

文章编号: 2095-2163(2021)12-0104-06

中图分类号: U216

文献标志码: A

# 基于YOLOv3改进算法的轨枕裂纹检测研究

童千倩, 李立明, 郑丹阳, 杨意, 汪晨曦, 郑树彬

(上海工程技术大学 城市轨道交通学院, 上海 201620)

**摘要:** 针对在现有轨枕裂纹检测中传统检测方法检测效率低和精度差的问题, 本文提出一种基于YOLOv3改进算法的轨枕裂纹检测模型。首先, 通过灰度投影法对原始有砟道床图像中的轨枕区域进行定位和分割; 其次, 将挤压与激活模块和空间金字塔池化引入改进的网络模型中, 有效抑制轨枕图像中其他冗余的特征干扰, 空间金字塔池化保证了原始图像的比例, 对分割得到混凝土轨枕区域的裂纹状态进行进一步的检测; 此外, 通过将改进网络模型中的损失函数回归预测部分改进为新的*CIoU*函数, 提高了模型的预测准确程度。经实验结果验证, 本检测方法与YOLOv3直接对轨枕进行定位相比, 对混凝土轨枕裂纹的检测, 精确度达到96.3%, 召回率达到91.2%, *MAP*达到91.5%, *FPS*达到76.6张/s。总体表现提升显著, 能够达到精确快速的检测效果。

**关键词:** 轨枕裂纹检测; 灰度投影法; YOLOv3改进; 空间金字塔池化

## Research on the improved algorithm of crack detection of sleeper based on YOLOv3

TONG Qianqian, LI Liming, ZHENG Danyang, YANG Yi, WANG Chenxi, ZHENG Shubin

(School of Urban Rail Transit, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** In order to solve the problems of low detection efficiency and poor precision of traditional detection method in the existing crack detection of sleeper, in this paper, a sleeper crack detection model based on the improved YOLOV3 algorithm is proposed. Firstly, the sleeper area in the original ballast track image is located and segmented by the gray projection method. Pyramid pooling and activation is introduced into the improved network model, which can effectively inhibit other redundancy in the sleeper image characteristics of interference. Space pyramid pooling ensures the proportion of the original image. The state of crack of concrete sleeper after segmentation is further tested. In addition, the loss function of the improved network model is changed to a new *CIoU* function, which improves the prediction accuracy of the model. The experimental results show that compared with YOLOv3 which directly locates the sleeper, the detection method in this experiment has an accuracy of 96.3%, a recall rate of 91.2%, a *MAP* of 91.5%, and an *FPS* of 76.6 / s in the detection of concrete sleeper cracks. The overall performance is improved significantly, which can achieve accurate and rapid detection.

**[Key words]** sleeper crack detection; gray projection method; improved YOLOv3; spatial pyramid pooling

## 0 引言

截止2020年底, 中国铁路总保有量达到 $146.3 \times 10^3$  km, 营业里程总数达到 $128 \times 10^3$  km, 铁路每万平方公里的路网密度达到152.3 km, 大规模的铁路建设对铁路基础建设和维护保障提出了较高的需求<sup>[1]</sup>。维持轨道的几何形位是铁路建设工程与维护保障中重要的部分, 轨道线路中, 轨枕充当着承载列车负荷、分散钢轨压力至道床的中间角色, 在列车运行中起到重要的作用。由于轨枕上截面的荷载弯矩大于轨枕的抗裂强度, 易导致混凝土轨枕表面产生裂纹, 若不及时检修, 会引起混凝土压溃, 导致轨枕失效, 严重地影响轨枕的使用寿命, 威胁轨道交通

的行车安全<sup>[2]</sup>。目前, 轨枕裂纹的检测方法主要依靠人工进行, 而人工巡检的主观性较强, 检测效率低<sup>[3]</sup>。近年来, 部分学者将无损检测, 如物理检测手段与计算机检测等手段运用到裂纹检测中。

物理检测通常从声、光、热、磁等角度提出检测方法, 对混凝土结构表面的微小异常进行判断, 典型的物理无损检测方法, 如: 声发射技术, 散斑干涉技术, 红外检测技术等。但是物理检测方法对环境要求较高, 检测成本高, 具有一定的局限性, 例如: 声发射技术需要一定的声耦合, 需要处于应力状态, 红外检测通常需要设备体积较大的仪器, 检测环境也有严格的限制<sup>[4]</sup>。计算机检测通常经过图像预处理, 图像特征提取, 图像特征处理, 最终得到裂纹检测结

**基金项目:** 国家自然科学基金(51975347, 51907117, 12004240); 国家级创新项目(202110856031); 市级创新项目(cs2010006)。

**作者简介:** 童千倩(2000-), 女, 本科生, 主要研究方向: 图像处理; 李立明(1982-), 男, 硕士, 高级实验师, 主要研究方向: 图像处理。

**通讯作者:** 李立明 Email: liliming@sues.edu.cn

**收稿日期:** 2021-07-02

果,具有低成本,高效率的特点。随着机器视觉技术与深度学习技术的发展,目标检测技术在轨道裂纹检测中应用广泛,例如 Faster R-CNN, SSD, YOLO 等,已经在裂纹图像检测方面取得了一定的检测成果。目前,Cha 等提出了一种基于 Faster R-CNN 的混凝土裂纹检测方法,该种优化后的 Faster R-CNN 网络能够在多分类的条件下快速地提取裂纹特征,达到较为精确的检测结果<sup>[5]</sup>;Mandal 等基于 DCNN 对道路混凝土裂纹提出了一种自动化的检测方法,但检测精度较低<sup>[6]</sup>;上述方法均针对路面裂纹提出,在有砟轨道的实际运用中取得的效果较差,Li 等基于改进的 YOLO 网络,叠加了原始 YOLO 中浅层卷积结构的低层特征,提取了深层卷积网络中的抽象语义特征,提高了轨道板裂纹的检测精度,增强

了小目标和重叠目标的检测能力,但是由于轨道板数据集的背景信息单一,该方法的通用性较差<sup>[7]</sup>;曹建斌通过结合 CF-Net 以及 FastCrack 模型的两阶段定位法,能满足自动检测轨枕裂纹的需求,但是其召回率与检测速度还需进一步的提升<sup>[8]</sup>。

针对小目标裂纹检测的方法存在通用性差,精度低,召回率低,检测速度慢的问题,本文设计了一种基于 YOLOv3 改进算法的轨枕裂纹检测模型,如图 1 所示。本模型首先使用灰度投影法结合经验值定位轨枕坐标并分割轨枕,然后将注意力机制与空间金字塔池化引入本文所述改进的 YOLOv3 模型中,并改善了裂纹检测损失函数,最终在保持速度的情况下,提高了检测的精度。

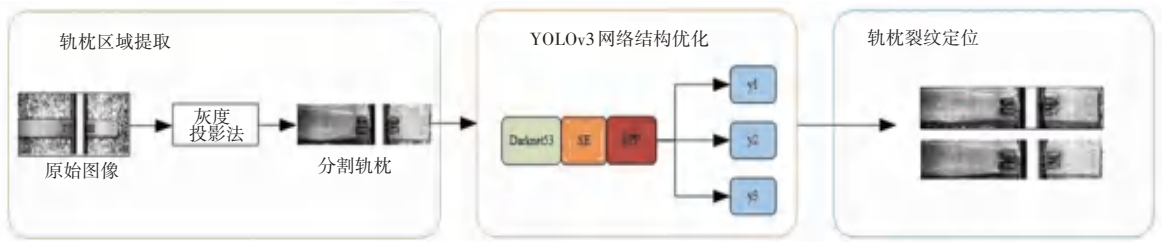


图 1 基于 YOLOv3 改进算法的轨枕裂纹检测模型示意图

Fig. 1 Schematic diagram of sleeper location segmentation and crack detection method

## 1 轨枕区域提取

由于道砟与轨枕的灰度存在差异,轨枕边缘特征明显,且轨枕水平放置于道床,垂直于钢轨,同时轨枕尺寸也有着严格的规定,可以利用灰度投影法结合经验值定位轨枕位置,具体流程图如图 2 所示。

灰度投影法针对复杂背景的物体边缘检测有着不错的效果,主要依赖波峰和波谷确定物体边缘位置的坐标<sup>[9]</sup>。针对原图中存在道砟背景的干扰,利用灰度投影法定位得到轨枕与光带边缘坐标,假设图像表示为  $f(x, y)$ , 在  $y$  方向的灰度投影函数为  $f_y(x)$ , 在  $x$  方向的灰度投影函数为  $f_x(x)$ , 图像中的像素点坐标为  $(x, y)$ , 则水平与垂直方向的灰度投影函数值为式(1)和式(2):

$$f_y(x) = \sum_y f(x, y) \quad (1)$$

$$f_x(x) = \sum_x f(x, y) \quad (2)$$

利用灰度投影法,计算出轨枕与光带边缘的位置,具体实验结果如图 2 所示。图 2(a) 表示有砟道床中钢轨与轨枕原图;图 2(b) 表示水平投影的波谷表示轨枕与道砟的接触边缘;图 2(c) 表示垂直投影

得到的波峰表示钢轨表面光带所在位置,根据灰度投影的实验结果能够得到光带与轨枕相交点的坐标。

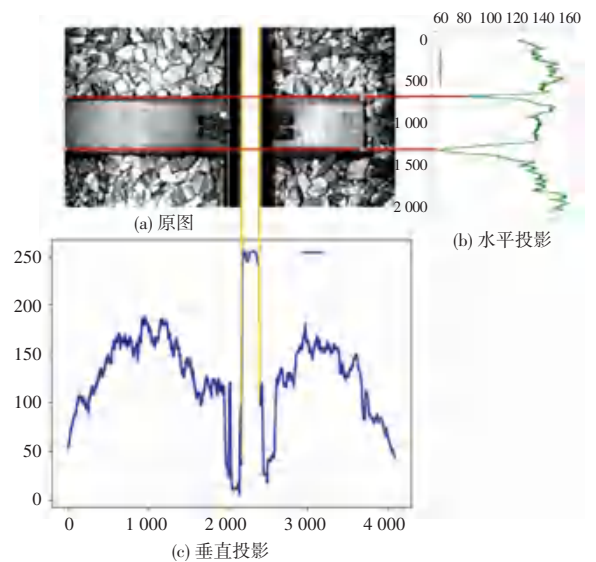


图 2 灰度投影实验结果图

Fig. 2 Grayscale projection experimental results

由于在采集的道床图片中,轨枕边缘的像素宽度相对固定,首先通过灰度投影法得到边缘坐标,然后叠加轨枕像素宽度的经验值,即可得到轨枕顶点

坐标,最终定位得到轨枕的位置并完成对轨枕的分割,如图 3 所示。



图 3 灰度投影分割效果图

Fig. 3 Grayscale projection segmentation effect

## 2 基于 YOLOv3 改进算法的轨枕裂纹检测模型

YOLOv3 主要改善了主干网络结构,融合多尺度特征,使用独立的逻辑分类器,YOLOv3 网络主要特征为:

(1) Darknet-53 模块的预测效率良好。基于 Darknet-53 构建的主干网络由于设计了类残差快捷连接(shortcut connections),实现高速的浮点运算,

使得预测效率更高<sup>[10]</sup>;

(2) 利用逻辑回归输出的多尺度预测。每个网格单元在 3 个尺度上预测得到 3 个物体框(bounding box),每个物体框包含  $(x, y, w, h, confidence)$  5 个基本参数,通过独立的逻辑分类器得出标签为裂纹的概率,根据上采样倍数的不同,输出不同尺度的张量尺寸,张量尺寸计算公式为  $N \times N \times [3 \times (5 + 1)]$ 。

由于 YOLOv3 网络在复杂背景下对轨枕裂纹类小目标检测精度低,本文设计了基于 YOLOv3 改进算法,使得 YOLOv3 在保证快速的前提下,进一步提升在混凝土轨枕裂纹检测方面的检测精度。在算法中引入注意力机制中的挤压与激励模块(Squeeze and Excitation, SE)和空间金字塔池化模块(Spatial Pyramid Pooling, SPP),改进后的 YOLOv3 网络结构如图 4 所示。

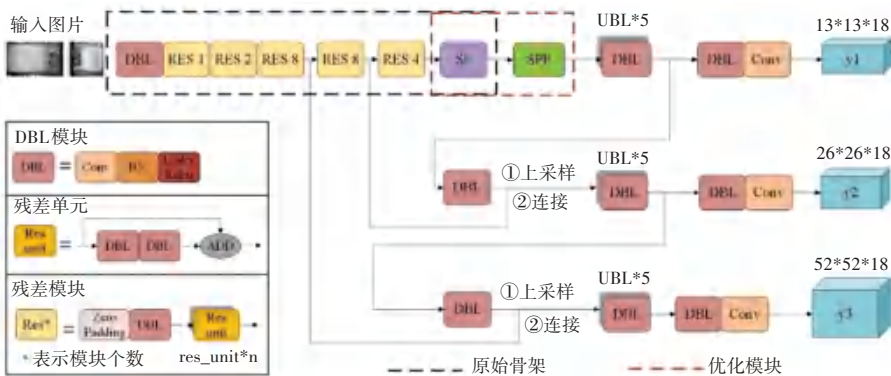


图 4 改进的 YOLOv3 结构图

Fig. 4 Improved YOLOv3 structure diagram

### 2.1 挤压与激励模块

由于轨枕裂纹特征存在微小细长、形态多样、与混凝土轨枕背景相似度高等的特点,传统算法存在提取轨枕裂纹特征精确度低等缺点,本算法在 YOLOv3 主干网络末端,引入注意力机制中的挤压与激励模块,改善对裂纹特征提取的精度,SE 模块的结构如图 5 所示。

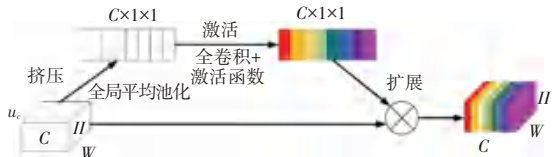


图 5 挤压与激励模块原理

Fig. 5 Principle of extrusion and excitation module

注意力机制的特点在于其经过特殊设计参数配置,有助于减少无效信息冗余,迅速提取关键特征,所以能够大幅度提高准确率<sup>[11]</sup>。引入的注意力机

制中的挤压与激励模块,能够快速、准确地提取到关键的轨枕裂纹特征。挤压与激励模块首先经过挤压操作(Squeeze),计算公式(3)如下:

$$u_c = V_c * X = \sum_{s=1}^{c'} v_c^s * x^s \quad (3)$$

其中,  $u_c$  表示输出轨枕图像中通道的特征;  $V_c$  表示轨枕图像中第  $c$  个学习的滤波器的参数;  $X$  表示输入通道特征;  $*$  表示卷积运算。

挤压与激励模块经过 1 个全局平均池化层,将轨枕裂纹图像转化为一个  $1 \times 1 \times 1$  024 的向量,该向量代表了轨枕图片中全局的特征通道分布情况,然后放入 2 个以全连接层和激活函数为单位的小型模块,针对轨枕裂纹的特征图进行激活(Excitation),最后统一对通道进行加权操作,输出轨枕裂纹特征图。此外,设有残差模块确保有效训练,使得全局特征聚焦于轨枕裂纹的信息特征,有效抑制轨枕图像

中其他冗余的特征干扰<sup>[12]</sup>。

## 2.2 空间金字塔池化

由于混凝土轨枕在采集的图像中的大小不尽相同,原有的池化层 (Pooling) 会丢失图片信息,导致裂纹检测效果不理想。如图 6 所示,在建立主干网络结构后,通过引入空间金字塔池化操作,能够最终输出固定的图片尺寸,不丢失主要特征,不受输入轨枕图片的尺寸影响,不破坏原始图像比例<sup>[13]</sup>。首先,将输入的轨枕图像经过一次卷积运算;其次,特征图经过  $5 \times 5, 9 \times 9, 13 \times 13$  这 3 个尺度的池化窗口 (pooling window);最后,输出一个同样长度的池化特征 (scale-invariance)。

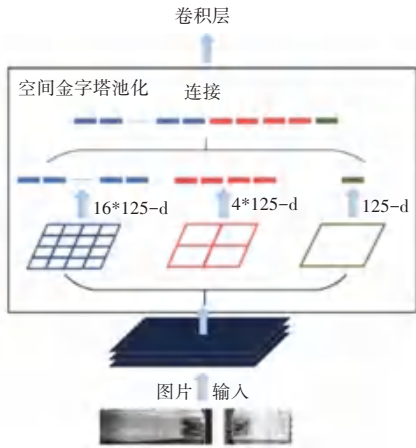


图 6 空间金字塔池化模块

Fig. 6 Pooling module of spatial pyramid

## 2.3 CIoU 损失函数

轨枕裂纹预测损失函数由回归预测误差  $L_{regression}$ , 置信度误差  $L_{confidence}$  以及分类误差  $L_{classification}$  3 部分组成<sup>[14]</sup>。其中,置信度误差由边框内存在轨枕裂纹和不存在轨枕裂纹两种情况叠加计算得到;轨枕裂纹分类误差是基于网格内存在裂纹的前提计算的,若边框内存在轨枕裂纹则分类误差值为 1,总体公式(4)如下:

$$Loss = L_{regression} + L_{confidence} + L_{classification} \quad (4)$$

在轨枕裂纹回归预测的损失函数部分, YOLOv3 使用二分交叉熵损失函数 (binary cross-entropy loss), 但该函数在误差较大的情况下,会因为小幅度地调整参数,使得训练速度更缓慢。此外,二分交叉熵损失函数较为复杂,运算量较大。

出于对模型收敛速度和精度的考虑,将 CIoU 损失函数 (Complete Intersection over Union Loss, CIoU Loss) 作为新的轨枕裂纹回归预测的损失函数,该函数的示意图如图 7 所示。

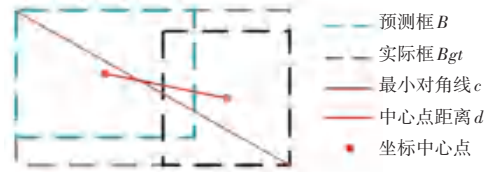


图 7 CIoU 损失函数示意图

Fig. 7 Schematic diagram of CIoU loss function

CIoU 部分的损失函数公式(5)为:

$$L_{regression} = IoU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \quad (5)$$

其中,  $IoU$  表示预测框和实际框的重叠度;  $\rho^2(\cdot)$  为预测框与实际框的欧式距离;  $b$  和  $b^{gt}$  分别表示预测框与实际框的中心点位置;  $c$  为预测框与实际框组合部分最小外矩形框的对角线;  $\alpha v$  是一个带有平衡参数  $\alpha$  且用于衡量宽高比相似性  $v$  的影响因子。

## 3 实验结果及分析

### 3.1 数据集与开发环境

本文的数据集的构建是基于有砟轨道上拍摄和采集的轨枕图片,轨枕状态分为有裂纹和无裂纹两种,本实验获取了共计 1 000 张裂纹数据集,根据轨枕状态与裂纹条数将其分类,其中含有轨枕裂纹的图像共 210 张,不含轨枕裂纹的图像 790 张,以 7:3 的训练集与测试集的比例开展实验。本实验是在 Win10 操作系统,双 CPU Intel Xeon Silver 4214 2.2 GHz, PyTorch 框架以及 PyCharm 的集成开发环境下进行。

### 3.2 轨枕裂纹检测实验

针对灰度投影法分割后的轨枕图像,本实验对训练参数初始值的设置值见表 1。

表 1 初始参数值的设定

Tab. 1 Setting of initial parameter values

参数	数值
Input size	128×608
Initial learning rate	0.1
Class	1
Batch size	6
Epochs	200

通过灰度定位法针对道床定位并分割轨枕,用注意力机制中的挤压与激活模块提取重要特征,引入空间金字塔池化,改进损失函数,最终利用改进的 YOLOv3 网络完成对轨枕裂纹的检测。评价指标主要从精确度与速度两个角度分析,用到的指标有精

精度(Precision, P), 召回率(Recall, R), 平均准确率(Mean Average Precision, MAP), 每秒传输帧数(Frames Per Second, FPS)。实验结果的精确度, 召回率和平均准确率的公式(6)~(8)如下:

$$P = \frac{T_p}{T_p + F_p} \quad (6)$$

$$R = \frac{T_p}{T_p + F_N} \quad (7)$$

$$M_{AP} = \frac{1}{N} \int_0^1 p(r) dr \quad (8)$$

在精确度与召回率的公式中,  $T_p$  (True Positive)表示人工标定为轨枕裂纹, 实际被判定存在轨枕裂纹的样本数量;  $T_N$  (False Positive)表示人工标定背景区域, 实际被判定存在轨枕裂纹的样本数量;  $F_N$  (False Negative)为人工标定背景区域, 实际被判定不存在轨枕裂纹的样本数量。在平均准确率的公式中,  $N$ 表示裂纹分类数量,  $p(r)$ 表示精确度与召回率的关系函数。

基于本算法实现的轨枕裂纹检测效果如图8所示。针对输入的整体道床图像与灰度投影法分割后的轨枕图像, 分别采用YOLOv3与本算法进行对比实验, 结果见表2。经过轨枕区域提取后裂纹检测在YOLOv3网络上精确度提升了35.0%, 召回率提升11.4%, MAP提升22.4%, 证明了轨枕区域提取对轨枕裂纹检测的必要性; 将灰度投影法提取的轨枕图像, 输入改进的YOLOv3裂纹检测网络中, 裂纹精确度较YOLOv3网络提升了16.9%, 召回率提升了3.5%, MAP提升了5.9%, 证明了本算法在轨枕裂纹检测精度具有显著的优越性。



图8 轨枕裂纹检测效果图

Fig. 8 Effects of sleeper crack detection

表2 实验量化结果对比

Tab. 2 Comparison of experimental results

算法模型	整体有砟道床图像输入				灰度投影法分割后的轨枕图像输入			
	Precision	Recall	MAP	FPS	Precision	Recall	MAP	FPS
YOLOv3	0.444	0.736	0.632	80.5	0.794	0.877	0.856	81.7
Ours	<b>0.469</b>	<b>0.792</b>	<b>0.659</b>	76.1	<b>0.963</b>	<b>0.912</b>	<b>0.915</b>	76.6

## 4 结束语

(1) 本文提出了一种基于YOLOv3改进算法的轨枕裂纹检测方法。首先, 通过灰度投影法对道床定位并针对轨枕进行了分割; 其次, 通过改进的YOLOv3算法对分割得到的轨枕图像进行轨枕裂纹的精确检测。

(2) 针对轨枕裂纹检测任务, 改进了YOLOv3算法。通过加入注意力机制中的挤压与激活模块, 提取轨枕裂纹的重要特征, 使得特征提取层部分能够对轨枕裂纹进行精确的定位; 引入空间金字塔池化, 实现输出相同尺寸的特征图, 改进损失函数, 最终实现对YOLOv3的改进。

(3) 经实验验证, 结果表明对轨枕裂纹的检测精确度达到96.3%, 召回率达到91.2%, MAP达到91.5%, FPS达到76.6张/s, 检测精确度、召回率和MAP在有砟轨道的轨枕裂纹数据集的表现提升显著。

## 参考文献

- [1] 发展和改革委员会. 中国国家铁路集团有限公司2020年统计公报[N]. 人民铁道, 2021-03-01(002).
- [2] 井国庆, 杜文博, 尤瑞林, 等. 基于DIC混凝土轨枕和复合轨枕受弯特征分析[J]. 铁道科学与工程学报, 2021, 18(1): 64-70.
- [3] 章欣, 王艳, 胡恒山, 等. 声发射技术在铁路系统检测中的研究和应用[J]. 应用声学, 2017, 36(3): 189-199.
- [4] 冉蓉, 徐兴华, 邱少华, 等. 基于深度卷积神经网络的裂纹检测方法综述[EB/OL]. 计算机工程与应用: 1-16[2021-04-09]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20210304.1549.002.html>.
- [5] CHA Y, CHOI W, BUYUKOZTURK O. Deep Learning Based Crack Damage Detection Using Convolutional Neural Networks[J]. Computer-aided Civil and Infrastructure Engineering, 2017, 32(5): 316-378.
- [6] MANDAL V, UONG L, ADU-GYSMFI Y. Automated road crack detection using deep convolutional neural networks[C]// 2018 IEEE International Conference on Big Data (Big Data), December 10-13, 2018, Seattle, USA. New York: IEEE, 2018: 5212-5215.
- [7] LI Wenju, SHEN Zihao, LI Peigang. Crack Detection of Track Plate Based on YOLO[J]. International Symposium on Computational Intelligence and Design, 2019: 15-18.