

基于 Attention-YOLOv3 的锈蚀区域检测与识别

吴之昊¹, 熊卫华^{1*}, 任嘉锋¹, 姜明²

(1. 浙江理工大学 机械与自动控制学院, 浙江 杭州 310018;

2. 杭州电子科技大学 计算机学院, 浙江 杭州 310018)

摘要: 电力设备的锈蚀检测作为危害电力系统安全运行的重要一环必须能被快速、准确地进行识别与检测并及时报警。为了提高电力设备锈蚀区域检测的时效性和可靠性, 基于 YOLOv3 目标检测算法, 结合注意力机制提出了一种改进的 Attention-YOLOv3 算法, 可以实现对锈蚀区域的快速可靠识别。首先, 利用深度可分离卷积对 YOLOv3 的特征提取网络进行轻量化处理来缩减模型的大小, 提高检测速度。其次, 为了弥补轻量化网络带来的精度损失, 提高特征的提取能力, 在上采样之后采用了空间注意力机制 (spatial-attention) 和通道注意力机制 (channel-attention) 结合的级联双注意力机制对特征进行融合筛选, 剔除冗余的无效特征。实验表明, 提出的锈蚀区域检测算法能有效地检测和识别出电力设备的锈蚀区域, 相比较标准 YOLOv3 可以做到在检测时间缩短近 46% 的情况下提升 9.06% 的检测精度, 在 RustDetection 数据集上可以达到 91.75% 的平均精度。

关键词: 目标检测; 锈蚀检测; 注意力机制; 特征提取网络; 轻量级网络

中图分类号: TP391.4

文献标识码: A

文章编号: 1673-629X(2020)11-0147-06

doi: 10.3969/j.issn.1673-629X.2020.11.027

Detection and Identification of Corrosion Area Based on Attention-YOLOv3

WU Zhi-hao¹, XIONG Wei-hua^{1*}, REN Jia-feng¹, JIANG Ming²

(1. Faculty of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China;

2. School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: Corrosion detection of power equipment, as an important part of the safety operation of the power system, must be quickly and accurately identified and detected and timely alarmed. In order to improve the timeliness and reliability of corrosion area detection of power equipment, we propose an improved Attention-YOLOv3 algorithm based on YOLOv3 target detection algorithm combined with attention mechanism, which can realize fast and reliable identification of rust area. Firstly, we reduce the model size and improve the detection speed by lightweighting the YOLOv3 backbone network with the depthwise convolution. Then, in order to compensate for the loss of precision caused by lightweight network and improve the feature extraction ability, we use a cascade of double-spatial-channel (channel-attention) and channel-attention (spatial-attention) after upsample operation. The attention mechanism combines and filters the features to eliminate redundant invalid features. Experiments show that the proposed rust area detection algorithm can effectively detect and identify the rust area of power equipment. Compared with the standard YOLOv3, the detection accuracy can be improved by 9.06% when the detection time is shortened by nearly 46%. The proposed algorithm can achieve an average precision of 91.75% at the RustDetection dataset.

Key words: object detection; corrosion detection; attention mechanism; feature extraction network; lightweight network

0 引言

电力设备锈蚀会导致电力变压器、输电线路发生故障, 甚至会影响到电力系统的安全运行。目前的锈蚀检测方法往往采用人工巡检的方式, 需要耗费大量的人力物力, 检测效果也不理想。为了增加电力设备

锈蚀检测的效率, 利用广泛架设的摄像头和机器视觉技术的检测方法已经成为了目前的主要发展方向, 宋伟^[1]、Liao^[2]、张洪才^[3]等根据锈蚀区域的颜色特征在 RGB、HSI 等不同的颜色空间中进行锈蚀检测, 卢树杰等^[4]根据单目视差原理改进了检测算法, 降低了误检

收稿日期: 2019-11-21

修回日期: 2020-03-25

基金项目: 国家自然科学基金 (61803339, 61503341); 浙江省自然科学基金 (LQ18F030011); 浙江省重点研发计划项目 (2019C03096)

作者简介: 吴之昊 (1995-), 男, 硕士研究生, 研究方向为深度学习、模式识别; 熊卫华, 博士, 副教授, 研究方向为信号处理、故障诊断。

测概率。此外,Chang^[5]采用了傅里叶变换、戴玉静等^[6]添加了纹理特征对锈蚀区域进行识别检测,取得了较好的检测效果。

由于电力设备一般架设在室外,受环境影响较大,因此基于传统图像处理的检测算法具有一定的局限性。而逐渐发展基于深度学习的目标检测方法可以有效地避免这个问题。Luca Petricca 等^[7]利用卷积神经网络对锈蚀图像进行分类,李辉等^[8]结合传统图像处理与深度学习,提高了检测的准确率。王达磊^[9]、周自强^[10]等则采用了深度学习的方法对锈蚀区域进行了分割和识别,有效地提取出了锈蚀区域。

目前电力设备的锈蚀区域检测存在目标样本少,实时性和可靠性要求高等特点,一般的通用目标检测算法不能很好地满足需求。因此该文基于目前表现最佳的 YOLOv3 算法,采用深度可分离卷积对模型进行压缩,并结合空间-维度双注意力机制提出一种 Attention-YOLOv3 检测方法,同时在构建的锈蚀检测数据集上进行训练,所获得的模型能有效地对锈蚀区

域进行检测和识别。

1 理论基础

1.1 YOLOv3 目标检测算法

目标检测算法分为二阶段算法 (two-stage) 和一阶段算法 (one-stage), 其中 YOLO (you only look once) 系列是一阶段算法中非常有代表性的算法,采用了在同一个网络中同时进行分类与回归的方式,相比较以 R-CNN^[11] 系列为代表的二阶段算法虽然损失了部分精度,但大幅度缩减了检测时间。YOLOv3 算法^[12] 在 v1 和 v2 的基础上增加了类似 FPN^[13] 的多尺度检测,增加了其在小目标检测上的能力。同时 YOLOv3 针对不同的数据集,利用 K-means 对训练集中的所有目标框进行聚类来得到候选框。所以 YOLOv3 算法可以很好地适应锈蚀目标大小、形状多样的特点,也可以满足电力设备检查的实时性需求。图 1 所示为标准 YOLOv3 的算法结构。

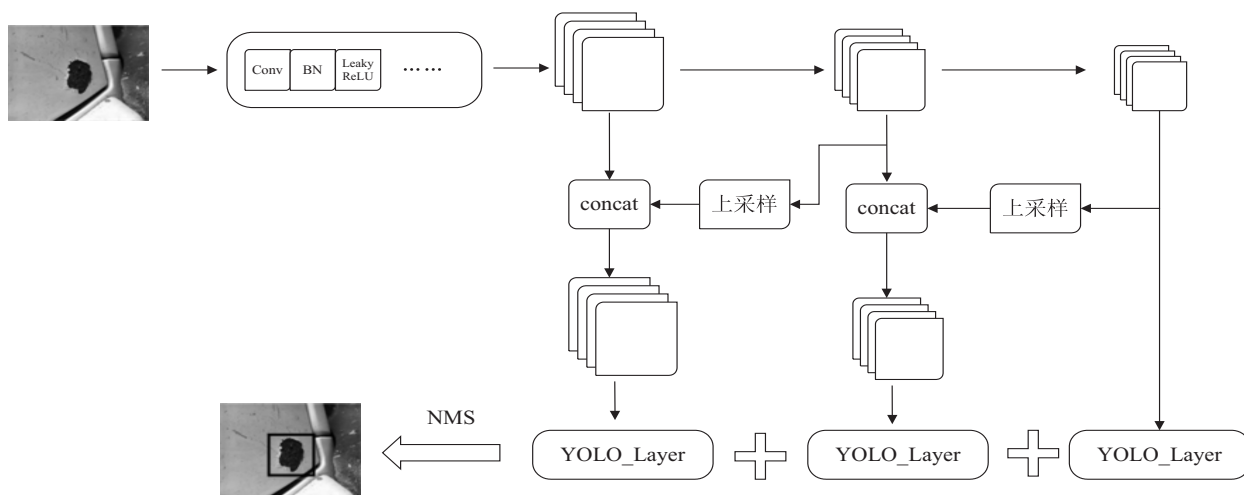


图 1 标准 YOLOv3 的算法结构

1.2 注意力模型 (attention model)

注意力模型来源于人眼的选择性注意机制,可以从大量信息中筛选出具有较高价值的主要信息,自从注意力模型最初被应用于循环神经网络(RNN)之后,该技术就被广泛地应用于图像处理、语音识别等多个深度学习领域。注意力模型是通过训练一个掩码层,通过掩码对该注意力域内的信息进行加权,利用分配不同的权重来使特征提取器忽略无效信息而关注有效信息。Hu 等^[14]在 SENet (squeeze-and-excitation networks) 中提出的 SE-Block 注意力模块和 Fu 等^[15]提出的应用于语义分割的 DANet 都采用了基于注意力机制的检测方法,并取得了较好的检测效果。图像领域的注意力一般集中于用于提取不同特征的通道域和建立像素之间关系的空间域内。

其中,通道域的注意力利用不同维度之间的依赖

关系进行建模,然后对原始特征进行加权重标定。通道注意力的算法模型结构如图 2(a) 所示。

而空间域的注意力则通过提取空间内各像素之间的关联性,形成一个空间上下文依赖度的关系掩码,并将其以权重的形式分配给特征图的每个像素位置,用于突出空间内的有效特征、抑制无效特征,从而达到对特征进行选择 and 融合的作用。通道注意力的算法模型结构如图 2(b) 所示。

2 文中方法

文中提出一种改进的 Attention-YOLOv3 算法锈蚀区域检测算法。首先采用 K-means 算法聚类得到适用于文中数据集的候选框,其次采用以深度可分离卷积为基本模块的 MobileNet-v3 网络替代 YOLOv3 的 Darknet 主干网络,在大量减少计算量的同时保证

一定的检测精度,最后结合空间注意力和维度注意力的级联双注意力模型对特征选择进一步优化得到最佳

的特征表达。图 3 为基于 Attention-YOLOv3 的锈蚀区域检测算法模型的具体架构。

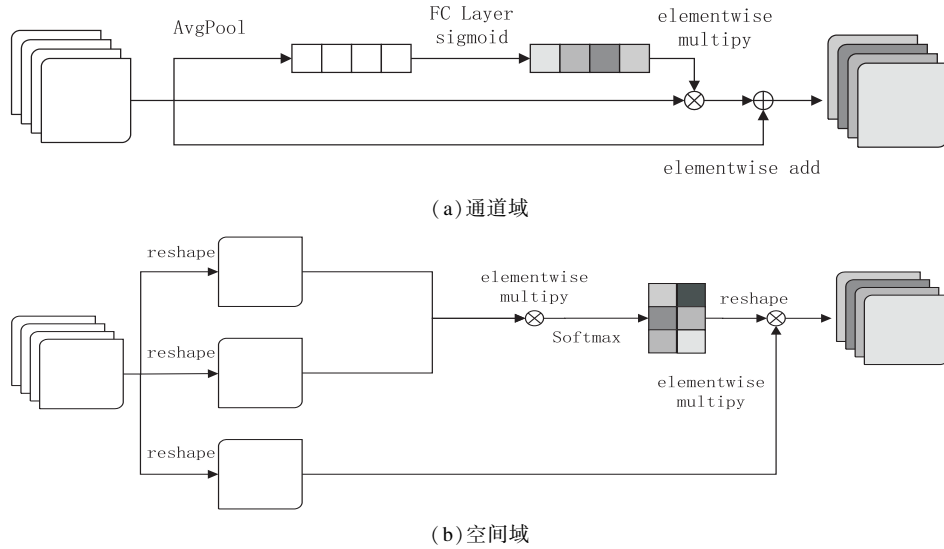


图 2 注意力机制

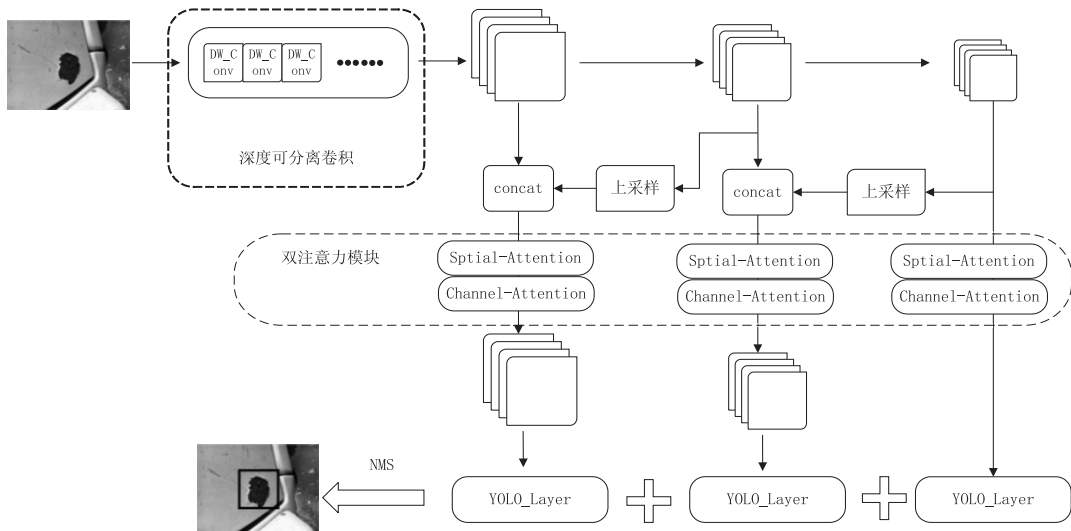


图 3 基于 Attention-YOLOv3 的锈蚀区域检测算法

2.1 轻量级 YOLOv3 主干网络

DarkNet 是由 YOLO 系列作者 Joseph Redmon 提出的一个开源神经网络框架。Darknet 在 YOLOv3 上的表现已经很优越,但是由于其庞大的网络结构,在实际运行中仍需要很大的计算量和推断时间。虽然 DarkNet 提出了轻量级网络 Tiny-DarkNet,但是只对网络结构做了简单的精简操作,没有考虑到特征提取能力下架会导致检测精度的大幅度下降的后果。

文中采用深度可分离卷积代替 DarkNet 中的标准卷积操作,并以此作为 YOLOv3 的特征提取主干网络。深度可分离卷积和标准卷积的密集连接的方式的不同之处是,首先深度可分离卷积采用了一个 1×1 的卷积增加通道数,增加特征的表达能力,其次使用 3×3 的单通道分组卷积作用于单特征图内的信息交流,最后再用 1×1 的卷积操作作用于通道间的信息交流,因此

可以在缩减参数的同时保证检测精度不会大幅下降。

2.2 双注意力模型

卷积操作实现了空间内的局部特征提取和通道间的信息交流,但是往往会丧失空间语义上的关联度,同时不会对通道的重要性进行区分。因此可以使用空间注意力模型对空间内的特征进行依赖性分析,同时通过通道注意力模型对不同通道的特征图进行加权得到更加有利于分类的特征图,所以该文提出一种级联的双注意力模型用于对特征进行选择。

空间注意力模型首先通过矩阵乘法构建每个像素点与其他像素点之间的关系,然后通过一个 softmax 函数对其进行关联性分析并组成权重掩码矩阵,再通过矩阵乘法对原始特征进行重标定得到空间注意力特征。随后将空间注意力特征作为通道注意力模型的输入,利用全局池化来表示全局特征,然后通过一个全连

接层和 sigmoid 函数对其进行权重标定,最后再通过点乘来得到完整的注意力特征。级联双注意力模型完整结构如图 4 所示。

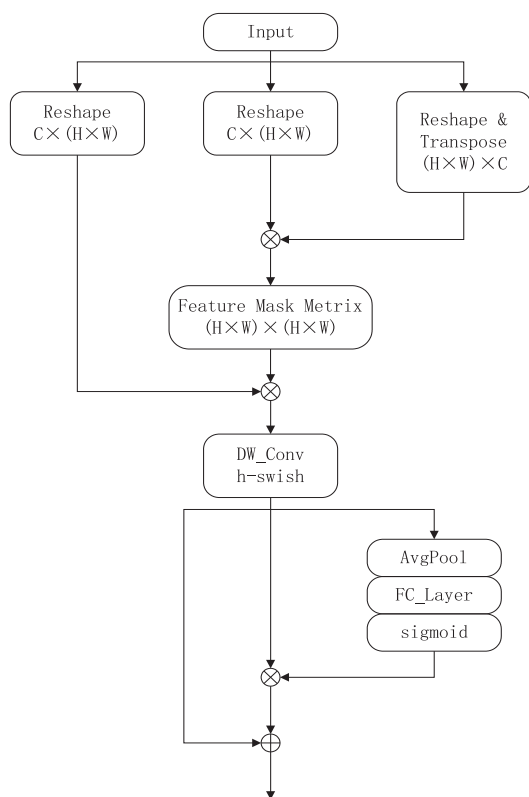


图 4 级联双注意力模型

图 5 为原图像经过上采样后得到的特征图与经过空间注意力与通道注意力之后的对比,其中图 5(a)为原图像,图 5(b)为卷积后的特征图,图 5(c)为空间注意力后的特征图,图 5(d)为通道注意力后的特征图。

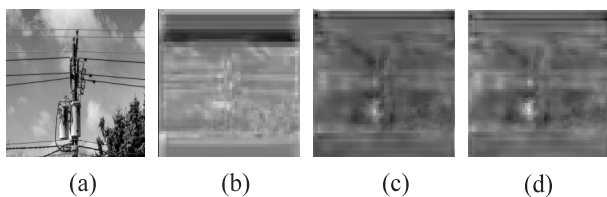


图 5 注意力模型前后的特征对比

由图 5 可知,经过上采样后的特征图存在大量的无效特征,而在经过空间注意力模块后,无效的特征被抑制了,电塔上的锈蚀特征则被提取出来,在此基础上再通过通道注意力模块对其进行加强,最后得到的特征图更加有利于目标的分类定位。

3 实验

3.1 数据集与前期工作

文中采用的 RustDetection 锈蚀目标检测数据以电缆、变压器、电表、电箱等多种电力设备锈蚀故障图片为基础,同时通过网络采集、实地拍摄等方式收集包含金属管道锈蚀、车辆锈蚀、桥梁锈蚀等共计 800 张锈蚀

图像,并以 3 : 1 的比例划分为训练集和测试集。并利用 LabelImg 标注软件对数据集进行标定,按照 COCO 格式建立标准的数据集。

随后结合 Faster R-CNN 的基于长宽比的 anchor 思想,提出一种改进的 K-means 聚类算法对训练集进行聚类,不仅采用了 IOU score,还增加了 ratio 作为参考依据用于更好地对细长型和扁平型目标框进行回归。改进的距离公式如式(1)所示,其中 λ_1 和 λ_2 分别取 0.8 和 0.2。

$$d(\text{box}, \text{centroid}) = 1 - (\lambda_1 \text{IOU} + \lambda_2 \text{ratio}) \quad (1)$$

共取得 Avg IOU 可以达到 68.98% 的 $[[13, 19], [23, 32], [28, 49], [43, 63], [77, 45], [69, 92], [99, 117], [45, 138], [192, 191]]$ 九种不同大小的预选框,并分别给三个尺度的特征图分配三种不同大小的预选框。

3.2 实验过程

算法模型的训练和测试过程均在 NVIDIA GTX 1080Ti GPU 上进行。首先对输入样本进行归一化,缩放为 416×416,然后输入算法模型进行训练。

训练过程采用迁移学习的方式,即整个网络首先在具有 17 125 张图片和 20 类的 VOC2012 公共数据集上进行预训练,等网络获得较好的检测效果后,将 20 类分类子网络改为文中实验所需的二分类子网络。随后将除该子网络之外的特征提取网络的参数冻结,利用提出的 RustDetection 数据集针对分类子网络进行专项训练,同时也防止因为数据集变化导致的整体网络性能下降,20 轮之后开放特征提取网络,对整个网络的特征提取能力进行训练。图 6 为采用了预训练模型和随机参数初始化的训练 loss 曲线,可以看到采用了预训练的模型的损失下降得更快,也可以更快地达到最小值。

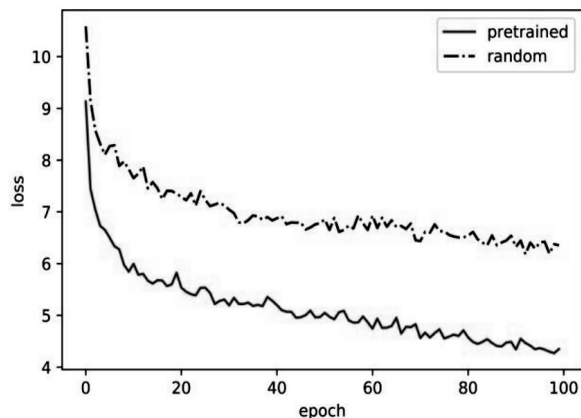


图 6 迁移学习的 Loss 曲线

3.3 实验结果分析

评价指标为 AP(average precision),采用 11-Point 方法进行计算得到,即采用 recall 值在 0 到 1 之间每隔

0.1 时取最大的 precision 值绘制 P-R 曲线,求得 precision 的平均值即为 AP 值。其中,准确率 (precision)、召回率 (recall)、AP 值的计算如下:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{AP}_{11\text{-point}} = \frac{1}{11} \left(\sum_{x \in \text{MaxPrecision}} x \right) \quad (4)$$

其中,TP 表示正例被判真的个数,FP 表示正例被判假的个数,FN 表示负例被判假的个数, x 为 11 个 recall 值下的最大 precision 值。在实际的检测中对于正样本需要被检测出来的要求更为严苛,因此召回率 (recall) 在

表 1 各模型的测试效果比较

model	precision/%	recall/%	AP/%
YOLOv3 (DarkNet)	69.20	83.72	82.69
Tiny-YOLOv3	45.83	53.48	50.52
YOLOv3 (DW_Conv)	63.41	83.87	71.89
Our Method	70.17	93.02	91.75

文中在此基础上添加了级联的双注意力模型,使得检测精度大幅度上升,这是由于注意力机制在上采样之后进行了特征的选择,剔除了无效的冗余特征,留下了对分类和定位更有利的特征信息,不会被无效特征所误导,因此可以更加精准地定位出锈蚀的区域。

表 2 各模型大小和检测时间比较

model	weight size/MB	time/s
YOLOv3 (DarkNet)	237	1.82
Tiny-YOLOv3	33	0.27
YOLOv3 (DW_Conv)	18	0.69
Our Method	20	0.98

由表 2 可知,虽然 Tiny-YOLOv3 有最快的推断时间,但是其权重文件比应用了 DW_Conv 的 YOLOv3 和文中算法都要大。这是因为深度可分离卷积相比较标准卷积有更多的矩阵运算,因此需要更多的时间用于推断计算,但其仍就比标准 YOLOv3 提升了 1.85 到 2.63 倍。而文中算法相较于应用了 DW_Conv 的 YOLOv3,由于加了级联双注意力模型,所以推断时间有 0.3 s 的延迟,相比较标准的 YOLOv3 算法可以缩短 46.7% 的检测速度。

在模型大小上,标准 YOLOv3 具有庞大的体量,需要很大的存储内存,不适合部署在工业设备上,应用了深度可分离卷积后减少了 12 倍,已经可以满足工业设备的存储需求。而文中方法增加了双注意模型的微量参数,增加了 2 MB 的内存需求,但是综合检测速度和检测精度,文中算法依旧是最佳的检测算法。

锈蚀目标检测中更为重要,也是最重要的指标之一。同时,由于 AP 值综合了准确率和召回率进行考量,所以 AP 值为判定模型是否能正确识别锈蚀区域的主要判定标准。表 1 为各种模型与文中算法在 RustDetection 测试集上的对比。可以看出, Tiny-YOLOv3 只是对网络结构进行简单的删减,并没有考虑其在精度上的变化,因此若直接采用 Tiny-YOLOv3 就会导致检测效果大幅度下降,不能满足现实中检测的需求。而采用深度可分离卷积来代替 DarkNet 的 YOLOv3 算法在检测效果上由于参数的下降带了一定程度的精度下降,仍在可接受的范围内,基本达到了与标准 YOLOv3 算法相同的水平。

表 2 为不同模型的权重文件大小和检测时间对比。由于实际应用场景中没有 GPU 进行检测的加速,往往只有 PC 机作为检测的主要设备,因此文中的检测过程均在 Intel Core i5-7200U CPU 上进行。

最终的检测效果如图 7 所示,可以较好地检测出检修电源箱、输电线、电力箱等上的锈蚀区域。



图 7 模型测试结果

4 结束语

提出的基于 Attention-YOLOv3 的锈蚀区域检测

算法得到了实验验证,实验结果表明该算法可以在轻量级的检测模型上得到更好的检测效果,若可以在专用的图形处理器上,该算法可以做到 10 fps 以上的检测效率,完全满足了实时监测的要求。在下一阶段,将进行模型的压缩和优化,并将其部署在移动设备上,并和报警设备进行结合,做到现场检测、现场报警、现场处理。

参考文献:

- [1] 宋 伟,左 丹,邓邦飞,等. 高压输电线防震锤锈蚀缺陷检测[J]. 仪器仪表学报,2016,37(S):113-117.
- [2] LIAO K W,LEE Y T. Detection of rust defects on steel bridge coatings via digital image recognition[J]. Automation in Construction,2016,71:294-306.
- [3] 张洪才. 直升机巡检输电线路锈蚀缺陷识别方法研究[D]. 大连:大连海事大学,2009.
- [4] 卢树杰,陈庆诚,文 祥,等. 基于无人机的钢结构锈蚀视觉检测[J]. 工业控制计算机,2019,32(7):79-81.
- [5] CHANG L M,SHEN H K,CHEN P H. Automated rust defect recognition method based on color and texture feature [J]. Automation in Construction,2013,31:338-356.
- [6] 戴玉静,吕东辉,郭松鸽. 基于颜色和纹理特征的输电线路锈蚀区域检测[J]. 工业控制计算机,2018,31(9):39-40.
- [7] PETRICCA L,MOSS T,FIGUEROA G,et al. Corrosion detection using AI: a comparison of standard computer vision techniques and deep learning model[C]//Proceedings of the sixth international conference on computer science,engineering and information technology. Vienna:ACM,2016:91-99.
- [8] 李 辉,钟 平,戴玉静,等. 基于深度学习的输电线路锈蚀检测方法的研究[J]. 电子测量技术,2018,41(22):54-59.
- [9] 王达磊,彭 博,潘 玥,等. 基于深度神经网络的锈蚀图像分割与定量分析[J]. 华南理工大学学报:自然科学版,2018,46(12):121-127.
- [10] 周自强,纪 扬,苏 烨,等. 基于迁移学习卷积神经网络的电缆隧道锈蚀识别算法[J]. 中国电力,2019,52(4):104-110.
- [11] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Columbus:IEEE,2014:580-587.
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas:IEEE,2016:779-788.
- [13] LIN T Y,DOLLÁR P,GIRSHICK R,et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Hawaii:IEEE,2017:2117-2125.
- [14] HU J,SHEN L,SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake: IEEE,2018:7132-7141.
- [15] FU J,LIU J,TIAN H,et al. Dual attention network for scene segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Long Beach:IEEE,2019:3146-3154.