与车辆工程学院; b.信息科学与工程学院,重庆 400074;2.西南交通大学
磁浮技术与磁浮列车教育部重点实验室,成都 610031)

摘要:针对基于目标检测方法的桥梁表观病害检测存在检测精度低、误检率和漏检率高的问题,提出一种改进 YOLOv3 的高准确率桥梁表观病害检测识别方法。为实现局部特征和全局特征有效融合,在 YOLOv3 的检测层中添加固定分块大小的池化模块,并在 YOLOv3 的特征提取网络中引入了 DenseNet 密集型连接网络结构以增强桥梁病害特征在网络中的传播和利用效率,提高检测效率,采用数据增强技术来扩充样本图像以解决现有桥梁病害数据集样本数量不足的问题。实验结果表明,改进后的 YOLOv3 在桥梁表观病害检测上的平均准确率比原 YOLOv3 提高了 3.0%,且模型训练时间减少了 33.2%,同时降低了对桥梁表观病害检测的误检率和漏检率。

关键词:目标检测;改进 YOLOv3;数据增强;平均准确率

中图分类号:TP391.41

文献标志码:A

文章编号:1000-582X(2022)06-121-10

Bridge apparent disease detection based on improved YOLOv3

ZHOU Qingsong^{1a}, DONG Shaojiang^{1a,2}, LUO Jiayuan^{1a}, QIN Yue^{1a}, XIA Zongyou^{1a}, YANG Jianxi^{1b}

(1a. School of Mechatronics and Vehicle Engineering; 1b. School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, P. R. China; 2. Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, Southwest Jiaotong University, Chengdu 610031, P. R. China)

Abstract: To solve the problems of low detection accuracy, high false detection rate, and high missed detection rate in bridge apparent disease detection based on target detection method, a recognition method with high accuracy of bridge apparent disease detection based on improved YOLOv3 is proposed. A pooling module with a fixed block size is added to the detection layer of YOLOv3 to realize effective fusion of local features and global features. To enhance the transmission and utilization efficiency of bridge disease features in the network and improve the detection efficiency, a DenseNet dense connection network structure is introduced in the feature extraction network of YOLOv3. To deal with the insufficient number of samples in the existing bridge disease data set, data enhancement technology is used to expand the sample images. The experimental results show that the mean accuracy precision (mAP) of the improved YOLOv3 on bridge apparent disease detection is increased by 3.0% and the model training time decreased by 33.2%, with a reduced false detection rate and a lower missed detection rate.

Keywords: object detection; improved YOLOv3; data enhancement; mean accuracy precision

收稿日期:2020-12-25 网络出版日期:2021-05-14

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)。

Supported by Natural Science Foundation of China (51775072), and the Science and Technology Innovation Leading Talent Support Program of Chongqing (CSTCCCXLJRC201920).

作者简介:周清松(1996—),男,硕士研究生,主要从事爬壁机器人与图像处理技术研究,(E-mail) 1374197415@qq.com。

通信作者:董绍江,男,教授,博士生导师,主要从事机电一体化技术研究,(E-mail) dongshaojiang100@163.com。

桥梁的可靠性和安全性对社会的福祉至关重要。因此有必要尽早发现桥上出现的缺陷,以防止桥的结构能力和耐用性进一步损失。在识别和监测缺陷的非破坏性评估技术中,人工目视检查是评估桥梁状况的主要手段^[1],但其结果是主观的且可能是不可靠的^[2]。在这种情况下,人们提出了基于计算机视觉的检测技术,借助爬壁机器人或无人机获取图像^[3-4]。基于机器学习的方法是更先进的检测方式,它利用提取图像中的特征来完成特定的任务,如 Nishikawa 等^[5]基于机器学习方法来研究桥梁表面裂缝。尽管机器学习技术比传统的图像处理技术在效率和鲁棒性方面有显著提高,但是这些方法仍然有基于手工制作的低级功能,并且需要进行预处理和后处理。因此,桥梁表面病害自动检测识别技术应运而生。

近年来,随着基于深度学习的目标检测算法不断改进创新,自动检测识别技术在行人检测、车辆检测等领域有很好的效果,但在桥梁表面病害检测方面效果较差。在目标检测算法的网络体系结构方面,这些对象检测器可分为两大类。一类以两级检测器 R-CNN^[6]系列为代表。R-CNN 方法使用选择性搜索^[7](SS)方法生成区域建议,然后分别提取每个潜在边界框的特征以进行分类和边框回归。但是,使用选择性搜索方法生成区域建议的步骤缓慢且复杂。为克服 R-CNN 存在的问题,Ren 等^[8]提出 Faster R-CNN 方法,用区域建议网络(RPN)代替选择性搜索方法生成区域建议。这种两级检测器被用于检测结构缺陷,如 Kim 等^[9]将 R-CNN 与形态学后处理相结合,以检测和量化混凝土桥梁中的裂缝;Cha 等^[10]使用 Faster R-CNN 架构检测混凝土和钢结构中的 5 种表面损伤类型。尽管上述研究证明了两级检测器可用于检测结构缺陷,但由于采用区域建议作为中间步骤,其检测速度不够理想。另一类是以单级检测器 SSD^[11]、YOLO^[12]系列为代表。由于 SSD 和 YOLO 都删除了生成区域建议步骤,并同时预测了多个边界框和类别概率,因此检测速度比两级检测器快。但是,SSD 算法的缺点明显,一是不能充分利用浅层的高分辨率特征图,二是候选框的尺寸比例需要人工根据经验设置。因此,本研究的目的是探讨先进的单级检测器 YOLOv3^[13]的适用性,用以识别桥梁表面的多种病害,并提高检测精度。

1 YOLOv3 算法介绍

目前在主流的目标检测网络中,YOLO 网络直接对图像中的目标进行回归检测,因此其检测速度比其他网络快。分析 YOLOv1 至 YOLOv3 网络的优缺点得出,初始 YOLOv1 网络检测精度差;YOLOv2^[14]在 YOLOv1 的基础上通过添加批量标准化、高分辨率的分类器、多尺度的训练等方法,在继续保持处理速度基础上,在预测更准确、速度更快、识别对象更多这 3 个方面进行了改进;YOLOv3 在 YOLOv2 的基础上取代了新的骨干网络,将单标签分类改进为多标签分类,采用了多尺度融合预测方法。YOLOv3 网络模型如图 1 所示,它的基本体系结构包括 3 个主要部分:特征提取网络层(DarkNet-53)、检测层和分类层。

特征提取网络层:它是被命名为 DarkNet-53 的特征提取算法。它由 53 个卷积层组成,每个卷积层都包含 1 个归一化层,每层的激活函数是 Leaky ReLU。它借鉴了残差网络^[15](ResNet)的做法,在一些层之间建立快捷链路,通过快捷链路跃层连接可以解决随着网络的逐步深入导致模型难以优化的问题,减小梯度爆炸风险,加强网络的学习能力并利用到更多的图像浅层特征信息。特征图的尺寸在某些卷积层之间减小了 2 倍,该算法总共将尺寸减小了 32 倍。为了 YOLOv3 网络的下一步工作,特征提取网络部分将输出 3 个不同尺寸的特征图作为下一个网络模块的输入,其尺寸分别比原始图像尺寸减小 8 倍、16 倍和 32 倍。

检测层:卷积核大小为 1×1 和 3×3 的卷积层交替出现,最后一层的卷积核大小必须为 1×1 。最小输入尺度的特征图仅在检测层进行处理。特征提取网络层输出的另外 2 个尺度的特征图在被发送到检测层之前,先与已经处理过的较低维映射进行拼接,然后输入检测层。

分类层:以检测层生成的包含多尺度特征信息的融合特征为输入,通过卷积核大小为 3×3 和 1×1 的卷积层产生模型的最终特征输出,且最后一层的通道数(Filters)大小为

$$\text{Filters} = 3 \times (4 + 1 + \text{classes}), \quad (1)$$

式中:3 代表 3 个预测尺度,4 和 1 分别为网络最终输出检测目标类别归一化后的中心坐标 (x, y, w, h) 和置信度,classes 代表检测目标类别数。本研究中检测类别有 4 类,因此 $\text{Filters} = 27$ 。

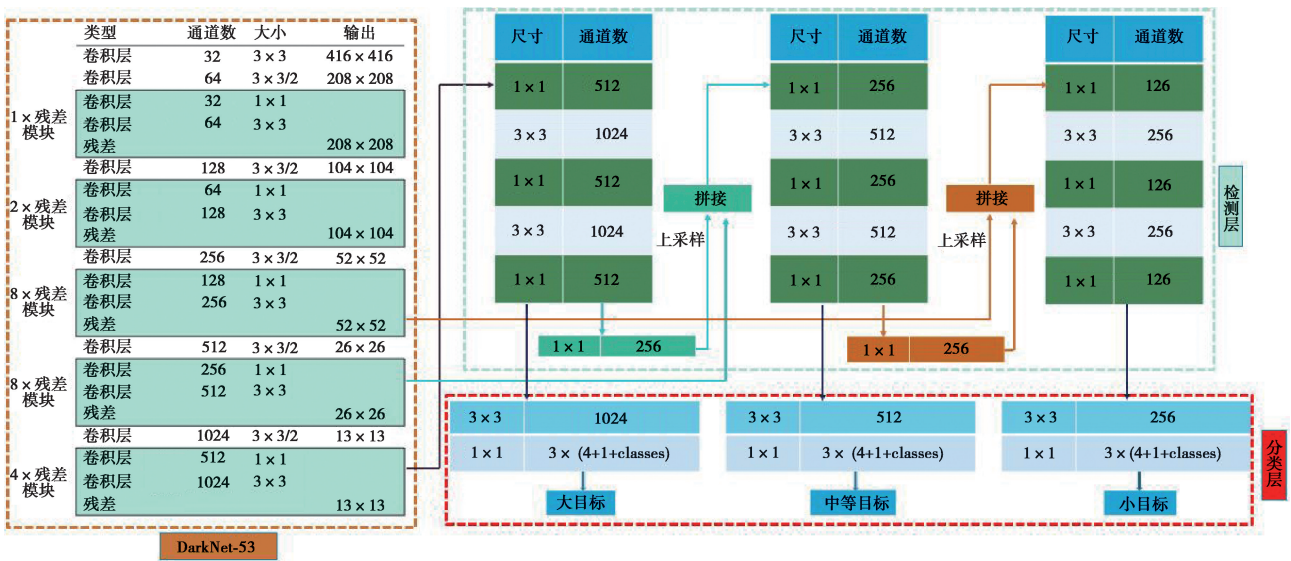


图 1 YOLOv3 网络结构图
Fig. 1 YOLOv3 network structure diagram

2 网络的改进

2.1 改进检测层网络结构

空间金字塔池化^[16] (SPP-Net)对图片进行不同的分块池化,把每个块中提取出的一个特征图作为一个维度,确保最后得到特征图的维度一致,从而解决信息丢失和尺度不统一的问题。此外,特征图中的通道数被扩展以提供有效的全局信息,因此它具有更强的细节特征描述能力,对不同类型目标的检测精度更高。病害图像的预处理和多尺度预测中可能产生信息丢失或者尺度不一致,从而影响最终的检测效果。因此本研究中参考空间金字塔池化方法,在 YOLOv3 检测层中融合空间金字塔池化的操作,不仅解决了在检测层中输入特征图的尺寸变化和图像失真的问题,也实现了特征图的局部特征和全局特征更好地融合。如图 2 所示,选择 3 个不同尺度($n_1 < n_2 < n_3$)的最大池化操作与预测分支的输入拼接构成空间金字塔池化。

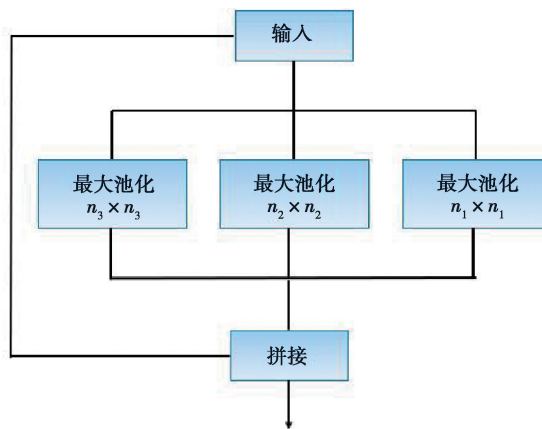


图 2 空间金字塔池化模块
Fig. 2 Spatial pyramid pooling module

虽然特征映射具有丰富的特征信息,但经过深层和浅层特征融合后存在一定的冗余。因此在空间金字塔池化结构中, n_1 、 n_2 和 n_3 这 3 个尺度的池化层分别对应 3 个不同程度的去冗余操作。尺度为 n_1 的池化层的池化区域最小,能够对特征图上的局部特征进行去冗余;尺度为 n_3 的池化层的池化区域最大,能够对特征图上的全局特征进行去冗余;尺度为 n_2 的池化层属于过渡操作,保留介于局部和全局之间的特征。完成冗余操作后,使用拼接操作聚合所有特征映射。经过去冗余再聚合操作的特征图的信息更有层次且表征性更强,检测器能够根据全局特征对缺陷进行定位,再根据局部特征对目标病害的类别进行判定。为实现局部特征和全局特征的特征图级别的融合,空间金字塔池化结构最大的池化核(n_3)要尽可能接近等于需要池化的特征图的大小。因此在 YOLOv3 预测大目标的检测层分支中嵌入空间金字塔池化,设置 $n_1=5$, $n_2=9$, $n_3=13$ 。

2.2 改进特征提取网络结构

DenseNet 是由 Huang 等^[17]提出的一种激进的、密集的连接机制,如图 3 所示。从功能上看,通过特征重用和旁路设置不仅能大幅减少网络参数的数量,而且能在一定程度上缓解梯度消失的问题。密集连接网络可以改善整个网络的信息流和梯度,其原理是:假设输入为 X_0 ,网络的每一层实现一个非线性变换函数 $H_i(\cdot)$, i 表示第 i 层,第 i 层的输出表示为 X_i ,则:

$$X_i = H_i([X_0, X_1, \dots, X_{i-1}]) \quad (2)$$

式中: $[\]$ 代表拼接; $H_i(\cdot)$ 是一个组合操作,包括一系列批归一化(BN)、激活函数(ReLU)、池化(Pooling)和卷积(Conv)操作。

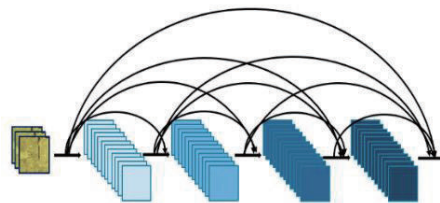


图 3 DenseNet 的密集连接机制

Fig. 3 Dense connection mechanism of DenseNet

密集连接网络通常包含多个密集模块,1 个密集模块由 n 个密集层组成,密集层中的非线性组合函数 $H_i(\cdot)$ 采用 BN+ReLU+ 3×3 Conv 的结构。与常见的活化机制不同,密集层使用了预活化机制,其批归一化层和激活函数层在卷积层之前,先进行激活操作,然后进行 3×3 卷积输出特征映射。其结构如图 4 所示。

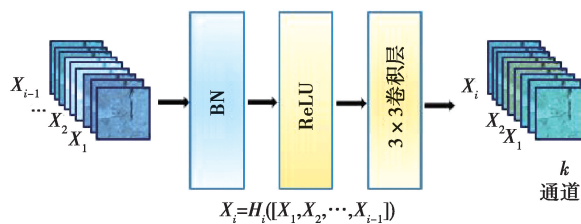


图 4 密集模块中的非线性转换结构

Fig. 4 Non-linear transformation structure in dense module

假设一个密集模块的输入 X_0 维度为 m ,每个密集层输出 k 个特征图。根据密集连接网络的原理,第 n 个密集层的输入是 $m + (n-1)k$ 个特征图,所以直接进行 3×3 卷积运算会带来大量的计算量,造成网络的负担。这时可以使用瓶颈结构来减少网络的计算量,主要方法是在原来的密集模块上增加一个 1×1 卷积层来减少特征的数量,即 BN+ReLU+ 1×1 Conv+BN+ReLU+ 3×3 Conv。在构建的瓶颈结构的密集层中,首先通过 1×1 卷积层可得 $2k$ 个特征图,然后通过 3×3 卷积层输出 k 个特征图。其结构如图 5 所示。

在使用瓶颈结构方法的基础上,为进一步减少网络计算量,在每两个密集模块之间使用卷积核大小分别为 1×1 的卷积层和 2×2 的平均池化层组成过渡层。假设一个密集模块的输出特征图数量为 p ,使用输出

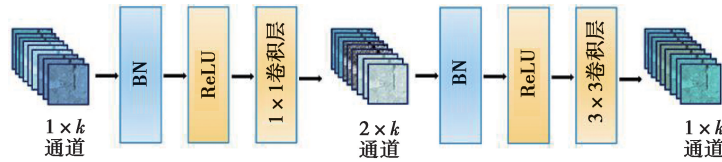


图 5 含有瓶颈结构的密集模块结构

Fig. 5 Dense module structure with a bottleneck structure

连接的过渡层输出 θp 个特征图,其中 $\theta(0 < \theta \leq 1)$ 表示压缩系数。在本研究中构建的过渡层中 $\theta = 0.5$,因此经过过渡层压缩操作后,输入到下一个密集模块的特征图的数量与尺度减少一半。

将 YOLOv3 网络与所构造的密集连接网络结合起来,提出了一种新型 YOLOv3 网络结构,如图 6 所示。为了平衡检测速度和准确率,将原网络的输出尺寸为 208×208 和 104×104 的残差模块保留,将输出为 52×52 、 26×26 和 13×13 的 3 组残差模块替换为密集模块。

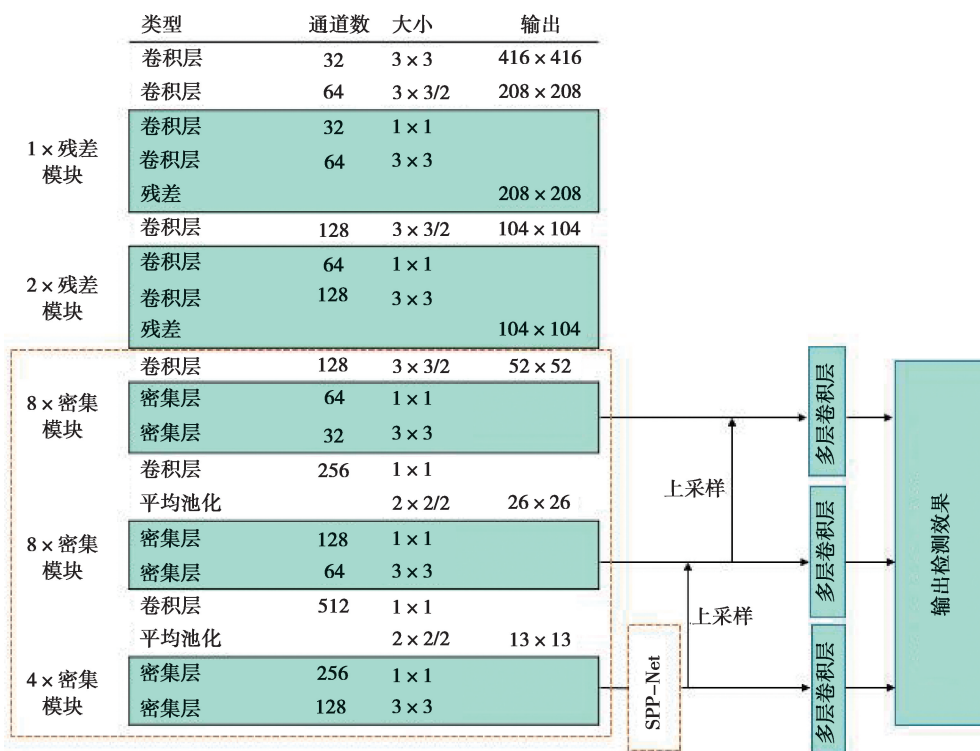


图 6 改进的 YOLOv3 网络结构图

Fig.6 Improved YOLOv3 network structure diagram

3 实验结果及分析

3.1 实验环境与数据集

实验环境为 Windows 操作系统,AMD Ryzen 5 3600 (CPU),32 GB 随机存取内存(RAM),RTX 2060-SUPER(GPU),16 GB 显示内存,深度学习框架为 pytorch。网络的初始化参数如表 1 所示。

表 1 初始化参数

Table 1 Initialization parameters

输入图片大小	批量大小	动量系数	学习率调度器	学习率	训练轮数
416×416	24	0.9	Step	0.01	500

数据集 CODEBRIM (COncrete DEfect BRidge Image dataset^[18], 混凝土缺陷桥图像数据集) 是 2019 年公布的, 用于计算机视觉和机器学习中的多目标病害的桥梁混凝土缺陷检测, 该数据集含有风化 (efflorescence)、钢筋外露 (exposed bars)、腐蚀 (corrosion stain) 和脱落 (spallation) 4 种桥梁病害。由于该数据集样本数量不足, 利用数据增强技术对数据进行样本扩充, 如表 2 所示。扩充后的部分图片存在病害标签失效现象, 其失效形式为原图像中的病害标签尺寸超过图像本身尺寸, 造成输入网络中归一化的值不属于 0~1 的范围, 因此需要对增强后的图片进行排查, 删除失效的图像。

表 2 数据增强方式

Table 2 Data enhancement method

增强方式	原始图片	旋转 180°	调整亮度	高斯模糊	总量
数量	1 052	978	1 048	1 037	4 115

扩充后的图片效果如图 7 所示。

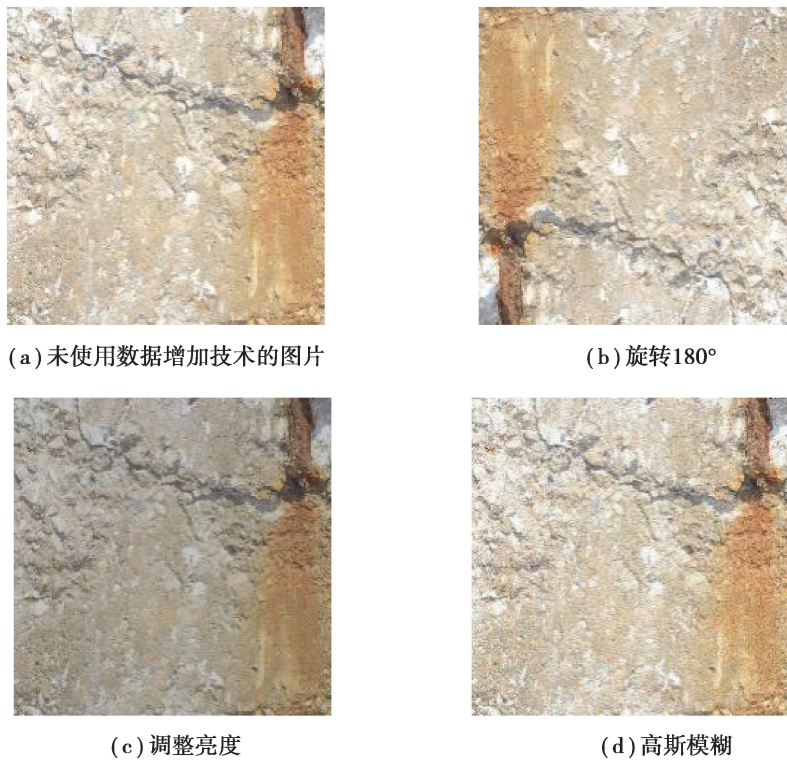


图 7 使用数据增强的图片对比

Fig. 7 Non-enhanced and data-enhanced images

3.2 评价指标

对目标检测算法, 平均准确率 (mAP) 是评价模型性能的标准度量。对二元分类问题, 根据其标记类和预测类的组合, 判断结果可分为 4 类: 真正例 (TP)、假正例 (FP)、真负例 (TN) 和假负例 (FN)。分类结果的混淆矩阵如表 3 所示。

表 3 二元分类的混淆矩阵
Table 3 Confusion matrix for binary classification

标记	预测	
	正 (Positive)	负 (Negative)
正(Positive)	TP	FP
负(Negative)	FN	TN

精确率和召回率计算公式如下:

$$\text{精确率}(P) = \frac{TP}{TP + FP} = \frac{TP}{\text{全部预测}}, \tag{3}$$

$$\text{召回率}(R) = \frac{TP}{TP + FN} = \frac{TP}{\text{全部标记框}}. \tag{4}$$

精确率是一个模型真实预测目标总数与所有预测目标总数比率,召回率是真实预测目标总数与数据集中目标总数的比率。以精确率为横轴,以召回率为纵轴,可以得到精确的校准曲线,称为 $P-R$ 曲线。 $P-R$ 曲线与坐标轴围成的面积为每个类的精度值(AP),mAP 则是计算所有类的 $P-R$ 曲线下面积的平均值(该面积通过黎曼求和来计算)。公式如下:

$$AP = \frac{\sum_{i=1}^N P_i}{N}, \tag{5}$$

$$mAP = \frac{\sum_{j=1}^M AP}{M}. \tag{6}$$

式中: N 为用于测试的图像数量, M 为检测的类别数。

3.3 实验结果与分析

以改进 YOLOv3 网络检测层 (YOLOv3-a),改进特征提取网络层 (YOLOv3-b1:使用瓶颈结构;YOLOv3-b2:使用瓶颈结构和过渡层)以及同时加入以上 2 种改进的网络 (YOLOv3-a-b1;YOLOv3-a-b2)与原 YOLOv3 进行对比实验。实验结果如表 4 所示。

表 4 改进的 YOLOv3 和 YOLOv3 实验结果详情
Table 4 Improved YOLOv3 and YOLOv3 experimental results

算法	AP/%				mAP/%	训练时间/h
	风化	钢筋外露	腐蚀	脱落		
YOLOv3	78.3	86.7	83.3	82.1	82.6	15.286
YOLOv3-a	80.3	88.5	85.1	83.0	84.2	15.651
YOLOv3-b1	79.7	88.9	84.9	84.4	84.4	12.565
YOLOv3-b2	80.5	89.0	84.5	83.9	84.5	10.044
YOLOv3-a-b1	81.2	89.1	85.4	84.1	85.0	12.964
YOLOv3-a-b2	81.4	89.2	85.6	86.0	85.6	10.212

通过上表可以看出,改进的 YOLOv3 比原 YOLOv3 在桥梁表观病害检测上的平均准确率有所提高。其中,YOLOv3-a 的平均准确率提升了 1.6%,由于融合空间金字塔池化操作,增加了网络计算参数量,因此训练时间增加 2.4%;YOLOv3-b1 的平均准确率提升了 1.8%,由于替换了原 YOLOv3 特征提取网络的最后 3 个残差模块,网络计算参数量减少 50%,因此训练时间减少 17.8%;而在 YOLOv3-b1 上做进一步改进的 YOLOv3-b2,与原 YOLOv3 相比,其平均准确率提升了 1.9%,网络计算参数量减少 59%,因此训练时间减

少 34.3%。融合两种改进 YOLOv3-a-b1 的平均准确率提升 2.4%，网络计算参数量减少 49%，导致训练时间减少 15.2%；YOLOv3-a-b2 的平均准确率的提升最大，达到 3.0%，网络计算参数量减少 57%，导致训练时间减少 33.2%。分析结果表明替换 YOLOv3 特征提取网络的最后 3 个残差模块为密集模块结构后，不仅使深层特征和浅层特征更好地融合，提高了网络的特征复用能力，还大大减少了网络模型的训练时间；添加空间金字塔池化结构虽增加了模型的复杂度，导致训练时间增加 2.4%、3.1% 和 1.7%，但平均准确率得到提升。分析结果证明了对 YOLOv3 的改进是有效的。

为进一步验证本研究中提出的模型在桥梁表面病害检测上优于其他目标检测算法，使用同样的桥梁表面病害数据集对 YOLOv3-a-b2 和 Faster R-CNN、SSD、RetinaNet 等目标检测算法进行病害识别任务的训练与测试。实验结果如表 5 所示。

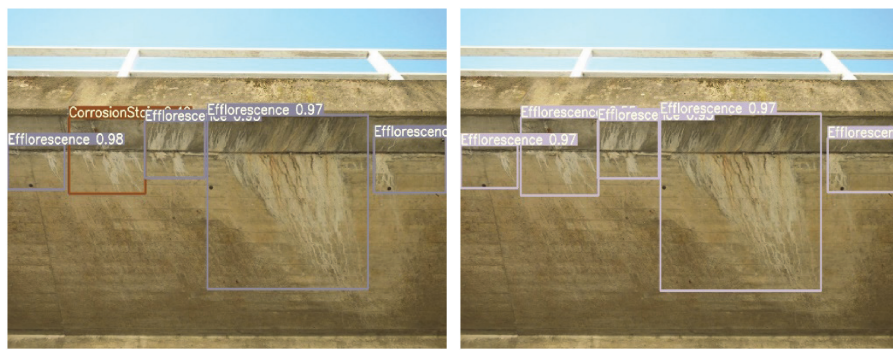
表 5 不同算法检测结果对比
Table 5 Comparison of detection results of different algorithms

算法	AP/%				mAP/%	训练时间/h
	风化	钢筋外露	腐蚀	脱落		
Faster R-CNN	72.0	80.3	77.2	73.7	75.8	22.852
SSD	69.7	77.5	74.3	73.6	73.6	16.327
RetinaNet-Resnet50	70.4	77.9	75.1	72.2	73.9	19.539
YOLOv3-a-b2	81.4	89.2	85.6	86.0	85.6	10.212

从表 5 可以看出无论是两级检测器 Faster R-CNN 还是单级检测器 SSD，其平均准确率均低于改进模型，且训练时间更长。证明了改进的 YOLOv3 在桥梁表面病害检测上有更好的检测效果。

3.4 图片检测效果

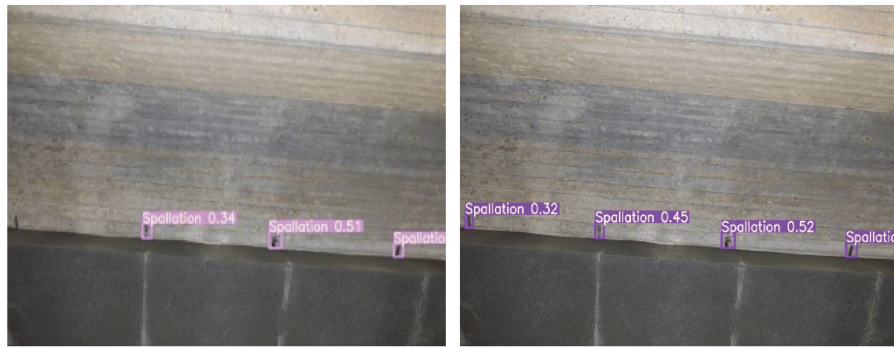
随机抽取测试集中的 3 张图片进行测试，结果如图 8 所示。



(a) 左:检测出腐蚀、风化; 右:检测出风化



(b) 左:检测出腐蚀; 右:检测出腐蚀、脱落



(c)左:检测出脱落;右:检测出脱落

图 8 YOLOv3(左)和 YOLOv3-a-b2(右)的检测结果

Fig. 8 YOLOv3 (left) and YOLOv3-a-b2(right) test results

从图 8(a)对比图片看出,左图中出现了将风化误检为腐蚀,而右图准确地检测出风化这一病害类型;图 8(b)中左图对病害部位的检测结果仅为腐蚀,而右图对病害部位的检测结果不仅有腐蚀,还有脱落(由于检测为脱落的结果准确率比腐蚀的结果准确率高,腐蚀的检测结果被覆盖了一部分,在右图中能看见腐蚀的结果准确率为 0.33);图 8(c)中左图出现了对脱落部位的漏检,右图则识别了全部脱落部位。从对检测结果的分析可以看出本研究中的模型的检测效果优于原 YOLOv3 检测效果,而且还降低了对目标病害的误检率与漏检率。

4 结 论

针对目前通用目标检测算法在桥梁表面病害检测上精度低、误检率和漏检率高的缺点,提出了基于单级检测器 YOLOv3 的桥梁表面病害检测方法。将本研究中构造的密集连接网络嵌入 YOLOv3 特征提取网络后,不仅增强了桥梁病害特征在网络层之间的传播和利用效率,使检测层得到更强的语义信息,为目标分类任务提供了更强的特征支持,还减少了网络的训练时间。引入空间金字塔池化结构使浅层与深层的特征信息更好地融合,提升了目标分类与定位的能力。经过实验结果分析,改进的网络模型对桥梁病害有更高的检测精度。下一步工作将研究基于 YOLOv3 改进算法的智能爬壁机器人进行病害检测实际应用,以机器人自主的目视检查代替以人眼目视的检查。

参考文献:

- [1] 高迪,付鹏飞. 探析道路桥梁施工中检测技术的应用[J]. 科技展望, 2014 (10):66.
Gao D, Fu P F. Analyze the application of detection technology in road and bridge construction [J]. Science and Technology Outlook, 2014 (10): 66. (in Chinese)
- [2] 滕亮. 桥梁结构损伤识别指标比选及损伤程度识别方法研究[D]. 长春: 吉林大学, 2014.
Teng L. Research on comparison of damage detection indices and damage severity identification for bridge structure[D]. Changchun: Jilin University, 2014. (in Chinese)
- [3] 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.
- [4] Prasanna P, Dana K, Gucunski N, et al. Computer-vision based crack detection and analysis[C]// Proceedings of SPIE 8345, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2012, San Diego, California, USA. Society of Photo-Optical Instrumentation Engineers (SPIE), 2012: 834542.
- [5] 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.
- [6] 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, June 23-28, 2014, Columbus, OH, USA. IEEE, 2014: 580-587.
- [7] Uijlings J R R, Sande K, Gevers T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [8] 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.
- [9] Kim I H, Jeon H, Baek S C, et al. Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle[J]. Sensors, 2018, 18(6): 1881.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] Redmon J, Farhadi A. YOLOv3: an incremental improvement [EB/OL]. 2018-04-08[2020-12-06]. <https://arxiv.org/abs/1804.02767>.
- [14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6517-6525.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [16] Purkait P, Zhao C, Zach C. SPP-net: deep absolute pose regression with synthetic views[EB/OL]. 2017-12-09[2020-12-06]. <https://arxiv.org/abs/1712.03452>.
- [17] Huang G, Liu Z, Laurens V D M, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 4700-4708.
- [18] Mundt M, Majumder S, Murali S, et al. Meta-learning convolutional neural architectures for multi-target concrete defect classification with the COConcrete 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.