

文章编号:1003-1995(2020)09-0123-05

# 基于 YOLOv3 算法的轨道扣件自动定位与检测

刘奇锋<sup>1,2</sup> 许贵阳<sup>1,2</sup> 白堂博<sup>1,2</sup>

(1.北京建筑大学 机电与车辆工程学院,北京 100044;

2.北京建筑大学 城市轨道交通车辆服役性能保障北京市重点实验室,北京 100044)

**摘 要** 针对目前轨道扣件人工检测效率低、准确率低等问题,提出了基于 YOLOv3 算法的轨道扣件自动定位及检测方法。采集有砟轨道和无砟轨道的扣件图像并进行标注,通过 K-means 聚类确定预设边界框大小;为了更好地检测到细粒度特征,采用 4 个不同尺度的特征图来进行对象检测;对 Darknet-53 网络进行改进,有利于解决深层次网络的梯度问题,增加轨道扣件目标识别模型的识别效果。试验结果表明,该方法对轨道扣件目标识别效果较好,检测准确率较高。

**关键词** YOLOv3 算法;轨道扣件;深度学习;自动定位与检测;特征图;细粒度

**中图分类号** U213.5+3 **文献标识码** A **DOI**:10.3969/j.issn.1003-1995.2020.09.29

轨道扣件的功能是将钢轨固定,使钢轨保持轨道间距,同时避免钢轨相对于轨枕的纵横向移动<sup>[1-2]</sup>。常用轨道扣件的紧固件分为六角螺栓和挂钩螺栓 2 种。在轨道巡检过程中,常见的轨道扣件异常状况包括扣件移位、扣件部分断裂、扣件丢失等,是列车正常运行的巨大隐患。

过去,由专门检测人员采用人工巡检的方式进行轨道扣件日常检查,需要逐个进行排查,存在许多缺点,如效率低下、工人对异常情况和潜在危险程度估计不准确等。目前,高速列车技术日益成熟,车速不断提高,对巡线的自动化与检测的智能化要求也越来越高。随着图像处理与深度学习等人工智能技术的迅猛发展,铁路部门出于缩减人工检测成本、提高运营安全等方面的考虑,在铁路巡检中越来越多地使用了自动铁路巡检装备及配套的图像处理技术<sup>[3-4]</sup>。

近年来,我国的轨道检测多采用将具有计算机视觉和图像处理技术的巡线设备安装在轨检车中,对轨道进行全自动检测。自动巡检过程中,对轨道进行拍摄,采集轨道图像,再基于神经网络或其他人工智能技术对轨道图像进行处理,实现对扣件状态的自动检测。这一方法摆脱了传统检测中依赖巡道工进行现场检测和维修的状况,成为扣件检测的新趋势<sup>[5]</sup>。

本文提出一种基于 YOLOv3 深度学习算法的扣件定位与检测的方法。该方法使用轨检小车采集轨道图像,通过 YOLOv3 网络进行训练与学习,达到自动识别轨道图像中扣件状态的效果<sup>[6]</sup>,并将其检测效果与人工检测的情况进行对比。

## 1 检测算法

### 1.1 深度学习算法简介

传统的目标检测技术利用检测窗口在图片上滑动来搜索检测目标。其技术优化主要表现为检测窗口的选择、特征的设计、分类器的设计等。检测速度慢,检测精度也不能保证。

随着深度学习算法在计算机视觉领域的应用,基于深度学习算法的目标检测技术得到迅猛发展。目前主要分为 2 大发展方向:①Two-Stage 目标检测算法,例如 Faster RCNN(Region Convolutional Neural Network)<sup>[7]</sup>, R-FCN(Region-Based Fully Convolutional Networks)<sup>[8]</sup>, FPN(Feature Pyramid Networks)<sup>[9]</sup>。首先将图像的前景和背景进行分类及边界框回归,在待检测目标可接受的召回率的前提下,候选窗口的数量可以控制在几千或几百个。再根据之前的输出窗口进行分析,而不是像传统的检测方法那样对整个图像的每个区域进行分析。②One-Stage 目标检测算法,例如 YOLO(You Only Look Once)系列算法<sup>[10-12]</sup>、SSD(Single Shot Multibox Detector)算法<sup>[12]</sup>。这类算法是将图像进行栅格化处理,直接定位检测目标的中心点,在指定栅格上使用不同层级、不同长宽比密集采样生成检测锚框,来直

收稿日期:2020-01-14;修回日期:2020-05-06

第一作者:刘奇锋(1996—),男,硕士研究生。

E-mail:liu13264319431@163.com

通信作者:许贵阳(1971—),男,教授,博士。

E-mail:xuguiyang@bucea.edu.cn

接进行分类和回归分析。与Two-Stage算法相比,One-Stage算法检测速度更快,能满足对物体进行实时检测的要求,但目标识别精度相对较低。在实际应用中,往往要求在精度不太低的情况下满足实时动态检测,所以One-Stage目标检测算法使用范围更广。YOLOv3算法借鉴了其他算法的长处,既保持了One-Stage目标检测算法的检测速度,又具有较高的目标检测精度。

## 1.2 YOLOv3 算法

相比之前的YOLO系列算法,YOLOv3算法主要在下列3部分进行了改进:

1)采用新的特征提取网络。使用Darknet-53替换YOLOv2使用的Darknet-19<sup>[13]</sup>,在满足检测实时性的同时具有更高的准确率。Darknet-53是一个53层的卷积网络,YOLOv3使用该卷积网络的前52层,所以整个YOLOv3的结构里面没有池化层和全连接层,有助于防止由池化导致的低级特征丢失。

2)利用多尺度特征对图像进行检测。YOLOv3算法采用3个不同尺寸预测对象,这3个不同尺寸来自不同层级的卷积层输出。该方法借鉴了Feature Pyramid Network的思想:由于卷积层每隔几层,Feature Mapping的宽和高就会减少而通道数会增加,随着网络层次的加深,特征映射组成的形状类似于金字塔,将不同层级的特征映射转换为最终的输出有助于提升模型在对象不同尺度大小上的表现,也就是提高模型从小目标到大目标的综合检测能力。

3)对象分类用Logistic函数取代Softmax函数。之前的YOLO系列算法,在分类模型中进行最后的输出前使用Softmax函数求出每个类别的概率,各类别是互斥的。YOLOv3算法的每个类别概率单独用逻辑回归函数即Sigmoid函数计算得出,所以各类别不必是互斥的,一个对象可以被预测出多个类别。

## 2 YOLOv3算法模型适应性改进

YOLOv3算法使用的特征提取网络Darknet-53网络采用全卷积结构,如图1所示。

利用YOLOv3算法模型输入图片后,使用双三次插值法对输入对象进行放缩,在图像放缩过程中进行以灰色为背景的贴图操作,即把放缩的图片贴在416×416的灰色背景图上,使算法模型输入的对象大小一致,均为416×416。算法实现过程中有5次下采样,每次均使用步长为2的卷积。因此,特征提取网络的最大步幅为2<sup>5</sup>。特征提取后,在YOLOv3算法模型网络中对生成的特征图进行上采样、Route等操作,达到分尺度采样的目的;分别在32倍、16倍、8倍降采样时进行检测,得到13×13、26×26、52×52这3个层次的输出

	Type	Filters	Size	Output
1×	Convolutional	32	3×3	256×256
	Convolutional	64	3×3/2	128×128
	Convolutional	32	1×1	
	Convolutional Residual	64	3×3	128×128
2×	Convolutional	128	3×3/2	64×64
	Convolutional	64	1×1	
	Convolutional	128	3×3	
	Residual			64×64
4×	Convolutional	256	3×3/2	32×32
	Convolutional	128	1×1	
	Convolutional	256	3×3	
	Residual			32×32
8×	Convolutional	512	3×3/2	16×16
	Convolutional	256	1×1	
	Convolutional	512	3×3	
	Residual			16×16
16×	Convolutional	1 024	3×3/2	8×8
	Convolutional	512	1×1	
	Convolutional	1 024	3×3	
	Residual			8×8

图1 Darknet-53网络

特征图。利用前面的低层特征,使后面的高层能够使用细粒度(Fine Grained)特征和更多的语义信息。最后1个尺寸输出使用前2个尺寸计算的特征映射,使最后的尺寸输出也能使用细粒度特征。

对应不同层次的特征图,模型对图片的视野感受也不同。在最小的13×13特征图(输出的第1层次)上有最大的感受视野,应使用较大的预设边界框,适于检测较大面积的对象;中等的26×26特征图(输出的第2层次)提供中等感受视野,应使用中等预设边界框,适于检测中等面积的对象;较大的52×52特征图(输出的第3层次)提供最小的感受视野,应使用较小的预设边界框,适于检测较小面积的对象。

相比轨道整体,扣件图片占据视野较小,模型训练与检测结果不太理想,须对YOLOv3算法模型进行适应性改进。在Darknet-53的基础上加1个步长为2的卷积核,同时添加4组残差组件,由4组1×1和3×3卷积构成,每组输入均由前面所有组的输出经过通道拼接而获得。将输入图片进行缩放处理,变成大小为512×512的图片。经过6次步长为2的卷积核下采样后,模型形成8×8、16×16、32×32、64×64这4个层次的特征图,增大了模型的感受视野,强化了对大物体检测的影响,提高了图片处理效果。

## 3 试验验证

### 3.1 数据集建立

利用某新型轨道巡检仪进行数据采集。该轨道巡检仪由电动巡检小车和轨道状态巡检系统2部分组成,其中巡检系统可以动态采集轨道表面数据,覆盖整个轨道结构表面,采集到的高清轨面图像像素分辨率最高可达0.3 mm,并具有海量图像数据实时压缩存储及显示加速功能。

对采集到的15 320张铁路轨面图片进行预处理,



选出其中的扣件信息并根据扣件状态将其分为4类:扣件正常(zc)、扣件丢失(ds)、扣件断裂(dl)、扣件移位(yw)。

采用Labeling软件对图像样本数据进行处理,将需要确定的对象(正常扣件)使用方框标注出来,将得到的标注信息生成.xml文件,得到感兴趣区域(Region of Interest, ROI),从而生成扣件图像的训练集和验证集数据,见图2。其中标注信息包括标注框左上角横坐标 $X_{\min}$ 、左上角纵坐标 $Y_{\min}$ 、右下角横坐标 $X_{\max}$ 、右下角纵坐标 $Y_{\max}$ 、标注物体种类 $C$ 。

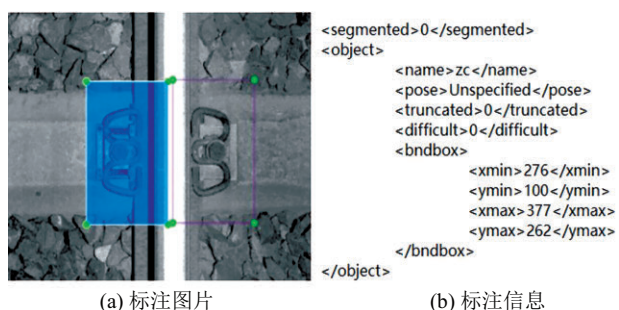


图2 数据集制作

### 3.2 预设边界框的选取与预测

对于YOLOv3模型,选取适合的预设边界框的参数不仅关系到其训练的收敛速度,而且对于目标的检测精度也会有影响。本文对扣件目标框采用K-means聚类的方法来确定先验框参数。

根据K-means算法的思想,将所有点作为1个簇,将该簇一分为二,再选择能最大程度降低聚类代价函数(即误差平方和)的簇并将其划分为2个簇,或采用选取最大的簇等其他选择方法,得出最优的聚类结果来计算平方误差和(Sum of the Squared Error)。平方误差和 $S_{SE}$ 的计算公式为

$$S_{SE} = \sum_{i=1}^K w_i (y_i - y^*)^2 \quad (1)$$

式中: $K$ 为簇的个数; $w_i$ 为权重; $y_i$ 为第 $i$ 个簇的中心点(质心); $y^*$ 表示该簇所有点的平均值。

依次进行下去,直到簇的数目等于用户给定的数目 $K$ 为止。本文进行4个尺度的特征图分析,每个层次赋予3个预设边界框。因此将 $K$ 设成12,得到数据的12组预设边界框分别为:(86,116),(96,170),(98,154),(100,161),(102,172),(102,134),(104,148),(106,165),(106,157),(113,131),(119,133),(127,145)。

而后,模型对图片进行栅格化处理,在扣件中心预设3个预测边界框,然后将预测出来的边界框的宽和高分别与预设边界框的宽和高绑定,这样无论开始时预测边界框输出的宽和高是多少,经过转化后都是与预设边界框的宽和高相关。经过多次惩罚训练后,

就可以对相应的边界框进行预测。

利用Sigmoid函数将预测偏移量缩放到0到1之间,使预设边界框的中心坐标固定在一个栅格中,加快网络收敛。实际输出的边界框参数为

$$b_x = \sigma(t_x) + c_x \quad (2)$$

$$b_y = \sigma(t_y) + c_y \quad (3)$$

$$b_w = p_w e^{t_w} \quad (4)$$

$$b_h = p_h e^{t_h} \quad (5)$$

式中: $b_x$ 和 $b_y$ 分别为最终预测输出目标边界框的中心横、纵坐标; $t_x$ 和 $t_y$ 分别为网络预测的边界框中心在横、纵坐标方向上的偏移量; $\sigma$ 为Sigmoid函数; $c_x$ 和 $c_y$ 分别为特征图上预设边界框中心的横、纵坐标值(即扣件中心栅格左上角与特征图的相对位置信息); $b_w$ 和 $b_h$ 分别为转换后预测边界框的实际宽和高,即最终预测输出目标边界框的宽和高; $t_w$ 和 $t_h$ 分别为宽和高的缩放比值; $p_w$ 和 $p_h$ 分别为特征图上预设边界框的宽和高。

将试验所用的扣件图像作为特征图进行边界框预测,如图3所示。

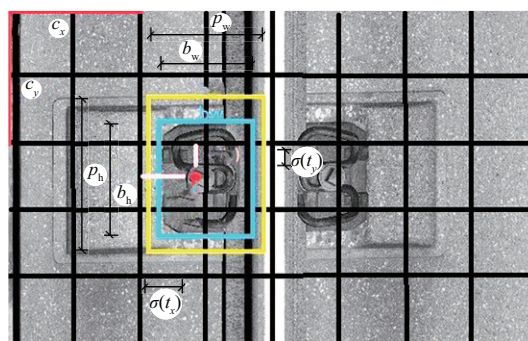


图3 预测相对位置

### 3.3 结果分析

将采集到的20 000张有砟和无砟轨道图片(其中包括预先挑选的负样本数据)作为测试数据集,使用训练好的模型对其进行检测分析。与人工检测结果对比,检出率(检测出的负样本数目占总体负样本数目的百分比)为91.5%,误报率(将正样本错误标注为负样本的数目占整体数据的百分比)为0.5%。选取部分有砟轨道和无砟轨道的扣件目标识别图像(图4、图5),分析得出扣件识别结果,见表1。

本试验预设的扣件状态种类的阈值(扣件种类判断置信度的最低值)为0.8。当网络判断扣件的种类置信度超过0.8时,认为它是相应种类,否则不识别此扣件状态。分析测试结果可知,此试验训练模型已初步具备对轨道扣件进行定位识别的能力,同时将正常或者异常状态进行标注,并将异常状态的图片进行汇总,再进行定点监测,可以减少巡线工作量,提高铁路轨道的巡线检测效率。

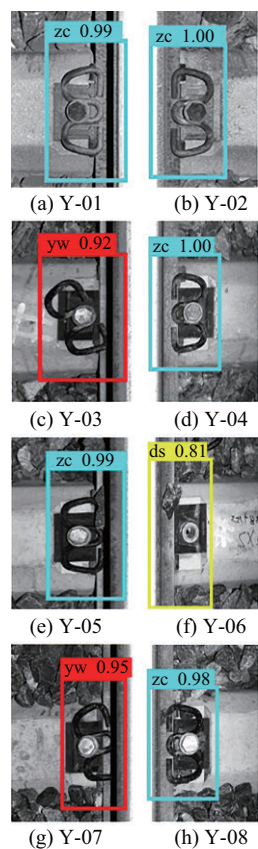


图4 部分有砟轨道扣件目标识别图像

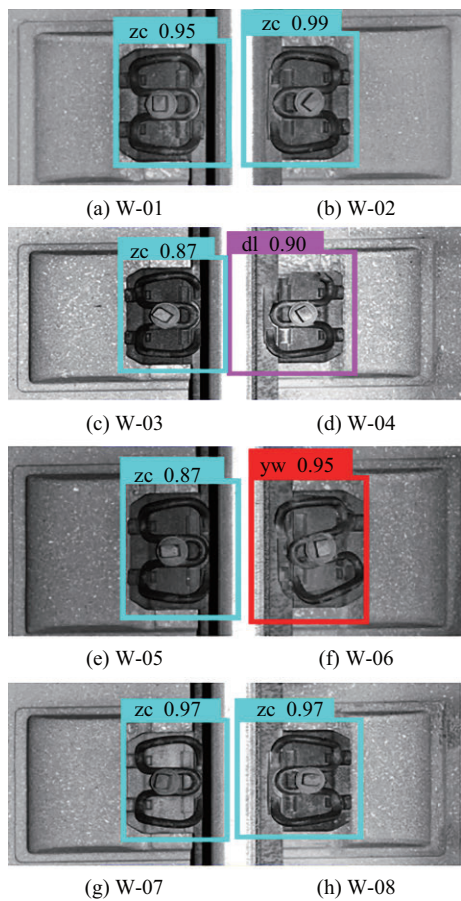


图5 部分无砟轨道目标识别图像

表1 YOLOv3算法扣件识别结果						
编号	扣件状态 种类	置信度	边界框参数			
			$X_{min}$	$Y_{min}$	$X_{max}$	$Y_{max}$
Y-01	zc	0.99	1 623	703	2 178	1 671
Y-02	zc	1.00	2 221	730	2 747	1 648
Y-03	yw	0.92	1 636	2 801	2 163	3 751
Y-04	zc	1.00	2 241	2 824	2 735	3 693
Y-05	zc	0.99	1 342	2 291	1 858	3 293
Y-06	ds	0.81	1 893	2 222	2 460	3 361
Y-07	yw	0.95	1 636	2 801	2 163	3 751
Y-08	zc	0.98	2 241	2 824	2 735	3 693
W-01	zc	0.95	1 234	2 086	1 825	2 842
W-02	zc	0.99	1 819	2 050	2 518	2 859
W-03	zc	0.87	1 559	526	2 163	1 259
W-04	dl	0.90	2 214	496	2 788	1 293
W-05	zc	0.87	1 237	2 124	2 476	2 882
W-06	yw	0.95	1 877	2 105	1 838	2 904
W-07	zc	0.97	1 540	2 674	2 186	3 371
W-08	zc	0.97	2 216	2 634	2 186	3 371

4 结语

本文研究了基于深度学习算法的轨道扣件自动定位与异常状态检测,利用检测速度相对较快且具有高目标检测精度的YOLOv3算法对轨道扣件进行自动检测,满足了快速准确识别扣件的需求,减少了人工检测的工作量,具有一定的实际意义。

试验过程中也暴露出一些问题,主要是轨道扣件负样本相对于正常扣件的数量较少,不能达到很好的训练效果,容易出现过拟合现象。因此,在后续工作中可进行如下改进:①继续使用轨道检测小车采集更多的轨道图像,找寻更多的负样本数据;②通过对抗生成网络来进行训练,自动生成异常状态扣件图像,如扣件移位、扣件断裂、扣件丢失等,达到一定的负样本训练数目,从而更加准确地检测轨道扣件状态。

参 考 文 献

[ 1 ]李欢,柴晓东,彭乐乐,等.轨道扣件检测系统设计与分析[J].计算机测量与控制,2019,27(2):43-46,64.

[ 2 ]林菲,杨子明,李永光,等.钢轨扣件检测技术研究[J].中国铁路,2019(6):103-110.

[ 3 ]苏靖棋.轨道检测技术的发展趋势[J].现代城市轨道交通,2019(8):112-114.

[ 4 ]付文龙,王培俊,李文涛,等.一种新型铁路轨道检测小车的设计与分析[J].中国机械工程,2019,30(3):300-305.

[ 5 ]王一军,张东,杨勇.一种基于结构光的扣件缺失判别新方法[J].科学技术与工程,2016,16(13):262-265.

- [6] 王兵水, 郑树彬, 李立明, 等. 基于YOLO改进算法的轨道扣件状态检测研究[J]. 智能计算机与应用, 2020, 10(1): 137-143.
- [7] 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.
- [8] DAI J F, LI Y, HE K M, et al. R-FCN: Object Detection via Region-Based Fully Convolutional Networks [C]//Proceedings of the International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2016.
- [9] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016.
- [11] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017.
- [12] REDMON J, FARHADI A. YOLOv3: an Incremental Improvement [EB/OL]. [2018-12-25]. <http://arxiv.org/abs/1804.02767>.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector [M]//LEIBE B, MATAS J, SEBE N, et al. Computer Vision-ECCV 2016. Cham: Springer, 2016.

## Automatic Positioning and Detection of Track Fastener Based on YOLOv3 Algorithm

LIU Qifeng<sup>1,2</sup>, XU Guiyang<sup>1,2</sup>, BAI Tangbo<sup>1,2</sup>

(1.School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 2.Beijing Key Laboratory of Performance Guarantee on Urban Rail Transit Vehicles, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

**Abstract** Aiming at the problems of low efficiency and low accuracy of manual detection for track fastener, a method of automatic positioning and detection of rail fastener based on YOLOv3 algorithm was proposed. The fastener images of ballasted track and ballastless track were collected and labeled, and the size of preset boundary box was determined by K-means clustering. In order to detect fine-grained features better, 4 feature maps with different scales were used for object detection. The improvement of Darknet-53 network is helpful to solve the gradient problem of deep network and increase the recognition effect of track fastener target recognition model. The experimental results show that the method has a good effect on track fastener target recognition, and the detection accuracy is high.

**Key words** YOLOv3 algorithm; track fastener; deep learning; automatic positioning and detection; feature map; fine grained

(编辑:苗蕾 校对:李付军)