

文章编号:1674-2869(2020)04-0462-05

# 基于改进 YOLO V3 的接触网绝缘子检测方法

陈智羽, 闵 锋\*

武汉工程大学计算机科学与工程学院, 湖北 武汉 430205

**摘 要:** 为代替人工对 4C 巡检车拍摄铁路接触网图像进行分析, 使检测的速度和准确率达到实用的要求, 本文提出一种基于改进 YOLO V3 的接触网绝缘子检测方法。该方法在 YOLO V3 的网络结构 Darknet-53 的第二个残差块和第三个残差块中间再增加一个新的对小目标友好的 4 倍降采样的残差块, 提高对小目标的检测准确率。并根据相似图像中绝缘子的位置大体相同的特点, 通过感知哈希算法分类图像, 对同类图像采用候选区域扫描策略加快检测速度。实验结果表明改进后的方法对绝缘子检测的准确率从 93.6% 提升至 99.2%, 同类图像的检测速度提升了 46%。

**关键词:** YOLO V3; 铁路接触网; 绝缘子检测; 残差块

中图分类号: TP391.4

文献标识码: A

DOI: 10.19843/j.cnki.CN42-1779/TQ.201906027

## Detection Method of Catenary Insulator Based on Improved YOLO V3

CHEN Zhiyu, MIN Feng\*

School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430205, China

**Abstract:** To replace the manual analysis of railway catenary images taken by 4C inspection vehicles and meet the practical requirements for the time efficiency and accuracy of detection, we proposed a detection algorithm for catenary insulators based on improved You Only Look Once V3 (YOLO V3). The YOLO V3 network Darknet-53 was optimized to improve the detection accuracy of small targets by adding an elaborated 4-fold downsampling residual block that is friendly to small targets between the second and the third residual blocks. According to the characteristic that catenary insulators may locate at the same positions in similar images, we divided catenary images into different categories based on the perceptual hash algorithm. Then, for each type of images, a candidate region scanning strategy was used to speed up the detection of insulators. The experimental results show that our approach can improve the accuracy of insulator detection from 93.6% to 99.2%, and the time efficiency is improved by 46%.

**Keywords:** YOLO V3; railway catenary; insulator detection; residual block

随着科技的发展,我国铁路里程数已经一跃成为世界第一,高铁营业里程已达  $2.9 \times 10^4$  km,超过世界高铁总里程的 2/3,成为世界上高铁里程最长、运输密度最高、成网运营场景最复杂的国家。在这种情况下,铁路的安全维护便成为了一个难题。现阶段,我国对于铁路接触网的维护和检修,

是依靠接触网悬挂状态检测监测装置(4C)车<sup>[1]</sup>对支柱杆进行拍摄,然后通过大量人工对每张图像进行查看,进行故障的排查。4C 系统采集的图像均为高分辨率图像,需要人工对图像进行多次缩放操作查看,这样的检测方式受限于技术人员的个人经验水平、注意力的集中程度等诸多不可控

收稿日期:2019-06-28

基金项目:湖北省技术创新重大专项基金(2019AAA045)。

作者简介:陈智羽,硕士研究生, E-mail: 438824368@qq.com

\*通讯作者:闵 锋,博士,副教授, E-mail: 123018298@qq.com

引文格式:陈智羽,闵锋. 基于改进 YOLO V3 的接触网绝缘子检测方法[J]. 武汉工程大学学报, 2020, 42(4): 462-466.

因素,且存在故障筛查时间长、漏检风险高等问题。对于我国铁路的里程数来说,人工的方式是无法及时发现故障,保证铁路安全的。

随着深度学习技术的不断发展<sup>[2]</sup>,各种基于深度学习的目标检测<sup>[3-4]</sup>和目标跟踪算法<sup>[5]</sup>被提出。深度卷积神经网络可以通过收集大量目标物体的图像样本,进行训练,得到目标物体的检测模型,所以神经网络被运用在各个领域中,如:人脸识别<sup>[6]</sup>、视网膜血管分割<sup>[7]</sup>等。而目前广泛使用的目标检测算法分为双步目标检测算法和单步目标检测算法两种。双步目标检测算法常用的有:快速的区域卷积神经网络(fast region-based convolutional neural networks, Fast R-CNN)<sup>[8]</sup>,加速的区域卷积神经网络(faster region-based convolutional neural networks, Faster R-CNN)<sup>[9]</sup>,分割掩模的区域卷积神经网络(mask region-based convolutional neural networks, Mask R-CNN)<sup>[10]</sup>等,该类算法是先生成一系列样本的候选框,再通过卷积神经网络在这些候选框中进行样本分类,以此完成对目标的检测和识别。而单步目标检测算法则不用产生候选框,直接将目标边框定位的问题转化为回归问题处理,因此检测速度较快。该类算法常用的有:单点多盒探测(single shot multi box detector, SSD)<sup>[11]</sup>,统一的实时目标检测(you only look once: unified, real-time object detection, YOLO)<sup>[12]</sup>,第三代统一的实时目标检测算法(you only look once: unified, real-time object detection V3, YOLO V3)<sup>[13]</sup>等。考虑到我国铁路接触网的里程长度,所需要检测的样本数量巨大,选用更加迅速的单步目标检测算法。YOLO是最先使用回归的方法,直接从一张图像中检测物体的边界框坐标与类别,采用视觉几何组16(visual geometry group 16, VGG16)作为基础网络的,其速度是Faster R-CNN的3倍,但由于YOLO<sup>[14]</sup>采用网格划分的思想,对每个单元格都预测其中包含物体的置信度及其类别分数,造成YOLO的检测精度偏低,对小物体及较密集的物体检测效果较差。YOLO V3具有较快的检测速度和较高的检测精度,所以本文选择YOLO V3对接触网绝缘子进行检测。

目前,对于绝缘子的研究集中在输电线上的绝缘子,对于铁路接触网的绝缘子研究较少。可以借鉴输电线路中绝缘子的研究方法,对铁路接触网中绝缘子检测作指导。由于输电线路上没有4C巡检车,所以输电线路的方法多是基于航拍的图像中的绝缘子进行研究,文献<sup>[15-16]</sup>为解决绝缘子

的错检、漏检等问题,将绝缘子分解成多个存在交集的可变型部件,并对其进行检测。在保证子目标检测精度与速度的前提下,利用各部件之间相交区域的特征及含义,对其进行聚合并重新定义,使检测到的目标区域更准确。由于铁路上有4C车对接触网进行拍摄,相较于无人机拍摄图像距离更近,设备更稳定,成像更清晰,获取更方便。所以在已有4C图像的情况下,直接以4C图像作为数据集进行训练是一种更好的选择。

## 1 改进YOLO V3模型

### 1.1 YOLO V3网络结构改进

由于目前对接触网绝缘子研究较少,没有公共数据集,需自己制作数据集。4C车拍摄铁路接触网图像如图1所示,从图1可以看到铁路接触网分为单杆和双杆,单杆较为简单,而双杆相对来说较为复杂。接触网上面比较重要的部件,有双耳套筒、抱箍、绝缘子和底座等。使用LABELIMG程序,对图像中的重要部件进行框选、分类、制作,用5 258张铁路接触网图像作为训练数据集,另外选取500张图像作为测试数据集。

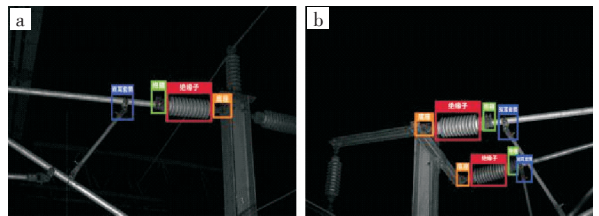


图1 接触网图像:(a)单杆,(b)双杆  
Fig. 1 Catenary image:(a)single rod ,(b)double rods

为了能清楚的表现各个部件的状态细节,4C拍摄的单张图像大小一般为6 576像素×4 384像素。而双耳套筒和绝缘子底座在图像中的像素只有不到500像素×500像素的大小,只占整个图像的0.9%不到。YOLO V3是以回归思想作为算法的基础,好处是计算速度快,但对小目标的检测准确率不够高。同时,YOLO V3在检测时,会出现重定位现象,如图2所示,接触网底座位置出现了多个检测框。

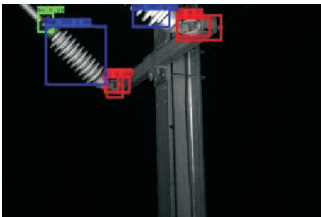


图2 YOLO V3检测结果图  
Fig. 2 Detection result image of YOLO V3

YOLO V3 采用的是 53 层的特征提取 (Darknet-53)<sup>[14]</sup> 网络结构, 该网络主要是由一系列的 1×1 和 3×3 的卷积层 (convolutional layers, CL) 组成, 因为此网络中有 53 个 CL, 所以叫做 Darknet-53 (53 层网络不包括 Residual 层)。

Darknet-53 借鉴了残差网络 (residual network, RN) 的想法, 由 5 个残差块 (residual block, RB) 组成, 每个残差块又是由多个残差单元 (resnet unit, RU) 组成。而上述结构是以输入图像为 256 像素×256 像素进行检测的, 更为常用的图像为 416 像素×416 像素, 均为 32 的倍数。输入 416 像素×416 像素的结构检测的 3 个特征层大小分别为 52×52, 26×26, 13×13。

其中 13×13 层负责检测大型目标, 26×26 层负责检测中型目标, 52×52 层负责检测小型目标。因此, 需要针对小目标检测, 对网络的结构进行修改。在 Darknet-53 的第 2 个残差块到第 3 个残差块之间增加一层残差块, 利用这个 4 倍降采样的特征融合目标检测层对小目标进行训练, 提高对小目标检测的准确率。增加残差块的网络结构如图 3 所示。

	类型	卷积数量	卷积大小	特征图大小
1x	卷积	32	3×3	416×416
	卷积	14	3×3/2	208×208
	卷积	32	1×1	208×208
	卷积	64	3×3	
	残差			
2x	卷积	128	3×3/2	104×104
	卷积	64	1×1	
	卷积	128	3×3	
4x	残差			52×52
	卷积	256	3×3/2	
	卷积	128	1×1	
	卷积	256	3×3	
	残差			
8x	卷积	256	3×3/2	26×26
	卷积	128	1×1	
	卷积	256	3×3	
	残差			
8x	卷积	512	3×3/2	13×13
	卷积	256	1×1	
	卷积	512	3×3	
	残差			
4x	卷积	1024	3×3/2	13×13
	卷积	512	1×1	
	卷积	1024	3×3	
	残差			

图 3 增加残差块的网络结构

Fig. 3 Network structure of added rsidual block

图 3 中, 前面的数字 4 表示 4 倍降采样, 128, 256 表示卷积过滤器的数量, 1×1 和 3×3 表示卷积的大小, 52×52 表示输出特征图的大小。

1.2 定位框去重与算法加速

针对重定位现象, 采用定位框去重算法。即在检测时对定位框进行判断, 如果该定位框与其他定位框的标签一样, 且重叠区域超过一定的阈值, 便寻找两个定位框的最大外接矩形, 让最大外接矩形作为新的定位框, 代替原本的两个定位框, 使得两个定位框进行融合。实验时, 阈值取自自身区域的 80%, 由于拍摄角度原因, 在拍双杆的时候, 可能存在同一部件相互遮挡的问题, 故阈值不能设置过低。

因同一路段的铁路接触网具有同质性, 所以在同一路段下拍摄的 4C 的图像就具有相似性, 绝缘子部件的位置大体相同。所以对于相似图像, 可以用上一张图像的检测结果, 作为下一张图像的检测候选区域, 只需要在候选区域附近搜索, 而不需要全图搜索, 可以极大的提高检测速度。判断两张图像的相似性, 采用感知哈希算法。该算法给每张图像生成一个“指纹”, 两张“指纹”字符串汉明距离越小, 说明两张图像越相似。通过指纹的相似度, 可对图像进行分类。当判断两张图像属于同一类时, 将第一张图像定位区域作为后一张图像的检测候选区域, 只在候选区域扫描, 避免了全图大范围扫描, 从而实现对检测算法的加速。

1.3 算法流程

YOLO V3 算法对输入的图像, 都会被放缩为 416 像素×416 像素大小, 然后进行训练。在 416 像素×416 像素的训练之后, 采用多尺度检测, 检测结果显示, 与训练大小保持一致, 即在 416 像素×416 像素的尺度下, 检测效果最好。

改进后的算法流程如图 4 所示: 首先用 LABELIMG 程序制作训练所需的数据集, 另选取 500 张图像用作测试; 然后用改进后的算法对制作的数据集进行训练, 得到训练后的模型; 之后检测图像, 检测前计算图像的哈希码, 通过哈希码判断该图像的相似性。如果相似则采用候选区域扫描策略进行检测, 否则采用全图扫描策略进行检测; 初步检测完成后, 进行检测框的去重; 最后完成检测, 输出检测后的图像。



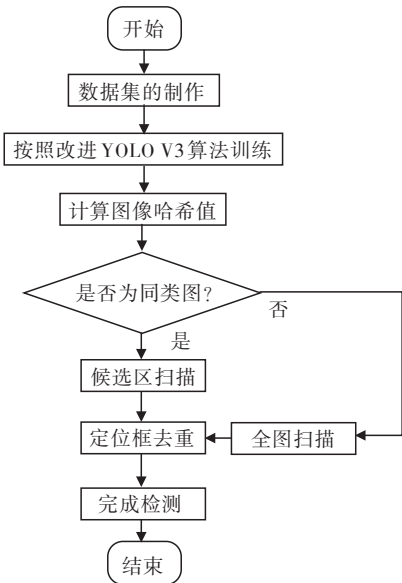


图4 绝缘子检测流程图  
Fig. 4 Flow chart of insulator detection

2 结果与讨论

使用5 258张4C接触网图像进行训练,并在416像素×416像素的尺度下检测。然后与YOLO V3,FAST R-CNN的检测结果进行对比。测试结果如表1所示,由于3种方法都不存在误识别,故不进行统计。

通过表1可以看出,对比FAST R-CNN,无论是YOLO V3还是改进后的算法,准确率都有很大提升,耗时也有大幅度的下降。改进后YOLO V3的漏检数有明显下降,准确率有较大的提升,绝缘子检测的准确率从93.6%提升至99.2%,证明增加一个对小目标友好的4倍降采样的残差块,提高了算法对于小目标检测的准确率。

表1 绝缘子检测结果统计表					
Tab. 1 Statistical table of insulator detection results					
方法	测试 图像数	正确 检测数	漏检 数	准确率/ %	耗时/ s
YOLO V3	500	468	32	93.6	537
改进的YOLO V3	500	496	4	99.2	384
FAST R-CNN	500	416	84	83.2	2 143

使用感知哈希的加速效果依赖于图像中待检区域的个数,当待检区域数量较少时,其候选区域少,加速效果更加明显。单杆图像平均耗时由1.32 s减少到0.53 s,双杆图像平均耗时由1.32 s减少到0.9 s。使用同一相机在同一条线路中连续拍摄的200张图像作为测试样本,测试结果如表2所示,使用全图扫描策略约耗时264 s,使用候选区域扫描约耗时143 s,其中146张利用了候选区域扫

描。通过感知哈希算法分类图像,对同类图像采用候选区域扫描策略加速,速度提升46%,如果连续处理的图像数量越多其速度提升越明显。

表2 加速效果统计表		
Tab. 2 Statistical table of time efficiency analysis		
扫描策略	图像数量	耗时 / s
全图扫描	200	264
候选区域扫描	200	143

改进YOLO算法检测结果如图5所示。从图5可以看出,改进后的算法不仅对小目标的检出率很高,而且置信度保持在很高的分数。在图5(a)中,接触网的上方,有一个与绝缘子一样的部件,但此部件并不在检测范围内,为干扰项。由于与绝缘子的角度不同,因此被算法排除;图5(b)中,有许多杆间支撑装置,与双耳套筒使用了相同的零件,该零件并不在检测范围内,为干扰项。由于此部件有弯折,因此被算法排除;图5(c)为双杆接触网图像,外部的支撑杆对内部杆上的抱箍造成了遮挡,使其像素信息不全,但该算法依然将其检出。通过这些例子可以看出,在复杂情况下,算法取得了较好的检测结果,表明改进后的算法有良好的检测能力与抗干扰能力。

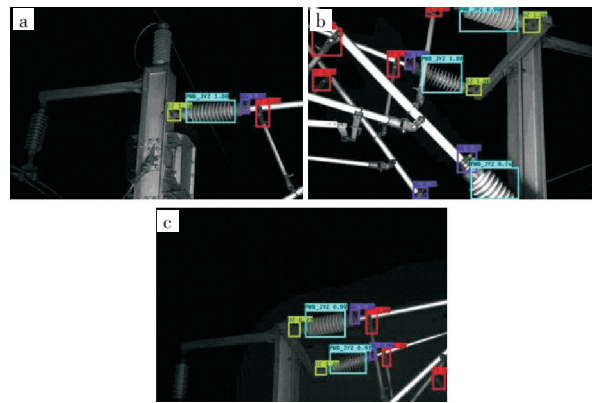


图5 检测结果图:(a)绝缘子干扰图,  
(b)双耳套筒干扰图,(c)抱箍遮挡图  
Fig. 5 Test results of images: (a) insulator blocked image,  
(b) clevis end holder for tubule blocked image,  
(c)holt hoop iron blocked image

3 结 论

综上所述,本文提出了一种基于改进YOLO V3的绝缘子检测方法,实验结果表明,该方法可以检测出铁路接触网上的绝缘子部件,不仅检出率高,而且速度快,抗干扰能力强,适用于大多数环境,基本满足铁路接触网检测要求的准确性和快速性。下一步工作将在绝缘子检测定位的基础

上,对绝缘子的缺陷检测展开研究。

参考文献

[1] 铁运[2012]136号,高速铁路供电安全检测监测系统(6C)系统总体技术规范[S]. 北京:中国标准出版社,2012.

[2] 张荣,李伟平,莫同. 深度学习研究综述[J]. 信息与控制,2018,47(4):385-397.

[3] 华夏,王新晴,王东,等. 基于改进 SSD 的交通大场景多目标检测[J]. 光学学报,2018,38(12):221-231.

[4] 王文秀,傅雨田,董峰,等. 基于深度卷积神经网络的红外船只目标检测方法[J]. 光学学报,2018,436(7):160-166.

[5] 罗海波,许凌云,惠斌,等. 基于深度学习的目标跟踪方法研究现状与展望[J]. 红外与激光工程,2017,46(5):6-12.

[6] 陈希彤,卢涛. 基于全局深度分离卷积残差网络的高效人脸识别算法[J]. 武汉工程大学学报,2019,41(3):276-282.

[7] 单玲玉,闵锋,李延达. 全局阈值与局部阈值相结合的视网膜血管分割方法[J]. 武汉工程大学学报,2015,37(3):66-71.

[8] GIRSHICK R. Fast R-CNN[C]//IEEE International Conference on Computer Vision and Pattern Recognition. Brussels: IEEE, 2015:1440-1448.

[9] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-Time object detection with region proposal

networks [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. Boston: IEEE.2017: 1137-1149.

[10] HE K M, GKIOXARI G. DOLLAR P, et al. Mask R-CNN [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence.Boston: IEEE. 2017: 386-397.

[11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//European Conference on Computer Vision.Cham:Springer.2016: 21-37.

[12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// Computer Vision and Pattern Recognition.Las Vegas:IEEE 2016: 779-788.

[13] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2018-4-8]. <https://arxiv.org/abs/1804.02767>.

[14] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Computer Vision and Pattern Recognition. Hawaii :IEEE. 2017: 6517-6525.

[15] 杨罡,张娜,晋涛,等. 基于 YOLO 算法的无人机航拍图片绝缘子缺失检测[J]. 科技资讯,2018,16(24):30-31.

[16] 高强,廉启旺. 航拍图像中绝缘子目标检测的研究[J]. 电测与仪表,2019,56(5):119-123.

本文编辑 陈小平