

改进 YOLOv8 的轨道扣件状态检测方法^{*}

范华琦 杨 柳

(西南交通大学信息科学与技术学院 成都 611756)

摘 要:针对现有轨道扣件状态检测算法对小目标物体与复杂形状物体的检测能力差而导致检测结果异常,以及小目标层特征冗余等问题,提出了一种改进 YOLOv8 的轨道扣件状态检测方法。在 YOLOv8 网络中增加可变形空间金字塔扩张卷积模块,以提高模型对小目标物体以及形变复杂物体的检测精度。同时增加小目标空间重构单元以减少小目标特征冗余,促进小目标特征的学习。根据采集到的轨道扣件数据集进行模型的训练和测试,并与多组轨道扣件状态检测算法进行对比,实验结果表明,相较于对比算法,所提算法精确度平均提升 3.20%,召回率平均提升 3.34%,平均精度平均提升 3.96%。实验证明所提算法能够有效进行轨道扣件状态检测,并且具有较强的泛化能力,可以部署于复杂交通场景。

关键词:图像处理;YOLOv8;轨道扣件;小目标检测

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.604

Improved YOLOv8 track fastener state detection method

Fan Huaqi Yang Liu

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: Aiming at the problems of poor detection ability of existing track fastenings state detection algorithms on small target objects and complex shapes, which lead to abnormal detection results, as well as feature redundancy of small target layer, an improved YOLOv8 track fastenings state detection method was proposed. A deformable space pyramid expansion convolutional module is added in YOLOv8 network to improve the detection accuracy of small target objects and deformable complex objects. At the same time, the reconstruction unit of small target space is added to reduce the redundancy of small target features and promote the learning of small target features. The model is trained and tested according to the collected data set of track fasteners, and compared with multiple groups of track fasteners state detection algorithms. The experimental results show that compared with the comparison algorithm, the accuracy of the proposed algorithm is increased by 3.20% on average, the recall rate is increased by 3.34% on average, and the average accuracy is increased by 3.96% on average. The experiments proves that the proposed algorithm can detect the state of rail fastenings effectively, and has strong generalization ability, and can be deployed in complex traffic scenarios.

Keywords: image processing; YOLOv8; track fastener; small target detection

0 引 言

近年来,随着中国城市轨道交通的快速发展,轨道安全问题变得越来越突出。扣件的主要功能就是把铁路线固定在铁路枕上,维护铁路线的间隔,防止铁路线的垂直或水平偏离。正确的扣件设置能够确保铁路线的完整性,从而确保火车的顺利、平稳地驶过。由于多种原因,轨道

扣件可能存在着移动、损坏或丢弃的情况,这些情况很有可能导致重大的交通事故。所以,正确地定位和检查轨道上的扣件非常必要。

在传统的轨道扣件状况辨认手段中,包括了人工巡逻测试法^[1]、振动信息测试法^[2]、激光测试法^[3]以及图像辨认法^[4-5]。然而,这些手段的测试效能较差,其精确性还会因为工作人员的巡逻经历而产生变化。由于铁道巡查手

收稿日期:2023-11-02

^{*} 基金项目:轨道交通工程信息化国家重点实验室(中铁一院)开放课题(SKJKZ22-02)、宜宾市双城市校协议专项科研经费科技项目(SWJTU2021020005)、朔黄铁路公司科研项目(SHYP-22-01)资助

段的自动化和智能化程度逐渐提高,基于图像采集与处理的扣件状态检测方式已成为高速铁路运行轨迹工作全过程自动化巡查的主要手段。

基于图像信息处理的测试方法是根据轨道现有的图像处理测试方法根据轨道扣件的特性和布局制定。刘甲等^[6]使用金字塔式梯度方向直方图和画面整体宏观肌理特性,利用支持向量机完成扣件划分。Gibert等^[7-8]采用多任务学习架构组合多检测器的扣件检测算法,通过贝叶斯框架组合极值理论进行优化。孙睿等^[9]针对运用图像处理技术在完成轨迹扣件位置的过程中易受到拍摄条件及复杂背景干扰的难题,给出了一套根据流行排序的轨迹扣片位置方案。袁文祥等^[10]通过预先处理所得的信息,采用十字交叉法来确定扣件部分的具体位置;同时,也利用了创建的快速归一化积相关算法的模型来精确地定位扣件部分的绝缘帽以及紧固螺丝,最后确认了其的中心坐标。根据所得的数据,构建了一套以特定点为依据的扣件状况辨认策略。赵建龙等^[11]成功地构建了一种用于准确识别轨道扣件缺陷的方法。王宇桐等^[12]利用这种方法,通过压缩图像并采用 Sobel 边缘检测技术,以及灰度投影技术,可以准确地识别出钢轨的位置,使得轨道扣件的质量更加准确可靠。根据扣件的安装距离,戴鹏等^[13]开发了一种新的方法来精确地定位扣件。该方法利用了一个叫做置信图的技术,可以在概率图上创造一个连接两个不同类别的图像,从而更加准确地找到扣件的中心。

尽管当前的图像处理手段已经能迅速且详尽地辨认出扣件的实际状态,但仍需针对每个扣件单独进行定位和图像属性的抽取,并且必须依照特定的被检查目标做出相应的算法调整,这使得它在面临复杂的实际环境时显得力不从心。虽然目前的图像处理手段已经能够辨认出扣件的实际状态,但是检测的准确率仍然较低。

由于新一代人工智能科学技术的发展,采用机器视觉的轨道扣件病害检验在理论和实践等领域方面得到了一些发展。刘奇锋等^[14]将 YOLOv3 原有的 3 个不同尺度特征层输入输出改为 4 个不同尺度特征层输入输出,并且根据扣件的不同采用 K-means 聚类方法来确定先验框的不同参量,以达到检验小目标的目的。王婧瑶等^[15]提出了一种改进 Mask R-CNN 算法,可以有效减少光照等环境干扰,同时减除图像采集时剧烈抖动的问题。白堂博等^[16]提出一种改进 Mask R-CNN 的轨迹扣件状态检测方法,使用对轨迹扣件边界像素点加以标记,Mask R-CNN 实例分割算子能够将轨道扣件与图像中的背景区分,只关心与扣件相关的元素数据,并且使用最小外接矩形法修改 Mask R-CNN 的输出层,计算结果得出扣件图像位置数据信息,进而确定偏移扣件。通过引入 K 均值聚类算法,大幅度优化了 Mask R-CNN 实例分割算法,以更加精细地识别和评估轨道上的零部件^[17]。

但是实际交通场景包含众多小目标,同时轨道扣件的

形状复杂,不利于常规卷积神经网络(CNN)提取特征,使用常规目标检测算法存在漏检严重问题。针对上述问题,本文提出一种改进 YOLOv8 算法,用 SD-Conv 模块代替每个跨步卷积层和每个池化层,使 YOLOv8 网络在低分辨率图像和小物体的更艰巨任务上具有较高的检测效果,同时卷积核的形状可以根据图像的内容自适应的发生改变,进一步提升对形变复杂物体的检测性能。同时引入小目标空间重构单元(small target layer-spatial reconstruction unit,ST-SRU),减少小目标层的冗余特征,并强化小目标层的代表性特征,使之多样化,进一步提升小目标物体的检测精度。

1 基于改进 YOLO 的扣件状态检测方法

目标检测模型的演变主要包括两阶段的目标检测算法与单阶段的目标检测算法,这两种算法都依赖于深度学习。前者采用了回归技术,用于对目标的识别,后者则是一种依赖于预设区域的策略。其中,基于候选区域的检测行为由两个阶段实现:第 1 阶段生成目标的候选区域,第 2 阶段对候选区域中的候选框大小和位置进行预测,最后生成预测框。该类算法在识别准确率以及定位精度上有着不错的效果,但网络模型复杂,运算速度慢。代表算法有 Faster R-CNN (faster region-based convolutional neural network)^[18],而基于单阶段的目标检测算法则是直接生成相应的候选框,将分类和回归任务都视作回归问题。相比于两阶段目标检测算法,其虽然并不具备更好的检测能力,但模型简单,运算速度较快。代表算法 YOLO^[19-20]系列。

为快速识别轨道扣件状态,本文选择基于 YOLO 的目标检测算法来进行轨道扣件状态检测。

1.1 YOLO 目标检测算法

YOLO 网络发展至今已有多个版本,其结构一般分为骨干网、特征融合层、3 个不同大小的预测头。其用 3 个不同大小预测头分别检测不同大小的目标,以在追求速度的同时保证精度。

2022 年,石欣等^[21]在 YOLOv4 的基础上引入浅层特征改进特征金字塔,提高了对小目标物体的检测精度。肖黎俊等^[22]使用提出了对数据集进行聚类分析以得到最佳锚框尺寸的改进 YOLOv5 算法,有效提高模型的泛化能力和定位精度。钱承山等^[23]提出了一种针对目标排列紧密,背景复杂的问题的改进 YOLOv5s 算法。

同年,文献^[24]提出 YOLOv7,主要是针对模型结构重参化和动态标签分配问题进行了优化。YOLOv7 提出了计划的模型结构重参化,并借鉴了 YOLOv5、Scale YOLOv4、YOLOX,“拓展”和“复合缩放”方法,以便高效的利用参数和计算量。引入了辅助训练模块一由粗到细(coarse-to-fine)引导标签分配策略。

2023 年,Reis 等^[25]提出 YOLOv8,相比 YOLOv5,最大的改动是将 Anchor Base 替换成 Anchor Free, TAL

(task-aligned-assigner)动态匹配的方式也是和 YOLOv6 很相似,但参数略有不同。相比于 YOLOv7, YOLOv8 采用了更先进的 backbone 和优化后的分类器。此外, YOLOv8 还加入了 mask 分支,能够有效地提高遮挡目标的检测准确率,大大提升了模型的适应性。在模型性能方面, YOLOv8 相比 YOLOv7 在目标检测的准确率和召回率方面有所提升。这意味着 YOLOv8 能够更准确地检测和识别目标物体。在速度与效率方面, YOLOv8 相比 YOLOv7 在运行速度和计算效率方面有所提升。YOLOv8 使用了一些优化技术,使得在保持准确率的前提下,能够更快地进行目标检测。

1.2 可变形空间金字塔扩张卷积结构 SD-CONV

CNN 可以解决许多复杂的计算机视觉问题,比如图像识别、对象检测等。由于目前的 CNN 框架存在一些问题,例如使用跨步卷积和池化层,可能会造成细节信息的

损耗,以及特征的过载。因此,需寻找一种新的方法来提高 CNN 检测的效果。Sunkara 等^[26]提出的 SPD-Conv (spatial pyramid dilated-Conv)卷积技术能够很好地解决跨步卷积层与池化层之间的差异,从而更准确地预测复杂的地理信息。不过,由于卷积神经网络本身的局限性,它仍会面临一些挑战。该模型使用的卷积和传统的卷积操作类似,对形变复杂的物体检测效果一般。可变性卷积 Deformable Conv 可以使卷积核的形状可以根据图像的内容自适应地发生改变,提高形变复杂物体的检测精度。基于此,在 SPD-Conv 模型的基础上对 YOLOv8 模型进行改进,在 YOLOv8 网络中添加 SD-CONV 模块,该模块由空间金字塔扩张 SPD 模块与可变形卷积 Deformable Conv 共同组成。使 YOLOv8 在提升小目标物体检测精度的同时提高对形变复杂物体的检测精度。经过改进的 YOLOv8 网络架构如图 1 所示。

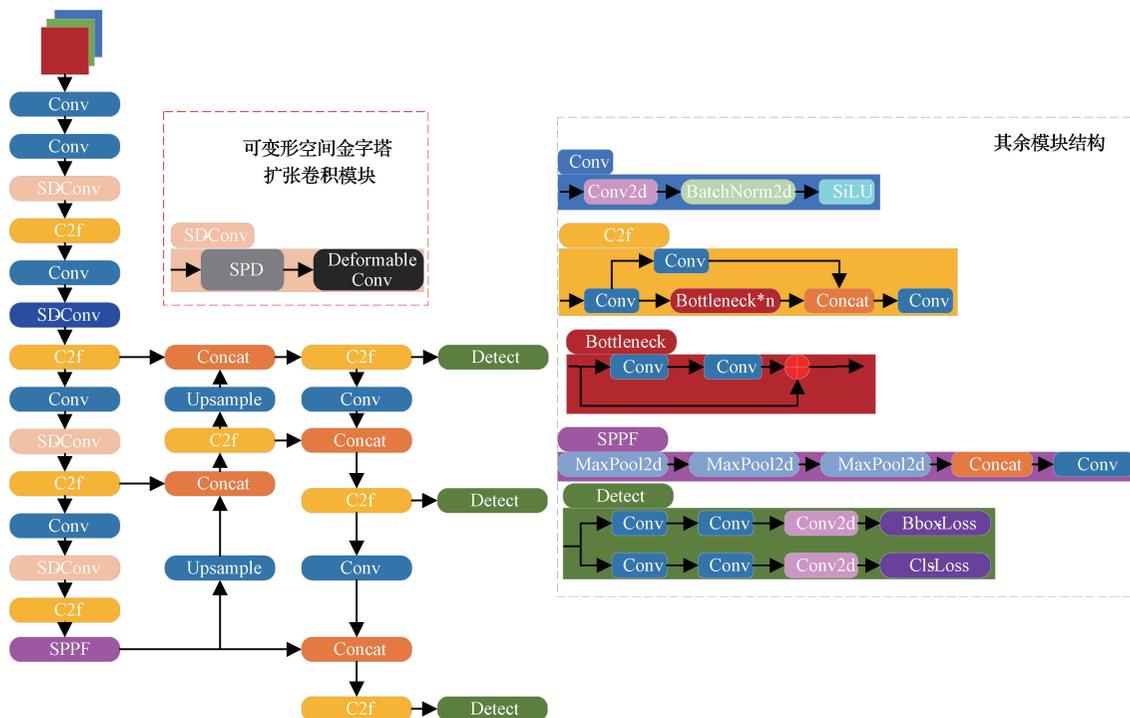


图 1 增加可变性空间金字塔扩张卷积后的 YOLO 网络结构

Fig. 1 YOLO network structure after adding small target space reconfiguration unit with variable spatial pyramid

空间金字塔扩张 SPD 层的作用是将输入特征图的每个空间维度降低到通道维度,同时保留通道内的信息。这可以通过将输入特征图的每个像素或特征映射到一个通道来实现。在这个过程中,空间维度的大小会减小,而通道维度的大小会增加。这种组合方式可以在不丢失信息的情况下减少空间维度的尺寸,同时保留通道内的信息,有助于提高 CNN 对低分辨率图像和小型物体的检测性能。

SPD 组件将原始图像转换技术推广到对 CNN 内部和整个 CNN 的特征图进行下采样,如下所示。

考虑任意大小为 $A \times A \times C_1$ 的中间特征映射 X ,将子特征映射序列进行切片,公式如下:

$$\begin{aligned}
 f_{0,0} &= X[0:A;S,0:A;S], \\
 f_{1,0} &= X[1:A;S,0:A;S], \dots, \\
 f_{s-1,0} &= X[S-1:A;S,0:A;S], f_{1,1}, \dots, \\
 f_{s-1,1} &= X[S-1:A;S,1:A;S]; \\
 &\vdots \\
 f_{0,s-1} &= X[0:A;S,S-1:A;S], f_{1,s-1}, \dots, \\
 f_{s-1,s-1} &= X[S-1:A;S,S-1:A;S]
 \end{aligned} \tag{1}$$

通常,给定任何(原始)特征图 X ,子图 $f_{x,y}$ 由 $i+x$ 和 $i+y$ 可按比例整除的所有条目 $X(i+y)$ 形成。因此,每个子图按比例因子对 X 下采样。图 2 所示为 $S=2$ 时的 SPD 示意图,得到了 4 个子图 $f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}$, 每个都具有形状 $(A/2, A/2, C_1)$, 并将 X 下采样 2 倍。

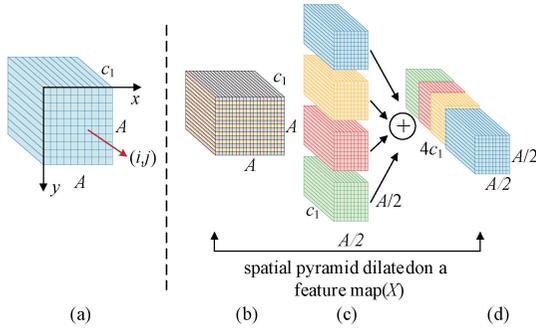


图 2 空间金字塔扩张卷积结构

Fig. 2 Spatial pyramid expansion convolution structure

沿着通道维度连接这些子特征图,从而获得特征图 X' ,他的空间维度减小了一个比例因子,通道维度增大了 2 倍。即 SPD 将特征图 $X(A, A, C_1)$ 转换为中间特征图 $X'(A/S, A/S, S^2C_1)$ 。

传统的卷积操作是将特征图分成一个个与卷积核大小相同的部分,然后进行卷积操作,每部分在特征图上的位置都是固定的。在处理形状复杂的物体,例如轨道扣件时,采取的卷积技术通常不能满足需求。为了解决这个问题,可以采取多项措施:收集大量的复杂形状的样本,利用 trick 技术进行模型优化,并结合自动化的特征提取与分析算法。采用数据集加以改进的方式往往会导致效率低下,并且必须使用更加精细的网格架构;相比之下,采用人为干预的特征提取技术则更具挑战性。

在空间金字塔扩张结构后使用传统卷积操作对轨道扣件状态的检测效果不佳。通过引入 SPD,能够让卷积核的外观和大小随着图像的不同而有所不同,而且还将其与 Deformable Conv^[27] 相连,这种 Conv 能够通过调整来实现对采样网格的灵活调整。不同卷积 3×3 采样位置示意图如图 3 所示,可以看到,在增加了一个新的卷积层后,可以更准确地识别出原始的特征。这样,就可以根据输入的信息,采用局域、紧凑且自动化的模型来处理。

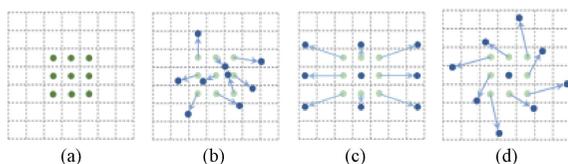


图 3 标准和可变形卷积 3×3 采样位置示意图

Fig. 3 Illustration of the sampling locations in 3×3 standard and deformable convolutions

图 3(a) 是一个按照一定原理进行采样的绿色网格。

图 3(b) 为改进后的卷积,纵横比发生了显著的改变,并且随着时间的推移而发生扩展,图 3(c) 和 (d) 都是不同的尺度、纵横比以及旋转。

为了更加有效地表达可变形卷积,将其应用于 2D 维度,而非特征图的 3D 维度。

2D 偏移是一个复杂的过程,需要两个复杂的过程包括两个步骤:首先通过 grid R 对输入特征图 x 中的数据进行采集;其次通过计算每个数据的权重 w 来实现 2D 偏移。grid R 定义了感受野尺寸和空洞。例如, $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$ 就定义了一个空洞率为 1 的 3×3 核。

对于输出特征图中的每一个位置 p_0 , 有:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \quad (2)$$

式中: p_n 枚举 R 中的位置。

在可变形卷积中,常规网络 R 被增加偏移 $\{\Delta p_n | n = 1, \dots, N\}$, 其中 $N = |R|$ 。式(2)就变成了如下形式:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3)$$

采样可以从 $p_n + \Delta p_n$ 的不规则位置开始,并以分数的形式进行,以此来实现双线性插值,公式如下:

$$x(p) = \sum_q G(q, p) \cdot x(q) \quad (4)$$

式中: p 代表任何一个位置, $p = p_0 + p_n + \Delta p_n$; q 则是从 x 的所有位置中抽取出来的; $G(\cdot, \cdot)$ 则是一个双线性插值的和函数,它具有二维结构,可以被分割成两个一维核。

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y) \quad (5)$$

其中 $g(a, b) = \max(0, 1 - |a - b|)$ 。式(4)的计算速度非常快是因为除了少数的 q_s 其余位置都为 0。

3×3 可变形卷积如图 4 所示,偏移是由一个卷积层产生的,它的卷积核和常规的卷积层没有区别。在这个说明里,输出的偏差范围和输入的特征值的大小是相同的。每条信号的 $2N$ 维度代表了 N 次 2D 的偏差。在训练阶段,需要同步学习用于产生输出特性的卷积核以及偏移。为了掌握偏移,采取了双线性操作并进行反向传播的方式,从而实现了产生偏移和输出特性的卷积核的同步训练。

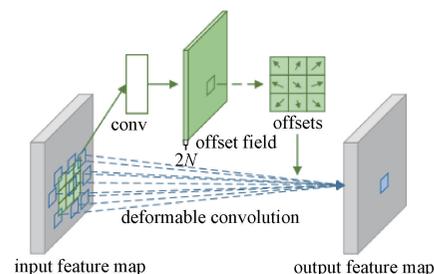


图 4 3×3 可变形卷积

Fig. 4 Illustration 3×3 deformable convolution diagram

1.3 ST-SRU

为进一步提升 YOLOv8 对小目标物体的检测精度,

减少冗余,使用 ST-SRU 模块,ST-SRU 模块由小目标检测层和空间重构单元组成。

YOLOv8 对象检测模型有 3 个输出层,为了提升对小目标的检测能力,添加第 4 个输出层。Backbone 部分的结果没有改变,但是 Neck 部分和 Header 部分模型结构

进行了改变。以提升小目标物体的检测精度。并在小目标检测层中插入空间重构单元(SRU)(可以理解作为一种注意力机制),减少小目标层的冗余特征,并强化小目标层的代表性特征,使之多样化。SRU 采用分离重构方法抑制信道空间冗余,其结构如图 5^[28]所示。

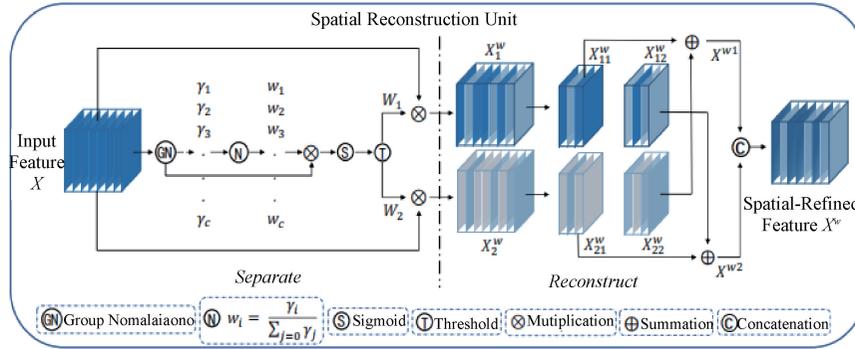


图 5 空间重构单元结构

Fig. 5 Spatial reconstruction unit structure

改进后的网络如图 6 所示,利用了分离和重构操作。分离操作的目的是将这些信息丰富的特征映射与空间内容对应的信息较少的特征映射分开。利用组标准化(GN)^[29]层中的缩放因子来评估不同特征图的信息内容。给定中间 feature map $X \in R^{N \times C \times H \times W}$,其中 N 为批处理轴, C 为通道轴, H, W 为空间高、宽轴。先用均值 μ 除以标准差 σ 对输入特征 X 进行标准化:

$$X_{out} = GN(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (6)$$

式中: μ 和 σ 为 X 的均值和标准差。

本文利用 GN 层中的可训练参数 $\gamma \in R^C$, 作为测量每个批次和通道的空间像素方差的一种方法。更丰富的空间信息反映了更多的空间像素变化,有助于更大的 γ 。归一化相关权重 $W_\gamma \in R^C$ 表示不同特征图的重要性,公式如下:

$$W_\gamma = \{w_i\} = \frac{\gamma_i}{\sum_{j=1}^c \gamma_j} \quad i, j = 1, 2, \dots, C \quad (7)$$

然后将经重新加权的特征映射的权值通过 sigmoid 型

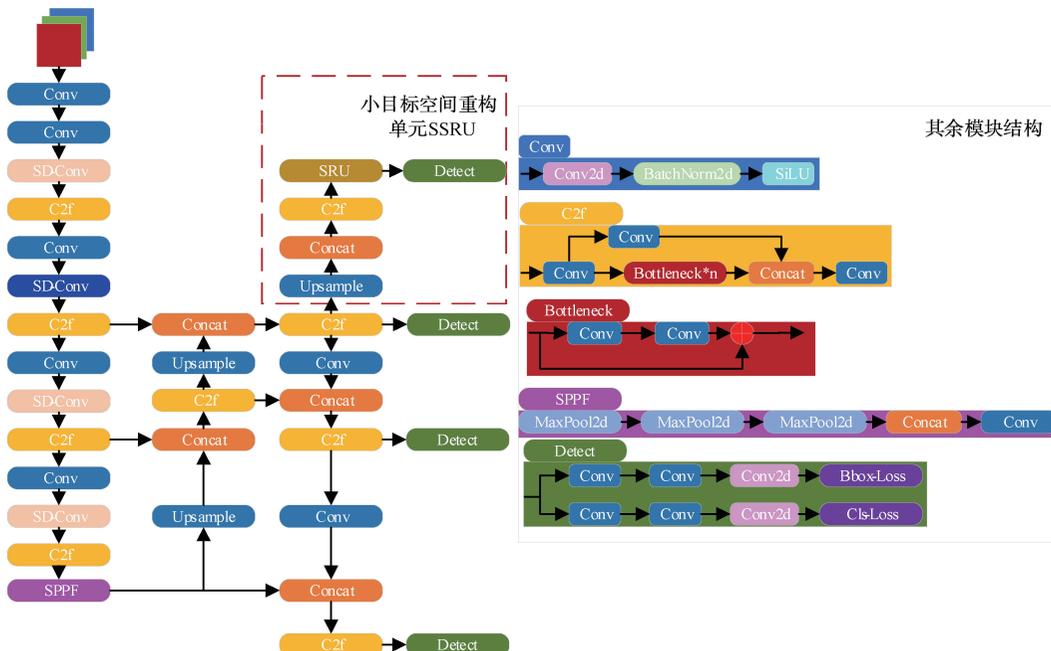


图 6 增加小目标空间重构单元后的 YOLO 网络结构

Fig. 6 YOLO network structure after adding small target space reconfiguration unit

函数映射到范围(0,1),并通过一个阈值进行控制。将高于阈值的权重设置为1,获得信息权重,将其设置为0,获得非信息权重(阈值设置为0.5)。获取W的整个过程公式如下:

$$W = Gate(sigmoid(W_\gamma(GN(X)))) \quad (8)$$

最后,将输入特征X分别乘以W₁和W₂,得到两个加权特征:信息丰富的特征和信息较少的特征。因此,成功地将输入特征分为两部分。

2 实验结果与分析

2.1 实验数据集

利用一种新型的轨道巡检仪来收集数据。采集到6371张铁路轨面图片,采集到的图片包括雨天、晴天、夜间等多种情况,其中雨天采集到图片数量为1246张,晴天采集到的图片为3197张,夜间采集到的图片数量为1928张,图片的大小都为2048×2048 pixels。按照6:2:2的比例随机将采集到的图像划分为训练集、验证集、测试集,将其像素大小压缩为512×512,并使用在线工具Make Sense标注了扣件的状态和位置,如图7所示。

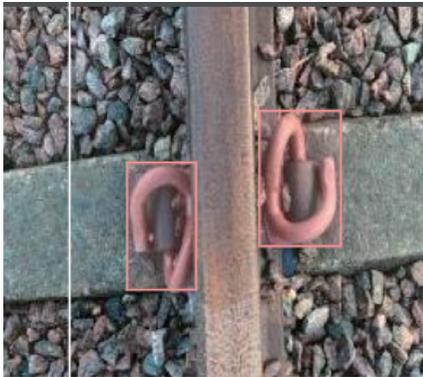


图7 Make Sense 标注信息

Fig. 7 Make Sense of the annotation information

扣件状态可分为e型扣件(fastener)、w型扣件(fastener-2)、w型扣件断裂(fastener_broken)、e型扣件断裂(fastener2_broken)、扣件丢失(missing)、扣件异物(trackbed_stuff)6类。不同类型轨道扣件状态样例如图8所示。

2.2 实验平台及参数配置

本文实验的实验环境如表1所示。

2.3 评价指标

所采用的评价指标包括精确度(precision),召回率(recall)和平均精度均值(mean average precision, mAP)。

Recall值可以反映样本中正确的情况的数量,其度量的是查全率,比如在轨道扣件异常状态场景中,要求模型有更高的recall,确保每一个异常轨道扣件状态都被检测出来。公式如下:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$



图8 轨道扣件状态示例

Fig. 8 Example diagram of track fastener status

表1 实验环境

Table 1 Experimental environment

环境	版本
CPU	AMD Ryzen9 5900HX
GPU	RTX3070 Laptop GPU
操作系统	Windows
平台	Pycharm
框架	Pytorch
CUDA	11.3

Precision是衡量预测为正样本中真正为正样本的比例。在轨道扣件状态检测中,需要高Precision来确保准确检测出每种轨道扣件状态。公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

式中:TP代表将正类鉴定为负类的次数;FP则表示将负类鉴定为正类的次数;FN则表示将正类鉴定为负类的次数;TN则指将负类鉴定为正类的次数。

mAP是一种广泛应用于多目标检测和多标签图像分类的评估指标,通过观测该指标更好地识别出多个标签,从而减少检测失误的可能性。此外,mAP的提升也可以有效地提升整体的检测效率。根据Precision-Recall曲线,可以计算出每个Recall值的平均值,从而得出一个数组形式的评估结果AP:

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) Pinter(r_i + 1) \quad (11)$$

式中: r_1, r_2, \dots, r_n 分别代表Precision插值段第1个位置的Recall,以升序顺次出现。

所有类别AP的均值就是mAP。mAP公式如下:

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (12)$$

最后添加每秒帧数(frame per second, FPS)评估模型每秒处理的帧数:

$$FPS = \frac{1\ 000}{\text{Processing time pre frame}} \quad (13)$$

其中, Processing time per frame 代表每帧处理时间。

2.4 网络训练

训练阶段的参数设置如表 2 所示。

表 2 训练参数

Table 2 Training parameter

参数	取值
优化器	SGD
Epoch	300
批量大小	16
学习率	0.01

为了确保模型的准确性和稳定性,实验将学习率调节至 0.01,这样可以有效地抑制小规模数据的过拟合,从而避免模型出现振荡现象。部分训练结果如图 9 所示。

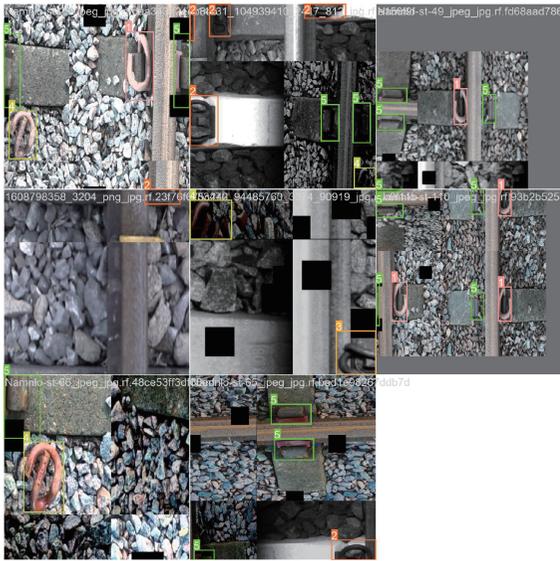


图 9 部分训练结果

Fig. 9 Part of the training results

2.5 对比实验

为了验证本文方法的有效性,使用本文的方法进行轨道扣件状态检测,并与 Faster R-CNN^[18]、Mask R-CNN^[30]、改进 Mask R-CNN^[16]、YOLOv7^[24]、YOLOv8^[25] 检测算法进行对比。为了保证对比实验的合理性,使用同一数据集训练上述模型并进行测试。不同算法训练过程的损失变化曲线如图 10 所示。并给出了训练过程中的精确度变化曲线,其结果如图 11 所示。从图 10、11 可以看出, Faster R-CNN 等双阶段目标检测模型在训练次数不到 150 次时就完成了收敛,收敛速度快于 YOLO 等单阶段目标检测算法,但是检测精确度较差。

选取数据集中一张复杂的轨道扣件图像进行检测结果的对比,如图 12 所示。

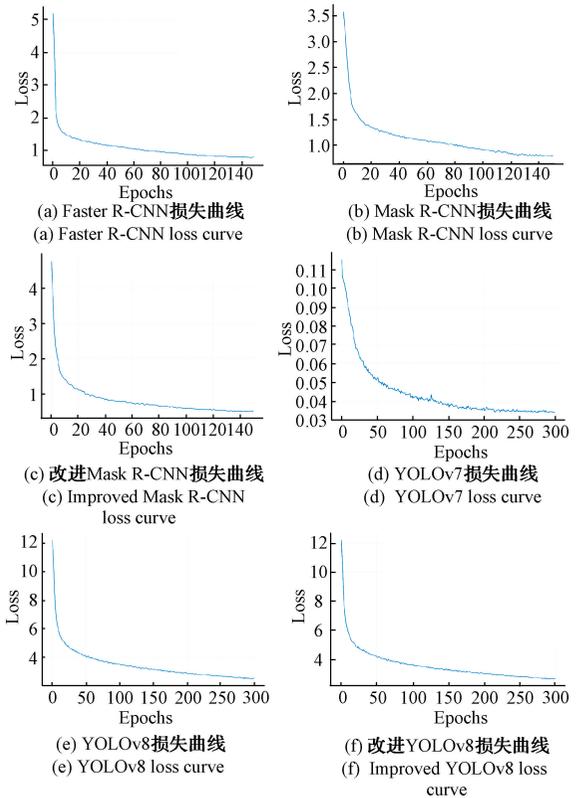


图 10 不同算法训练损失变化曲线

Fig. 10 Training loss for different algorithms change curve

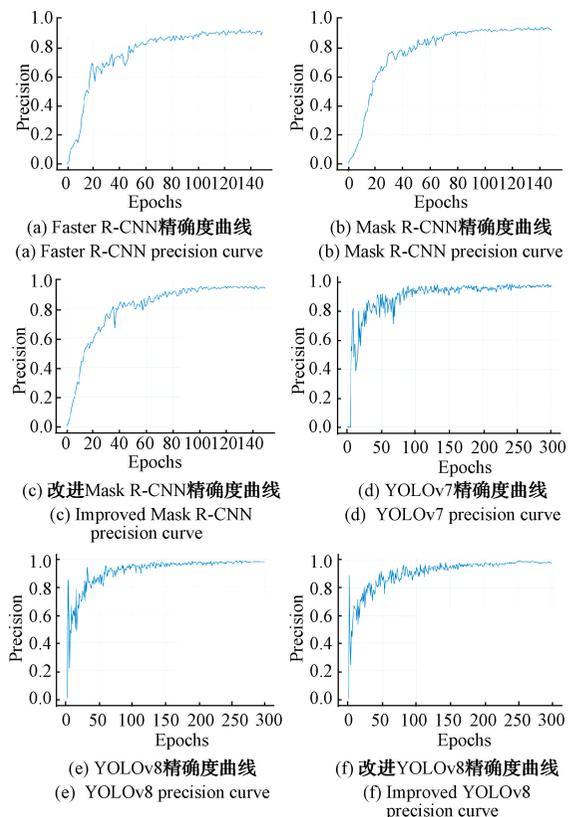


图 11 不同算法训练精确度变化曲线

Fig. 11 Different algorithm training precision change curve

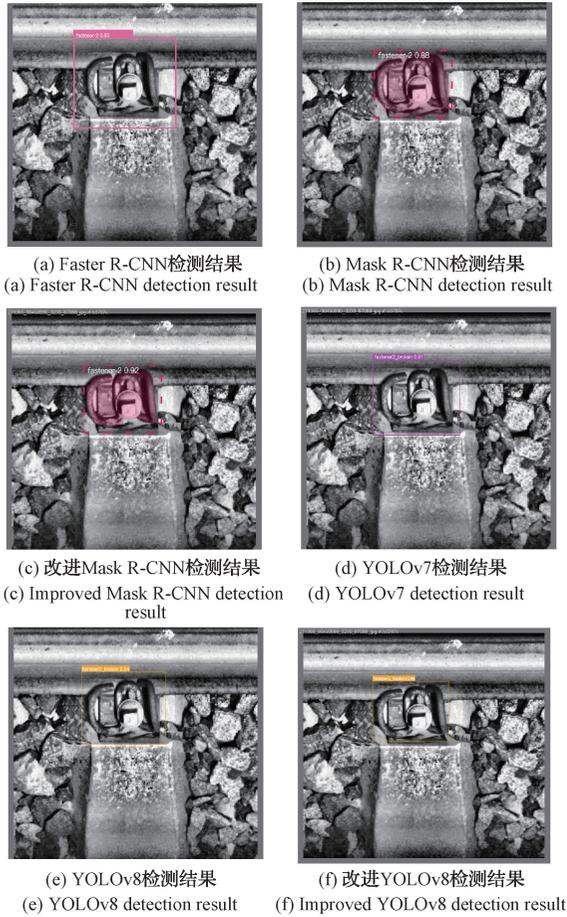


图 12 不同算法测试结果

Fig. 12 Test results of different algorithms

从图 12 可以看出,这些算法并没有出现漏检的情况。Faster R-CNN、Mask R-CNN 算法、改进 Mask R-CNN 算法出现了误判的问题,把损坏的 W 扣件识别为了正常的样本(将 fastener2_broken 识别为了 fastener-2)。YOLOv7、YOLOv8 算法相较于前面 3 种算法,没有出现误判的现象,但是其对于物体的定位效果较差,精确程度明显低于本文所提算法。

为了进一步对检测结果进行客观的定量评价,给出不同算法评价指标的对比,如表 3 所示。

表 3 不同算法扣件状态检测结果

Table 3 Different algorithm fastener status detection results

Model	Precision	Recall	mAP	检测速度/fps
Faster R-CNN ^[18]	0.921	0.884	0.915	8.1
Mask R-CNN ^[30]	0.936	0.914	0.928	5.3
改进 Mask R-CNN ^[16]	0.951	0.936	0.946	13.1
YOLOv7 ^[24]	0.964	0.941	0.956	58.0
YOLOv8 ^[25]	0.973	0.943	0.962	66.2
改进 YOLOv8	0.991	0.957	0.981	57.4

由表 3 可以看出,本文提出的算法相较于其他算法,检测效果更好,Precision 平均提升 3.20%,Recall 平均提升 3.34%,mAP 平均提升 3.96%。虽然检测速率相比原始的 YOLOv8 算法有所下降,检测速度下降了 8.8 fps。但其检测速率基本与 YOLOv7 持平,高于 Faster R-CNN、Mask R-CNN、改进 Mask R-CNN 等算法。由于实时检测一般要求检测帧率大于 25 fps,故本文算法同样可以满足实时性要求,可有效实时判定出轨扣件的状态,显著提高轨道扣件状态检测的准确率。

2.6 消融实验

本文提出的改进方法分别为插入 SD-Conv 模块,即使用空间金字塔扩张 SPD 与可变性卷积 Deformable Conv;引入小目标空间重构单元 ST-SRU 模块。

为进一步验证方法的有效性,本文对两组消融实验进行了深入研究,包括以原始的 YOLOv8 为基础,分别只增加一种改进方法,验证每一种方法的有效性;以增加一种改进方法后的算法为基础,再次增加一种改进方法,验证两种方法配合使用的有效性。实验结果如表 4 所示。

表 4 消融实验结果

Table 4 Ablation results

实验	方法	Precision	Recall	mAP	检测速度/fps
1	YOLOv8	0.973	0.943	0.962	66.2
2	YOLOv8+SD-CONV	0.982	0.949	0.968	69.4
3	YOLOv8+ST-SRU	0.980	0.948	0.966	53.4
4	YOLOv8+SD-CONV+ST-SRU	0.991	0.957	0.981	57.4

从表 4 可以看出,实验 1 是不包括 3 种优化策略的 YOLOv8 原模型,实验 2 是只加上 SD-Conv 卷积构建块的模型,可以看出增加 SPD-Conv 卷积构建块后对比原模型 Precision、Recall、mAP 分别提高了 0.9%、0.6%、0.9%,检测速度提高了 3.2 fps。实验 3 是只加上 ST-SRU 的模型。可以看出增加 SD-Conv 卷积构建块后对比原模型 Precision、Recall、mAP 分别提高了 0.7%、0.7%、0.4%。实验 4 为在同时增加 SD-Conv 与 ST-SRU 的模型。对比实验 2 模型,实验 4 的模型的 Precision、Recall、mAP 分别提高了 0.9%、0.8%、1.3%,检测速度提高了 3.2 fps。对比实验 3 模型,实验 4 的模型的 Precision、Recall、mAP 分别提高了 1.1%、0.9%、1.4%。对比原始模型,实验 4 的模型的 Precision、Recall、mAP 分别提高了 1.8%、1.4%、1.9%。

因此得出结论单独插入可变形空间金字塔扩张 SD-Conv 或小目标空间重构单元 ST-SRU 均可以有效提高模

型的检测效果。同时插入 SD-Conv 与 ST-SRU 的完整模型检测效果最好。并且无论插入哪种模型,检测速率都大于 25 fps,满足实时性的要求。

为进一步验证插入 SD-Conv 的有效性。用 YOLOv8, YOLOv8+SPD-Conv, SD-Conv 进行实验对比,结果如表 5 所示,插入 SPD-Conv 卷积构建块后模型的 Precision、Recall、mAP 分别提高了 0.5%、0.4%、0.2%,FPS 下降了 4.1 fps。插入 SD-Conv 卷积构建块后,对比 SPD-Conv 模型,模型的 Precision、Recall、mAP 分别提高了 0.4%、0.2%、0.4%,检测速度提高了 7.3 fps。可见本文提出的 SD-Conv 模块检测效果优于 Sunkara 等^[26]提出的 SPD-Conv 模块。

表 5 引入 SD-CONV 实验对比
Table 5 The experimental comparison of SD-CONV was introduced

实验	方法	Precision	Recall	mAP	检测速度/fps
1	YOLOv8	0.973	0.943	0.962	66.2
2	YOLOv8+SPD-CONV	0.978	0.947	0.964	62.1
3	YOLOv8+SD-CONV	0.982	0.949	0.968	69.4

为进一步验证使用 ST-SRU 模块的有效性,使用 YOLOv8, YOLOv8+SRU, YOLOv8+Small Target 检测层, YOLOv8+ST-SRU 进行实验,结果如表 6 所示。

表 6 引入 ST-SRU 实验对比
Table 6 The experimental comparison of ST-SRU was introduced

实验	方法	参数量/($\times 10^6$)	Precision	检测速度/fps
1	YOLOv8	301.22	0.973	66.2
2	YOLOv8+SRU	301.98	0.978	58.3
3	YOLOv8+Small Target 检测层	292.75	0.975	62.1
4	YOLOv8+ST-SRU	292.94	0.980	53.4

由表 6 可得增加小目标检测层可以减少模型参数,并带来少量的精度提升。使用空间重构单元 SRU 可以在基本不增加模型大小的前提下取得明显精度提升。同时增加小目标检测层与空间重构单元即使用 ST-SRU 的结构不仅可以有效地降低网络参数的数量,而且还能够更好地优化网络结构,从而极大地提升了小目标物体检测的效率。因此得出结论:引入 Small Target 检测层或者引入 SRU 均可以带来有效的精度提升并且无论增加哪一部分都只会带来少量的速率下降,并不会对检测速率有较大的

影响,将两部分进行结合得到的的效果最好。

3 结论

本文提出一种改进 YOLOv8 算法用于轨道扣件状态检测,得出如下结论。

1)通过插入 SD-Conv 可以有效细粒度信息的丢失和特征冗余,提高模型对小目标模型的检测精度,并且可以提升模型对形变复杂物体的检测精度,从而进一步提高模型的检测精度。

2)使用 ST-SRU 模块可以有效减少小目标层的冗余特征,强化了小目标层的代表性特征,使小目标检测层的特征多样化,在减少网络参数的同时提升检测精度。

3)通过多组的对比实验,进一步验证了本文算法的有效性。相较于 Faster R-CNN、Mask R-CNN、改进 Mask R-CNN、YOLOv7、YOLOv8 等算法,本文所提算法在轨道扣件状态检测任务中各方面指标均有所提高,具有更好的泛化性。更适合部署于复杂交通场景进行轨道扣件状态检测。

本文算法在实验过程中并未形成完整的分类系统,在后续的工作中会尝试把实验的所有过程开发成完整的系统。

参考文献

- [1] 王建柱,李清勇,张靖,等. 轨道病害视觉检测:背景方法与趋势[J]. 中国图象图形学报. 2021, 26(2): 287-296.
WANG J ZH, LI Q Y, ZHANG J, et al. Visual inspection of rail defects: Background, methodologies, and trends[J]. Journal of Image and Graphics, 2021, 26(2): 287-296.
- [2] 周文静,刘冲,韦佳宏,等. 基于小波包能量谱的钢轨扣件松脱检测研究[J]. 传感器与微系统, 2017, 36(5): 40-43.
ZHOU W J, LIU CH, WEI J H, et al. Research on rail fastener loose detection based on wavelet packet energy spectrum [J]. Sensors and Microsystems, 2017, 36(5): 40-43.
- [3] 余喆琦,朱洪涛,王志勇. 激光测量在高速铁路扣件扣压力检测的适用性研究[J]. 铁道标准设计, 2016, 60(12): 41-45.
YU Z Q, ZHU H T, WANG ZH Y. Study on the applicability of laser measurement in the pressure detection of fastener in high-speed railway[J]. Railway Standard Design, 2016, 60(12): 41-45.
- [4] FENG H, JIANG Z, XIE F, et al. Automatic fastener classification and defect detection in visionbased railway inspection systems[J]. IEEE Transactions on Instrumentation and Measurement, 2014, 63(4):

- 877-888.
- [5] 侯云,范宏,熊鹰,等. 基于代价敏感卷积神经网络的扣件缺陷检测算法[J]. 中国铁道科学, 2021, 42(1): 26-31.
HOU Y, FAN H, XIONG Y, et al. Flaw detection algorithm of fastener based on cost sensitive convolutional neural network[J]. Chinese Railway Science, 2021, 42(1): 26-31.
- [6] 刘甲甲,李柏林,罗建桥,等. 融合 PHOG 和 MSLBP 特征的铁路扣件检测算法[J]. 西南交通大学学报, 2015, 50(2): 256-263.
LIU J J, LI B L, LUO J Q. Railway fastener detection algorithm integrating PHOG and MSLBP features[J]. Journal of Southwest Jiaotong University, 2015, 50(2): 256-263.
- [7] GIBERT X, PATEL V M, CHELLAPPA R. Deep multitask learning for railway track inspection[C]. IEEE Transactions on Intelligent Transportation Systems. Piscataway. IEEE, 2017: 153-164.
- [8] GIBERT X, PATEL V M, CHELLAPPA R. Sequential score adaptation with extreme value theory for robust railway track inspection[C]. IEEE International Conference on Computer Vision Workshop. IEEE, 2015: 131-138.
- [9] 孙睿,陈兴杰,李立明,等. 基于流行排序的轨道扣件定位方法研究[J]. 铁道科学与工程学报, 2020, 17(2): 288-296.
SUN R, CHEN X J, LI L M, et al. Research on positioning method of track fastener based on popular ordering[J]. Journal of Railway Science and Engineering, 2020, 17(2): 288-296.
- [10] 袁文祥,柴晓冬,李立明,等. 基于图像处理的潘得路快速弹条扣件状态识别[J]. 计算机测量与控制, 2019, 27(6): 177-181, 185.
YUAN W X, CHAI X D, LI L M, et al. State recognition of Pantelo fastener based on image processing[J]. Computer Measurement and Control, 2019, 27(6): 177-181, 185.
- [11] 赵建龙,顾桂梅. 基于引导滤波和积分投影算法的轨道扣件定位[J]. 兰州交通大学学报, 2019, 38(3): 31-36.
ZHAO J L, GU G M. Positioning of track fastener based on guided filtering and integral projection algorithm[J]. Journal of Lanzhou Jiaotong University, 2019, 38(3): 31-36.
- [12] 王宇桐,许贵阳,周素霞. 基于垂向定位与灰度投影的地铁扣件定位方法[J]. 铁道建筑, 2019, 59(6): 112-116.
WANG Y T, XU G Y, ZHOU S X. Subway fastener positioning method based on vertical positioning and gray projection [J]. Railway Construction, 2019, 59(6): 112-116.
- [13] 戴鹏,王胜春,杜馨瑜,等. 基于半监督深度学习的无砟轨道扣件缺陷图像识别方法 [J]. 中国铁道科学, 2018, 39(4): 43-49.
DAI P, WANG SH CH, DU X Y, et al. Image recognition method of ballastless track fasteners based on semi-supervised deep learning[J]. Chinese Railway Science, 2018, 39(4): 43-49.
- [14] 刘奇锋,许贵阳,白堂博. 基于 YOLOv3 算法的轨道扣件自动定位与检测[J]. 铁道建筑, 2020, 60(9): 123-127.
LIU Q F, XU G Y, BAI T B. Automatic positioning and detection of track fasteners based on YOLOv3 algorithm[J]. Railway Construction, 2020, 60(9): 123-127.
- [15] 王婧瑶,王红军. 基于 Mask R-CNN 与 SG 滤波的手势识别关键点特征提取方法[J]. 电子测量与仪器学报, 2021, 35(9): 41-48.
WANG J Y, WANG H J. Gesture key point extraction method based on Mask R-CNN and SG filter[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(9): 41-48.
- [16] 白堂博,高嘉琳,杨建伟,等. 基于改进 Mask R-CNN 的铁路扣件状态检测方法研究[J]. 铁道建筑, 2021, 61(6): 4.
BAI T B, GAO J L, YANG J W, et al. Research on state detection method of railway fasteners based on improved Mask R-CNN [J]. Railway Construction, 2021, 61(6): 4.
- [17] 许贵阳,李金洋,白堂博,等. 基于改进 Mask R-CNN 的轨道扣件状态检测方法[J]. 中国铁道科学, 2022, 43(1): 44-51.
XU G Y, LI J Y, BAI T B, et al. State detection method of track fastener based on improved Mask R-CNN [J]. Chinese Railway Science, 2022, 43(1): 44-51.
- [18] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [19] BOCHKOVSKIY A, WNG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint ArXiv: 2004. 10934, 2020.
- [20] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial appli-

- cations[J]. ArXiv Preprint ArXiv: 2209.02976, 2022.
- [21] 石欣, 卢灏, 秦鹏杰, 等. 一种远距离行人小目标检测方法[J]. 仪器仪表学报, 2022, 43(5):136-146.
SHI X, LU H, QIN P J, et al. A long-distance pedestrian small target detection method[J]. Chinese Journal of Scientific Instrument, 2022, 43(5): 136-146.
- [22] 肖黎俊, 潘睿志, 李超, 等. 基于改进YOLOv5s绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.
XIAO C J, PAN R ZH, LI CH, et al. Research on defect detection technology based on improved YOLOv5s insulator[J]. Electron Measure Technology, 2022, 45(24):137-144.
- [23] 钱承山, 沈有为, 孙宁, 等. 改进YOLOv5s的遥感图像检测研究[J]. 国外电子测量技术, 2022, 41(11): 57-66.
QIAN CH SH, SHEN Y W, SUN N, et al. Improved YOLOv5s remote sensing image detection research[J]. Foreign Electronic Measurement Technology, 2022, 41(11): 57-66.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. ArXiv Preprint ArXiv: 2207.02696, 2022.
- [25] REIS D, KUPEC J, HONG J, et al. Real-time flying object detection with YOLOv8[J]. ArXiv Preprint ArXiv: 2305.09972, 2023.
- [26] SUNKARA R, LUO T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects[J]. ArXiv Preprint ArXiv: 2208.03641, 2022.
- [27] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 764-773.
- [28] LI J, WEN Y, HE L. SCConv: Spatial and channel reconstruction convolution for feature redundancy [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6153-6162.
- [29] WU Y, HE K. Group normalization[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [30] HE K M, GKIOXARI G, DOLLAR P, et al. Mask RCNN [C]. IEEE International Conference on Computer Vision. IEEE, 2017: 2961-2969.

作者简介

范华琦, 硕士研究生, 主要研究方向为目标检测。

E-mail: a18792989791@163.com

杨柳, 博士, 高级工程师, 主要研究方向为移动通信与工程信息化。

E-mail: yangliu@swjtu.edu.cn