

基于改进 YOLO 的双网络桥梁表面病害快速检测算法

彭雨诺^{1,2} 刘敏^{1,2} 万智³ 蒋文博^{1,2} 何文轩^{1,2} 王耀南^{1,2}

摘要 桥梁表面病害检测是确保桥梁安全的关键步骤。然而,桥梁表面病害类型多样,不同病害间外观差异显著且病害之间可能发生重叠,现有算法无法实现快速且准确的桥梁多病害检测。针对这一问题,对 YOLO (You only look once) 进行了改进,提出了 YOLO-lump 和 YOLO-crack 以提高网络检测多病害的能力,进而形成基于双网络的桥梁表面病害快速检测算法。一方面,YOLO-lump 在较大的滑动窗口图像上实现块状病害的检测。在 YOLO-lump 中,提出了混合空洞金字塔模块,其结合了混合空洞卷积与空间金字塔池化,用于提取稀疏表达的多尺度特征,同时可以避免空洞卷积造成的局部信息丢失;另一方面,YOLO-crack 在较小的滑动窗口图像上实现裂缝病害的检测。在 YOLO-crack 中,提出了下采样注意力模块,利用 1×1 卷积和 3×3 分组卷积分别解耦特征的通道相关性和空间相关性,可以增强裂缝在下采样阶段的前景响应,减少空间信息的损失。实验结果表明,该算法能够提高桥梁表面病害检测的精度,同时可实现病害的实时检测。

关键词 桥梁表面病害检测,深度卷积神经网络,空间金字塔模块,注意力机制

引用格式 彭雨诺,刘敏,万智,蒋文博,何文轩,王耀南. 基于改进 YOLO 的双网络桥梁表面病害快速检测算法. 自动化学报, 2022, 48(4): 1018–1032

DOI 10.16383/j.aas.c210807

A Dual Deep Network Based on the Improved YOLO for Fast Bridge Surface Defect Detection

PENG Yu-Nuo^{1,2} LIU Min^{1,2} WAN Zhi³ JIANG Wen-Bo^{1,2} HE Wen-Xuan^{1,2} WANG Yao-Nan^{1,2}

Abstract Surface defect detection is a critical step to ensure bridge safety. However, there are various types of bridge surface defects, different defects have a wide range of variation in appearance and generally overlap with each other. The existing algorithms cannot efficiently and precisely detect such defects. To solve this problem, we improve the YOLO (You only look once) to enhance the performance of the network to detect multiple defects, YOLO-lump and YOLO-crack are proposed to form a dual deep network for fast bridge surface defect detection. On the one hand, the YOLO-lump can realize the detection of the lump defects on larger sub-images, by employing a hybrid dilated pyramid module based on the hybrid dilated convolution and the spatial pyramid pooling to extract multi-scale features and to avoid losing local information caused by the dilated convolution. On the other hand, the YOLO-crack can realize the detection of the crack defects on smaller sub-images, by proposing a downsampling attention module which uses the 1×1 convolution and the 3×3 group convolution to respectively map cross-channel correlation and spatial correlation of features, enhancing the foreground response of the crack in the downsampling stage and reducing the loss of spatial information. Experimental results show that the proposed algorithm can improve the detection accuracy of the bridge surface defects and realize real-time detection.

Key words Bridge surface defect detection, deep convolutional neural network, spatial pyramid module, attention mechanism

Citation Peng Yu-Nuo, Liu Min, Wan Zhi, Jiang Wen-Bo, He Wen-Xuan, Wang Yao-Nan. A dual deep network based on the improved YOLO for fast bridge surface defect detection. *Acta Automatica Sinica*, 2022, 48(4): 1018–1032

收稿日期 2021-03-06 录用日期 2021-11-17

Manuscript received March 6, 2021; accepted November 17, 2021

国家自然科学基金 (61771189, 62073126, 62027810), 湖南省自然科学基金杰出青年基金 (2020JJ2008), 湖南省交通运输厅科技计划项目 (201734, 202138) 资助

Supported by National Natural Science Foundation of China (61771189, 62073126, 62027810), Natural Science Foundation of Distinguished Young Scholars of Hunan Province (2020JJ2008), Science and Technology Progress and Innovation Plan of Department of Transportation of Hunan Province (201734, 202138)

本文责任编辑 胡清华

Recommended by Associate Editor HU Qing-Hua

1. 湖南大学电气与信息工程学院 长沙 410082 2. 湖南大学机器人视觉感知与控制技术国家工程研究中心 长沙 410082 3. 湖南桥梁康智能科技有限公司 长沙 410021

1. College of Electrical and Information Engineering, Hunan University, Changsha 410082 2. National Engineering Research Center for Robot Visual Perception and Control Technology, Hunan University, Changsha 410082 3. Human Qiaokang Intelligent Technology Company Limited, Changsha 410021

中国桥梁数量位居世界首位^[1], 大量的桥梁改善了交通状况, 同时, 为保障桥梁的安全运营, 需要对桥梁进行定期的检查和养护. 桥梁在建造以及使用的过程中, 受到施工材料、建筑工艺、极端天气、车辆超载等因素的影响, 不可避免地出现结构性或非结构性的损伤, 进而形成蜂窝、漏筋、孔洞、裂缝等表观病害. 若是能够及时地对桥梁进行检查并修复损伤, 将大大减少桥梁的维护成本, 延长桥梁的使用寿命. 到目前为止, 桥梁健康状况的评估大多仍是由人工目视进行^[2], 工人在升降装置的协助下, 到达桥梁各个位置对病害进行测量和统计. 然而, 这种传统方法检测效率低下且结果受到工人主观经验的影响, 因此, 实现桥梁表观病害的自动化检测具有重要意义.

随着人工智能技术的发展, 近年来越来越多的机器人被用于桥梁的自动化检查中, 文献 [3-4] 利用无人机获取桥梁和钢索的高清图像并进行病害的检测, 文献 [5] 使用移动机器人对桥梁进行检查, 在文献 [6] 开发了一种水下机器人对桥墩的水下部分进行检测. 由湖南桥康智能科技有限公司研发的轻量化桥梁智能检测机器人 BIR-X-LITE, 可海量地采集桥梁底部的高分辨率表观图像数据. 为了高效且准确地评估桥梁健康状况, 需要利用桥梁表观病害检测算法对图像进行快速有效的分析. 过去许多传统病害检测算法中, 边缘检测器^[7]和阈值检测方法^[8-9]被用于裂缝检测. 文献 [10] 利用基于局部熵的阈值算法检测桥梁蜂窝病害. 如果病害具有高对比度以及良好的连续性, 传统算法可实现高精度的检测. 但在实际应用中, 受到光照强度变化、噪声、背景干扰等因素的影响, 部分病害图像特征不明显, 从而导致传统算法无法实现准确地病害检测.

近年来, 深度学习在计算机视觉等领域表现出了良好的性能^[11-14], 利用该技术在复杂条件下对桥梁表观病害图像进行检测已成为现实. 在文献 [15] 利用卷积神经网络 (Convolutional neural networks, CNN) 对 16×16 像素的图像进行分类, 来判断该图像中是否存在裂缝病害. 而 Kim 等^[16]将区域卷积神经网络 (Region-CNN, R-CNN) 与形态学后处理相结合, 以检测和分割桥梁裂缝病害. 但是大部分深度学习方法只针对一种类型的病害进行检测, 不能对桥梁的健康状况进行全面地评估.

随着多尺度检测网络的发展, 桥梁多病害检测算法被相继提出. Zhang 等^[17]利用 YOLO (You only look once) 网络实现了对桥梁裂缝、剥落、蜂窝和漏筋 4 种病害的检测. 该方法在 YOLOv3 的基础上, 引入了迁移学习方法和批量正则化等方法.

Li 等^[18]基于 Dense-net 网络和迁移学习方法, 提出了新的全卷积神经网络 (Fully convolutional networks, FCN), 对裂缝、蜂窝、风化和孔洞 4 种病害进行检测和分割. 然而, 在上述网络训练所使用的数据集中, 各类病害从不相等的距离被拍摄, 从而缩小了不同病害间的大小差异.

在实际应用中, 如 Yang 等^[19]、Mundt 等^[20]和 Hühthwohl 等^[21]指出, 不同类型的病害外观差异显著, 且病害之间可能发生重叠, 利用单个网络完成所有桥梁表观病害的检测或分类是一个巨大的挑战. 针对上述问题, 文献 [19] 通过引入距离加权系数对蜂窝和裂缝病害图像进行缩放对齐; 文献 [20] 在 224×224 像素步长的滑动窗口上对桥梁多病害进行分类, 但可能导致大型病害整体结构信息的丢失; 文献 [21] 提出了三阶段分类器, 可将不健康的桥梁区域分为特定病害类型, 以不同步长的滑动窗口对图像进行了多次的分类, 提高了分类的准确性.

不同类型的病害外观差异显著, 且病害之间可能发生重叠, 加上光照强度变化、噪声、背景干扰等影响, 现有的目标检测算法无法在固定步长的滑动窗口上实现桥梁多病害的快速鲁棒检测. 针对上述问题, 本文对 YOLO^[22-23]进行了改进, 提出了 YOLO-lump 和 YOLO-crack, 以提高网络检测多病害的能力, 进而形成基于双网络的桥梁表观病害快速检测算法. 1) YOLO-lump 在 1280×1280 像素步长的滑动窗口上实现块状病害 (包括蜂窝、漏筋和孔洞病害) 的检测, 针对长宽比和大小多变的块状病害, 提出了混合空洞金字塔模块, 采用空间金字塔池化^[24]的并行采样结构, 并在不同的通道上利用混合空洞卷积^[25]提取不同尺度的特征. 该模块可用于提取稀疏的多尺度特征, 能在不显著增加计算复杂度的情况下扩展网络的感受野, 同时可以避免空洞卷积造成的局部信息丢失. 2) YOLO-crack 在 640×640 像素步长的滑动窗口上实现裂缝病害的检测, 由于常规的下采样方法可能会导致裂缝像素损失, 受文献 [26-27] 启发, 提出了下采样注意力模块, 其在额外的下采样通道中, 利用 1×1 卷积和 3×3 分组卷积分别解耦特征的通道相关性和空间相关性^[28], 以找到特征中重要的信息, 然后叠加到原始的下采样结果上. 该模块能够增强下采样过程中裂缝的前景响应, 有利于裂缝细节特征的保留. 为了提高网络的检测速度, 基于文献 [29-30] 对 YOLO-crack 进行了轻量化的调整. 此外, 为了进一步提升本文算法的性能, 条件生成式对抗网络^[31]被用于生成新的训练样本以提升病害检测的鲁棒性. 而 focal loss 损失函数^[32]被用于解决正负样本不平衡的问题以

优化网络训练过程. 本文提出的算法可配合 BIR-X-LITE 机器人实现桥梁多病害的快速鲁棒检测, 算法整体流程如图 1 所示. 综上所述, 本文主要贡献如下:

1) 建立了桥梁表观图像数据库, 共有 169 621 张高分辨率图像, 包括蜂窝、漏筋、孔洞和裂缝 4 种常见病害类型.

2) 提出了混合空洞金字塔模块, 结合了混合空洞卷积和空间金字塔池化, 能有效扩展网络的感受野并保护特征连续性, 提高网络的多尺度检测性能.

3) 提出了下采样注意力模块, 利用 1×1 卷积和 3×3 分组卷积分别解耦特征的通道相关性和空间相关性, 能加强下采样阶段病害的前景响应, 减少空间信息损失.

4) 根据桥梁病害的实际情况, 提出了双网络桥梁表观病害快速检测算法, 其双网络分别为 YOLO-lump 和 YOLO-crack 网络, 配合 BIR-X-LITE 机器人可实现桥梁多病害的检测. 实验结果证明, 在本文数据集上, 该算法性能优于其他最新方法, 且检测一张 5120×5120 像素的图像仅花费 0.995 秒.

1 基于 GAN 网络的数据增广

桥梁健康状况评估是多阶段的工作^[33], 由于桥梁机器人可海量地采集桥梁底部的高分辨率表观图

像数据, 为了提高桥梁健康状况评估效率, 首先需要利用检测网络快速地筛选出病害区域. 因此, 检测阶段的性能会极大地影响桥梁健康状况评估的结果. 为了全面准确地掌握桥梁健康状况, 需要实现鲁棒性高的病害检测. 在图像采集的过程中, 受到光照强度变化、噪声、机械振动等因素的影响, 少部分图像成像质量不佳从而形成困难样本, 导致病害检测性能下降. 为了增强网络的鲁棒性, 提高网络对于困难样本的检测能力, 利用生成式对抗网络^[34] (Generative adversarial networks, GAN) 自动生成困难样本以对训练数据进行增广. 近年来, GAN 网络被广泛应用于各类图像处理任务中. 刘建伟等^[35-39]通过实验证明, GAN 网络生成的数据具有一定的实用性.

利用条件 GAN 网络^[31] 对桥梁病害图像进行数据增广, 该网络可以将桥梁表观病害语义分割图像合成为真实病害图像. 原始数据集包含 597 张块状病害图像和 516 张裂缝病害图像以及对应的语义分割图像. 网络的训练参数设置如下: 初始学习率设置为 0.0001, 动量设置为 0.5, 使用的优化器为 Adam, 批量大小设置为 1, 训练迭代次数为 100 轮. 在新生成的病害图像中, 挑选出 444 张块状病害图像以及 427 张裂缝病害图像, 部分生成图像见图 2.

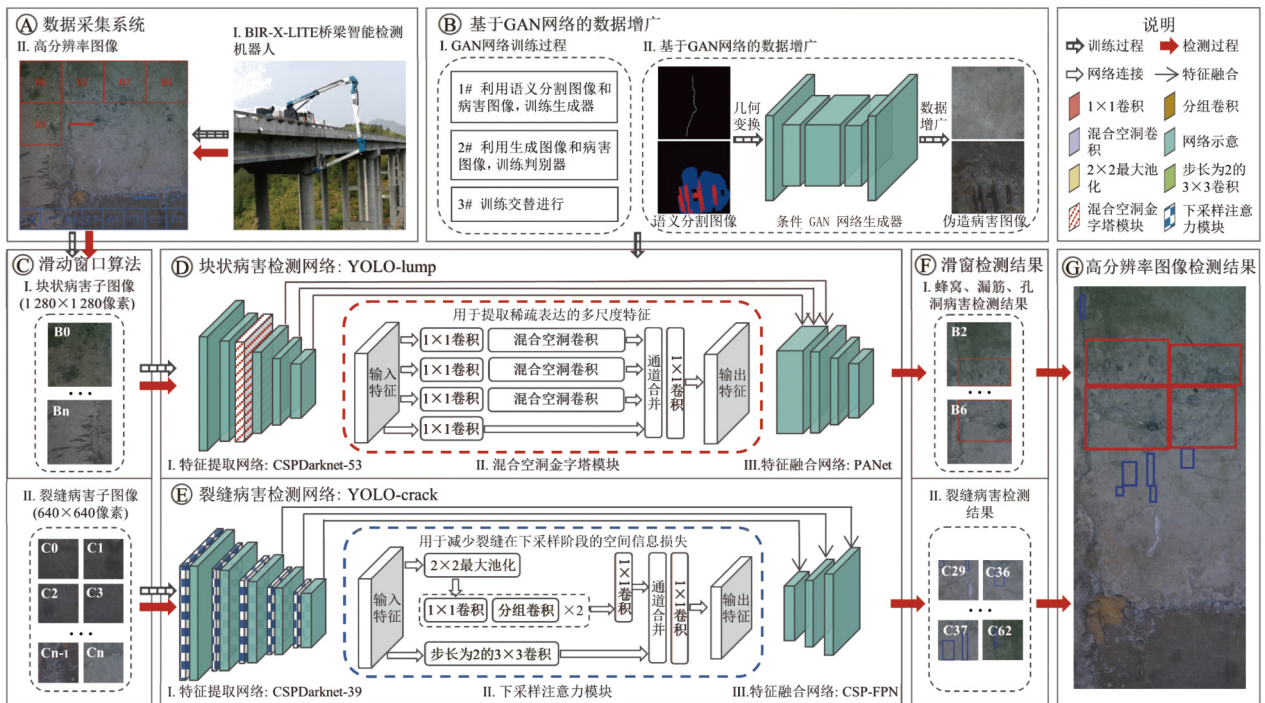


图 1 双网络桥梁表观病害快速检测算法整体框架

Fig. 1 Overview of the dual deep network for fast bridge surface defect detection

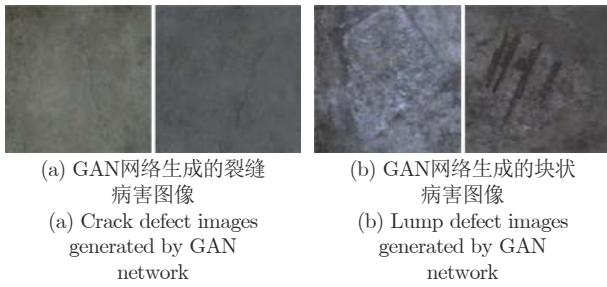


图 2 GAN 网络生成的桥梁表面病害图像

Fig.2 Bridge surface defect images generated by GAN network

2 双网络桥梁表面病害快速检测算法

2.1 混合空洞金字塔模块

桥梁病害的长宽比及大小差异显著, 其中, 漏筋病害特征细长, 而蜂窝病害的尺寸几乎可以任意变化, 不同的蜂窝病害其面积差异可达到 10 倍以上. 由于 YOLO 需要根据聚类结果先行设定锚框, 然后网络在特征图上以设定好的锚框对目标进行搜索与检测. 桥梁病害长宽比及大小的差异对聚类结果的准确性造成了影响, 当目标的长宽比及大小与聚类结果相差较大时, 会影响检测网络的定位准确性. 为解决上述问题, 需要增强网络的多尺度检测性能, 使网络能够更好地提取和识别不同桥梁病害的特征.

在过去工作中, 空洞空间金字塔池化^[40](Atrous spatial pyramid pooling, ASPP) 被用于解决分割领域物体尺度变化大的问题, 将空洞卷积添加到空间金字塔池化的框架中, 以有效地提取多尺度上下文信息, 但是空洞卷积的膨胀系数设置过大, 不利于桥梁病害的检测. 为了对桥梁病害多尺度特征进行提取, 将 ASPP 中高膨胀系数的空洞卷积分解为多个低膨胀系数的空洞卷积, 空洞卷积膨胀系数的设置与文献 [25], 这样可以避免空洞卷积级联后造成的局部信息丢失. 为了减少计算量, 在空洞卷积之前额外加入了 1×1 卷积来减少特征通道数量. 另一方面, 由于 1×1 卷积后引入了 Mish 激活函数^[41], 可以增强模块的非线性表达, 提升其特征提取能力. 本文提出的混合空洞金字塔模块如图 3 所示, 其中 r 代表空洞卷积的膨胀系数.

在混合空洞金字塔模块中, 输入特征被采样到 4 条相互独立的通道之中, 组成空间金字塔进行并行采样. 首先, 经过 1×1 卷积将特征通道数减少至输入特征通道数的 $1/4$. 在第 1 条通道内, 特征图不进行额外的处理. 而在另外 3 条通道内, 通过不同

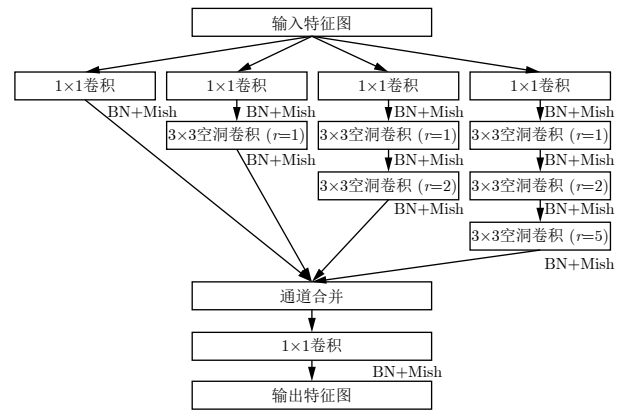


图 3 混合空洞金字塔模块

Fig.3 The hybrid dilated pyramid module

数量和膨胀系数的空洞卷积叠加, 提取包含不同感受野的特征图. 从整体上看, 每条特征提取通道内的空洞卷积都只与输入特征的一部分相连接, 这有助于提取特征级的稀疏信息, 提高特征的多样性. 最后, 将 4 条通道中的特征图进行合并, 然后经过 1×1 卷积, 得到稀疏的多尺度输出特征. 此外, 批量归一化和 Mish 激活函数被添加到每一次卷积操作之后.

2.2 下采样注意力模块

在深度神经网络中, 常使用步长为 2 的 3×3 卷积实现图像的下采样, 这在一定程度上丢失了上下文中包含的特征, 造成特征图分辨率的降低和病害空间信息的损失. 为了减少下采样过程中的信息损失, 利用注意力机制加强病害的前景响应.

挤压与激发 (Squeeze and excitation, SE)^[26] 注意力模块和卷积块注意力模块 (Convolutional block attention module, CBAM)^[27] 是典型的包含注意力机制的模块设计. 在这些注意力模块内, 首先利用全局池化将全局信息编码到通道维度上或者是空间维度上, 然后利用多层感知器和 3×3 卷积分别解耦特征的通道相关性和空间相关性. 但是全局池化的使用压缩了特征的维度, 造成细节信息的丢失, 这不利于裂缝病害注意力特征图的提取. 为此, 本文提出了下采样注意力模块.

文献 [28] 指出, 常规的卷积层在三维空间中学习提取特征, 为了简化任务, 可以利用 1×1 的卷积和 3×3 分组卷积分别映射特征的通道相关性和空间相关性. 受此启发, 本文在原有的下采样通道上, 额外的增加了一条通道, 在该通道内, 通过 1×1 卷积和 3×3 分组卷积实现特征间通道相关性和空间相关性的解耦, 以判断特征内不同通道及空间位置信息的重要性; 然后在原采样特征的基础上, 对信息进行增强, 来加强裂缝的前景响应. 下采样注意

力模块可以方便地添加到任何网络框架下,其结构如图4所示。

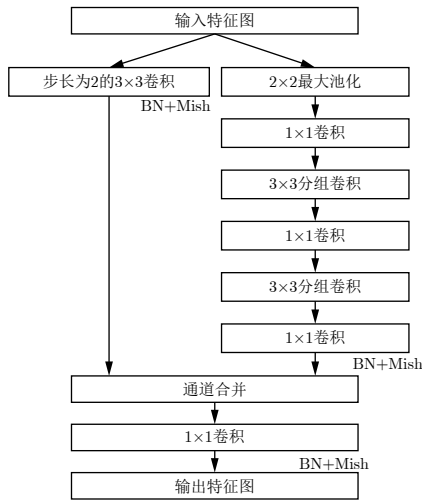


图4 下采样注意力模块

Fig.4 The downsampling attention module

在注意力通道内,先使用 2×2 最大池化对特征进行下采样,之后通过两个连续的 1×1 卷积和 3×3 分组卷积分别实现特征间通道相关性和空间相关性的解耦.在解耦特征相关性的过程中,每次卷积操作后均不添加批量归一化,同时使用线性激活函数,以避免破坏兴趣流形^[42].最后,将注意力通道的下采样结果和原始的下采样结果合并,得到下采样注意力模块的输出。

2.3 双网络桥梁表面病害检测算法

在实际应用中,首先利用BIR-X-LITE机器人对桥梁表面图像进行采集.该机器人由车身主体、相机装置和控制室组成,分别如图5(a1)~(a3)所示,工作方式如图5(b)所示,典型桥梁表面病害见图5(c).

由于不同病害以相等的距离进行拍摄,部分病害如蜂窝和裂缝,在图像内表现出巨大的外观差异.如图6所示,在 5120×5120 像素的高分辨率图像上,仅有巨大的蜂窝病害可被观察到;在 1280×1280 像素的子图像上,蜂窝和漏筋病害表现出明显的特征,但裂缝病害仍不能被很好地识别;在 640×640 像素的子图像上,裂缝病害可以被更好地观察到,但这有可能导致蜂窝病害整体结构信息的丢失.为了对桥梁多病害进行快速鲁棒地检测,本文对YOLO进行了改进,提出了YOLO-lump和YOLO-crack网络,进而形成基于双网络的桥梁表面病害快速检测算法。

2.3.1 YOLO-lump 网络

YOLO-lump 在 1280×1280 像素步长的滑

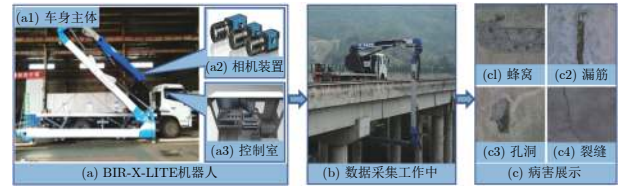


图5 BIR-X-LITE 机器人数据采集过程

Fig.5 The process of data acquisition by the BIR-X-LITE robot

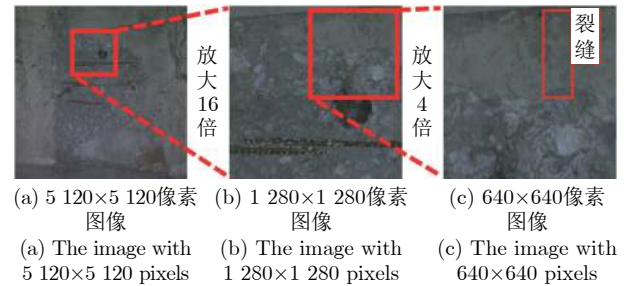


图6 不同病害之间大小比较

Fig.6 Comparison of defects with different sizes

动窗口上实现蜂窝、漏筋和孔洞3种病害的检测.YOLO-lump 由特征提取网络 CSPDarknet-53^[23] 和特征融合网络 PANet^[43] 组成.混合空洞金字塔模块被添加到网络第3个下采样层之前,用于提取稀疏的多尺度特征.正常情况下,正负样本的比例应该都是比较接近的,现有的检测模型正是基于这一假设.然而在桥梁表面病害检测问题中,病害前景的区域(即正样本)是远小于桥梁背景的区域(即负样本)的.负样本的梯度累计,可能会掩盖病害前景的作用,使得网络在训练中无法充分学习到病害的特性信息.为了解决这一问题,focal loss 损失函数被用于计算置信度损失以及分类损失。

2.3.2 YOLO-crack 网络

YOLO-crack 在 640×640 像素步长的滑动窗口上实现裂缝病害的检测,在这一过程中将产生大量的滑窗图像,网络无法实现高效率地检测.因此本文对网络进行了轻量化调整,以实现检测精度和检测速度的平衡.YOLO-crack 由特征提取网络 CSPDarknet-39 和跨阶段局部特征金字塔(Cross stage partial feature pyramid networks, CSP-FPN) 特征融合网络组成.与 CSPDarknet-53 相比,CSPDarknet-39 网络的深度减少了25%,特征通道数减少了19%.而对于标准的CNN网络来说,下采样是必不可少的,常规的下采样方法可能会造成裂缝空间信息的丢失,本文提出的下采样注意力模块,被添加到每一个下采样阶段,以加强裂缝的前景响应,减少空间信息的损失。

为了进一步减少网络参数, 本文使用特征金字塔网络^[44] (Feature pyramid networks, FPN) 作为 YOLO-crack 的特征融合网络, 并在 FPN 中 1×1 卷积和 3×3 卷积级联的部分增加了跳跃连接^[30], 改进后的网络被称为 CSP-FPN. 这可以将具有更多细节信息的底层特征和更多语义信息的高层特征融合, 加强了特征的复用, 有利于保留裂缝细节特征. 在网络输入大小为 640×640 像素时, YOLO-crack 的计算量相比 YOLOv4 减少了 52%. 此外, YOLO-crack 的损失函数与 YOLO-lump 一致.

3 实验结果与分析

本节首先介绍实验设定、数据集和评价指标, 然后对实验结果进行展示和分析. 实验共分为 4 个部分. 第 1 部分为改进前后块状病害检测网络性能对比实验, 比较 YOLO-lump 与其他目标检测算法在块状病害检测上的性能差异; 第 2 部分用于验证本文提出的改进对于裂缝病害检测的有效性, 比较 YOLO-crack 与其他模型在裂缝病害检测上的性能差异; 第 3 部分验证了 YOLO-lump 和 YOLO-crack 的实际应用性能, 在高分辨率图像上对检测结果进行了分析. 第 4 部分用于验证双网络算法对于解决块状病害与裂缝病害差异过大问题的有效性.

3.1 实验说明

3.1.1 实验设定

本文实验程序运行环境为 Ubuntu18.04, 深度学习显卡为 NVIDIA RTX 2080Ti, CUDA 版本为 11.1. 训练过程中所有网络均不使用预训练模型, 初始学习率设置为 0.0005, 动量设置为 0.94, 衰减系数设置为 0.0005, 使用的优化器为 SGD, 批量大小设置为 2, 训练迭代次数为 100 轮. 使用的数据增强方法有 HSV 随机变换、缩放变换、旋转变换和翻转变换.

3.1.2 数据集介绍

本文使用的数据由 BIR-X-LITE 机器人以统一标准采集得到. 所建立的桥梁表观图像数据库中, 包含不同地区共计 10 座桥梁的表观图像数据, 其中有 5120×5120 像素的高分辨率图像共 169621 张, 总计大小 792.7 GB, 具体如表 1 所示. 由于大多数桥梁健康状况良好, 仅有少数图像中包含了桥梁病害, 从上述图像中, 挑选出 1151 张块状病害图像和 643 张裂缝病害图像, 并人工对病害位置进行了标注.

实验中使用 806 张块状病害图像、450 张裂缝病害图像和 118735 张无病害图像进行网络的训练

表 1 桥梁表观图像数据库

采集时间	桥梁名称	图像数目	数据大小
2018-09	东临路大桥	2126张	6.4 GB
	红旗二号桥	10470张	31.4 GB
	荒唐亭大桥	6090张	18.2 GB
	马家河大桥	1402张	4.2 GB
2018-10	南川河大桥	4055张	13.6 GB
	宁家冲大桥	17119张	48.4 GB
	天马大桥	3614张	11.9 GB
2018-11	铜陵长江大桥	19961张	116.0 GB
2019-07	新庆大桥	25784张	225.5 GB
2019-04	广东潮汕大桥	79000张	317.1 GB
总计	10座大桥	169621张	792.7 GB

和性能验证, 剩下的高分辨率图像用于网络性能的测试. 然而受到计算资源的限制, 无法在网络中直接训练高分辨率图像. 为了对网络进行训练, 依据前文所做的分析, 将高分辨率图像中的块状病害裁剪为 1280×1280 像素大小的子图像, 将裂缝病害裁剪为 640×640 像素大小的子图像, 建立的数据集见表 2. 部分训练数据见图 7, 图像经不同比例缩放以方便展示. 图 7(a) 为大型蜂窝病害, 图 7(b) 为小型蜂窝病害, 图 7(c) 为孔洞病害, 图 7(d) 为漏筋和蜂窝病害, 图 7(e) 为严重漏筋病害, 图 7(f) 为小型漏筋病害, 图 7(g) 为显著裂缝, 图 7(h) 为细小裂缝, 图 7(i) 为潮湿裂缝, 图 7(j) ~ (k) 为背景.

3.1.3 评价指标

本实验中, 评价指标采用召回率、准确率、F1 (F1 Score) 和 mAP^[45] (Mean average precision). 其中, 召回率用于描述正确检测到的病害数占应该被正确检测到的病害数的比例, 准确率用于描述正确检测到的病害数占所有检测到的病害数的比例, F1 为召回率和准确率二者加权调和平均, mAP 体现了目标检测网络的综合性能.

3.2 块状病害检测实验分析

3.2.1 网络输入大小对比实验

网络输入大小通常会对块状病害检测结果有较大影响, 在 YOLOv4 网络的训练阶段, 本文将块状病害子图像缩放到不同大小, 缩放大小分别设置为 416×416 像素、 512×512 像素、 608×608 像素和 704×704 像素. 表 3 为不同输入大小下块状病害的检测结果. 由表 3 可知, 受到网络设计的限制, 网络往往具有感受野上限, 当检测物体分辨率过大时, 网络无法捕捉到物体的全局信息. 而过低的输入分

表 2 训练/验证/测试数据集
Table 2 Training/validation/testing datasets

类型	训练集 (正/负样本)	验证集 (正/负样本)	测试集 (正/负样本)
块状病害	7668张 (2611/5057)	2924张 (978/1946)	51231张 (345/50886)
裂缝病害	5643张 (3283/2360)	1453张 (873/580)	51079张 (193/50886)

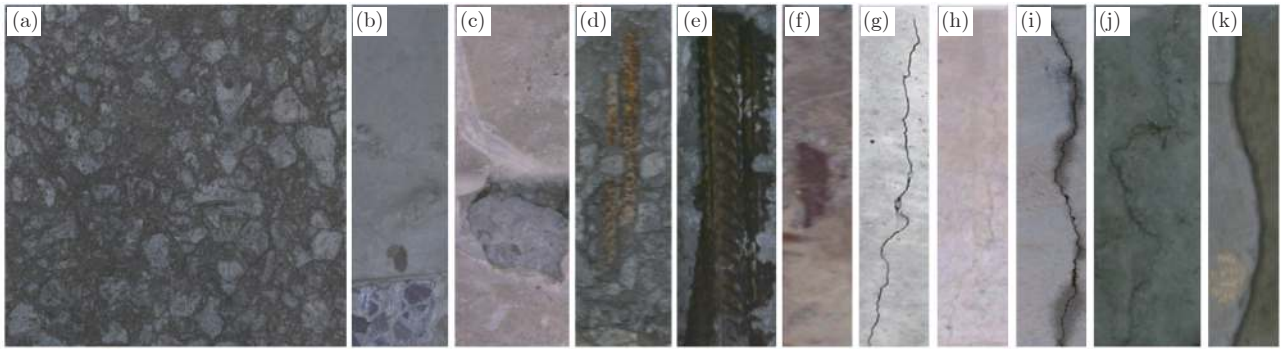


图 7 训练数据示例

Fig.7 Examples of the training dataset

分辨率会导致细节信息的丢失,这都会造成块状病害检测性能的下降.当输入大小为 512×512 像素时,块状病害检测 mAP 为 88.6%,相比于输入大小为 608×608 像素时 mAP 降低了 0.6%,但是检测时间缩短为了 18.8 ms,仅为后者的 76.1%,因此综合考虑检测效率和检测性能,本文选择将块状病害图像缩小至 512×512 像素进行检测.

3.2.2 YOLO-lump 网络结构消融实验

YOLO-lump 网络结构消融实验,验证了本文提出的模块和改进对于块状病害检测网络性能的影响.以 YOLOv4 网络为基础,基于 GAN 网络生成新的样本加入网络训练中,得到的网络被称为 YOLO-lump-A.将 YOLO-lump-A 网络的损失函数修改为 focal loss,所得到的网络被称为 YOLO-lump-B.然后将 ASPP 模块加入到 YOLO-lump-B 网络中,得到的网络被称为 YOLO-lump-C.将膨胀系数全部设置为 3 的混合空洞金字塔模块添加到 YOLO-lump-B 网络中,得到的网络被称为 YOLO-lump-D.将膨胀系数依次设置为 1、2、5 的混合空洞金字塔模块添加到 YOLO-lump-B 网络中,得到的网络被称为 YOLO-lump.实验结果如表 4 和图 8 所示.

通过 YOLOv4 和 YOLO-lump-A 的结果对比可知,利用 GAN 网络生成的困难样本可以提升网络的鲁棒性.由 YOLO-lump-A 与 YOLO-lump-B 的结果对比可知, focal loss 损失函数能够有效解决块状病害检测中正负样本不平衡的问题,提升网络性能.由 YOLO-lump-B 与 YOLO-lump-C 的结果

表 3 不同输入大小下块状病害检测结果对比
Table 3 Results of lump defect detection with different input sizes

输入大小	召回率	准确率	F1	mAP	检测时间
704×704	80.6%	79.3%	79.9%	85.6%	37.8 ms
608×608	86.5%	83.9%	85.2%	89.2%	24.7 ms
512×512	85.8%	84.5%	85.1%	88.6%	18.8 ms
416×416	80.7%	77.4%	79.0%	82.1%	16.6 ms

表 4 YOLO-lump 网络消融实验
Table 4 Ablation experiment on the YOLO-lump

网络模型	召回率	准确率	F1	mAP	检测时间
YOLOv4	85.8%	84.5%	85.1%	88.6%	18.8 ms
YOLO-lump-A	86.1%	84.8%	85.4%	89.3%	18.8 ms
YOLO-lump-B	87.2%	84.4%	85.8%	90.7%	18.8 ms
YOLO-lump-C	79.3%	68.1%	73.3%	74.9%	26.7 ms
YOLO-lump-D	84.4%	83.3%	83.8%	87.7%	20.1 ms
YOLO-lump	86.4%	89.7%	88.0%	92.7%	20.4 ms
YOLO-lump-E	88.7%	89.5%	89.1%	93.5%	24.3 ms

对比可知,ASPP 中每条通道上空洞卷积膨胀系数过大,使得 ASPP 在检测网络底层结构中不能很好地提取特征,导致了网络性能下降.由 YOLO-lump-B 与 YOLO-lump-D 的结果对比可知,在混合空洞金字塔模块内使用同一膨胀系数的空洞卷积,会导致局部信息的丢失和特征连续性的破坏,进而造成病害检测性能的下降.由 YOLO-lump-D 与 YOLO-lump 的结果对比可知,膨胀系数设置合理的混合空洞金字塔模块可以增强网络获取多尺度信息的能

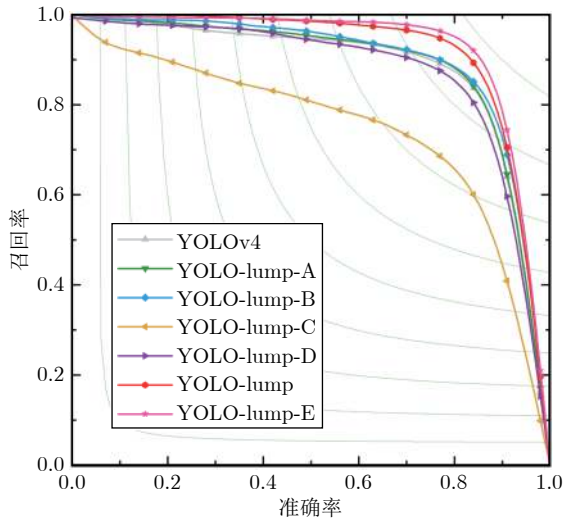


图 8 不同块状病害检测网络的 PR 曲线

Fig.8 Precision-recall curves of different detectors on the lump dataset

力, 并避免了空洞卷积造成的局部信息丢失, 能有效提高块状病害的检测性能. 最终本文提出的 YOLO-lump 网络, 在块状病害检测上 mAP 达到了 92.7%.

此外, 在 YOLO-lump 的基础上加入下采样注意力模块, 得到的网络被称为 YOLO-lump-E. 由 YOLO-lump 与 YOLO-lump-E 的结果对比可知, 下采样注意力模块可以增强下采样过程中块状病害的前景响应, 提升检测性能. 然而 YOLO-lump 不是轻量化的网络, 随着下采样注意力模块中分组卷积数量的增加, 计算效率也是成倍的下降. YOLO-lump-E 检测时间相较 YOLO-lump 增加了 19.1%, 且块状病害检测更依赖于多尺度的全局信息, 下采样注意力模块对 YOLO-lump 网络性能提升相对有限. 因此在 YOLO-lump 网络中使用下采样注意力模块不利于实现检测精度和检测速度的平衡. 根据实际问题需求, 本文使用 YOLO-lump 检测块状病害.

3.2.3 块状病害检测算法对比

本实验将 YOLO-lump 网络的性能与 SSD^[46]、Faster-RCNN^[47]、RetinaNet^[32]、FCOS^[48]、EfficientDet^[29]、YOLOv3、Improved-YOLOv3^[17] 以及 YOLO-v4 网络进行了比较. 实验结果如表 5 所示. 在对比的网络内, 本文算法有最好的性能, mAP 相比其他网络至少保持着 3.1% 的领先, 同时检测时间相较于 YOLOv4 仅增加了 1.6 ms.

图 9 为不同网络在桥梁表观图像上的检测结果. 图像 I~IV 是指大型蜂窝病害图像、蜂窝和漏筋病害图像、细长漏筋病害图像、蜂窝和孔洞病害图像. 其中, A 代表蜂窝病害, B 代表漏筋病害, C 代

表 5 块状病害检测网络对比实验
Table 5 Comparison of different detectors on the lump dataset

网络模型	特征提取网络	mAP	检测时间
SSD	VGG-16	85.1%	30.3 ms
Faster-RCNN	ResNet-101	86.9%	34.9 ms
RetinaNet	ResNet-101	89.5%	41.5 ms
FCOS	ResNet-101	87.9%	28.8 ms
EfficientDet	EfficientNet	89.6%	22.3 ms
YOLOv3	Darknet-53	87.6%	15.4 ms
Improved-YOLOv3	Darknet-53	89.3%	15.4 ms
YOLOv4	CSPDarknet-53	88.6%	18.8 ms
YOLO-lump	CSPDarknet-53	92.7%	20.4 ms

表孔洞病害, D 代表裂缝病害. 由图 9 可知, SSD 和 Faster-RCNN 网络在块状病害检测中出现了较多漏检和误检目标框, 且定位准确度不高, 说明其在特征区分度上低其他算法; 而 RetinaNet、FCOS 和 EfficientDet 网络在检测与正样本相似的背景干扰时, 可能会出现错误检测现象, 例如将图 V 中墙面的黑色痕迹错误地识别为漏筋病害; 在 YOLO 网络 (YOLOv3、Improved-YOLOv3、YOLOv4) 上的检测结果显示, 相比于以上算法, YOLO 网络能够更好地识别背景干扰, 但是在部分场景下对大型蜂窝病害的定位不够准确, 如在图像 II 中, 当蜂窝与漏筋病害发生重叠时, 网络对蜂窝病害进行检测时丢失了左右两侧的信息; 而本文提出的 YOLO-lump 网络与 YOLO 网络相比, 有更强的多尺度性能和更大的感受野, 因此在大型块状病害的检测上有更好的表现, 同时, 能够较准确地识别背景干扰.

3.3 裂缝病害检测实验分析

3.3.1 YOLO-crack 网络结构消融实验

通过 YOLO-crack 网络结构消融实验, 验证了本文提出的模块和改进对于裂缝病害检测网络性能的影响. 以 YOLOv4 网络为基础, 基于 GAN 网络生成新的样本加入网络训练中, 得到的网络被称为 YOLO-crack-A. 将 YOLO-crack-A 网络的损失函数修改为 focal loss, 得到的网络被称为 YOLO-crack-B. 对 YOLO-crack-B 进行轻量化调整, 使用 CSPDarknet-39 作为特征提取网络, 使用 FPN 作为特征融合网络, 得到的网络被称为 YOLO-crack-C. 将 YOLO-crack-C 的特征融合网络改进为 CSP-FPN, 得到的网络被称为 YOLO-crack-D. 在 YOLO-crack-D 中加入下采样注意力模块, 提出的网络被称为 YOLO-crack. 实验结果如表 6 和图 10 所示.

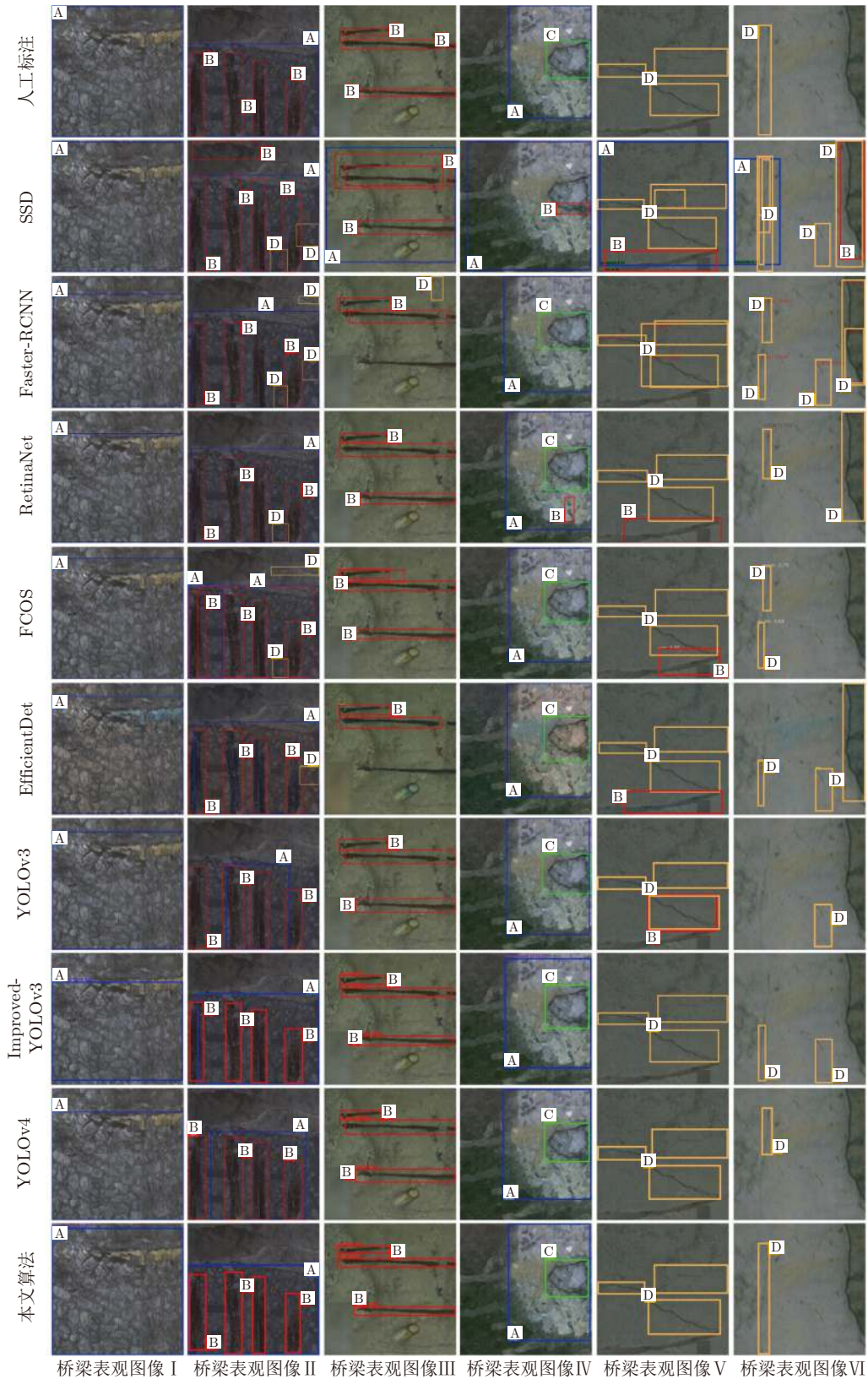


图 9 本文方法和其他方法在不同桥梁表面图像上的测试结果

Fig.9 Results of the proposed method and other methods on various bridge surface images

表 6 YOLO-crack 网络消融实验
Table 6 Ablation experiment on the YOLO-crack

网络模型	召回率	准确率	F1	mAP	检测时间
YOLOv4	80.8%	79.4%	80.2%	84.5%	29.7 ms
YOLO-crack-A	77.5%	82.6%	80.0%	85.0%	29.7 ms
YOLO-crack-B	85.6%	76.7%	81.0%	85.7%	29.7 ms
YOLO-crack-C	78.7%	79.1%	78.9%	83.8%	17.1 ms
YOLO-crack-D	79.0%	79.5%	79.2%	84.6%	16.5 ms
YOLO-crack	80.2%	81.2%	80.7%	86.2%	17.6 ms
YOLO-crack-E	77.9%	80.9%	79.4%	82.5%	17.7 ms

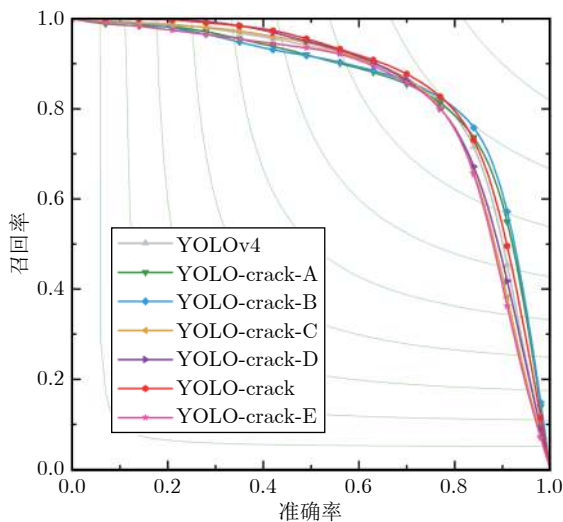


图 10 不同裂缝病害检测网络的 PR 曲线

Fig.10 Precision-Recall curves of different detectors on the crack dataset

通过 YOLOv4 和 YOLO-crack-A 的结果对比可知, 利用 GAN 网络生成的困难样本可以提升网络的鲁棒性和网络对正负样本的识别能力, YOLO-crack-A 在裂缝病害检测上准确率最高. 由 YOLO-crack-A 和 YOLO-crack-B 的结果对比可知, 在 YOLO-crack-A 中加入 focal loss 损失函数后, 网络在正样本上获得充分的训练, 能够更全面地检测裂缝病害, 同时背景干扰信息也更容易被误判为正样本, YOLO-crack-B 在裂缝病害检测上召回率最高. 由 YOLO-crack-B 与 YOLO-crack-C 的结果对比可知, 轻量化的网络在不显著降低裂缝检测性能的前提下, 大幅减少了检测时间, 有利于实现检测精度和检测速度的平衡. 由 YOLO-crack-C 与 YOLO-crack-D 的结果对比可知, CSP-FPN 相较于 FPN, 可以将具有更多细节信息的底层特征和更多语义信息的高层特征融合, 加强了特征的复用, 有利于裂缝病害的检测. 由 YOLO-crack-D 与 YOLO-crack 的结果对比可知, 本文提出的下采样

注意力模块可以减少下采样过程中的信息损失, 加强裂缝的前景响应, 提高裂缝检测的性能. 最终本文提出的 YOLO-crack 网络在裂缝病害检测上 mAP 达到了 86.2%. 分析 F1 和 mAP 指标可知, YOLO-crack-B 和 YOLO-crack 在裂缝病害检测上综合性能优于其他网络, 但是 YOLO-crack 检测时间仅为 YOLO-crack-B 的 59.3%, 可以实现高精度高效率的裂缝病害检测. 根据实际问题, 本文使用 YOLO-crack 检测裂缝病害.

此外, 在 YOLO-crack 网络的下采样注意力模块内, 使用批量归一化和非线性激活函数, 得到的网络被称为 YOLO-crack-E. 由 YOLO-crack 与 YOLO-crack-E 的结果对比可知, 在解耦通道注意力和空间注意力过程中, 使用批量归一化和非线性激活函数会破坏兴趣流形, 造成下采样注意力模块性能的下降.

3.3.2 注意力模块对比实验

为验证本文提出的下采样注意力模块的有效性, 本文将 YOLO-crack-D 网络与注意力模块有 SE 注意力模块、CBAM 注意力模块、下采样注意力模块进行对比实验. 实验结果如表 7 所示, 此外, 利用 Grad-CAM++^[49] 算法生成了网络的激活热力图见图 11.

表 7 注意力模块对比实验

Table 7 Comparison of different attention modules

网络模型	mAP	检测时间
YOLO-crack-D	84.6%	16.5 ms
YOLO-crack-D+SE注意力模块	84.9%	16.9 ms
YOLO-crack-D+CBAM注意力模块	85.7%	17.4 ms
YOLO-crack-D+下采样注意力模块	86.2%	17.6 ms

Grad-CAM++ 算法可以清晰地展示网络关注的区域. 由实验结果可知, YOLO-crack-D 加下采样注意力模块的组合在裂缝病害检测上有最好的性能. 对比 SE 和 CBAM 注意力模块, 下采样注意力模块不需要对特征进行全局池化, 因此能够更好地保留细节信息. 加入下采样注意力模块后, 能够更有效地减少裂缝像素在下采样阶段的损失, 有助于裂缝边缘轮廓信息的保留, 增强网络对裂缝附近特征信息的提取.

3.3.3 裂缝病害检测算法对比

YOLO-crack 与 SSD、Faster-RCNN、RetinaNet、FCOS、EfficientDet、YOLOv3、Improved-YOLOv3、YOLOv4 和 YOLOv4-crack 进行对比实验, 实验结果如表 8 所示. 本文提出的模型有着最

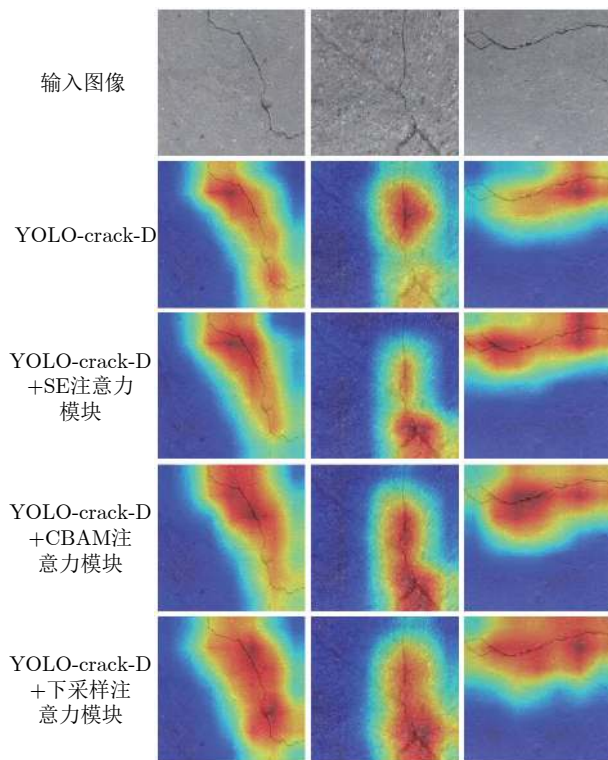


图 11 Grad-CAM++可视化结果

Fig. 11 Grad-CAM++ visualization results

表 8 裂缝病害检测网络对比实验

Table 8 Comparison of different detectors on the crack dataset

网络模型	特征提取网络	mAP	检测时间
SSD	VGG-16	79.8%	45.2 ms
Faster-RCNN	ResNet-101	81.2%	54.7 ms
RetinaNet	ResNet-101	82.9%	58.4 ms
FCOS	ResNet-101	83.4%	42.9 ms
EfficientDet	EfficientNet	83.5%	27.4 ms
YOLOv3	Darknet-53	82.3%	23.8 ms
Improved-YOLOv3	Darknet-53	84.1%	23.8 ms
YOLOv4	CSPDarknet-53	84.5%	29.7 ms
YOLO-crack	CSPDarknet-39	86.2%	17.6 ms

高的精度和最快的检测速度, 相较于其他的目标检测算法, YOLO-crack 在 mAP 上至少保持着 1.7% 的领先, 检测时间仅需要 17.6 ms. 这是因为与其他网络中采用的常规下采样方法相比, 本文提出的下采样注意力模块能增强裂缝的前景响应, 有助于提升裂缝检测的性能. 而轻量化调整可以有效减少网络参数, 提高检测速度.

图 9 中图像 V 和 VI 是指显著裂缝病害图像和细小裂缝病害图像, 裂缝病害位置用字母 D 标记. 由图 9 可知, SSD 和 Faster-RCNN 网络在裂缝检

测上表现相对较差, 出现较多的错误检测案例, 同时在单个病害上有多个重叠的检测框; 对于 RetinaNet、FCOS 和 EfficientDet 网络, 在背景干扰与裂缝较为相似时, 网络可能不能做出准确地识别, 如在图像 II 中, 有一小段外漏的细钢丝, 而上述 3 个网络将其错误识别成裂缝病害. YOLO 网络的检测结果显示, 其能有效减少裂缝病害的错误检测现象, 但 YOLO 网络可能会丢失细小裂缝的信息从而导致漏检; 而本文对 YOLO 进行改进, 提出了 YOLO-crack 网络, 能减少下采样阶段裂缝空间信息的损失, 更完整地提取裂缝的细节特征, 因此在裂缝病害检测上有更好的表现.

3.4 实际应用性能测试

为了评估本文所提出的双网络算法在实际应用中的性能, 对 345 张块状病害图像、193 张裂缝病害图像和 50 886 张无病害图像的高分辨率图像进行检测. 在检测过程中, 利用 TensorRT 对 YOLO-lump 和 YOLO-crack 网络进行了部署, TensorRT 是一个高性能的深度学习前向传播优化器, 可以有效加快检测速度. 由于正负样本的极度不平衡 (1:189), 相较于准确率, 召回率能更好地反映算法性能, 因此在实验结果中, 本文主要对召回率进行计算, 并对误检区域面积与检测图像总面积的比值进行了统计. 实验结果统计见表 9 中, 其中 *GT* 代表图像中包含病害的总数量, *TP* 代表被网络检测到并正确分类的病害数量, *FN* 代表属于病害但是没有被正确检测出来的病害数量, *FP* 代表被错误识别成病害的背景区域数量.

由实验结果可知, 在高分辨率桥梁表面图像上, 本文算法能够在 1 s 以内完成桥梁多病害的检测, 配合 BIR-X-LITE 机器人可实现病害的实时检测. 同时, 对块状病害和裂缝病害检测的平均召回率达到 95% 以上. 而在没有病害的图像上, 误检区域面积仅占检测图像总面积的 1.655%. 实验结果证明, 本文算法可以实现高召回率的桥梁表面病害检测, 并能够在海量的桥梁表面图像中快速筛除大面积的无病害区域, 提高桥梁健康状况评估的效率.

3.5 双网络算法与单网络性能对比实验

在本文的双网络算法中, 使用了不同的滑动窗口大小, 分别对块状病害和裂缝病害进行检测. 实验重新训练了 YOLOv4 网络使其能检测所有桥梁表面病害, 以验证双网络算法对于检测性能的影响. 为了使单网络检测过程中病害尺度保持一致, 将高分辨率图像中的块状与裂缝病害均裁剪为 640×640 像素大小的子图像. 双网络算法与单网络检测的实验结果见表 10.

表 9 实际应用测试结果
Table 9 Results of the practical application

测试数据集	图像数量	块状病害检测					裂缝病害检测					检测时间
		GT	TP	FN	FP	召回率	GT	TP	FN	FP	召回率	
东临路大桥	872	8	8	0	907	100%	6	5	1	1478	83.3%	995 ms/张
红旗二号桥	3265	26	25	1	2132	96.2%	17	17	0	13582	100%	995 ms/张
荒唐亭大桥	2929	22	19	3	3295	86.4%	26	25	1	10115	96.2%	994 ms/张
马家河大桥	836	11	9	2	1041	81.8%	7	7	0	3569	100%	993 ms/张
南川河大桥	2617	20	20	0	4238	100%	23	21	2	5331	91.3%	996 ms/张
宁家冲大桥	2453	28	27	1	3145	96.4%	28	26	2	9504	92.9%	997 ms/张
天马大桥	7107	65	62	3	7294	95.4%	90	86	4	22383	95.6%	996 ms/张
铜陵长江大桥	5962	46	45	1	8505	97.8%	57	55	2	26237	96.5%	996 ms/张
新庆大桥	6194	63	61	2	5598	96.8%	46	45	1	15394	97.8%	995 ms/张
广东潮汕大桥	19189	130	124	6	19869	95.4%	186	179	7	41908	96.2%	995 ms/张
总计	51424	419	400	19	56024	95.5%	486	466	20	149501	95.9%	995 ms/张

表 10 双网络算法与单网络性能比较
Table 10 Comparison of performance between the dual deep network and the single network

检测策略	蜂窝病害				漏筋病害				孔洞病害				裂缝病害				检测时间
	召回率	准确率	F1	mAP	召回率	准确率	F1	mAP	召回率	准确率	F1	mAP	召回率	准确率	F1	mAP	
双网络	85.2%	84.6%	84.9%	86.7%	87.1%	86.5%	86.8%	89.8%	87.3%	85.2%	86.2%	89.3%	80.8%	79.4%	80.1%	84.5%	34.8 ms
单网络	78.5%	77.2%	77.8%	80.6%	83.8%	84.3%	84.0%	84.4%	84.4%	83.1%	83.7%	85.0%	78.7%	76.8%	77.7%	80.1%	30.3 ms

由表 10 分析可知, 与单网络检测相比, 双网络算法在蜂窝、漏筋、孔洞和裂缝病害的检测上, mAP 分别提高了 6.1%、5.4%、4.3% 和 4.4%。这是因为, 块状病害体积较大, 以 640×640 像素对图像进行滑窗时, 往往无法得到完整的病害图像, 这可能会造成病害整体结构信息的丢失, 导致块状病害检测性能的下降。比如高分辨率蜂窝病害的内部看起来可能像是健康的混凝土表面。此外, 在双网络算法中, 块状病害和裂缝病害锚框的平均值分别为 320×304 像素和 178×225 像素。在单网络算法中, 锚框的平均值为 338×336 像素。锚框的变化会对裂缝病害检测性能造成较大的影响。而双网络算法在不同的滑窗大小下对块状病害和裂缝病害进行检测, 可以更加有效地提取不同病害的特征。同时, 双网络算法可以针对不同病害分别设定锚框, 可以有效解决块状病害与裂缝病害差异过大的问题。在速度方面, 双网络算法检测病害所用的加权平均时间相比单网络仅多了 4.5 ms。总之, 双网络算法能够提升桥梁表观病害检测的性能。

4 结束语

本文基于改进 YOLO 提出了一个双网络桥梁表观病害快速检测算法, 其双网络分别为 YOLO-

lump 和 YOLO-crack, 配合 BIR-X-LITE 机器人可实现桥梁多病害的自动鲁棒检测。首先本文利用机器人对数十座桥梁进行拍摄, 建立了一个桥梁表观图像数据库。针对长宽比和大小多变的块状病害, 提出了混合空洞金字塔模块, 该模块可以有效提高网络的多尺度性能, 扩展网络感受野。而针对细小的裂缝病害, 提出了下采样注意力模块, 通过对特征相关性的解耦, 能减少裂缝在下采样阶段的信息损失。将本文算法与目前最新方法进行对比, 实验结果证明了本文方法的有效性, 并适合移植到工业检测环境。在未来的工作中, 可从双网络之间的特征共享角度做适当改进, 进而达到更好的检测结果。

References

- 1 Ma Jian, Sun Shou-Zeng, Yang Qi. Review on China's bridge engineering research: 2014. *China Journal of Highway and Transport*, 2014, **27**(5): 1-96
(马建, 孙守增, 杨琦. 中国桥梁工程学术研究综述: 2014. 中国公路学报, 2014, **27**(5): 1-96)
- 2 Chen Rong-Feng, Xu Qun-Li, Qin Kai-Qiang. Research progress of intelligent bridge crack detection system. *Highway*, 2019, **64**(5): 101-105
(陈榕峰, 徐群丽, 秦凯强. 桥梁裂缝智能检测系统的研究新进展. 公路, 2019, **64**(5): 101-105)
- 3 Zhong Xin-Gu, Peng Xiong, Shen Ming-Yan. Study on the feasibility of intelligent bridge crack detection system based on YOLO-crack and YOLO-lump. *Journal of Highway and Transport*, 2019, **32**(12): 1-6

- ibility of identifying concrete crack width with images acquired by unmanned aerial vehicles. *China Civil Engineering Journal*, 2019, **52**(4): 52–61
(钟新谷, 彭雄, 沈明燕. 基于无人飞机成像的桥梁裂缝宽度识别可行性研究. 土木工程学报, 2019, **52**(4): 52–61)
- 4 Lin W G, Sun Y C, Yang Q N. Real-time comprehensive image processing system for detecting concrete bridges crack. *Computers and Concrete*, 2019, **23**(6): 445–457
- 5 Sutter B, Lelevé A, Pham M T, Gouin O, Jupille N, Kuhn M. A semi-autonomous mobile robot for bridge inspection. *Automation Construction*, 2018, **91**(7): 111–119
- 6 Hirai H, Ishii K. Development of dam inspection underwater robot. *Journal of Robotics, Networking and Artificial Life*, 2019, **6**(1): 18–22
- 7 Zhong Fan, Zhou Ji-Liu, Lang Fang-Nian, He Kun, Huang Mei. Adaptive scale filtering for edge detection. *Acta Automatica Sinica*, 2007, **33**(8): 867–870
(钟帆, 周激流, 郎方年, 何坤, 黄梅. 边缘检测滤波尺度自适应选择算法. 自动化学报, 2007, **33**(8): 867–870)
- 8 Win M, Bushroa A R, Hassan M A. A contrast adjustment thresholding method for surface defect detection based on mesoscopy. *IEEE Transactions on Industrial Informatics*, 2017, **11**(3): 642–649
- 9 Kamaliardakani M, Sun L, Ardakani M K. Sealed-crack detection algorithm using heuristic thresholding approach. *Journal of Computing in Civil Engineering*, 2016, **30**(1): 04014110
- 10 German S, Brilakis I, Desroches R. Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments. *Advanced Engineering Informatics*, 2012, **26**(4): 846–858
- 11 Zhang Hui, Wang Kun-Feng, Wang Fei-Yue. Advances and perspectives on applications of deep learning in visual object detection. *Acta Automatica Sinica*, 2017, **43**(8): 1289–1305
(张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望. 自动化学报, 2017, **43**(8): 1289–1305)
- 12 Shi Y, Cui L, Qi Z. Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems*, 2016, **17**(12): 3434–3445
- 13 Gou Hong-Ye, Yang Biao, Hua Hui, Xie Rui. Research progress of bridge informatization and intelligent bridge in 2019. *Journal of Civil and Environmental Engineering*, 2020, **42**(5): 14–27
(勾红叶, 杨彪, 华辉, 谢蕊. 桥梁信息化及智能桥梁2019年度研究进展. 土木与环境工程学报, 2020, **42**(5): 14–27)
- 14 Zou Q, Zhang Z, Li Q, Qi X, Wang Q, Wang S. Deepcrack: Learning hierarchical convolutional features for crack detection. *IEEE Transactions on Image Processing*, 2019, **28**(3): 1498–1512
- 15 Li Liang-Fu, Ma Wei-Fei, Li Li, Lu Cheng. Research on detection algorithm for bridge cracks based on deep learning. *Acta Automatica Sinica*, 2019, **45**(9): 1727–1742
(李良福, 马卫飞, 李丽, 陆铖. 基于深度学习的桥梁裂缝检测算法研究. 自动化学报, 2019, **45**(9): 1727–1742)
- 16 Kim I H, Jeon H, Baek S C, Hong W H. Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. *Sensors*, 2018, **18**(6): 1881
- 17 Zhang C B, Chang C, Maziar J. Concrete bridge surface damage detection using a single-stage detector. *Computer - aided Civil and Infrastructure Engineering*, 2020, **35**(4): 389–409
- 18 Li S Y, Zhao X F, Zhou G Y. Automatic pixel - level multiple damage detection of concrete structure using fully convolutional network. *Computer - aided Civil and Infrastructure Engineering*, 2019, **34**(7): 616–634
- 19 Yang L, Li B, Li W. Deep concrete inspection using unmanned aerial vehicle towards CSSC database. In: Proceeding of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, Canada: IEEE, 2017. 24–28
- 20 Mundt M, Majumder S, Murali S. Meta-learning convolutional neural architectures for multi-target concrete defect classification with the concrete defect bridge image dataset. In: Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019. 11188–11197
- 21 Hütthwohl P, Lu R D, Brilakisa I. Multi-classifier for reinforced concrete bridge defect. *Automation in Construction*, 2019, **105**(SEP.): 102824.1–102826.15
- 22 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 779–788
- 23 Bochkovskiy A, Wang C Y. Yolov4: Optimal speed and accuracy of object detection [Online], available: <https://arxiv.org/abs/2004.10934>, April 23, 2020
- 24 He K M, Zhang X, Ren S. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **37**(9): 1904–1916
- 25 Wang P, Chen P, Yuan Y. Understanding convolution for semantic segmentation. In: Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe, USA: IEEE, 2018. 1451–1460
- 26 Hu J, Shen L, Surr G. Squeeze-and-excitation networks. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 7132–7141
- 27 Woo S, Park J, Lee J Y. CBAM: Convolutional block attention module. In: Proceedings of the 2018 European Conference on Computer Vision. Munich, Germany: Springer, 2018. 3–19
- 28 Chollet F. Xception: Deep learning with depth-wise separable convolutions. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2017
- 29 Tan M X, Le Q V. EfficientNet: Rethinking model scaling for convolutional neural networks. In: Proceedings of the 2019 International Conference on Machine Learning. Los Angeles, USA: IEEE, 2019. 97: 6105–6114
- 30 Wang C Y, Mark-Liao H Y, Wu Y H, Chen P Y. CSPNet: A new backbone that can enhance learning capability of CNN. In: Proceedings of the 2020 IEEE Conference on Computer Vision and Pattern Recognition Workshop. Seattle, USA: IEEE, 2020. 390–391

- 31 Wang T H, Liu M Y, Zhu J Y, Tao A, Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional GANs. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 8798–8807
- 32 Lin T Y, Goyal P, Girshick R. Focal loss for dense object detection. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 2999–3007
- 33 Jiang W B, Liu M, Peng Y N, Wu L H, Wang Y N. HDCB-Net: A neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges. *IEEE Transactions on Industrial Informatics*, 2021, **17**(8): 5485–5494
- 34 Goodfellow I, Mirza M, Xu B, Courville A, Bengio Y. Generative adversarial networks. In: Proceedings of the 2014 Conference and Workshop on Neural Information Processing Systems. Montreal, Canada: 2014.
- 35 Liu Jian-Wei, Xie Hao-Jie, Luo Xiong-Lin. Research progress on application of generative adversarial networks in various fields. *Acta Automatica Sinica*, 2020, **46**(12): 2500–2536 (刘建伟, 谢浩杰, 罗雄麟. 生成对抗网络在各领域应用研究进展. *自动化学报*, 2020, **46**(12): 2500–2536)
- 36 Lin Yi-Lun, Dai Xing-Yuan, Li Li, Wang Xiao, Wang Fei-Yue. The new frontier of AI research: Generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 775–792 (林懿伦, 戴星原, 李力, 王晓, 王飞跃. 人工智能研究的新前线: 生成式对抗网络. *自动化学报*, 2018, **44**(5): 775–792)
- 37 Zheng Z, Zheng L, Yang Y. Unlabeled samples generated by GAN improve the person reidentification baseline in vitro. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. 3774–3782
- 38 Yang Q, Yan P, Zhang Y. Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging*, 2018: 1348–1357
- 39 Nie D, Trullo R, Lian J, Wang L, Petitjean C. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, 2018, **65**(12): 2720–2730
- 40 Chen L, Papandreou G, Kokkinos I. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(4): 834–848
- 41 Misra D. Mish: A self regularized non-monotonic activation function. In: Proceedings of the 2020 British Machine Vision Virtual Conference. Virtual Event, UK: 2020.
- 42 Sandler M, Howard A, Zhu M, Zhmoginov A. MobileNetV2: Inverted residuals and linear bottlenecks. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 4510–4520
- 43 Liu S, Qi L, Qin H F, Shi J P, Jia J Y. Path aggregation network for instance segmentation. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 8759–8768
- 44 Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Venice, Italy: IEEE, 2017. 2117–2125
- 45 Dollár P, Zitnick C L. Fast edge detection using structured forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **37**(8): 1558–1570
- 46 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S. SSD: Single shot multi-box detector. In: Proceedings of the 2016 European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016. 21–37
- 47 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137–1149
- 48 Tian Z, Shen C H, Chen H, He T. FCOS: Fully convolutional one-stage object detection. In: Proceedings of the 2019 IEEE International Conference on Computer Vision. Seoul, Korea: IEEE, 2019. 9627–9636
- 49 Chattopadhyay A, Sarkar A, Howlader P. Grad-CAM++: Improved visual explanations for deep convolutional networks. In: Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe, USA: IEEE, 2018.



彭雨诺 湖南大学电气与信息工程学院硕士研究生. 2019 年获得湖南大学学士学位. 主要研究方向为深度学习, 图像处理.

E-mail: pengyunuo@hnu.edu.cn

(PENG Yu-Nuo Master student at the College of Electrical and Information Engineering, Hunan University. He received his bachelor degree from the Hunan University in 2019. His research interest covers deep learning and image processing.)



刘敏 湖南大学电气与信息工程学院教授. 2012 年获美国加州大学河滨分校博士学位. 主要研究方向为模式识别和机器学习. 本文通信作者.

E-mail: liu_min@hnu.edu.cn

(LIU Min Professor at the College of Electrical and Information Engineering, Hunan University. He received his Ph.D. degree from University of California, Riverside in 2012. His research interest covers pattern recognition and machine learning. Corresponding author of this paper.)

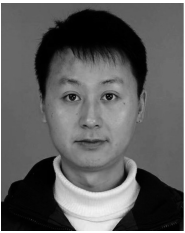


万 智 湖南桥康智能科技有限公司总工程师。1999 年获湖南大学学士学位, 2010 年获中南大学博士学位。主要研究方向为公路交通和环保监测。

E-mail: xbh0n3@163.com

(**WAN Zhi** Chief engineer at Hunan Qiaokang Intelligent Technology Company Limited.

He received his bachelor's degree from Hunan University in 1999 and Ph.D. degree from the Central South University in 2010. His research interest covers highway traffic and environmental protection monitoring.)

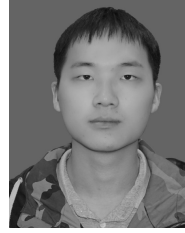


蒋文博 湖南大学电气与信息工程学院硕士研究生。2018 年获桂林电子科技大学学士学位。主要研究方向为图像处理 and 模式识别。

E-mail: jiang_wenbo@hnu.edu.cn

(**JIANG Wen-Bo** Master student at the College of Electrical and Information Engineering, Hunan University.

He received his bachelor degree from the Guilin University of Electronic Technology in 2018. His research interest covers image processing and pattern recognition.)



何文轩 湖南大学电气与信息工程学院硕士研究生。2020 年获昆明理工大学学士学位。主要研究方向为图像处理和模式识别。

E-mail: hwx@hnu.edu.cn

(**HE Wen-Xuan** Master student at the College of Electrical and Information Engineering, Hunan University.

He received his bachelor degree from Kunming University of Science and Technology in 2020. His research interest covers image processing and pattern recognition.)



王耀南 中国工程院院士, 湖南大学电气与信息工程学院教授。1995 年获湖南大学博士学位。主要研究方向为机器人学, 智能控制和图像处理。

E-mail: yaonan@hnu.edu.cn

(**WANG Yao-Nan** Academician at Chinese Academy of Engineering, professor at College of Electrical and Information Engineering, Hunan University.

He received his Ph.D. degree from the Hunan University in 1995. His research interest covers robotics, intelligent control and image processing.)