

列车轮对踏面缺陷的轻量级检测算法优化设计

侯涛, 任祎坤, 牛宏侠

(兰州交通大学自动化与电气工程学院, 730070, 兰州)

摘要: 针对列车轮对踏面缺陷检测算法计算量、参数量较大, 且高精度与低计算复杂度难以匹配等问题, 提出一种基于 YOLO11n 算法的列车轮对踏面缺陷的轻量级检测算法。构建轻量化的高效幽灵开端(Ghost-Inceptiona)特征提取模块替代 YOLO11n 骨干网络中的特征提取模块, 降低模型的参数量与计算量, 提高轮对踏面缺陷检测的精度和效率; 在颈部网络中采用改进的带小卷积核的跨阶段部分连接瓶颈模块(C3k2_Faster-EMA)作为特征提取模块, 提升算法对不同尺度下目标区域的特征提取融合能力, 进而提高轮对踏面缺陷的检测精度; 优化设计轻量化检测头, 通过加入深度可分离卷积实现算法参数量、计算量的大幅降低; 设计基于交并比(IoU)的改进损失函数(Focaler-PIoU), 加强算法对不同尺度缺陷的适应性, 提高踏面缺陷检测能力。实验结果表明: 在自制列车轮对踏面缺陷的数据集上, 所提改进算法的参数量减少了 18.2%, 计算量下降了 28.6%, 平均精度均值提升了 2.3%。改进算法在提升踏面缺陷检测精度的前提下实现算法的轻量化改进, 在轮对踏面缺陷检测中具有广泛的应用前景。

关键词: 列车轮对踏面; 缺陷检测; YOLO11n 算法; 轻量化改进

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202603021 **文章编号:** 0253-987X(2026)03-0220-13

Optimized Design of Lightweight Detection Algorithm for Train Wheelset Tread Defects

HOU Tao, REN Yikun, NIU Hongxia

(School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: To address the issues of excessive computational load, large number of parameters, and the difficulty in achieving a satisfactory balance with accuracy in existing algorithms for detecting train wheelset tread defects, this study proposes an YOLO11n-based lightweight detection algorithm. A lightweight and efficient Ghost-Inceptiona feature extraction module is constructed to replace the feature extraction module in the backbone network of YOLO11n, reducing the model's parameters and computational load while improving the precision and efficiency of wheelset tread defect detection. In the neck network, a cross-stage partial connection bottleneck structure with small convolution kernels (C3k2-Faster-EMA) is used as the feature extraction module, enhancing the algorithm's ability to extract and fuse features of target regions at different scales, thus increasing detection accuracy. A lightweight detection head is optimally designed through the incorporation of depthwise separable convolutions, significantly decreasing the algorithm's

parameter size and computational load. An improved loss function based on intersection over union (IoU), named Focaler-PIoU, is designed to strengthen the algorithm's adaptability to defects of varying scales and enhance the performance of tread defect detection. The experimental results show that on a self-collected dataset of train wheelset tread defects, the improved YOLO11n algorithm reduces parameters by 18.2%, decreases computation by 28.6%, and improves mean average precision by 2.3%. The improved algorithm achieves lightweight optimization while improving detection accuracy, demonstrating broad application prospects in train wheelset tread defect detection.

Keywords: train wheelset tread; defect detection; YOLO11n algorithm; lightweight improvements

列车轮对作为连接机车车辆和轨道的关键走行部部件^[1],承担着支撑车体重量、传递牵引力、保持行驶平稳等重要任务,其稳定工作是列车安全舒适运行的核心因素之一。轮对作为走行部的滚动部件,在行驶一定里程后,可能会因制动摩擦、长途行驶等原因形成踏面擦伤;在运行过程中会因为与轨道内异物的碰撞、轨道不规则等原因导致踏面出现坑洞;在行驶过程中会因轨道雨水滞留、异常黏附等致使踏面锈蚀。这些缺陷的发生会导致轮对出现不同程度的局部和环向损伤^[2],严重的缺陷会导致列车转向架系统剧烈振动,极端情况下出现列车脱轨事故,对乘客的安全造成严重威胁^[3]。因此,及时检测出缺陷可以降低维护成本并降低危害的发生。

目前,常用的轮对踏面缺陷检测方法分为人工检测、声学检测、电磁检测与光学检测等。人工方法需要列车回库和落轮处理进行人工辨别与处理^[4],其工作量大、人工成本高且检测速度较慢。声学检测多指超声波探伤方法,其基本原理是利用超声波在材料中的传播速度和衰减特性来检测轮对中的缺陷,该方法发展较早,并且已经足够成熟,但易受环境噪声的干扰。电磁检测方法主要利用电磁感应原理来检测轮对缺陷产生的电流波动信号,分析不同信号的特征进而确定缺陷的发生。Cavuto 等^[5]利用基于超声波激发器的高功率激光源,而设计的激光-超声波检测系统是非接触式技术,但该方法往往包含复杂的噪声,易受环境的影响,并且需要部署专业设备,成本较高。Sun 等^[6]基于经典的压电超声技术,提出了一种超声探头组合阵列方法,以满足列车车轮深裂纹在线动态探测的需求,但该方法只适用于轮对踏面裂纹缺陷检测,检测种类较少。

光学方法主要是采用光学传感器结合相关处理算法进行缺陷检测,该方法能够直观反映缺陷的分布情况。主要分为激光位移方法、热成像法,以及基于计算机视觉的方法。基于计算机视觉的检测方法

主要使用图像处理技术进行自动检测和缺陷分类,该技术可以识别多种踏面缺陷并准确定位和量化。刘二林等^[7]使用传统的图像检测方法,利用灰度-梯度共生矩阵和聚类算法对特征进行提取和分类,实现对踏面缺陷图像的分类检测,但该方法对图像质量要求高、适用性较差。Zhang 等^[8]提出了一种基于多尺度特征融合的改进 YOLOv4 高速列车轮对踏面检测算法,以实现轮对踏面缺陷类别、语义和位置信息的更多获取,但该方法的训练数据较少、检测速度较低,需要进一步优化。Zhang 等^[9]针对轮对踏面噪声干扰大、缺陷形式多样、定位效果差等问题,提出了一种可变形的残差注意力网络,并采用多路径聚合网络特征融合方法加强对各种形式缺陷的特征提取能力,同时构建可变形双分支检测头,进一步优化模型的整体性能,但存在参数量较大、检测速度较低等问题。Xing 等^[10]通过网络结构优化和损失函数的设计,实现了列车轮对踏面 4 种缺陷检测的精度提升。Liu 等^[11]通过构建基于局部推理约束网络的轮对踏面缺陷检测方法,对小样本导致的特征空间提取不足的问题进行了改善。

上述轮对踏面缺陷检测算法针对特定场景做了大量的研究工作,但多数只考虑了精度提升或者轻量化的单一目标,网络结构的轻量化与检测精度难以实现良好的平衡。为解决上述问题,以常见的擦伤(abrasion)、坑洞(hole)及锈蚀(rusting)3种踏面缺陷作为检测目标,提出一种基于 YOLO11n 的改进算法。YOLO11 算法以算法的深度和宽度为标准,由小到大可分为 n、s、m、l 和 x 共 5 个版本,YOLO11n 是 YOLO11 系列中最轻量化的版本。通过对 YOLO11n 算法网络结构的改进优化实现算法参数量与计算量的降低、增强算法的检测能力,有效实现列车轮对踏面缺陷检测算法的轻量化改进。

1 踏面缺陷检测问题描述与解决方法

随着轨旁嵌入式设备部署的需求和人工检测成本的增加,列车轮对踏面缺陷的智能化检测面临两大核心挑战。如何在有限计算资源内实现高精度实时检测,算法轻量化设计成为技术刚需;轮对踏面缺陷因采集视角差异与种类多样性等原因,呈现出空间分布不均匀和尺度差距较大等复杂特征,这要求检测算法具备更优的多尺度特征解析能力,以提升检测精度。然而,现有的踏面缺陷检测算法多数存在检测精度提升与轻量化设计难以协同优化等问题,对此,本研究基于轻量化的基准架构 YOLO11n 进行设计。YOLO11n 算法在 YOLOv8n 算法的基础上,通过改进算法结构,增强了特征提取能力、减少了算法参数量与计算量,在实现轻量化设计的同时

时保持了良好的检测精度。

虽然 YOLO11n 具有良好的检测性能和轻量化的模型,但在面对踏面缺陷尺度变化较大时仍存在特征提取困难等问题。因此,为了增强缺陷的特征提取能力并进一步减少计算资源的消耗,便于现场边缘设备的部署,对原算法进行改进优化。改进的 YOLO11n 算法通过构建幽灵开端(Ghost-Inceptiona)特征提取模块,减少算法的计算量与参数量;通过构建改进的带小卷积核的跨阶段部分连接瓶颈特征提取模块(C3k2_Faster-EMA),有效优化缺陷特征提取过程,更好地融合缺陷的多尺度特征并提升检测精度;设计轻量化检测头结构,实现算法的轻量化设计;通过改进的损失函数对权重进行分配,使算法对不同缺陷具有更好的适应性,实现检测精度的提升。改进算法网络结构如图 1 所示。

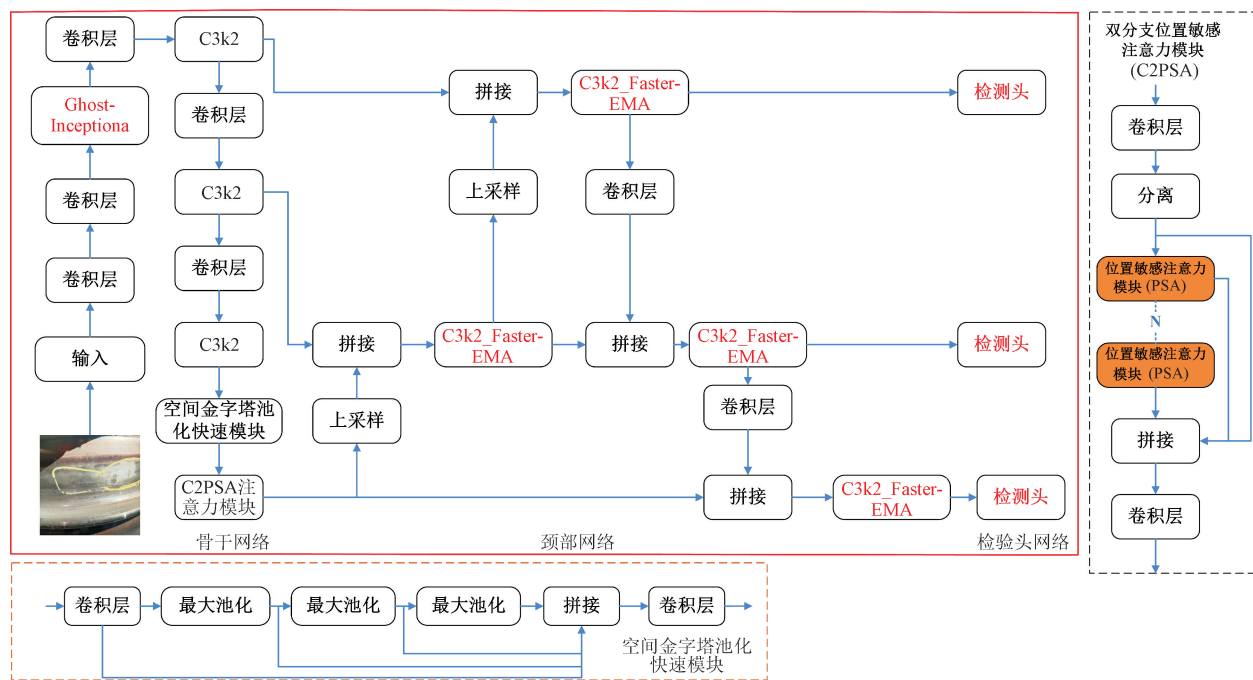


图 1 改进的 YOLO11n 网络结构

Fig. 1 Improved YOLO11n network structure diagram

2 改进 YOLO11 算法的轻量化设计

2.1 构建 Ghost-Inceptiona 特征提取模块

YOLO11n 采用了带小卷积核的跨阶段部分连接瓶颈模块(C3k2)作为骨干网络中特征提取模块,该模块主要通过标准卷积操作进行特征提取,其在一般情况下表现良好,但标准卷积使用参数较多,在处理大批量高分辨率的图像时会造成较大的参数计算量,可能导致模型在资源受限的环境中难以高效

运行。在列车轮对踏面检测场景中,算法模型常部署于移动端且对计算量要求较小,因此 C3k2 结构难以满足相应要求,在此设计 Ghost-Inceptiona 轻量化特征提取模块。

Ghost-Inceptiona 主要结合了 InceptionV2a^[12]和 Ghost 卷积^[13],其网络结构如图 2 所示。通过使用 Ghost 卷积替换 InceptionV2a 中标准卷积,使得算法能够在保持性能的同时降低计算复杂度,以实现计算效率的提高和参数量的减少,更有利于移动端部署。

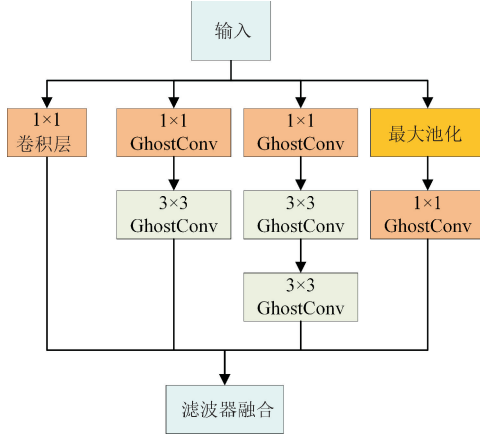


图2 Ghost-Inceptiona 网络结构

Fig. 2 Network structure diagram of Ghost-Inceptiona

InceptionV2a 通过在各个层级中使用大小不同的卷积核来提取多尺度特征。这种多尺度特征提取结构使得模型能够捕捉到不同尺寸的信息,进而实现模型对复杂场景理解能力的优化。同时该结构在前代的基础上采取将 5×5 标准卷积分解为两个 3×3 标准卷积。可以通过感受野计算公式得出,该设计在理论上保持相同感受野的同时显著减少了计算量和参数量,从而提高了模型的效率。

InceptionV2a 广泛使用 1×1 卷积层,不仅实现了网络深度的拓宽,还实现了计算复杂度的大幅下降,同时保持了特征信息提取的丰富性。通过结合多个并行的卷积路径和池化层,InceptionV2a 模块能够训练到多元化的特征,提高了算法的泛化能力,能够在一定程度上降低过拟合的风险。

GhostConv 模块网络结构如图3所示,通过生成更多的特征图来减少计算量,相比于标准卷积,能够在保持性能的同时显著降低计算资源的消耗。给定一个输入 $F \in \mathbf{R}^{c \times w \times h}$,其中 c 为通道数, w 和 h 分别为输入特征图的宽度和高度,经过 $n \times k \times k$ 的卷积核得到输出特征图 $F' \in \mathbf{R}^{n \times W \times H}$, n 为卷积核数, W 和 H 分别为输出特征图的宽度与高度, $k \times k$ 为卷积核大小。标准卷积的参数量为 nck^2 ,计算量为 $HWnck^2$ 。GhostConv 的参数量为 $nck^2/s + (s-1)d^2(n/s)$,计算量为 $HWnck^2/s + (s-1)HWd^2/s$ 。设标准卷积和 GhostConv 模块的计算量比值为 r_1 ,参数量比值为 r_2 ,其计算公式如下

$$r_1 = nHWck^2 / \left(\frac{n}{s}HWck^2 + (s-1) \frac{n}{s}HWd^2 \right) \approx \frac{sc}{s+c-1} \approx s \quad (1)$$

$$r_2 = nck^2 / \left(\frac{n}{s}ck^2 + (s-1) \frac{n}{s}d^2 \right) \approx \frac{sc}{s+c-1} \approx s \quad (2)$$

式中: n/s 为第1次变换时的输出通道数; s 为线性变换次数。因为采用恒等映射所以减少了一次线性变换计算,故式中为 $s-1$ 而不是 s ,其中 $d \times d$ 为线性变换的卷积核大小^[14]。

由式(1)和式(2)可得,线性变换次数 s 直接影响计算量和参数量之比,通过合理调整 s 实现本特征生成量的减少、GhostConv 操作次数的增加,最终实现算法的计算量、参数量的降低。GhostConv 模块具有降低计算负载、高效利用资源、易于适应的特点,替换 InceptionV2 的标准卷积模块,构建 Ghost-Inceptiona 模块,极大地减少了参数量与计算量。

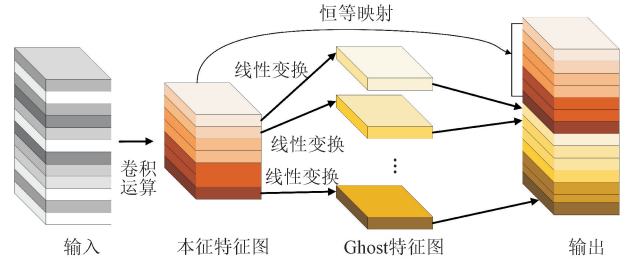


图3 GhostConv 模块网络结构

Fig. 3 Network structure diagram of GhostConv module

2.2 构建轻量化检测头

YOLO11 的检测头在 YOLOv8 的解耦头基础上进行改进,YOLOv8 的解耦头通过独立处理分类和回归定位任务,提升模型的检测性能。尽管原 YOLOv8 的解耦头提高了模型的检测能力,但仍然存在一些不足之处。首先,这种设计大大增加了网络的参数量,导致推理速度下降;其次,由于特征表示的互补性,解耦头无法充分利用骨干网络提取的高级特征^[15],因此在处理复杂背景和多任务的情况下,检测精度和速度会降低。YOLO11 将分类检测头增加了两个深度卷积(DWConv),实现了参数量和计算量的减少,但在定位检测头中仅采用两个标准卷积,使得算法的参数量与计算量仍然较大。

在多数检测中,分类任务是为每个检测到的对象分配相应的类别标签。在该任务中,模型需要判断图像中每个边界框内的对象属于何种类别。定位任务的目标是预测每个对象的边界框,通常用边界框的中心点坐标以及宽度和高度的值进行确定,定位任务对坐标回归和精度两者都有要求。模型依靠坐标的回归实现准确预测边界框的位置和大小,依

靠识别精度实现边界框与实际对象的匹配,微小的偏差可能导致边界框与实际对象不匹配,影响检测效果。这些需求导致列车轮对踏面缺陷检测中,定位任务通常比简单的分类任务更具挑战性。

在列车轮对踏面缺陷检测中,缺陷数据被分为 3 类,分类任务相对简单,而定位任务较为复杂。为了进一步实现轻量化检测,本文参照一种轻量级非对称性检测头^[16],在 YOLO11 的检测头基础之上,设计轻量化检测头。图 4 是改进的检测头网络结构。在该网络中,采用 3×3 深度可分离卷积(DSConv)^[17]来替代每个分支中的 3×3 标准卷积。与标准卷积相比,深度可分离卷积的优势在于通过将卷积操作分解为深度卷积和逐点卷积,大大减少了参数量。

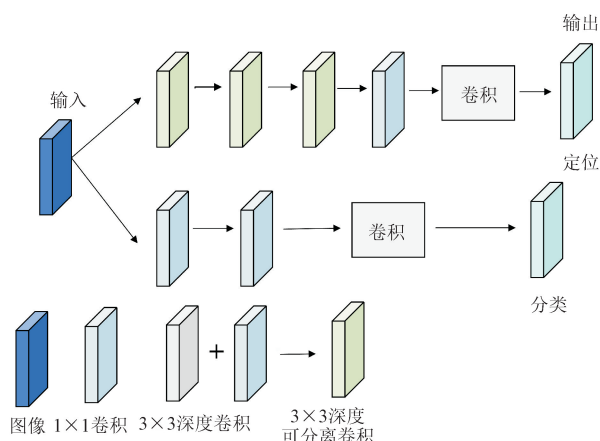


图 4 轻量化检测头网络结构

Fig. 4 Network structure diagram of lightweight detection head

输入特征图尺寸为 $h \times w \times M$, 标准卷积核大小为 $k \times k$, 输入通道数为 M , 输出通道数为 N , 输出的特征图尺寸为 $H \times W \times N$, 则通过标准卷积处理后, 可以得出标准卷积相应的参数量 Q_1 和计算量 G_1 , 其计算式如下

$$Q_1 = k^2 MN \quad (3)$$

$$G_1 = k^2 MNHW \quad (4)$$

深度卷积的卷积核尺寸为 $k \times k$, 所以参数量为 $k^2 M$; 在计算量方面, 每个深度卷积都要做 $H \times W$ 次运算, 所以计算量为 $k^2 MHW$ 。深度卷积的参数量 Q_2 、计算量 G_2 的表达式为

$$Q_2 = k^2 M \quad (5)$$

$$G_2 = k^2 MHW \quad (6)$$

深度卷积仅保留空间滤波, 无通道混合。深度可分离卷积不仅提取了空间特征, 还通过逐点卷积

融合了通道信息, 增强了特征的表达能力。逐点卷积的卷积核尺寸为 1×1 , 所以参数量为 MN , 在计算量方面每个都要做 HW 次运算, 所以计算量为 $MNHW$ 。因此, 深度可分离卷积的参数量 Q_3 、计算量 G_3 表达式为

$$Q_3 = k^2 M + MN \quad (7)$$

$$G_3 = k^2 MHW + MNWH \quad (8)$$

深度可分离卷积与标准卷积的参数量之比 Q' 、计算量之比 G' 分别为

$$Q' = \frac{k^2 M + MN}{k^2 MN} = \frac{1}{N} + \frac{1}{k^2} \quad (9)$$

$$G' = \frac{k^2 MHW + MNWH}{k^2 MNHW} = \frac{1}{N} + \frac{1}{k^2} \quad (10)$$

深度可分离卷积通过分解空间滤波和通道混合, 显著降低了检测任务的参数量和计算量, 同时保持较好的特征提取能力。在定位任务采用多个深度可分离卷积的堆叠使得改进的检测头能够在不同的空间尺度上提取丰富的特征, 提高了整体的检测精度。逐点卷积通过在不同通道之间应用 1×1 卷积核来混合通道间的信息, 不改变特征图的空间维度, 只改变通道的数量。这种方法可以进一步提高检测效果。在分类任务前使用简单的卷积, 实现结构的简单化, 提升分类速度。采用改进的轻量化检测头结构取代 YOLO11n 网络中的解耦头后, 不仅提升了检测精度, 还实现了算法模型的轻量化。这使得改进的检测头成为在性能和效率之间取得良好平衡的选择, 适合于现代目标检测任务的需求。

3 提升检测精度的优化设计

轮对踏面缺陷通常存在大小尺度不一、分布不均匀的情况, 如擦伤和锈蚀较大, 而坑洞通常较小, 同时轻量化的结构设计会使算法提取能力受限。因此为解决该问题, 设计新的颈部特征提取网络并优化损失函数, 实现检测精度的提升和轻量化的平衡。

3.1 优化颈部特征提取网络

YOLO11 的颈部网络采用 C3k2 模块进行特征的提取, 该模块使用了标准卷积, 计算量与参数量较大且特征提取能力有限。为了解决该问题, 引入 FasterNet Block 模块^[18]和高效多尺度注意力模块(EMA)^[19]注意力机制, 并以两者构建新特征提取网络。

FasterNet Block 模块主要通过设计一个部分卷积(PCConv)对标准卷积操作进行替换。为了连续的内存访问, 只采用第一个或最后一个连续 c_p 个通

道作为计算整个特征映射的代表。在不丧失一般性的情况下,只考虑输入和输出特征映射具有相同数量的通道的情況^[20],其中部分卷积只对特征图的 c_p 个通道进行规则卷积,其余通道保持不变。部分卷积的计算量 G_4 和内存访问量 Z 可以根据下式计算

$$G_4 = h w k^2 c_p^2 \quad (11)$$

$$Z = 2 h w c_p + k^2 c_p^2 \approx 2 h w c_p \quad (12)$$

式中: c_p 表示连续选取的通道数。当卷积操作选取常用的 $1/4$ 通道数时,可以得出部分卷积的计算量仅为标准卷积的 $1/16$,内存访问量为标准卷积的 $1/4$ 。采取FasterNet能够在复杂的踏面缺陷场景下提升算法的目标检测效果。与其他网络相比,FasterNet在发挥提取空间特征方面的优势的同时实现更精炼的结构和更低的复杂度,更好地利用计算资源。

3.1.1 构建Faster-EMA网络

通过借鉴FasterNet轻量化网络与EMA注意力机制的思想,构建的Faster-EMA的网络结构如图5所示,实现在不同尺度的缺陷检测中提升检测性能。图5中输入输出特征图以宽、高为 w, h 的长方体表征,其由宽、高为 w, h 且厚度为1的单通道特征图堆叠形成,通道数对应图5中长方体的厚度尺寸,选取 c_p 个通道即对应选取厚度为 c_p 的长方体尺寸。

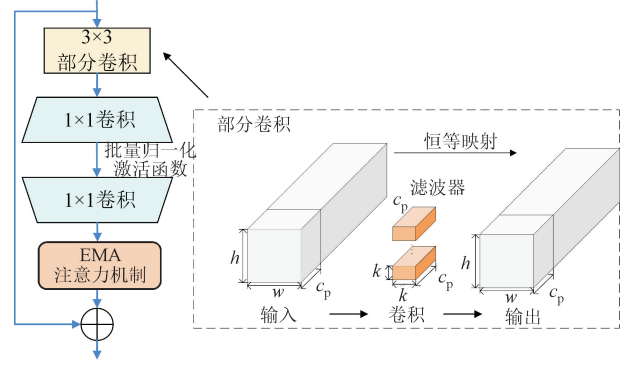


图5 Faster-EMA网络结构

Fig. 5 Network structure diagram of Faster-EMA network

3.1.2 构建C3k2_Faster-EMA模块

C3k2_Faster-EMA通过继承C3k2和Faster-EMA模块化设计,提供了一种灵活的方式来组合不同的卷积模块。通过设定参数 f_{C3k} ,可以动态决定使用哪种类型的模块,在C3k2_Faster-EMA中,若设定 f_{C3k} 为真,则使用C3k_Faster-EMA代替原来的C3k,以实现更快的计算。若设定 f_{C3k} 为假,则使用Faster-EMA代替原来的瓶颈层以提高性能,这使得模型在不同场景下具有更好的适应性。通过引入Faster-EMA相关的模块,C3k2_Faster-EMA在性能上进行了优化,实现了更好的检测效果,同时减少了参数量、计算量。C3k2_Faster-EMA模块网络结构如图6所示。

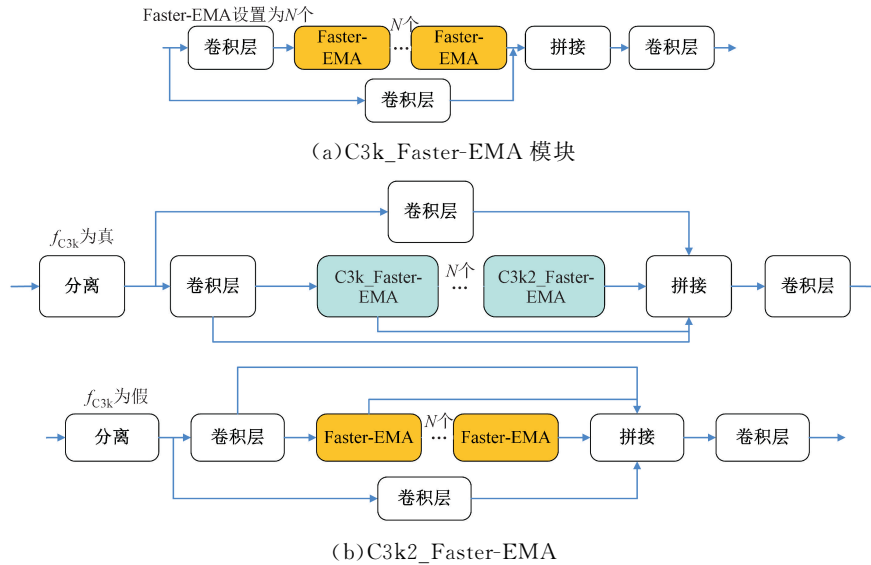


图6 C3k2_Faster-EMA模块网络结构

Fig. 6 Network structure diagram of C3k2_Faster-EMA module

3.2 优化损失函数

检测算法通过对每个目标的边界框坐标和置信度精度进行预测并结合分类信息实现目标检测任

务,因此边界框损失函数的作用至关重要,它通过优化边界框的预测,促进模型的收敛、处理多尺度问题,并与分类损失结合,提高整体检测精度,能为目

标检测模型带来显著的性能提升。通过精确的边界框预测, YOLO算法能够在目标检测任务中表现出色, 满足实时检测的需求。YOLO11采用了完整交并比(CIoU)作为边界框回归损失函数^[21], 其相关公式为

$$L_{\text{CIoU}} = 1 - I_{\text{IoU}} + \frac{\rho^2(B, A)}{o^2} + \alpha v \quad (13)$$

$$I_{\text{IoU}} = (S \cap T) / (S \cup T) \quad (14)$$

$$\alpha = \frac{v}{(1 - I_{\text{IoU}}) + v} \quad (15)$$

$$v = \frac{4}{\pi^2} \left(\arctan\left(\frac{b_1}{b_2}\right) - \arctan\left(\frac{a_1}{a_2}\right) \right)^2 \quad (16)$$

式中: L_{CIoU} 为 CIoU 损失函数的损失值; I_{IoU} 为预测框和真实框的交并比; A 点为算法输出的预测框 S 的中心点, B 点为人工标注的真实框 T 的中心点; ρ 为两个中心点的欧氏距离; o 表示能够同时包含预测框和真实框的最小闭包区域的对角线距离; α 为平衡参数, 通过 I_{IoU} 来赋予优先级, 当预测框和真实框 I_{IoU} 越大时, α 越大; v 是衡量预测框和真实框的宽高比一致性的惩罚项; b_2 、 b_1 表示真实框的高和宽, a_2 和 a_1 表示预测框的高和宽, l_h 和 l_w 表示最小闭包区域的高和宽, 相关参数如图 7 所示。

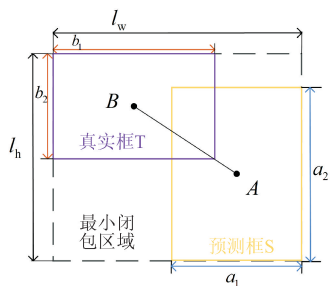


图7 CIoU 相关参数示意图

Fig. 7 Schematic diagram of parameters in CIoU

CIoU 仍存在不足之处, 该损失函数通过边界框相互的几何定位关系来提高性能, 但并未考虑难易缺陷样本分布会对边界框回归的影响。针对该问题, 常规训练过程中采用对困难样本进行采样和重新加权进行缓解, 但效果并不明显。基于 IoU(交并比)的改进损失函数(Focaler-PIoU)^[22], 通过线性区间映射来重构 IoU 损失, 实现对不同的回归样本进行聚焦, 来提高不同检测任务中算法的性能, 实现专注不同的回归样本, 再利用不同样本在边界框回归中的分布对回归结果的影响进行改进。同时, 该损失函数通过其特有的方法来弥补现有边界框回归

方法的不足, 从而在不同的检测任务中进一步提高检测性能。

为了能够在检测任务中聚焦不同的缺陷样本, 使用线性区间映射的方法来重构损失, 使得边框回归效果得到提升, 相关计算公式为

$$I_{\text{focaler}} = \begin{cases} 0, & I_{\text{IoU}} < d \\ \frac{I_{\text{IoU}} - d}{u - d}, & d \leq I_{\text{IoU}} \leq u \\ 1, & I_{\text{IoU}} > u \end{cases} \quad (17)$$

式中: $[d, u] \in [0, 1]$; I_{focaler} 为重构后的交并比, 可以根据原交并比的值进行调整。当 I_{IoU} 小于下限阈值 d 时, 数值为 0; 当 I_{IoU} 大于上限阈值 u 时, 数值为 1; 而当 I_{IoU} 处于 d 和 u 之间时, 数值是一个根据 I_{IoU} 值线性递增的函数。通过调整 d 和 u 的值, 可以使 I_{focaler} 关注不同难度的回归样本。该设计允许损失函数在一定区间内对 I_{IoU} 值进行操作, 从而能够更专注于预测边界框与真实边界框重叠度中等的样本, 有助于模型更好地学习从中等困难的样本中提取特征, 而不是仅仅专注于最容易或最困难的样本。Focaler-IoU 损失函数的损失 L_{focaler} 定义如下

$$L_{\text{focaler}} = 1 - I_{\text{focaler}} \quad (18)$$

在原 CIoU 中存在两个惩罚项, 反映了中心点距离和纵横比的差异。然而, 它并没有直接捕捉目标框和锚框之间的形状差异, 可能会出现较差的收敛效果。此外, 两个惩罚项均未反映目标框大小的变化。引入的强大的交并比损失函数(Powerful-IoU)引导锚框沿着有效路径回归, 从而比现有的基于 IoU 的损失更快地收敛, 并将分类损失和回归损失相结合以提高模型的整体性能。当前多数损失函数都采用锚框和目标框的最小外部边界框的对角线长度或面积等维度信息作为损失因子的分母^[23]。相反, Powerful-IoU 损失函数仅使用目标框的边缘长度作为惩罚因子的分母, 并且与锚框和目标框的最小边界区域的大小无关, 在训练过程中锚框的大小变化不会影响惩罚因子, 表现出对目标大小的适应性, 相关参数如图 8 所示。

为了解决锚框扩大的问题和现有基于 IoU 的损失函数的局限性, 根据 Powerful-IoU 提出一个适应目标大小的惩罚因子 p , 定义如下

$$p = \left(\frac{d_{w1}}{b_3} + \frac{d_{w2}}{b_3} + \frac{d_{h1}}{b_4} + \frac{d_{h2}}{b_4} \right) / 4 \quad (19)$$

式中: d_{w1} 、 d_{w2} 和 d_{h1} 、 d_{h2} 是锚框和目标框的相应边之间距离的绝对值, 并且 b_3 和 b_4 表示目标框的宽度

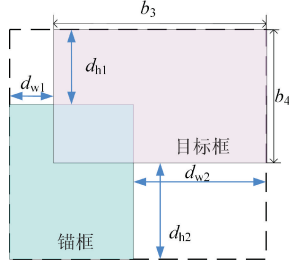


图8 Powerful-IoU 相关参数示意图

Fig. 8 Schematic diagram of parameters in Powerful-IoU

和高度。惩罚系数 $f(x)$ 的计算式如下

$$f(x) = 1 - e^{-x^2} \quad (20)$$

将惩罚因子 p 代入式(20)中的变量 x 可得出惩罚系数 $f(p)$

$$I_{\text{PIoU}} = I_{\text{IoU}} - f(p), \quad -1 \leq I_{\text{PIoU}} \leq 1 \quad (21)$$

进一步应用于下式计算出相应损失

$$L_{\text{IoU}} = 1 - I_{\text{IoU}} \quad (22)$$

$$L_{\text{PIoU}} = 1 - I_{\text{PIoU}} = L_{\text{IoU}} + f(p), \quad 0 \leq L_{\text{PIoU}} \leq 2 \quad (23)$$

式中: I_{PIoU} 为新构建的交并比; L_{PIoU} 为 PIoU 损失函数的损失值。将 Focaler-IoU 损失函数二次应用于基于 Powerful-IoU 的边界框回归损失函数,可得到性能更优的 Focaler-PIoU 损失函数,其损失定义为

$$L_{\text{Focaler-PIoU}} = L_{\text{PIoU}} + I_{\text{IoU}} - I_{\text{focaler}} \quad (24)$$

经实验验证, Focaler-IoU 结合 Powerful-IoU 即可拥有更加优秀的定位识别能力,实现比原 CIoU 更好的检测效果。

4 实验验证

4.1 实验环境与过程

为了验证和评估本文方法的有效性,针对不同缺陷进行了多组实验,包括自然状态下的擦伤、坑洞、锈蚀,以及经过色彩变换、添加噪声、裁剪缩放和翻转后的增强图像,图像主要来源于网络收集,少部分来源于国铁集团某机务段,共 7 518 张缺陷图像。将 5 697 张作为训练集,693 张作为验证集,1 128 张作为测试集。通过图像增强模拟不同检测环境下的复杂场景,测试本文算法的性能,以证明网络的鲁棒性和可靠性。实验运行的硬件为:显卡型号为 Nvidia GeForce RTX3080 $\times$ 2。模型框架是 Pytorch 2.0.0, python 3.8, CUDA 11.8。在本实验中,输入图像的尺寸是 640×640 像素,训练批次

大小为 32,初始学习率设为 0.01,权重的衰减设为 0.000 5,训练轮次为 300,选随机梯度下降作为本网络的优化器。

列车轮对踏面缺陷检测的具体过程为:利用收集的缺陷图像进行训练,训练过程采用了改进的 Ghost-InceptionA 特征提取模块、C3k2_Faster-EMA 特征提取模块、轻量化检测头、Focaler-PIoU 损失函数,并将训练好的改进算法经验证集验证,再在测试集上进行推理,实现轮对踏面缺陷的精确检测。

4.2 性能评价指标

本文采用精确率 P 、召回率 R 、平均精度 P_{AP} 、平均精度均值 P_{mAP} 、交并比阈值取 0.5 时所有检测目标类别的平均精度均值 $P_{\text{mAP}@0.5}$ 、算法模型的参数量、计算量(F_{FLOPs})、帧率作为评价算法模型检测性能的指标,部分计算式如下

$$P = \frac{T_p}{T_p + F_p} \quad (25)$$

$$R = \frac{T_p}{T_p + F_N} \quad (26)$$

$$P_{\text{mAP}} = \left(\sum_{j=1}^{N_y} P_{\text{AP}_j} \right) / N_y = \left(\sum_{j=1}^{N_y} \int_0^1 P_j(R_j) dR_j \right) / N_y \quad (27)$$

式中: P_{AP_j} 为第 j 个检测目标的平均精度; T_p 为模型正确识别且种类为正的个数; F_p 为模型错误识别且种类为正的个数; F_N 为模型错误识别且种类为负的个数; N_y 为缺陷总种类数。

4.3 性能评价与对比实验

4.3.1 不同骨干网络改进下的性能评价

采用不同特征提取模块的性能评价结果见表 1。Ghost-InceptionA 相较于 Ghost-InceptionA 使用了平均池化替代最大池化; Ghost-Inceptionv1 则在 Inceptionv1 基础上融合了 GhostConv; Ghost-InceptionA 相较于以上特征提取模块可以在低参数量、计算量的同时拥有更优的检测精度。相较于 Ghost-InceptionA,增加卷积门控线性单元(Additive-CGLU)、幽灵动态卷积(Ghost-Dynamic)特征提取模块,虽然使算法的参数量、计算量分别下降了 2.0%、12.9%与 1.7%、8.5%,但检测精度下降了 1.8%与 4.6%,用少量计算需求的减少导致检测精度的大幅下降,不利于实时检测的准确性。因此对比表 1 内网络模块, Ghost-InceptionA 实现了各项指标的优化与平衡。

表 1 不同骨干网络的性能评价结果

Table 1 Performance evaluation results of different backbone networks

算法	$P_{mAP@0.5}$ / %	参数量 / 10^6	F_{LOPs} / 10^9	帧率 / (帧 · s ⁻¹)
YOLO11n	85.4	2.58	6.3	74
+Ghost-InceptionA	84.9	2.56	5.9	76
+Ghost-Dynamic	80.9	2.23	5.4	40
+Additive-CGLU	83.7	2.51	5.8	63
+Ghost-Inceptionv1	84.3	2.56	5.9	83
+Ghost-Inceptiona	85.5	2.56	5.9	76

注:实验以 YOLO11n 为基准算法,“+”表示添加改进模块。

4.3.2 不同 Neck 网络改进的性能评价

在 Ghost-Inceptiona 基础上,采用不同特征提取、融合模块的性能评价结果见表 2。可以看出,添加 C3k2_Faster-EMA 模块后, $P_{mAP@0.5}$ 精度相较于原模型精度提升了 2.1%,同时参数量、计算量也有所下降。对比 FreqFFPN 特征拼接模块^[24]虽然计算量略大,但有更优的检测精度和更少的参数量;对比 C3k2_ContextGuided^[25]和 C3k2_Faster-CGLU 特征提取模块,算法的参数量和计算量略有上升,但检测精度大幅提升,检测效果更好。

表 2 不同 Neck 网络改进的性能评价结果

Table 2 Performance evaluation results of different Neck network improvements

算法	$P_{mAP@0.5}$ / %	参数量 / 10^6	F_{LOPs} / 10^9	帧率 / (帧 · s ⁻¹)
+C3k2_Faster-EMA	87.5	2.41	5.7	61
+FreqFFPN	84.9	2.45	5.6	67
+C3k2_ContextGuided	84.8	2.16	5.1	62
+C3k2_Faster-CGLU	85.3	2.38	5.5	92

注:实验以表 1 第 6 行改进算法为基准,“+”表示添加相应模块。

4.3.3 不同检测头的性能评价

采用不同检测头的性能评价结果见表 3。可以看出,改进的检测头虽然在前两个改进方案之上牺牲了少量精度,但大幅降低了算法的参数量与计算量,较改进前提升了检测的帧率,同时相较于表内其他检测头实现了轻量化和精度的良好平衡,证明该检测头的检测效果更优,实现了良好的轻量化改进。

表 3 不同检测头的性能评价结果

Table 3 Performance evaluation results of different detection heads

算法	$P_{mAP@0.5}$ / %	参数量 / 10^6	F_{LOPs} / 10^9	帧率 / (帧 · s ⁻¹)
+轻量化检测头	86.4	2.11	4.5	69
+Detect_Efficient	83.7	2.14	4.5	58
+Detect_MultiSEAM	86.3	4.42	5.4	46
+Detect_SEAM	85.7	2.06	4.9	59
+Detect_RSCD	82.8	2.00	4.7	62
+Detect_LSDECD	84.8	2.09	5.4	71

注:实验以表 2 第 1 行改进算法为基准,“+”表示添加检测头。

4.3.4 不同损失函数的缺陷检测性能评价

在前 3 部分改进的基础上,采用不同损失函数的性能评价结果见表 4。由表 3 可以看出,相较于改进前,改进的 Focaler-PIoU 损失函数在不改变算法的参数量与计算量的前提下,实现了检测精度的提升。同时,相较其他损失函数实现了最高的检测精度,证明该损失函数的改进效果更优。

表 4 不同损失函数的性能评价结果

Table 4 Performance evaluation results of different loss functions

算法	$P_{mAP@0.5}$ / %	参数量 / 10^6	F_{LOPs} / 10^9	帧率 / (帧 · s ⁻¹)
+Focaler-PIoU	87.7	2.11	4.5	72
+ShapeIoU	85.4	2.11	4.5	72
+SIoU	83.4	2.11	4.5	67
+Focaler-CIoU	85.8	2.11	4.5	67
+Inner-CIoU	84.4	2.11	4.5	69
+Inner-PIoU	86.1	2.11	4.5	74
+Focaler-WIoU	85.7	2.11	4.5	68
+Focaler-EIoU	84.8	2.11	4.5	70

注:实验以表 3 第 1 行改进算法为基准,“+”表示添加损失函数。

4.3.5 消融实验

为了验证和评估各模块的功能性,本文设计了消融实验。以 YOLO11n 为基准,将改进的 YOLO11n 中的 Ghost-Inceptiona 模块、C3k2_Faster-EMA 特征提取模块、轻量化检测头、Focaler-PIoU 损失函数分别进行对比实验,针对不同的检测模型实验得到的性能见表 5。

表5 不同模块的消融实验性能评价结果

Table 5 Performance evaluation results of ablation experiments of different modules

组号	Ghost-Inceptiona	C3k2_Faster-EMA	检测头	Focaler-PIoU	$P_{AP}/\%$			$P_{mAP@0.5}/\%$	参数量/ 10^6	$F_{LOPs}/10^9$	帧率/(帧 $\cdot s^{-1}$)
					擦伤	坑洞	锈蚀				
1					94.7	83.9	77.6	85.4	2.58	6.3	74
2	✓				96.7	81.7	78.2	85.5	2.56	5.9	76
3		✓			98.1	82.6	77.5	86.1	2.43	6.1	60
4			✓		97.1	84.4	75.9	85.8	2.28	5.2	63
5				✓	98.2	82.3	79.0	86.5	2.58	6.3	69
6	✓	✓			97.1	84.8	80.4	87.5	2.41	5.7	61
7	✓	✓	✓		97.3	83.9	77.9	86.4	2.11	4.5	69
8	✓	✓	✓	✓	96.6	86.0	80.5	87.7	2.11	4.5	72

注:实验以原始 YOLO11n 为基准,“✓”表示改进模块的添加。

由表5可以看出,构建的 Ghost-Inceptiona 模块的设计极大地减少了网络计算量,同时因为网络的深度拓宽,使得模型更聚焦关键特征,因此缺陷面积较大的擦伤和锈蚀的检测精度得到小幅提升;在 Ghost-Inceptiona 基础上,构建 C3k2_Faster-EMA 模块提升了算法对轮对踏面缺陷不同尺度的特征提取能力,同时减少了部分参数量与计算量;加入构建的轻量化检测头,实现了计算量与参数量的大幅下降,并提升了检测的帧率;利用 Focaler-PIoU 损失函数,在不改变网络参数的基础上,使得网络对缺陷细节提取能力有所提高,同时优化了检测效果。

通过4种方法的结合,实现了算法轻量化设计,提升了检测的精度。为了进一步证明改进的有效性,在此给出 YOLO11n 算法和改进的 YOLO11n 算法的精确率和召回率曲线图的对比。通过阈值的变化来计算得到不同阈值下各缺陷的精确率和召回率并绘制成一条曲线,并对3种缺陷的精确率和召回率取平均得到3种缺陷总体的变化情况。通过式(27)计算对应阈值的平均精度。算法改进前、后精确率和召回率曲线的对比如图9所示,图9中各项精度均为交并比阈值取0.5时所得。由图9可以看出,改进算法具有更好的检测效果。

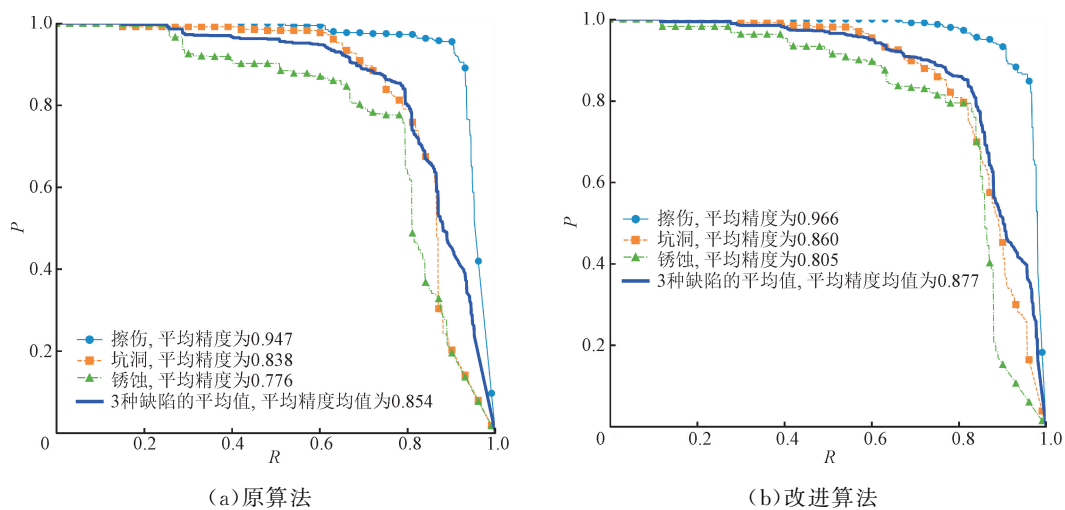


图9 算法改进前、后精确率和召回率曲线的对比

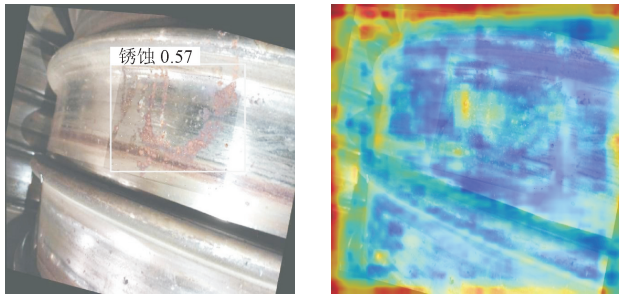
Fig. 9 Comparison of PR curves before and after algorithm improvement

4.3.6 踏面缺陷检测的可视化结果

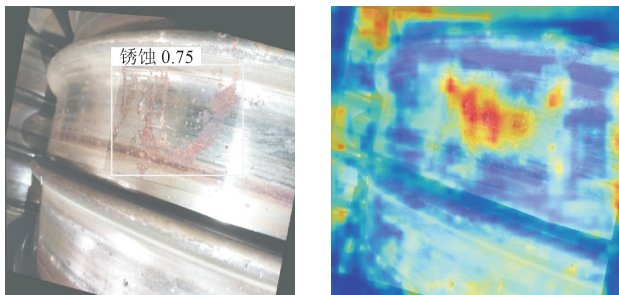
图10给出了改进 YOLO11n 算法与 YOLO11n

算法的可视化检测结果与相应的热力图,通过观察热力图的亮度、深浅与分布情况,可以直观地反映算

法对不同区域的关注程度。可以明显看出,相较于原算法,改进算法对轮对踏面缺陷定位更加准确,具有更高的热力值,因此改进算法的检测效果更优。



(a) YOLO11n 算法可视化结果 (b) YOLO11n 检测结果热力图



(c) 改进算法可视化结果 (d) 改进算法检测结果热力图

图 10 算法改进前后可视化结果对比

Fig. 10 Comparison of visualization results before and after algorithm improvement

4.3.7 不同算法处理结果对比

采用不同目标检测算法性能评价结果见表 6。由表 6 可以看出,本文算法相较于其他单阶段 YOLO 算法的轻量化版本在 $P_{mAP@0.5}$ 上表现最优。与 YOLOv10n、YOLOv9t^[26]、YOLOv8n、YOLOv7-tiny、YOLOv5n 算法相比,本文算法的参数量、计算量分别减少了 21.8% 和 34%、19.2% 和 57.9%、29.7% 和 44.4%、64.8% 和 65.4%、15.6% 和 36.6%,同时 $P_{mAP@0.5}$ 分别提高了 4.7%、3.1%、2.6%、4.0%、4.6%,在大幅减小算法模型的同时实现了更好的检测精度。同时相较于原 YOLO11n 算法的参数量减少了 18.2%、计算量下降了 28.6%,并提高了 2.3% 的平均精度均值,更适合在移动端部署。对比可知,改进算法的检测效果更优。与同为单阶段的 SSD 算法相比,SSD 精度略高,但改进算法大幅减少了参数量、计算量,实现了更好的算法轻量化设计。与两阶段算法 Faster R-CNN 相比,改进算法的检测性能得到提升。与 RT-DETR-R18 算法^[27]相比,虽然检测帧率下降,但检测精度提升了 3.7%,且参数量、计算量较小。

表 6 不同算法的性能评价结果

Table 6 Performance evaluation results of different algorithms

算法	$P_{mAP@0.5} / \%$	参数量 / 10^6	$F_{LOPs} / 10^9$	帧率 / (帧 · s ⁻¹)
YOLO11n	85.4	2.58	6.3	74
本文	87.7	2.11	4.5	72
YOLOv8n	85.1	3.00	8.1	76
YOLOv5n	83.1	2.50	7.1	67
YOLOv7-tiny	83.7	6.0	13.0	75
YOLOv10n	83.0	2.27	6.9	80
YOLOv9t	84.6	2.62	10.7	73
SSD	88.3	26.30	62.7	70
Faster R-CNN	80.1	137.10	370.2	48
RT-DETR-R18	84.0	19.88	57.1	78

选取表 6 中改进的 YOLO11n 算法、YOLO11n 算法和同为 YOLO 算法小模型且检测精度较高的 YOLOv9t、YOLOv8n 以及参数量、计算量较小的 YOLOv10n 算法进行检测效果对比,结果如图 11 所示。由图可见,原算法在坑洞、擦伤检测时存在重复检测的情况,而其他算法并未出现,且本文改进算法的检测精度最高。在锈蚀检测中,本文改进算法检测精度最高,表现最为优秀,因此本文改进算法的检测效果更好。

5 结 论

针对列车轮对踏面缺陷检测困难、精度与轻量化难以平衡的问题,开展了列车轮对踏面缺陷检测研究。完成了 Ghost-Inceptiona 特征提取模块、C3k2_Faster-EMA 特征提取模块的构建,轻量化检测头、Focaler-PIoU 损失函数的设计,使得改进算法相较于原算法的模型参数量减少了 18.2%,计算量下降了 28.6%,平均精度均值提升了 2.3%,实现了列车轮对踏面缺陷检测算法的轻量化改进与检测精度的提升。验证了所提算法的可行性和有效性,为列车轮对踏面缺陷的轻量级检测算法设计提供了一种可靠且实用性强的解决方案。

本文仅对常见的 3 种踏面缺陷问题进行了研究,而在列车的实际运行过程中,由于列车行驶的环境和采集设备的差异,可能会造成多种缺陷的重叠以及图像的质量损失,所以下一步计划对重叠缺陷以及缺陷图像质量优化进行研究,并进一步将算法部署于移动端进行检测验证。



图 11 不同算法检测结果对比

Fig. 11 Comparison of detection results of different algorithms

参考文献:

- [1] 胡新亮. 基于卷积神经网络的高速列车轮对踏面缺陷诊断方法研究 [D]. 株洲: 湖南工业大学, 2022.
- [2] YE Yunguang, LI Haoqian, WANG Qunsheng, et al. Fault diagnosis of railway wheelsets: a review [J]. Measurement, 2025, 242, Part E: 116169.
- [3] VAKKALAGADDA M R K, VINEESH K P. Causes and failure forms of railway wheels [J]. Materials Today: Proceedings, 2024, 98: 97-101.
- [4] 刘应桃. 轮对踏面检测系统设计与缺陷智能识别技术研究 [D]. 成都: 西南交通大学, 2023.
- [5] CAVUTO A, MARTARELLI M, PANDARESE G, et al. FEM based design of experiment for train wheelset diagnostics by laser ultrasonics [J]. Ultrasonics, 2021, 113: 106368.
- [6] SUN Zhilin, LU Jingui, SUN Jie. Online dynamic detection of hazardous cracks in train wheels by probe combination: based on classical piezoelectric ultrasonic technology [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-10.
- [7] 刘二林, 刘成刚, 姜香菊, 等. 基于灰度-梯度共生矩阵的车轮踏面缺陷聚类分析 [J]. 光电子·激光, 2022, 33(1): 53-60.
- [8] ZHANG Changfan, HU Xinliang, HE Jing, et al. Yolov4 high-speed train wheelset tread defect detection system based on multiscale feature fusion [J]. Journal of Advanced Transportation, 2022, 2022: 1172654.
- [9] ZHANG Changfan, XU Yifu, SHENG Zhenwen, et al. Deformable residual attention network for defect detection of train wheelset tread [J]. The Visual Computer, 2024, 40(3): 1775-1785.
- [10] XING Zongyi, ZHANG Zhenyu, YAO Xiaowen, et al. Rail wheel tread defect detection using improved YOLOv3 [J]. Measurement, 2022, 203: 111959.
- [11] LIU Jianhua, JIANG Shiyi, WANG Zhongmei, et al. Detection of train wheelset tread defects with small samples based on local inference constraint network [J]. Electronics, 2024, 13(11): 2201.
- [12] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 1-9.
- [13] HAN Kai, WANG Yunhe, TIAN Qi, et al. GhostNet: more features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 1577-1586.

- [14] 王楠, 侯涛, 牛宏侠. 多尺度特征融合的铁轨异物入侵检测研究 [J]. 西安交通大学学报, 2024, 58(9): 139-153.
WANG Nan, HOU Tao, NIU Hongxia. Research on railway track foreign object intrusion detection based on multi-scale feature fusion [J]. Journal of Xi'an Jiaotong University, 2024, 58(9): 139-153.
- [15] 杨志渊, 罗亮, 吴天阳, 等. 改进 YOLOv8 的轻量级光学遥感图像船舶目标检测算法 [J]. 计算机工程与应用, 2024, 60(16): 248-257.
YANG Zhiyuan, LUO Liang, WU Tianyang, et al. Improved lightweight ship target detection algorithm for optical remote sensing images with YOLOv8 [J]. Computer Engineering and Applications, 2024, 60(16): 248-257.
- [16] ZHANG Jiarui, CHEN Zhihua, YAN Guoxu, et al. Faster and lightweight: an improved YOLOv5 object detector for remote sensing images [J]. Remote Sensing, 2023, 15(20): 4974.
- [17] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 1800-1807.
- [18] CHEN Jierun, KAO S H, HE Hao, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 12021-12031.
- [19] OUYANG Daliang, HE Su, ZHANG Guozhong, et al. Efficient multi-scale attention module with cross-spatial learning [C]//ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ, USA: IEEE, 2023: 1-5.
- [20] 王满利, 杨爽, 张长森. 基于改进 YOLOv8n 的立井刚性罐道接头错位检测算法 [J]. 煤炭科学技术, 2024, 52(S2): 236-248.
WANG Manli, YANG Shuang, ZHANG Changsen. An improved YOLOv8n based dislocation detection algorithm for shaft rigid tank channel joint [J]. Coal Science and Technology, 2024, 52(S2): 236-248.
- [21] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2020: 12993-13000.
- [22] ZHANG Hao, ZHANG Shuaijie. Focaler-IoU: more focused intersection over union loss [EB/OL]. (2024-01-19)[2025-05-23]. <https://arxiv.org/abs/2401.10525>.
- [23] LIU Can, WANG Kaige, LI Qing, et al. Powerful-IoU: more straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism [J]. Neural Networks, 2024, 170: 276-284.
- [24] CHEN Linwei, FU Ying, GU Lin, et al. Frequency-aware feature fusion for dense image prediction [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 10763-10780.
- [25] WU Tianyi, TANG Sheng, ZHANG Rui, et al. CG-Net: a light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [26] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: learning what you want to learn using programmable gradient information [C]//Computer Vision-ECCV 2024. Cham: Springer Nature Switzerland, 2025: 1-21.
- [27] ZHAO Yian, LV Wenyu, XU Shangliang, et al. DETRs beat YOLOs on real-time object detection [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2024: 16965-16974.

(编辑 武红江)