

# 基于图像增强与改进 Cascade R-CNN 的钢轨表面缺陷检测

罗晖, 李健\*, 贾晨

华东交通大学信息工程学院, 江西 南昌 330013

**摘要** 针对钢轨表面缺陷检测中, 钢轨表面图像存在背景不均匀、缺陷尺度变化大且样本数据不足的问题, 提出一种基于图像增强与改进 Cascade R-CNN 的钢轨表面缺陷检测方法。首先, 采用改进 Retinex 算法处理钢轨表面图像, 增强缺陷与背景的对比度。然后, 采用改进 Cascade R-CNN 对钢轨表面缺陷进行检测, 并应用交并比 (IoU) 平衡采样、感兴趣区域对齐和完全交并比 (CIoU) 损失分别解决训练样本 IoU 分布与困难样本 IoU 分布不平衡、感兴趣区域池中取整量化导致的感兴趣区域与提取的特征图不匹配和回归损失 Smooth L1 对于预测边框回归不准确的问题。最后, 采用翻转变换、随机剪裁、亮度变换和生成对抗网络等方法增广钢轨表面缺陷图像数据集, 消除样本数据不足导致的网络训练过拟合现象。实验结果表明, 该方法以 ResNet-50 作为特征提取器, 平均精度可达 98.75%, 相对于未改进的 Cascade R-CNN 提高了 2.52%, 且检测时间缩短了 24.2 ms。

**关键词** 图像处理; 钢轨表面缺陷; 图像增强; Cascade R-CNN; 感兴趣区域对齐; 完全交并比; 数据增广

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.2212001

## Rail Surface Defect Detection Based on Image Enhancement and Improved Cascade R-CNN

Luo Hui, Li Jian\*, Jia Chen

School of Information Engineering, East China JiaoTong University, Nanchang, Jiangxi 330013, China

**Abstract** In the rail surface defect detection, the rail surface image has the problem of uneven background, large variation of defect scale, and insufficient sample data. Therefore, this paper proposes a rail surface defect detection method based on image enhancement and improved Cascade R-CNN. First, the improved Retinex algorithm is used to process the rail surface image to enhance the contrast between the defects and the background. Then, an improved Cascade R-CNN is adopted to detect rail surface defects, and the intersection over union (IoU) balanced sampling, region of interest align and complete intersection over union (CIoU) loss are applied to solve the imbalance between training sample IoU distribution and the difficult sample IoU distribution, the misalignment between region of interest and extracted feature map caused by rounding quantization in region of interest pooling, and the inaccuracy of the regression loss Smooth L1 for the regression of predicted bounding box. Finally, the dataset of rail surface defect images is expanded using methods such as flipping transformation, random cropping, brightness transformation, and generative adversarial networks, so as to solve the phenomenon of over-fitting of network training caused by insufficient sample data. Experimental results show that the average accuracy of the proposed method, using ResNet-50 as the feature extractor, can reach 98.75%, which is 2.52% higher than the unimproved Cascade R-CNN, and the detection time is reduced by 24.2 ms.

**Key words** image processing; rail surface defect; image enhancement; Cascade R-CNN; region of interest align; complete intersection over union; data augmentation

收稿日期: 2020-12-07; 修回日期: 2021-01-04; 录用日期: 2021-01-21

基金项目: 江西省重点研发计划项目(20202BBEL53001)

通信作者: \*lj\_hjdj@163.com

OCIS codes 100.2000; 100.2980; 100.3008; 100.4996

## 1 引言

在列车运行过程中,由于列车对钢轨产生挤压、冲击和摩擦等作用而引起的钢轨表面缺陷严重威胁着铁路的安全运营,因此,对钢轨表面缺陷进行准确、快速检测已成为我国铁路发展亟待解决的关键问题。

在很长一段时间中,钢轨表面的缺陷都是由经验丰富的检查员来检测,他们的高强度工作只能得到低效率的主观结果。与人工检测相比,先进的无损检测更加高效、准确。目前,主要的检测技术有涡流检测<sup>[1-2]</sup>、超声检测<sup>[3-4]</sup>、声发射检测<sup>[5-6]</sup>和激光检测<sup>[7-8]</sup>,然而,这些方法虽然可以达到检测的目的,但是由于各种外部干扰,其信号难以处理,可能会导致检测存在盲点,并且这些方法的实时性较差。

近年来,机器视觉技术凭借其检测精度高、速度快和无需接触等优势,被广泛应用于钢轨表面缺陷检测。

在传统的机器视觉检测方法中,Li 等<sup>[9]</sup>采用局部归一化(LN)算法增强了钢轨图像的对比度,并利用基于投影轮廓的缺陷定位算法(DLBP)对缺陷进行检测。Li 等<sup>[10]</sup>采用 Michelson-like 对比度(MLC)算法来增强钢轨图像,显著地增强了缺陷和背景的对比度,同时提出了一种基于比例最大熵的自动阈值算法(PEME)对钢轨表面缺陷进行分割。Zhang 等<sup>[11]</sup>设计了一种曲率滤波器以消除噪声,然后建立了一种基于马尔可夫随机场的快速鲁棒高斯混合模型,用于钢轨表面缺陷的准确检测。Gan 等<sup>[12]</sup>提出了一种由粗到细的缺陷提取器(CFE),基于均值漂移算法的粗提取器定位包括真实缺陷和其他噪声点在内的离群异常值,然后利用结合纵向背景信息和横向先验信息的细提取器来滤除噪声点,提高了检测精度,但该方法计算复杂度较高。曹义亲等<sup>[13]</sup>对钢轨表面图像进行背景差分处理,并利用最大熵阈值分割和连通区域标记法进行缺陷检测。这些方法在一定程度上推动了钢轨表面缺陷检测技术的发展,但这些方法不能较好地提取缺陷特征,检测精度较低。

在基于深度学习的机器视觉检测方法中,Song 等<sup>[14]</sup>提出了一种基于 YOLOv3 算法的钢轨表面缺陷检测算法,较好地完成了钢轨表面缺陷检测。苏焯等<sup>[15]</sup>利用 Radon 变换实现了钢轨表面区域的提

取,利用 Faster R-CNN 算法对钢轨表面缺陷进行检测,但未针对钢轨表面缺陷特征进行改进。Yuan 等<sup>[16]</sup>设计了一种基于改进 YOLO 的钢轨表面缺陷检测算法 MOLO,结合多尺度结构,对三类钢轨表面缺陷进行准确、快速检测。梁波等<sup>[17]</sup>利用平移、旋转、缩放变换等方法增广钢轨表面数据集,并在 U-Net 网络的收缩路径和扩展路径中增加一个分支结构,用于辅助生成预测边框,提高了检测精度。Jin 等<sup>[18]</sup>针对复杂多样的钢轨表面,设计了一种深度多模型钢轨检测系统 DMRIS,该系统由概率图模型 FRGMM 生成分割方案,并采用 Faster-R-CNN 实现目标定位,联合命中目标为真实缺陷,进一步提高了钢轨表面缺陷的检测精度。

基于深度学习的钢轨表面缺陷检测方法具有较好的检测性能,但对于钢轨表面缺陷检测还存在诸多挑战:1)背景不均匀,钢轨表面图像中存在光照不均的问题,并且由于钢轨轨头呈弧形以及钢轨表面外侧铁锈的存在,其表面各区域反射特性不同,这使得钢轨表面图像具有不均匀的背景。2)钢轨表面缺陷尺度变化大,钢轨表面缺陷具有多尺度特征,且形状各异。3)数据集不足,钢轨表面缺陷图像数据较少,容易在训练过程中导致过拟合问题。因此,在分析钢轨图像特征的基础上,本文提出了一种基于图像增强和改进 Cascade 区域卷积神经网络(R-CNN)<sup>[19]</sup>的钢轨表面缺陷检测方法,该方法通过改进 Retinex 算法处理钢轨表面图像,增强了缺陷与背景的对比度,并应用交并比(IoU)平衡采样<sup>[20]</sup>、感兴趣区域对齐(RoIAlign)<sup>[21]</sup>和完全交并比(CIoU)损失<sup>[22]</sup>对 Cascade R-CNN 进行改进,同时,采用翻转变换、随机剪裁、亮度变换、生成对抗网络等方法增广钢轨表面缺陷图像数据集,以实现钢轨表面缺陷的准确、快速检测。

## 2 方法整体结构

本文所提方法的整体结构如图 1 所示,主要包括数据集建立、图像增强和钢轨表面缺陷检测三个阶段。首先,本文采用翻转变换、随机剪裁、亮度变换和生成对抗网络等方法增广钢轨表面缺陷图像数据集,减少因样本数据不足导致的网络训练过程中的过拟合现象的发生。其次,提出一种改进 Retinex 算法增强钢轨表面图像,该方法在多尺度 Retinex (MSR)算法的基础上,进一步对图像进行局部非线

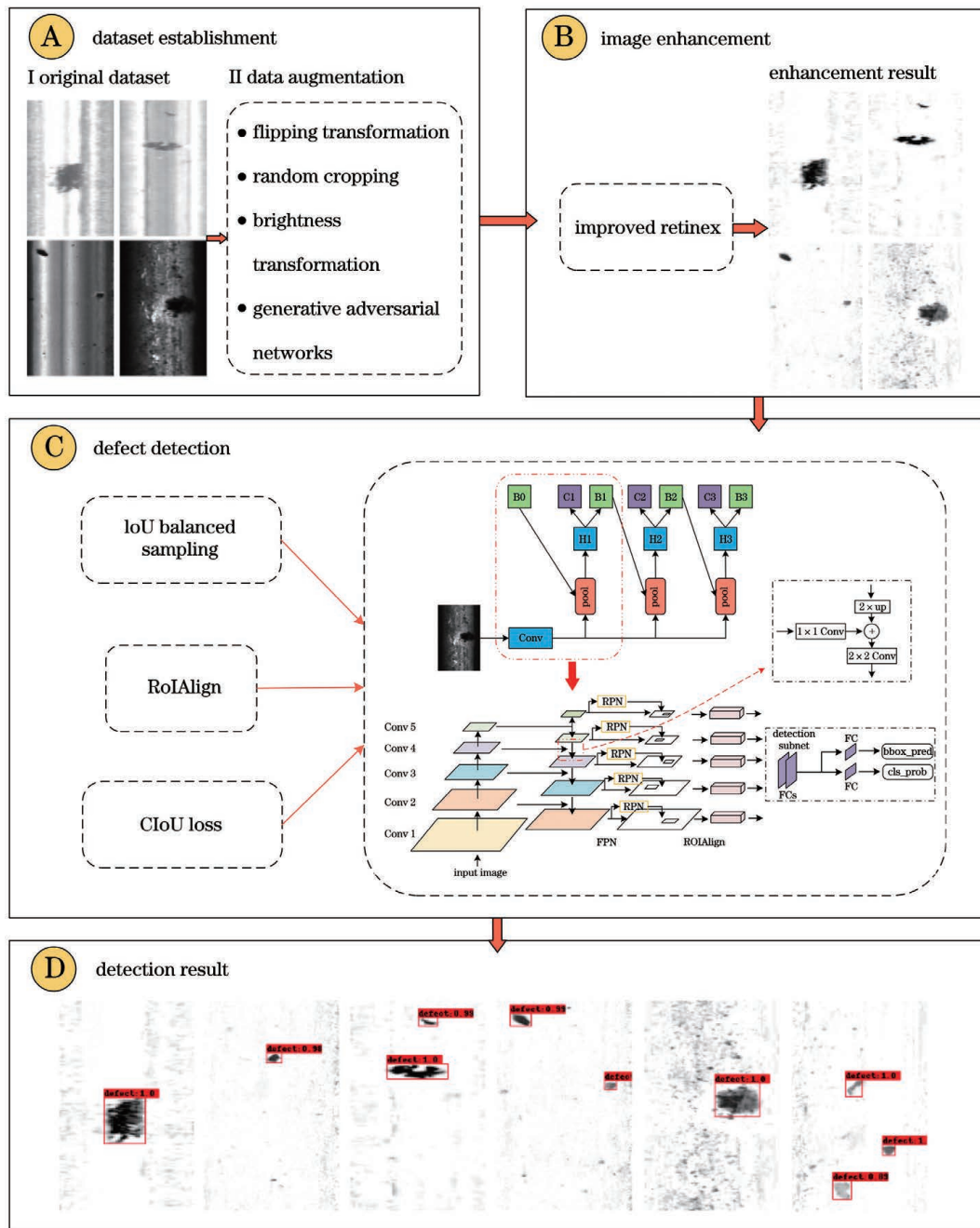


图 1 本文方法的整体结构

Fig. 1 Overall structure of the method in this paper

性对比度增强<sup>[10]</sup>,显著增强了缺陷和背景的对比度,并使钢轨表面图像具有均匀的背景。然后,采用改进 Cascade R-CNN 对应用改进 Retinex 算法处理过后的钢轨表面图像进行检测,针对训练样本 IoU 分布与困难样本 IoU 分布不平衡、感兴趣区域池化(RoI Pooling)中取整量化导致的感兴趣区域(RoI)与提取的特征图不匹配和回归损失 Smooth L1 对于预测边框回归不准确的问题,本文分别应用 IoU 平衡采样、RoIAlign 和 CIoU 损失对算法进行改进,以提高钢轨表面缺陷的检测精度。

### 3 钢轨表面图像增强

Retinex 理论认为,一幅图像  $I(x, y)$  可看作是由照射分量和反射分量组成,且可表示为二者的乘积,其表达式为

$$I(x, y) = L(x, y) \cdot R(x, y), \quad (1)$$

式中: $x$  表示像素点横坐标; $y$  表示像素点纵坐标;照射分量  $L(x, y)$  表示图像中像素所能达到的动态范围;反射分量  $R(x, y)$  表示图像中与光照无关的本质特征。



为了得到图像中的反射分量,采用基于 Retinex 理论的 MSR 算法<sup>[23-24]</sup>将图像中的照射分量去除,即,

$$\ln R(x, y) = \sum_{k=1}^n \omega_k \{ \ln I(x, y) - \ln [I(x, y) * G(x, y)] \}, \quad (2)$$

式中: $n$  为尺度总数; $\omega_k$  为第  $k$  个尺度所对应的权重,且满足  $\sum_{k=1}^n \omega_k = 1$ ;  $G(x, y)$  为高斯函数。其定义为

$$G(x, y) = K \exp \left[ -\frac{(x^2 + y^2)}{c^2} \right], \quad (3)$$

式中: $K$  为由归一化函数  $\iint G(x, y) dx dy = 1$  确定的常数; $c$  为高斯函数的标准差。对于单个尺度, $c$  越小,算法的动态压缩能力越强,图像中的细节可以得

到更好地增强,但容易产生颜色失真,反之,算法的颜色保真度越高,但动态压缩能力会减弱。

MSR 算法可以在一定程度上增强钢轨表面图像,但无法消除钢轨表面各区域反射特性不同造成的背景不均匀现象,因此,对采用 MSR 算法提取到的反射分量进行局部非线性对比度增强,以进一步增强钢轨表面缺陷和背景的对比如。

本文提出的改进 Retinex 钢轨表面图像增强算法框架如图 2 所示。

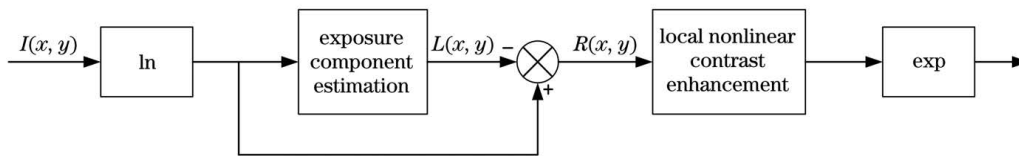


图 2 改进 Retinex 钢轨表面图像增强算法框架

Fig. 2 Improved Retinex rail surface image enhancement algorithm framework

假定存在包含像素点  $(x, y)$  在内的尺寸为  $w \times h$  的小窗口  $W$ ,  $w$  代表  $W$  的宽度,  $h$  代表  $W$  的长度,局部非线性对比度增强算法表示为

$$C(x, y) = \begin{cases} \frac{R(x, y) - U_w}{R(x, y) + U_w}, & \text{if } R(x, y) < U_w \\ 0, & \text{if } R(x, y) \geq U_w \end{cases}, \quad (4)$$

式中: $U_w$  为  $W$  中所有像素的灰度均值。当  $R(x, y) < U_w$  时,  $C(x, y)$  在  $[-1, 0)$  区间内变化,结合缺陷区域灰度低于背景区域的特征,可得出该式将缺陷区域的灰度映射到  $[-1, 0)$ , 而将背景区域的灰度映射为 0。

钢轨表面在列车运行方向上的反射特性除缺陷区域外基本是均匀的,而在垂直于列车运行方向上的反射特性可能会有很大的变化,且光照不均主要表现在垂直于列车运行方向。因此,选择尺寸为  $1 \times h$  的线形  $W$ , 其光照和反射特性的变化很小,  $h$  为列车运行方向上的钢轨表面图像长度。

综上所述,局部非线性对比度增强算法归纳如下:

- 1) 将图像分成互不重叠的区域,每一块区域的大小即为  $W$ , 并计算每块区域的  $U_w$ ;
- 2) 扫描  $W$  中的每一个像素点,按照(4)式计算每个像素点的  $C(x, y)$ ;

3) 重新将  $C(x, y)$  映射到  $0 \sim 255$  灰度区间。

## 4 基于 Cascade R-CNN 的钢轨表面缺陷检测

Cascade R-CNN 扩展了两阶段检测架构 Faster R-CNN<sup>[25]</sup>, 首先,应用区域建议网络(RPN)生成候选区域。然后,级联不同 IoU 阈值的检测器进行训练和回归,每个检测器的输入都是上一个检测器进行边框回归之后的具有良好分布的样本。当以这种方式运行时,可以保证在样本数不减少的情况下训练出高质量的检测器,从而提高钢轨表面缺陷的检测精度。

### 4.1 网络结构

本文采用 ResNet-50<sup>[26]</sup> 和特征金字塔网络<sup>[27]</sup> (FPN)相结合的结构进行特征提取。ResNet-50 通过残差学习可以有效解决深层网络训练出现的退化问题。FPN 结合网络的低层特征图和高层特征图,生成多尺度高语义特征图,可以显著提高小目标的检测精度。在 ResNet-50 和 FPN 的基础上,结合级联检测器,本文提出的钢轨表面缺陷检测算法整体结构如图 3 所示。

### 4.2 FPN

在进行特征提取的过程中,低层特征图具有较

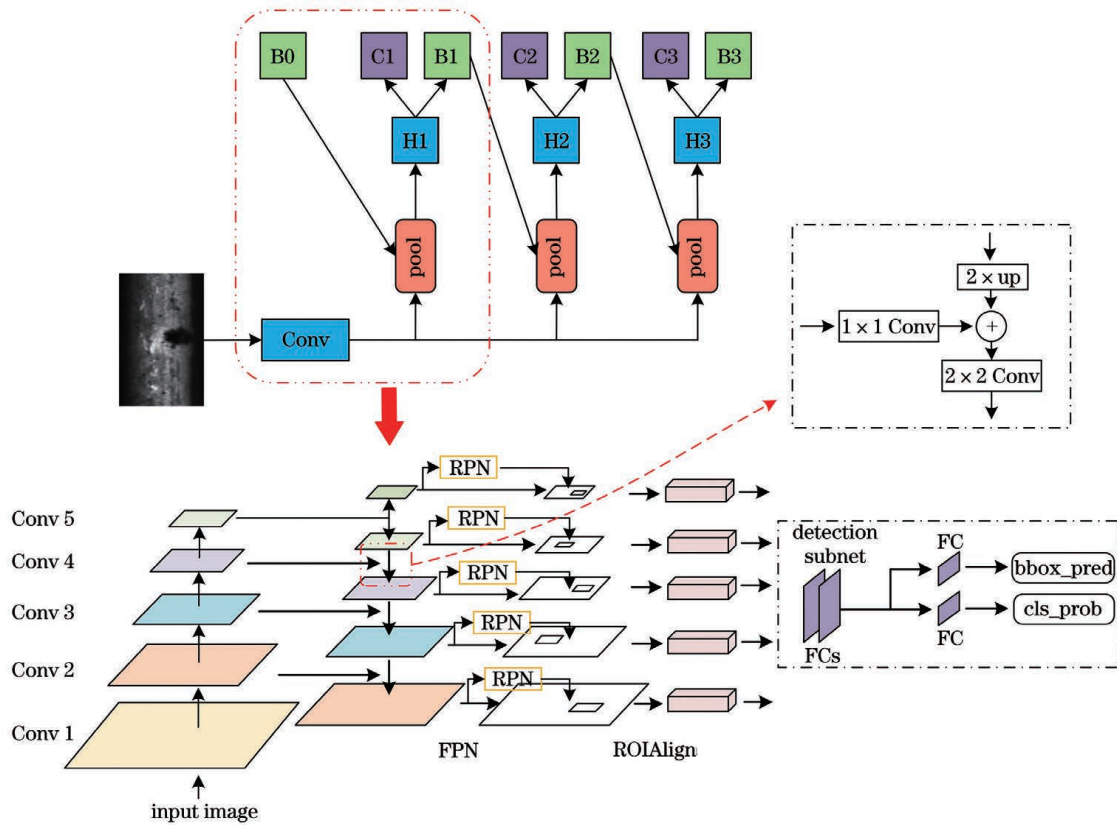


图 3 钢轨表面缺陷检测算法整体结构图

Fig. 3 Overall structure diagram of rail surface defect detection algorithm

准确的位置信息但是语义信息缺乏,而高层特征图具有丰富的语义信息。FPN 将低层特征图和高层特征图进行融合,可生成不同尺度高语义特征图。

FPN 采用自底向上、自顶向下和横向连接相结合的结构。自底向上即网络的特征提取过程,分别采用 Conv2、Conv3、Conv4、Conv5 中的最后一层进行特征金字塔构建。自顶向下通过 2 倍上采样进行,横向连接将自底向上生成的特征图和上采样生成的相同大小的特征图进行融合。融合生成的特征图分别记为 P2、P3、P4、P5,并对 P5 进行两倍下采样得到 P6。

将 FPN 与 RPN 结合用于生成候选区域,原来的 RPN 是在特征提取网络的最后一个卷积层上生成锚点(anchor),现在对于 FPN 生成的 P2、P3、P4、P5、P6,设置 anchor 的大小分别为  $32^2$ 、 $64^2$ 、 $128^2$ 、 $256^2$ 、 $512^2$ ,且每个特征图的 anchor 都应用 1:2、1:1 和 2:1 三种长宽比,因此,整个特征金字塔生成 15 种 anchor。

钢轨表面缺陷尺度变化大,且具有较多小目标缺陷,应用 FPN 结构,在不同尺度特征图生成候选区域进行检测,可以提高钢轨表面缺陷的检测精度。

### 4.3 级联检测器

Cascade R-CNN 最核心的部分是级联检测器,所有级联的检测器具有相同的架构,但 IoU 阈值 ( $u_t$ )不相同。在每一个阶段,每一个  $u_t$  都对应一个分类器( $h_t$ )和一个回归器( $f_t$ )。其中,级联的回归器表示为

$$f(x, b) = f_T \circ f_{T-1} \circ \dots \circ f_1(x, b), \quad (5)$$

式中:  $T$  表示级联阶段的总数;  $\circ$  表示回归器级联; 每个回归器  $f_t$  对应于该阶段的样本分布 ( $b^t$ ),而不是初始阶段样本分布 ( $b^1$ ),即用一个阶段的输出来训练下一个阶段的回归器。由于候选区域在经过回归器后的 IoU 基本上都是增加的,那么再经过一个更大阈值训练的回归器,可以保证在不减少训练样本的情况下,得到更高质量的候选区域。

本文采用 3 阶段级联,其对应的  $u_t$  分别设置为 0.5、0.6、0.7,检测器的损失函数设计为

$$L(x^t, g) = L_{cls} [h_t(x^t), y^t] + \lambda (y^t \geq 1) \times L_{loc} [f_t(x^t, b^t), g], \quad (6)$$

式中:分类损失  $L_{cls}$  应用交叉熵损失函数;回归损失  $L_{loc}$  应用 Smooth L1 损失函数;  $t$  表示级联的阶段;  $g$  是  $x^t$  对应的真实边框;  $b^t = f_{t-1}(x^{t-1}, b^{t-1})$ ;  $\lambda$  是

权重系数; ( $y' \geq 1$ ) 表示负样本不参与回归损失计算;  $y'$  是给定  $u_i$  时  $x'$  对应的类别标签。

通过训练级联检测器,可以得到一系列质量不断提高的检测器,并对更高质量的候选区域进行检测,从而可以有效提高钢轨表面缺陷的检测精度。

#### 4.4 对算法的改进

##### 4.4.1 候选区域采样方式的改进

在 Cascade R-CNN 中,对于候选区域的采样采用随机采样的方式进行,假设采样正负样本总数为  $A$ ,随机采样的过程如下:

1) 在所有正样本中以随机采样的方式采样  $\mu A$  个正样本,  $\mu$  为设定的超参数,如果正样本总数少于  $\mu A$ ,则取所有的正样本;

2) 在所有负样本中以随机采样的方式采样  $A - I^*$  个负样本,  $I^*$  为采样的正样本个数。

以这种方式得到的训练样本会遗漏很多最终将预测错误的样本,这样的样本称之为困难样本,有实验表明<sup>[20]</sup>,随机采样得到的训练样本只有 30% 的 IoU 大于 0.05,而超过 60% 的困难样本的 IoU 大于 0.05,因而会产生训练样本 IoU 的分布与困难样本 IoU 分布不平衡的问题。因此,本文采用 IoU 平衡采样代替随机采样,以提高训练样本中困难样本所占的比例,进而提高检测精度。

IoU 平衡采样对于正样本和负样本采用不同的方案,且主要关注于负样本。假设要从候选区域中采样  $N$  个负样本, IoU 平衡采样表示为

$$p_k = \frac{N}{K} \cdot \frac{1}{M_k}, k \in [0, K), \quad (7)$$

式中:  $K$  表示根据 IoU 将采样间隔划分为  $K$  个区间;  $M_k$  表示第  $k$  个采样区间的候选区域数量;  $p_k$  表示在第  $k$  个采样区间采样的概率。通过这种方式,将在  $K$  个采样区间上均匀的选取样本,进而提高了高 IoU 样本所占的比例,也即提高了困难样本所占的比例。

在对正样本的采样方案中,由于正样本所对应的候选区域较少,因此,采用对于每一个真实边框采样相等数量的正样本的方式进行。

##### 4.4.2 RoI 池化方式的改进

在利用 RPN 生成候选区域后, Cascade R-CNN 采用 RoI Pooling 的方式提取每个 RoI 对应的固定大小的特征图,在这个过程中存在两次取整量化的操作:

1) 将浮点数形式的 RoI 坐标映射到特征图上,并进行取整量化;

2) 在划分池化区域为  $k \times k$  个单元时,对单元边界区域进行取整量化。

这些量化操作会导致 RoI 与提取的特征图存在不匹配的现象,尤其对于钢轨表面缺陷检测中的小目标缺陷更为明显。因此,本文提出采用 RoIAlign 替代 RoI Pooling 的优化方案,以提高钢轨表面缺陷的检测精度。

RoIAlign 避免了取整量化的操作,图 4 对 RoIAlign 进行了说明,其具体执行过程如下:

1) 在将 RoI 坐标映射到特征图上时保留浮点数边界,直接采用  $x/16$  的方式进行,16 为特征提取过程中特征图缩小的比例;

2) 将池化区域平均划分为  $k \times k$  个单元,每个单元的边界不执行取值量化;

3) 将每个单元划分为四个部分,取每个部分的中心点作为采样点,并利用双线性插值计算采样点的值,最后取四个采样点的最大值作为该单元的最终池化值。

通过这样的方式, RoIAlign 可以使 RoI 与提取的特征图具有更高的匹配度。

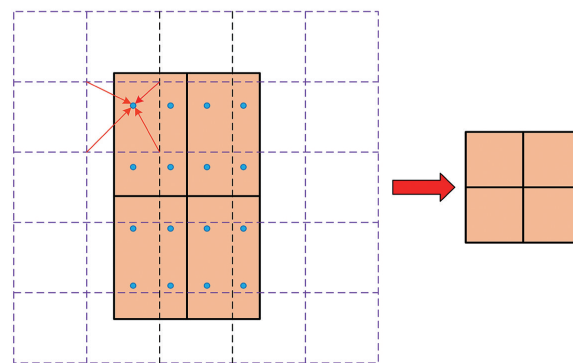


图 4 RoIAlign 图解

Fig. 4 RoIAlign diagram

##### 4.4.3 损失函数的改进

在采用 Smooth L1 损失作为回归损失函数时,只考虑了预测边框与真实边框之间的坐标距离,没有较好地考虑到两框之间的重叠度,且将用于表示边框位置的四个变量作为独立变量进行优化,不利于预测边框的精准回归,另外,广义交并比(GIoU)损失利用预测边框和真实边框的 IoU 以及最小外接矩形面积进行优化,可以较好地解决以上问题,但还存在以下不足:

1) 当预测边框与真实边框具有包含关系时,无法区分其相对位置关系;

2) 收敛速度较慢,需要进行足够多的迭代次数。



因此,本文采用 CIoU 损失作为回归损失函数, CIoU 损失同时考虑到预测边框与真实边框之间的重叠度、中心点距离以及边框的纵横比,相比于 Smooth L1 损失和 GIoU 损失具有更好的性能。对于给定预测边框  $B(x, y, w, h)$  和真实边框  $B^{gt}(x^{gt}, y^{gt}, w^{gt}, h^{gt})$ , CIoU 损失 ( $L_{CIoU}$ ) 定义为

$$L_{CIoU} = 1 - I_{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha\nu, \quad (8)$$

式中:  $I_{IoU}$  为  $B$  和  $B^{gt}$  的 IoU;  $b, b^{gt}$  分别表示  $B, B^{gt}$  的中心点;  $\rho(\cdot)$  表示欧氏距离;  $c$  是  $B$  和  $B^{gt}$  的最小外接矩形的对角线长度;  $\alpha$  为权重参数;  $\nu$  是用来衡量长宽比一致性的参数。相关计算公式为

$$I_{IoU} = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|}, \quad (9)$$

$$\alpha = \frac{\nu}{(1 - I_{IoU}) + \nu}, \quad (10)$$

$$\nu = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right). \quad (11)$$

## 5 实验结果与分析

### 5.1 数据集与评价指标

本文以北京交通大学李清勇教授团队公开的 RSDDs 钢轨表面缺陷图像数据集来验证所提出方法的性能,该数据集包括 195 张具有挑战性的钢轨表面缺陷图像,其中 67 张尺寸为  $160 \text{ pixel} \times 1260 \text{ pixel}$ , 128 张尺寸为  $55 \text{ pixel} \times 1250 \text{ pixel}$ 。在实验过程中,首先,经过分割以及调整图像大小处理得到了 347 张尺寸为  $160 \text{ pixel} \times 250 \text{ pixel}$ , 且至少包含一个缺陷的钢轨表面缺陷图像,建立初始钢轨表面缺陷图像数据集。然后,在所建立的初始数据集基础上,采用翻转变换、随机剪裁、亮度变换、生成对抗网络等方法增广数据集,最终生成 3000 张钢轨表面缺陷图像,并随机选取 70% 作为训练集、10% 作为验证集、20% 作为测试集。本文所采用初始钢轨表面缺陷图像以及利用数据增广方法生成的图像示例如图 5 所示。

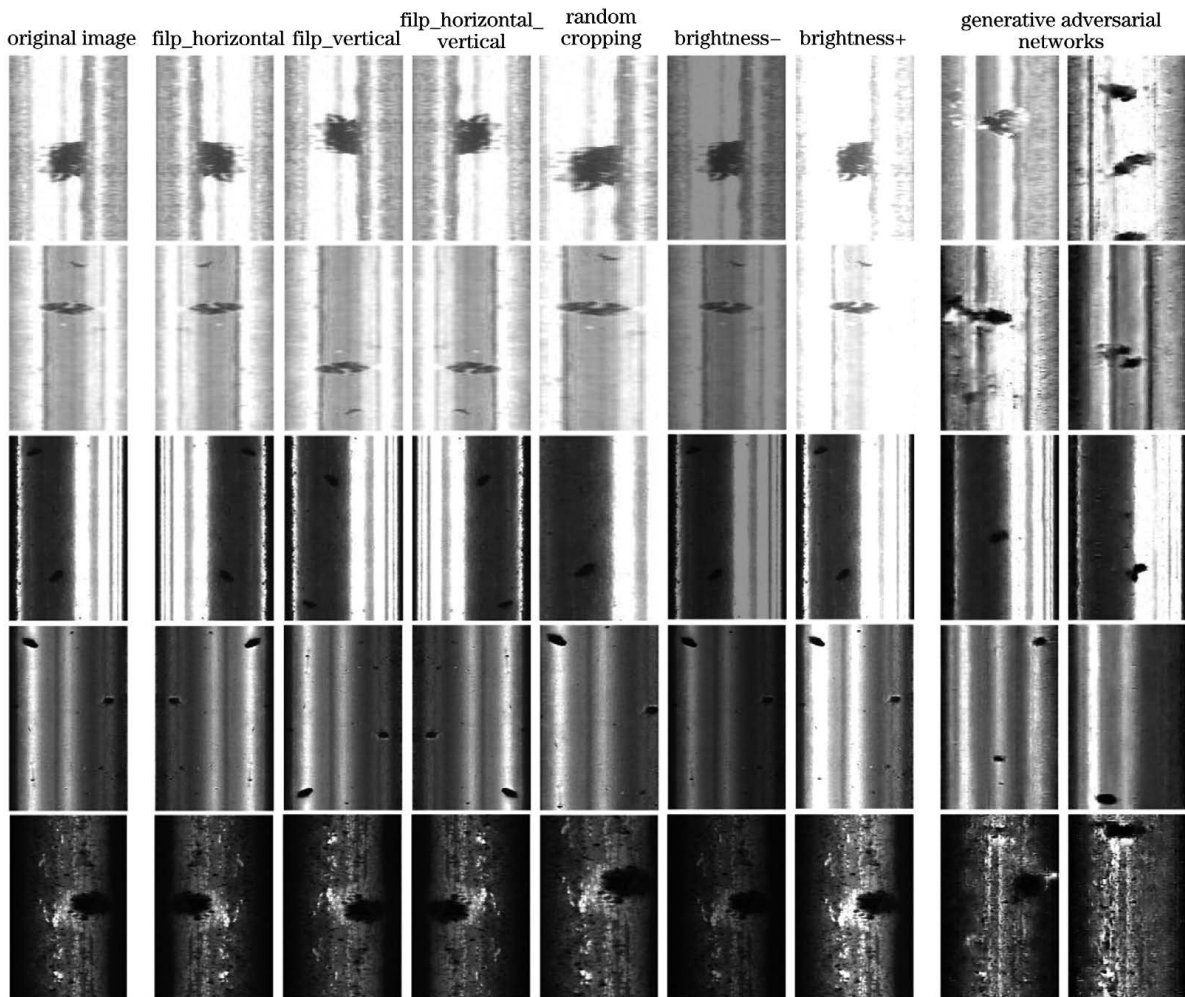


图 5 钢轨表面缺陷图像数据增广示例

Fig. 5 Examples of rail surface defect image data augmentation

从图 5 中可以看出,钢轨表面缺陷形状各异且钢轨表面缺陷图像具有不均匀的背景,这给钢轨表面缺陷检测带来极大的挑战,同时可以看到,本文利用翻转变换、随机剪裁、亮度变换、生成对抗网络等方法增广数据集,有效地增加了钢轨表面缺陷图像的数量和多样性。

本文采用的评价指标为目标检测领域通用的平均精度(AP),AP 是精确率( $P_{Precision}$ )、召回率( $P_{Recall}$ )曲线所包围的面积, $P_{Precision}$ 、 $P_{Recall}$  定义为

$$P_{Precision} = T_P / (T_P + F_P), \quad (12)$$

$$P_{Recall} = T_P / G_T, \quad (13)$$

式中: $T_P$  表示检测正确的缺陷数量; $F_P$  表示检测错误的缺陷数量; $G_T$  为标注的缺陷总数。

## 5.2 实验结果

### 5.2.1 实验平台及训练过程

本文基于 Python 和 TensorFlow 框架,在 Inter i7-7800X 3.5 GHz 处理器、32 GB 内存、NVIDIV GeForce RTX 2080 显卡的实验环境下进行实验。

在 Cascade R-CNN 的训练过程中,采用动量法对网络进行训练,动量设置为 0.9。输入图像经线性插值调整大小后,实际输入网络的图像大小为 400 pixel×625 pixel。设置 batch 大小为 1、初始学习率为 0.001、迭代次数为 70000 次。然后,在迭代 30000 次和 50000 次时,学习率分别乘以 0.1。期间,采用早停法和 Dropout 策略防止过拟合。

基于本文提出的改进 Cascade R-CNN,以 ResNet-50 作为特征提取器,网络训练损失值的变化曲线如图 6 所示,训练经过 39000 次迭代后,损失值低于 0.1,训练损失大约在 50000 次迭代后趋于稳定,由于采用早停法策略,网络在 66000 次迭代时停止训练。由此说明,本文网络训练各项超参数设

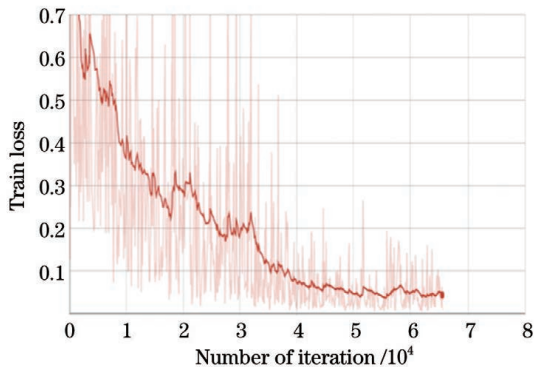


图 6 基于改进 Cascade R-CNN+ResNet-50 的训练损失值变化

Fig. 6 Changes of training losses based on improved Cascade R-CNN+ResNet-50

置合理,训练效果理想。

### 5.2.2 数据增广有效性实验对比

本文首先采用未改进的 Cascade R-CNN 网络,使用 ResNet-50 作为特征提取器,针对数据增广的有效性进行实验,在使用初始钢轨表面缺陷图像数据集进行实验时,300 张作为训练集、47 张作为测试集,迭代次数为 15000 次,实验结果如表 1 所示。由于初始钢轨表面缺陷图像数量较少,网络训练容易出现过拟合的现象,故在使用初始钢轨表面缺陷图像数据集进行实验时,AP 值只有 91.21%。在进行数据增广之后,图像数量增加,图像的多样性得到较大提高,检测网络可以更好地学习缺陷特征,进而 AP 值达到 95.74%。

表 1 数据增广有效性实验结果

Table 1 Experimental results of data augmentation effectiveness

Experiment dataset	AP / %
Original dataset	91.21
Augmentation dataset	95.74

### 5.2.3 图像增强实验对比

在本文提出的改进 Retinex 算法中,对于 MSR 算法,本文取三个尺度进行加权平均,令  $w_1 = w_2 = w_3$ ,并综合考虑算法的动态压缩能力和颜色保真度,分别将其对应的  $c_1$ 、 $c_2$ 、 $c_3$  的取值设置为 15、80、200,对于线性  $W$ ,  $h$  取值为 250。

为了验证本文所提出的改进 Retinex 算法的有效性,将其与直方图均衡化(HE)、对比度受限自适应直方图均衡化(CLAHE)<sup>[28]</sup>和 MSR 算法进行了实验对比,实验结果如图 7 所示。从图 7 中可以看出,HE 算法对于部分图像出现了过增强的现象。CLAHE 算法在一定程度上对钢轨表面图像的背景进行了平滑处理,但无法有效增强缺陷和背景的对比度。MSR 算法在一定程度上增强了缺陷和背景的对比度,但钢轨表面图像仍然具有不均匀的背景。而本文所提出的改进 Retinex 算法实现了较好的增强效果,在 MSR 算法的基础之上,进一步增强了缺陷和背景的对比度,并使钢轨表面图像具有均匀的背景。

本文采用未改进的 Cascade R-CNN 网络,使用 ResNet-50 作为特征提取器,对分别应用 4 种图像增强算法处理过后的钢轨表面图像进行检测,检测结果如表 2 所示。从表 2 中可以看出,HE、CLAHE 算法并没有提高钢轨表面缺陷的检测精度,说明



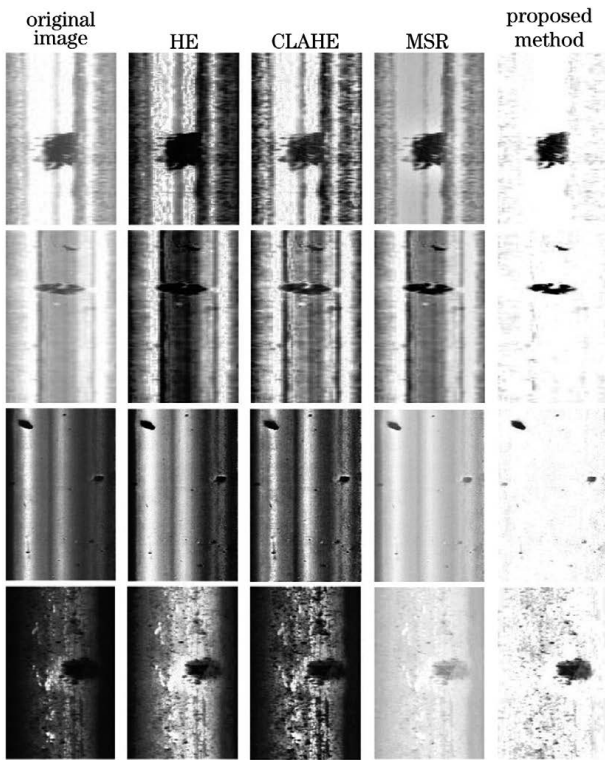


图 7 钢轨表面图像增强效果对比图

Fig. 7 Contrast diagrams of rail surface image enhancement effect

HE、CLAHE 算法不适用于钢轨表面缺陷检测的应

表 3 不同架构以及改进算法实验结果

Table 3 Experimental results of different architectures and improved algorithms

Backbone	Image enhancement	IoU balanced sampling	RoIAlign	CIoU	AP / %	Time / ms
ResNet-50	-	-	-	-	95.74	151.8
	✓	-	-	-	96.61	145.0
	✓	✓	-	-	97.03	145.1
	✓	-	✓	-	97.66	146.2
	✓	-	-	✓	97.83	145.0
	✓	✓	✓	-	97.85	146.3
	✓	✓	✓	✓	98.19	145.1
	✓	-	✓	✓	98.42	146.2
	✓	✓	✓	✓	<b>98.75</b>	<b>146.3</b>
ResNet-101	-	-	-	-	96.23	170.5
	✓	-	-	-	96.95	163.2
	✓	✓	-	-	97.34	163.3
	✓	-	✓	-	97.82	164.5
	✓	-	-	✓	98.23	163.2
	✓	✓	✓	-	98.12	164.6
	✓	✓	✓	✓	98.41	163.3
	✓	-	✓	✓	98.74	164.5
	✓	✓	✓	✓	<b>98.96</b>	<b>164.6</b>

用场景。改进 Retinex 算法取得了最好的检测结果,相对于不应用图像增强算法和 MSR 算法,AP 值分别提高了 0.87 和 0.48 个百分点,证明了本文所提算法可以有效提高钢轨表面缺陷的检测精度。

表 2 不同图像增强算法检测性能实验结果

Table 2 Experimental results of detection performance for different image enhancement algorithms

Algorithm	AP / %
-	95.74
HE	92.56
CLAHE	95.62
MSR	96.13
Improved Retinex	96.61

(Note: - means that the image enhancement algorithm is not applied)

#### 5.2.4 算法改进实验对比

此部分分别选取 ResNet-50、ResNet-101 作为特征提取器,验证 Cascade R-CNN 结合钢轨图像增强、IoU 平衡采样、RoIAlign、CIoU 损失之后的性能,实验结果如表 3 所示,表中 Backbone 表示特征提取器。

从表 3 中数据可以看出,以 ResNet-50 作为特征提取器,当引入图像增强之后,AP 值提高了 0.87 个百分点,并且可以发现,在引入图像增强后,检测时间缩短 6.8 ms,有利于钢轨表面缺陷的快速检测。在对钢轨表面图像进行增强的基础上,Cascade R-CNN 结合 IoU 平衡采样、RoIAlign、CIoU 损失之后,AP 值分别进一步提高了 0.42、1.05、1.22 百分点,且当同时引入三种方法时,AP 值提高 2.14 百分点,验证了 IoU 平衡采样、RoIAlign、CIoU 损失三种改进方法的有效性。表 3 中使用 ResNet-101 作为特征提取器的检测精度整体略高于使用 ResNet-50 作为特征提取器的检测精度,但是检测速度较慢,同样在同时引入图像增强、IoU 平衡采样、RoIAlign、CIoU 损失的条件下,使用 ResNet-101 作为特征提取器相对于使用 ResNet-50 作为特征提取器,AP 值提高 0.21 百分点,但检测时间增加了 18.3 ms,因此,结合检测精度和检测速度,ResNet-50 更适合作为本文钢轨表面缺陷检测算法的特征提取器。

图 8 展示了检测效果图,从图中可以看出,利用 Cascade R-CNN 较好地检测出了大部分缺陷,但对于小目标缺陷依然存在漏检的情况。在利用本文所

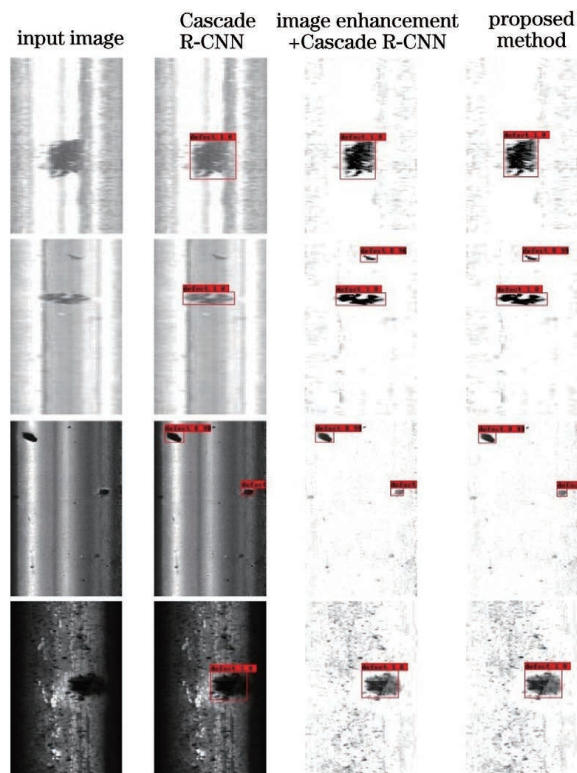


图 8 钢轨表面缺陷检测效果对比图

Fig. 8 Comparison diagrams of rail surface defect detection result

提图像增强方法对钢轨表面图像进行增强之后,Cascade R-CNN 准确检测出了漏检的缺陷,并从图中可以看出,进一步利用本文改进的 Cascade R-CNN 算法进行检测,预测边框的准确度明显提高,取得了更好的检测效果。

### 5.2.5 不同算法实验对比

本文将所提算法与其他一些经典算法针对钢轨表面缺陷检测的性能进行实验对比,表 4 展示了实验结果,表中 Proposed method 表示本文结合钢轨图像增强、IoU 平衡采样、RoIAlign、CIoU 损失对 Cascade R-CNN 进行改进之后的改进 Cascade R-CNN 算法。从表 4 可以观察到,本文所提算法的 AP 值明显高于其他算法,相对于 SSD<sup>[29]</sup>、YOLOv3<sup>[30]</sup>、YOLOv4<sup>[31]</sup>、Faster R-CNN、R-FCN<sup>[32]</sup>以及未改进的 Cascade R-CNN,AP 值分别提高了 17.52、13.21、8.38、4.32、11.12、2.52 百分点,SSD、YOLOv3、YOLOv4 的检测精度较低,这主要是因为无区域提名法对于钢轨表面缺陷检测的应用场景,不能很好地生成候选区域,另外 R-FCN 也没有取得良好的检测效果,这主要是因为钢轨表面图像具有很多小目标缺陷,而 R-FCN 将 RoI 划分为多个子区域,根据各个子区域是否出现目标的某个部分来综合判定目标,这种方式对钢轨表面的小目标缺陷检测效果不佳。在检测速度方面,SSD、YOLOv3 和 YOLOv4 速度较快,但其精度较低,本文所提算法的检测时间相对于 Faster R-CNN 和未改进的 Cascade R-CNN 分别缩短了 7.3 ms 和 24.2 ms,可以在保证一定检测速度的前提下,AP 值达到 98.75%,这对于钢轨表面缺陷检测具有更高的理论和实用价值。

表 4 不同算法性能比较

Algorithm	Backbone	AP /%	Time /ms
SSD	ResNet-101	81.23	58.3
YOLOv3	Darknet-53	85.54	39.5
YOLOv4	CSPDarknet-53	90.37	38.4
Faster R-CNN	ResNet-101	94.43	153.6
R-FCN	ResNet-101	87.63	95.8
Cascade R-CNN	ResNet-101	96.23	170.5
Proposed method	ResNet-50	<b>98.75</b>	<b>146.3</b>

## 6 结 论

本文针对钢轨表面缺陷检测过程中,由于钢轨

表面图像存在背景不均匀、缺陷尺度变化大且样本数据不足导致的缺陷检测困难的问题,提出了一种改进 Retinex 算法对钢轨表面图像进行处理,提高了缺陷与背景的对比度,并应用 IoU 平衡采样、RoIAlign 和 CIoU 损失对 Cascade R-CNN 算法进行改进,分别取得了一定的检测精度提升,验证了改进方法的有效性。同时,采用翻转变换、随机剪裁、亮度变换、生成对抗网络等方法增广钢轨表面缺陷图像数据集,有效地提高了图像数量和图像的多样性。与其他算法相比,该方法在保证一定检测速度的前提下,可以获得更高的检测精度,AP 值可达 98.75%,具有较高的理论和实用价值。

### 参 考 文 献

- [1] Gao B, Bai L B, Woo W L, et al. Automatic defect identification of eddy current pulsed thermography using single channel blind source separation[J]. IEEE Transactions on Instrumentation and Measurement, 2014, 63(4): 913-922.
- [2] Huang F Y. Quantitative evaluation method for eddy current testing of rail surface crack[J]. China Railway Science, 2017, 38(2): 28-33.  
黄凤英. 钢轨表面裂纹涡流检测定量评估方法[J]. 中国铁道科学, 2017, 38(2): 28-33.
- [3] Cruz F C, Filho E F S, Albuquerque M C S, et al. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing[J]. Ultrasonics, 2017, 73: 1-8.
- [4] Zhang J L, Wang Z C, Chen X, et al. High-speed railway rail ultrasound flaw detection systems based on phased array[J]. China Mechanical Engineering, 2019, 30(3): 339-344.  
张俊岭, 王子成, 陈潇, 等. 基于相控阵的高速铁路钢轨超声探伤检测系统[J]. 中国机械工程, 2019, 30(3): 339-344.
- [5] Zhang X, Feng N Z, Wang Y, et al. Acoustic emission detection of rail defect based on wavelet transform and Shannon entropy[J]. Journal of Sound and Vibration, 2015, 339: 419-432.
- [6] Song Y, Wu F, Liu D K, et al. Rail damage detection method based on acoustic emission and wavelet singularity[J]. Journal of Vibration and Shock, 2017, 36(2): 196-200.  
宋阳, 吴凡, 刘德扣, 等. 基于声发射及小波奇异性的钢轨损伤检测[J]. 振动与冲击, 2017, 36(2): 196-200.
- [7] Sui H, Gao X R, Luo L, et al. Finite element simulation and experiment on interaction of surface waves excited by laser point or line source with rail defects[J]. Laser & Optoelectronics Progress, 2019, 56(8): 081201.  
隋皓, 高晓蓉, 罗林, 等. 激光点源与线源激发表面波与钢轨缺陷作用的有限元仿真和实验[J]. 激光与光电子学进展, 2019, 56(8): 081201.
- [8] Tan Z W, Guo Y Y F. Study on rail wear detection system of laser triangulation method[J]. Transducer and Microsystem Technologies, 2016, 35(5): 20-24.  
谭周文, 郭雁一夫. 激光三角法钢轨磨损检测系统研究[J]. 传感器与微系统, 2016, 35(5): 20-24.
- [9] Li Q Y, Ren S W. Areal-time visual inspection system for discrete surface defects of rail heads[J]. IEEE Transactions on Instrumentation and Measurement, 2012, 61(8): 2189-2199.
- [10] Li Q Y, Ren S W. A visual detection system for rail surface defects[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012, 42(6): 1531-1542.
- [11] Zhang H, Jin X T, Wu Q M J, et al. Automatic visual detection system of railway surface defects with curvature filter and improved Gaussian mixture model[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(7): 1593-1608.
- [12] Gan J R, Li Q Y, Wang J Z, et al. A hierarchical extractor-based visual rail surface inspection system[J]. IEEE Sensors Journal, 2017, 17(23): 7935-7944.
- [13] Cao Y Q, Liu L B. Rail surface defect detection method based on background differential with defect proportion limitation[J]. Journal of Computer Applications, 2020, 40(10): 3066-3074.  
曹义亲, 刘龙标. 基于缺陷比例限制的背景差分钢轨表面缺陷检测方法[J]. 计算机应用, 2020, 40(10): 3066-3074.
- [14] Song Y N, Zhang H, Liu L, et al. Rail surface defect detection method based on YOLOv3 deep learning networks[C]//2018 Chinese Automation Congress (CAC), November 30-December 2, 2018, Xi'an, China. New York: IEEE Press, 2018: 1563-1568.
- [15] Su Y, Li J, Yang H M, et al. Research on rail surface defect recognition based on Faster R-CNN[J]. Electronic Science and Technology, 2020, 33(9): 63-68.  
苏焯, 李筠, 杨海马, 等. 基于 Faster R-CNN 的钢轨表面缺陷识别研究[J]. 电子科技, 2020, 33(9): 63-68.
- [16] Yuan H, Chen H, Liu S W, et al. A deep convolutional neural network for detection of rail surface defect [C]//2019 IEEE Vehicle Power and Propulsion Conference (VPPC), October 14-17,



- 2019, Hanoi, Vietnam. New York: IEEE Press, 2019: 1-4.
- [17] Liang B, Lu J, Cao Y. Rail surface damage detection method based on improved U-net convolutional neural network[J]. *Laser & Optoelectronics Progress*, 2021, 58(2): 021509.  
梁波, 卢军, 曹阳. 基于改进 U-Net 卷积神经网络的钢轨表面损伤检测方法[J]. *激光与光电子学进展*, 2021, 58(2): 021509.
- [18] Jin X T, Wang Y N, Zhang H, et al. DM-RIS: deep multimodel rail inspection system with improved MRF-GMM and CNN [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69 (4): 1051-1065.
- [19] Cai Z W, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6154-6162.
- [20] Pang J M, Chen K, Shi J P, et al. Libra R-CNN: towards balanced learning for object detection [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 821-830.
- [21] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386-397.
- [22] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34 (7): 12993-13000.
- [23] Hu Y B, Tang C, Tao S. Enhancement of optical Interferometry fringe patterns based on multiscale retinex algorithm[J]. *Laser & Optoelectronics Progress*, 2019, 56(24): 241006.  
胡一冰, 唐晨, 陶珊. 基于多尺度 Retinex 算法的光干涉条纹图像增强[J]. *激光与光电子学进展*, 2019, 56(24): 241006.
- [24] Yu P, Hao C C. Foggy image enhancement by combined fractional differential and multi-scale retinex[J]. *Laser & Optoelectronics Progress*, 2018, 55(1): 011012.  
余萍, 郝成成. 基于分数阶微分和多尺度 Retinex 联合的雾霭图像增强算法[J]. *激光与光电子学进展*, 2018, 55(1): 011012.
- [25] 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.
- [26] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [27] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [28] Zhang F J, Peng W G, Yu C J, et al. Algorithm for image enhancement of wood texture based on contrast-limited adaptive histogram equalization [J]. *Science Technology and Engineering*, 2020, 20(21): 8629-8636.  
张发军, 彭文刚, 虞成俊, 等. 对比度受限下直方图均衡化处理的木板纹路图像增强算法[J]. *科学技术与工程*, 2020, 20(21): 8629-8636.
- [29] 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. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [30] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2020-12-06]. <https://arxiv.org/abs/1804.02767>.
- [31] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2020-12-06]. <https://arxiv.org/abs/2004.10934>.
- [32] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks [EB/OL]. (2016-05-20) [2020-12-06]. <https://arxiv.org/abs/1605.06409>.